Biological Journal of the Linnean Society, 2017, **XX**, 1–21. With 6 figures.

Deep and concordant subdivisions in the self-fertilizing mangrove killifishes (*Kryptolebias*) revealed by nuclear and mtDNA markers

ANDREY TATARENKOV[1]*, SERGIO M. Q. LIMA[2], RYAN L. EARLEY[3], WALDIR M. BERBEL-FILHO[2,4], FRANS B. M. VERMEULEN[5], D. SCOTT TAYLOR[6], KRISTINE MARSON[3], BRUCE J. TURNER[7] and JOHN C. AVISE[1]

[1]*Department of Ecology and Evolutionary Biology, University of California, Irvine, CA 92697, USA*
[2]*Laboratório de Ictiologia Sistemática e Evolutiva, Departamento de Botânica e Zoologia, Universidade Federal do Rio Grande do Norte, Av. Senador Salgado Filho 3000, 56078-970 Natal, RN, Brazil*
[3]*Department of Biological Sciences, University of Alabama, Tuscaloosa, AL 35487, USA*
[4]*Department of BioSciences, College of Science, Swansea University, Swansea, Wales SA2 8PP, UK*
[5]*Tanki Leendert 194 c, Noord, Aruba*
[6]*Brevard County Environmentally Endangered Lands Program, Melbourne, FL 32904, USA*
[7]*Department of Biological Sciences, Virginia Tech, Blacksburg, VA 24061, USA*

Received 6 May 2017; revised 27 July 2017; accepted for publication 28 July 2017

We use extensive geographical sampling and surveys of nuclear microsatellite and mitochondrial DNA loci to investigate the phylogeographic structure of the only recognized self-fertilizing vertebrates, the mangrove killifishes, currently thought to comprise two cryptic species, *Kryptolebias marmoratus* and *Kryptolebias hermaphroditus*. All genetic markers revealed three concordant main clades. The Northern clade includes populations from Florida, northern Cuba, Bahamas, Belize and Honduras and corresponds to *K. marmoratus*. The Southern clade encompasses populations from Brazil and corresponds to *K. hermaphroditus*. This species was considered endemic to southeastern Brazil, but molecular data corroborate its occurrence in northeastern Brazil. The Central clade, not previously resolved with genetic data, includes populations from Panama and Antilles. Despite the geographic proximity of the Northern and Central clades, the latter is genetically closer to the Southern clade. The discovery of the Central clade raises some taxonomic issues – it can either be considered a distinct species or united with the Southern clade into a single species with two subspecies. Another possible taxonomic solution is a single selfing species, *K. marmoratus*, with three subspecies. We show that the Central and Southern clades are highly selfing (97–100%), whereas selfing rates of the Northern clade populations vary geographically (39–99%). Genetic patterns indicate that populations in SE Brazil are recent, contrary to expectations based on the known distributions of related species.

ADDITIONAL KEYWORDS: Caribbean – cryptic species – Cynolebiidae – mangrove rivulus – microsatellites – selfing – species concept – species delimitation.

INTRODUCTION

The mangrove ecosystem of the western Atlantic is extensive, occupying 3.2 million ha along the coasts of the Americas (Luther & Greenberg, 2009). At the same time, mangroves are a vulnerable ecosystem, declining at a faster rate than inland tropical forests and coral reefs (Giri *et al.*, 2011). Despite their prominence, there have been few genetic studies of mangrove representatives, compared, for example, to the study of organisms associated with coral reefs (Diaz-Ferguson *et al.*, 2010).

One of the few species exclusively associated with mangals along the Atlantic shores of the Americas is mangrove rivulus killifish *Kryptolebias marmoratus* (or, to be more precise, the species complex '*K. marmoratus*'). Locally, *K. marmoratus* occupies habitat formed by the red mangrove *Rhizophora mangle*, and at a broad geographic range, the northern and southern

*Corresponding author. E-mail: tatarenk@uci.edu

limits of these species nearly coincide, stretching from central Florida (29°N) to São Paulo state in Brazil (23°S) (Taylor, 2012; Costa, 2016). Such tight association suggests that contemporary population and phylogeographic structure of the killifish was influenced by past diversification and distribution of the red mangrove. Therefore, patterns of the population structure of the mangrove rivulus may parallel with those of the red mangrove and of the other mangrove dwellers, and thus be broadly informative about the recent evolution of the mangrove ecosystem.

The mangrove rivulus, *K.marmoratus* (Cynolebiidae), is a small fish that is best known as the world's only self-fertilizing hermaphroditic vertebrate (Harrington, 1961; Avise, 2008). Indeed, the only other known selfing vertebrate (*Kryptolebias hermaphroditus*; Costa, 2011) has recently been recognized as a distinct species. *Kryptolebias marmoratus* and *K. hermaphroditus* are considered members of *K. marmoratus* species complex, together with a third species, *Kryptolebias ocellatus*, better known under its junior synonym *Kryptolebias caudomarginatus* (Costa, 2011; Avise & Tatarenkov, 2015). Unlike the former two species, *K. ocellatus* apparently does not reproduce by self-fertilization, even though its populations consist of hermaphrodites and males (Tatarenkov *et al.*, 2009; Costa, Lima & Bartolette, 2010). Of the three species, *K. ocellatus* is the most diverged (Murphy, Thomerson & Collier, 1999; Kanamori *et al.*, 2016). All other known *Kryptolebias* species are phylogenetically more distant and are dioecious, implying that hermaphroditism arose in the common ancestor of all three species of the *K. marmoratus* species complex, with selfing evolving later on in the *marmoratus*/*hermaphroditus* common ancestor (Costa *et al.*, 2010).

Taxonomy of the *K. marmoratus* species complex is rather confused. *Kryptolebias ocellatus* was described by Hensel (1868) using a fish from Rio de Janeiro, Brazil. *Kryptolebias marmoratus* was described a decade later (Poey, 1880) based on the specimens from Cuba or the USA. For nearly a hundred years, mention of these species was largely limited to the taxonomic literature, but this changed with the discovery of self-fertilization in *K. marmoratus* (Harrington, 1961), which quickly brought this species into the limelight. Cassel (1982) reported a selfing hermaphroditic species from southern Brazil and proposed that it was in fact *K. marmoratus*. Seegers (1984) agreed that Brazilian and Caribbean selfing fish were the same species but, due to misidentification, incorrectly applied to it the name *ocellatus* (Costa, 2011). In the same study, Seegers (1984) redescribed *ocellatus* as *caudomarginatus*. Since Seegers' (1984) study, the name *ocellatus* became firmly (but incorrectly, according to Costa, 2011) associated with a self-fertilizing hermaphroditic fish from Brazil, until it was described as *K. hermaphroditus* (Costa, 2011).

Before Costa's proposal that *K. hermaphroditus* and *K. marmoratus* were distinct species (Costa, 2006, 2011), many authors considered selfing *Kryptolebias* a single species and referred to it either as *K. ocellatus* (*sensu* Seegers), or *K. marmoratus*, or sometimes by trinomens such as *K. ocellatus marmoratus* (Goodwin & Grizzle, 1994). Precise geographic ranges of *K. marmoratus* and *K. hermaphroditus* are not known, but the combined range is broad: peninsular Florida, most Caribbean islands, including the Bahamas, and the Atlantic coast from Yucatan to southeastern Brazil (Taylor, 2000, 2012; Costa *et al.*, 2010; Tatarenkov, Lima & Avise, 2011; Lira *et al.*, 2015). This distribution is assumed to be continuous, but in fact there are major gaps in collections. Records east and south of Venezuela are extremely sparse (Huber, 1992). In Brazil, until recently, there were no reports of *K. hermaphroditus* to the north of Rio de Janeiro, prompting Costa to describe it as endemic to southeastern Brazil (Costa, 2011). However, this situation changed with recent findings of putative *K. hermaphroditus* in the Brazilian states of Espírito Santo, Alagoas, Rio Grande do Norte and Pará (Sarmento-Soares *et al.*, 2014; Lira *et al.*, 2015; Guimarães-Costa, Schneider & Sampaio, 2017).

Kryptolebias marmoratus was thought to occupy most of the Caribbean, from Florida to Guianas (Costa, 2006; Taylor, 2012), but Tatarenkov *et al.* (2010) found that laboratory lineages originating from single fishes caught in southern Cuba and Panama were genetically closer to *K. hermaphroditus* than they were to *K. marmoratus*, calling for a major reassessment of the *K. marmoratus* complex distribution and taxonomy (Tatarenkov *et al.*, 2011). Furthermore, at various times, several species and subspecies were described in the Caribbean (such as *Rivulus heyei* Nichols, *Rivulus marmoratus bonairensis* Hoedeman and *Rivulus garciai* de la Cruz & Dubitsky), which are currently synonymized with *K. marmoratus*, but in the absence of genetic comparisons, their true affinities remain uncertain. Intriguingly, Costa (2016) recorded the presence of *K. marmoratus* in Rio Grande do Norte and Alagoas, in the same areas where Lira *et al.* (2015) collected *K. hermaphroditus*. However, as Costa (2016) acknowledges himself, colour patterns used to distinguish two species are difficult to analyse and may be subjective. Clearly, in the absence of reliable morphological characters, a genetic confirmation is necessary.

Detailed population genetic studies of the *K. marmoratus* species group have so far been primarily limited to natural populations from Florida and Belize, with other areas – such as Bahamas and Brazil – represented poorly, by one to two populations each (Mackiewicz *et al.*, 2006b, c; Tatarenkov *et al.*, 2007, 2009, 2011, 2012, 2015; Ellison *et al.*, 2012).

During the last several years, we and our colleagues embarked on multiple collection trips in the Caribbean and South America. Collecting in mangrove forests is inherently difficult, and this difficulty is amplified when looking for small, inconspicuous fish that hide most of the time in the muddy waters of crab burrows. Our surveys resulted in samples covering most of the geographic distribution of the *K. marmoratus* species complex, including several geographic areas where these species have not previously been reported. Identification of species in the *K. marmoratus* complex is difficult using morphological characters (Taylor, 2003), but the three valid species of the complex (*K. ocellatus*, *K. hermaphroditus* and *K. marmoratus*) are easily distinguishable with genetic markers such as microsatellites or mitochondrial DNA (mtDNA) (Tatarenkov *et al.*, 2009, 2011). Therefore, instead of trying to fit our collections into a particular species based on morphology, we chose to first determine major genetic groupings of the selfing mangrove rivulus and only then evaluate how these groups agree with current taxonomy.

Our objectives are to (1) describe population genetic structure of the selfing species of the *K. marmoratus* complex over its whole geographic distribution and define main genetic lineages; (2) map the geographic distributions of the main genetic lineages; (3) find out whether and how well the main genetic lineages agree with current taxonomy; (4) verify the status of specimens from NE Brazil that were a matter of some controversy; (5) estimate selfing/outcrossing rates of newly sampled populations that have not been studied before; and (6) describe major trends in genetic diversity (heterozygosity, allelic richness) that could be informative about the origin and evolution of the *K. marmoratus* species complex.

MATERIAL AND METHODS

Samples

We used 33 microsatellite loci and a mitochondrial gene *cox1* (encoding subunit I of cytochrome oxidase) to determine major genetic lineages and population structure of the *K. marmoratus* species complex over its entire geographic distribution. Microsatellites were analysed in a total of 734 specimens of *K. marmoratus* species complex from 33 localities in the Caribbean and Brazil, stretching from central Florida (29°N) to southern Brazil (23°S) (Fig. 1; Table 1; Supporting Information, Fig. S1). Sixteen populations were sampled for the first time (N = 316), whereas the rest of the microsatellite data were taken from our previous publications (Table 1). *cox1* was sequenced in a subsample of this material (N = 119), representing main geographic areas, as well as the outgroup *K. ocellatus*. Due to difficulties in collecting mangrove rivulus, some populations are represented by a single or only a few individuals; whereas such samples provide little insight about genetic diversity, they are useful for identifying major lineages and their distributions. For this purpose, we also included literature material: one specimen from Guanabo River (N. Cuba) and another specimen from Turks and Caicos; for these two individuals, only mtDNA sequences are available. Therefore, altogether 35 geographic localities are considered in this study. Fish were captured from temporary pools or from the burrows of great land crabs (*Cardisoma guanhumi*) or mangrove land crabs (*Ucides cordatus*) using cup traps, wire minnow traps, miniature hook and line, or dip nets.

