Title

Genetic Subdivision and Variation in Selfing Rates Among Central American Populations of the Mangrove Rivulus, Kryptolebias marmoratus

Journal

Journal of Heredity, 106(3)

ISSN

0022-1503

Authors

Tatarenkov, Andrey
Earley, Ryan L
Perlman, Benjamin M
et al.

Publication Date

2015-05-01

DOI

10.1093/jhered/esv013

American Genetic Association

Journal of Heredity, 2015, 276–284
doi:10.1093/jhered/esv013
Original Article
Advance Access publication March 25, 2015

OXFORD

Original Article

Genetic Subdivision and Variation in Selfing Rates Among Central American Populations of the Mangrove Rivulus, *Kryptolebias marmoratus*

Andrey Tatarenkov, Ryan L. Earley, Benjamin M. Perlman, D. Scott Taylor, Bruce J. Turner, and John C. Avise

From the Department of Ecology and Evolutionary Biology, University of California, Irvine, CA 92697 (Tatarenkov and Avise); the Department of Biological Sciences, University of Alabama, Tuscaloosa, AL (Earley); the Department of Biology, Wake Forest University, Winston-Salem, NC (Perlman); the Brevard County Environmentally Endangered Lands Program, Melbourne, FL (Taylor); and the Department of Biological Sciences, Virginia Polytechnic Institute and State University, Blacksburg, VA (Turner).

Address correspondence to Andrey Tatarenkov at the address above, or e-mail: tatarenk@uci.edu.

Data deposited at Dryad: http://dx.doi.org/doi:10.5061/dryad.528nm

Received November 10 2014; First decision December 17, 2014: Accepted February 16, 2015.

Corresponding Editor: Stewart Grant

Abstract

We used 32 polymorphic microsatellite loci to investigate how a mixed-mating system affects population genetic structure in Central American populations ($N = 243$ individuals) of the killifish *Kryptolebias marmoratus* (mangrove rivulus), 1 of 2 of the world's only known self-fertilizing vertebrates. Results were also compared with previous microsatellite surveys of Floridian populations of this species. For several populations in Belize and Honduras, population structure and genetic differentiation were pronounced and higher than in Florida, even though the opposite trend was expected because populations in the latter region were presumably smaller and highly selfing. The deduced frequency of selfing (s) ranged from $s = 0.39$–0.99 across geographic locales in Central America. This heterogeneity in selfing rates was in stark contrast to Florida, where $s > 0.9$. The frequency of outcrossing in a population ($t = 1 - s$) was tenuously correlated with local frequencies of males, suggesting that males are one of many factors influencing outcrossing. Observed distributions of individual heterozygosity showed good agreement with expected distributions under an equilibrium mixed-mating model, indicating that rates of selfing remained relatively constant over many generations. Overall, our results demonstrate the profound consequences of a mixed-mating system for the genetic architecture of a hermaphroditic vertebrate.

Subject area: Population structure and phylogeography
Key words: dispersal, gene flow, heterozygosity, microsatellites, outcrossing

Mating systems can profoundly impact genetic variation and population structure. Consistent self-fertilization is an extreme form of inbreeding causing a cascade of population genetic consequences (Charlesworth 2003; Cutter 2006). Selfing reduces effective population size such that N_e in a population of strict selfers is only 50% that of constitutive cross-fertilizers, all else being equal (Nordborg 1997). Since equilibrium levels of genetic variation at neutral loci are proportional to N_e, populations of selfers will generally

be genetically less variable than those of outcrossers. One major effect of inbreeding is the increased levels of individual homozygosity, which within several generations can reach nearly 100% in a population of predominant self-fertilizers (Hedrick 2005). Although meiosis and crossing-over go on unimpeded in selfers, the elevated homozygosity reduces the number of realized recombination events, in effect tightening linkage and heightening gametic-phase disequilibrium throughout the genome (Charlesworth 2003). Tight linkage between loci in selfers enhances opportunities for natural selection to "see" both beneficial and deleterious mutations and thereby to drive selective sweeps and promote background selection, eventually resulting in further reductions of within-population genetic diversity (Cutter 2006; Andersen et al. 2012). Furthermore, because the impact of genetic drift is higher with smaller N_e, population structure should be magnified by inbreeding (Maruyama and Tachida 1992; Charlesworth and Pannell 2001), as has been demonstrated in several species of selfing plants and animals (Hamrick and Godt 1996; Jarne 1995; Jarne and Auld 2006). Moreover, whereas genetic variation in local populations is expected to be low in selfing species, genetic diversity over the entire species' range may be exceptionally high due to pronounced population subdivision that conserves distinct genetic variants (Charlesworth and Pannell 2001; Ingvarsson 2002). Finally, the ability of a single selfing hermaphrodite to establish a new population may further contribute to low intra- and high inter-population variation via genetic bottlenecking (Baker 1955).

The mangrove rivulus, *Kryptolebias marmoratus* (Rivulidae) is a small fish that is most famous as the world's only self-fertilizing hermaphroditic vertebrate (Harrington 1961). Indeed, the only other known selfing vertebrate (*K. hermaphroditus*; Costa 2011) has only recently been recognized as a distinct species. Because *K. marmoratus* previously was not distinguished from this sibling species, their respective exact geographic ranges are unknown. However, the combined ranges are broad: peninsular Florida, most Caribbean islands, including the Bahamas, and the Atlantic coast from Yucatan to southeastern Brazil (Taylor 2000, 2012; Tatarenkov et al. 2011). In addition to selfing, members of the mangrove rivulus occasionally outcross, apparently by spawning with gonochoristic males that occur in some populations. Thus, *K. marmoratus* also provides an example of androdioecy, a rare reproductive system involving hermaphrodites plus pure males (Avise 2011).

With the exception of 1 surveyed population in Twin Cays, Belize (Turner et al. 1992a), males appear to be rare (< ~2%) in most populations of *K. marmoratus*. There is currently no satisfactory explanation why males comprise about 20% of the Twin Cays population, while other locations in Belize and Honduras reportedly have male frequencies as low as 1–2%, similar to those observed in Florida and the Bahamas (Taylor et al. 2001; Turner et al. 2006; Ellison et al. 2012).

