Centrality to the Metapopulation is more Important for Population Genetic Diversity than Habitat Area or Fragmentation

Anthony A. Snead[1,4], Andrey Tatarenkov[2], D. Scott Taylor[3], Kristine Marson[1], Ryan L Earley[1]

[1]Department of Biological Sciences, University of Alabama, 300 Hackberry Lane, Tuscaloosa, AL 35487, USA

[2]Department of Ecology and Evolutionary Biology, University of California, Irvine, CA 92697, USA

[3] 32620 W. Haverford Dr., Franklin MI 48025

[4]Current address: Department of Biology, New York University, New York, NY 10003, USA

Corresponding/Contact Author:

Anthony A. Snead

Email: anthony.snead@nyu.edu

Department of Biology

New York University

New York, NY 10003, USA

Running Title: Ocean Current Driven Connectivity

Keywords: Metapopulation Structure, Genetic Diversity, Gene Flow, Graph Theory, Ocean Currents

Abstract

Drift and gene flow affect genetic diversity. Given that the strength of genetic drift increases as population size decreases, management activities have focused on increasing population size through preserving habitats to preserve genetic diversity. Few studies have empirically evaluated the impacts of drift and gene flow on genetic diversity. *Kryptolebias marmoratus*, henceforth 'rivulus', is a small killifish restricted to fragmented New World mangrove forests with gene flow primarily associated with ocean currents. Rivulus form distinct populations across patches, making them a well-suited system to test the extent to which habitat area, fragmentation, and connectivity are associated with genetic diversity. Using over 1,000 individuals genotyped at 32 microsatellite loci, high resolution landcover data, and oceanographic simulations with graph theory, we demonstrate that centrality (connectivity) to the metapopulation is more strongly associated with genetic diversity than habitat area or fragmentation. By comparing models with and without centrality standardized by the source population's genetic diversity, our results suggest that metapopulation centrality is critical to genetic diversity regardless of the diversity of adjacent populations. While we find evidence that habitat area and fragmentation are related to genetic diversity, centrality is always a significant predictor with a larger effect than any measure of habitat configuration.

Introduction

Genetic diversity is one of the three internationally recognized levels of biological diversity by the United Nations' Convention on Biological Diversity and can impact the evolutionary trajectory of species[1,2]. Given that natural selection acts on heritable phenotypic variation, genetic diversity dictates the ability of populations to evolve in response to environmental conditions[3,4]. Preserving genetically distinct populations (i.e., preserving species-level genetic diversity) has been championed to maintain the potential for peripheral populations to rescue declining populations with low genetic diversity through natural (i.e., evolutionary rescue) or augmented (i.e., genetic rescue) gene flow[5,6]. Decreases in genetic diversity are associated with increased extinction risk[7,8]. While all genetic diversity ultimately originates from mutation, observed levels of genetic diversity result from previous episodes of gene flow, genetic drift, and selection[9]. Thus, genetic diversity is the product of past evolutionary forces while also dictating future evolutionary responses.

Genetic diversity often is attributed to gene flow and drift. Gene flow, the exchange of genetic material between populations, can either increase or decrease genetic diversity, depending on whether the focus is on the population or species-level. Gene flow can increase population genetic diversity through introduction of alleles from adjacent populations[10,11]; however, gene flow can reduce species-level genetic diversity by homogenizing allele frequencies and driving the loss of private alleles[12]. The pattern of gene flow across a group of populations is the result of immigration and emigration within a group of populations followed by successful reproduction (i.e., metapopulation structure)[14,15] . This genetic connectivity (i.e., gene flow) is impacted by the distance[16,17] and environmental conditions between populations[18]. Hence, environmental spatial heterogeneity between populations and the distribution of

66 populations impacts patterns of gene flow and, subsequently, patterns of genetic diversity.
67 However, population genetic diversity is likely not the product of incoming gene flow from one
68 population, but it is the sum of all incoming connections.

69 Genetic drift refers to stochastic changes in allele frequencies unrelated to fitness and
70 opposes local genetic diversity. However, independent bouts of drift between populations can
71 maintain species-level genetic diversity through retention of private alleles in isolated
72 populations[19]. Because drift is stronger in smaller populations[20,21], demographic declines
73 decrease genetic diversity[22,23] and limit the population's ability to recover from or respond to
74 environmental change[24,25]. Environmental changes that impact population size such as habitat
75 loss[26,27] and fragmentation[28,29] increase drift and decrease genetic diversity[30,31]. Gene flow and
76 drift can operate simultaneously[32], therefore, identifying environmental drivers of genetic
77 diversity requires concurrently evaluating the abiotic conditions that influence both patterns of
78 gene flow and the strength of drift.

79 *Kryptolebias marmoratus*[33], hereafter rivulus, is a cryptic self-fertilizing androdiecious
80 killifish[34] that inhabits the highly threatened and fragmented mangrove forests in North America,
81 Central America, the Caribbean, and the Bahamas[35,36,37]. As anthropogenic activities such as
82 greenhouse gas emission and land development continue, rivulus will face reduced habitat
83 availability[38] and may be ill-prepared to evolve in response to these novel environmental
84 conditions because populations often have low genetic diversity[39,40]. Rivulus dispersal is likely
85 passive through eggs attached, via adhesive filaments, to flotsam or adults rafting within
86 debris[40,41], thus limiting rivulus' ability to leave unsuitable habitats. Gene flow between rivulus
87 populations is generally low and asymmetric, with asymmetries associated with ocean currents[41],

resulting in complex patterns of gene flow that may limit the introduction of adaptive alleles and genetic rescue.

Given rivulus' limited genetic diversity, restricted gene flow, increased habitat loss, and the future impact of climate change on rivulus' distribution[38], preserving existing genetic diversity is essential for the persistence of rivulus populations. By quantifying abiotic factors that impact the strength of drift (i.e., habitat area, fragmentation) and patterns of gene flow (i.e., oceanic connectivity), we can explicitly evaluate their independent contributions to genetic diversity while comparing their relative impacts. Because population genetic diversity is likely associated with gene flow patterns across the range, we use a network approach to quantify each population's centrality to the metapopulation. Centrality refers to statistics that characterize how populations are connected within a directed network (see Supplementary Materials – *Network Centrality Measures*). We hypothesized that centrality to the metapopulation via ocean currents and patch qualities (e.g., habitat area) would influence population-level genetic diversity (H1). We predicted that increased centrality to the metapopulation would increase genetic diversity while decreases in habitat area and increased fragmentation would decrease genetic diversity.

