

SUPPORTING INFORMATION

Functional connectivity in sympatric spiny rats reflects different dimensions of Amazonian forest-association

Jeronymo Dalapicolla, Joyce Rodrigues do Prado, Alexandre Reis Percequillo, and L. Lacey Knowles

APPENDIX S1

Table S1.1. Information about the locality, species, institution of origin of the samples, as well as the number of retained reads after demultiplexing. Ambrose Monell Cryo Collection, American Museum of Natural History, New York, USA, **AMNH-AMCC**; Field Museum of Natural History, Chicago, Illinois, USA, **FMNH**; Laboratório de Mamíferos da Escola Superior de Agricultura “Luiz de Queiroz, Universidade de São Paulo, Piracicaba, São Paulo, Brazil, **LMUSP**; Museum of Southwestern Biology, Albuquerque, New Mexico, USA, **MSB**; Museum of Vertebrate Zoology, Berkeley, California, USA, **MVZ**; Museu de Zoologia da Universidade de São Paulo, São Paulo, Brazil, **MZUSP**; National Museum of Natural History, Washington, D.C., USA, **USNM**; Texas Tech University, Lubbock, Texas, USA, **TTU**.

Table S1.2. Predictor variables used for creating resistance matrices for MLPE analyses (for details see methods section and Appendix 2.2 for full references). Note that because of resolution differences, the Tropical and Subtropical wetlands distribution map was resampled at 1Km² (231m²) to create variables with the same resolution. Methods column shows the R package used in analysis followed by :: and the function name in italics.

Resistance Matrix	Variables	Methods	Source
<i>Geographic Distance</i>	Longitude and Latitude from samples	<i>geosphere::distm</i>	Goslee & Urban 2007
<i>Habitat Distance</i>	Tropical and Subtropical wetlands distribution map	<i>gdistance::commuteDistance</i>	Gumbrecht <i>et al.</i> 2017 van Etten 2018
<i>Productivity Distance</i>	12 monthly temperature variables from WorldClim 12 monthly precipitation variables from WorldClim 6 Potential Evapotranspiration (PET) variables from ENVIREM	Species Distribution Models (SDM) <i>biomod2::BIOMOD_Modeling</i>	Fick & Hijmans 2017 https://www.worldclim.org/ Title & Bemmels 2018 https://envirem.github.io/ Thuiller <i>et al.</i> , 2020
<i>River Distance</i>	river shapefile “ <i>ne_10m_rivers_lake_centerlines</i> ” from Natural Earth Data	<i>riverdist::riverdistancemat</i>	Tyers 2020 https://www.naturalearthdata.com
<i>Topographic distance</i>	SRTM elevation data from WorldClim 2.1	<i>topoDistance::topoDist</i>	Wang 2020 Fick & Hijmans 2017 https://www.worldclim.org/

Table S1.3: Validated co-occurrence points for the sympatric species *P. simonsi* and *P. steerei* from Western Amazon used in SDM analyses for creation of the habitat productivity distance for IBR analyses.

	<i>P. simonsi</i>		<i>P. steerei</i>	
	Longitude	Latitude	Longitude	Latitude
1	-77.71047	-1.13044	-71.62369	-4.53573
2	-75.33333	0.46667	-70.73359	-6.80001
3	-77.751	-4.022	-70.85003	-6.75
4	-73.26836	-4.02398	-68.76667	-6.46669
5	-77.817	-0.983	-68.91651	-6.53356
6	-76.81724	-0.43999	-66.58528	-5.6364
7	-64.71989	-3.3539	-66.2333	-3.2833
8	-71.39377	-12.82978	-66	-3.31667
9	-71.26166	-12.68001	-65.31321	-10.84689
10	-68.57543	-12.39863	-64.84993	-15.37092
11	-68.91681	-11.3501	-65.15511	-12.43147
12	-69.23741	-12.59659	-65.06785	-11.906
13	-69.26065	-8.84739	-65.63494	-14.98761
14	-72.7633	-8.83389	-69.26065	-8.84739
15	-75.21626	-9.86708	-72.78304	-8.66663
16	-68.88173	-12.95664	-72.81662	-8.36666
17	-73.16208	-5.2495	-72.85094	-8.60044
18	-62.26745	-4.42802	-67.2833	-9.8
19	-72.94991	-11.58344	-60.11667	-3.15
20	-69.07289	-12.6	-60.55351	-3.2883
21	-71.21667	-10.13333	-60.52022	-3.08039
22	-68.76672	-6.46666	-65.70879	-2.21689
23	-68.89219	-6.58282	-64.67127	-2.53438
24	-72.81662	-8.36666	-66.21667	-11.01667

25	-66.01666	-3.31666	-66.31668	-10.98328
26	-69.20697	-12.72047	-67.2	-11.38333
27	-73.3258	-11.7667	-68.57543	-12.39863
28	-68.7667	-13.5833	-67.45	-10.7
29	-69.6122	-13.1472	-74.56879	-8.39313
30	-66.7333	-10.7667	-75.21626	-9.86708
31	-75.08802	-8.39409	-73.04723	-3.50143
32	-64.82581	-9.4455	-72.13332	-3.43328
33	-65.43986	-9.63406	-73.81561	-10.68165
34	-72.78304	-8.66663	-75.09959	-6.73333
35	-70.85003	-6.75	-67.56023	-11.49004
36	-73.66744	-4.90625	-68.84981	-11.35059
37	-76.1094	-5.89482	-66.77966	-11.74947
38	-64.6795	-9.27685	-66.09131	-10.95254
39	-68.74681	-10.99838	-73.16208	-5.2495
40	-68.67144	-9.07918	-69.07289	-12.6
41	-65.70879	-2.21689	-71.0451	-12.15612
42			-71.11199	-12.09053
43			-68.7667	-13.5833
44			-75.08802	-8.39409
45			-76.1094	-5.89482
46			-73.2282	-3.91344
47			-65.77163	-7.66384

Table S1.4. Summary statistics for each of the univariate MLPE regression models for *P. simonsi* and for *P. steerei* used to identify which level of habitat resistance (i.e., VH, H, M, and L) would be included in the full model. All models measured the degree of association between relatedness coefficients of individuals and the resistance distance, with 5 degrees of freedom. The best models for both species were L (0.6 for resistance and 0.4 for conductance).

	Intercept	Coefficients	logL	AIC	ΔAIC	Weighted AIC
<i>P. simonsi</i>						
Habitat L	-0.087	-0.098	67.536	-125.072	0.000	0.997
Habitat M	-0.087	-0.124	61.564	-113.127	11.945	0.003
Habitat H	-0.087	-0.160	54.880	-99.759	25.313	0.000
Habitat VH	-0.087	-0.049	48.206	-86.411	38.661	0.000
<i>P. steerei</i>						
Habitat L	-0.130	-0.070	119.758	-229.516	0.000	0.997
Habitat M	-0.134	-0.067	113.951	-217.902	11.614	0.003
Habitat H	-0.138	-0.068	107.454	-204.908	24.607	0.000
Habitat VH	-0.141	-0.067	99.795	-189.590	39.925	0.000

Table S1.5. Genomic datasets used in analyses of two sympatric species of *Proechimys* from Western Amazon. The number of individuals retained after processing reads, and the number of SNPs retained after filtering of the trimmed 132 bp sequenced loci, as well as the average percent missing data by individual and maximum by locus, and the average coverage depth considering retained individuals are shown (see Appendix S2.1 for details of filtering and SNPs calling).

	<i>P. simonsi</i>	<i>P. steerei</i>
Initial individuals	21	20
Retained individuals	17	19
Initial SNPs	47,009	99,740
LD Filter	24,016	38,090
Quality Filter	13,310	15,252
Neutral Filter	12,784	13,971
Missing data by locus	17.65%	15.79%
Missing data by individual	16.81%	18.68%
Depth coverage	31.98X	34.33X

Table S1.6. Tracy-Widom tests of eigenvalues for *P. simonsi* and *P. steerei*; significant PCs are in bold.

PC	<i>P. simonsi</i>			<i>P. steerei</i>		
	Eigenvalues	TW stats	<i>p</i> -values	Eigenvalues	TW stats	<i>p</i> -values
1	28740	2.462	0.005	71190	2.202	0.007
2	17080	-2.024	0.732	37740	4.021	0.000
3	15800	-2.095	0.752	16700	2.812	0.002
4	13240	-3.736	0.985	11250	-1.291	0.507
5	13120	-3.124	0.944	9964	-2.422	0.832
6	11550	-4.013	0.993	8526	-4.53	0.999
7	11360	-3.55	0.977	7670	-5.763	1.000
8	10260	-3.954	0.991	7425	-5.728	1.000
9	10100	-3.481	0.973	7352	-5.322	1.000
10	9964	-2.913	0.919	7332	-4.797	0.999
11	9825	-2.252	0.792	7006	-4.721	0.999
12	9649	-1.428	0.551	6981	-4.161	0.995
13	7002	-2.168	0.771	6919	-3.611	0.980
14	6308	-1.791	0.664	6813	-3.054	0.937
15	5534	-1.281	0.504	6438	-2.669	0.881
16	4467	-0.5552	0.291	5992	-2.227	0.786
17				5690	-1.547	0.589
18				5468	-0.5552	0.291

Table S1.7. Significant F_{ST} -values (** p -value < 0.001) between genetic clusters (GC) in *Proechimys steerei* (the seasonal floodplain forests/*várzea* species) in lower diagonal, and 95% confidence intervals in upper diagonal. Because a single genetic cluster was estimated in *P. simonsi*, F_{ST} 's were not estimated.