Early studies involving fin grafts (Harrington and Kallman 1968) and DNA fingerprinting (Turner et al. 1992b) implied that *Kryptolebias* lineages in Florida are homozygous across the genome and produce progeny that are genetically identical to one another and to the selfing parent. In contrast, all examined fish from Twin Cays, Belize were heterozygous, as inferred from segregant progeny, apparently due to crossbreeding between hermaphrodites and males (Lubinski et al. 1995; Taylor et al. 2001). Later, all of these results were confirmed and extended by genotyping large numbers of microsatellite loci (Mackiewicz et al. 2006a,b,c). Moreover, the use of microsatellites permitted estimates of selfing rates (s), which were found to be high (s = 0.91–1.0) in Florida and the Bahamas but considerably lower (s = 0.40–0.45) in Twin Cays, Belize. Subsequent genetic studies showed that natural populations were highly subdivided (Tatarenkov et al. 2007, 2012), detected peculiarities in the distribution of isogenic lineages within and among sites (Tatarenkov et al. 2012), found high selfing rates and low intrapopulational variation in *K. hermaphroditus* in southeastern Brazil (Tatarenkov et al. 2009, 2011), and demonstrated random mating (s = 0) in another androdioecious congener, *K. ocellatus* (=*K. caudomarginatus* [Seegers]) (Tatarenkov et al. 2009).

Because most prior genetic analyses of *K. marmoratus* have focused on Floridian populations, with only 2 locations assayed in the Bahamas and Belize (Tatarenkov et al. 2007; Ellison et al. 2011, 2012), population structure and outcrossing rates ($t = 1 - s$) remain virtually unknown across the remainder of the species' vast range. Meanwhile, geographic areas can vary greatly in population sizes and demography, correspondingly affecting population structure and selection on mating systems. Some of the factors shaping distinct population structure in Florida compared with other regions may be due to the fact that Florida is at the northern limit of the *K. marmoratus* distribution. Peripheral populations are generally smaller and often ephemeral, thus increasing the role of stochastic processes (genetic drift) resulting in more pronounced genetic structure (Eckert et al. 2008). Marginal populations of many organisms are characterized by increased clonal reproduction because clonality eliminates the risk of not finding a mate in low-density populations and makes possible colonization by a single individual (Kawecki 2008). When applied to mangrove rivulus, this may translate into higher selfing rates in Florida populations. Accordingly, here we extend our genetic analyses of *K. marmoratus* to additional sites in Belize and Honduras. Our objectives are to: 1) estimate selfing/outcrossing rates at new locales and evaluate the correlation of these rates with sexual composition of the populations; 2) compare observed distributions of heterozygosity to those expected under mixed-mating in steady state populations; 3) characterize local genetic differentiation and establish whether it follows an isolation-by-distance (IBD) model; 4) compare population structure in Florida and Central America; and 5) explore whether the patterns of genetic diversity and differentiation under selfing follow theoretical predictions and, in particular, whether a region with greater selfing has higher differentiation among populations than one with a higher incidence of outcrossing.

Materials and Methods

Samples

We analyzed a total of 243 specimens of *K. marmoratus* from several mangrove islands in Central America (Figure 1; Supplementary Table S1 online).

To compare population structure in Central America versus the Florida Keys, we used a previously reported dataset (Tatarenkov et al. 2012) with 2 small modifications: 1) we expanded the sample size for site CRWL by adding 17 fish; and 2) we added an additional sample (SOB, N = 11) from Sugarloaf Key (24°36'5.22"N, 81°34'34.08"W). The 2 geographic regions are similar in size and the microsatellite data are directly comparable (the same allele scoring system is used).

Frequency of males was estimated using all available specimens ever collected in particular locales.The sex of each specimen (hermaphrodite vs. male) was assessed by body coloration and availability of a black ocellus on the caudal fin; males are orange in color and lack the caudal ocellus.

Microsatellite Genotyping

Genomic DNA was extracted from fish fin clips preserved in 95% ethanol or DMSO solution using proteinase K method, as described in Tatarenkov et al. (2012).

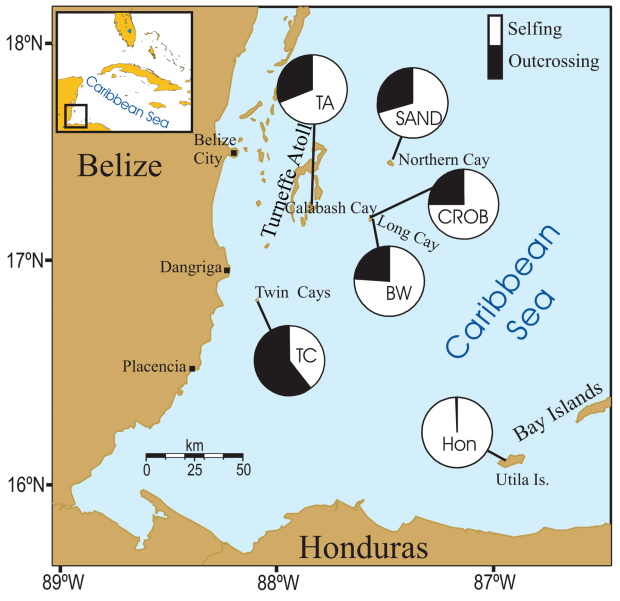

Figure 1. Sampling sites and rates of selfing and outcrossing (white and black, respectively) in populations of *Kryptolebias marmoratus* from Central America. Populations are abbreviated as in Table 1.

We used 32 microsatellite loci developed for *K. marmoratus* (Mackiewicz et al. 2006a), which were amplified in several multiplex reactions, as described previously (Tatarenkov et al. 2012). Alleles were separated in an GA 3100 instrument and scored using Genemapper 4.0 (Applied Biosystems). Binning of alleles followed Tatarenkov et al. (2010). We have deposited the primary data underlying these analyses with Dryad following data archiving guidelines (Baker 2013).

