Supporting information

Landscape genetics and species delimitation in the Andean Palm Rocket Frog (Aromobatidae, *Rheobates*).

Gabrielle Genty[1*], Carlos E. Guarnizo[1], Juan P. Ramírez[1,2], Lucas S. Barrientos[1,3], Andrew J. Crawford[1]

[1] Department of Biological Sciences, Universidad de los Andes, Bogotá, 111711, Colombia.

[2] Department of Biology, San Diego State University, San Diego, CA, 92182, USA.

[3] Universidad del Rosario, Departamento de Biología, Bogotá, 111221, Colombia.

The Supplementary Materials contain sample data, primers used for DNA sequence amplification, evolutionary genetic analyses performed, and substitution models selected for each gene. We provide figures for the maximum likelihood (ML) phylogenetic inference trees using RAxML applied to either the concatenated mitochondrial genes or nuclear genes only, and species tree analysis applied to all genes using StarBEAST2.

Supporting material includes Table S1, Table S2, and Table S3; Figure S1, Figure S2, Figure S3 and Figure S4, plus details of the species tree and BPP analyses.

Table S1 Species, field, and institutional numbers, locality, geographical coordinates, GenBank accession numbers for ingroups and outgroups samples, and the new data amplified for this study (represented by the GenBank accession numbers starting with 'MW' and in bold). Abbreviations for field series: AAV, Álvaro Andrés Velásquez; AJC, Andrew J. Crawford; CG, Carlos E. Guarnizo; LSB, Lucas Santiago Barrientos; MAR, Marco Antonio Rada. Acronyms for museums are: ANDES-A and ANDES-T, Amphibians Collection and Tissues Collection, Museo de Historia Natural C.J. Marinkelle, Universidad de los Andes, Bogotá, Colombia; MHUA, Museo de Herpetología de la Universidad de Antioquia, Colombia. Explanation of letter codes: eEC (eastern flank Eastern Cordillera), CC (Central Cordillera), wEC (western flank Eastern Cordillera).

Species	Locality (m a.s.l.)	Geographic region	Clade	Latitude	Longitude	Field number	Institutional number	*COI*	*16S*	*POMC*	*SF232*	*SF328*	*SF412*
Rheobates palmatus	Santa María (870)	eEC	SM	4.848	-73.272	AJC 4235	ANDES-A 1486	KJ130682	MW596297	KJ130740	**MW596289**	**MW596300**	**MW596247**
	Santa María (870)	eEC	SM	4.848	-73.272	AJC 4239	ANDES-A 1487	KJ130683	KJ130718	KJ130742	**MW596271**	**MW596301**	**MW596250**
	Santa María (870)	eEC	SM	4.858	-73.264	AJC 4232	ANDES-A 1485	-	**MW596298**	-	**MW596272**	**MW596322**	**MW596248**
	Santa María (870)	eEC	SM	4.858	-73.264	AJC 4233	-	**MW596296**	**MW596299**	**MW596292**	**MW596290**	**MW596302**	**MW596249**
Rheobates pseudopalmatus	Amalfi (1550)	CC	CC	6.800	-75.086	MHUA 4732	MHUA 4732	KJ130692	KJ130726	KJ130746	**MW596288**	**MW596303**	**MW596251**
	San Rafael (1250)	CC	CC	6.390	-75.011	LSB 337	Voucher lost	KJ130679	KJ130715	KJ130738	-	-	**MW596252**
	Maceo (575)	CC	CC	6.547	-74.644	MHUA 4348	MHUA 4348	KJ130678	KJ130713	KJ130737	**MW596276**	**MW596318**	**MW596253**
	Anorí (1530)	CC	CC	6.978	-75.111	MHUA 5162	MHUA 5162	KJ130694	KJ130727	KJ130747	**MW596277**	**MW596319**	**MW596254**
	Anorí (1530)	CC	CC	6.978	-75.111	MHUA 5357	MHUA 5357	KJ130695	KJ130728	**MW596293**	**MW596287**	**MW596320**	**MW596255**
Rheobates palmatus	San Vicente (300)	wEC	EC	7.079	-73.552	AJC 3526	ANDES-A 1481	KJ130670	KJ130706	KJ130733	**MW596278**	**MW596309**	**MW596258**
	Virolín (1748)	wEC	EC	6.105	-73.199	CG 001	ANDES-T 2350	KJ130681	KJ130716	KJ130739	**MW596283**	**MW596317**	**MW596259**
	Puente Nacional (1623)	wEC	EC	5.882	-73.678	AJC 3398	ANDES-A 1476	KJ130665	KJ130701	KJ130729	**MW596284**	**MW596314**	**MW596260**
	Puente Nacional (1623)	wEC	EC	5.882	-73.678	AJC 3403	ANDES-A 1478	KJ130690	KJ130724	**MW596294**	**MW596285**	**MW596315**	**MW596261**
	Puente Nacional (1623)	wEC	EC	5.882	-73.678	AJC 3404	ANDES-A 1479	KJ130676	KJ130711	**MW596295**	**MW596286**	**MW596316**	**MW596262**
	Piedecuesta (940)	wEC	EN	6.783	-73.017	AAV153	ANDES-A 1472	KJ130685	KJ130720	KJ130743	**MW596279**	**MW596304**	**MW596263**
	Piedecuesta (940)	wEC	EN	6.783	-73.017	AAV 154	ANDES-A 1473	KJ130688	KJ130722	KJ130744	**MW596281**	**MW596306**	**MW596265**