	GC 1	GC 2	GC 3
GC 1	–	0.535 – 0.547	0.557 – 0.568
GC 2	0.542**	–	0.266 – 0.280
GC 3	0.563**	0.273**	–

Table S1.8. Genetic diversity indices for the two sympatric *Proechimys* species from the Western Amazon, which differ in their habitat associations (i.e., *P. simonsi* occurs in the non-flooded forests, the *terra-firme*, whereas *P. steerei* occurs in from seasonal floodplain forests, the *várzea*). Estimates of diversity are also presented for each of the three genetic clusters (GC) in *P. steerei*; *P. simonsi* is represented by a single genetic cluster. The 95% confidence intervals are in parentheses.

Species	H_{EXP}	π	F_{IS}
<i>P. simonsi</i>	0.346 (0.345 – 0.347)	0.347 (0.345 – 0.349)	0.310 (0.228 – 0.392)
<i>P. steerei</i>	0.310 (0.309 – 0.310)	0.311 (0.309 – 0.313)	0.498 (0.440 – 0.556)
<i>P. steerei</i> GC 1	0.452 (0.451 – 0.453)	0.128 (0.124 – 0.131)	0.190 (0.133 – 0.248)
<i>P. steerei</i> GC 2	0.390 (0.389 – 0.392)	0.217 (0.214 – 0.221)	0.280 (0.163 – 0.398)
<i>P. steerei</i> GC 3	0.371 (0.370 – 0.373)	0.218 (0.214 – 0.221)	0.139 (0.09 – 0.187)

H_{EXP} = Expected Heterozygosity;

π = Nucleotide Diversity;

F_{IS} = Inbreeding Coefficient.

Table S1.9. Tukey's tests of significant mean differences between species and genetic clusters (GC) using individual-based genetic diversity indices (expected heterozygosity, H_{EXP} , inbreeding coefficient; F_{IS}) for the sympatric species *Proechimys simonsi*, SIM (from the non-flooded forests/*terra-firme*), and *P. steerei*, STE (from the seasonal floodplain forests/*várzea*).

	Difference	Lower	Upper	Adjusted <i>p</i> -value
<i>H_{EXP}</i>				
GC 1 – SIM	0.106	0.104	0.108	< 0.001
GC 2 – SIM	0.044	0.042	0.046	< 0.001
GC 3 – SIM	0.025	0.023	0.027	< 0.001
STE – SIM	-0.036	-0.038	-0.035	< 0.001
GC 2 – GC 1	-0.062	-0.064	-0.059	< 0.001
GC 3 – GC 1	-0.081	-0.083	-0.078	< 0.001
STE – GC 1	-0.142	-0.144	-0.140	< 0.001
GC 3 – GC 2	-0.019	-0.021	-0.017	< 0.001
STE – GC 2	-0.081	-0.083	-0.079	< 0.001
STE – GC 3	-0.062	-0.063	-0.060	< 0.001
<i>F_{IS}</i>				
GC 1 – SIM	-0.120	-0.298	0.058	0.326
GC 2 – SIM	-0.030	-0.196	0.136	0.986
GC 3 – SIM	-0.171	-0.321	-0.021	0.018
STE – SIM	0.187	0.071	0.304	< 0.001
GC 2 – GC 1	0.090	-0.122	0.301	0.751
GC 3 – GC 1	-0.051	-0.250	0.148	0.949
STE – GC 1	0.307	0.132	0.483	< 0.001
GC 3 – GC 2	-0.141	-0.330	0.048	0.230
STE – GC 2	0.217	0.054	0.381	0.004
STE – GC 3	0.359	0.211	0.506	< 0.001

Table S1.10. Tests of association between genetic and geographic distance for *P. simonsi* (non-flooded forests, *terra-firme*) and *P. steerei* (seasonal floodplain forests, *várzea*). We performed Mantel tests using two different distances matrices at individual level: Euclidean distance and distance considering the river channel (Along Rivers). Tests for *P. steerei* were conducted separately as well, eliminating each one of the three genetic cluster (GC) by turn. Pearson’s correlation coefficient, r , measures the strength of the association in the Mantel test, and t_0 measures the strength of the association in the Procrustes analysis; significant values are shown in bold.

	Mantel Test (Euclidean)		Mantel Test (Along Rivers)		Procrustes Analysis	
	r	p -value	r	p -value	t_0	p -value
Species						
<i>P. simonsi</i>	0.619	< 0.001	0.091	0.161	0.799	< 0.001
<i>P. steerei</i>	0.572	< 0.001	0.512	< 0.001	0.765	< 0.001
Dropped cluster						
<i>P. steerei</i> GC 1	0.629	< 0.001	0.808	< 0.001	0.791	< 0.001
<i>P. steerei</i> GC 2	0.681	< 0.001	0.793	< 0.001	0.834	< 0.001
<i>P. steerei</i> GC 3	0.519	< 0.001	0.149	0.081	0.845	< 0.001

Table S1.11. Results of sequential drop-out tests to evaluate the robustness of Procrustes results ($t_0 = 0.764$, $p < 0.001$) to the inclusion of particular genetic clusters of *P. steerei* (the seasonal floodplain forests/*várzea*) species, as well as other statistics from the Procrustes analysis (specifically, the strength of association between geographic and genetic maps without the genetic cluster, t'' , and its p -value (in parentheses), and the impact of dropping the cluster relative to t_0 , $t_0 - t''$; the angle of rotation between the genetic and geographic maps without the cluster $\theta_{t''}$; the strength of the association between the original genetic PCA and the genetic PCA without the genetic cluster, t' ; and the angle of rotation of the genetic matrices to minimize the sum-squared differences $\theta_{t'}$).

Dropped cluster	t''	$t_0 - t''$	$\theta_{t''}$	t'	$\theta_{t'}$
GC 1	0.791 ($p < 0.001$)	0.027	87.121	0.857	80.716
GC 2	0.834 ($p < 0.001$)	0.069	22.929	0.935	11.938
GC 3	0.845 ($p < 0.001$)	0.081	37.730	0.942	-25.923

Table S1.12. Summary statistics for 16 uncorrelated models using MLPE regression and five predictor variables in *P. simonsi*, the non-flooded forests/*terra-firme* species; 16 models were retained (among the 32 models generated consider all possible combinations between the five predictor variables) since these were the only models for which the correlation between the combined predictor variables, r^2 , was < 0.6. The best models ($\Delta AIC < 2$) are in bold. Models identification and coefficients for all variables are shown, along with the degrees of freedom, df, log-likelihood, logL, Akaike's Information Criterion, AIC, the difference in AIC compared with best model, ΔAIC , and the model importance considering all uncorrelated models (that is, the weighted AIC).

Model	Intercept	Productivity	Habitat	River	Euclidean	Topography	df	logL	AIC	ΔAIC	Weighted AIC
1	-0.087	-0.020	-0.091	.	.	.	6	68.578	-125.157	0.000	0.333
2	-0.087	.	-0.098	.	.	.	5	67.536	-125.072	0.084	0.319
3	-0.087	.	-0.095	-0.008	.	.	6	67.676	-123.352	1.805	0.135
4	-0.087	-0.021	-0.092	0.003	.	.	7	68.601	-123.202	1.955	0.125
5	-0.087	-0.027	.	.	-0.083	.	6	65.585	-119.169	5.987	0.017
6	-0.087	-0.027	.	.	.	-0.083	6	65.450	-118.901	6.256	0.015
7	-0.087	.	.	-0.026	-0.083	.	6	65.357	-118.715	6.442	0.013
8	-0.087	.	.	-0.025	.	-0.083	6	65.257	-118.514	6.643	0.012
9	-0.087	-0.019	.	-0.016	-0.081	.	7	66.062	-118.124	7.032	0.010
10	-0.087	-0.018	.	-0.015	.	-0.081	7	65.922	-117.844	7.312	0.009
11	-0.087	.	.	.	-0.091	.	5	63.585	-117.170	7.987	0.006
12	-0.087	-0.091	5	63.544	-117.088	8.068	0.006
13	-0.087	-0.034	.	-0.031	.	.	6	50.433	-88.866	36.291	0.000
14	-0.087	-0.052	5	48.890	-87.781	37.376	0.000
15	-0.087	.	.	-0.051	.	.	5	48.564	-87.129	38.028	0.000
16	-0.087	4	42.777	-77.553	47.603	0.000

Table S1.13. Summary statistics for 7 uncorrelated models using MLPE regression and five predictor variables for *P. steerei*, the seasonal floodplain forests/*várzea* species; 7 models were retained (among the 32 models generated consider all possible combinations between the five predictor variables) since these were the only models for which the correlation between the combined predictor variables, r^2 , was < 0.6 . The best models ($\Delta AIC < 2$) are in bold. Models identification and coefficients for all variables are shown, along with the degrees of freedom, df, log-likelihood, logL, Akaike's Information Criterion, AIC, the difference in AIC compared with best model, ΔAIC , and the model importance considering all uncorrelated models (that is, the weighted AIC).