Statistical Analyses

GDA software (Lewis and Zaykin 2001) was used to calculate basic descriptive statistics (H_E, H_O, F_{IS}) and to find 95% confidence intervals for F_{IS} by bootstrapping of loci. Rates of selfing (s) and outcrossing ($t = 1 - s$) were estimated from the empirical fixation index using the equation $s = 2F_{IS}/(1 + F_{IS})$. We also used the program RMES (David et al. 2007) to estimate selfing rate using the maximum likelihood option; we refer to this estimate as s_{ml} to distinguish it from s. Unlike estimation of s from F_{IS}, the approach implemented in RMES does not use heterozygote deficiencies and, therefore, the 2 estimates of selfing rates are independent. Statistical significance of differences in selfing rates was determined by comparing unconstrained and constrained likelihoods, as described in RMES documentation. Pairwise and overall estimates of population differentiation F_{ST} were computed in FSTAT (ver. 2.9.3.2; Goudet 1995), and significances were evaluated by performing 1000 permutations of genotypes among samples. Genepop (ver. 4.2; Rousset 2008) was used to test for IBD with a Mantel test and 10 000 permutations. Hierarchical analysis of genetic differentiation was conducted in HIERFSTAT (Goudet 2005), with statistical significance of differences between sites within regions ($F_{Sites/Regions}$) and between regions ($F_{Regions/Total}$) being assessed with 1000 randomizations. The number of heterozygous loci (32 loci genotyped for each individual) was counted with the help of Microsatellite Analyser (MSA; Dieringer and Schlötterer 2003). These "individual heterozygosity" scores were used to build distributions of individual heterozygosity for each population. Note that the average of individual heterozygosities divided by the total number of loci corresponds to the observed heterozygosity (H_O) of a population. In organisms with a mixed-mating system each individual has a selfing history characterized by the number of generations of self-fertilization until the most recent event of outcrossing. Since each generation of selfing reduces heterozygosity by a factor of 2, the distribution of individual heterozygosity should show discrete peaks formed by individuals of a particular selfing generation and, depending on the populational selfing rate, a peak formed by individuals resulting from outcrossing. If outcrossing is random, then the latter peak would correspond to gene diversity (H_E). Comparison of the observed distribution of individual heterozygosities with its expected distribution can be informative about mating patterns in natural populations. The theoretical distributions of heterozygosities under a mixed mating model were calculated as described in Mackiewicz et al. (2006c). In addition to the assumption of random mating, this model also assumes constant allele frequencies and selfing rates across generations, no migration, and an absence of natural selection with respect to the level of individual heterozygosity (i.e., no inbreeding depression, no heterosis). The overall genotypic associations of individuals were displayed with a factorial correspondence analysis (FCA) using the procedure implemented in GENETIX (ver. 4.04; Belkhir et al. 2003). The program STRUCTURE (ver. 2.3.4; Falush et al. 2003) was used to assign individual fish to a specified number (K) of clusters (presumed populations). The assignment is probabilistic, so that an individual may have joint membership in multiple populations, with membership coefficients summing to one. STRUCTURE was run under an admixture model assuming correlated allele frequencies for K ranging from 1 to 10. Ten independent chains were run for each value of K, each chain consisting of 50 000 burn-in iterations and 100 000 MCMC iterations. The method by Evanno et al. (2005) was used to determine the most likely value of K, using the Structure Harvester web service (http://taylor0.biology.ucla.edu/structureHarvester/; Earl and vonHoldt 2012). STRUCTURE in its original formulation was designed to infer subdivision of randomly mating populations. As *K. marmoratus* may violate model assumptions in STRUCTURE we also used the program INSTRUCT (Gao et al. 2007) to repeat the clustering analyses. INSTRUCT is an extension of STRUCTURE that estimates selfing rates and takes them into account when inferring population structure. We ran INSTRUCT as described above for STRUCTURE, except that each independent chain had 100 000 burn-in iterations. The Deviance Information Criterion calculated in INSTRUCT was used to choose the appropriate K. Graphical output of the results of clustering was produced with help from the program DISTRUCT (Rosenberg et al. 2004). INSTRUCT was also used to estimate the number of generations of selfing in the genealogy of the individual since an outcrossing event last took place.

Results

Variation Within Populations

A mean of 40.5 individuals was genotyped per population in Central America (Table 1). Gene diversity (expected heterozygosity, H_E) ranged from 0.426 in SAND to 0.688 in TC; observed heterozygosity (H_O) varied from 0.005 in HON to 0.520 in TC. All populations showed a significant deficiency of heterozygosity, expressed as positive F_{IS}. Taking into account confidence intervals of inbreeding coefficients and corresponding rates of selfing (Figure 1), 3 broad population groups can be distinguished: 1) HON with a high F_{IS} (=0.991, corresponding to a selfing rate of $s = 0.995$); 2) Calabash Caye (TA), Long Caye (BW and CROB), and Sandbore Caye (SAND) with

intermediate F_{IS} values (range 0.527–0.615, corresponding to selfing rates s = 0.690–0.762); and 3) Twin Cays (TC) with low F_{IS} (=0.245, corresponding to s = 0.394). Maximum likelihood estimates of selfing s_{ml} were somewhat lower than s (Table 1), but the differences were not statistically significant in 5 populations where tests were possible. s_{ml} could not be calculated for the HON population due to an insufficient number of usable loci, because RMES excludes variable loci without heterozygotes. In comparison, 26 loci variable in HON ensured firm estimate of s based on heterozygote deficiency in this highly selfing population. A test of heterogeneity of selfing rates confirmed that TC had a significantly lower rate of selfing than the other studied populations in Central America. Differences in selfing rates s_{ml} were not statistically significant among TA, BW, CROB, and SAND.

Only 3 pairs of isogenotypes were found in this study: 1 pair each at CROB, SAND, and HON. With the exception of these 3 pairs of genetically identical fish, all other fish in our samples were distinct from one another, having at least 11 different alleles (among 64 total) and some pairs of individuals having 54 different alleles (Supplementary Figure S1 online).

The new or updated samples from Florida had variation typical for that region, with high H_E (SOB: 0.495, CRWL: 0.448), low H_O (SOB: 0.023, CRWL: 0.028), and high s (SOB: 0.978, CRWL: 0.968).

Distribution of Individual Heterozygosity

Figure 2 plots the observed distribution of individual heterozygosities within each of the 6 Central American populations. Also shown are the expected distributions of individual heterozygosity under a mixed-mating model given the particular levels of gene diversity and rates of selfing that characterize each sample. In these expected curves, the distinct peaks correspond to classes of individuals resulting from outcrossing (rightmost peak), and various generations of selfing (peaks to the left).

With the number of loci used in our study and given the level of variation, it was generally possible to visually assign each fish to 1 of 4 classes produced by: outcrossing (F_O), first (F_1), and second (F_2) generations of selfing, and 3 or more (F_{3+}) generations of selfing. For high selfing rates, the observed and expected distributions are skewed to the left, such that many individuals have much lower H_O than H_E. In the extreme (at s = 1), all individuals in a population are completely homozygous (barring de novo mutation). The HON sample (s = 0.995) closely approximates this outcome, with nearly all fish being homozygous at 32 loci and the remaining few fish being heterozygous at only one locus. On the other hand, given that gene diversity in HON is H_E = 0.497, under random mating these fish should have been heterozygous at 16 loci on average.

