SCIENTIFIC REPORTS

natureresearch

OPEN

Tempo and mode of allopatric divergence in the weakly electric fish *Sternopygus dariensis* in the Isthmus of Panama

Celestino Aguilar [1,2,3,4], Matthew J. Miller[1,5], Jose R. Loaiza [1,3,6], Rigoberto González[3], Rüdiger Krahe [7,8] & Luis F. De León [1,3,4*]

Spatial isolation is one of the main drivers of allopatric speciation, but the extent to which spatially-segregated populations accumulate genetic differences relevant to speciation is not always clear. We used data from ultraconserved elements (UCEs) and whole mitochondrial genomes (i.e., mitogenomes) to explore genetic variation among allopatric populations of the weakly electric fish *Sternopygus dariensis* across the Isthmus of Panama. We found strong genetic divergence between eastern and western populations of *S. dariensis*. Over 77% of the UCE loci examined were differentially fixed between populations, and these loci appear to be distributed across the species' genome. Population divergence occurred within the last 1.1 million years, perhaps due to global glaciation oscillations during the Pleistocene. Our results are consistent with a pattern of genetic differentiation under strict geographic isolation, and suggest the presence of incipient allopatric species within *S. dariensis*. Genetic divergence in *S. dariensis* likely occurred *in situ*, long after the closure of the Isthmus of Panama. Our study highlights the contribution of spatial isolation and vicariance to promoting rapid diversification in Neotropical freshwater fishes. The study of spatially-segregated populations within the Isthmus of Panama could reveal how genetic differences accumulate as allopatric speciation proceeds.

The closure of the Isthmus of Panama is one of the main drivers of Neotropical diversification. On the one hand, the rise of the Isthmus resulted in immediate reduction in gene flow between marine organisms on the two sides of the Isthmus[1–4]. On the other, the exposure of the land bridge facilitated dispersal and colonization in both terrestrial[5] and freshwater organisms[6,7]. In addition, global events such as the Pleistocene glaciations[8] have facilitated the expansion and contraction of local populations via changes in sea level[9–11]. Together, these events have defined the current spatial structure of the Isthmian populations[6,12–14], with implications for allopatric divergence across isolated populations.

Accordingly, spatial isolation of populations along and across the Isthmus of Panama is expected to facilitate the accumulation of genetic differences leading to the formation of new species[15–19]. However, the extent to which spatially segregated populations accumulate genetic differences relevant to allopatric divergence is not always clear. In addition, if divergence occurs, its historical context might be difficult to define, given the dynamic nature of the rise of the Isthmus of Panama. For instance, allopatric divergence might proceed randomly across the range of segregated populations, or it might be driven by multiple dispersal and colonization events[6–21]. Furthermore, divergence of seemingly isolated populations might be influenced by selective (i.e., local adaptation) and random processes (i.e., drift) that are difficult to disentangle if relying on low numbers of molecular markers[22]. Here, we explore these issues by quantifying genetic variation at both nuclear loci linked to ultraconserved elements

[1]Centro de Biodiversidad y Descubrimiento de Drogas, Instituto de Investigaciones Científicas y Servicios de Alta Tecnología (INDICASAT AIP), P. O. Box 0843-01103, Panamá, República de Panamá. [2]Department of Biotechnology, Acharya Nagarjuna University, Guntur, India. [3]Smithsonian Tropical Research Institute, Balboa Ancón, P.O. Box 0843-03092, Panamá, República de Panamá. [4]Department of Biology, University of Massachusetts Boston, Boston, MA, USA. [5]Sam Noble Oklahoma Museum of Natural History and Department of Biology, University of Oklahoma, Norman, OK, USA. [6]Programa Centroamericano de Maestría en Entomología, Universidad de Panamá, Panamá, República de Panamá. [7]Institut für Biologie, Humboldt-Universität zu Berlin, Berlin, Germany. [8]Department of Biology, McGill University, Montreal, QC, Canada. *email: luis.deleonreyna@umb.edu

Sample ID	Species	Region	River	Genbank		Reference
				BioSample	mtDNA	
PB05	*Sternopygus dariensis*	eastern	Chucunaque	SAMN09637294	MK530706	Present study
PB06	*Sternopygus dariensis*	eastern	Chucunaque	SAMN09637295	MH399590	Aguilar *et al.* 2019
PB07	*Sternopygus dariensis*	eastern	Chucunaque	SAMN09637296	—	Present study
Tu218	*Sternopygus dariensis*	western	Tumaganti	SAMN09637297	MK530707	Present study
Tu219	*Sternopygus dariensis*	western	Tumaganti	SAMN09637298	MH605315	Present study
Tu221	*Sternopygus dariensis*	western	Tumaganti	SAMN09637299	—	Present study
RSM01	*Sternopygus dariensis*	western	Santa María	SAMN09637300	MH605309	Present study
RSM02	*Sternopygus dariensis*	western	Santa María	SAMN09637301	MH605310	Present study
RSM03	*Sternopygus dariensis*	western	Santa María	SAMN09637302	MH605311	Present study
RSP02	*Sternopygus dariensis*	western	San Pablo	SAMN09637303	MH605312	Present study
RSP03	*Sternopygus dariensis*	western	San Pablo	SAMN09637304	MH605313	Present study
RSP04	*Sternopygus dariensis*	western	San Pablo	SAMN09637305	MH605314	Present study
Tu227	*Eigenmannia humboldtii*	outgroup	Tumaganti	SAMN09637306	—	Present study
PB01	*Eigenmannia humboldtii*	outgroup	Chucunaque	SAMN09637307	—	Present study
—	*Sternopygus arenatus*	outgroup	—	—	KX058571	Elbassiouny *et al.* 2016
—	*Sternopygus macrurus*	outgroup			MH263671	Rincón-Sandoval *et al.* 2018
—	*Eigenmannia humboldtii*	outgroup			MH263668	Rincón-Sandoval *et al.* 2018
—	*Eigenmannia sp.*	outgroup	—	—	AB054131	Saitoh *et al.* 2003
—	*Apteronotus rostratus*	outgroup			MH399592	Aguilar *et al.* 2019

Table 1. Sample information for *S. dariensis* and other Gymnotiformes included in the present study. Genbank short-read archive (BioSample) for ultraconserved elements (UCEs) and accession numbers for mitogenomes (mtDNA) are also provided.

(UCEs[23,24]; and whole mitochondrial genomes across geographically segregated populations of the weakly electric fish, *Sternopygus dariensis*, in the Isthmus of Panama.

Sternopygus dariensis (Meek & Hildebrand, 1916) is a geographically unique species within the Blue-green knifefish (*S. aequilabiatus*; Humboldt, 1805) species complex. Similar to other Neotropical electric fishes[25], this *Sternopygus* complex originated in South America, but it has since colonized the Isthmus of Panama. Although it has not been resolved whether *S. aequilabiatus* and *S. dariensis* are allospecies or distinct species, here we consider *S. dariensis* as an independent species following Hulen *et al.*[26] and Albert[27]. However, our inferences do not change if we consider *S. dariensis* as synonym of *S. aequilabiatus* as suggested by Maldonado *et al.*[28]. *Sternopygus dariensis* is narrowly distributed from the Pacific slope of Colombia to the Tabasará River in western Panama[26], which represents the effective range limit of the species. As with other weakly electric fishes, *S. dariensis* has an elongate eel-shaped body and the ability to produce electric organ discharges (EOD) used for electrolocation and communication[29,30]. Interestingly, despite evidence for genetic variation in other species of the genus *Sternopygus*[31,32] no population-level analysis has been performed within the Isthmus of Panama. Thus, quantifying genetic variation across Isthmian populations of *S. dariensis* will help inform the tempo and mode of allopatric divergence in Neotropical freshwater fishes.