Models	Intercept	Productivity	Habitat	River	Euclidean	Topography	df	logL	AIC	ΔAIC	Weighted AIC
1	-0.127	.	-0.032	-0.111	.	.	6	172.178	-332.355	0.000	1.000
2	-0.128	.	.	-0.135	.	.	5	158.560	-307.120	25.236	0.000
3	-0.120	-0.120	5	136.049	-262.098	70.258	0.000
4	-0.120	.	.	.	-0.119	.	5	135.953	-261.906	70.449	0.000
5	-0.113	-0.135	5	134.948	-259.897	72.458	0.000
6	-0.130	.	-0.070	.	.	.	5	119.758	-229.516	102.840	0.000
7	-0.134	4	88.031	-168.062	164.293	0.000

Table S1.14. Parameter estimates for the best regression models ($\Delta AIC < 2$) using MLPE for *P. simonsi* and *P. steerei*, with the 95% of confidence intervals shown in parentheses.

Predictors	Estimates
<i>P. simonsi</i>	
Habitat	-0.109 (-0.136 to -0.083)
Productivity	-0.001 (-0.060 to 0.008)
River Distance	0.000 (-0.036 to 0.024)
<i>P. steerei</i>	
River Distance	-0.111 (-0.128 to -0.093)
Habitat	-0.032 (-0.044 to -0.021)

Table S1.15. Importance of predictors variables in MLPE regression for *P. simonsi* (the non-flooded forests/*terra-firme* species) and *P. steerei* (the seasonal floodplain forests/*várzea* species), considering the best models (i.e., with $\Delta\text{AIC} < 2$), and coefficients of determination (R_{Σ}^2) for the predictor variable in the best model with confidence intervals (95% C.I.) in parentheses. R_{Σ}^2 corresponds to the proportion of generalized variance explained by the fixed predictors in the best model.

Predictors	<i>P. simonsi</i>		<i>P. steerei</i>	
	$\Delta\text{AIC} < 2$	R_{Σ}^2	$\Delta\text{AIC} < 2$	R_{Σ}^2
Habitat	1	0.349 (0.467 – 0.234)	1	0.099 (0.194 - 0.031)
Productivity	0.502	.	.	.
River	0.285	.	1	0.565 (0.645 – 0.479)
Topography
Euclidean

APPENDIX S2

Appendix S2.1. Details on library preparation, sequencing, SNPs calling and filtering steps.

Genomic DNA from liver and muscle samples were extracted with DNeasy Blood and Tissue Kit (Qiagen, Valencia, CA), following the manufacturer's recommendations, except for the DNA elution step where DNA was eluted in double distilled water (ddH₂O). Extracted DNA was quantified by Qubit fluorometer (Life Technologies, Grand Island, NY, USA), and diluted with ddH₂O or concentrated in a SpeedVac (ThermoFisher Scientific, Waltham, MA, USA) at 43° C (medium temperature) to a concentration of 17.6 ng/μL. Genomic library using the ddRAD-Seq technique was prepared following Peterson et al. (2012). Specifically, 300 ng of genomic DNA (*i.e.*, 17 μL of extracted DNA) were digested using two restriction enzymes: *Eco*-RI and *Mse*-I, cleaned with commercial Ampure XP Beads (Beckman Coulter, Brea, CA, USA), and quantified in a Qubit fluorometer (Life Technologies, Grand Island, NY, USA). Illumina adapters and a unique bar code were ligated to 50 ng of the digested DNA in a volume of 33 μL per sample. Samples were pooled and cleaned with commercial Ampure XP Beads (Beckman Coulter, Brea, CA, USA), and DNA fragments were size-selected (between 350 and 450 bp) using Pippin Prep (Sage Science, Beverly, MA, USA) and amplified by PCR. The library was cleaned with the beads again, quantified and sequenced in one lanes of a HiSeq2500 (Illumina, San Diego, CA, USA) at the Center for Applied Genomics in Toronto, Canada to generate 150 bp, single-end reads.

Genomic DNA was processed using the STACKS 2.41 pipeline (Rochette et al., 2019). Samples were demultiplexed, and reads with low quality, more than 2 mismatches by barcode, or with uncalled nucleotides (Ns), were removed, as were the adapters and barcodes. The 140 bp reads were stacked and aligned without a reference genome (*i.e.*, *de novo* alignment) for stacks with 6 or more reads (parameter *-m*) and 3 nucleotides of distance between stacks (*-M*) (Paris et al., 2017). A catalog of loci for all individuals was made, with up to 3 fixed differences between individuals (*-n*) (Paris et al., 2017).

Data were filtered using a customized R script and the R package “r2vcftools” v0.0.0.9 (Pope, 2020), removing highly variables loci (> 6 segregating sites per read with 132 pb) and those with coverage less than 20X, retaining loci with 5% minimum allele frequency that also conformed to Hardy-Weinberg equilibrium (HWE, $p < 0.0001$; O’Leary et al., 2018) and were not in linkage disequilibrium (LD; $r^2 < 0.8$; Larson et al., 2014). Loci with more than 20% missing data, or potentially under selection were removed. Putative loci under selection were determined from an outlier analysis based on principal component analysis (PCA) with Mahalanobis distance to identify outliers in a

multidimensional space with $\alpha = 0.05$ (Lotterhos et al., 2017; Luu et al., 2017) that does not require assignment of individuals to discrete populations (Privé et al., 2020) using the package “pcadapt” v4.3.3 (Luu et al., 2017); different K -values (1-10) were tested and the best K was chosen for tests of selection based on Cattell's Rule in the PCA scree plot (Cattell, 1966), with false positive rates adjusted based on the distribution of p -values and the genomic inflation factor (GIF) using the Benjamini-Hochberg algorithm (François et al., 2016).

Cattell, R. B. (1966). The scree test for the number of factors. *Multivariate Behavioral Research*, 1(2), 245–276.

https://doi.org/10.1207/s15327906mbr0102_10

François, O., Martins, H., Caye, K., & Schoville, S. D. (2016). Controlling false discoveries in genome scans for selection.

Molecular Ecology, 25(2), 454–469. <https://doi.org/10.1111/mec.13513>

Larson, W. A., Seeb, L. W., Everett, M. V., Waples, R. K., Templin, W. D., & Seeb, J. E. (2014). Genotyping by sequencing resolves shallow population structure to inform conservation of Chinook salmon (*Oncorhynchus tshawytscha*).

Evolutionary Applications, 7(3), 355–369. <https://doi.org/10.1111/eva.12128>

Lotterhos, K. E., Card, D. C., Schaal, S. M., Wang, L., Collins, C., & Verity, B. (2017). Composite measures of selection can improve the signal-to-noise ratio in genome scans.

Methods in Ecology and Evolution, 8(6), 717–727.

<https://doi.org/10.1111/2041-210X.12774>

Luu, K., Bazin, E., & Blum, M. G. B. (2017). pcadapt: an R package to perform genome scans for selection based on principal component analysis.

Molecular Ecology Resources, 17(1), 67–77. <https://doi.org/10.1111/1755-0998.12592>

O’Leary, S. J., Puritz, J. B., Willis, S. C., Hollenbeck, C. M., & Portnoy, D. S. (2018). These aren’t the loci you’e looking for: principles of effective SNP filtering for molecular ecologists.

Molecular Ecology, 27(16), 3193–3206.

<https://doi.org/10.1111/mec.14792>

Paris, J. R., Stevens, J. R., & Catchen, J. M. (2017). Lost in parameter space: a road map for stacks. *Methods in Ecology and Evolution*, 8(10), 1360–1373.

<https://doi.org/10.1111/2041-210X.12775>

Peterson, B.K., Weber, J.N., Kay, E.H., Fisher, H.S. & Hoekstra, H.E. (2012) Double Digest RADseq: An Inexpensive Method for De Novo SNP Discovery and Genotyping in Model and Non-Model Species L. Orlando (Ed). *PLoS ONE* 7, e37135.

Pope, N. (2020). *r2vcftools: An R interface for vcftools*. R package version 0.0.0.9000.

Privé, F., Luu, K., Vilhjálmsson, B. J., Blum, M. G. B., & Rosenberg, M. (2020). Performing highly efficient genome scans for local adaptation with R package pcadapt version 4.

Molecular Biology and Evolution, 37(7), 2153–2154.

<https://doi.org/10.1093/molbev/msaa053>

Rochette, N. C., Rivera-Colón, A. G., & Catchen, J. M. (2019). Stacks 2: Analytical methods for paired-end sequencing improve RADseq-based population genomics.

