# The influence of environmental variation on the genetic structure of a poison frog distributed across continuous Amazonian rainforest

Ferreira, A, Lima, AP, Jehle, R, Ferrao, M and Stow, A

http://dx.doi.org/10.1093/jhered/esaa034

# The influence of environmental variation on the genetic structure of a poison frog distributed across continuous Amazonian rainforest

ANTHONY S. FERREIRA[1*], ALBERTINA P. LIMA[2], ROBERT JEHLE[3], MIQUÉIAS FERRÃO[4] and ADAM STOW[5]

[1] Programa de Capacitação Institucional, Instituto Nacional de Pesquisas da Amazônia, Manaus, Amazonas, Brazil

[2] Coordenacão de Biodiversidade, Instituto Nacional de Pesquisas da Amazônia, Manaus, Amazonas, Brazil

[3]School of Science, Engineering and Environment, University of Salford, M5 4WT, Salford, UK

[4]Museum of Comparative Zoology, Harvard University, Cambridge, Massachusetts, USA

[5]Department of Biological Sciences, Macquarie University, Sydney, NSW, Australia

[*]Corresponding author: anthonyyferreira@gmail.com.br

**Abstract**

Biogeographic barriers such as rivers have been shown to shape spatial patterns of biodiversity in the Amazon basin, yet relatively little is known about the distribution of genetic variation across continuous rainforest. Here, we characterize the genetic structure of the brilliant-thighed poison frog (*Allobates femoralis*) across an 880 km long transect along the Purus-Madeira interfluve south of the Amazon river, based on 64 individuals genotyped at 7 609 SNP loci. A population tree and clustering analyses revealed four distinct genetic groups, one of which was strongly divergent. These genetic groups were concomitant with femoral spot coloration differences, which was intermediate within a zone of admixture between two of the groups. The location of these genetic groups did not consistently correspond to current ecological transitions between major forest types. A multi-model approach to quantify the relative influence of isolation-by-distance (IBD) and isolation-by-environmental resistance (IBR) nevertheless revealed that, in addition to a strong signal of IBD, spatial genetic differentiation was explained by IBR primarily linked to dry season intensity ($r^2$ = 8.4%) and canopy cover ($r^2$ = 6.4%). We show significant phylogenetic divergence in the absence of obvious biogeographical barriers and that finer-scaled measures of genetic structure show patterns that are associated with environmental variables also known to predict the density of *A. femoralis*.

*Keywords*: RADseq, genetic clusters, landscape genetics, Amazonia, amphibians

## Introduction

A key goal in ecology and evolutionary studies is to understand the processes that explain contemporary patterns of genetic diversity. Based on the classic allopatric speciation model, genetic divergence is a consequence of geographic isolation (Wallace 1852; Mayr 1963; Coyne and Orr 2004). However, divergence can also arise when isolation is incomplete, under scenarios that may include ecologically-mediated selection triggered by environmental heterogeneity (Nosil 2012; Shafer and Wolf 2013; see also Endler 1977 for an early 'gradient diversification hypothesis'). Recent evidence that incipient diversification along environmental clines is often associated with secondary contact of already existing ancient lineages (e.g., Dean et al. 2019; Marques et al. 2019) further suggests that, when species' range expand and contract over time, allopatric and sympatric diversification models are not necessarily mutually exclusive.

Neutral genetic population structure arises through the interplay of drift, mutation and migration. Disentangling the legacy of historical events on patterns of genetic structure from more contemporary effects needs to account for the sensitivity of the molecular assays, the analytical approaches employed, as well as recognizing the time required for causal processes to shape genetic structure (Stow et al. 2001; Anderson et al. 2010; Epps and Keyghobadi 2015). While isolation by geographic distance (IBD, Wright 1943; Slatkin 1987) is revealed by most empirical studies (for summaries see e.g. Jenkins et al. 2010; Sexton et al. 2014), gene flow can be further influenced by the landscape matrix where habitat heterogeneity results in different levels of resistance to migration (Manel et al. 2003; Storfer et al. 2010). Because patterns of isolation-by-environmental resistance (IBR) are influenced by species-specific life-history attributes and ecological preferences, such as propensity and ability for migration through given environments, they reveal essential information about