By integrating available data from mitogenomes as well as UCEs, we examine i) the historical factors driving allopatric divergence across spatially segregated populations, and ii) the genetic consequences of allopatric divergence in the weakly electric fish *S. dariensis* across the Isthmus of Panama.

Results

Ultraconserved elements (UCEs). We recovered 150 UCE loci (32 were invariant) that had an average length of 861 bp shared across individuals of *S. dariensis*. The 118 variant loci contained 285 single nucleotide polymorphisms (SNPs), ranging from 1 to 8 SNPs per locus. The complete dataset including the outgroup *Eigenmannia humboldtii* had 98 enriched UCE loci with a total alignment length of 85 028 bp. UCE raw read data are available on NCBI SRA SRP071703 (BioProject PRJNA480353, see Table 1 for BioSample numbers). Overall, our RAxML phylogeny analysis using 98 UCEs showed high bootstrap support (100%) for one western and one eastern clade, comprising samples from San Pablo, Santa María, and Tumaganti, and samples from Chucunaque, respectively (Fig. 1C).

Similar results were obtained with the Bayesian clustering analyses (based on the 118 SNPs), which revealed high posterior probability for two well-defined independent clusters (K = 2), as identified by maximum ΔK value (Fig. 1B). These two clusters corresponded to western (Santa María, San Pablo and Tumaganti), and eastern (Chucunaque) Panama, respectively.

When examining genetic divergence between clades, we identified a total of 91 (77%) UCE loci containing 150 fixed SNPs. All UCE loci were successfully mapped to the available genomic scaffolds of the electric eel (*Electrophorus electricus*), and 83 loci (containing 129 fixed SNPs) were mapped to the channel catfish (*Ictalurus punctatus*) genome (Fig. 2; Supplementary Table S1). These 129 fixed SNPs were located in 22 of the 29 channel
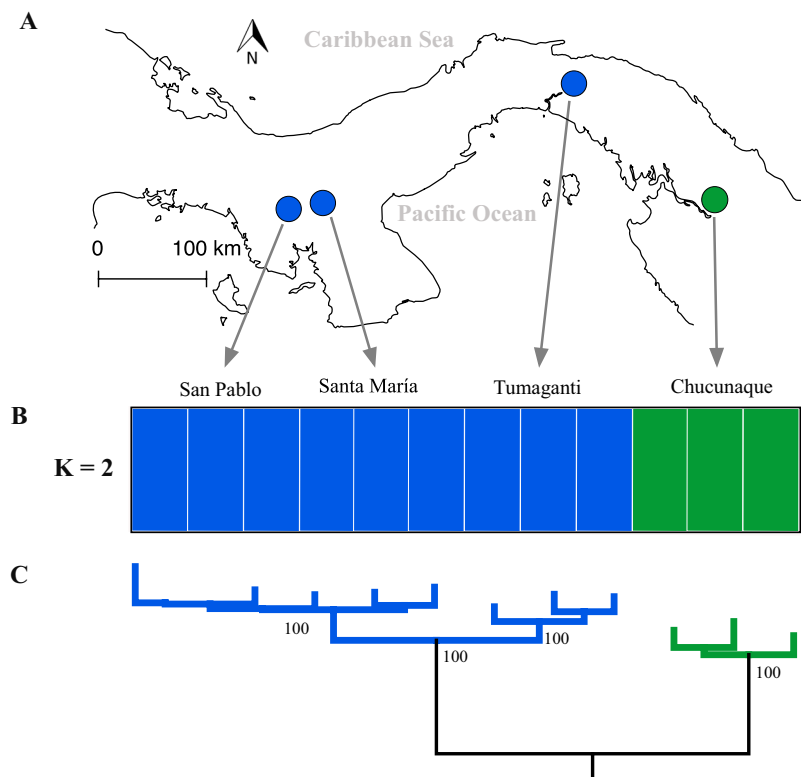
Figure 1. Phylogenetic reconstruction of allopatric populations of *S. dariensis* based on UCE data. The figure shows sampling sites (**A**), population structure with K = 2 representing the highest posterior probabilities as implemented in STRUCTURE and CLUMPPAK, for 118 SNPs (**B**), and maximum likelihood tree inferred by RAxML, for 98 UCEs sequences. Numbers at the nodes are bootstrap values (**C**).
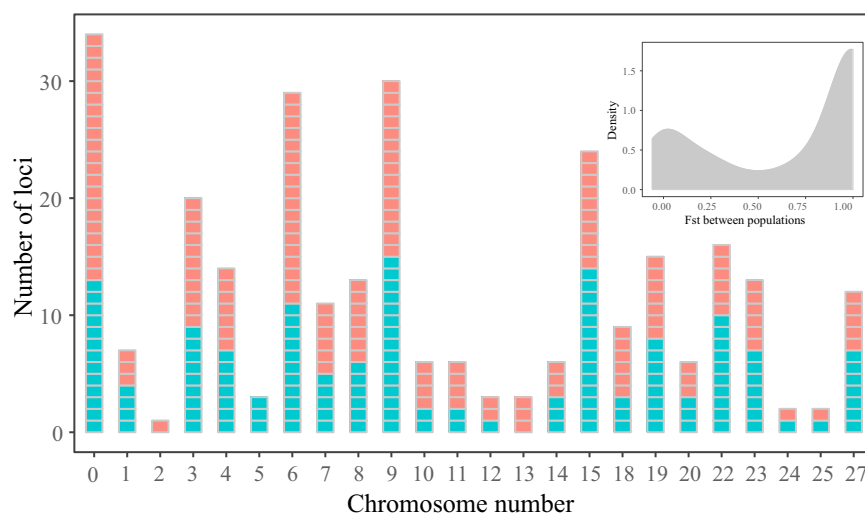

Figure 2. Chromosomal position of SNPs linked to UCEs in *S. dariensis*. Pink bars show highly differentiated SNPs between eastern and western populations within the Isthmus of Panama. Blue bars show non-differentiated SNPs. Chromosome mapping was done using the channel catfish reference genome. Chromosome "0" indicates a set of SNPs that did not map to any of the known chromosomes. The upper right inset shows the frequency distribution of F_{ST} values between eastern and western populations across the entire SNP dataset.