	Piedecuesta (940)	wEC	EN	6.783	-73.017	AJC 3860	ANDES-A 1482	KJ130 666	KJ130 702	KJ130 730	-	**MW5 96305**	**MW5 96266**
	Suratá (1740)	wEC	EN	7.367	-72.983	CG 003	ANDES-T 2352	KJ130 675	KJ130 709	KJ130 735	**MW5 96280**	**MW5 96307**	**MW5 96264**
	Suratá (1740)	wEC	EN	7.367	-72.983	CG 004	ANDES-T 2353	KJ130 672	KJ130 708	-	**MW5 96282**	**MW5 96308**	**MW5 96267**
	San Francisco (1749)	wEC	ES	4.996	-74.270	AAV 167	AAV 167	KJ130 669	-	KJ130 732	-	**MW5 96321**	**MW5 96270**
	Icononzo (1285)	wEC	ES	4.185	-74.547	MAR217 5	MAR2175	-	KJ130 717	KJ130 741	**MW5 96273**	**MW5 96310**	**MW5 96268**
	Cáqueza (1515)	eEC	ES	4.414	-73.948	CG 007	ANDES-T 2354	KJ130 684	KJ130 719	KJ871 617	**MW5 96274**	**MW5 96311**	**MW5 96256**
	Cáqueza (1515)	eEC	ES	4.414	-73.948	CG 013	ANDES-T 2355	KJ130 677	KJ130 712	KJ130 736	**MW5 96291**	**MW5 96312**	**MW5 96269**
	Las Brisas (2000)	eEC	ES	4.437	-73.919	AJC 2106	ANDES-A 1474	KJ130 668	KJ130 704	KJ130 731	**MW5 96275**	**MW5 96313**	**MW5 96257**
Allobates aff. *juanii*	Sabanalarga (320)	-	Out group	4.773	-73.038	AJC 3383	ANDES-A 1073	KJ130 661	KJ130 697	KJ871 619	-	-	-
Allobates femoralis	El Once (90)	-	Out group	4.117	-69.950	AJC 3598	ANDES-A 1471	KJ130 660	KJ130 696	KJ871 618	-	-	-

Table S2 Conditions used for PCR amplification. For the mitochondrial and POMC markers, the same conditions used by Muñoz-Ortíz *et al*. (2015) were implemented here (Appendix S1; Table S2). To amplify the other three nuclear genes, we used an initial denaturing step of 2 min at 95 ºC followed by 30 amplification cycles (30 s at 95 ºC, 30 s at 51 ºC, 60 s at 73 ºC) and a final extension of 20 min at 72 ºC. PCR products were cleaned with exonuclease I and SAP enzymes (Werle *et al.* 1994), and sequenced directly with Sanger technology.