habitat relationships of the studied taxa (Balkenhol et al. 2017; Armansin et al. 2020). The spatial scale of sampling is an especially important consideration when testing for IBD and IBR. If the scale of sampling is too small relative to the scale of gene flow of the target species, gene flow from beyond the study area may overwhelm patterns of genetic structure mediated by local environmental variables (Anderson et al. 2010). On the other hand, observed genetic discontinuities may also have arisen from past events rather than contemporary landscapes, due to a time lag between demographic processes and their consequences for population genetic structure (Epps and Keyghobadi 2015).

For the world's largest area of continuous rainforest in Amazon basin, the main processes responsible for spatial patterns of biodiversity remain debated (Moritz et al. 2000; Hoorn et al. 2010; Ribas et al. 2012; Leite and Rogers 2013). The majority of empirical studies demonstrate that the retraction of past environmental barriers in the Holocene resulted in range expansions of lineages that diverged in isolation up to about 0.8 million years ago (Ma), with major rivers often acting as local biogeographic boundaries (e.g. Naka et al. 2012; Nazareno et al. 2017; Ribas et al. 2018; Thom et al. 2020). The vast, forested areas between major rivers of the Amazon basin are however also characterized by gradual environmental variation, for which patterns of IBD and, possibly, IBR might be expected for broadly distributed taxa. However, difficulties in systematically sampling the vast, often inaccessible terrain of the Amazon basin has resulted in the gradient hypothesis receiving little attention (Beheregaray et al. 2015).

Amphibians are well suited to detect environmental and geographic influences on genetic divergence because they have low dispersal abilities and are sensitive to ecological conditions (e.g., Zeisset and Beebee 2008; Pabijan et al. 2020). The brilliant-thighed poison frog *Allobates femoralis* (Dendrobatoidea: Aromobatidae, Grant et al. 2017) is a small (~ 33 mm), ground-dwelling, iteroparous diurnal frog commonly

distributed throughout primary forest in the Amazon basin (Silverstone 1975; Amézquita et al. 2009), and likely comprises cryptic taxa (Grant et al. 2006, 2017; Fouquet et al. 2007; Santos et al. 2009; Simões et al. 2010). It prefers clay-rich soils and is more abundant in open forest than in forest with closed canopies (Ferreira et al. 2018). Males exhibit territorial behavior and signal territory ownership by calling from elevated positions on the forest floor (Roithmair 1994; Montanarin et al. 2011), with their mating success possibly correlated to territory size (about 200 $m^2$ maximally, Kaefer et al. 2012). Females lay egg clutches under leaf litter during the rainy season, and tadpoles are usually transported by males to ephemeral puddles in order to complete their development (Ringler et al. 2013). Both sexes are highly polygamous (Ursprung et al. 2011), and life-time dispersal rates are generally low (about 100 m, Ringler et al. 2009; Pašukonis et al. 2016). Populations across Amazonia vary in the coloration of a conspicuous femoral spot, which is both an aposematic signal through mimicry with syntopic toxic species as well as sexually selected trait (Amézquita et al. 2009, 2017; Ferreira et al. unpublished).

Here, we assess environmental and historical influences on the spatial genetic structure of *A. femoralis* along an ˜880 km long transect in the Purus-Madeira interfluve (PMI) south of the Amazon river. We explore the existence of local genetic structure along the transect using clustering techniques, and assess whether the genetic structure of *A. femoralis* conforms to previous studies on other taxa along the same transect (Ortiz et al. 2018; De Abreu et al. 2018). In parallel, we employ landscape genetic inferences to compare the relative contribution of IBD and IBR, predicting that genetic structure will be influenced by landscape variables that have previously been shown to determine the occurrence and abundance of *A. femoralis* along this transect (land cover, silt content, temperature seasonality, and intensity of the dry season; Ferreira et al.