catfish chromosomes, and ranged from 1 to 18 SNPs per chromosome, with chromosome 6 showing the highest number of fixed SNPs (Fig. 2). In addition, 21 fixed SNPs did not map to any of the known channel catfish chromosomes. Across the complete data matrix (285 SNPs), only seven chromosomes (5, 16, 17, 21, 26, 28 and 29) did not show fixed SNPs (Fig. 2). The distribution of fixed SNPs across the electric eel scaffolds was widespread and appeared to show similar positions with respect to their location on the channel catfish chromosomes
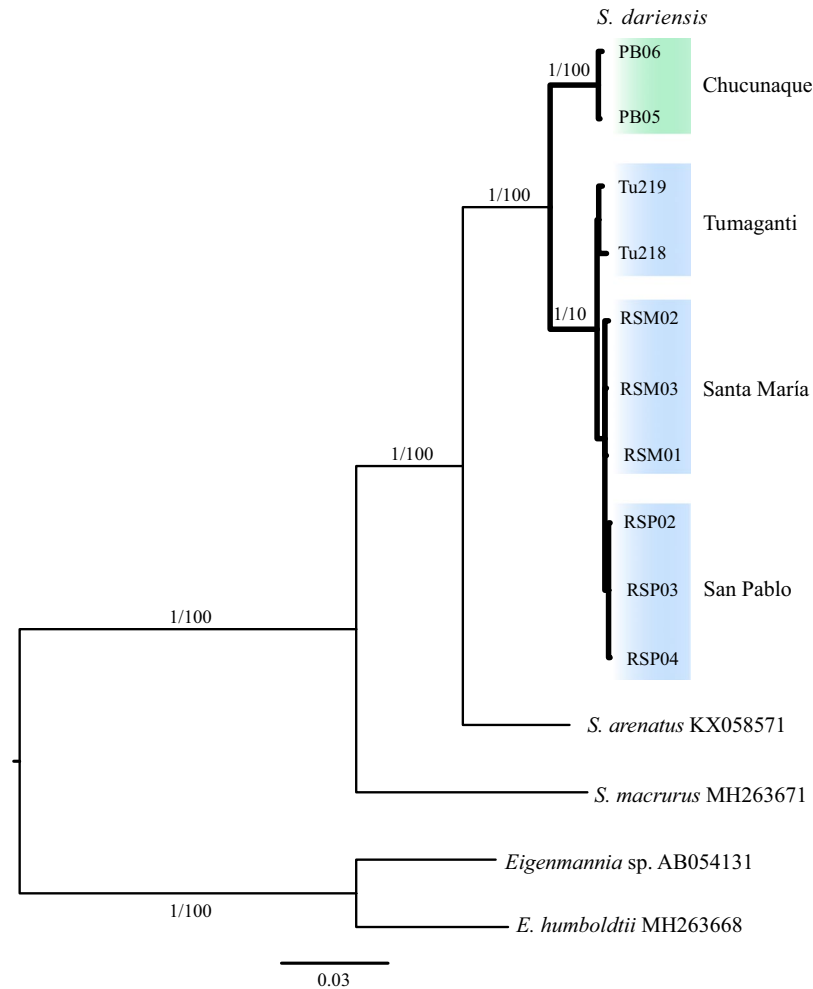
Figure 3. Phylogenetic relationships among *S. dariensis* based on RAxML and MrBayes. The phylogeny represents the best-scoring maximum likelihood tree based on 12 PCGs. Branch lengths are mean estimates. Numbers at the nodes are posterior probabilities and bootstrap values. The scale bar indicates relative branch lengths.

(Supplementary Table S1). In addition, the frequency distribution of F_{ST} estimates between clades was highly skewed toward large values (Mean = 0.65, Median = 1.0, Skewness = −0.62; Fig. 2). Finally, our analysis of outlier loci using BayeScan 2.1[33] failed to detect loci under directional selection (Mean q-val = 0.90). By contrast, the PCAdapt analysis identified 12 potential outlier loci associated with differences between the two lineages, at FDR of 0.05 (Fig. S1). These SNPs occurred over 11 separate chromosomes of the channel catfish genome.

Mitogenomes. We recovered a total of 9 complete mitogenomes: 8 from the present study (GenBank accession nos. MH605309-MH605315 and MK530706) and one assembled previously (MH399590[34]) (Table 1). We also were able to retrieve a partial mitogenome from 1 individual from Tumaganti (GenBank accession no. MK530707). Both maximum likelihood and Bayesian phylogenetic analyses based on a concatenated dataset of 12 protein-coding genes (PCGs) derived from whole mitogenomes yielded a monophyletic relationship among samples of *S. dariensis* collected in Panama (Fig. 3). In agreement with the UCE results, we detected two highly supported (BS, 100%; PP, 1.0) phylogenetic clades within *S. dariensis* (Fig. 1C). One clade comprised samples from the three western populations (San Pablo, Santa María and Tumaganti rivers), and the other encompassed samples from the eastern population of the Chucunaque River (Fig. 3). Furthermore, the most western populations of *S. dariensis*, San Pablo and Santa María and Tumaganti, showed similar genetic distances among them (~0.18%), but the greatest genetic distance (2.83%, SE = 0.25%) from the eastern population of the Chucunaque River (Table 2). Across species, we found high levels of genetic divergence between *S. dariensis* and its most closely related species, *S. arenatus* (6.6%, SE = 0.91%) (Table 3).

TCS[35] haplotype networks were characterized by two centrally shared haplotypes, corresponding to the eastern and western populations, which were separated by up to 13 mutational events. In addition, there were three peripheral haplotypes. The most common haplotype was shared by the western populations (Santa María and San Pablo) and represented up to 45% of all sampled individuals. The second-most common haplotype was unique to the eastern population (Chucunaque River; Supplementary Fig. S2). These results suggest the presence of two distinct genetic groups spanning the four geographical locations sampled.

Locality	Chucunaque	Tumaganti	Santa María
Tumaganti	0.02807		
Santa María	0.02845	0.00366	
San Pablo	0.02839	0.00354	0.00049

Table 2. Pairwise uncorrected p-distances between populations of *Sternopygus dariensis* based on 12 PCGs.

Species	*S. dariensis*	*S. arenatus*	*S. xingu*
S. arenatus	0.06599		
S. xingu	0.12034	0.10404	
S. macrurus	0.14441	0.12422	0.13509

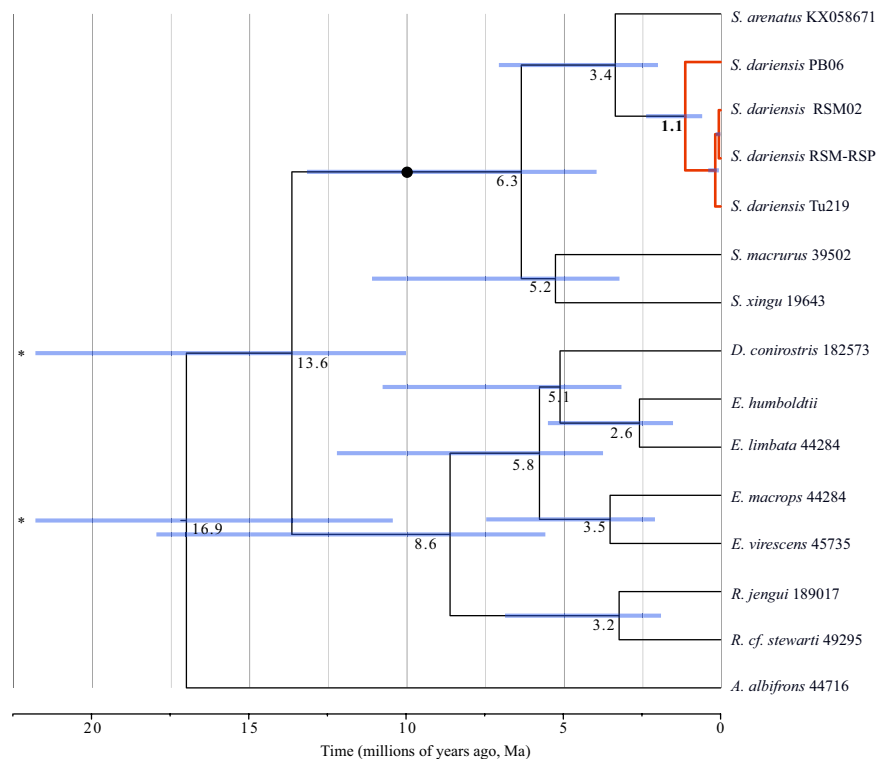**Table 3.** Pairwise uncorrected p-distances between *Sternopygus* species based on *cox1* gene.

Figure 4. Divergence time estimates among *Sternopygus* species. Results are based on 2217 bp of concatenated *cox1*, *cytb* and *16S* with posterior probabilities from Bayesian Inference (BI) shown in grey, reconstructed using BEAST. Nodes are individually numbered, with gray bars indicating 95% confidence intervals on divergence time estimates (in My before present). Calibrated node is indicated with a black circle. Asterisks indicate that confidence interval bars are truncated.