Gene region	Primer	Primer sequence (5'-3')	Source
Mitochondrial *COI* (658 bp)	dgLCO-1490	GGT CAA CAA ATC ATA AAG AYA TYG G	(Meyer & Paulay, 2005)
	dhHCO-2198	TAA ACT TCA GGG TGA CCA AAR AAY CA	
	Chmf4	TYT CWA CWA AYC AYA AAG AYA TCG G	(Che et al., 2012)
	Chmr4	ACY TCR GGR TGR CCR AAR AAT CA	
Mitochondrial *16S* (569 bp)	Sar-L	CGC CTG TTT ATC AAA AAC AT	(Kessing et al., 2004)
	Sbr-H	CCG GTC TGA ACT CAG ATC ACG T	
Nuclear *POMC* (472 bp)	POMC_DRV_R1	GGR RTT YTT GAA WAG AGT CAT TAG WGG	(Vieites et al., 2007)
	POMC_DRV_F1	ATA TGT CAT GAS CCA YTT YCG CTG GAA	
Nuclear *SF232* (375 bp)	SF232 F	AGT CAT AAT GGT GCC ACT AAA AG	(Tezuka et al., 2012)
	SF232 R	TGT GGT CCT TGT ATG GGT TG	
Nuclear *SF328* (404 bp)	SF328 F	CCC AAA AGA AGT TTT GCT GA	(Tezuka et al., 2012)
	SF328 R	GCC TCA CAA ACA ACC ACA GA	
Nuclear *SF412* (352 bp)	SF412 F	ACC ATC CTC ACT GTG ACA CC	(Tezuka et al., 2012)
	SF412 R	TCT TTG GCA AAC TGG ACC TT	

Table S3 Best-fit partitioning scheme and models of nucleotide substitution for each of the partitions of the genes included in this study, as inferred by the corrected Akaike information criterion (AICc) implemented in PartitionFinder 2. The abbreviation 'pos.' refers to codon position.

Partition number	**Partition**	**Best model**
1	16S	GTR + Γ
2	COI pos. 3	TRN + Γ
3	COI pos. 1	TRNEF +Γ
4	COI pos. 2	F81 + Γ
5	POMC pos. 2	TRN + I
6	POMC pos. 3	SYM
7	POMC pos. 1	HKY
8	SF412 pos. 3	HKY + Γ
9	SF412 pos. 1, SF412 pos. 2	F81 + I
10	SF328	K81 + I + Γ
11	SF232 pos. 3	TVMEF + I
12	SF232 pos. 1	F81
13	SF232 pos. 2	F81 + I

Figure S1. ML concatenated tree inferred for the mitochondrial genes, 16S and COI, of the populations of *Rheobates* spp., as inferred using the software RAxML. The phylogenetic tree was constructed using the standard GTRGAMMA substitution model for each of the partitions recommended by PartitionFinder2. Tree searching involved 1000 replicate searches to find the optimal ML tree, followed by a bootstrap analysis using the autoMRE option, which automatically determines a sufficient number of bootstrap replicates. Numbers next to each node indicate bootstrap support values. The scale bar represents the inferred number of substitutions per site or branch length. Specimens are indicated by their field or museum voucher number, as summarized in supporting Table S1.