In the Belizean populations, individual heterozygosities were generally dispersed (Figure 2), with some fish being heterozygous at up to 27 loci and others being completely homozygous. As selfing rates decrease, the expected distributions of heterozygosity become less skewed (as in BW), and may even become skewed to the right (as in TC) with many individuals having high heterozygosities corresponding to expectations under random mating (generation F_O). In populations from Calabash, Long, and Sandbore Cayes, (s = 0.69–0.76), at least 10% of the fish reached complete homozygosity, meaning that particular lineages apparently had passed through at least 4–5 generations of selfing. In Twin Cays (s = 0.39), not a single fish reached complete homozygosity. On the opposite end of the heterozygosity distribution for this locality, 37 fish displayed heterozygosities in the portion of the curve expected under random mating (rightmost peak), meaning that at least 63% of the fish in Twin Cays had resulted from outcrossing. This rough visual method for evaluating the number of selfing generations based on individual heterozygosity broadly corresponds to estimates produced by the program INSTRUCT, when the number of selfing generations is small (1–2), but it fails to correctly estimate larger numbers of selfing generations because resulting heterozygosities are lumped into single peak (F_{3+}). Numbers of selfing generations as estimated by INSTRUCT are shown on Supplementary Figure S2 online. The estimates of the number of selfing generations were obtained at K = 5, which was determined as the most likely number of populations by Evanno's method (see later). The program INSTRUCT estimates that nearly 63% of fish in TC resulted from outcrossing; corresponding figures of F_O generation ranged from 18% to 30% in 4 Belizean populations that had similar selfing rates (TA, CROB, BW, and SAND), and there were no fish produced by outcrossing in HON. INSTRUCT confirmed that no fish in TC selfed for more than 3 generations, whereas about 18–43% of fish in the other populations from Belize were a result of 4 or more generations of selfing. The fully homozygous fish in these populations were produced by 6–7 generations of selfing. In HON selfing is very high (>99%); according to INSTRUCT, homozygous fish sampled in this population had their last ancestor generated by outcrossing 22–23 generations back, and 3 fish that were heterozygous at 1 locus trace back to outcrossed ancestors for 7 generations of selfing (Figure 2, Supplementary Figure S2 online).

Table 1. Summary of genetic variation in *Kryptolebias marmoratus* from Central America

Sites	Location	Coordinates	N	P_{95}	A	A_R	H_E	H_O	F_{IS}	s	s_{ml}
BW	Long Caye, Belize	17°13'04.1"N, 87°35'25.6"W	40	0.94	5.78	5.17	0.608	0.236	0.615	0.762	0.737
CROB	Long Caye, Belize	17°13'03 "N, 87°35'37"W	50	0.94	5.97	5.09	0.590	0.239	0.597	0.748	0.676
SAND	Sandbore Caye, Belize	17°27'49.7"N, 87°29'16.0"W	44	0.84	3.16	2.76	0.426	0.195	0.546	0.706	0.610
HON	Utila Is., Honduras	16°06'N, 86°56'W	20	0.81	3.97	3.97	0.497	0.005	0.991	0.995	n.d.
TC	Twin Cays, Belize	16°49'N, 88°06'W	59	0.97	9.28	7.24	0.688	0.520	0.245	0.394	0.332
TA	Calabash Caye, Belize	17°16'N, 87°48'W	30	0.94	6.19	5.63	0.591	0.282	0.527	0.690	0.644
Mean			40.5	0.91	5.72	4.98	0.567	0.246	0.512	0.677	0.612

N, sample size; P_{95}, proportion of polymorphic loci (95% criterion); A, average number of alleles; A_R, allelic richness; H_E, gene diversity; H_O, observed heterozygosity; F_{IS}, coefficient of inbreeding; s and s_{ml} are 2 independent estimates of selfing—s is based on F_{IS} and s_{ml} is found by maximizing the log-likelihood of the multilocus heterozygosity structure of the sample (see Materials and Methods section). All F_{IS} values are highly significant ($P < 0.0001$) as evaluated by randomization in FSTAT (Goudet 1995). s and s_{ml} are not significantly different in each population. Mean value of s_{ml} does not include HON population, the selfing rate of which could not be determined by this method.