Methods

Genetic and Environmental Data

We used 1,245 published genetic samples[39,40,41,42] genotyped at 32 microsatellite markers[43] previously collected from 56 sites across Central America, the Bahamas, the Caribbean, and North America between 1994 and 2014 (for details see Supplementary Material – *Genetic Data*). Using R version 4.3.1[44], we grouped sites into larger populations, filtered samples, and estimated oceanographic connectivity between each pair of populations following Snead et al. (2023) (for details see Supplementary Information – *Biophysical Modeling*) resulting

111 in 17 populations and 1,120 individuals (Figure 1, Table.S1). Using the biophysical model
112 (CMS) output files and the Global Land Cover and Land Use 2019 dataset[45], measures of
113 oceanographic connectivity and habitat configuration were calculated. Oceanographic
114 connectivity was calculated following Snead at al. (2023) (for details see SI– *Biophysical*
115 *Modeling*). The Global Land Cover and Land Use 2019 dataset[45] was downloaded at ~30 m^2
116 resolution and reclassified into either suitable habitat – wetlands (except salt pans) and water - or
117 unsuitable habitat – all remaining classifications resulting in a binary measure of potentially
118 suitable habitat. Total habitat area, cohesion, edge density, and the number of patches were
119 calculated for suitable habitat within each population cluster buffer using *landscapemetrics*[46].
120 Total habitat area is the measure of all suitable habitat within the population buffer, while the
121 number of patches is a count of the number of disconnected clusters of suitable habitats.
122 Cohesion is a measure of aggregation between zero and 100 that characterizes the continuity of
123 habitat within the buffer[47], and edge density represents the configuration of the landscape by
124 calculating the number of edges (i.e., where suitable habitat meets unsuitable habitat) and

standardizing it by the total area of within the buffer[48].

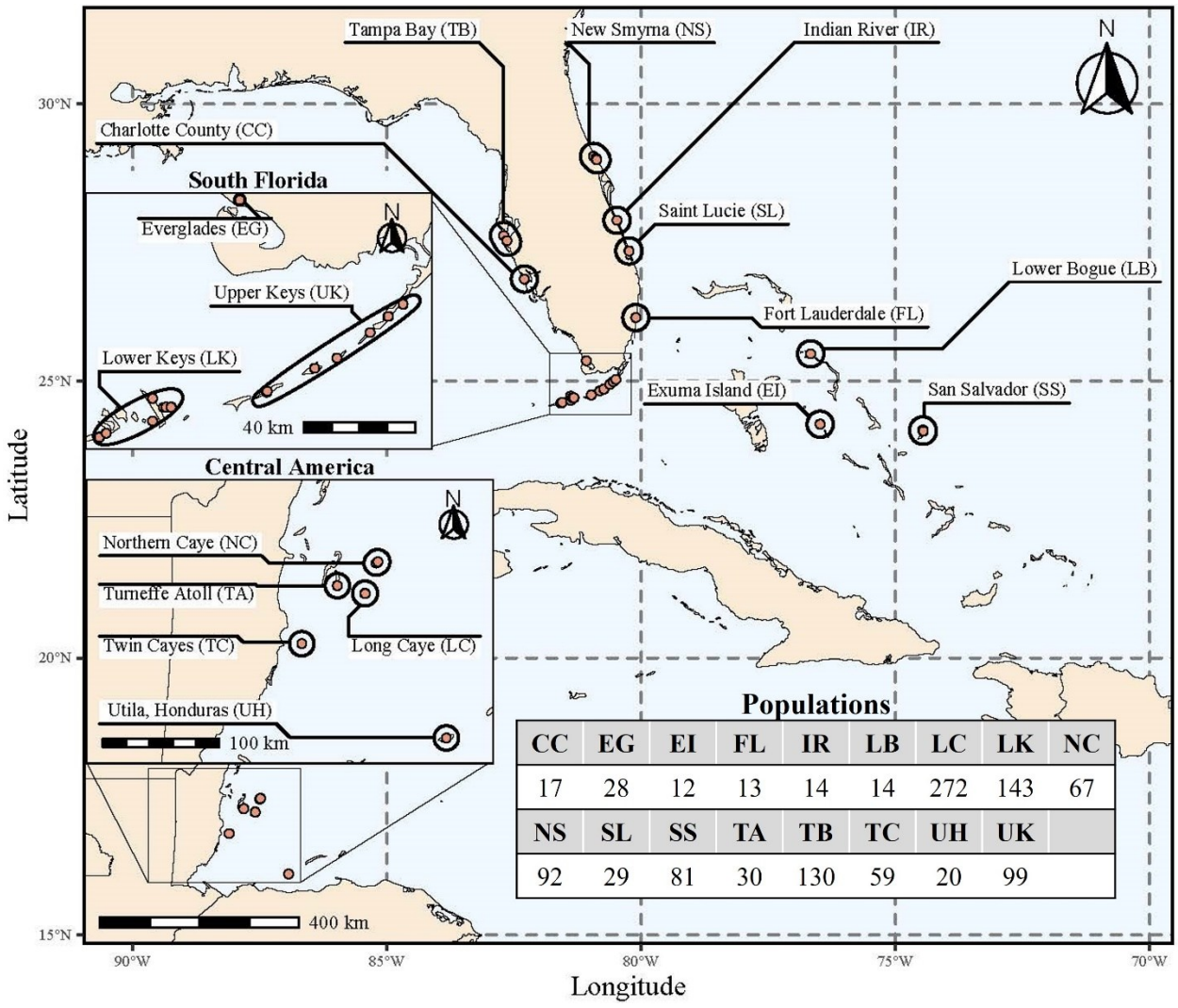

Figure 1 A map of all the sampling locations of *Kryptolebias marmoratus* in Florida and the Caribbean. The sampling locations were grouped in populations as shown on the map and described in the text. The number of samples in each population is shown in the table.