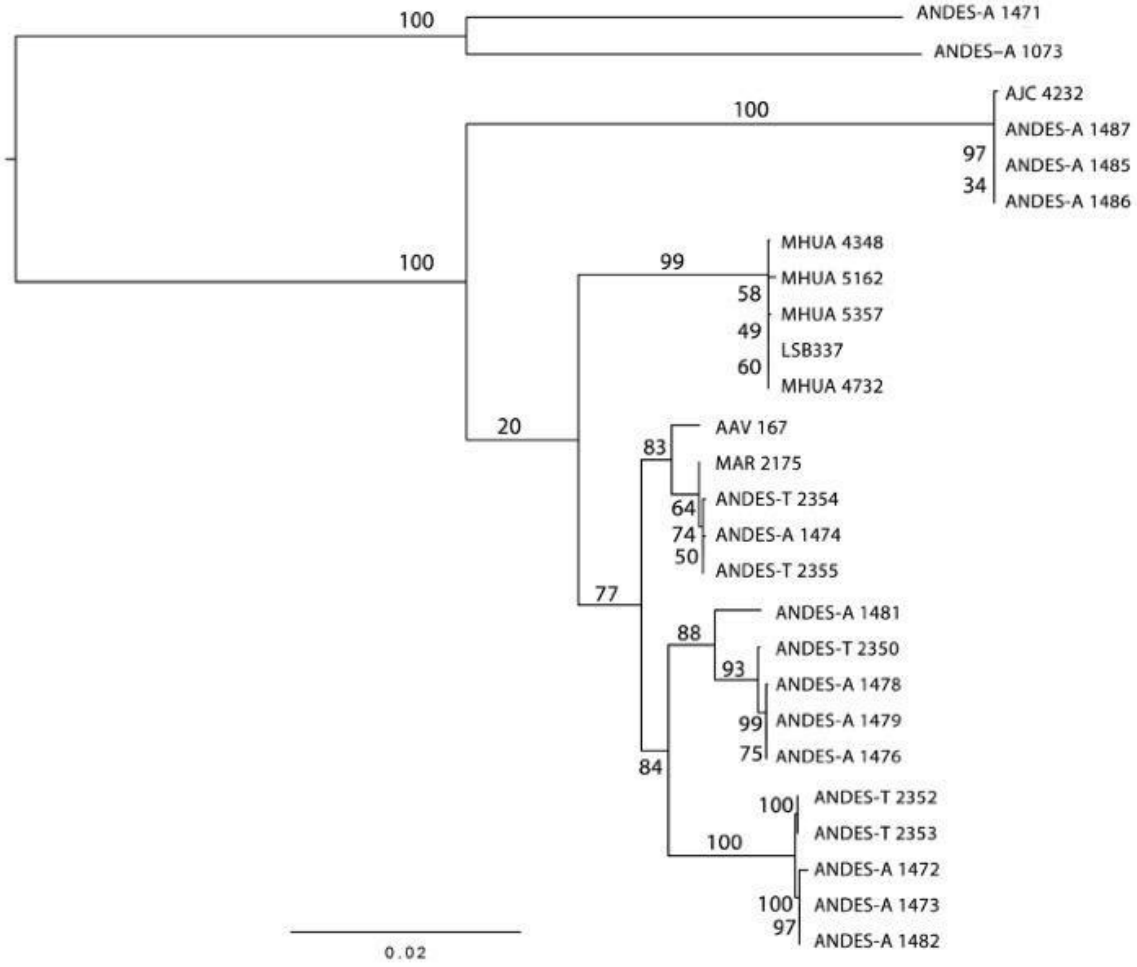


Figure S2. ML concatenated gene tree inferred for the four nuclear genomic markers, POMC, SF232, SF328, and SF412 from populations of *Rheobates* spp., as inferred using the software RAxML. The phylogenetic tree was constructed using the standard GTRGAMMA substitution model for each of the partitions recovered by PartitionFinder2. Tree searches involved 1000 replicate searches to find the optimal ML tree, followed by a bootstrap analysis using the autoMRE option, which automatically determines a sufficient number of bootstrap replicates. Numbers above each branch indicate bootstrap support values for the descending node. The scale bar represents the average number of substitutions per site or branch length. Specimens are indicated by their field or museum voucher number, as summarized in supporting Table S1.