Molecular Ecology, 28(21), 4737–4754.

<https://doi.org/10.1111/mec.15253>

Appendix S2.2. Details on the methodologies of the Procrustes analyses and the calculation of the resistance matrices used in isolation by resistance (IBR) analyses and Mantel tests. All analyses were performed in R 3.6.3.

Procrustes Analyses

Procrustes analyses were also used to test the similarity between taxa in the structuring of genetic variation, under the expectations of isolation by distance (IBD) (Knowles et al., 2016; Wang et al., 2010). Unlike traditional IBD tests, a Procrustes analysis retains information about the relative positions of populations both latitudinally and longitudinally, providing visualizations of the magnitude and the direction of deviations of individuals from IBD (Knowles et al., 2016; Papadopoulou & Knowles, 2015; Peres-Neto & Jackson, 2001). Specifically, the association between geography and genetic variation was quantified using the first two principal components of the genetic PCA (see in Material and Methods section on the Mantel tests) and the geographical coordinates (i.e., longitude and latitude) of the sampling localities, minimizing the differences between the geo-genetic maps in the multivariate space (Peres-Neto & Jackson, 2001; Wang et al., 2012). The strength of the association is quantified by the t_0 -statistic that ranges from 0 (for no association) to 1 (a perfect association) (Wang et al., 2010), and the robustness of t_0 was evaluated using a sequential removal and replacement procedure to identify sensitivities to particular geographic localities (measured by t''), or individuals (as measured by t' , following the nomenclature of Knowles et al., 2016; Prado et al., 2019; Wang et al., 2012). Analyses were performed using the function *protest* in the R package “vegan” with 10,000 permutations.

Geographic (Euclidean) Distance

The geographic distance in km was calculated for the IBR analyses and the Mantel tests. We estimated the distances between individuals using the *dism* function in the package “geosphere” v1.5-10 (Hijmans, 2019).

Topographic Distance

Topographic distance was calculated using SRTM elevation data from WorldClim 2.1 (Fick & Hijmans, 2017; <https://www.worldclim.org/>) and “topoDistance” v.1.0.1 with the package *topoDist* function (Wang, 2020).

River Network Distance

River network distance was estimated using the package “riverdist” v0.15.3 (Tyers, 2020), using the shapefile “ne_10m_rivers_lake_centerlines” from Natural Earth Data (<https://www.naturalearthdata.com/downloads>). The *cleanup* function was used to identify and dissolve segments, remove duplicate lines, and insert vertices in every 500 m in the river network before applying the *xy2segvert* function to snap sampling points to the closest vertex available in the network; river distances were calculated in Km using the *riverdistancemat* function.

Habitat Resistance Distance

The habitat resistance distance matrix was generated from the Tropical and Subtropical wetlands distribution map (Gumbricht et al., 2017), a binary map of wetlands and non-wetlands. However, these wetlands classifications also include permanent wetlands and swamps (open areas) where *Proechimys* does not occur (Wittmann et al., 2010). Because of that, we reclassified the cells in the map to differentiate flooded areas (permanent or not) from non-flooded areas (permanently dry) where only *P. simonsi* occurs, and thus can be used as a proxy for habitat preferences. Specifically, different resistance values were assigned to wetlands and non-wetlands habitats in a species-specific manner since *P. simonsi* occurs in the non-flooded forests (*terra-firme*), whereas *P. steerei* occurs in from seasonal floodplain forests (*várzea*). Resistance distances were conducted in each species using a range of resistance values for each wetland versus non-wetland habitat representing very high resistance (VH), high resistance (H); moderate resistance (M), and low resistance (L). Because we had an insufficient number of sampling localities and classes to optimize resistance values (e.g., using *ResistanceGA* function in R; Peterman, 2018), we instead explored a range of resistance values (see below) for both wetland and non-wetland habitats. This approach is admittedly arbitrary, but are nonetheless useful because we are primarily interested in whether the same or different models fit the sympatric taxa (as opposed to identifying the optimal resistance value in a single species). For *P. steerei*, the species from seasonal floodplain forests (*várzea*), cells with non-wetland habitats (i.e., the 0 values from the binary map from Gumbricht et al., 2017) were replaced with higher levels of resistance (i.e., 0.9 for the VH scenario; 0.8 for the H scenario; 0.7 for the M scenario; 0.6 for the L scenario) and cells with wetlands (i.e., the 1 values from the binary map from Gumbricht et al., 2017) were replaced with low levels of resistance (i.e., 0.1, 0.2, 0.3, and 0.4 for the VH, H, M, and L scenarios, respectively). For *P. simonsi*, the species from the non-flooded forests (*terra-firme*), the high levels of resistance were assigned to wetlands, and the low levels of resistance were assigned to the non-wetlands (i.e., the reverse assignment to cells as for *P. steerei*). Resistance distances between individuals were calculated based on random-walk commute time using the functions: *transition*,

geoCorrection and *commuteDistance* for “*gdistance*” v1.3-6 (van Etten, 2018), using the *mean* of 16 directions for cells connections to calculate the transition values of the grids. Among the four scenarios, we selected only the best model ($\Delta AIC = 0$) through MLPE univariate models and *model.sel* function in “MuMIn” R package (Bartoń, 2020). The best MLPE univariate model for habitat distance was the VH scenario (with resistance values of 0.6 and 0.4 for movements in wetlands for *P. simonsi* and *P. steerei*, respectively; Table S1.4).

Habitat Productivity Distance

For the species distribution models (SDM) 286 occurrence points were obtained by the authors during visits to scientific collections, using only specimens identified based on morphology following Patton & Leite (2015). All occurrence points were verified using Google Maps (<http://maps.google.com.br>), SpeciesLink (<http://slink.cria.org.br/>), and the ornithological gazetteers of the Neotropics (Stephens & Traylor, 1983; Paynter & Traylor, 1991a, b) in order confirm that the name and the coordinates corresponded to those designated in the museum catalog and in the specimen tags. After, filtering the occurrence points using a 10 Km buffer, the remaining dataset had a total of 88 unbiased occurrence points (*P. simonsi*, $n = 41$; *P. steerei*, $n = 47$), and were used to build the models (Table S1.3).

Temperature, precipitation, and potential evapotranspiration (PET) variables were used as proxy for habitat productivity (Field et al., 1998; Hawkins et al., 2003; Li et al., 2011; Yang et al., 2018; Moura et al., 2019). Seasonal floodplain forests show higher net primary production than the non-flooded forest (Junk and Piedade, 1993; Wittmann et al., 2010), and so both forest types might respond differently to variation in precipitation and temperature, as determined by tree-ring chronology data (Schöngart et al., 2010). For the temperature and precipitation variables, environmental values extracted from 12 monthly layers available in WorldClim 2.1 (Fick & Hijmans, 2017; <https://www.worldclim.org/>), and for the PET variables, values were extracted from six layers based on ENVIREM variables (Title & Bemmels, 2018; <https://envirem.github.io/>) (Table S1.2), all 30 layers were in a 1Km² resolution or 30 arc-seconds.

We employed the function *vif* in “*usdm*” v1.1-18 (Naimi et al., 2014) to select the environmental variables by variance inflation factor (VIF) to avoid collinearity (Dormann et al., 2013). For each species, we calculated the VIF for all variables, eliminated the variable with the highest VIF, and performed a new VIF estimation without the variable with the highest VIF. We repeated this approach of sequentially dropping one variable until we had 4 variables, at least one variable for each group (temperature, precipitation, and PET), following the rule of 10 observations by a predictor (Breiner et al., 2015; Harrison et al., 2018). In the end, all remaining variables had a $VIF \leq 2$, with four

variables for each target species (*P. simonsi* = PET in the Coldest Quarter, PET in the Warmest Quarter, Precipitation in November (Prec_11), and Average Temperature in January (tavg_01); *P. steerei* = PET in the Coldest Quarter, Annual PET, Precipitation in September (Prec_09), and Average Temperature in December (tavg_12).