During our collections, the sex of fish (hermaphrodite vs. male) was assessed by body coloration and the presence/absence of a black ocellus on the caudal fin; males are orange in colour and lack the caudal ocellus or, in some cases, display a faded ocellus. It is thought that hermaphrodites do not mate with each other, and outcrossing results from matings between males and hermaphrodites (Furness, Tatarenkov & Avise, 2015).

Molecular markers

Nuclear genetic markers used in this study were from a set of 33 microsatellite loci developed for *K. marmoratus* (Mackiewicz *et al.*, 2006a). DNA preparations, genotyping protocol and binning of alleles followed Tatarenkov *et al.* (2010, 2012).

A 704 bp region of the mitochondrial cytochrome oxidase subunit I gene (*cox1*) was amplified with FishCOI-F (5′-TCAACYAATCAYAAAGACATYGGCAC-3′) and FishCOI-R (5′-ACTTCYGGGTGTCCRAARAAYCA-3′) primers. These primers map to regions commonly used in barcoding studies of fishes (e.g. Ward *et al.*, 2005; Weigt *et al.*, 2012), but were optimized by us to better accommodate variation in Cynolebiidae. PCR amplification was conducted under the following conditions: initial denaturation at 95 °C for 5 min; 30 cycles of denaturation at 95 °C for 40 s, annealing at 52 °C for 40 s and extension at 72 °C for 1 min; and a final extension step of 7 min at 72 °C. Both DNA strands were sequenced, and the forward and reverse sequences were assembled and edited using software SeqMan Pro version 12.0.0 (DNAStar, Inc., http://www.dnastar.com). Trimmed 655 bp-long sequences were deposited in GenBank under accession numbers MF554974–MF555093. Two additional *cox1* sequences were obtained from GenBank (Table 1), one of them originating from Guanabo River in northern Cuba (Lara *et al.*, 2010).

Fish from Turks and Caicos was sequenced at *cytB* (MF555094; Weibel, Dowling & Turner, 1999) and thus was not directly comparable with our data. We determined the placement of this fish into the major genetic lineages by selecting 32 representative specimens from main clades on the *cox1* phylogenetic tree and sequencing

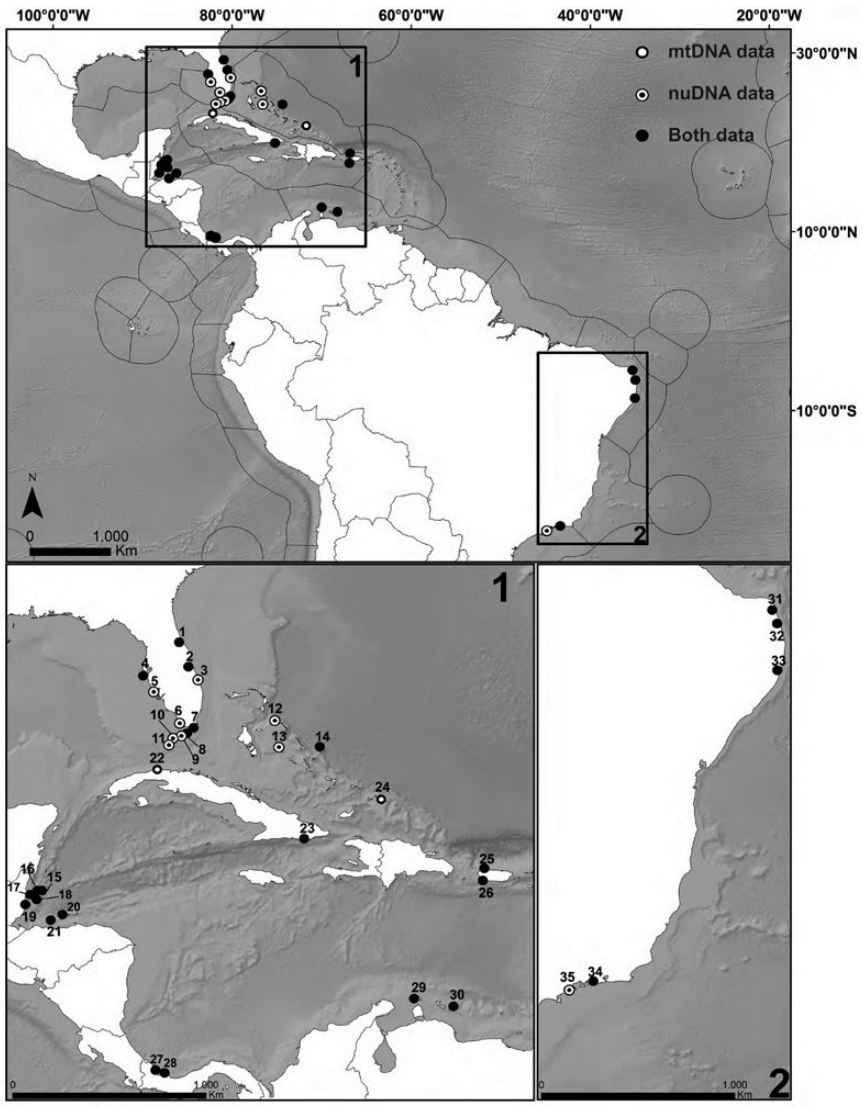

Figure 1. Sampling locations of *Kryptolebias marmoratus* species complex. The locations are labelled as in Table 1. Pattern of circles, designating the locations, indicates genetic markers – microsatellites, mtDNA or both – studied in the specific population. Thin lines outline marine ecoregions (*sensu* Spalding *et al.*, 2007).

them at *cytB* (GenBank MF554942–MF554973). Incidentally, a *cytB* sequence was also available for the above-mentioned fish from Guanabo River (Cuba) (LK022680; Ponce de Leon *et al.*, 2014) and included in our analyses (Supporting Information, Table S1; Table 1).

Statistical analyses

GDA software (Lewis & Zaykin, 2001) was used to calculate basic descriptive statistics (H_E, H_O, F_{IS}). Rates of selfing (s) and outcrossing ($t = 1 - s$) were estimated from the empirical fixation index using the equation $s = 2F_{IS}/(1 + F_{IS})$ (Wright, 1969: 195). Redelings *et al.* (2015) showed that the F_{IS}-based method produces unbiased estimates of selfing and is preferable to the maximum-likelihood method implemented in the program RMES (David *et al.*, 2007) at high selfing rates (which are typical for most *K. marmoratus* populations). Pairwise and overall estimates of population differentiation F_{ST} were computed in FSTAT (version 2.9.3.2; Goudet, 1995), and their significance was evaluated by performing 1000 permutations of genotypes among samples. Pairwise F_{ST} values were based on 30 loci that were screened in all populations. FSTAT was also used to calculate allelic richness (A_R) to control for the effect of sample size on allelic composition. Hierarchical analysis of genetic variation for microsatellite and mitochondrial data was conducted in ARLEQUIN (ver. 3.5.2.2; Excoffier & Lischer, 2010). For the microsatellite data set, matrices of pairwise

Table 1. *Kryptolebias marmoratus* species complex sampling locations and codes

Sample ID*	Location	Major area	Ecoregion	Latitude	Longitude	Msats†	*cox1*/*cytB*‡	Reference§
1. IRN^{N}	Indian River North	Florida, USA	42	29°03′23.8″N	080°56′10.8″W	22	13/0	This study
2. PC^{N}	Pepper Cove, Brevard County	Florida, USA	70	27°54′07.1″N	080°28′30.7″W	14	3/0	This study
3. NUKR^{N}	Nuclear Power Plant	Florida, USA	70	27°21′00.2″N	080°14′22.5″W	29	0/4	This study
4. TBP^{N}	Tampa Bay	Florida, USA	70	27°32′00.7″N	082°38′30.8″W	24	17/1	This study
5. CC^{N}	Charlotte County	Florida, USA	70	26°50′27.4″N	082°17′44.9″W	17		Tatarenkov *et al.* (2007)
6. SR^{N}	Shark River	Florida, USA	70	25°20′46.2″N	081°07′10.1″W	30		Tatarenkov *et al.* (2007)
7. DC^{N}	Dove Creek/Key Largo	Florida Keys, USA	70	25°01′45.6″N	080°29′49.2″W	26	4/4	Tatarenkov *et al.* (2012)
8. LK^{N}	Long Key	Florida Keys, USA	70	24°49′23.3″N	080°48′40.3″W	31	4/4	Tatarenkov *et al.* (2012)
9. CRWL^{N}	Crawl Key	Florida Keys, USA	70	24°44′55.0″N	080°58′41.2″W	21		Tatarenkov *et al.* (2012)
10. BP^{N}	Big Pine Key	Florida Keys, USA	70	24°41′48.1″N	081°20′51.3″W	40		Tatarenkov *et al.* (2012)
11. SOB^{N}	Sugarloaf Key	Florida Keys, USA	70	24°36′05.2″N	081°34′34.1″W	11		Tatarenkov *et al.* (2015)
12. NEL^{N}	North Eleuthera	Bahamas	63	25°29′27.3″N	076°39′57.1″W	14		This study
13. EI^{N}	Exuma Island	Bahamas	63	24°13′14.4″N	076°28′48.6″W	12		Tatarenkov *et al.* (2007)
14. SSN^{N}	San Salvador Island, North	Bahamas	63	24°06′49.3″N	074°27′22.1″W	41	7/1	This study
15. SAND^{N}	Sandbore Caye	Belize	68	17°27′49.7″N	087°29′16.0″W	44	4/4	Tatarenkov *et al.* (2015)
16. NC^{N}	Northern Caye	Belize	68	17°27′10.6″N	087°30′06.4″W	23	4/0	This study
17. TA^{N}	Calabash Caye, Turneffe Atoll	Belize	68	17°16′44.7″N	087°48′57.1″W	30	4/4	Tatarenkov *et al.* (2015)
18. LC^{N}	Long Caye	Belize	68	17°13′04.1″N	087°35′25.6″W	40	4/4	Tatarenkov *et al.* (2015)
19. TC^{N}	Twin Cayes	Belize	68	16°49′46.5″N	088°06′12.7″W	59	5/0	Tatarenkov *et al.* (2015)
20. HONR^{N}	Roatán Island	Honduras	68	16°23′03.2″N	086°23′05.4″W	1	1/1	Tatarenkov *et al.* (2010)
21. HONU^{N}	Utila Island	Honduras	68	16°06′00.0″N	086°56′00.0″W	20	4/4	Tatarenkov *et al.* (2015)
22. GNB^{N}	Guanabo River	Cuba	65	23°10′22.4″N	082°07′01.5″W		1/1	n/a
23. GIT^{C}	Guantanamo Bay	Cuba	65	19°54′38.6″N	075°11′44.5″W	1	1/1	Tatarenkov *et al.* (2010)
24. T&C^{C}	Turks and Caicos	Turks and Caicos	63	21°47′25.2″N	071°40′11.9″W		0/1	n/a
25. PRN^{C}	Punta Maracayo	Puerto Rico	65	18°29′28.4″N	066°47′40.3″W	1	1/0	This study
26. PRS^{C}	Puerto Rico, Southwest	Puerto Rico	65	17°56′57.0″N	066°52′06.0″W	3	3/0	This study
27. ALM^{C}	Almirante Bay	Panama	67	09°21′12.2″N	082°15′31.5″W	13	4/1	This study
28. CHI^{C}	Chiriquí Lagoon	Panama	67	09°00′40.7″N	081°47′27.1″W	28	12/0	This study
29. ARB^{C}	Parkietenbos Dumpsite	Aruba	66	12°29′31.5″N	070°00′30.8″W	8	4/0	This study
30. BNR^{C}	Lac Bay	Bonaire	66	12°06′37.7″N	068°13′02.4″W	11	5/0	This study

Table 1. *Continued*

Sample ID*	Location	Major area	Ecoregion	Latitude	Longitude	Msats†	*cox1*/*cytB*‡	Reference§
31. CEA[S]	Ceará-Mirim River, Extremoz	NE Brazil	75	05°40′25.9″S	035°14′14.5″W	38	4/0	This study
32. CUR[S]	Curimataú River, Baía Formosa	NE Brazil	75	06°21′32.5″S	035°01′54.8″W	30	4/0	This study
33. MPE[S]	Maracaípe River, Ipojuca	NE Brazil	75	08°31′06.6″S	035°00′36.8″W	17	4/0	This study
34. GUA[S]	Piracão River, Guaratiba	SE Brazil	180	23°00′06.7″S	043°34′51.6″W	10	4/2	Tatarenkov *et al.* (2011)
35. PIC[S]	Fazenda River, Picinguaba	SE Brazil	180	23°22′01.0″S	044°50′13.4″W	25		Tatarenkov *et al.* (2011)
All						734	121/37	