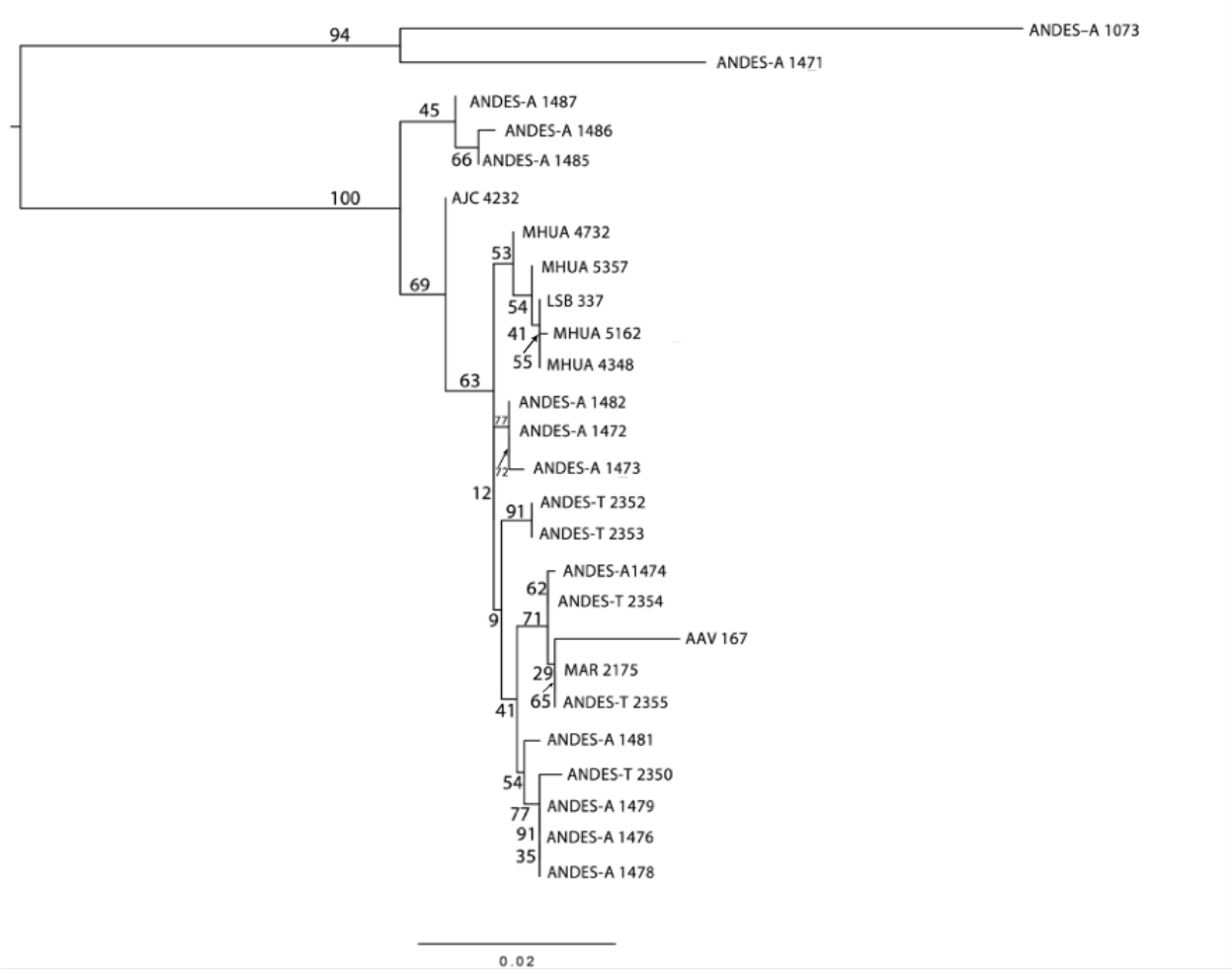


Figure S3. StarBEAST2 multispecies coalescent tree of the five major clades of the frog genus *Rheobates* inferred in this study, as obtained from the sequences of two mitochondrial genes (16S and COI) and four nuclear genes (POMC, SF232, SF328, and SF412). See main text for details of the analysis. Numbers to the right of each node correspond to the Bayesian posterior probabilities of the subtended clade. The scale bar represents the average number of substitutions per site or branch length. Specimens are indicated by their field or museum voucher number, as summarized in supporting Table S1.