Models were created in the program MaxEnt v3.4.1 (Phillips et al., 2006), implemented in the “biomod2” v3.5.1 with 50 replicates, using bootstrap and 30% of occurrence points as testing samples. To check the importance of the variables, we performed 100 replications. Study area were selected using convex hulls on the occurrence points plus 100 Km buffer. We eliminated areas of the Andes Mountains where the genus *Proechimys* did not occur and that showed high environmental heterogeneity which can bias our models. Complexity and the feature class (FC) for modeling were chosen in “ENMeval” v0.3.1 (Muscarella et al., 2014), with *ENMevaluate* function. We tested eight FC combinations: L, LQ, LQP, H, T, LQH, LQHP, and LQHPT (L = linear, Q = quadratic, H = hinge, T = threshold, and P = product); and regularization multiplier (RM) values from 0.5 to 3.0 with increments of 0.5, totaling 48 models. We used a lower Akaike information criterion value (i.e. ΔAIC) and other parameters to select the best models: higher Area Under the Curve (AUC) values, lower standard deviation in replicates, and lower difference between train and test AUC. Best model for *P. simonsi* was FC = L and RM = 0.5; and for *P. steerei* was FC = LQ, and RM = 0.5. We calculated the Area Under the Curve (AUC) True Skill Statistics (TSS) to validate the models. We created an average model with 20 replicates with higher TSS values to create our resistance matrix for habitat productivity. We found good accuracy metrics (>0.7 for AUC and >0.4 for TSS) for the average of replicates for AUC test (*P. simonsi* = 0.730; and for *P. steerei* = 0.713) and for TSS (*P. simonsi* = 0.448; and for *P. steerei* = 0.398). In *P. simonsi* the most important variable was PET in the Warmest Quarter (49.72%), followed by Average Temperature in January (35.04%), Precipitation in November (7.70%), and PET in the Coldest Quarter (7.54%). For *P. steerei* most important variable was Precipitation in September (58.78%) followed by Average Temperature in December (24.65%), Annual PET (15.69%), and PET in the Coldest Quarter (0.88%). Response curves for each variable are showed in Fig. S3.1. The raw values for suitability for each species were inverted (1-SDM) following Chan et al. (2011). After we applied the functions: *transition*, *geoCorrection* and *costDistance* for “gdistance” to calculate the cost distance between individuals with the mean of 16 directions for cells connections to calculate the transition values of the grids. All R scripts are available at: https://github.com/jdalapicolla/SDM_biomod2.

- Bartoń, K. (2020). *MuMIn: multi-model inference* (p. R package version 1.43.17).
<https://cran.r-project.org/package=MuMIn>
- Breiner, F. T., Guisan, A., Bergamini, A., & Nobis, M. P. (2015). Overcoming limitations of modelling rare species by using ensembles of small models. *Methods in Ecology and Evolution*, 6(10), 1210–1218.
<https://doi.org/10.1111/2041-210X.12403>
- Chan, L. M., Brown, J. L., & Yoder, A. D. (2011). Integrating statistical genetic and geospatial methods brings new power to phylogeography. *Molecular Phylogenetics and Evolution*, 59(2), 523–537.
<https://doi.org/10.1016/j.ympev.2011.01.020>
- Dormann, C. F., Elith, J., Bacher, S., Buchmann, C., Carl, G., Carré, G., Marquéz, J. R. G., Gruber, B., Lafourcade, B., Leitão, P. J., Münkemüller, T., McClean, C., Osborne, P. E., Reineking, B., Schröder, B., Skidmore, A. K., Zurell, D., & Lautenbach, S. (2013). Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. *Ecography*, 36(1), 27–46. <https://doi.org/10.1111/j.1600-0587.2012.07348.x>
- Fick, S. E., & Hijmans, R. J. (2017). WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. *International Journal of Climatology*, 37(12), 4302–4315. <https://doi.org/10.1002/joc.5086>
- Field, C. B., Behrenfeld, M. J., Randerson, J. T., & Falkowski, P. (1998). Primary production of the biosphere: integrating terrestrial and oceanic components. *Science*, 281(5374), 237–240. <https://doi.org/10.1126/science.281.5374.237>
- Gumbrecht, T., Roman-Cuesta, R. M., Verhot, L., Herold, M., Wittmann, F., Householder, E., Herold, N., & Murdiyarto, D. (2017). An expert system model for mapping tropical wetlands and peatlands reveals South America as the largest contributor. *Global Change Biology*, 23(9), 3581–3599. <https://doi.org/10.1111/gcb.13689>
- Harrison, X. A., Donaldson, L., Correa-Cano, M. E., Evans, J., Fisher, D. N., Goodwin, C. E. D., Robinson, B. S., Hodgson, D. J., & Inger, R. (2018). A brief introduction to mixed effects modelling and multi-model inference in ecology. *PeerJ*, 2018(5), e4794. <https://doi.org/10.7717/peerj.4794>
- Hawkins, B. A., Porter, E. E., & Diniz-Filho, J. A. F. (2003). Productivity and history as predictors of the latitudinal diversity gradient of terrestrial birds. *Ecology*, 84(6), 1608–1623. [https://doi.org/10.1890/0012-9658\(2003\)084\[1608:PAHAPO\]2.0.CO;2](https://doi.org/10.1890/0012-9658(2003)084[1608:PAHAPO]2.0.CO;2)
- Hijmans, R. J., & van Etten, J. (2015). raster: Geographic Data Analysis and Modeling. In *R package version 2.3-40*.
- Hijmans, R. J. (2019). *geosphere: spherical trigonometry* (p. R package version 1.5-10).
<https://cran.r-project.org/package=geosphere>
- Junk, W. J., & Piedade, M. T. (1993). Herbaceous plants of the Amazon floodplain near Manaus: Species diversity and adaptations to the flood pulse. *Amazoniana: Limnologia et Oecologia Regionalis Systematis Fluminis Amazonas*, 12(3/4), 467–484.
- Knowles, L. L., Massatti, R., He, Q., Olson, L. E., & Lanier, H. C. (2016). Quantifying the similarity between genes and geography across Alaska's alpine small mammals. *Journal of Biogeography*, 43(7), 1464–1476. <https://doi.org/10.1111/jbi.12728>
- Li, W., Zhang, P., Ye, J., Li, L., & Baker, P. A. (2011). Impact of two different types of El Niño events on the Amazon climate and ecosystem productivity. *Journal of Plant Ecology*, 4(1–2), 91–99. <https://doi.org/10.1093/jpe/rtq039>
- Moura, M. M., dos Santos, A. R., Pezzopane, J. E. M., Alexandre, R. S., da Silva, S. F., Pimentel, S. M., de Andrade, M. S. S., Silva, F. G. R., Branco, E. R. F., Moreira, T. R., da Silva, R. G., & de Carvalho, J. R. (2019). Relation of El Niño and La Niña phenomena to precipitation, evapotranspiration and temperature in the Amazon basin. *Science of the Total Environment*, 651, 1639–1651. <https://doi.org/10.1016/j.scitotenv.2018.09.242>

- Muscarella, R., Galante, P. J., Soley-Guardia, M., Boria, R. A., Kass, J. M., Uriarte, M., & Anderson, R. P. (2014). ENMeval: An R package for conducting spatially independent evaluations and estimating optimal model complexity for Maxent ecological niche models. *Methods in Ecology and Evolution*, 5(11), 1198–1205. <https://doi.org/10.1111/2041-210X.12261>
- Naimi, B., Hamm, N. A. S., Groen, T. A., Skidmore, A. K., & Toxopeus, A. G. (2014). Where is positional uncertainty a problem for species distribution modelling? *Ecography*, 37(2), 191–203. <https://doi.org/10.1111/j.1600-0587.2013.00205.x>
- Papadopoulou, A., & Knowles, L. L. (2015). Species-specific responses to island connectivity cycles: refined models for testing phylogeographic concordance across a Mediterranean Pleistocene Aggregate Island Complex. *Molecular Ecology*, 24(16), 4252–4268. <https://doi.org/10.1111/mec.13305>
- Paynter, R. A., & Traylor, M. A. (1991a). *Ornithological gazetteer of Brazil A-M*. Bird Department, Museum of Comparative Zoology, Harvard University. <https://www.biodiversitylibrary.org/item/50711>
- Paynter, R. A., & Traylor, M. A. (1991b). *Ornithological gazetteer of Brazil N-Z*. Bird Department, Museum of Comparative Zoology, Harvard University. <https://www.biodiversitylibrary.org/item/50708>
- Peres-Neto, P. R., & Jackson, D. A. (2001). How well do multivariate data sets match? The advantages of a procrustean superimposition approach over the Mantel test. *Oecologia*, 129(2), 169–178. <https://doi.org/10.1007/s004420100720>
- Peterman, W. E. (2018). ResistanceGA: an R package for the optimization of resistance surfaces using genetic algorithms. *Methods in Ecology and Evolution*, 9(6), 1638–1647. <https://doi.org/10.1111/2041-210X.12984>
- Phillips, S. J., Anderson, R. P., & Schapire, R. E. (2006). Maximum entropy modeling of species geographic distributions. *Ecological Modelling*, 190(3), 231–259. <https://doi.org/10.1016/j.ecolmodel.2005.03.026>
- Prado, J. R., Percequillo, A. R., Thomaz, A. T., & Knowles, L. L. (2019). Similar but different: Revealing the relative roles of species-traits versus biome properties structuring genetic variation in South American marsh rats. *Journal of Biogeography*, 46(4), 770–783. <https://doi.org/10.1111/jbi.13529>
- Schöngart, J., Wittmann, F., & Worbes, M. (2010). Biomass and net primary production of Central Amazonian floodplain forests. In Wolfgang J. Junk, M. T. F. Piedade, F. Wittmann, J. Schöngart, & P. Parolin (Eds.), *Amazonian floodplain forests: ecophysiology, biodiversity and sustainable management* (pp. 347–388). Springer. https://doi.org/10.1007/978-90-481-8725-6_18
- Stephens, L., & Traylor, M. A. (1983). *Ornithological gazetteer of Peru*. Bird Department, Museum of Comparative Zoology, Harvard University. <https://www.biodiversitylibrary.org/item/50707>
- Title, P. O., & Bemmels, J. B. (2018). ENVIREM: an expanded set of bioclimatic and topographic variables increases flexibility and improves performance of ecological niche modeling. *Ecography*, 41(2), 291–307. <https://doi.org/10.1111/ecog.02880>
- Tyers, M. (2020). *riverdist: river network distance computation and applications* (p. R package version 0.15.3). <https://cran.r-project.org/package=riverdist>
- van Etten, J. (2018). *gdistance: distances and routes on geographical grids* (p. R package version 1.2-2). <https://cran.r-project.org/package=gdistance>
- Velasco, J. A., & González-Salazar, C. (2019). Akaike information criterion should not be a “test” of geographical prediction accuracy in ecological niche modelling. *Ecological Informatics*, 51, 25–32. <https://doi.org/10.1016/J.ECOINF.2019.02.005>
- Wang, C., Szpiech, Z. A., Degnan, J. H., Jakobsson, M., Pemberton, T. J., Hardy, J. A., ... Rosenberg, N. A. (2010). Comparing spatial maps of human population-genetic variation using Procrustes analysis. *Statistical Applications in Genetics and Molecular Biology*, 9(1), 1–22. <https://doi.org/10.2202/1544-6115.1493>
- Wang, C., Zöllner, S., & Rosenberg, N. A. (2012). A quantitative comparison of the similarity between genes and geography in worldwide human populations. *PLoS Genetics*, 8(8), e1002886. <https://doi.org/10.1371/journal.pgen.1002886>