Time to the most recent common ancestor (TMRCA). Mitochondrial-based dating of TMRCA placed the first split within Sternopygidae around 13.6 Ma (95% HPD: 10–28.4 Ma; Fig. 4), separating members of Eigenmanninae and *Sternopygus*. Two clades separated 6.3 Ma (95% HPD: 2.58–12.28 Ma) were identified within *Sternopygus*, the first one including *S. arenatus* and *S. dariensis*; the second clade included *S. macrurus* and *S. xingu*. The split within the first clade (*S. arenatus* and *S. dariensis*) occurred during the Pliocene period, approximately 3.4 Ma (95% HPD: 1.2–6.7 Ma), while divergence between western and eastern clades of *S. dariensis* took place in the Pleistocene, approximately 1.1 Ma (95% HPD: 0.5–2.6 Ma).

Discussion

Allopatric divergence most often occurs when populations accumulate random genetic differences as a byproduct of spatial isolation[19,36,37]. However, the extent to which spatially segregated populations show progress towards allopatric divergence is not always obvious. Here, we explore this issue by using a panel of over 285 UCE, SNPs and complete mitogenome sequences in the weakly electric fish *S. dariensis* in the context of the biogeographic history of the Isthmus of Panama.

Overall, our analysis of population structure based on UCEs revealed two divergent clusters (without evidence for admixture) that corresponded to eastern (i.e., Chucunaque/Tuira) and western populations (Santa María, San Pablo and Tumaganti; Fig. 1B). Indeed, over 77% of UCE loci, distributed across 22 of 29 chromosomes of the channel catfish and across available scaffolds of the electric eel, contained SNPs differentially fixed between these two genetic clades (Fig. 2). In addition, the frequency distribution of F_{ST} estimates across the entire dataset was highly skewed toward large values (Fig. 2). Thus, our results are consistent with a pattern of genetic differentiation under strict geographic isolation (i.e., a model of allopatric divergence[38]) and suggest that *S. dariensis* is composed of two genetic lineages, likely representing incipient allopatric species.

Tempo and mode of allopatric divergence in *S. dariensis*.

Allopatric divergence in primary freshwater fishes, including weakly electric fishes, in the context of Lower Mesoamerica is often associated with multiple colonization events during the dynamic formation of the Isthmus of Panama[6,12]. However, we showed that substantial progress towards allopatric speciation could also occur *in situ* following the closure of the Isthmus. This was supported by our phylogenetic analysis, which indicated that the Isthmian populations of *S. dariensis* constitute a monophyletic group (Fig. 3), likely derived from the South American *S. arenatus* over 3 Ma ago (Fig. 4). Thus, allopatric divergence in *S. dariensis* is likely the result of a single colonization event, followed by segregation of populations via colonization and dispersal. In addition, our analyses of mitogenome sequences revealed that divergence between eastern and western populations of *S. dariensis* is substantial (2.8%; Table 2) when compared to divergence across established *Sternopygus* species (6.6% between *S. dariensis* and *S. arenatus*; Table 3). This suggests an early geographic split in the mitochondrial genome between these populations, which occurred ~1.1 Ma (Fig. 4).

The historical context of this divergence is consistent with a scenario of early divergence between populations from Tuira/Chucunaque and Tumaganti watersheds, followed by more recent dispersal from Tumaganti to the western watersheds (Santa María and San Pablo). This scenario could explain the low genetic variation across the three western populations (i.e., Tumaganti, Santa María and San Pablo), possibly due to a strong bottleneck (or founder effect) or a limited time for the accumulation of genetic differences[39–41]. This expansion and contraction of populations was likely facilitated by changes in sea level during the Pleistocene glaciation oscillations[42]. For instance, during periods of low sea level, large portions of the eastern Pacific seabed remained exposed[43–45], likely facilitating the exchange of freshwater fishes between eastern and western watersheds[39,44,46]. By contrast, during periods of high sea level, riverine populations likely became spatially isolated due to the intrusion of seawater[39,47,48]. It is important to notice that analyses of behavioral and genetic incompatibility may be necessary to confirm if these allopatric populations represent reproductively isolated populations/species within *S. dariensis*. For example, previous studies on electric fishes have highlighted the role of the temporal (shape) properties of the EODs in maintaining prezygotic isolation and speciation[49,50]. Cytogenetic analyses have also been used to test for genetic incompatibility Gymnotiformes[51,52]. Overall, however, our findings indicate that substantial genetic divergence has accumulated within the Isthmus of Panama. This also suggests that the diversity of *Sternopygus* species might have been underestimated, possibly due to the existence of cryptic species across the Isthmian range of the genus.

Exploring the genetic consequences of allopatric divergence in *S. dariensis*.

An expected consequence of allopatric divergence is the accumulation of genetic differences across large portions of the species' genome due to random drift[38,53,54]. However, allopatric divergence may also be affected by adaptive processes (e.g., local adaptation[55,56]), which might lead to genomic islands of divergence[57,58]. Distinguishing between these possibilities requires a large number of genetic markers that can be mapped across the genome[59,60]. Although our analysis only included a few individuals per population and a reduced representation of the species' genome (285 SNPs), the high number of fixed loci between populations, and the broad distribution of these loci across the majority of the species' chromosomes, is consistent with a model of strict allopatric divergence[38,53]. In addition, the low number of loci under selection also points towards the role of non-adaptive processes as the predominant driver of divergence in *S. dariensis*. For instance, Picq et al.[6] found sizeable evidence of the role of drift in EOD evolution. However, further work is needed to disentangle the contribution of both factors to *in situ* divergence in *S. dariensis* post-colonization of the Isthmus of Panama. This work could benefit from inclusion of additional sampling sites, and a larger number of individuals to examine genetic structure across the entire range of the species. In addition, future work should explore the geographic origins of *S. dariensis* in South America and its pattern of colonization of the Isthmus of Panama.

In summary, our results suggest that allopatric divergence in *S. dariensis* progressed *in situ* post-colonization of the Isthmus of Panama, and that genetic divergence is likely to occur widely across the species' genome. Thus, our findings support the pattern of genetic differentiation expected under the classical model of allopatric divergence. Although reproductive isolation between allopatric populations is difficult to determine in nature, we suggest that integrating UCEs and mitogenome data is a useful approach to assess the evolution of genetic incompatibilities between geographically isolated populations. Our findings also underscore the role of vicariance and spatial isolation in promoting diversification in weakly electric fishes. Although more work is needed to disentangle the demographic and phylogeographic history of *S. dariensis* across its range, the study of spatially segregated populations within the context of the Isthmus of Panama could provide further insights into the accumulation of genetic differences that drive allopatric diversification in freshwater fishes.

Methods

Taxon sampling and ethical statement. Fish were first detected using a wire electrode connected to a mini-amplifier (Radioshack, Fort Worth, TX), and then collected using a dip-net. Fish were then euthanized with an overdose of eugenol ($C_{10}H_{12}O_2$) derived from clove oil. Twelve specimens of *Sternopygus dariensis* were collected from four sites in the Republic of Panama (Fig. 1A): San Pablo River (SP: 8°11′24″ N, −81°15′0″ W), Santa María River (SM: 8°13′50″ N, −80°58′12″ W), Tumaganti River (TU: 9°13′12″ N, −78°52′47″ W) and Quebrada La Hoya stream, which flows into the Chucunaque River (PB: 8°15′0″ N–77°43′11″ W). Two individuals of *Eigenmannia humboldtii* were also sampled and used as outgroups in the phylogenetic analysis.