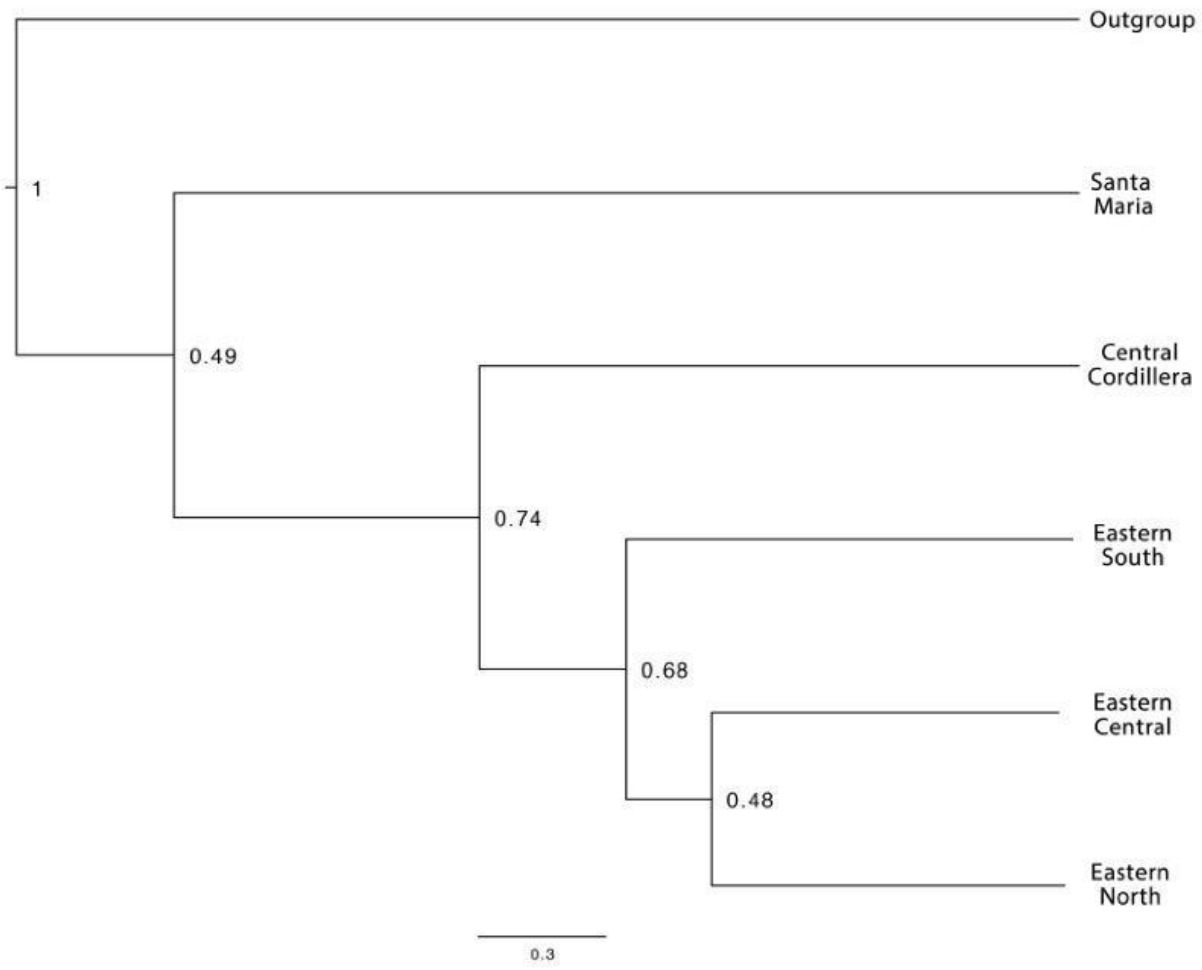