Major geographic area, marine ecoregion (*sensu* Spalding *et al.*, 2007), latitude and longitude are shown for each sampling locality. Localities are ordered from north to south, and sample numbers correspond to those of Figure 1. n/a, not applicable.
*Superscript letter indicates clade determined in genetic analyses: [N], Northern clade; [C], Central clade; [S], Southern clade.
†Number of specimens used in microsatellite analyses.
‡Number of mitochondrial *cox1* and *cytB* gene sequences, respectively. All *cox1* sequences are new (GenBank MF554974–MF555093), with exception of GNB (FN544255) and one TC sample (JQ840547). Sources for the *cytB* sequences are shown in Supporting Information, Table S1; GenBank accession numbers for the newly deposited sequences are MF554942–MF554973 and MF555094–MF555095.
§Original source of microsatellite data.

differences between populations (1000 bootstrap replicates) were constructed using MICROSATELLITE ANALYSER ver. 4.05 (Dieringer & Schlötterer, 2003) based on Cavalli-Sforza & Edwards (1967) chord distance. These distance matrices were processed with modules NEIGHBOR and CONSENSE of the PHYLIP software ver. 3.695 (Felsenstein, 1993) to construct a neighbor-joining (NJ) tree with bootstrap support. MICROSATELLITE ANALYSER was also used to estimate pairwise genetic differences between individuals with the D_{PS} distance metric (Bowcock *et al.*, 1994) based on the proportion of shared alleles. Values of D_{PS} can range from zero (genetic identity) to one (no shared alleles). The differences between individuals were summarized with the NJ tree constructed in PHYLIP as described above. The program STRUCTURE (version 2.3.4; Falush, Stephens & Pritchard, 2003) was used to assign individual fish to a specified number (K) of clusters (presumed populations). STRUCTURE implements a Bayesian model-based clustering method that assigns the individuals based on distinct allele frequencies. 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 various numbers of K, depending on the subset considered (see below). Ten independent chains were run for each value of K, each chain consisting of 50 000 burn-in iterations and 100 000 Markov chain Monte Carlo (MCMC) iterations. The method by Evanno, Regnaut & Goudet (2005) was utilized to determine the most likely value of K, using the STRUCTURE HARVESTER web service (http://taylor0.biology.ucla.edu/structureHarvester/; Earl & 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 verified results with the program INSTRUCT (Gao, Williamson & Bustamante, 2007). 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 clustering results was produced with the program DISTRUCT (Rosenberg, 2004).

Phylogenetic reconstruction

The phylogenetic analysis was carried out by Bayesian coalescent reconstruction using BEAST v.1.75 (Drummond *et al.*, 2012). The data set included all the *cox1* haplotypes from the *K. marmoratus* and *K. hermaphroditus*, and one individual of *K. ocellatus* as an outgroup. The appropriate model of nucleotide substitution required for the Bayesian inference (BI) analysis was selected by a hierarchical likelihood-ratio test using MEGA 6 (Tamura *et al.*, 2013). The substitution model selected was HKY. We performed 10^6 MCMC runs, sampling every 10^3 run. We assessed convergence of the MCMC runs and effective sample sizes (≥ 200) using TRACER 1.6 (Drummond & Rambaut, 2007). The first 20% of the trees were removed as the

burn-in period, and a consensus tree assessing the posteriori probability values of each clade was constructed with the TREEANNOTATOR 1.6.1 software (Drummond *et al.*, 2012). An uncorrelated lognormal relaxed molecular clock that allows rate variation among lineages was implemented using 0.009 per site per million mutation rate based on the cyprinodontiform Goodeidae *cox1* calibration adjusted for fossil record and geological data (Webb *et al.*, 2004; García *et al.*, 2012).

Phylogeographic analyses

ARLEQUIN 3.5 (Excoffier & Lischer, 2010) was used to calculate haplotype (h) and nucleotide (π) diversities, pairwise fixation indices (Φ_{ST}), and to perform neutrality tests. The divergence between sequences (p-distance) was estimated using MEGA 6 (Tamura *et al.*, 2013).

We used analyses of molecular variance (AMOVAs; 1000 permutations) to test whether the partitioning of *cox1* sequence variation in the *K. marmoratus* species complex could be best explained by taxonomy – major clades recovered in the BI phylogeny and microsatellite analyses, or geography – marine ecoregions according to Spalding *et al.* (2007).

In addition, a haplotype network was generated with the program POPART (Leigh & Bryant, 2015) to visualize haplotype distribution and mutational steps. Network ambiguities were solved according to Crandall & Templeton (1993).

RESULTS

Major lineages recovered in phylogenetic analyses of microsatellites and mtDNA

An NJ tree based on the microsatellite data set shows that populations are split into two major lineages; one lineage (Northern clade) includes populations from Florida, Bahamas, Belize and Honduras, while the other lineage encompasses populations from the rest of the Caribbean (southern Cuba, Puerto Rico, Panama, Aruba, Bonaire) and Brazil (Fig. 2). The second lineage is further divided into a Central clade (Caribbean populations) and a Southern clade (Brazilian populations). The Northern clade can similarly be divided into two constituent parts: Florida + Bahamas, and Belize + Honduras, with a cautionary note that the HONR population is not monophyletic with the other Belizean and Honduran populations.

The network constructed with haplotypes of the *cox1* gene shows a pattern congruent to that of the microsatellite tree, but with some differences in the relative extent of divergence between clades (Figs 3, 4). On the microsatellite tree, divergence between the Central and Southern clades is of the same magnitude

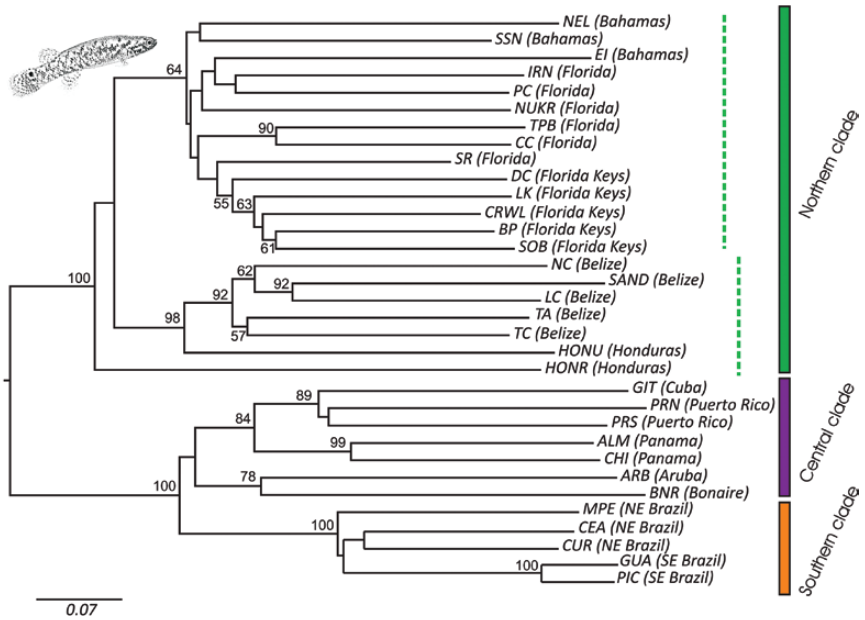

Figure 2. A neighbour-joining tree showing the genetic relationships of 33 populations of *Kryptolebias marmoratus* species complex. The tree is constructed using Cavalli-Sforza & Edwards' (1967) chord distance based on allele frequencies of 33 microsatellite loci. Bootstrap support, shown at the nodes, is based on 1000 replicates; only bootstrap values above 50% are shown. Coloured thick lines outline populations forming three main clades. Broken green lines delineate subclades of the Northern clade.

as divergence between subclades of the Northern clade, whereas the *cox1* gene shows much stronger difference between the Central and Southern clades. The mtDNA *cox1* data set included one specimen from Guanabo River, northern Cuba (GNB) that was not studied at microsatellites. The GNB fish clearly belongs to the Northern clade, as it has the same haplotype (Hap1) as the majority of fish from Florida and the Bahamas. Remarkably, another fish from southern Cuba (Guantanamo Bay, GIT) belongs to the genetically distinct Central clade, having a genotype identical to fish from Puerto Rico and Panama.

To determine the phylogenetic position of a fish from Turks and Caicos (T&C), which was previously studied at the *cytB* gene, we determined sequences of *cytB* in a subset of fish representing the major clades outlined above. The *cytB* tree confirmed the presence of the three clades (Supporting Information, Fig. S2) and showed that the T&C fish belongs to the Central clade. Similar to the situation with the *cox1* gene, the Central and Southern clades were less diverged from each other (mean *p*-distance = 2.54%) than they were from the Northern clade (4.8–5.2%). Relative to diversity between clades, variability within clades was much smaller (0–0.26%).

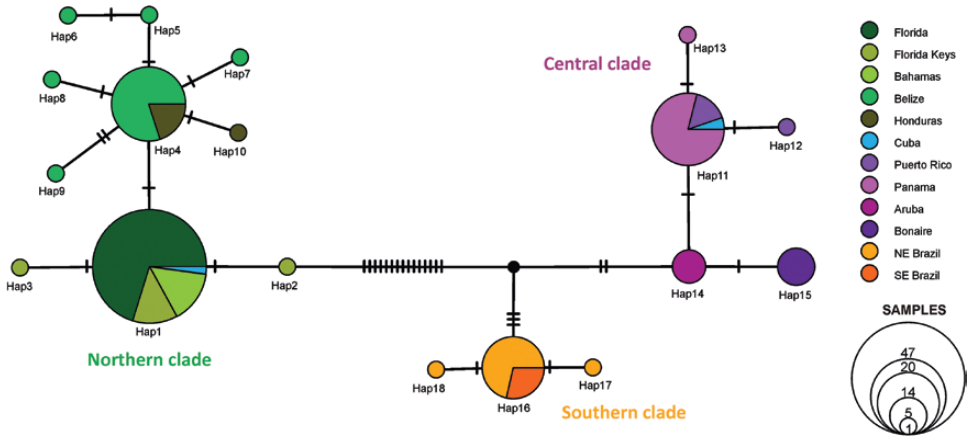

Figure 3. An mtDNA haplotype network based on 655 bp of *cox1* gene from 121 specimens of *Kryptolebias marmoratus* species complex. Each circle represents a haplotype with its size proportional to the frequency of the haplotype. Ticks on branches connecting the haplotypes indicate nucleotide mutations. Major clades are also shown on the network.

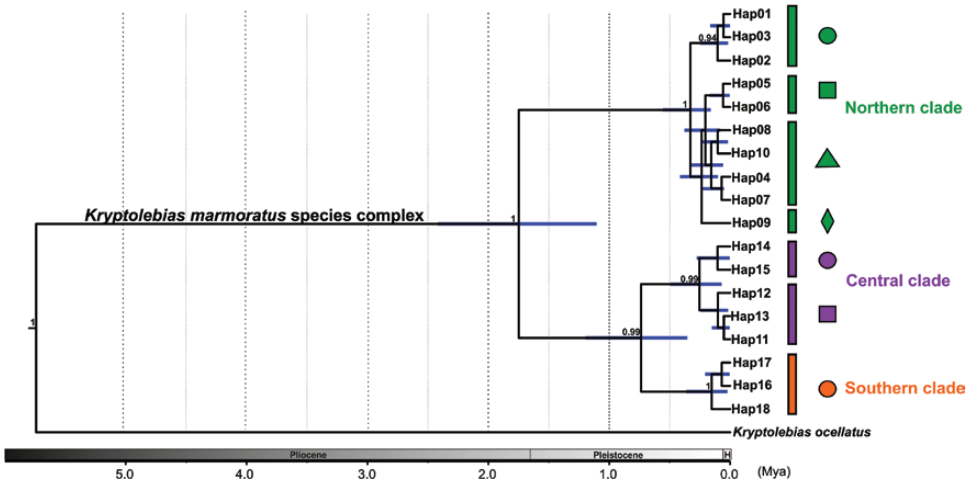

Figure 4. A Bayesian phylogenetic tree of 18 haplotypes based on 655 bp of *cox1* gene from 121 specimens of *Kryptolebias marmoratus* species complex. *Kryptolebias ocellatus* is used as an outgroup. Posterior probabilities are shown above the nodes. Major clades are indicated. The tree is calibrated with the cyprinodontiform Goodeidae *cox1* molecular clock adjusted for fossil record and geological data (see text). Geographic distribution of haplotypes is shown on Supporting Information, Figure 1.