UCE library preparation and sequencing. We extracted DNA from frozen-preserved tissues following the Qiagen DNeasy Blood and Tissue Kit (Qiagen, Valencia, CA.) protocol. We quantified all DNA extracts with a Qubit Fluorometer (Life Technologies, Inc.), assessed each extract for quality by electrophoresis, and then sheared extracts by sonication to a target size of 600 bp on a Covaris S220 instrument (Covaris, Woburn, Massachusetts, USA). Approximately, 500 ng of genomic DNA was used to prepare twelve 300 bp paired-end libraries, following the protocol Kapa Kit (New England Biolabs, Ipswich, MA, USA), and the UCE protocol described in Faircloth *et al.*[61], available online from http://ultraconserved.org. We enriched libraries for UCE targets using Arbor Biosciences UCE Capture Kits (myBaits UCE Actinopterygians 0.5Kv1) designed to target 500 highly conserved loci across fishes. We determined the size of enriched and purified pools with a Bioanalyzer (Agilent Technologies, Inc.), and quantified enriched libraries by qPCR (Kapa Biosystems) prior to sequencing. Genomic libraries were multiplexed before sequencing 300 bp from both ends on the Illumina MiSeq platform at Naos Molecular Laboratory of the Smithsonian Tropical Research Institute (STRI), in Panama City, Panama.

UCEs bioinformatics. Raw sequence data were converted to FASTQ before demultiplexing, trimming and cleaning using Illumiprocessor[62], which works with Trimmomatic[63]. We then followed the standard PHYLUCE[64] pipeline (http://phyluce.readthedocs.io/en/latest) for processing target-enriched UCEs data. Trinity version r2013-02-25[65,66] was used to assemble reads using the script *assemblo_trinity.py*. We then used the *match_contigs_to_probes.py* script to map assembled contigs to the UCE probes, which allowed us to assemble contigs representing enriched UCE loci from each species. We created two FASTA datasets: the first containing both the ingroup and outgroup taxa, and the second with only the ingroup, to increase the number of shared loci. We then followed the PHYLUCE pipeline to produce MAFFT[67] alignments across all loci from both datasets. We removed locus names from each alignment, and created a complete dataset, with each locus containing data for all the individuals. Alignments of the first dataset (containing both ingroup and outgroup) were concatenated in PHYLIP format for subsequent analyses of phylogenetic structure (see below).

We chose the sample with the most UCE contigs recovered within the ingroup as reference sequence to call SNPs. We mapped reads (per individual) to this reference using the program BWA[68]. We converted SAM files to Binary Alignment Map (BAM) files format using the SAMtools[69], and Picard (http://broadinstitute.github.io/picard/) to identify and remove PCR duplicates. We added read groups for each individual using Picard, and merged the BAM files across individuals with the SAMtools. We used the Genome Analysis Toolkit (GATK; McKenna *et al.*[70] to identify and realign indels, to call and annotate SNPs and indels, and to mask SNP calls around indels, following the population genomics pipeline for UCEs developed by Faircloth and Harvey (https://github.com/mgharvey/seqcap_pop). This included restricting data to high-quality SNPs (Q30), and read-back phasing in GATK. At the end of the pipeline we created a Variant Call File (VCF) with nuclear SNPs. Finally, we used Python scripts from the seqcap_pop pipeline to convert the phased VCF into an input file for subsequent population analyses.

UCE analyses. To quantify phylogenetic structure across species (i.e., the first dataset), we performed unpartitioned concatenated maximum-likelihood (ML) analyses, using RAxML 8.0.19[71]. Support for the best ML topology was assessed by performing 1000 nonparametric bootstraps in CIPRES Science Gateway[72]. For this analysis we used the GTR GAMMA site-rate substitution model for the best ML tree obtained from CIPRES Science Gateway[72], using JModelTest 2[73].

To estimate population structure across *S. dariensis* populations, we used a Bayesian clustering algorithm as implemented in STRUCTURE 2.3.4[74]. One random SNP from each locus was selected using the script *structure_from_vcf.py* (github.com/mgharvey/seqcap_pop) to create STRUCTURE input file, to minimize linkage disequilibrium. The number of subpopulations (ΔK) was determined using the ad-hoc statistical method, based on the rate of change in the log probability of data between successive K values. Ten independent runs for K values ranging from 1 to 4 were performed with a burn-in length of 50,000, followed by 500,000 interactions. Best K results were analyzed on Structure Harvester[75] and Clumpak[76].

To further examine divergence across *S. dariensis* populations, we quantified the proportion of SNPs that were differentially fixed (F_{ST})[77] between the two major clades recovered from the STRUCTURE analysis (see results). These analyses were performed in the R package adegenet version 3.2.2[78,79], and was performed on the 285 SNPs data set using the script *adegent_from_vcf.py* (github.com/mgharvey/seqcap_pop). To determine the genomic distribution of differentiated SNPs, we mapped each fixed SNP (UCE locus) to the available scaffolds of the electric eel, *Electrophorus electricus*[80] genome using BLAST on SequenceServer (http://www.sequenceserver.com) implemented in EFISH genomics (https://efishgenomics.integrativebiology.msu.edu/blast_search/). We also mapped the fixed SNPs to the channel catfish, *Ictalurus punctatus*[81] chromosomes, using the NCBI Genome Workbench version 2.12.8. In cases in which we obtained multiple hits, we retained the hits with >90% sequence identity and the highest query coverage. Finally, to examine the pattern of genetic divergence between clades, we estimated the frequency distribution, including skewness, of F_{ST} values across the entire SNP dataset.

Outlier analyses. We quantified outlier loci with two methods: BayeScan 2[33] and PCAdapt[82]. BayeScan uses differences in allele frequencies between populations, and estimates the probability that each locus is subject to selection using a Bayesian method. BayeScan was run under default settings. In PCAdapt, population structure is defined with PCA, and outliers are detected with respect to their contribution to population structure. Cattell's graphical rule was used to choose the number of principal components (K) that identify potential SNPs under selection. Outliers were selected by performing the *q*-value procedure at a false discovery rate (FDR) of 0.05 using the R package qvalue[83]. PCAdapt was run assuming three genetic clusters after graphical evaluation of the eigenvalues according to Luu *et al.*[82]. To account for population structure, we retained the first two PC axes that explained most variation. For both analyses, the input file was created using the previously generated VCF file, and it was converted to other formats, as needed, using PGD Spider[84] for BayeScan and Plink v1.9[85] for PCAdapt.

Recovery of mitogenomes. We identified mitogenomes from UCEs *off-target* reads by following the same methodology described in Aguilar *et al.*[34]. The complete sequences of mitogenomes were annotated in Geneious version 11.1.4[86] using the complete mitochondrial genome sequence of *S. dariensis* (GenBank accession no. MH399590) as a reference.

Phylogenetic analyses of complete mitochondrial genomes. In order to compare mitochondrial gene sequences, we extracted the protein coding genes (*nad6* and stop codons excluded). We aligned them using Multiple Alignment using Fast Fourier Transform (MAFFT)[87] in Geneious version 11.1.4[86]. *Sternopygus arenatus*[88] and *Eigenmannia sp.* mitogenomes[89] were used as outgroup (Table 1). We manually checked the alignments to correct for annotation errors based on consensus, and removed positions with long gaps, as well as regions with uncertain alignment, ambiguous portions and stop codons. A General Time Reversible (GTR) model, with a proportion of invariable sites (I) and heterogeneity of substitution rates among sites using gamma distribution (G), was selected as the preferred model of nucleotide sequence evolution by jModelTest2 on XSEDE[73], performed in the CIPRES Science Gateway[72], with corrected Akaike information criterion (AIC). Maximum likelihood (ML) analyses were performed using RAxML. The resulting tree topology was evaluated by a rapid bootstrap analysis with 1000 replicates. Bayesian inference (BI) analyses were performed in MrBayes on XSEDE version 3.2.6 on CIPRES Science Gateway[72]. We ran two independent runs of 2,000,000 generations for each of the four chains. Each chain was sampled every 2000 generations with a burn-in of 25%. Trees inferred prior to stationarity were discarded as burn-in, and the remaining trees were constructed using a 50% majority-rule consensus tree with posterior probabilities.