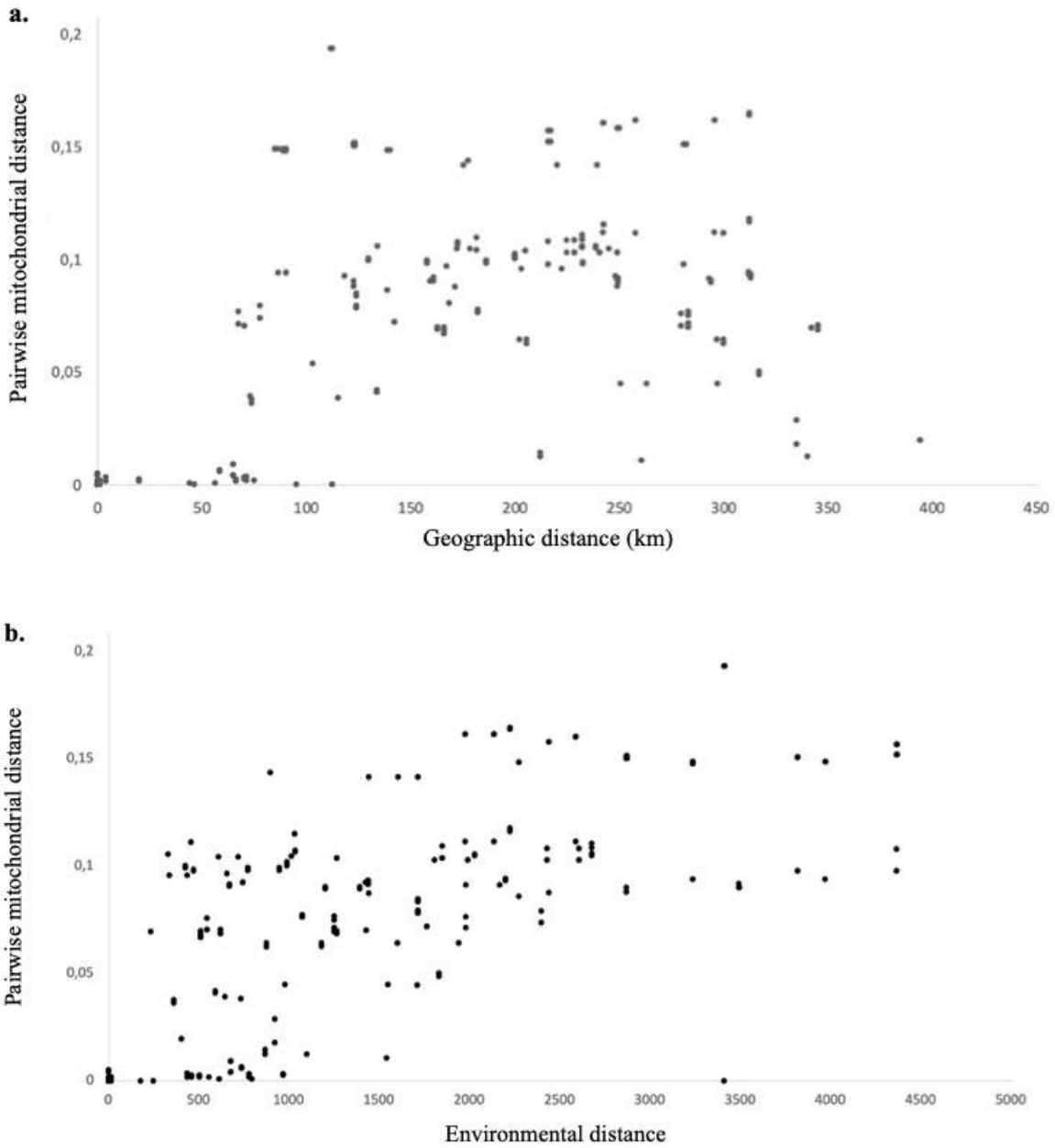