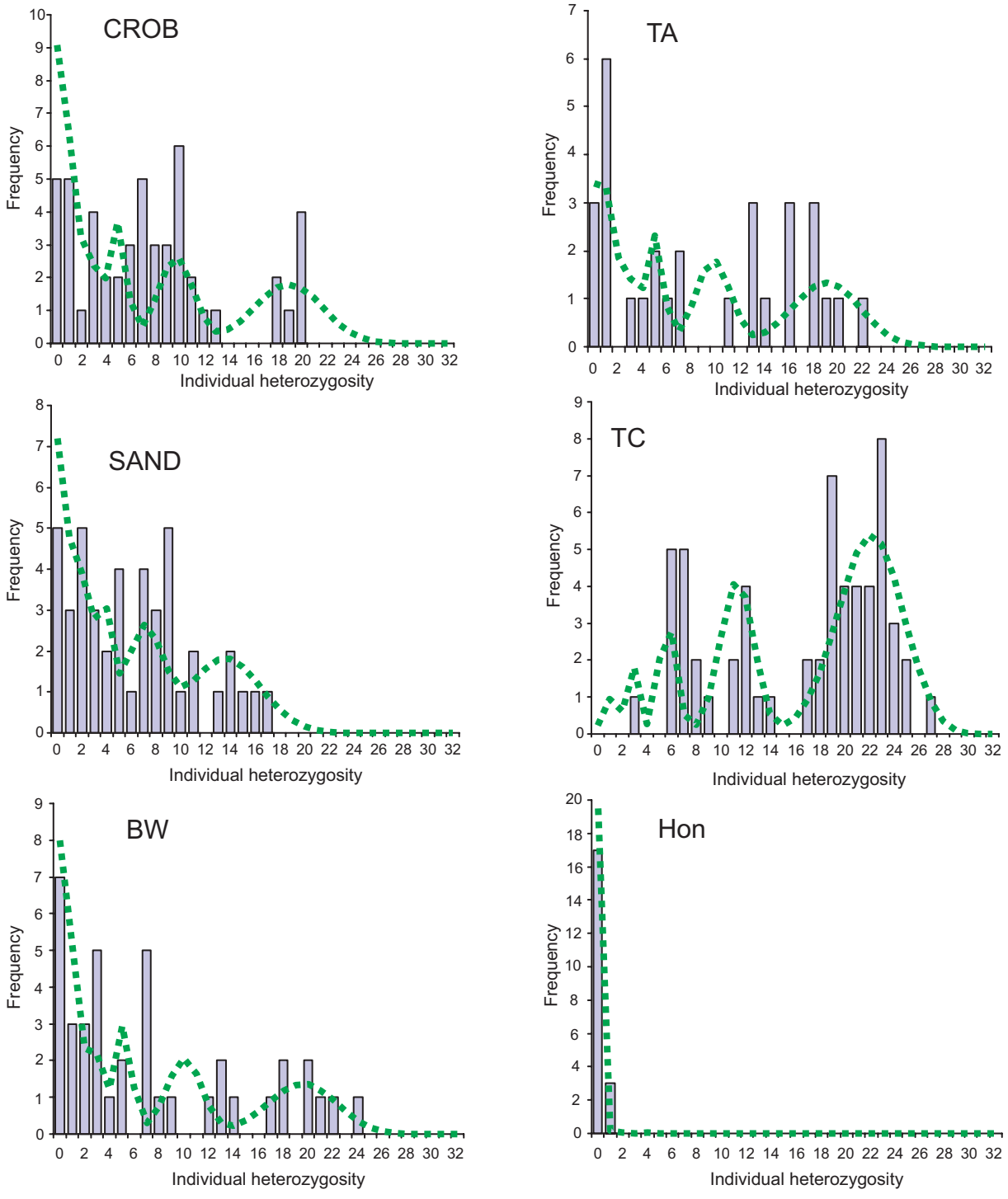

Figure 2. Histograms showing frequency distributions of individual heterozygosities in 6 Central American populations of *Kryptolebias marmoratus*. The *x* axis shows the number of heterozygous loci. Solid bars represent observed values. Dashed lines represent the distributions expected under a mixed-mating model given the particular level of heterozygosity and selfing in each population specified in Table 1.

Geographical Differentiation

Genetic differentiation between sample locations was statistically significant in each pairwise comparison as well as overall (after Bonferroni correction for multiple testing). The smallest value of F_{ST} (0.018) was between BW and CROB, which are only 400 m apart on the same mangrove island (Long Caye). Differentiation between samples from different cayes averaged F_{ST} = 0.245, but varied considerably in pairwise comparisons (range 0.126–0.430). The geographically remote HON site was genetically the most distinct from all other samples (mean F_{ST} = 0.307). Among Belizean locales, SAND was most strongly differentiated from the rest (F_{ST} = 0.227–0.289) despite its relative proximity to TA and Long Caye. Likewise, TA was genetically closer to TC than to Long Caye despite being geographically closer to the latter. Overall, there was an indication of a

weak isolation by distance in the area studied (Mantel test: $0.05 < P < 0.10$).

Analysis of genetic relationships at the individual level using 2 different approaches—FCA and clustering analysis implemented in the programs STRUCTURE and INSTRUCT—further confirmed strong population structure in *K. marmoratus* (Figure 3). In the FCA plot, 4 rather compact genetic groups or clusters were formed: HON, TA + TC, BW + CROB, and SAND. Within the mixed clusters individuals from adjacent BW and CROB broadly overlapped, whereas a majority of individuals from TA and TC were generally well separated, except for some overlap due to a penetration of a few TC individuals deep into the TA cluster. Interestingly, although nearly all fish from the SAND locality formed a tight genetic cluster (Figure 3), 1 individual (SAND24) from this site appeared in the middle of the BW + CROB genetic group, thus suggesting a recent migration event from Long Caye to Sandbore. A closer inspection of SAND24 reveals that this individual is probably a first generation migrant. High H_O in SAND24 (17 of 32 loci) clearly indicates that this fish was produced by outcrossing between 2 distinct lineages. If SAND24 had resulted from an outcross between a local SAND fish and a migrant from another population, it would occupy an intermediate position between the clusters corresponding to those 2 populations. Instead, SAND24 is firmly embedded within the BW + CROB cluster. This genetic evidence points to this specimen as being a first generation outcross migrant to SAND, rather than the descendent of a migrant.

Application of Evanno's method to STRUCTURE analyses suggested that the most probable number of clusters is $K = 5$. Under this configuration, individuals are firmly assigned to only one population (Figure 3B), thus grouping fish of the same islands together with high probability. The only exception was fish SAND24, which, although collected from Sandbore Caye, was strongly assigned to Long Caye. This confirms results of FC analysis and indicates that SAND24 is a migrant. Noteworthy, STRUCTURE analysis (same as FCA) was unable to distinguish between fish from 2 locations (BW and CROB) of the same island; application of higher K values resulted either in individuals of mixed membership or in strong assignments that did not coincide with geographic separation between BW and CROB. On the other hand, STRUCTURE analysis reliably distinguished between fish from TA and TC, whereas separation of these fish in FCA was less conclusive. Clustering analysis using INSTRUCT produced virtually indistinguishable clusters from those of STRUCTURE (Supplementary Figure S3 online); 5 clusters were apparent, each cluster corresponding to a single caye population. Fish SAND24 was confirmed as a migrant from Long Caye (or that area).