We also assessed the evolutionary distance (uncorrected p-distance) among *S. dariensis* populations across all of the genes (concatenated) using uncorrected p-distances among unique haplotypes with MEGA version 7[90]. In addition, we calculated pairwise distance across *Sternopygus* species (*S. dariensis*, *S. arenatus*, *S. xingu* and *S. macrurus*), using the *cox1* gene. Standard error estimates were obtained by a 1000 bootstrap replicates under a maximum likelihood model[91]. Phylogenetic relationships between haplotypes were determined by constructing a Templeton, Crandall and Sing (TCS) haplotype network[35] based on *cox1* sequences in PopART 1.7 (Population Analysis with Reticulate Trees[92]). One previously published *cox1* (645 bp) sequence[6] from Tuira River from Eastern Panama, was aligned independently to the newly generated mitogenomes.

Time to the most recent common ancestor. We estimated divergence time among populations and species using BEAST2 on XSEDE v2.4.8[93] on the CIPRES Science Gateway[72] under a relaxed molecular clock with uncorrelated lognormal distribution of rates, to allow for lineage-specific rate variation, and using a Yule speciation model as tree prior. Sequences of *cox1*, *cytb* and *16S* were concatenated on a 2217 bp alignment. Nucleotides were first grouped into four different partitions: the 1st, 2nd, and 3rd codon positions respectively for *cox1* and *cytb*; and the *16S*. Nucleotide substitution models were estimated for each partition using the Bayesian Information Criterion in Partition Finder[94]. The best-fit models for the three genes were TVM + I + G (for the *16S* region, and the 1st and 2nd codon positions), and the TRN + G (for the 3rd codon position of the coding genes). Bayesian posterior distributions of evolutionary rates were estimated using Markov Chain Monte Carlo (MCMC) procedure. Four independent MCMC chains were run for 4×10^7 generations and were sampled every 1,000 generations, after discarding the first 20% as burn-in. We used TRACER version 1.6[95] to test for convergence of the chains to the stationary distribution, which was determined by an effective size (ESS) of over 200[95]. The four independent runs were combined using LogCombiner v1.8, and the dates of divergence along with their 95% confidence intervals (HPD) were estimated using Tracer v1.6. The resulting phylogeny and the 95% HPD for the dates of divergence for the major clades were visualized using FigTree v1.4[96].

To calibrate the molecular clock, we used a fossil of *Humboldtichthys kirschbaumi* dated to c. 10 Ma, which shares morphological characters (i.e., opercle) with extant *Sternopygus* species[97]. The fossil age was placed on the stem node of the genus *Sternopygus*, using an uniform prior distribution with a maximum age equal to the maximum tree root height, following Picq *et al.*[6]. We used the three sequences (*cox1*, *cytb* and *16S*) to include all *Sternopygus* species available in Genbank and four genera of Gymnotiformes as outgroups (Supplementary Table S2).

Ethics statement. Sampling permit was obtained from the Panamanian Ministry of Environment (Permit number SE/A-100-14). This research was approved by the Institutional Animal Care and Use Committee (IACUC) at the Instituto de Investigaciones Científicas y Servicios de Alta Tecnología (INDICASAT AIP), and all methods were performed according to the guidelines and regulations of the approved protocol (IACUC-16-001).