Genetic Analysis

Snead et al. (2023) use the same microsatellite data to explore local and regional population structure with an Analysis of MOlecular Variance[49], a Discriminant Analysis of Principal Components[50], TESS3[51], sNMF[52], STRUCTURE[53], and InStruct[54], while patterns of gene flow were investigated using G_{ST}[55], G'_{ST}[56], Jost's D[57], R_{ST}[58], and BayesAss[59]. Similarly to

Snead et al. (2023), deviations from Hardy-Weinberg Equilibrium (HWE) were tested for each microsatellite locus at the population level and for the entire dataset with the package *pegas*[60]. With the entire dataset, all loci deviated significantly from HWE, as expected due to nonrandom mating; however, no locus deviated from HWE in all populations. Therefore, all loci were retained. While there is a large temporal range across the samples, previous work found low of genetic differentiation (F_{ST} = 0.023) between samples collect over ten years apart in Twin Cayes, Belize (a population with more males and higher genetic diversity), and even lower patterns of isolation by time in three populations across the Florida Keys (F_{ST} = 0.002) which predominantly self-fertilize with few males.[40,42] Therefore, previous results suggest little change in genetic diversity across the sampling period.

Unique to this experiment, the rarefied number of multilocus genotypes (eMLG; the average number of unique multilocus genotypes after randomly subsampling ten individuals across 1000 iterations), Zahl's unbiased estimator (Z)[61,62], rarefied Stoddart and Taylor's index (G)[63], expected heterozygosity (H_{exp})[64], observed heterozygosity (H_{obs}), and the mean rarefied allelic richness (Ar) were calculated for each population with the packages *poppr*[65], *PopGenReport*[66], *adegenet*[67]. To account for uneven sampling across populations, both Ar and G were calculated with rarefaction. There were no significant correlations between sample size and genetic diversity (eMLG[r = 0.13, p = 0.61], Z[r = 0.37, p = 0.15], G[r = 0.1, p = 0.71], H_{exp}[r = 0.38, p = 0.14], H_{obs}[r = 0.12, p = 0.64], Ar[r = 0.37, p = 0.14]). This combination of metrics was chosen to facilitate comparisons between typical population genetic diversity metrics that lack strong assumptions (i.e., Z, H_{obs}, Ar) with H_{exp}, which assumes random mating, and a measure of genotypic diversity specifically developed for mixed mating systems (G)[63]. While H_{obs}, H_{exp}, and Ar are common metrics of genetic diversity in other mixed mating systems such as plant[68,69],

159 comparing results with Z and G enables us to evaluate whether our inference is robust to metric
160 choice and mating system by comparing across metrics with different assumptions.

161 *Statistics*

162 Snead et al. (2023) used measures of genetic differentiation and oceanic connectivity to
163 demonstrate that patterns of gene flow were primarily associated with ocean currents. Novel to
164 this experiment, oceanic connectivity values were used to calculate two measurements of
165 network centrality (closeness and strength). Closeness is the inverse average distance from any
166 node or vertex in the network to the target node, while strength is the sum of all oceanographic
167 connectivity estimates to a given vertex[70,71]. Models were constructed with centrality calculated
168 in two ways: with or without standardization of oceanographic connectivity by source population
169 genetic diversity. Comparing these models enabled us to determine whether source genetic
170 diversity modulates the impact of ocean connectivity on sink population genetic diversity (for
171 details see SI– *Network Centrality Measures*). All variables were scaled and centered prior to
172 variable reduction and modeling. The number of variables was reduced using a Variance
173 Inflation Factor (VIF) threshold of 5 before being further reducing to retain a metric of area (total
174 area), fragmentation (number of patches), and the two centrality measures (closeness and
175 strength). The VIF variable reduction was an iterative process where the variable with the
176 highest VIF was removed before the VIF for all variables were recalculated until no variables
177 had a VIF greater than 5. In fact, no variables had a VIF greater than 2.1 with the maximum
178 absolute correlation coefficient being between closeness and total area (-0.53) while the
179 minimum absolute correlation coefficient was between strength and total area (0.013).

180 Linear models were run separately using genetic diversity metrics as response variables
181 and every combination of landscape metrics, centrality measures, and all two-way interactions

between landscape metrics and centrality measures using the package *MuMIn*[72]. Models were run once with centrality measures standardized by source population genetic diversity and once without. To meet normality assumptions Z, G, and H_{exp} were raised to the 2nd, 3rd, and 3rd power, respectively, while Ar and H_{obs} were left untransformed. Models were compared via AICc per Burnham and Anderson (2004).

Results

Genetic Diversity

The rarefied number of multilocus genotypes (eMLG), Zahl's estimator (Z), Stoddart and Taylor's index (G), expected heterozygosity (H_{exp}), observed heterozygosity (H_{obs}), and allelic richness (Ar) varied considerably across populations. These metrics segregated largely on a regional basis, with few exceptions. Populations in North Florida (CC, IR, NS, SL, TB) and the Bahamas (EI, LB, SS) had fewer eMLGs, lower G, and lower genetic diversity (H_{exp} , H_{obs}, Ar) than populations in South Florida (EG, FL, LK, UK) and Central America (NC, LC, TA, TC, UH). Notable exceptions included that Honduran populations (UH) were less diverse than Belizean populations (NC, LC, TA, TC), low diversity in the southeastern-most population in peninsular Florida (FL) was more like Bahamas and Northern Florida populations than the other south Florida populations - Keys (LK, UK) and Everglades (EG), and two of the most genetically diverse populations, one from south Florida (LK) and another from Belize (LC), showed fewer eMLGs relative to other populations from the same regions (Table 1, Table S2).

Habitat Metrics

Total suitable habitat area ranged from 244.26 m^2 to 20,605.72 m^2 and number of patches from 8 to 1413 across rivulus populations. The general trend was for Belizean populations (LC, NC, TA, TC) to have much less area, but more contiguous area than most populations from Florida and the Bahamas. Exceptions included the southeastern-most population on the Florida peninsula (FL) and one Bahamas population (EI) having low area and the Everglades (EG) population having fewer patches compared to other non-Central American populations (Table 1, Table S2).