Microsatellite variation within populations

On average, 22.2 individuals were genotyped per population (Table 2), but sample size varied considerably (from 1 to 59). Overall, we screened 33 microsatellite loci, but some loci did not amplify in certain populations. Thus, locus *R34* could not be amplified in individuals from 12 populations forming the Central and Southern clades, whereas it worked perfectly in populations of the Northern clade (549 individuals). Locus *R112* could not be amplified in both populations from SE Brazil, even though it was reliable in all other populations, including three genetically similar populations of the same clade (Southern) from NE Brazil. For both loci, the non-amplification is most likely caused by mutation(s) in the priming sites. One set of primers amplified two distinct loci (named *R22* and *R22-Dup*). *R22-Dup* was not recorded in populations from Belize, Florida and the Bahamas because it was invariably monomorphic for the same allele, but we started to record this locus systematically once variation was detected in individuals from the Southern clade.

Overall, there was a high level of variation and heterogeneity at the microsatellite loci. Average intrapopulation allelic richness (A_R) was 2.77 (Table 2). In the Northern clade, high heterogeneity in A_R was apparent both within and among regions. Highest values of A_R were in the Belizean populations (as high as 5.45, and three other populations had A_R above 4.1), although a population from the tiny Sandbore Caye (SAND) was only 2.45. In Florida, A_R was high in the Florida Keys (range 3.0–3.5) and Shark River (3.71), and it dropped more than two-fold in central Florida at the northern limit of *Kryptolebias* distribution. A_R was low in the Bahamas (1.36–2.22), especially in the remote Exuma Island (EI). A_R values in populations of the Central clade were also highly heterogeneous (1.53–3.44), although sample sizes were admittedly small to draw firm conclusions. No geographical trends in the distribution of A_R were apparent here. In the Southern clade, the highest A_R was in the northeastern Brazil (2.61–2.89), and it dropped two-fold in southeastern Brazil, at the southern extreme of *Kryptolebias* distribution.

The pattern of gene diversity (i.e. expected heterozygosity, H_E) generally followed that described for A_R. The highest H_E was in the Belizean populations, followed by the Florida Keys and Panama. The lowest H_E values were observed in the marginal populations in central Florida, southeastern Brazil and Exuma Island in the Bahamas. There was noticeable heterogeneity in the distribution of H_E both within and among major regions. All populations showed significant deficiencies of heterozygotes, expressed as positive values of F_{IS}. The finding that observed heterozygosity H_O was significantly lower than H_E is not surprising considering that the *K. marmoratus* and *K. hermaphroditus* reproduce by selfing, an extreme form of inbreeding. Rates of selfing estimated from the inbreeding coefficient F_{IS} had clear geographical trends. Florida and the Bahamas had high rates of selfing (up to 100% in several populations), although it was as low as 91% in Key Largo (DC), the northernmost island of the Florida Keys. The Belizean populations had noticeably lower rates of selfing that were also highly heterogeneous, ranging from 77% in Northern Caye (NC) to only 39% in Twin Cayes (TC). Utila Island (HONU) in Honduras, only 150 km from Twin Cayes, showed 100% selfing. All studied populations of the Central clade were 100% selfing. The Southern clade was also highly selfing (97–100%).

Genetic differentiation at microsatellite loci

Overall genetic differentiation between populations was high and statistically significant (F_{ST} = 0.489; $P < 0.001$). After excluding samples with less than three individuals, nearly all pairwise comparisons (404 out of 406 conducted tests) were also significant after Bonferroni correction for multiple testing. The lowest detected value of F_{ST} (0.094) was between Big Pine (BP) and Sugarloaf Key (SOB), which are from keys only 25 km apart. Despite the relatively low F_{ST}, the differentiation between these populations was statistically significant.

A summary of the distribution of genetic variation among samples obtained with analysis of variance for a three-level hierarchy (*sensu* Weir, 1996) is shown in Table 3. AMOVA of the data set partitioned according to taxonomy (i.e. the three groups formed in accordance with clades revealed in the microsatellite and mtDNA analyses) revealed significant differentiation at all hierarchical levels ($P < 0.001$). Differences between clades explained 29% of the total variation; differences between populations within clades explained 28%; and the intrapopulation variation was highest, at 42%. In another AMOVA test, the data were partitioned geographically and ecologically, with groups of populations corresponding to ecoregions (Spalding *et al.*, 2007). Differences between ecoregions accounted for 30% of the total variation, with another 22% being differences between populations within ecoregions, and 48% of the variation was held within populations (variation at all hierarchical levels was statistically significant; $P < 0.001$).

Analysis of genetic relationships at the individual level using clustering analysis implemented in the programs STRUCTURE and INSTRUCT further confirmed strong population structure in *K. marmoratus* species complex and the existence of three to four major clades (Fig. 5). For the whole data set, application of Evanno's method to STRUCTURE analyses suggested that the most probable number

Table 2. Descriptive statistics of genetic variation at microsatellite loci in *Kryptolebias marmoratus* species complex

Population	N	L	P_{99}	A	A_R	H_E	H_O	F_{IS}	s	t
IRN	22	32	0.28	1.53	1.48	0.122	0.000	1.000	1.00	0.00
PC	14	32	0.47	1.50	1.44	0.068	0.000	1.000	1.00	0.00
NUKR	29	32	0.75	2.13	2.00	0.286	0.027	0.907	0.95	0.05
TPB	24	32	0.41	1.56	1.52	0.170	0.000	1.000	1.00	0.00
CC	17	32	0.66	2.13	1.85	0.250	0.023	0.911	0.95	0.05
SR	30	32	0.88	4.59	3.71	0.452	0.002	0.995	1.00	0.00
DC	26	32	0.91	4.38	3.54	0.451	0.076	0.835	0.91	0.09
LK	31	32	0.84	4.19	3.48	0.486	0.061	0.875	0.93	0.07
CRWL	21	32	0.81	3.38	3.03	0.448	0.028	0.938	0.97	0.03
BP	40	32	0.91	4.34	3.55	0.489	0.016	0.968	0.98	0.02
SOB	11	32	0.84	3.44	3.37	0.498	0.023	0.956	0.98	0.02
NEL	14	32	0.75	2.25	2.22	0.362	0.018	0.952	0.98	0.02
EI	12	32	0.34	1.38	1.36	0.139	0.000	1.000	1.00	0.00
SSN	41	32	0.63	2.59	2.09	0.283	0.002	0.995	1.00	0.00
SAND	44	32	0.91	3.16	2.45	0.426	0.195	0.546	0.71	0.29
NC	23	32	0.91	4.97	4.05	0.548	0.207	0.627	0.77	0.23
TA	30	32	0.97	6.19	4.42	0.591	0.282	0.527	0.69	0.31
LC	40	32	0.94	5.78	4.23	0.608	0.236	0.615	0.76	0.24
TC	59	32	0.97	9.28	5.45	0.688	0.520	0.245	0.39	0.61
HONR	1	32	0.00	1.00	n/a	0.000	0.000	n/a	n/a	n/a
HONU	20	32	0.81	3.97	3.55	0.497	0.005	0.991	1.00	0.00
GIT	1	32	0.00	1.00	n/a	0.000	0.000	n/a	n/a	n/a
PRN	1	32	0.06	1.06	n/a	0.062	0.062	n/a	n/a	n/a
PRS	3	32	0.44	1.50	n/a	0.250	0.000	1.000	1.00	0.00
ALM	13	32	0.75	3.38	3.18	0.446	0.002	0.995	1.00	0.00
CHI	28	32	0.75	4.22	3.44	0.463	0.001	0.998	1.00	0.00
ARB	8	32	0.31	1.50	1.53	0.165	0.000	1.000	1.00	0.00
BNR	11	32	0.59	2.44	2.43	0.310	0.003	0.991	1.00	0.00
CEA	38	32	0.59	3.31	2.61	0.331	0.003	0.990	0.99	0.01
CUR	30	32	0.59	3.75	2.86	0.316	0.007	0.977	0.99	0.01
MPE	17	32	0.66	3.28	2.89	0.369	0.002	0.995	1.00	0.00
GUA	10	31	0.13	1.29	1.26	0.060	0.003	0.949	0.97	0.03
PIC	25	31	0.26	1.52	1.30	0.052	0.001	0.976	0.99	0.01
Mean	22.2		0.61	3.09	2.77	0.324	0.055	0.775	0.91	0.09

N, sample size; L, number of loci; P_{99}, proportion of polymorphic loci (99% criterion); A, average number of alleles; A_R, allelic richness based on eight individuals and 30 loci genotyped in all populations; H_E, gene diversity; H_O, observed heterozygosity; F_{IS}, coefficient of inbreeding. All F_{IS} values are highly significant ($P < 0.0001$) as evaluated by randomization in FSTAT (Goudet, 1995). s and t are estimates of selfing and outcrossing, respectively. See Table 1 for population codes. n/a, not applicable.

of clusters was $K = 3$. Under this configuration, the first group was formed by populations from Florida and the Bahamas, the second group included Belize and Honduras and the third group was composed of populations from a vast area of the southern Caribbean and Brazil, including southern Cuba, Puerto Rico, Panama, Aruba, Bonaire and northeastern and southeastern Brazil (i.e. Central and Southern clades combined). At $K = 5$, the Central and Southern clades were consistently distinguished. At $K = 10$, which was the highest number that we explored, the groupings generally made sense considering the geographical divisions.

To gain further insight into finer divisions among populations, we divided all our microsatellite data into smaller data sets and explored them individually. The smaller data sets were formed based on patterns revealed in the phylogenetic and clustering analyses discussed above and included (A) populations of Florida and the Bahamas (##1–14 in Table 1, range of $K = 1$–18); (B) populations of Belize and Honduras (##15–21, $K = 1$–10); (C) populations forming the Central and Southern clades on the microsatellite and mtDNA phylogenetic trees (##23–35; $K = 1$–14); (D) populations of the Central clade (##23–30, $K = 1$–10); and (E) populations of the Southern

Table 3. AMOVA using microsatellite loci for various levels of hierarchical population structure for *Kryptolebias marmoratus* species complex

Source of variation	d.f.	Sum of squares	Variance components	% of variation	Significance (*P*)
Test 1 (structure: taxonomy; 3 clades)					
Among groups (F_{CT})	2	3036.972	4.38406 Va	29.45	< 0.001
Among populations within groups (F_{SC})	30	5849.947	4.20698 Vb	28.26	< 0.001
Among individuals within populations (F_{IS})	701	7791.252	4.82052 Vc	32.39	< 0.001
Genes within individuals (F_{IT})	734	1081.5	1.47343 Vd	9.9	< 0.001
Total	1467	17 759.671	14.88499		
F_{ST}: 0.577, F_{SC}: 0.401, F_{CT}: 0.295, F_{IS}: 0.766, F_{IT}: 0.901					
Test 2 (structure: geography; 9 ecoregions)					
Among groups (F_{CT})	8	5587.734	3.94515 Va	30.08	< 0.001
Among populations within groups (F_{SC})	24	3299.186	2.87691 Vb	21.93	< 0.001
Among individuals within populations (F_{IS})	701	7791.252	4.82052 Vc	36.75	< 0.001
Genes within individuals (F_{IT})	734	1081.5	1.47343 Vd	11.23	< 0.001
Total	1467	17 759.671	13.11602		
F_{ST}: 0.520, F_{SC}: 0.314, F_{CT}: 0.301, F_{IS}: 0.766, F_{IT}: 0.888					

Groups (clades or ecoregions) are defined in Table 1. AMOVA, analysis of molecular variance; d.f., degrees of freedom.

clade encompassing all Brazilian samples (##31–35, K = 1–10). Figure 5 shows results obtained with the program STRUCTURE for the most probable number of Ks according to Evanno's method.