References

1. Lessios, H. A. The great American schism: divergence of marine organisms after the rise of the Central American Isthmus. *Annu. Rev. Ecol. Evol. Syst.* **39**, 63–91 (2008).
2. Miura, O., Torchin, M. E. & Bermingham, E. Molecular phylogenetics reveals differential divergence of coastal snails separated by the Isthmus of Panama. *Mol. Phylogenet. Evol.* **56**, 40–48 (2010).
3. Galván-Quesada, S. *et al.* Molecular phylogeny and biogeography of the amphidromous fish genus *Dormitator* Gill 1861 (Teleostei: Eleotridae). *PLoS One* **11**, e0153538 (2016).
4. O'Dea, A. *et al.* Formation of the Isthmus of Panama. *Sci Adv* **2**, e1600883 (2016).
5. Woodburne, M. O. The Great American Biotic Interchange: dispersals, tectonics, climate, sea level and holding pens. *J. Mamm. Evol.* **17**, 245–264 (2010).
6. Picq, S., Alda, F., Krahe, R. & Bermingham, E. Miocene and Pliocene colonization of the Central American Isthmus by the weakly electric fish *Brachyhypopomus occidentalis* (Hypopomidae, Gymnotiformes). *J. Biogeogr.* **41**, 1520–1532 (2014).
7. McMahan, C. D. *et al.* Pleistocene to holocene expansion of the black-belt cichlid in Central America, *Vieja maculicauda* (Teleostei: Cichlidae). *PLoS One* **12**, e0178439 (2017).
8. Roberts, T. R. Ecology of fishes in the Amazon and Congo basins. *Bull Mus Comp Zool*, Vol 143, No 2, p 117–147, 1972. (1972).
9. Lambeck, K., Yokoyama, Y. & Purcell, T. Into and out of the Last Glacial Maximum: sea-level change during oxygen isotope stages 3 and 2. *Quat. Sci. Rev.* **21**, 343–360 (2002).
10. Dias, M. S. *et al.* Global imprint of historical connectivity on freshwater fish biodiversity. *Ecol. Lett.* **17**, 1130–1140 (2014).
11. Swartz, E. R., Chakona, A., Skelton, P. H. & Bloomer, P. The genetic legacy of lower sea levels: does the confluence of rivers during the last glacial maximum explain the contemporary distribution of a primary freshwater fish (*Pseudobarbus burchelli*, Cyprinidae) across isolated river systems? *Hydrobiologia* **726**, 109–121 (2014).
12. Bermingham, E. & Martin, A. P. Comparative mtDNA phylogeography of neotropical freshwater fishes: testing shared history to infer the evolutionary landscape of lower Central America. *Mol. Ecol.* **7**, 499–517 (1998).
13. Reeves, R. G. & Bermingham, E. Colonization, population expansion, and lineage turnover: phylogeography of Mesoamerican characiform fish. *Biol. J. Linn. Soc. Lond.* **88**, 235–255 (2006).
14. Ingley, S. J., Reina, R. G., Bermingham, E. & Johnson, J. B. Phylogenetic analyses provide insights into the historical biogeography and evolution of Brachyrhaphis fishes. *Mol. Phylogenet. Evol.* **89**, 104–114 (2015).
15. Faulks, L. K., Gilligan, D. M. & Beheregaray, L. B. Islands of water in a sea of dry land: hydrological regime predicts genetic diversity and dispersal in a widespread fish from Australia's arid zone, the golden perch (*Macquaria ambigua*). *Mol. Ecol.* **19**, 4723–4737 (2010).
16. Loxterman, J. L. & Keeley, E. R. Watershed boundaries and geographic isolation: patterns of diversification in cutthroat trout from western North America. *BMC Evol. Biol.* **12**, 38 (2012).
17. Hung, C.-M., Drovetski, S. V. & Zink, R. M. Recent allopatric divergence and niche evolution in a widespread Palearctic bird, the common rosefinch (*Carpodacus erythrinus*). *Mol. Phylogenet. Evol.* **66**, 103–111 (2013).
18. Taugbøl, A., Junge, C., Quinn, T. P., Herland, A. & Vøllestad, L. A. Genetic and morphometric divergence in threespine stickleback in the Chignik catchment, Alaska. *Ecol. Evol.* **4**, 144–156 (2014).
19. Mayr, E. *Systematics and the origin of Species, from the Viewpoint of a Zoologist.* (Harvard University Press, 1942).
20. Alda, F., Reina, R. G., Doadrio, I. & Bermingham, E. Phylogeny and biogeography of the *Poecilia sphenops* species complex (Actinopterygii, Poeciliidae) in Central America. *Mol. Phylogenet. Evol.* **66**, 1011–1026 (2013).
21. Turchetto-Zolet, A. C., Pinheiro, F., Salgueiro, F. & Palma-Silva, C. Phylogeographical patterns shed light on evolutionary process in South America. *Mol. Ecol.* **22**, 1193–1213 (2013).
22. Avise, J. Intraspecific Phylogeography: The mitochondrial-DNA bridge between population genetics and systematics. *Annu. Rev. Ecol. Syst.* **18**, 489–522 (1987).
23. McCormack, J. E. *et al.* Ultraconserved elements are novel phylogenomic markers that resolve placental mammal phylogeny when combined with species-tree analysis. *Genome Res.* **22**, 746–754 (2012).
24. Faircloth, B. C. *et al.* Ultraconserved elements anchor thousands of genetic markers spanning multiple evolutionary timescales. *Syst. Biol.* **61**, 717–726 (2012).
25. Albert, J. S. & Crampton, W. G. R. Diversity and phylogeny of Neotropical electric fishes (Gymnotiformes). In *Electroreception* (eds Bullock, T. H., Hopkins, C. D., Popper, A. N. & Fay, R. R.) 360–409, https://doi.org/10.1007/0-387-28275-0_13 (Springer New York, 2005).
26. Hulen, K. G., Crampton, W. G. R. & Albert, J. S. Phylogenetic systematics and historical biogeography of the Neotropical electric fish Sternopygus (Teleostei: Gymnotiformes). *System. Biodivers.* **3**, 407–432 (2005).
27. Albert, J. S. Family Sternopygidae. Checklist of the freshwater fishes of South and Central America. *Edipucrs, Porto Alegre* 493–497 (2003).
28. Ocampo, J. A. M., Buckup, P. A., Gomes, J. A. & Lovejoy, N. R. Sternopygus species relationships, molecular and morphological evidence. *Front. Behav. Neurosci.* **6** (2012).
29. Moller, P. Electric fishes: history and behavior. **17**, Chapman & Hall (1995).
30. Bullock, T. H., Fay, R. R., Hopkins, C. D. & Popper, A. N. Electroreception. *Springer Handbook of Auditory Research.* (Springer, 2005).
31. Rincón-Sandoval, M., Betancur-R, R. & Maldonado-Ocampo, J. A. Mitochondrial genomes of the South American electric knifefishes *Eigenmannia humboldtii* (Steindachner 1878), *Eigenmannia limbata* (Schreiner and Miranda Ribeiro 1903), *Sternopygus aequilabiatus* (Humboldt 1805) and *Sternopygus macrurus* (Bloch and Schneider 1801), (Gymnotiformes, Sternopygidae). *Mitochondrial DNA Part B* **3**, 572–574 (2018).
32. Rincon-Sandoval, M., Betancur-R, R. & Maldonado-Ocampo, J. A. Comparative phylogeography of Trans-Andean freshwater fishes based on genome-wide nuclear and mitochondrial markers. *Mol. Ecol.* **28**:1096–1115, https://doi.org/10.1111/mec.15036 (2019).
33. Foll, M. & Gaggiotti, O. A genome-scan method to identify selected loci appropriate for both dominant and codominant markers: a Bayesian perspective. *Genetics* **180**, 977–993 (2008).
34. Aguilar, C., Miller, M. J., Loaiza, J. R., Krahe, R. & De León, L. F. Mitogenomics of Central American weakly-electric fishes. *Gene* **686**, 164–170, https://doi.org/10.1016/j.gene.2018.11.045 (2019).
35. Templeton, A. R., Crandall, K. A. & Sing, C. F. A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping and DNA sequence data. III. *Cladogram estimation. Genetics* **132**, 619–633 (1992).
36. Bush, G. L. Modes of animal speciation. *Annu. Rev. Ecol. Syst.* **6**, 339–364 (1975).
37. Kenney, A. M. & Sweigart, A. L. Reproductive isolation and introgression between sympatric Mimulus species. *Mol. Ecol.* **25**, 2499–2517 (2016).
38. Feder, J. L., Flaxman, S. M., Egan, S. P., Comeault, A. A. & Nosil, P. Geographic mode of speciation and genomic divergence. *Annu. Rev. Ecol. Evol. Syst.* **44**, 73–97 (2013).
39. Albert, J. S. & Reis, R. *Historical biogeography of Neotropical freshwater fishes.* (University of California Press, 2011).
40. Seehausen, O. & Wagner, C. E. Speciation in freshwater fishes. *Annu. Rev. Ecol. Evol. Syst.* **45**, 621–651 (2014).
41. Souza, C. da S. de, Oliveira, C. & Pereira, L. H. G. *Knodus moenkhausii* (Characiformes: Characidae): one fish species, three hydrographic basins - a natural or anthropogenic phenomenon? *DNA Barcodes* **3** (2015).
42. Clark, P. U. *et al.* The Last Glacial Maximum. *Science* **325**, 710–714 (2009).