Figure S4. a. Relationship between geographic distance in kilometers (km) and mitochondrial DNA distance (concatenated COI and 16S genes). **b.** Relationship between environmental distance based on WorldClim layers and mitochondrial distances (COI+16S).



Species tree inference

We estimated the posterior distribution of trees, including dates and rates, by assuming a birth-death model tree prior (a diversification model that considers both speciation and extinction; Gernhard, 2018) and a random local clock model of evolution (Drummond & Suchard, 2010). Our prior distribution on the substitution rate of the mitochondrial genes assumed a lognormal distribution with a mean of 0.0155 substitutions per lineage per million years (Gehara, Summers & Brown, 2013) with upper and lower bounds of the central 95% prior density of 0.0131 and 0.0182, respectively. We set StarBEAST2 to estimate the substitution rate of each nuclear gene under a 1/X distribution with an offset of 0.

Species delimitation analysis (BPP)

To test for the species delimitation we used Bayesian Phylogenetics and Phylogeography (BPP) version 3.2 (Yang, 2015). We assumed three contrasting sets of prior distributions on demographic parameters representing three types of histories: small population size with shallow divergence among species $\theta \sim \Gamma(2, 2000)$, $\tau 0 \sim \Gamma(2, 2000)$; large population size with deep divergence $\theta \sim \Gamma(1, 10)$, $\tau 0 \sim \Gamma(1, 10)$; and large population size with shallow divergences $\theta \sim \Gamma(1, 10)$, $\tau 0 \sim \Gamma(2, 2000)$, where θ is the population mutation rate (proportional to population size), τ is the population divergence time, and Γ indicates a gamma distribution with (shape, scale) parameters (Flouri, Jiao, Rannala & Yang, 2018; Yang, 2015). For other divergence time parameters, we assigned a Dirichlet prior (Yang & Rannala, 2010: equation 2). We used the 'A10' analysis (species delimitation = 1, species tree = 0) which uses the reversible-jump MCMC algorithm and used the uniform rooted tree (species model prior = 1), and left the automatic fine-tuning of the MCMC. Considering that estimating different mutation rates would

not affect the Bayesian comparison and would instead be more of a computational burden, we decided to assume the same mutation rate for all loci (locus rate = 0). Given that the reversible jump MCMC algorithm can exhibit mixing problems in some datasets (Yang, 2015; Rannala & Yang, 2013; Yang & Rannala, 2010), we repeated the analysis under each scenario three times, with different starting seeds, to confirm consistency among runs.

References

Drummond, A. J., & Suchard, M. A. (2010). Bayesian random local clocks, or one rate to rule them all. *BMC Biology*, *8*(1), 1-12. DOI: 10.1186/1741-7007-8-114

Flouri, T., Jiao, X., Rannala, B., & Yang, Z. (2018). Species tree Inference with BPP using genomic sequences and the multispecies coalescent. *Molecular Biology and Evolution*, *35*(10), 2585–2593. DOI: 10.1093/molbev/msy147

Gehara, M., Summers, K., & Brown, J. L. (2013). Population expansion, isolation and selection: novel insights on the evolution of color diversity in the strawberry poison frog. *Evolutionary Ecology*, *27*(4), 797-824. DOI: 10.1007/s10682-013-9652-0

Gernhard, T. (2008). The conditioned reconstructed process. *Journal of Theoretical Biology*, *253*(4), 769–778. DOI: 10.1016/j.jtbi.2008.04.005

Rannala, B., & Yang, Z. (2013). Improved reversible jump algorithms for Bayesian species delimitation. *Genetics*, *194*(1), 245–253. DOI: 10.1534/genetics.112.149039

Yang, Z. (2015). The BPP program for species tree estimation and species delimitation. *Current Zoology*, *61*(5), 854–865. DOI: 10.1093/czoolo/61.5.854

Yang, Z., & Rannala, B. (2010). Bayesian species delimitation using multilocus sequence data. *Proceedings of the National Academy of Sciences*, *107*(20), 9264–9269. DOI: 10.1073/pnas.0913022107