STRUCTURE analysis of set A showed that many populations from central Florida and the Bahamas had unique genetic profiles, whereas populations from the Florida Keys were largely indistinguishable, being composed of multiple genotypes and individuals with mixed ancestries. Samples from some locales, such as Tampa Bay (TPB) and San Salvador Island (SSN), consisted of two main genetic lineages, which may reflect microgeographic variation or the co-existence of distinct clonal lineages in those locales.

The characteristic feature of clustering in set B – Belize and Honduras – was that individuals were firmly assigned to only one cluster, which matched their respective populations. The only exception was fish SAND24, which, although collected from Sandbore Caye, was strongly assigned to another island within the Lighthouse Reef Atoll (Long Caye – LC), suggesting a recent case of a 35 km migration.

Application of Evanno's method to set C – the Central and Southern clades – resulted in six clusters. Three of these clusters corresponded to sampled populations (ARB, BNR, CEA), and the other three clusters contained fish from different populations. One such cluster combined individuals from Cuba, Puerto Rico and Panama, which are rather remote geographic locales. There was evidence for recent migration events: two fish from Curimataú River (CUR) had a genotypic composition characteristic of Ceará-Mirim River (CEA) populations (these two locales are separated by over 100 km). Further division of set C into sets D (Central clade) and E (Southern clade) overall did not result in concomitant fine-tuning of the revealed clusters. However, exploration of the Southern clade at higher Ks showed that local populations of NE Brazil corresponded to unique clusters, whereas the SE Brazilian populations (GUA and PIC) could not be distinguished, indicating their high genetic similarity.

Clustering analyses using INSTRUCT produced virtually indistinguishable clusters from those of STRUCTURE but differed in the most probable number of clusters K. Compared with Evanno's method, the number of clusters suggested by deviance information criterion implemented in INSTRUCT was much higher and, in some cases, equalled the highest Ks that were considered by us. However, visual inspection of the bar plots obtained in the clustering analyses did not render support for such high K values (if we take the values of K as proxies for geographical populations). Instead, it appears that INSTRUCT was assigning distinct population status to individual isogenic ('clonal') lineages.

To gain additional insight about the interrelationships among individuals and populations, we constructed an NJ tree using individual multilocus microsatellite genotypes. This analysis was conducted only for the Central and Southern clades (Fig. 6). We did not do such analysis for the Northern clade for two reasons: (1) the substantially larger number of individuals (N = 549),

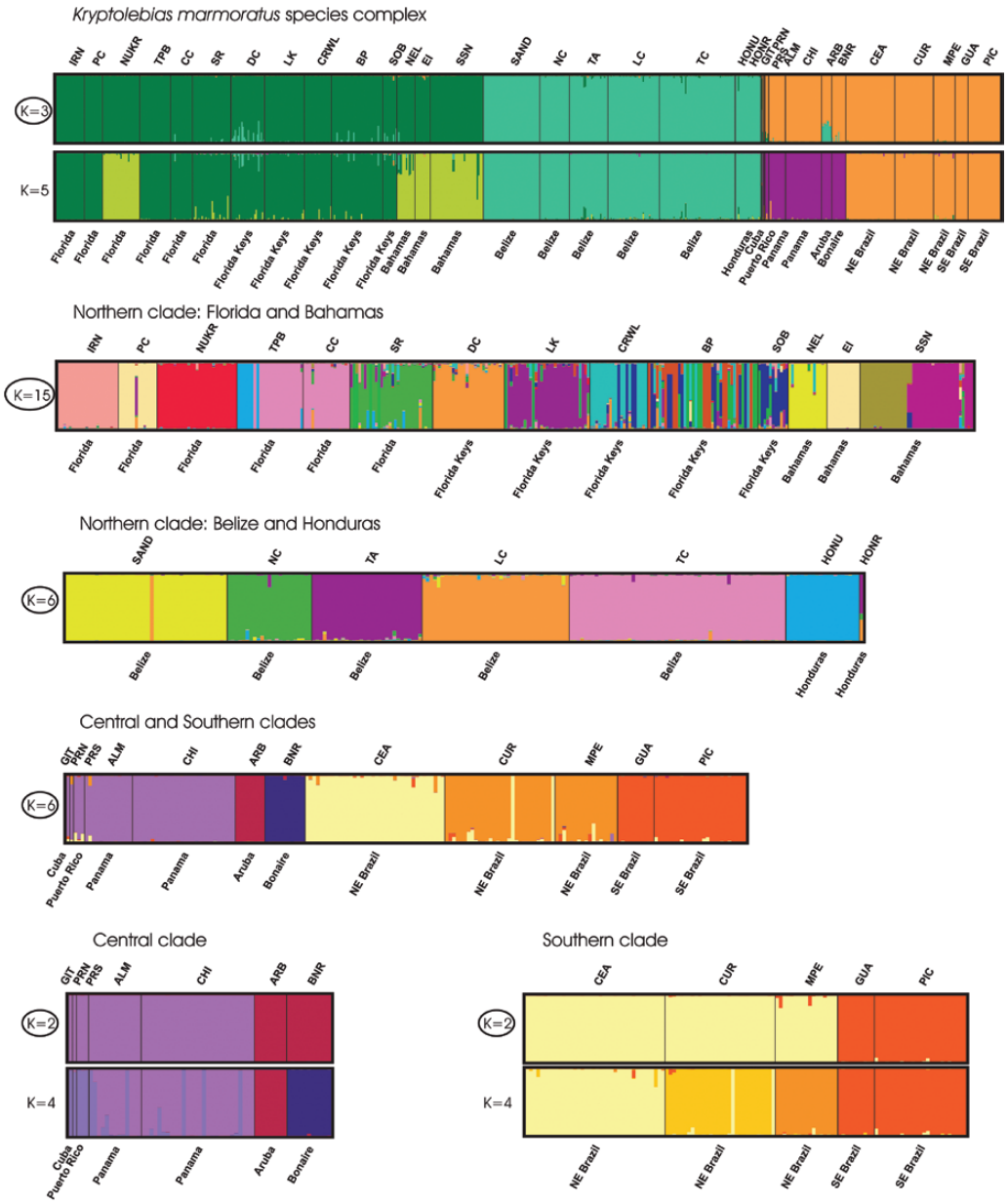

Figure 5. Genotypic clustering of the mangrove rivulus as assessed in Structure. Each killifish specimen is represented by a thin bar, often partitioned into coloured segments each representing an individual's proportionate genetic membership in a given *K*th cluster. Black lines separate sample locations. Specific sample sites (coded as in Table 1) are indicated above each panel, and major geographical areas are shown at the bottom. The panels depict the highest probability outcomes at each level of *K* indicated to the left. The most likely value of *K*, as determined by Evanno *et al.* (2005), is circled.

making their simultaneous depiction impracticable, and (2) similar analysis was previously conducted by us on a smaller representative data set (Tatarenkov *et al.*, 2007). The NJ tree of the individuals from the Central and Southern clades largely confirms results of the analysis by STRUCTURE, but reveals some additional details. One interesting point is that on the NJ tree, individuals cluster nearly perfectly in correspondence with their origin, even when STRUCTURE failed to reveal it. Thus, fish from southeastern Brazil fall into two distinct clusters corresponding to the sampled locales (GUA and PIC), although STRUCTURE could not distinguish them. Similarly, fish from the Greater Antilles (GIT, PRN, PRS) formed a distinct cluster from Panama fish (ALM and CHI), while in STRUCTURE, the division was not so clear. In NE Brazil, the majority of CUR fish formed a compact cluster, but two fish from this population (CUR19 and CUR30) were genetically more similar

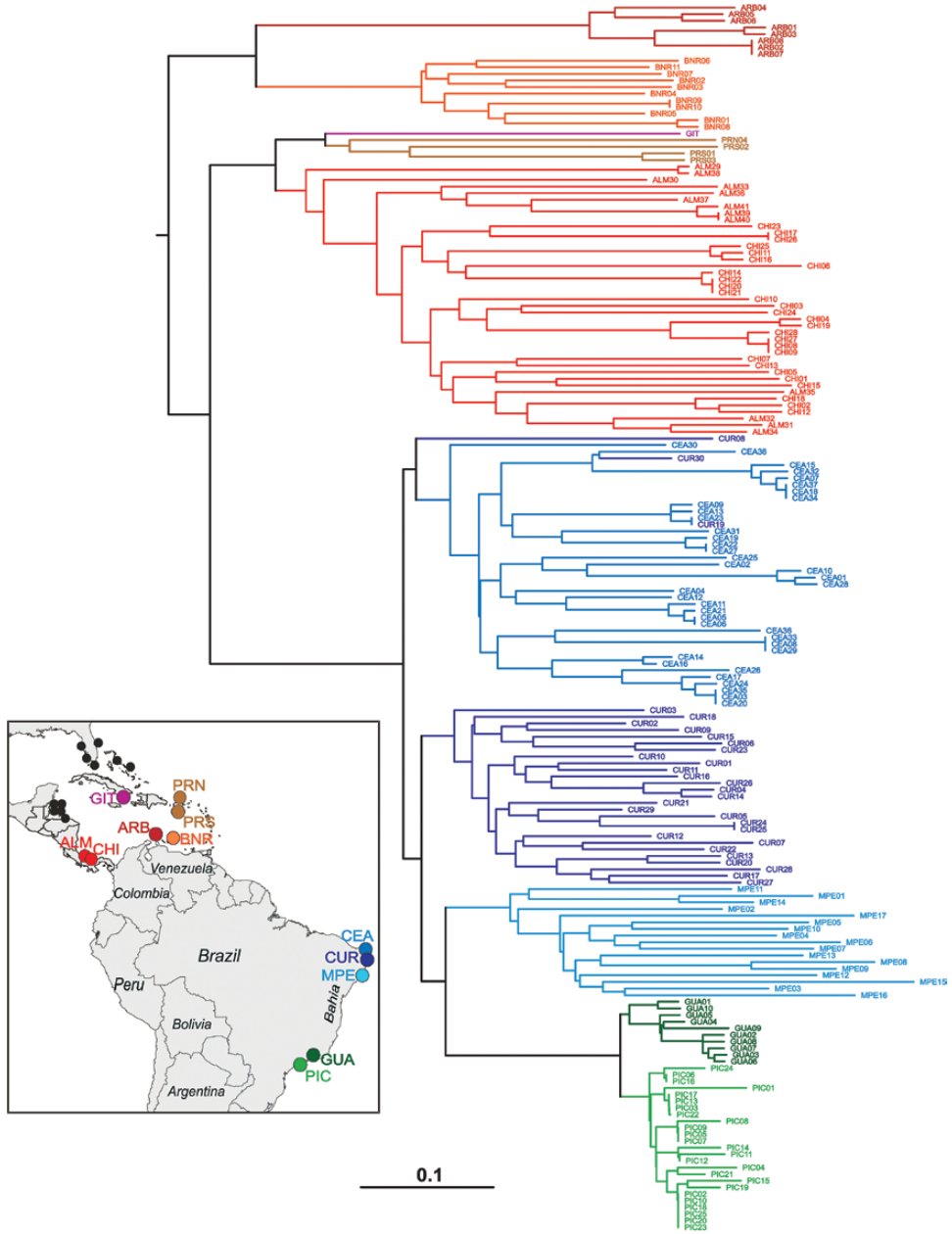

Figure 6. Neighbour-joining tree for 185 *Kryptolebias* killifish forming the Central and Southern clades. The tree is constructed from D_{PS} distances calculated from microsatellite data. Specimens are labelled according to their sampling locations (as in Table 1), followed by individual-specific index. The tree is rooted using as outgroup 11 specimens of *Kryptolebias marmoratus* from the Northern clade; populations of origin of outgroup specimens are shown on the map as small black circles. The outgroup is not shown on the tree.

to CEA fish, illustrated by their deep embedment in the CEA cluster. In fact, fish CUR19 from Curimataú River had genotypes identical to fish CEA23 from Ceará-Mirim at all 32 microsatellite loci (the two fish therefore being effectively clonal). The closer affiliation of CUR19 and CUR30 to the Ceará-Mirim population was also indicated in STRUCTURE, strongly suggesting that these fish are recent migrants (or descendants of migrants) from Ceará-Mirim into Curimataú.