Network Variables

Network closeness ranged from 0.04 to 0.21, and network strength from 1.4×10^{-7} to 0.35. The Florida Keys (LK, UK) and larger islands off the coast of Belize (LC, TA, TC) had the highest closeness values, and showed some of the highest values for strength as well. Two populations with the highest area – Everglades (EG) in south Florida and New Smyrna (NS) in north Florida – had relatively low centrality. Populations on the southern fringe of island systems in Central America (UH) and the northern fringe of island systems in the Bahamas (LB) have some of the lowest measures of centrality. The Exuma Island (EI) population was the only one to show considerable disagreement in the two measures of centrality, closeness and strength; this population showed moderate-to-low closeness but high strength indicating that the population receives a large number of immigrants from a few adjacent populations but is not well connected to the entire metapopulation (Table1, Table S2).

Table 1 The rarefied number of multilocus genotypes (eMLG), Stoddart and Taylor's index (G), Simpson's index (λ), expected heterozygosity (H_{exp}), observed heterozygosity (H_{obs}), mean allelic richness (Ar), total habitat area (A), number of patches (NP), network closeness (C), and network strength (S) for each population along with the population abbreviation.

Population	N	eMLG	Z	G	H_{exp}	H_{obs}	Ar	A (m^2)	NP	C	S

Charlotte County (CC)	17	8.56	0.42	7.81	0.25	0.02	1.92	5907.72	677	0.1	0.03
Everglades (EG)	28	11.65	0.97	24.5	0.45	0.002	3.86	20,605.72	142	0.05	$6.28e^{-7}$
Exuma Island (EI)	12	9	0.21	8	0.14	0	1.37	806.16	289	0.08	0.05
Fort Lauderdale (FL)	13	12	0.78	13	0.44	0	2.79	645.06	515	0.1	$6.01e^{-3}$
Indian River (IR)	14	2.7	0.14	1.34	0.07	0	1.45	5156.91	905	0.08	$2.12e^{-3}$
Lower Bogue (LB)	14	5.43	0.6	3.5	0.36	0.02	2.22	7278.16	474	0.04	$1.44e^{-7}$
Long Caye (LC)	272	10.95	1.25	38.69	0.6	0.16	4.59	244.26	14	0.15	0.01
Lower Keys (LK)	143	11.31	1.19	53.96	0.54	0.01	4.59	10,959.34	711	0.14	0.21
Northern Caye (NC)	67	12	1.12	65.06	0.55	0.2	4.1	272.7	8	0.14	$8e^{-3}$
New Smyrna (NS)	92	9.85	0.34	20.35	0.18	0.001	1.93	112,717.53	1413	0.06	$9.42e^{-5}$
Saint Lucie (SL)	29	11.51	0.47	24.03	0.29	0.03	2	4,061.9	704	0.08	0.01
San Salvador (SS)	81	9.27	0.77	14.1	0.39	0.01	3.05	5,906.9	265	0.06	$5.2e^{3}$
Turneffe Atoll (TA)	30	12	1.29	30	0.59	0.28	4.82	3,026.49	51	0.18	0.03
Tampa Bay (TB)	130	5.49	0.45	4.12	0.28	0.001	1.89	3,682.99	1069	0.09	$7.26e^{-3}$
Twin Cayes (TC)	59	12	1.61	59	0.69	0.52	6.13	287.14	40	0.21	0.04
Utila, Honduras (UH)	20	11.65	0.99	18.18	0.5	0.004	3.66	3,435.69	154	0.01	$1.43e^{-4}$
Upper Keys (UK)	99	11.92	1.28	88.3	0.56	0.05	5.15	3,733.57	793	0.14	0.35

Statistical Models

The model rankings, coefficient estimates, and R^2 were similar between models that used centrality measures calculated with or without standardizing oceanic connectivity by the genetic diversity of the source population (Table 2, Table S3); therefore, models without centrality standardized are reported and discussed. Regardless of the diversity measure (Z, G, H_{exp}, H_{obs}, Ar) used as the response variable, closeness was always included within the best model and was significant ($p < 0.05$). In the set of models within 2 AICc units of the best model, habitat area was included in at least one of the best fit models for H_{exp} and Ar . The number of patches was included in the set of best fit models for all but diversity metrics. Strength was included only in the Ar set and was not significant. Habitat area, number of patches, and closeness were all significance ($0.05 < p < 0.1$) for at least one model in each set except G (H_{exp}, H_{obs}, Ar) with closeness being the only significant predictor of G (Table 2.

Table 2 A table with the formula, coefficient values, standard errors, significance, AICc, AICc weight ($AICc_w$), and adjusted R-squared (R^2) for all models using **unstandardized** centrality measures within 2 AICc units of the best model for all the diversity metrics (Stoddart and Taylor's Index = G, Expected Heterozygosity = H_{exp}, Observed Heterozygosity = H_{obs}, Allelic Richness = Ar). Covariates are symbolized by their abbreviations (Total Area = A, Number of

Patches = NP, Closeness = C, Strength = S) with interactions between variables indicated with an *x* between the two covariates and the intercept is reported for all models. Estimates shown in italics are significant at $0.01 \leqq P < 0.05$, and those shown in bold are significant at $P \leq 0.01$. If a cell is blank, it indicates that the covariate was not included in the best fit model(s).

		Habitat		Centrality					
Response	**Intercept**	**A**	**NP**	**C**	**S**	**A x C**	**AICc**	**AICc$_w$**	**R^2**
Z ~	**0.85 ± 0.16** p < 0.0001	**0.27 ± 0.19** p = 0.01	*-0.24 ± 0.18* p = 0.01	**0.67 ± 0.2** p < 0.0001			18.26	0.41	0.84
G ~	**1014.5 ± 285.578** p < 0.0001			*404.8 ± 0294.4* p = 0.01			268.66	0.37	0.32
G ~	**1014.5 ± 279.889** p < 0.0001		-196.2 ± 312.65 p = 0.2	***329.2 ± 312.68*** p = 0.04			270.06	0.16	0.36
H_{exp} ~	**0.1 ± 0.02** p < 0.0001	*0.02 ± 0.02* p = 0.04	**-0.03 ± 0.02** p = 0.006	**0.08 ± 0.02** p < 0.0001			-54.67	0.37	0.85
H_{exp} ~	**0.1 ± 0.02** p < 0.0001		*-0.03 ± 0.02* p = 0.02	**0.07 ± 0.02** p < 0.0001			-53.04	0.17	0.81
H_{obs} ~	**0.05 ± 0.03** p =0.005		**-0.05 ± 0.03** p = 0.004	**0.06 ± 0.04** p = 0.005		**-0.07 ± 0.04** p = 0.001	-39	0.49	0.84
Ar ~	**3.27 ± 0.35** p < 0.0001	**0.68 ± 0.43** p =0.005	*-0.5 ± 0.4* p =0.02	**1.31 ± 0.45** p < 0.0001			45.61	0.38	0.79
Ar ~	**3.27 ± 0.33** p < 0.0001	**0.6 ± 0.41** p = 0.008	**-0.63 ± 0.4** p = 0.005	**1.09 ± 0.49** p = 0.0004	0.34 ± 0.41 p = 0.09		46.34	0.26	0.82