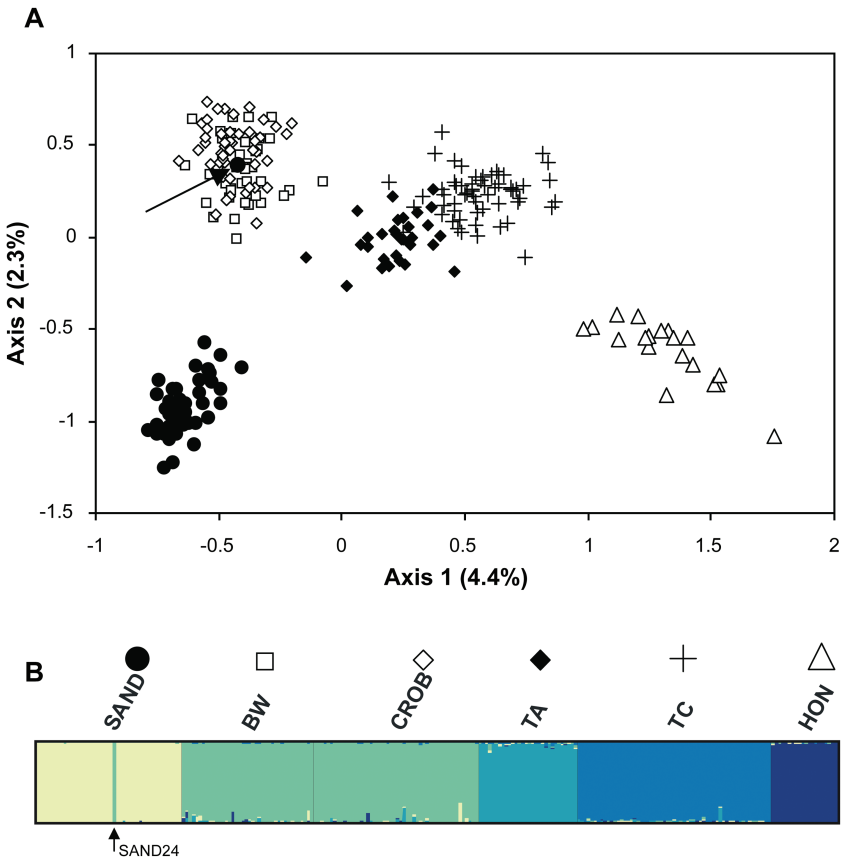

Figure 3. Genotypic association of of individual *Kryptolebias marmoratus* from 6 local populations in Central America. (**A**) FCA showing genetic relationships among killifish. Axis 1 and 2 are first and second principle factors of variability that explain the shown percentages of total genetic variation. (**B**) Genotypic clustering of the mangrove rivulus as assessed by STRUCTURE for $K = 5$. Each killifish specimen is represented by a thin bar, often partitioned into coloured segments each representing an individual's proportionate genetic membership in a given *K*th cluster. Black lines separate geographical locations of the samples, which are shown at the top (coded as in Table 1). In both panels the arrow points to fish SAND24, which was captured in Sandbore Caye but was genetically affiliated with the Long Caye population, from which it probably migrated.

Discussion

We found high variation in rates of selfing/outcrossing in *K. marmoratus* from Central America. Selfing rates ranged from $s = 0.99$ in Honduras, to $s = 0.39$–0.76 across several proximate sites in Belize, all separated at most by 160 km (Figure 1). This variation in selfing rates between surveyed sites in Central America is much higher than the variation previously reported in Florida, where all populations showed $s > 0.90$. The higher selfing rates in Florida may be advantageous for mangrove rivulus as the ability to self ensures reproduction in the absence of mating partner, which is important in low-density marginal populations (Kawecki 2008).

As a result of intense inbreeding, most fish in Florida were homozygous across all loci (Tatarenkov et al. 2012) and only rarely did a fish have 2 parents from an outcross event. In Central America, only the Honduran population was similar to Florida in this regard. Observed and expected (under a mixed-mating model) distributions of individual heterozygosity generally showed good correspondence in all of the surveyed populations in Central America (Figure 2). This suggests that despite the spatial heterogeneity in selfing rates among populations, selfing rates *within* populations were relatively stable through time. Constancy of the selfing rates was previously demonstrated in Twin Cays, where samples from the same population 14 years apart yielded similar estimates of s (Tatarenkov et al. 2007). The strong correspondence between observed and expected distributions of heterozygosity also indicates that the probabilities of reproducing either by selfing or by outcrossing are approximately equal for each fish in a given population. In other words, there is no evidence that some subset of fish is more likely to reproduce by selfing while other fish are inherently far more likely to outcross consistently.

Outcrossing in *K. marmoratus* is thought to be mediated by males; matings between hermaphrodites have never been described in the lab or in nature (Sakakura and Noakes 2000), similar to what has been also reported for other androdioecious animals such as nematodes and branchiopods (Weeks et al. 2006). Thus, we ask whether an empirical relationship exists between the abundance of males and outcrossing rates across populations. We do indeed find a broad correspondence for our Floridian and Central American locales (Figure 4). However, much of the trend was due to the presence of the TC sample, in which the outcrossing rate and male frequency were both exceptionally high. If the TC population is ignored for a moment, then the overall relationship becomes somewhat vague. For example, males comprised only 2.5% of the population in Calabash Caye, yet outcrossing there was about $t = 0.30$. In comparison, the frequency of males in Honduras (2.1%) was only slightly lower than in Calabash, yet the outcrossing rate in Honduras was $t < 0.01$. The absence of a more robust relationship between male frequency and outcrossing rate suggests that other factors may be involved, lending credence to the possibility that hermaphrodites can mate with each other (Mackiewicz et al. 2006c), particularly given the fact that spermatozoons in the mangrove rivulus have morphological structure typical of externally fertilizing fishes (Kweon et al. 1998). Another possibility is that a certain "threshold" number of males is necessary to significantly drive rates of outcrossing.

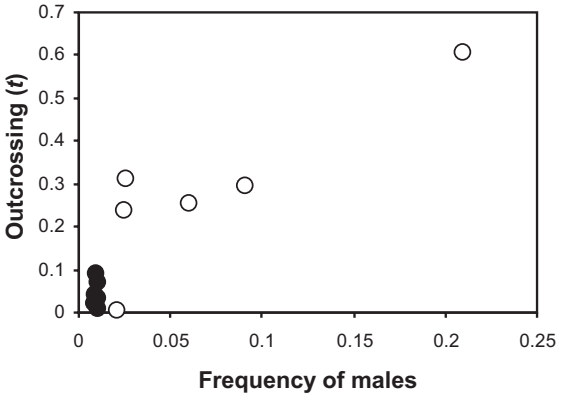

Figure 4. Overall relationship between the incidence of males and outcrossing rates in Central American (open circles) and Floridian (closed circles) populations of *Kryptolebias marmoratus*.

An even more puzzling question is what determines the frequency of males. Twin Cays is not far from the other studied Belizean cays and ostensibly has a similar environment, yet the abundance of males on this mangrove island is much higher than at the other sites. In a common garden experiment, Turner et al. (2006) showed that the production of males in *K. marmoratus* is influenced by both genetics and environment, but much further study is needed to clarify this issue.

In our study, the distances between local collections varied from about 400 m, between 2 sites in Long Caye, to 160 km, between Utila in Honduras and Sandbore Caye in Belize. The pairwise value of genetic divergence between 2 sites belonging to the same cay yielded a low but statistically significant value of $F_{ST} = 0.018$. Ellison et al. (2012) reported a somewhat higher level of differentiation (F_{ST} up to 0.075) between sites from Calabash Caye that were separated by distances of up to 1 km, but if we disregard 1 sample in Ellison's dataset that had very few individuals, then the earlier estimate becomes approximately $F_{ST} = 0.025$, which is quite similar to the F_{ST} values at this spatial scale in our current study of Central American sites. By contrast, differentiation between populations on different cays was pronounced in Central America, reaching $F_{ST} = 0.43$ in some comparisons and being $F_{ST} = 0.25$ on average. Not surprisingly, there was pronounced genetic differentiation between Central America and the Florida Keys, as evaluated by a hierarchical F_{ST} approach (Supplementary Table S1 online). Overall F_{ST} was 0.335, with about equal amounts being attributable to "differences between regions" and to "differences between sites within regions." Differentiation at all levels was highly significant statistically.