MITOCHONDRIAL *COX1* VARIATION

cox1 sequence data from the *K. marmoratus* species complex revealed 34 variable sites and 18 haplotypes

(Table 4). Overall haplotype diversity was h = 0.787. The haplotype diversity was highest in the Central clade (0.570), closely followed by the Northern clade (0.542), and lowest in the Southern clade (0.242). The pattern of nucleotide diversity paralleled that of haplotype diversity, with overall π = 0.0149, and with diversity within clades ranging from 0.0004 to 0.0013. Nucleotide diversity was four times lower in the Southern clade compared to the Northern and Central clades, but this may have been influenced by a smaller sample size in the former (N = 16 vs. N = 75 and N = 30). Nucleotide diversity in the major geographic areas (defined in Table 1) was generally low ($\pi \leq 0.001$), except for Cuba where π = 0.029. The high nucleotide diversity in Cuba was due to the presence of fish belonging to the divergent Northern and Central clades.

Interpopulational variation at *cox1* gene

Results of pairwise comparisons using Φ_{ST} between major geographic areas are shown in Table 5. If we disregard Cuba, which was represented by only two individuals, the majority of pairwise comparisons (50 among 55) were statistically significant. Differentiation at the *cox1* gene was non-significant between the Bahamas and Florida (including the Florida Keys), between Honduras and Belize, between Panama and Puerto Rico and between southeastern and northeastern Brazil; the Φ_{ST} values in these comparisons were correspondingly low, ranging from 0 to 0.165. The Florida Keys area was significantly different from mainland Florida (Φ_{ST} = 0.212), even though these areas shared a major haplotype (Hap1). The remaining pairwise comparisons showed high values of Φ_{ST}, ranging from 0.619 to 1.0.

The AMOVAs of mitochondrial data sets partitioned in accordance with taxonomy (three clades), or geography (nine ecoregions), showed that variation within populations was low (≤ 1.75%). Most of the *cox1* variation (> 92%) was apportioned to differences between groups, be these groups major clades or marine ecoregions (Table 6). All Φ_{CT} values expressing differences between groups were highly significant. Differences between populations within groups accounted for 3.3–5.8% of the variation at *cox1*.

Table 4. Population diversity indices for mtDNA *cox1* gene in *Kryptolebias marmoratus* species complex

	Population diversity indices				
	N	H	S	h	$\pi \times 100$
Clades					
Northern clade	75	10	10	0.542	0.116
Central clade	30	5	4	0.570	0.131
Southern clade	16	3	2	0.242	0.038
Major areas*					
Florida^{N}	33	1	0	0	0
Florida Keys^{N}	8	3	2	0.464	0.076
Bahamas^{N}	7	1	0	0	0
Belize^{N}	21	6	6	0.428	0.100
Honduras^{N}	5	2	1	0.400	0.061
Cuba^{N,C}	2	2	19	1.000	2.900
Puerto Rico^{C}	4	2	1	0.500	0.076
Panama^{C}	16	2	1	0.125	0.019
Aruba^{C}	4	1	0	0	0
Bonaire^{C}	5	1	0	0	0
NE Brazil^{S}	12	3	2	0.318	0.050
SE Brazil^{S}	4	1	0	0	0
Total	121	18	34	0.787	1.492

H, number of haplotypes; S, number of polymorphic sites; h, haplotype diversity; π, nucleotide diversity. mtDNA, mitochondrial DNA.
*Clade to which samples in the area were assigned in genetic analyses: ^{N}, Northern; ^{C}, Central; ^{S}, Southern.

Concordance of mtDNA and microsatellite variation

Microsatellite and mtDNA markers were concordant in recovering three major clades of mangrove rivulus as well as in establishing their phylogenetic relationships. One important distinction between the mtDNA and microsatellite patterns of variation is in its distribution among hierarchical levels. If we consider partitioning according to taxonomy, then the largest proportion of microsatellite variation is concentrated within populations; the remaining variation (~58%) is evenly divided into differences among populations within clades, and differences among clades. In contrast, 95.6% of variation at the *cox1* gene is attributed to differences among clades, with another 3.3% accounted for by differences among populations, and variability among individuals within populations accounting for only 1.1%. Thus, mtDNA is highly informative about membership of individual fish in a certain clade and has some power in distinguishing main areas or ecoregions, but is unable to ascribe individuals to local populations. In comparison, the combined power of multiple microsatellite loci ensures solid assignment of individual fish at nearly all levels, including the level of local populations. One consequence of the extremely high intrapopulation microsatellite variation is that it makes it difficult to establish the hierarchical order of relationships among populations (as is reflected in the low bootstrap support on an NJ tree of populations; Fig. 2). In other words, while we can confidently assign individual fish to one local population or another, we cannot reliably say precisely how populations in a given clade or area are related.

Table 5. Pairwise divergence Φ_{ST} values for mtDNA *cox1* gene among major areas of *Kryptolebias marmoratus* species complex

	1	2	3	4	5	6	7	8	9	10	11
1. Florida	–										
2. Florida Keys	0.212*	–									
3. Bahamas	0	–0.018	–								
4. Belize	0.799*	0.618*	0.657*	–							
5. Honduras	0.958*	0.682*	0.863*	–0.026	–						
6. Cuba	0.888*	0.573	0.588	0.760*	0.482	–					
7. Puerto Rico	0.997*	0.974*	0.991*	0.967*	0.977*	0.360	–				
8. Panama	0.997*	0.987*	0.995*	0.977*	0.990*	0.770	0.165	–			
9. Aruba	1*	0.980*	1*	0.969*	0.987*	0.384	0.800*	0.904*	–		
10. Bonaire	1*	0.983*	1*	0.971*	0.989*	0.537*	0.903*	0.952*	1*	–	
11. NE Brazil	0.995*	0.979*	0.988*	0.972*	0.981*	0.792*	0.942*	0.965*	0.949*	0.960*	–
12. SE Brazil	1*	0.981*	1*	0.970*	0.988*	0.556	0.960*	0.982*	1*	1*	–0.128

*$P < 0.05$.

Table 6. Results from the AMOVA tests evaluating different hypotheses for mtDNA *cox1* gene in *Kryptolebias marmoratus* species complex

Hypothesis (number of groups)	Among groups (%)	Among populations, within groups (%)	Within populations (%)	Φ_{CT}
Major clades (3)	95.55	3.31	1.13	0.955*
Marine ecoregions (9)	92.42	5.84	1.75	0.924*

Φ_{CT} is *F*-statistic explaining differences between groups. Groups are defined in Table 1. AMOVA, analysis of molecular variance; mtDNA, mitochondrial DNA.
*$P < 0.05$.

DISCUSSION

The *K. marmoratus* species complex includes two currently recognized self-fertilizing species, *K. marmoratus* and *K. hermaphroditus*, but due to their morphological similarity, confusing taxonomical history and inadequate sampling, their respective geographic distributions were poorly delineated. The combined range of these species encompasses the entire Tropical Western Atlantic and extends further into the temperate Northwestern and Southwestern Atlantic provinces (Tatarenkov *et al.*, 2011; Taylor, 2012). Our current study included 35 populations covering the known geographic distribution of self-fertilizing *Kryptolebias*, allowing us to assess the phylogeographic structure of this complex and also to determine rates of self-fertilization over a wide area. Quite remarkably, we discovered a deep genetic division between various populations in the Caribbean (an area formerly thought to be the exclusive domain of *K. marmoratus*) and established that many of these populations are genetically more closely related to populations of *K. hermaphroditus* from Brazil. Below we discuss some taxonomic implications of these findings, followed by a discussion of the pattern of population structure and the genetically assessed rates of self-fertilization.

Both types of genetic markers that we used (many nuclear microsatellite loci and mitochondrial gene sequences) indicated the presence of three distinct genetic lineages here termed Northern, Central and Southern. The Northern clade encompasses populations of the Floridian, Bahamian and Western Caribbean ecoregions, as well as northern Cuba (Greater Antilles) and north-central Florida (Carolinian ecoregion). The Central clade occupies the Greater Antilles, the Southern and Southwestern Caribbean ecoregions, and also found in Turks and Caicos (Bahamian ecoregion). The Southern clade includes the Northeastern and Southeastern Brazilian ecoregions; furthermore, recently discovered populations in Eastern Brazil and

Amazonia (Sarmento-Soares *et al.*, 2014; Guimarães-Costa *et al.*, 2017) also belong to this clade. Of the three lineages, the Central and Southern clades genetically are more similar to each other [Kimura-2-parameter (K2P) distance = 0.98% at *cox1*], with the Northern clade being genetically more distant from them (K2P distance = 3%).

The finding of two genetically strongly diverged clades – Northern and Central – in the Caribbean region was unexpected; it generally had been assumed that all of the Caribbean was the domain of *K. marmoratus*. Earlier genetic studies did not detect subdivisions within Caribbean populations of *Kryptolebias*. Thus, Vrijenhoek (1985) reported that populations in Florida and Curaçao had identically fixed alleles at 31 allozyme loci, and only a Yucatan population was slightly different. Even more puzzling is the higher relatedness of the Caribbean Central clade to the geographically distant Southern clade from Brazil. The discovery of the Central clade requires some consideration of the taxonomy of *Kryptolebias*, as well as the origin of these genetic clades. Currently, two selfing *Kryptolebias* species are valid: *K. marmoratus* and *K. hermaphroditus* (Costa, 2006, 2011). Clearly, the Northern clade corresponds to *K. marmoratus* and the Southern clade is *K. hermaphroditus*. Where does the Central clade then fit?

Hoedeman (1958) in a morphological analysis of *K. marmoratus* proposed that specimens from Bonaire, Aruba and Los Roques represent a distinct subspecies (*K. marmoratus bonairensis*). Our study showed that specimens from Bonaire and Aruba belong to the Central clade, thus establishing its correspondence to *K. marmoratus bonairensis* (Hoedeman). This subspecies has been synonymized with *K. marmoratus* (Costa, 2003), but considering the genetic distinctiveness of the Central clade, this taxonomic unit should now be re-evaluated. The Central clade can also be potentially identified with *Kryptolebias heyei* described from Saona island in the Dominican Republic (Nichols, 1914) and currently synonymized with *K. marmoratus*. Our study does not include specimens from Saona, but geographic position of this island suggests that mangrove rivulus there probably belong to the Central clade, considering its occurrence on Cuba in the north and Puerto Rico in the south. The resurrection of either *bonairensis* or *heyei*, however, poses a taxonomic conundrum. If we accept two currently recognized selfing species (*K. marmoratus* and *K. hermaphroditus*), then the Central clade would belong to the latter. However, availability of the earlier names requires abandonment of the name *K. hermaphroditus* (or downgrading its rank to the subspecies, e.g. *K. heyei hermaphroditus*) in accordance with principle of priority. The second approach is to recognize three selfing species: *K. marmoratus*, *K. heyei*/*bonairensis* and *K. hermaphroditus*. Yet another approach is to accept the existence of a single (super)species, *K. marmoratus*, with three subspecies: *K. marmoratus marmoratus*, *K. marmoratus heyei*/*bonairensis* and *K. marmoratus hermaphroditus*.

Division of selfing *Kryptolebias* in two species appears to us the least desirable because it further complicates the already convoluted taxonomy of the group. The second approach (three species) is consistent with the existence of three genetic lineages but depends a lot on future studies confirming such distinctiveness in the area between Venezuela and NE Brazil (and also on further knowledge about intrinsic reproductive boundaries between these taxa). The third approach (a single species with three subspecies) seems to be conservative and practical because, on the one hand, it recognizes the existence of three genetic lineages and, on the other hand, takes into account their high morphological and genetic similarity. Currently, there are no diagnostic morphological characters that allow unambiguous assignment of mangrove rivulus specimens to their respective clades. Furthermore, there is some indication that fish of the Northern and Central clades are genetically compatible (and thus are potentially not isolated reproductively). Nakamura *et al.* (2008) and Kanamori *et al.* (2016) used artificial fertilization to produce viable F1 hybrids between laboratory strains originating from Panama (Central clade) and those from Florida and Belize (Northern clade, *marmoratus*). These hermaphroditic hybrids then self-fertilized, resulting in a viable F2 generation. Although conducted in the laboratory, this study does demonstrate a potential avenue for gene exchange between the Central and Northern clades. If genetic similarity is an indicator of genetic compatibility, then viability of the crosses between the Southern and Central clades is highly probable because they are genetically closer to one another than either of them is to the Northern clade.