Discussion

The spatial distribution of genetic variation is the product of drift, gene flow, natural selection, and mutation[74,75]. Because decreases in habitat area[26,27] and increases in fragmentation often decrease population size[28,29] and because the strength of drift increases as population sizes decline[16,20], habitat area and configuration are frequently prioritized when attempting to maintain genetic diversity[76,77]. However, comparing the relative importance of habitat measures against connectivity is uncommon. In this study, we combined over a thousand genetic samples from across rivulus' range, ocean current simulations, and land classification data within a network framework to test the role of habitat area, fragmentation, and connectivity in maintaining genetic variation. While our models show that both habitat configuration and connectivity dictate genetic

variation, connectivity was repeatedly identified as the most important determinant with the largest effect size.

Considering that mating system impacts genetic diversity[78], rivulus' status as a self-fertilizing vertebrate may spark warranted apprehension regarding the applicability of this study to other species, while variation in outcrossing and selfing rates across rivulus populations may raise concern regarding the determinants of genetic diversity. However, mixed mating systems are extremely common in plant studies using the same genetic diversity metrics[68,69]. Research suggests that mixed-mating systems can maintain genetic diversity at similar levels to purely outcrossing populations[79,80]. Within this study, there are examples of populations that primarily self and have low genetic diversity (North Florida) along with populations that primarily self and have high genetic diversity (South Florida). Populations with high genetic diversity and in which self-fertilization is the predominant mode of reproduction[39,40,42] also have high centrality to the metapopulation (Table 1, Table S2). Studies suggest that the genetic diversity metrics applied within this study and the comparison across populations with different outcrossing rates are robust and can be applied to other systems. However, mating systems should still be considered when designing management plans and interpreting patterns of genetic variation because mating systems have large impacts on genetic diversity.

Habitat area and fragmentation are often significantly associated with decreased genetic diversity, a finding that has inspired many management decisions[81]. While we found evidence for habitat area or fragmentation impacting the distribution of genetic variation for rivulus (Table 2), these variables were not always within the best model, nor did they have the largest effect size. When habitat area and fragmentation were included within the model, habitat area was positively associated with genetic diversity, while fragmentation was negatively associated with

282 genetic diversity, supporting previous studies in plants and mammals[30,31]. When testing our
283 genotypic measure of diversity (G), neither habitat area nor fragmentation were important
284 determinants. Hence, we find support for habitat configuration dictating genetic diversity but not
285 genotypic diversity (H1).

286 Drift and gene flow are regularly described as antagonistic, with drift decreasing and
287 gene flow increasing population-level genetic diversity[10,30]. We find that closeness (i.e., the
288 number and magnitude of incoming connections) was a significant predictor for all measures of
289 genetic diversity (i.e., Z, G, H_{exp}, H_{obs}, Ar) (H1; Table 2). We ran the analysis with and without
290 scaling measures of connectivity (used to calculate closeness and strength) by the source
291 populations' genetic pool (i.e., rarefied number of multilocus genotypes). Given that the results
292 of the two analyses were similar (Table 2; Table S3), genetic diversity may be more impacted by
293 centrality to the metapopulation than the specific genetic source pools of immigrants. While
294 there has been recent interest in preserving populations with high emigration that harbor genetic
295 diversity to facilitate natural genetic rescue[82], our results indicate that, for rivulus, genetic
296 diversity is linked more tightly with metapopulation structure than the level of genetic diversity
297 within connected populations or local habitat configuration.

298 While this research uses connectivity and measures of habitat configuration as proxies for
299 gene flow and drift, gene flow and drift are complex evolutionary forces that cannot be reduced
300 to any single environmental measure. Patterns of genetic variation are the product of historical
301 changes such as demography[74,75] that may not necessarily be represented in current
302 environmental conditions. Hence, the use of habitat configuration and connectivity as proxies for
303 drift and gene flow, respectively, should not be misconstrued as proposing equivalency because
304 current environmental patterns may not represent past patterns of evolutionary forces.

Furthermore, this study does not include all populations of rivulus across the range meaning that some aspects of connectivity may have been missed. The sampling does represent populations from all major areas across the range (i.e., Caribbean, Central America, South Florida, East Florida, and West Florida) which suggests that our estimated patterns of oceanic connectivity are representative even without some of the unsampled populations.

Anthropogenic activities are increasing fragmentation, decreasing habitat area, and exposing species to novel environmental stressors[83]. Hence, understanding the determinants of genetic variation, which is essential for the evolvability of populations[24,25], is critical to mitigate population extirpation. Using a network approach, we calculated connectivity with respect to the entire metapopulation and compared inferences with and without standardizing connectivity by source genetic diversity. While previous research emphasized associations between habitat configuration and genetic diversity, we found that patterns of connectivity - the population's location within the metapopulation network - is more important for genetic variation than the amount of habitat area or fragmentation suggesting that range-wide connectivity assessments are essential for designing effective management plans that not only protect populations in the present but preserve the evolvability of populations under future environmental change.

Acknowledgements

We want to acknowledge all the students and collaborators that helped sample. We also acknowledge John Avise for hosting the majority of the molecular work and Bruce Turner for his contributions to sampling.