43. Coates, A. G. *et al*. Closure of the Isthmus of Panama: The near-shore marine record of Costa Rica and western Panama. *Geol. Soc. Am. Bull.* **104**, 814–828 (1992).
44. Coates, A. G. & Obando, J. A. The geologic evolution of the Central American Isthmus. *Evolution and environment in tropical America* 21–56 (1996).
45. Clark, P. U. & Mix, A. C. Ice sheets and sea level of the Last Glacial Maximum. *Quat. Sci. Rev.* **21**, 1–7 (2002/1).
46. Schaefer, S. A. *et al*. Phylogeny and classification of Neotropical fishes. *Phylogeny and classification of Neotropical fishes* (1998).
47. Ochoa, L. E. *et al*. Genetic structure and historical diversification of catfish *Brachyplatystoma platynemum* (Siluriformes: Pimelodidae) in the Amazon basin with implications for its conservation. *Ecol. Evol.* **5**, 2005–2020 (2015).
48. Mondin, L. A. *et al*. Genetic pattern and demographic history of *Salminus brasiliensis*: Population expansion in the Pantanal Region during the Pleistocene. *Front. Genet.* **9**, 1 (2018).
49. Crampton, W. G. R., Lovejoy, N. R. & Waddell, J. C. Reproductive character displacement and signal ontogeny in a sympatric assemblage of electric fish. *Evolution* **65**, 1650–1666 (2011).
50. Crampton, W. G. R., Rodríguez-Cattáneo, A., Lovejoy, N. R. & Caputi, A. A. Proximate and ultimate causes of signal diversity in the electric fish Gymnotus. *J. Exp. Biol.* **216**, 2523–2541 (2013).
51. Milhomem, S. S. R. *et al*. *Gymnotus capanema*, a new species of electric knife fish (Gymnotiformes, Gymnotidae) from eastern Amazonia, with comments on an unusual karyotype. *J. Fish Biol.* **80**, 802–815 (2012).
52. Machado, M. A. *et al*. Extensive karyotype reorganization in the fish *Gymnotus arapaima* (Gymnotiformes, Gymnotidae) highlighted by Zoo-FISH analysis. *Front. Genet.* **9**, 8 (2018).
53. Wu, C.-I. The genic view of the process of speciation. *J. Evol. Biol.* **14**, 851–865 (2001).
54. Marques, L. A. & Renesto, E. Genetic variability and evidence of founder effect in *Hemiodus orthonops* (Characiformes: Hemiodontidae) from the upper Paraná River basin. *Brazil. Acta Scientiarum. Biological Sciences* **39**, 53–58 (2017).
55. Westley, P. A. H., Ward, E. J. & Fleming, I. A. Fine-scale local adaptation in an invasive freshwater fish has evolved in contemporary time. *Proc. Biol. Sci.* **280**, 20122327 (2013).
56. Tiffin, P. & Ross-Ibarra, J. Advances and limits of using population genetics to understand local adaptation. *Trends Ecol. Evol.* **29**, 673–680 (2014).
57. Feder, J. L., Egan, S. P. & Nosil, P. The genomics of speciation-with-gene-flow. *Trends Genet.* **28**, 342–350 (2012).
58. Chaves, J. A. *et al*. Genomic variation at the tips of the adaptive radiation of Darwin's finches. *Mol. Ecol.* **25**, 5282–5295 (2016).
59. Brown, R. P., Paterson, S. & Risse, J. Genomic signatures of historical allopatry and ecological divergence in an island lizard. *Genome Biol. Evol.* **8**, 3618–3626 (2016).
60. Kang, L. *et al*. Genomic signatures of speciation in sympatric and allopatric Hawaiian picture-winged Drosophila. *Genome Biol. Evol.* **8**, 1482–1488 (2016).
61. Faircloth, B. C., Sorenson, L., Santini, F. & Alfaro, M. E. A Phylogenomic perspective on the radiation of ray-finned fishes based upon targeted sequencing of Ultraconserved Elements (UCEs). *PLoS One* **8**, e65923 (2013).
62. Faircloth, B. C. Illumiprocessor: a trimmomatic wrapper for parallel adapter and quality trimming. doi: 10.6079 (2013).
63. Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120 (2014).
64. Faircloth, B. C. PHYLUCE is a software package for the analysis of conserved genomic loci. *Bioinformatics* **32**, 786–788 (2016).
65. Grabherr, M. G. *et al*. Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat. Biotechnol.* **29**, 644–652 (2011).
66. Marçais, G. & Kingsford, C. A fast, lock-free approach for efficient parallel counting of occurrences of k-mers. *Bioinformatics* **27**, 764–770 (2011).
67. Katoh, K., Misawa, K., Kuma, K.-I. & Miyata, T. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res.* **30**, 3059–3066 (2002).
68. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).
69. Li, H. *et al*. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
70. McKenna, A. *et al*. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* **20**, 1297–1303 (2010).
71. Stamatakis, A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30**, 1312–1313 (2014).
72. Miller, M. A., Pfeiffer, W. & Schwartz, T. Creating the CIPRES Science Gateway for inference of large phylogenetic trees. In *2010 Gateway Computing Environments Workshop (GCE)* 1–8 (IEEE), https://doi.org/10.1109/GCE.2010.5676129
73. Darriba, D., Taboada, G. L., Doallo, R. & Posada, D. jModelTest 2: more models, new heuristics and parallel computing. *Nat. Methods* **9**, 772 (2012).
74. Pritchard, J. K., Wen, X. & Falush, D. Documentation for STRUCTURE software, version 2.3. University of Chicago, Chicago, IL. (2010).
75. Earl, D. A. & vonHoldt, B. M. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conserv. Genet. Resour.* **4**, 359–361 (2012).
76. Kopelman, N. M., Mayzel, J., Jakobsson, M., Rosenberg, N. A. & Mayrose, I. Clumpak: a program for identifying clustering modes and packaging population structure inferences across K. *Mol. Ecol. Resour.* **15**, 1179–1191 (2015).
77. Weir, B. S. & Cockerham, C. C. Estimating F-statistics for the analysis of population structure. *Evolution* (1984).
78. R Core Team R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/ (2013).
79. Jombart, T. & Ahmed, I. Adegenet 1.3-1: ne w tools for the analysis of genome-wide SNP data. *Bioinformatics* **27**, 3070–3071 (2011).
80. Gallant, J. R. *et al*. Genomic basis for the convergent evolution of electric organs. *Science* **344**, 1522–1525 (2014).
81. Liu, Z. *et al*. The channel catfish genome sequence provides insights into the evolution of scale formation in teleosts. *Nat. Commun.* **7**, 11757 (2016).
82. Luu, K., Bazin, E. & Blum, M. G. B. pcadapt: an R package to perform genome scans for selection based on principal component analysis. *Mol. Ecol. Resour.* **17**, 67–77 (2017).
83. Storey, J. D., Bass, A. J., Dabney, A. & Robinson, D. qvalue: Q-value estimation for false discovery rate control. R package version 2.6. 0. *Available at github. com/jdstorey/qvalue. Accessed April* 14, 2017 (2015).
84. Lischer, H. E. L. & Excoffier, L. PGDSpider: an automated data conversion tool for connecting population genetics and genomics programs. *Bioinformatics* **28**, 298–299 (2012).
85. Purcell, S. *et al*. PLINK: A tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* **81**, 559–575 (2007).
86. Kearse, M. *et al*. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* **28**, 1647–1649 (2012).
87. Katoh, K. & Standley, D. M. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* **30**, 772–780 (2013).
88. Elbassiouny, A. A. *et al*. Mitochondrial genomes of the South American electric knifefishes (Order Gymnotiformes). *Mitochondrial DNA Part B* **1**, 401–403 (2016).

89. Saitoh, K., Miya, M., Inoue, J. G., Ishiguro, N. B. & Nishida, M. Mitochondrial genomics of ostariophysan fishes: perspectives on phylogeny and biogeography. *J. Mol. Evol.* **56**, 464–472 (2003).
90. Kumar, S., Stecher, G. & Tamura, K. MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for Bigger Datasets. *Mol. Biol. Evol.* **33**, 1870–1874 (2016).
91. Tamura, K., Nei, M. & Kumar, S. Prospects for inferring very large phylogenies by using the neighbor-joining method. *Proc. Natl. Acad. Sci. USA* **101**, 11030–11035 (2004).
92. Leigh, J. W. & Bryant, D. popart: full-feature software for haplotype network construction. *Methods Ecol. Evol.* **6**, 1110–1116 (2015).
93. Drummond, A. J., Suchard, M. A., Xie, D. & Rambaut, A. Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Mol. Biol. Evol.* **29**, 1969–1973 (2012).
94. Lanfear, R., Calcott, B., Ho, S. Y. W. & Guindon, S. Partitionfinder: combined selection of partitioning schemes and substitution models for phylogenetic analyses. *Mol. Biol. Evol.* **29**, 1695–1701 (2012).
95. Rambaut, A. & Drummond, A. J. Tracer, version 1.5. (2007).
96. Rambaut, A. & Drummond, A. FigTree: Tree figure drawing tool, v1. 4.2. *Institute of Evolutionary Biology, University of Edinburgh* (2012).
97. Albert, J. S. & Fink, W. L. Phylogenetic relationships of fossil neotropical electric fishes (Osteichthyes: Gymnotiformes) from the upper Miocene of Bolivia. *J. Vert. Paleontol.* **27**, 17–25 (2007).

Acknowledgements

Author contributions

C.A. and L.F.D. conceived the study. L.F.D., M.J.M., J.R.L. and C.A. designed the research; L.F.D., R.K and R.G. conducted fieldwork; C.A. conducted the laboratory experiments and analyzed the data; C.A. and L.F.D. wrote the first draft of the manuscript. All authors read, revised and approved the final manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41598-019-55336-y.

Correspondence and requests for materials should be addressed to L.F.D.L.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.