Currently, however, interbreeding between genetic lineages of the *K. marmoratus* species complex in nature is not known and potentially impeded by predominant reproduction by self-fertilization. In fact, the highly selfing nature of *Kryptolebias*, resulting in low levels of genetic exchange and recombination, can be an argument that the biological species concept (BSC) may not be quite applicable to this fish and that the phylogenetic species concept (PSC) is more appropriate. Each of the three phyletic lineages that our data reveal qualifies as a distinct species under the PSC. Divergence between the Central and Southern clades is not particularly high (K2P distance is 1% at *cox1*), but about 3% of valid fish species have divergence less than 1% (Ward, Hanner & Hebert, 2009). The decision on taxonomic status of the three identified lineages of *Kryptolebias* is thus somewhat subjective and largely

determined by the adherence of a researcher to one or another school of thought, BSC or PSC. Anecdotally, even within the small group of co-authors of this study, we could not come to consensus on this subject. Some of us favoured the idea of a single species with subspecies status for all three lineages, others preferred three selfing *Kryptolebias* species and yet others favoured synonymizing of the Central and Southern clades, resulting in two selfing *Kryptolebias* species – *K. marmoratus* and *K. heyei*/*bonairensis*.

It is tempting to contemplate on the time of the origin of the three clades and on the possible isolating barriers (physiographical, hydrological, geological, ecological) that triggered their formation. The mtDNA tree (Fig. 4) suggests that the Northern clade separated from the other two clades about 1.75 Mya, whereas the Central and Southern clades went their separate courses about 0.75 Mya. Unfortunately, due to complicated geologic formation of the Caribbean, no particular events can be assigned with certainty to these periods. However, the boundary between the Northern and Central clades coincides nearly perfectly with the deep Cayman trench, which is paralleled to the south by an extensive shelf running from the coast of Nicaragua and Honduras to Jamaica and then extending onto Hispaniola. Deep water bodies are often strong isolating barriers for coral reef fishes (Taylor & Hellberg, 2006) and may also be a substantial barrier for *Kryptolebias*. Furthermore, the Pleistocene was characterized by frequent and considerable oscillations of sea level; in times of low sea level, the northern part of the Caribbean (Yucatan and Cayman Basins) would be nearly completely separated by exposed shelf from the southern part (the Colombian and Venezuelan Basins), resulting in geographic isolation of populations that now comprise the Northern clade.

A parsimonious interpretation of our phylogenetic tree suggests that the common ancestor of the Central and Southern clades resided in the Caribbean, from where it may have colonized the region to the south. This idea is further supported by the higher genetic diversity in Caribbean populations forming the Central clade, compared to the Brazilian populations. However, sampling of the populations of the Central and Southern clades is sparse, allowing for other scenarios. One possibility is that the most recent common ancestor of the Central and Southern clades resided more to the south (e.g. Guianas or northern Brazil), from where it expanded both to the north and to the south. On the whole, our data support the surprising conclusion that the self-fertilizing *Kryptolebias* lineage did not originate in southeastern Brazil even though *K. ocellatus* occurs in that region. The latter, a hermaphroditic (but not selfing) species (Tatarenkov *et al.*, 2009; Costa *et al.*, 2010), is the closest known relative of the selfing *Kryptolebias* species and the only other congener that inhabits brackish/marine mangals. Our data quite unambiguously show that populations of *K. hermaphroditus* in southeastern Brazil are derived (Tatarenkov *et al.*, 2011). This proposal is reinforced in the present study by observation that populations in northeastern Brazil harbour much higher diversity than those in southeastern Brazil, suggesting that *K. hermaphroditus* spread from northern Brazil to the south.

For some time, it was believed that *K. hermaphroditus* is endemic to SE Brazil (Costa, 2011). *Kryptolebias* populations in NE Brazil were only reported 2 years ago (Lira *et al.*, 2015), and now we show that these populations are genetically highly similar to those in SE Brazil and unequivocally belong to *K. hermaphroditus* (in accordance with description by Lira *et al.*, 2015), even though the two areas are separated by over 2400 km. Another study (Guimarães-Costa *et al.*, 2017) reported *K. hermaphroditus* even further north, just south of Amazon river. These findings cast doubts on a recent report of *K. marmoratus* in NE Brazil (Costa, 2016) and suggest that genetic confirmation of species identities is necessary in the absence of reliable morphological characters.

Results of this and previous studies indicate that *Kryptolebias* is capable of occasional long distance migrations, perhaps via adhesive embryos attached to floating matter (Taylor *et al.*, 2008; Turko & Wright, 2015). These cases of long distance transfer – 35 km to 100 km – include Florida Keys (Tatarenkov *et al.*, 2012), Belize (Tatarenkov *et al.*, 2015) and southeastern Brazil (this study). In all of these cases, the direction of migration was consistent with the direction of predominant sea currents (Supporting Information, Fig. S1). Furthermore, in all mentioned cases, the migration must have been very recent (not more than a few generations ago), judging by the absence of genetic differences between fish at microsatellite loci with high rates of mutations (Tatarenkov *et al.*, 2010). The ability to traverse long distances coupled with ability of a single selfing individual to found new population are consistent with the enormous geographic range of mangrove rivulus, in fact the widest distribution of any inshore-dwelling coastal fish species on American continents (Taylor, 2012).

Previous population genetic studies of *Kryptolebias* uncovered high selfing rates (91–100%) in Florida and the Bahamas and generally lower (but highly geographically heterogeneous) selfing rates in Belize (39–77%) (Mackiewicz *et al.*, 2006b; Tatarenkov *et al.*, 2007, 2012, 2015). Additional populations from those areas included in the present study confirm those findings. Natural populations of the Central clade were studied for the first time. These populations showed some of the highest selfing rates, effectively approaching 100%. In Florida, where selfing rates are also high,

we infrequently found fish that were heterozygous at multiple loci, suggesting recent outcrossing events. In the Central clade, almost all fish were completely homozygous (only a few fish displayed heterozygosity, and that was limited to one or two loci). Such levels of heterozygosity indicate that the last outcrossing events producing such fish occurred ~20–25 generations ago (Redelings *et al.*, 2015). Populations of *K. hermaphroditus* in Brazil are highly selfing in both studied areas (range 97–100%).

Occasional outcrossing in *K. marmoratus* has been linked to the presence of males; areas with high outcrossing rates generally have higher male frequencies (Turner, Davis & Taylor, 1992; Taylor, Fisher & Turner, 2001; Turner *et al.*, 2006). Males in *K. hermaphroditus* were unknown until recently, but about a year ago, a male was found in northeastern Brazil (Berbel-Filho, Espírito-Santo & Lima, 2016). This is consistent with the very high, but incomplete, deduced rates of selfing in this species. In the Central clade, males appear to be infrequent. On Curaçao (an island halfway between Aruba and Bonaire), Kristensen (1970) during the years 1960–1965 found males only in 1960. No males could be detected there in 1974 (Lubinski *et al.*, 1995). Nevertheless, despite the very low individual heterozygosity, the high diversity of multilocus genotypes in populations of the Central clade strongly suggests that occasional outcrossing does occur. Previous studies showed that in the absence of outcrossing, some individuals will be identical (belong to the same 'clone'), whereas even low rates of outcrossing drastically change the pattern so that a great majority of specimens remains almost completely homozygous, but the occurrence of identical individuals becomes exceedingly rare in populations of moderate size (Mackiewicz *et al.*, 2006c; Tatarenkov *et al.*, 2015).

Our finding of two highly distinct Caribbean lineages – Northern and Central – raises intriguing questions. Are there significant (but hitherto unrecognized) biological differences between the lineages? Are they (or were they ever) in direct contact? Does introgression or displacement occur in putative contact zones? The reclusive *K. marmoratus* is still far from revealing all of its secrets.

ACKNOWLEDGEMENTS

We thank Lucas Medeiros, Mateus Lira and Roney Paiva for help during field collections. We are grateful to Luiz Rulff, Pablo Robles and David Mikkelsen for Puerto Rico specimens and to Carole McIvor for collections in West Florida. Our sincere thanks to Mrs Sabine Engel, Caribbean Marine Biological Institute (CARMABI) for help in the field and to Mr Frank van Slobbe, Government of Bonaire, for arranging licences. We also 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 in the field. Portions of this work were approved by the following permits to Ryan Earley – 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. Collections in Brazil were taken under permit #30532-1/2011 issued by ICMbio/SISBIO. Sergio Lima was granted a postdoctoral abroad scholarship (PDE) from Science without Borders Program/CNPq (process #203476/2014-0) and Waldir Berbel-Filho receives a PhD scholarship from the same program (process #233161/2014–7). Olga Tatarenkova made a drawing of the mangrove rivulus, and Andrew Furness commented on the manuscript. Molecular genetic analyses were supported by funds from the University of California at Irvine. We thank four anonymous reviewers for their helpful comments.

DATA ACCESSIBILITY

Cytochrome oxidase (*cox1*) gene sequences: GenBank accessions MF554974–MF555093. Cytochrome B (*cytB*) gene sequences: GenBank accessions MF554942–MF554973 and MF555094–MF555095.

REFERENCES

Avise JC. 2008. *Clonality: the genetics, ecology, and evolution of sexual abstinence in vertebrate animals*. New York: Oxford University Press.

Avise JC, Tatarenkov A. 2015. Population genetics and evolution of the mangrove rivulus *Kryptolebias marmoratus*, the world's only self-fertilizing hermaphroditic vertebrate. *Journal of Fish Biology* **87:** 519–538.

Berbel-Filho WM, Espírito-Santo HMV, Lima SMQ. 2016. First record of a male of *Kryptolebias hermaphroditus* Costa, 2011 (Cyprinodontiformes: Cynolebiidae). *Neotropical Ichthyology* **14:** e160024.

Bowcock AM, Ruiz-Linares A, Tomfohrde J, Minch E, Kidd JR, Cavalli-Sforza LL. 1994. High resolution of human evolutionary trees with polymorphic microsatellites. *Nature* **368:** 455–457.

Cassel A. 1982. *Rivulus marmoratus* in Brazilien? *Deutsche Killifisch Gemeinschaft Journal* **14:** 101–103.

Cavalli-Sforza LL, Edwards AWF. 1967. Phylogenetic analysis: models and estimation procedures. *American Journal of Human Genetics* **19:** 233–257.

Costa WJEM. 2003. Family Rivulidae (South American annual fishes). In: Reis RE, Kullander SO, Ferraris CJ Jr,

eds. *Check list of the freshwater fishes of South and Central America*. Porto Alegre: EDIPUCRS, 527–549.

Costa WJEM. 2006. Redescription of *Kryptolebias ocellatus* (Hensel) and *K. caudomarginatus* (Seegers) (Teleostei: Cyprinodontiformes: Rivulidae), two killifishes from mangroves of south-eastern Brazil. *Journal of Ichthyology and Aquatic Biology* **11:** 5–12.

Costa WJEM. 2011. Identity of *Rivulus ocellatus* and a new name for a hermaphroditic species of *Kryptolebias* from south-eastern Brazil (Cyprinodontiformes: Rivulidae). *Ichthyological Exploration of Freshwaters* **22:** 185–192.

Costa WJEM. 2016. Colouration, taxonomy and geographical distribution of mangrove killifishes, the *Kryptolebias marmoratus* species group, in southern Atlantic coastal plains of Brazil (Cyprinodontiformes: Rivulidae). *Ichthyological Exploration of Freshwaters* **27:** 183–192.

Costa WJEM, Lima SMQ, Bartolette R. 2010. Androdioecy in *Kryptolebias* killifish and the evolution of self-fertilizing hermaphroditism. *Biological Journal of the Linnean Society* **99:** 344–349.

Crandall KA, Templeton AR. 1993. Empirical tests of some predictions from coalescent theory with applications to intraspecific phylogeny reconstruction. *Genetics* **134:** 959–969.

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

Diaz-Ferguson E, Haney RA, Wares JP, Silliman BR. 2010. Population genetics of a trochid gastropod broadens picture of Caribbean Sea connectivity. *PLoS ONE* **5:** e12675.

Dieringer D, Schlötterer C. 2003. MICROSATELLITE ANALYSER (MSA): a platform independent analysis tool for large microsatellite data sets. *Molecular Ecology Notes* **3:** 167–169.

Drummond AJ, Rambaut A. 2007. BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evolutionary Biology* **7:** 214.