Funding

Sampling and genotyping were funded by an EPSCor GRSP and an FSBI research grant to RLE as well as a Howard Hughes Medical Institute Undergraduate Science Education grant. AAS is supported by the NSF Postdoctoral Research Fellowships in Biology Program under Grant No. 2305939.

Author Contributions

AS: Conceptualization, methodology, software, formal analysis, data curation, writing—original draft preparation, writing—review and editing, visualization; AT: Methodology, investigation, resources, writing—review and editing; DT: Investigation, resources, writing—review and editing; KM: Investigation, resources, writing—review and editing; RE: Investigation, resources, writing—review and editing, supervision. All authors contributed to the article and approved the submitted version.

Competing Interests Statement

The authors declare no commercial or financial relationships that could be considered a potential conflict of interest related to this study.

Data Accessibility Statement

All code and microsatellite data for the manuscript is provided at https://github.com/anthonysnead/Rivulus-Genetic-Diversity. While the raw biophysical modeling data files were too large to provide, the Rdata file and an additional copy of the code for the manuscript is available at in a Figshare repository (https://figshare.com/s/200b01482fdbb5597ef5).

References

1. McClure MM, Carlson SM, Beechie TJ, Pess GR, Jorgensen JC, Sogard SM, et al. Evolutionary consequences of habitat loss for Pacific anadromous salmonids. Evolutionary Applications. 2008;1(2):300–18.

2. O'Donnell DR, Parigi A, Fish JA, Dworkin I, Wagner AP. The Roles of Standing Genetic Variation and Evolutionary History in Determining the Evolvability of Anti-Predator Strategies. PLOS ONE. 2014 Jun 23;9(6):e100163.

3. Vilas A, Pérez-Figueroa A, Quesada H, Caballero A. Allelic diversity for neutral markers retains a higher adaptive potential for quantitative traits than expected heterozygosity. Molecular Ecology. 2015;24(17):4419–32.

4. de Villemereuil P, Rutschmann A, Lee KD, Ewen JG, Brekke P, Santure AW. Little Adaptive Potential in a Threatened Passerine Bird. Current Biology. 2019;4;29(5):889-894.e3.

5. Lesica P, Allendorf FW. When are Perifpheral Populations Valuable for Conservation? Conservation Biology. 1995;9(4):753-760.

6. Bell, DA, Robinsion ZL, Funck WC, Fitzapatrick SW, Allendorf FW, Tallmon DA, Whiteley AR. The Exciting Potential and Remaining Uncertainties of Genetic Rescue. Trends in Ecology & Evolution. 2019;34(12):1070-1079.

7. Benson JF, Mahoney PJ, Sikich JA, Serieys LEK, Pollinger JP, Ernest HB, et al. Interactions between demography, genetics, and landscape connectivity increase extinction probability for a small population of large carnivores in a major metropolitan area. Proceedings of the Royal Society B: Biological Sciences. 2016 Aug 31;283(1837):20160957.

8. Weeks AR, Stoklosa J, Hoffmann AA. Conservation of genetic uniqueness of populations may increase extinction likelihood of endangered species: the case of Australian mammals. Front Zool. 2016 Jul 8;13(1):31.

9. Ellegren H, Galtier N. Determinants of genetic diversity. Nat Rev Genet. 2016 Jul;17(7):422–33.

10. Morrissey MB, de Kerckhove DT. The Maintenance of Genetic Variation Due to Asymmetric Gene Flow in Dendritic Metapopulations. The American Naturalist. 2009 Dec;174(6):875–89.

11. Smith AL, Hodkinson TR, Villellas J, Catford JA, Csergő AM, Blomberg SP, et al. Global gene flow releases invasive plants from environmental constraints on genetic diversity. Proceedings of the National Academy of Sciences. 2020 Feb 25;117(8):4218–27.

12. Garant D, Forde SE, Hendry AP. The Multifarious Effects of Dispersal and Gene Flow on Contemporary Adaptation. Functional Ecology. 2007;21(3):434–43.

13. Hanksi, I. Single-species metapopulation dynamics: concepts, models and observations. Biological Journal of the Linnean Society. 1991 Jan 1;42(1–2), pp. 17–38.

14. Botigué LR, Henn BM, Gravel S, Maples BK, Gignoux CR, Corona E, et al. Gene flow from North Africa contributes to differential human genetic diversity in southern Europe. Proceedings of the National Academy of Sciences. 2013 Jul 16;110(29):11791–6.

15. Cortés AJ, Waeber S, Lexer C, Sedlacek J, Wheeler JA, van Kleunen M, et al. Small-scale patterns in snowmelt timing affect gene flow and the distribution of genetic diversity in the alpine dwarf shrub Salix herbacea. Heredity. 2014 Sep;113(3):233–9.

16. Wright S. Isolation by Distance. Genetics. 1943 Mar;28(2):114–38.

17. Jenkins DG, Carey M, Czerniewska J, Fletcher J, Hether T, Jones A, et al. A meta-analysis of isolation by distance: relic or reference standard for landscape genetics? Ecography. 2010;33(2):315–20.

18. van Rees CB, Reed JM, Wilson RE, Underwood JG, Sonsthagen SA. Landscape genetics identifies streams and drainage infrastructure as dispersal corridors for an endangered wetland bird. Ecology and Evolution. 2018;8(16):8328–43.

19. Hill P, Dickman CR, Dinnage R, Duncan RP, Edwards SV, Greenvill A, Sarre SD, Stringer EJ, Wardle GM, Gruber B. Episodic Population Fragmentation and Gene Flow Reveal a Trade-Off Between Heterozygosity and Allelic Richness. Molecular Ecology. 2023;32(24):6766-6776.

20. Wright S. Evolution in Mendelian Populations. Genetics. 1931 Mar;16(2):97–159.

21. Wright S. Evolution and the Genetics of Populations, Volume 2: Theory of Gene Frequencies. Chicago, IL: University of Chicago Press; 1984.

22. Stoffel MA, Humble E, Paijmans AJ, Acevedo-Whitehouse K, Chilvers BL, Dickerson B, et al. Demographic histories and genetic diversity across pinnipeds are shaped by human exploitation, ecology and life-history. Nat Commun. 2018 Nov 16;9(1):4836.