For surveyed sites in Central America, we detected only a mild correspondence between genetic divergence and geographic separation ($P < 0.1$). Still, 47% of the variance in genetic differentiation could be attributed to the variance in geographic distance between samples, and F_{ST} increased by a value of 0.20 for every 100 km of spatial separation (Figure 5). Comparable analyses for the Florida Keys, excluding 1 artificial population from a drainage ditch, similarly yielded evidence for mild isolation by distance in *K. marmoratus* (Mantel test $P < 0.05$, Figure 5).

Theory predicts that inbreeding facilitates the formation of population structure because genetic drift is greater when N_e is small (Maruyama and Tachida 1992; Charlesworth and Pannell 2001). This relationship has generally been confirmed experimentally in several species of plants (Hamrick and Godt 1996) and animals (Jarne 1995). Comparison of population structure in *K. marmoratus* in 2 areas, Florida and Central America, with distinct selfing rates does not follow the prediction. In fact, genetic differentiation among Central American populations, which have lower selfing rates, is slightly higher than in Florida Keys at the same geographical distances (Figure 5); this relationship holds even if HON (population with high selfing rate) is not considered. Taking into account that population densities of mangrove rivulus in Central America are not

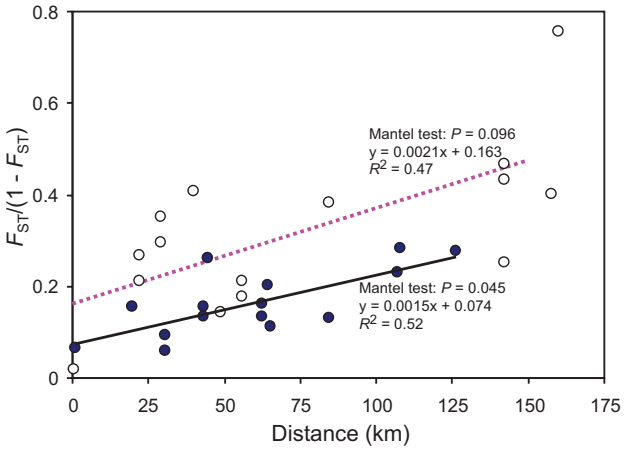

Figure 5. Relationships of genetic differentiation and geographic distance between populations of *Kryptolebias marmoratus* in Central America (open circles, dashed line) and Florida Keys (closed circles, solid line).

lower than in Florida (pers. obs.), such a pattern of differentiation suggests that dispersal in Central America may be hampered by some barriers. A possible impediment could be very deep and relatively wide channels with strong currents that separate reef ridges upholding cayes of our study. In contrast, the Florida Keys are surrounded by shallow waters and form a nearly continuous shoreline, with the Florida Current along it enhancing connectivity. A systematic survey that includes samples within and between distinct reef systems is necessary to better understand dispersal and gene flow in *K. marmoratus*.

Supplementary Material

Supplementary material can be found at http://www.jhered.oxfordjournals.org/

Funding

Smithsonian Institution, Caribbean Coral Reef Ecosystem (CCRE) Program and University of California at Irvine.

Acknowledgments

The authors thank the Keys Marine Laboratory (Long Key, FL), Blake Ross and Sarah Edwards (Lighthouse Reef, Long Caye, Belize); Phillip Hughes (National Key Deer Refuge, Big Pine Key, FL); and Yvonne Wielhouwer for logistical support. This work was approved by FFWCC Special Activity License SAL-09-1132-SR; Florida State Parks Permit 03071220; National Key Deer Refuge Permit 2010-008; and Belize Fisheries Permit #000039-11.

References

Andersen EC, Gerke JP, Shapiro JA, Crissman JR, Ghosh R, Bloom JS, Félix MA, Kruglyak L. 2012. Chromosome-scale selective sweeps shape *Caenorhabditis elegans* genomic diversity. *Nat Genet*. 44:285–290.

Avise JC. 2011. *Hermaphroditism: the biology, ecology, and evolution of dual sexuality*. New York (NY): Columbia University Press.

Baker CS. 2013. Journal of heredity adopts joint data archiving policy. *J Hered*. 104:1.

Baker HG. 1955. Self-compatibility and establishment after "long-distance" dispersal. *Evolution*. 9:347–349.

Belkhir K, Borsa P, Chikhi L, Raufaste N, Bonhomme F. 2003. *GENETIX, logiciel sous WindowsTM pour la génétique des populations*. Montpellier (France): Laboratoire Génome, Populations, Interactions CNRS UMR 5000, Université de Montpellier II.

Charlesworth D. 2003. Effects of inbreeding on the genetic diversity of populations. *Philos Trans R Soc Lond B Biol Sci*. 358:1051–1070.

Charlesworth D, Pannell JR. 2001. Mating system and population structure in the light of coalescence theory. In: Silvertown J, Antonovics J, editors. *Integrating ecology and evolution in a spatial context*. London: Blackwell Scientific. p. 73–95.

Costa WJEM. 2011. Identity of *Rivulus ocellatus* and a new name for a hermaphroditic species of *Kryptolebias* from southeastern Brazil (Cyprinodontiformes: Rivulidae) *Ichthyol Explor Freshwaters*. 22:185–192.

Cutter AD. 2006. Nucleotide polymorphism and linkage disequilibrium in wild populations of the partial selfer *Caenorhabditis elegans*. *Genetics*. 172:171–184.

David P, Pujol B, Viard F, Castella V, Goudet J. 2007. Reliable selfing rate estimates from imperfect population genetic data. *Mol Ecol*. 16:2474–2487.

Dieringer D, Schlötterer C. 2003. Microsatellite analyser (MSA): a platform independent analysis tool for large microsatellite data sets. *Mol Ecol Notes*. 3:167–169.

Earl DA, vonHoldt BM. 2012. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conserv Genet Resour*. 4:359–361.

Eckert CG, Samis KE, Lougheed SC. 2008. Genetic variation across species' geographical ranges: the central-marginal hypothesis and beyond. *Mol Ecol*. 17:1170–1188.

Ellison A, Cable J, Consuegra S. 2011. Best of both worlds? Association between outcrossing and parasite loads in a selfing fish. *Evolution*. 65:3021–3026.

Ellison A, Wright P, Taylor DS, Cooper C, Regan K, Currie S, Consuegra S. 2012. Environmental diel variation, parasite loads, and local population structuring of a mixed-mating mangrove fish. *Ecol Evol*. 2:1682–1695.

Evanno G, Regnaut S, Goudet J. 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol Ecol*. 14:2611–2620.

Falush D, Stephens M, Pritchard JK. 2003. Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. *Genetics*. 164:1567–1587.

Gao H, Williamson S, Bustamante CD. 2007. A Markov chain Monte Carlo approach for joint inference of population structure and inbreeding rates from multilocus genotype data. *Genetics*. 176:1635–1651.