Drummond AJ, Suchard MA, Xie D, Rambaut A. 2012. Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Molecular Biology and Evolution* **29:** 1969–1973.

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

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. *Ecology and Evolution* **2:** 1682–1695.

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

Excoffier L, Lischer HE. 2010. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources* **10:** 564–567.

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.

Felsenstein J. 1993. *PHYLIP (Phylogeny Inference Package), Version 3.5c*. Seattle: Department of Genetics, University of Washington.

Furness AI, Tatarenkov A, Avise JC. 2015. A genetic test for whether pairs of hermaphrodites can cross-fertilize in a selfing killifish. *Journal of Heredity* **106:** 749–752.

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.

García G, Gutiérrez V, Vergara J, Calviño P, Duarte A, Loureiro M. 2012. Patterns of population differentiation in annual killifishes from the Paraná-Uruguay-LaPlata Basin: the role of vicariance and dispersal. *Journal of Biogeography* **39:** 1707–1719.

Giri C, Ochieng E, Tieszen LL, Zhu Z, Singh A, Loveland T, Masek J, Duke N. 2011. Status and distribution of mangrove forests of the world using earth observation satellite data. *Global Ecology and Biogeography* **20:** 154–159.

Goodwin AE, Grizzle JM. 1994. Oncogene expression in hepatic and biliary neoplasms of the fish *Rivulus ocellatus marmoratus*: correlation with histologic changes. *Carcinogenesis* **15:** 1993–2002.

Goudet J. 1995. FSTAT (Version 1.2): a computer program to calculate F-statistics. *Journal of Heredity* **86:** 485–486.

Guimarães-Costa A, Schneider H, Sampaio I. 2017. New record of the mangrove rivulid *Kryptolebias hermaphroditus* Costa, 2011 (Cyprinodontiformes: Cynolebiidae) in the Pará state, northern Brazil. *Check List* **13:** 2093.

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

Hensel R. 1868. Beiträge zur kenntniss der wirbelthiere Südbrasiliens. *Archiv für Naturgeschichte* **34:** 323–375.

Hoedeman JJ. 1958. Rivulid fishes of the Antilles. *Studies on the Fauna of Curaçao and other Caribbean Islands* **8:** 112–126.

Huber JH. 1992. *Review of* Rivulus: *ecobiogeography – relationships*. Paris: Société Française d'Ichtyologie.

Kanamori A, Sugita Y, Yuasa Y, Suzuki T, Kawamura K, Uno Y, Kamimura K, Matsuda Y, Wilson CA, Amores A, Postlethwait JH, Suga K, Sakakura Y. 2016. A genetic map for the only self-fertilizing vertebrate. *G3: Genes, Genomes, Genetics* **6:** 1095–1106.

Kim HS, Hwang DS, Hagiwara A, Sakakura Y, Lee JS. 2016. Complete mitochondrial genome of the mangrove killifish *Kryptolebias hermaphroditus* (Cyprinodontiformes, Rivulidae). *Mitochondrial DNA B: Resources* **1:** 540–541.

Kristensen I. 1970. Competition in three Cyprinodont fish species in the Netherlands Antilles. *Studies on the Fauna of Curaçao and other Caribbean Islands* **32:** 82–101.

Lara A, Ponce de Leon JL, Rodriguez R, Casane D, Cote G, Bernatchez L, Garcia-Machado E. 2010. DNA barcoding of Cuban freshwater fishes: evidence for cryptic species and taxonomic conflicts. *Molecular Ecology Resources* **10:** 421–430.

Leigh JW, Bryant D. 2015. POPART: full-feature software for haplotype network construction. *Methods in Ecology and Evolution* **6:** 1110–1116.

Lewis PO, Zaykin D. 2001. *Genetic data analysis: computer program for the analysis of allelic data*. Available at: http://phylogeny.uconn.edu/software/

Lira MGS, Paiva REC, Ramos TPA, Lima SMQ. 2015. First record of *Kryptolebias hermaphroditus* Costa, 2011 (Cyprinodontiformes: Rivulidae) in the extreme north Atlantic Forest mangroves, Rio Grande do Norte state, Brazil. *Check List* **11:** 1656.

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

Luther DA, Greenberg R. 2009. Mangroves: a global perspective on the evolution and conservation of their terrestrial vertebrates. *Bioscience* **59:** 602–612.

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

Mackiewicz M, Tatarenkov A, Taylor DS, Turner BJ, Avise JC. 2006b. Extensive outcrossing and androdioecy in a vertebrate species that otherwise reproduces as a self-fertilizing hermaphrodite. *Proceedings of the National Academy of Sciences of the United States of America* **103:** 9924–9928.

Mackiewicz M, Tatarenkov A, Turner BJ, Avise JC. 2006c. A mixed-mating strategy in a hermaphroditic vertebrate. *Proceedings of the Royal Society of London B: Biological Sciences* **273:** 2449–2452.

Murphy WJ, Thomerson JE, Collier GE. 1999. Phylogeny of the Neotropical killifish family Rivulidae (Cyprinodontiformes, Aplocheiloidei) inferred from mitochondrial DNA sequences. *Molecular Phylogenetics and Evolution* **13:** 289–301.

Nakamura Y, Suga K, Sakakura Y, Sakamoto T, Hagiwara A. 2008. Genetic and growth differences in the outcrossings between two clonal strains of the self-fertilizing mangrove killifish. *Canadian Journal of Zoology* **86:** 976–982.

Nichols JT. 1914. *Gobiosoma longum* and *Rivulus heyei*, new fishes from the West Indian fauna. *Bulletin of the American Museum of Natural History* **33:** 143–144.

Poey F. 1880. Revisio piscium cubensium. *Anales de la Sociedad Española de Historia Natural* **9:** 243–261.

Ponce de Leon JL, Leon G, Rodriguez R, Metcalfe CJ, Hernandez D, Casane D, Garcia-Machado E. 2014. Phylogeography of Cuban *Rivulus*: evidence for allopatric speciation and secondary dispersal across a marine barrier. *Molecular Phylogenetics and Evolution* **79:** 404–414.

Redelings BD, Kumagai S, Tatarenkov A, Wang LY, Sakai AK, Weller SG, Culley TM, Avise JC, Uyenoyama MK. 2015. A Bayesian approach to inferring rates of selfing and locus-specific mutation. *Genetics* **201:** 1171–1188.

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

Sarmento-Soares LM, Ingenito LFS, Duboc LF, Martins Pinheiro RF, Borçato R, Silva JP. 2014. Primeiros registros de *Kryptolebias ocellatus* (Hensel) (Cyprinodontiformes, Rivulidae) para riachos de Mata Atlântica no Espírito Santo. *Boletim Sociedade Brasileira de Ictiologia* **111:** 15–19.

Seegers L. 1984. Zur revision der *Rivulus* – arten südost-Brasiliens, mit einer neubeschreibung von *Rivulus luelingi* n. sp. und *Rivulus caudomarginatus* n. sp. (Pisces, Cyprinodontidae: Rivulinae). *Zoologische Beiträge* **28:** 271–320.

Spalding MD, Fox HE, Halpern BS, McManus MA, Molnar J, Allen GR, Davidson N, Jorge ZA, Lombana AL, Lourie SA, Martin KD, McManus E, Molnar J, Recchia CA, Robertson J. 2007. Marine ecoregions of the world: a bioregionalization of coastal and shelf areas. *Bioscience* **57:** 573–583.

Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. 2013. MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. *Molecular Biology and Evolution* **30:** 2725–2729.

Tatarenkov A, Earley RL, Perlman BM, Taylor DS, Turner BJ, Avise JC. 2015. Genetic subdivision and variation in selfing rates among Central American populations of the mangrove rivulus, *Kryptolebias marmoratus*. *Journal of Heredity* **106:** 276–284.

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

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*). *Molecular Ecology* **16:** 2701–2711.

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. *Journal of Fish Biology* **79:** 2095–2105.

Tatarenkov A, Lima SMQ, Taylor DS, Avise JC. 2009. Long-term retention of self-fertilization in a fish clade. *Proceedings of the National Academy of Sciences of the United States of America* **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.

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

Taylor DS. 2003. Meristic and morphometric differences in populations of *Rivulus marmoratus*. *Gulf of Mexico Science* **21:** 145–158.

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? Integrative and Comparative Biology* **52:** 724–736.

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

Taylor DS, Turner BJ, Davis WP, Chapman BB. 2008. A novel terrestrial fish habitat inside emergent logs. *American Naturalist* **171:** 263–266.

Taylor MS, Hellberg ME. 2006. Comparative phylogeography in a genus of coral reef fishes: biogeographic and genetic concordance in the Caribbean. *Molecular Ecology* **15:** 695–707.

Turko AJ, Wright PA. 2015. Evolution, ecology and physiology of amphibious killifishes (Cyprinodontiformes). *Journal of Fish Biology* **87:** 815–835.

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

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*. *Evolutionary Ecology Research* **8:** 1475–1486.

Vrijenhoek RC. 1985. Homozygosity and interstrain variation in the self-fertilizing hermaphroditic fish, *Rivulus marmoratus*. *Journal of Heredity* **76:** 82–84.

Ward RD, Hanner R, Hebert PDN. 2009. The campaign to DNA barcode all fishes, FISH-BOL. *Journal of Fish Biology* **74:** 329–356.

Ward RD, Zemlak TS, Innes BH, Last PR, Hebert PDN. 2005. DNA barcoding Australia's fish species. *Philosophical Transactions of the Royal Society B: Biological Sciences* **360:** 1847–1857.

Webb SA, Graves JA, Macias-Garcia C, Magurran AE, Foighil DO, Ritchie MG. 2004. Molecular phylogeny of the livebearing Goodeidae (Cyprinodontiformes). *Molecular Phylogenetics and Evolution* **30:** 527–544.

Weibel AC, Dowling TE, Turner BJ. 1999. Evidence that an outcrossing population is a derived lineage in a hermaphroditic fish (*Rivulus marmoratus*). *Evolution* **53:** 1217–1225.

Weigt LA, Baldwin CC, Driskell A, Smith DG, Ormos A, Reyier EA. 2012. Using DNA barcoding to assess Caribbean reef fish biodiversity: expanding taxonomic and geographic coverage. *PLoS ONE* **7:** e41059.

Weir BS. 1996. *Genetic data analysis II: methods for discrete population genetics data*. Sunderland: Sinauer Associates.

Wright S. 1969. *Evolution and the genetics of populations. II. The theory of gene frequencies*. Chicago: University of Chicago Press.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article at the publisher's web-site:

Figure S1. Geographic distribution of major genetic clades of *Kryptolebias marmoratus* species complex. The sampling locations are labelled as in Table 1. Colour of shapes designates major genetic clades: green – Northern, purple – Central, orange – Southern. Pattern of filled shapes corresponds to *cox1* haplotypes on Figure 4. Assignment of populations depicted by open circles is based on microsatellite loci and *cytB*. Thin lines outline marine ecoregions (*sensu* Spalding *et al.*, 2007). Arrows indicate direction of ocean surface currents (after http://oceancurrents.rsmas.miami.edu/site-map.html).

Figure S2. Analysis of mitochondrial cytochrome B (*cytB*) gene of *K. marmoratus* species complex. The analysis was conducted primarily to establish the phylogenetic position of a *Kryptolebias* specimen from Turks and Caicos (T&C, GenBank MF555094), whose sequence was obtained from the Appendix in Weibel *et al.* (1999). Additional *cytB* sequences extracted from literature were for fish from Guanabo River, Cuba (GNB, LK022680; Ponce de Leon *et al.*, 2014), Panama (ALM, KX268503, Kim *et al.*, 2016), and Rio de Janeiro (RIO, MF555095, Weibel *et al.*, 1999). The remaining *cytB* sequences were obtained by us from individuals representing major clades as determined in the analyses of microsatellite loci and the mitochondrial *cox1* gene. Trees were constructed with neighbor-joining method and Kimura-2-parameter distances. a) Phylogenetic tree based on 378 bp fragment of *cytB* (N = 37); b) Phylogenetic tree based on 1085 bp fragment of *cytB* (N = 34); c) Map showing sampling locations of specimens used in the analysis – for detailed information see Supporting Informations, Table S1; d) Phylogenetic tree and genetic distances based on complete sequence of *cytB* gene.

Table S1. Cytochrome B (*cytB*) sequences and references.