- Wang, I. J. (2020). Topographic path analysis for modelling dispersal and functional connectivity: Calculating topographic distances using the *topoDistance* r package. *Methods in Ecology and Evolution*, *11*(2), 265–272. <https://doi.org/10.1111/2041-210X.13317>
- Wittmann, F., Schöngart, J., & Junk, W. J. (2010). Phytogeography, species diversity, community structure and dynamics of Central Amazonian floodplain forests. In W. J. Junk, M. T. F. Piedade, F. Wittmann, J. Schöngart, & P. Parolin (Eds.), *Amazonian floodplain forests: ecophysiology, biodiversity and sustainable management* (pp. 61–102). Springer.
- Yang, J., Tian, H., Pan, S., Chen, G., Zhang, B., & Dangal, S. (2018). Amazon drought and forest response: largely reduced forest photosynthesis but slightly increased canopy greenness during the extreme drought of 2015/2016. *Global Change Biology*, *24*(5), 1919–1934. <https://doi.org/10.1111/gcb.14056>

APPENDIX S3

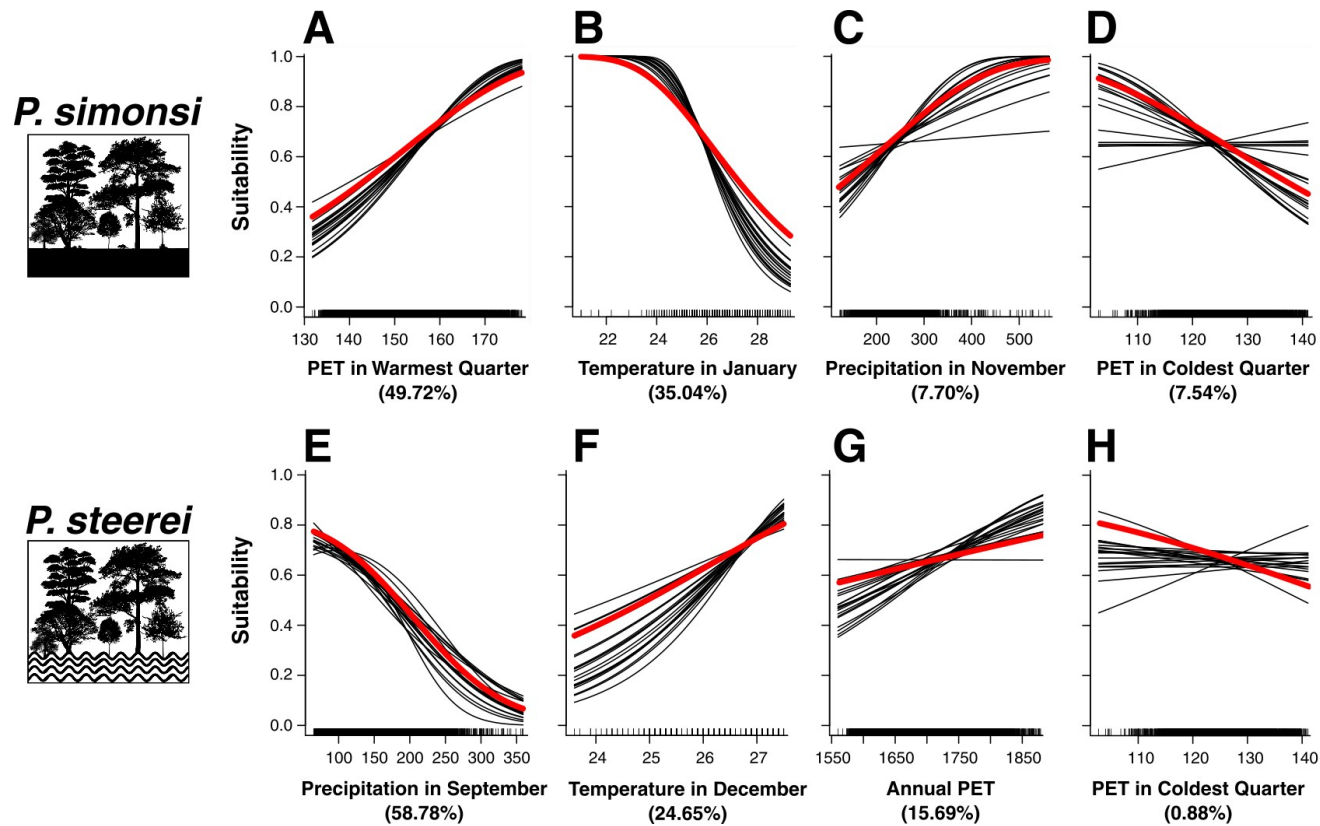


Figure S3.1. Response curves to explain the suitability of the four variables used in the species distribution models (SDM) in *P. simonsi* (A-D) and in *P. steerei* (E-H); percentages in parenthesis give the importance of variables for the model after 100 replications. Units in PET variables are in mm/month; Precipitation in mm, and Temperature in °C. Black lines represent the curves for all 20 model replicates and the red line for the replicate with better accuracy metric (AUC and TSS). Also shown are icons for the non-flooded forests/*terra-firme* in plots A-D, and the seasonal floodplain forests/*várzea* in plots E-H.