23. Espregueira Themudo G, Rey-Iglesia A, Robles Tascón L, Bruun Jensen A, da Fonseca RR, Campos PF. Declining genetic diversity of European honeybees along the twentieth century. Sci Rep. 2020 Jun 29;10(1):10520.

24. Westemeier RL, Brawn JD, Simpson SA, Esker TL, Jansen RW, Walk JW, et al. Tracking the Long-Term Decline and Recovery of an Isolated Population. Science. 1998 Nov 27;282(5394):1695–8.

25. Aspi J, Roininen E, Ruokonen M, Kojola I, Vilà C. Genetic diversity, population structure, effective population size and demographic history of the Finnish wolf population. Molecular Ecology. 2006;15(6):1561–76.

26. Bender DJ, Contreras TA, Fahrig L. Habitat Loss and Population Decline: A Meta-Analysis of the Patch Size Effect. Ecology. 1998;79(2):517–33.

27. Fletcher Jr. RJ, Reichert BE, Holmes K. The negative effects of habitat fragmentation operate at the scale of dispersal. Ecology. 2018;99(10):2176–86.

28. Wiegand T, Revilla E, Moloney KA. Effects of Habitat Loss and Fragmentation on Population Dynamics. Conservation Biology. 2005;19(1):108–21.

29. Brinker A, Chucholl C, Behrmann-Godel J, Matzinger M, Basen T, Baer J. River damming drives population fragmentation and habitat loss of the threatened Danube streber (*Zingel streber*): Implications for conservation. Aquatic Conserv: Mar Freshw Ecosyst. 2018 Jun;28(3):587–99.

30. Schlaepfer DR, Braschler B, Rusterholz HP, Baur B. Genetic effects of anthropogenic habitat fragmentation on remnant animal and plant populations: a meta-analysis. Ecosphere. 2018;9(10):e02488.

31. Lino A, Fonseca C, Rojas D, Fischer E, Pereira MJR. A meta-analysis of the effects of habitat loss and fragmentation on genetic diversity in mammals. Mamm Biol. 2019 Jan 1;94(1):69–76.

32. Monaghan MT, Spaak P, Robinson CT, Ward JV. Population Genetic Structure of 3 Alpine Stream Insects: Influences of Gene Flow, Demographics, and Habitat Fragmentation. Journal of the North American Benthological Society. 2002 Mar;21(1):114–31.

33. Costa WJEM. Kryptolebias, a substitute name for Cryptolebias Costa, 2004 and Kryptolebiatinae, a substitute name for Cryptolebiatinae Costa, 2004 (Cyprinodontiformes: Rivulidae). Neotrop ichthyol. 2004 Jun;2:107–8.

34. Harrington RW. How Ecological and Genetic Factors Interact to Determine When Self-Fertilizing Hermaphrodites of Rivulus marmoratus Change into Functional Secondary Males, with a Reappraisal of the Modes of Intersexuality among Fishes. Copeia. 1971;1971(3):389–432.

35. Blanco-Libreros JF, Estrada-Urrea EA. Mangroves on the Edge: Anthrome-Dependent Fragmentation Influences Ecological Condition (Turbo, Colombia, Southern Caribbean). Diversity. 2015 Sep;7(3):206–28.

36. Tatarenkov A, Lima SMQ, Earley RL, Berbel-Filho WM, Vermeulen FBM, Taylor DS, et al. Deep and concordant subdivisions in the self-fertilizing mangrove killifishes (Kryptolebias) revealed by nuclear and mtDNA markers. Biological Journal of the Linnean Society. 2017 Oct 25;122(3):558–78.

37. Taylor DS. Twenty-Four Years in the Mud: What Have We Learned About the Natural History and Ecology of the Mangrove Rivulus, Kryptolebias marmoratus? Integrative and Comparative Biology. 2012 Dec 1;52(6):724–36.

38. Snead AA, Earley RL. Predicting the in-between: Present and future habitat suitability of an intertidal euryhaline fish. Ecological Informatics. 2022 May 1;68:101523.

39. Avise JC, Tatarenkov A. Population genetics and evolution of the mangrove rivulus Kryptolebias marmoratus, the world's only self-fertilizing hermaphroditic vertebrate. Journal of Fish Biology. 2015;87(3):519–38.

40. Tatarenkov A, Earley RL, Taylor DS, Avise JC. Microevolutionary Distribution of Isogenicity in a Self-fertilizing Fish (Kryptolebias marmoratus) in the Florida Keys. Integrative and Comparative Biology. 2012 Dec 1;52(6):743–52.

41. Snead AA, Tatarenkov A, Avise JC, Taylor DS, Turner BJ, Marson K, et al. Out to sea: ocean currents and patterns of asymmetric gene flow in an intertidal fish species. Frontiers in Genetics. 2023 June 28;14

42. Tatarenkov A, Gao H, Mackiewicz M, Taylor DS, Turner BJ, Avise JC. Strong population structure despite evidence of recent migration in a selfing hermaphroditic vertebrate, the mangrove killifish (Kryptolebias marmoratus). Molecular Ecology. 2007;16(13):2701–11.

43. Mackiewicz M, Tatarenkov A, Perry A, Martin JR, Elder JF, Bechler DL, et al. Microsatellite Documentation of Male-Mediated Outcrossing between Inbred Laboratory Strains of the Self-Fertilizing Mangrove Killifish (Kryptolebias Marmoratus). Journal of Heredity. 2006 Sep 1;97(5):508–13.

44. R Core Team. A Language and Environment for Statistical Computing. Foundation for Statistical Computing. 2023. Vienna, Austria.

45. Hansen MC, Potapov PV, Pickens AH, Tyukavina A, Hernandez-Serna A, Zalles V, et al. Global land use extent and dispersion within natural land cover using Landsat data. Environ Res Lett. 2022 Mar;17(3):034050.

46. Hesselbarth MHK, Sciaini M, With KA, Wiegand K, Nowosad J. landscapemetrics: an open-source R tool to calculate landscape metrics. Ecography. 2019;42(10):1648–57.

43. 1.

47. Schumaker NH. Using Landscape Indices to Predict Habitat Connectivity. Ecology. 1996;77(4):1210–25.

48. McGarigal K, Cushman SA, Ene E. FRAGSTATS v4: Spatial Pattern Analysis Program for Categorical Maps. [Internet]. 2023. Available from: https://www.fragstats.org

49. Excoffier, L., Smouse, P.E. and Quattro, J. Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. Genetics. 1992;131(2), pp.479-491.