Goudet J. 1995. FSTAT (version 1.2): a computer program to calculate F-statistics. *J. Hered*. 86:485–486.

Goudet J. 2005. HIERFSTAT, a package for R to compute and test hierarchical F-statistics. *Mol Ecol Notes*. 5:184–186.

Hamrick JL, Godt MJW. 1996. Effects of life history traits on genetic diversity in plant species. *Philos Trans R Soc Lond B-Biol Sci*. 351:1291–1298.

Harrington RW. 1961. Oviparous hermaphroditic fish with internal self-fertilization. *Science*. 134:1749–1750.

Harrington, R.W, Kallman, K.D. 1968. The homozygosity of clones of the self-fertilizing hermaphrodite fish *Rivulus marmoratus* Poey (Cyprinodontidae, Atheriniformes). *Am. Nat*. 102: 337–343.

Hedrick PW. 2005. *Genetics of populations, 3rd ed*. Boston (MA): Jones & Bartlett.

Ingvarsson PK. 2002. A metapopulation perspective on genetic diversity and differentiation in partially self-fertilizing plants. *Evolution*. 56:2368–2373.

Jarne P. 1995. Mating system, bottlenecks and genetic polymorphism in hermaphroditic animals. *Genet Res*. 65:193–207.

Jarne P, Auld JR. 2006. Animals mix it up too: the distribution of self-fertilization among hermaphroditic animals. *Evolution*. 60:1816–1824.

Kawecki TJ. 2008. Adaptation to marginal habitats. *Ann Rev Ecol Evol Syst*. 39:321–342.

Kweon HS, Park EH, Peters N. 1998. Spermatozoon ultrastructure in the internally self-fertilizing hermaphroditic teleost, *Rivulus marmoratus* (Cyprinodontiformes, Rivulidae). *Copeia*. 1998:1101–1106.

Lewis PO, Zaykin D. 2001. Genetic data analysis: computer program for the analysis of allelic data. [cited November 9 2014]. Available from: http://hydrodictyon.eeb.uconn.edu/people/plewis/software.php

Lubinski BA, Davis WP, Taylor DS, Turner BJ. 1995. Outcrossing in a natural population of a self-fertilizing hermaphroditic fish. *J Hered*. 86:469–473.

Mackiewicz M, Tatarenkov A, Perry A, Martin JR, Elder JF Jr, Bechler DL, Avise JC. 2006a. Microsatellite documentation of male-mediated outcrossing between inbred laboratory strains of the self-fertilizing mangrove killifish (*Kryptolebias Marmoratus*). *J Hered*. 97:508–513.

Mackiewicz M, Tatarenkov A, Turner BJ, Avise JC. 2006b. A mixed-mating strategy in a hermaphroditic vertebrate. *Proc R Soc B-Biol Sci*. 273:2449–2452.

Mackiewicz M, Tatarenkov A, Taylor DS, Turner BJ, Avise JC. 2006c. Extensive outcrossing and androdioecy in a vertebrate species that otherwise reproduces as a self-fertilizing hermaphrodite. *Proc Natl Acad Sci U S A*. 103:9924–9928.

Maruyama K, Tachida H. 1992. Genetic variability and geographical structure in partially selfing populations. *Jpn J Genet*. 67:39–51.

Nordborg M. 1997. Structured coalescent processes on different time scales. *Genetics*. 146:1501–1514.

Rosenberg NA. 2004. DISTRUCT: a program for the graphical display of population structure. *Mol Ecol Notes*. 4:137–138.

Rousset F. 2008. genepop'007: a complete re-implementation of the genepop software for Windows and Linux. *Mol Ecol Resour*. 8:103–106.

Sakakura Y, Noakes DLG. 2000. Age, growth, and sexual development in the self-fertilizing hermaphroditic fish *Rivulus marmoratus*. *Env Biol Fish*. 59:309–317.

Tatarenkov A, Gao H, Mackiewicz M, Taylor DS, Turner BJ, Avise JC. 2007. Strong population structure despite evidence of recent migration in a selfing hermaphroditic vertebrate, the mangrove killifish (*Kryptolebias marmoratus*). *Mol Ecol*. 16:2701–2711.

Tatarenkov A, Lima SM, Taylor DS, Avise JC. 2009. Long-term retention of self-fertilization in a fish clade. *Proc Natl Acad Sci U S A*. 106:14456–14459.

Tatarenkov A, Ring BC, Elder JF, Bechler DL, Avise JC. 2010. Genetic composition of laboratory stocks of the self-fertilizing fish *Kryptolebias marmoratus*: a valuable resource for experimental research. *PLoS One*. 5:e12863.

Tatarenkov A, Lima SM, Avise JC. 2011. Extreme homogeneity and low genetic diversity in *Kryptolebias ocellatus* from south-eastern Brazil suggest a recent foundation for this androdioecious fish population. *J Fish Biol*. 79:2095–2105.

Tatarenkov A, Earley RL, Taylor DS, Avise JC. 2012. Microevolutionary distribution of isogenicity in a self-fertilizing fish (*Kryptolebias marmoratus*) in the Florida Keys. *Integr Comp Biol*. 52:743–752.

Taylor DS. 2000. Biology and ecology of *Rivulus marmoratus*: new insights and a review. *Florida Sci*. 63:242–255.

Taylor DS. 2012. Twenty-four years in the mud: what have we learned about the natural history and ecology of the mangrove rivulus, *Kryptolebias marmoratus*? *Integr Comp Biol*. 52:724–736.

Taylor DS, Fisher MT, Turner BJ. 2001. Homozygosity and heterozygosity in three populations of *Rivulus marmoratus*. *Environ Biol Fish*. 61:455–459.

Turner BJ, Davis WP, Taylor DS. 1992a. Abundant males in populations of a selfing hermaphroditic fish—*Rivulus marmoratus*—from some Belize Cays. *J Fish Biol*. 40:307–310.

Turner BJ, Elder JF, Laughlin TF, Davis WP, Taylor DS. 1992b. Extreme clonal diversity and divergence in populations of the self-fertilizing killifish, *Kryptolebias marmoratus*. *Proc Natl Acad Sci U S A*. 89:10643–10647.

Turner BJ, Fisher MT, Taylor DS, Davis WP, Jarrett BL. 2006. Evolution of "maleness" and outcrossing in a population of the self-fertilizing killifish, *Kryptolebias marmoratus*. *Evol Ecol Res*. 8:1475–1486.

Weeks SC, Benvenuto C, Reed SK. 2006. When males and hermaphrodites coexist: a review of androdioecy in animals. *Integr Comp Biol*. 46:449–464.