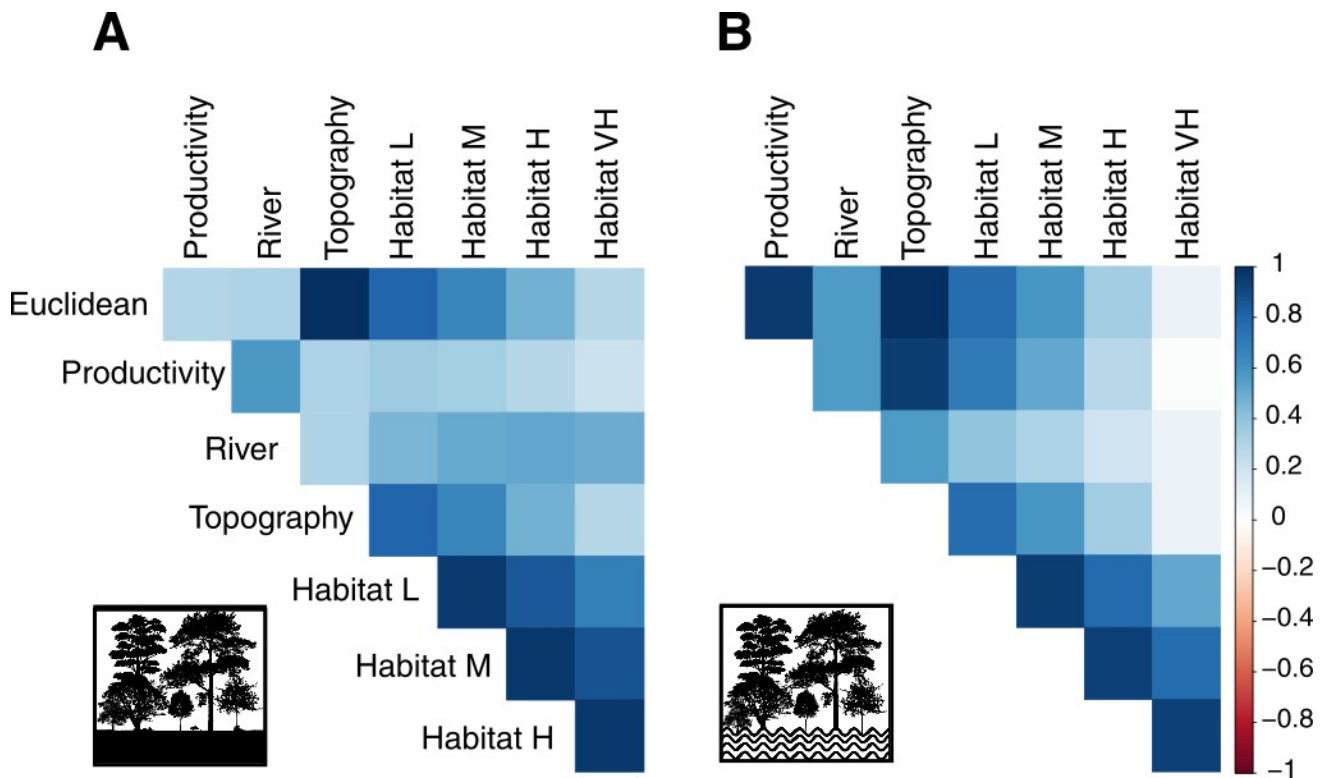


Figure S3.2. Pairwise correlations between resistance predictor variables used in MLPE analyses in (A) *P. simonsi* and (B) *P. steerei*, as measured by Pearson's correlation coefficient, r^2 , with the values color-coded and absolute values of the correlations (between 0 and 1) shown as the shaded proportion of the squares. Also shown are icons for the non-flooded forests/*terra-firme* in plot A, and the seasonal floodplain forests/*várzea* in plot B.

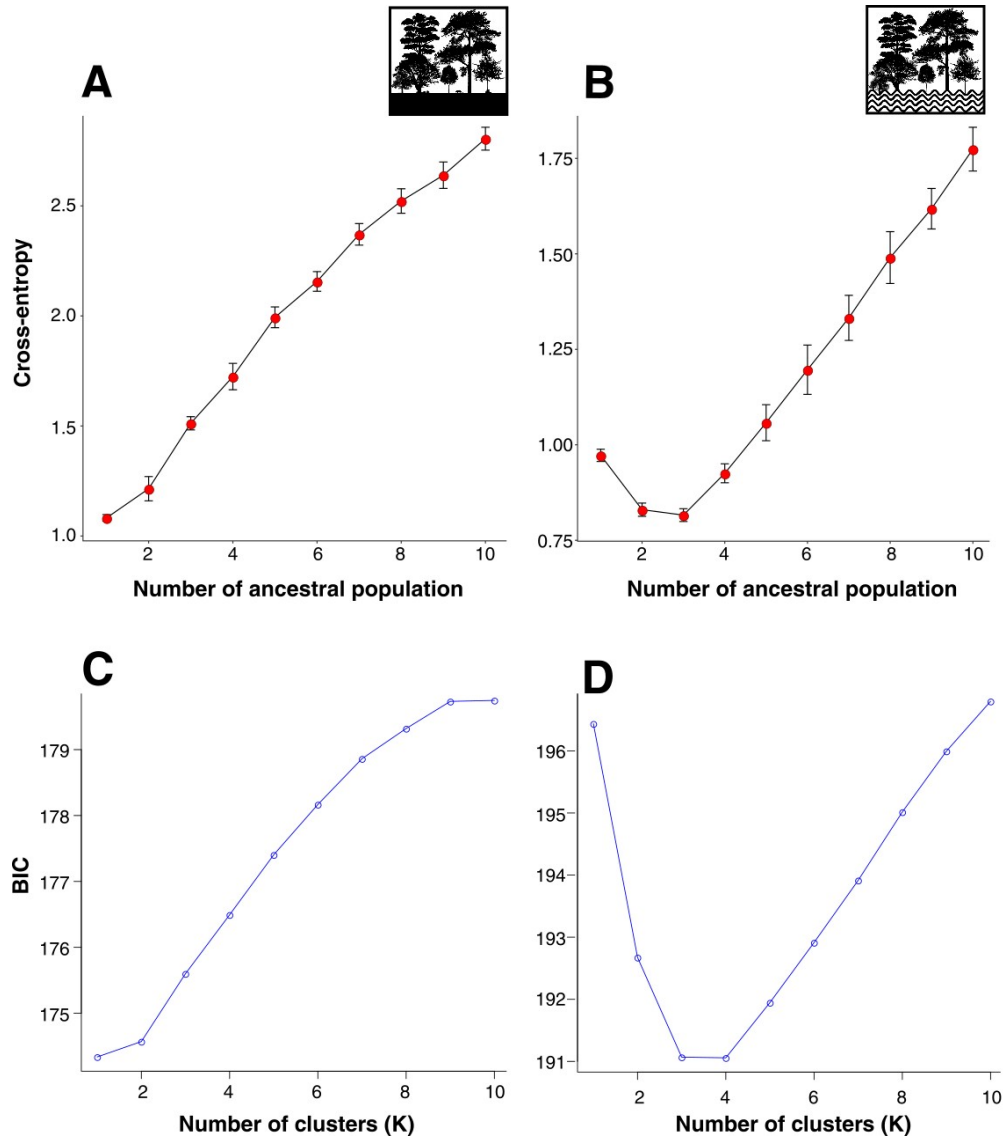


Figure S3.3. Number of ancestral populations (A,B) and number of genetic clusters (C,D) estimated with sNMF and DAPC for *P. simonsi* (on the left) and *P. steerei* (on the right), respectively, where the best number is the one with the smallest cross-entropy (sNMF) or BIC-values (DAPC). For cross-entropy the number of populations was consistent across different regularization parameters, showing plots with lower cross-entropy values $\alpha = 4000$ (A), and $\alpha = 4000$ (B). Note that based on the DAPC results, we focus on 3 genetic clusters in *P. steerei* throughout manuscript. Also shown are icons for the non-flooded forests/*terra-firme* in plot A, and the seasonal floodplain forests/*várzea* in plot B.

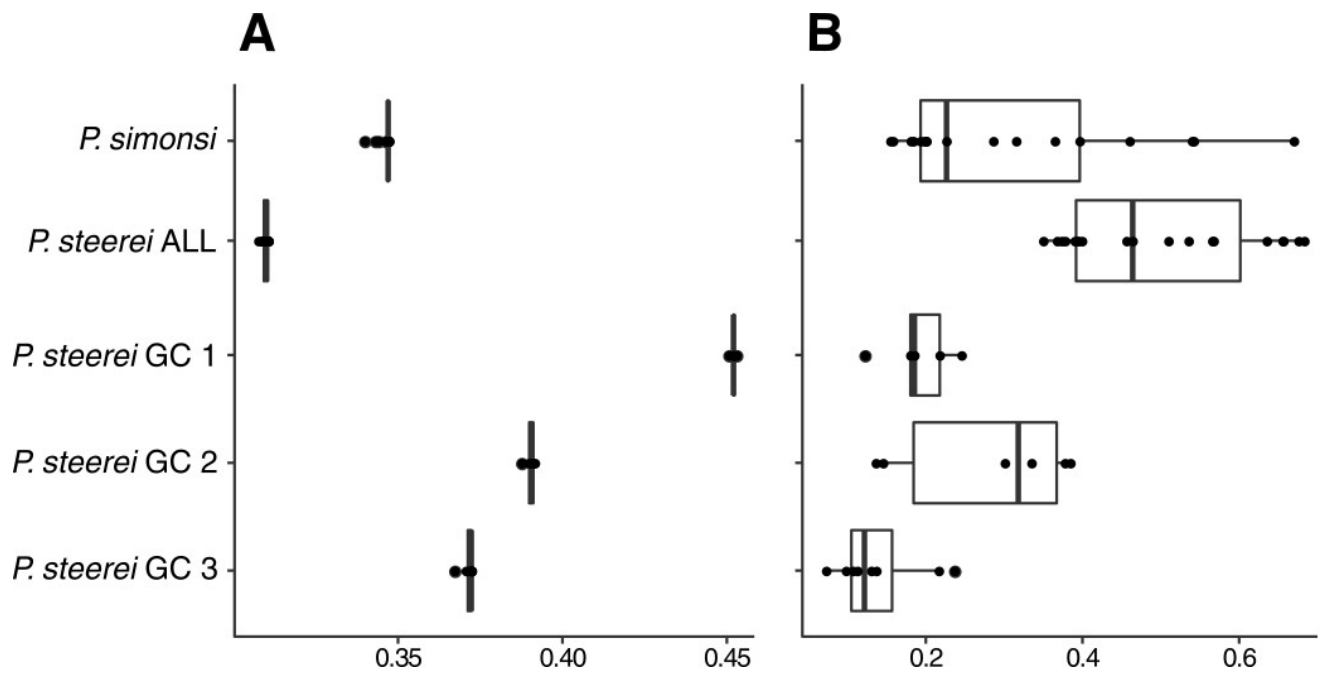


Figure S3.4. Comparison of individual-based (A) genetic diversity indices expected heterozygosity, H_{EXP} , (B) and the inbreeding coefficient, F_{IS} , for *P. simonsi* (the non-flooded forests/*terra-firme* species) and *P. steerei* (the seasonal floodplain forests/*várzea* species), as well as the three genetic clusters (GC) in *P. steerei* recovered using sNMF. *P. simonsi* is represented by one single cluster in the sNMF approach; all individuals were used for diversity estimates for species wide estimates.