50. Jombart, T., Devillard, S., & Balloux, F. Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. BMC genetics. 2010;11, 1-15.

51. Caye, K., Deist, T. M., Martins, H., Michel, O., & François, O. TESS3: fast inference of spatial population structure and genome scans for selection. Molecular Ecology Resources. 2016;16(2), 540-548.

52. Frichot, E., & François, O. LEA: An R package for landscape and ecological association studies. Methods in Ecology and Evolution, 2015;6(8), 925-929.

53. Pritchard, J. K., Stephens, M., & Donnelly, P. Inference of population structure using multilocus genotype data. Genetics. 2000;155(2), 945-959.

54. Gao, H., Williamson, S., & Bustamante, C. D. A Markov chain Monte Carlo approach for joint inference of population structure and inbreeding rates from multilocus genotype data. Genetics. 2007;176(3), 1635-1651.

55. Nei, M., & Chesser, R. K. Estimation of fixation indices and gene diversities. Annals of human genetics. 1983;47(3), 253-259.

56. Hedrick, P. W. A standardized genetic differentiation measure. Evolution. 2005;59(8), 1633-1638.

57. Jost, L. GST and its relatives do not measure differentiation. Molecular ecology. 2008;17(18), 4015-4026.

58. Slatkin, M. A measure of population subdivision based on microsatellite allele frequencies. Genetics. 1995;139(1), 457-462.

59. Wilson, G. A., & Rannala, B. Bayesian inference of recent migration rates using multilocus genotypes. Genetics. 2003;163(3), 1177-1191.

60. Paradis E. pegas: an R package for population genetics with an integrated–modular approach. Bioinformatics. 2010 Feb 1;26(3):419–20.

61. Zahl, S. Jackknifing an index of diversity. Ecology; 1997;58(4), 907-913.

62. Konopiński, M. K. Shannon diversity index: a call to replace the original Shannon's formula with unbiased estimator in the population genetics studies. PeerJ. 2020;8, e9391.

63. Stoddart JA, Taylor JF. Genotypic Diversity: Estimation and Prediction in Samples. Genetics. 1988 Apr;118(4):705–11.

64. Nei, M. Estimation of average heterozygosity and genetic distance from a small number of individuals. Genetics. 1978;89(3), 583-590.

65. Kamvar ZN, Tabima JF, Grünwald NJ. Poppr: an R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction. PeerJ. 2014 Mar 4;2:e281.

66. Adamack, A. T., & Gruber, B. PopGenReport: simplifying basic population genetic analyses in R. Methods in Ecology and Evolution. 2014;5(4), 384-387.

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

68. Kennedy, J. P., Pil, M. W., Proffitt, C. E., Boeger, W. A., Stanford, A. M., & Devlin, D. J. Postglacial expansion pathways of red mangrove, Rhizophora mangle, in the Caribbean Basin and Florida. American Journal of Botany. 2016;103(2), 260-276.

69. López-Villalobos, A., & Eckert, C. G. Consequences of multiple mating-system shifts for population and range-wide genetic structure in a coastal dune plant. Molecular Ecology. 2018;27(3), 675-693.

70. Barrat A, Barthélemy M, Pastor-Satorras R, Vespignani A. The architecture of complex weighted networks. Proceedings of the National Academy of Sciences. 2004 Mar 16;101(11):3747–52.

71. Scott J. Social Networks: Critical Concepts in Sociology. Taylor & Francis; 2002. 416 p.

72. Bartoń K. MuMIn: Multi-model inference. 2009.

73. Burnham KP, Anderson DR. Multimodel Inference: Understanding AIC and BIC in Model Selection. Sociological Methods & Research. 2004 Nov 1;33(2):261–304.

74. Blanquart F, Gandon S, Nuismer SL. The effects of migration and drift on local adaptation to a heterogeneous environment. Journal of Evolutionary Biology. 2012;25(7):1351–63.

75. McDonald TK, Yeaman S. Effect of migration and environmental heterogeneity on the maintenance of quantitative genetic variation: a simulation study. Journal of Evolutionary Biology. 2018;31(9):1386–99.

76. Moritz C. Strategies to Protect Biological Diversity and the Evolutionary Processes That Sustain It. Systematic Biology. 2002 Mar 1;51(2):238–54.

77. Azman A, Ng KKS, Ng CH, Lee CT, Tnah LH, Zakaria NF, et al. Low genetic diversity indicating the threatened status of Rhizophora apiculata (Rhizophoraceae) in Malaysia: declined evolution meets habitat destruction. Sci Rep. 2020 Nov 5;10(1):19112.

78. De Kort, H., Prunier, J. G., Ducatez, S., Honnay, O., Baguette, M., Stevens, V. M., & Blanchet, S. Life history, climate and biogeography interactively affect worldwide genetic diversity of plant and animal populations. Nature Communications. 2021;12(1), 516.

79. Laenen, B., Tedder, A., Nowak, M. D., Toräng, P., Wunder, J., Wötzel, S., ... & Slotte, T. Demography and mating system shape the genome-wide impact of purifying selection in Arabis alpina. Proceedings of the National Academy of Sciences. 2018;115(4), 816-821.

80. Stenøien, H. K., & Såstad, S. M. Genetic variability in bryophytes: does mating system really matter?. Journal of Bryology. 2001;23(4), 313-318.

81. Hoban S, Bruford M, D'Urban Jackson J, Lopes-Fernandes M, Heuertz M, Hohenlohe PA, et al. Genetic diversity targets and indicators in the CBD post-2020 Global Biodiversity Framework must be improved. Biological Conservation. 2020 Aug 1;248:108654.

82. Minter M, O'Brien D, Cottrell J, Ennos R, Hill JK, Hall J. Exploring the potential for 'Gene Conservation Units' to conserve genetic diversity in wild populations. Ecological Solutions and Evidence. 2021;2(2):e12061.

83. Stone DA, Hansen G. Rapid systematic assessment of the detection and attribution of regional anthropogenic climate change. Clim Dyn. 2016 Sep 1;47(5):1399–415.