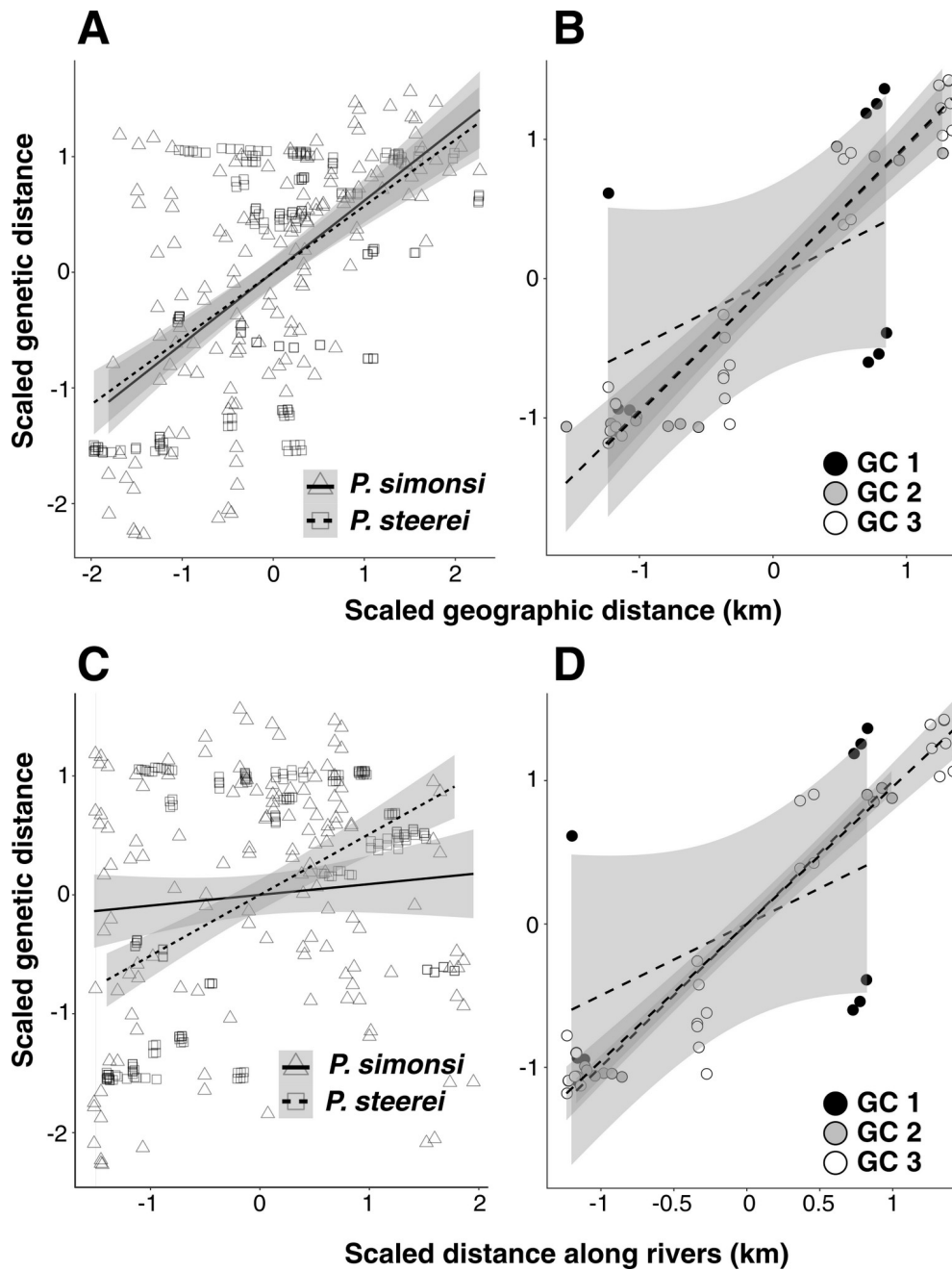


Figure S3.5. Tests of association between genetic and geographic distance based on Mantel tests for (A-C) both sympatric species using all individuals (triangles and solid line for *P. simonsi*; squares and dashed for *P. steerei*), and for (B-D) each of the three genetic cluster (GC) recovered using sNMF (marked by different colored circles). Lines represent the relationships between genetic distance (based on PCA distance among pairs of individuals) and geographical (Euclidean) distances (scaled values) are shown in A and B and distances along the rivers are in C and D, with the 95% confidence intervals under the linear models in the shaded gray areas. Pearson's correlation coefficient, r , a measure of the strength of the association, and p -values are given in Table S1.10.

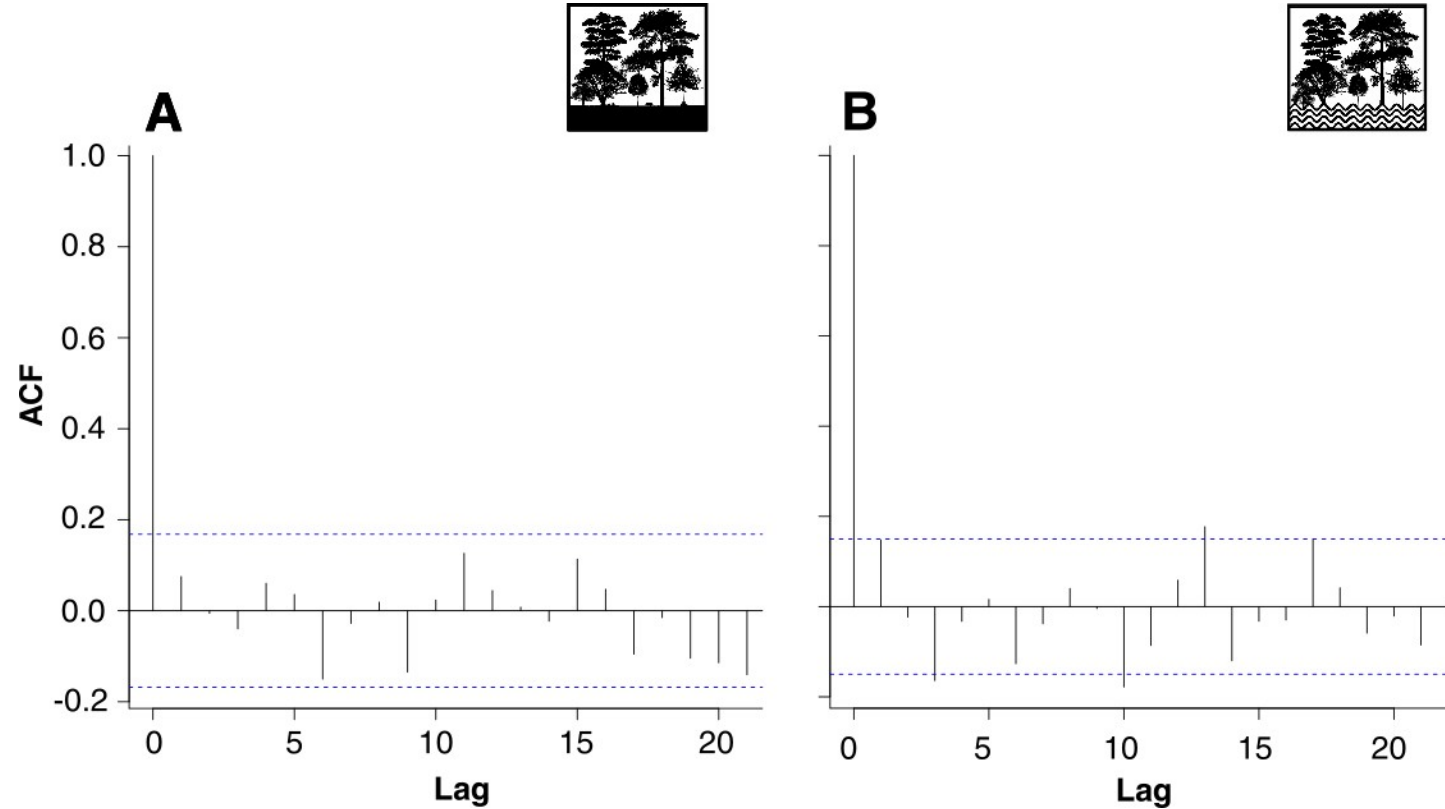


Figure S3.6. Auto-correlation function (ACF) plots showing the degree of spatial autocorrelation based on MLPE model residuals from the best model in (A) *P. simonsi* and (B) *P. steerei*, which have weak spatial autocorrelation among samples with lag values within or close to the confidence band (blue dashed lines). Also shown are icons for the non-flooded forests/*terra-firme* in plot A, and the seasonal floodplain forests/*várzea* in plot B.