126 2018). We also test whether there are genetic signals for selection associated with these
127 variables. Finally, we examine whether patterns of femoral spot coloration are
128 congruent with distinct genetic lineages and whether there is any evidence of lineage
129 admixture.
130
131 **Materials and Methods**
132 *Study area and sampling*
133 The Purus-Madeira interfluve (PMI) is situated south of the Amazon river and covers
134 approximately 15.4 million hectares, with vegetation, soil and climatic conditions
135 gradually changing along a latitudinal gradient (Cintra et al. 2013; Schietti et al. 2016).
136 The mean annual precipitation varies from 2200 mm to 2800 mm, and is highest in
137 central areas (Alvares et al. 2013; Fick and Hijmans 2017). The northeast of the PMI is
138 characterized by dense lowland rainforest with a mean tree basal area of 56.45 $m^2$ $ha^{-1}$,
139 plinthosols with a predominance of silt, and a complex hydrography with large
140 seasonally flooded areas (Fan and Miguez-Macho 2010; Schietti et al. 2016).
141 Southwestern and central parts are characterized by open lowland rainforest with a
142 mean tree basal area of 19.31 $m^2$ $ha^{-1}$, podzolic soils with high clay content, and small
143 temporary rivers filled during the rainy season (Cintra et al. 2013; Ferreira et al. 2018).
144 Considerable areas of savanna are also present between these two forested regions
145 (IBGE 1997; Figure 1).
146 Between November and March 2010-2015, we collected a total of 66 *A.*
147 *femoralis* individuals from 13 localities along an established 880 km transect which
148 runs in parallel to a federal highway (BR-319), and spans the entire length of the PMI
149 (Figure 1; Table S1). Sampling was carried on regularly spaced biodiversity monitoring
150 plots (modules) constructed by the Rapid Assessment for Long Duration Ecological

Projects system (RAPELD; for details see Magnusson et al. 2013). The same sampling design has previously been used to quantify environmental correlates for the occurrence and abundance of *A. femoralis* (Ferreira et al. 2018), and revealed that the species is present in all but three modules (M3-5, see Figure 1). *Allobates femoralis* was sampled by acoustic and visual surveys during the daily periods of peak vocalization (7:00-10:00 a.m. and 14:00-18:00 p.m.). We captured frogs by hand and maintained them in sealed plastic bags until arrival in the laboratory, where they were sacrificed and fixed after tissue (leg muscle) was removed for genetic analyses and stored in 96% ethanol. For each captured individual, the femoral spot coloration was noted as yellow, red, or orange.

*DNA extraction, genotyping and initial filtering*

Extraction of DNA and SNP discovery was carried out at Diversity Arrays Technology sequencing Pty. Ltd. (DArTseq) facility (Canberra, Australia; more detail in Supplementary Information Text S1). A modified double-digest restriction-site associated DNA (ddRAD) sequencing protocol was performed on libraries prepared using a combination of *Pst*l-H*pa*ll restriction enzymes (Kilian et al. 2012). The *Pst*l enzyme adaptor also contained an Illumina adaptor sequence, a primer sequence and a variable-length barcode as described by Elshire et al. (2011). The H*pa*ll adaptor contained an Illumina flow cell attachment and overhang sequence. Following enzymatic digestion, fragments were amplified and sequenced on an Illumina HiSeq2500. DNA sequences were aligned via BLAST using the *Nanorana parkeri* reference genome (Sun et al. 2015). To check for contamination, sequences were also blasted to bacterial and fungal genomes (NCBI).

A raw dataset of 147 595 SNPs was filtered for missing data using the *filter_dart*

function of the *R* package RADIATOR v. 0.010 (Gosselin 2017). Only individuals and loci with ≥ 95% SNPs genotyped were retained. SNPs were also screened for allele coverage, with any SNPs displaying a local and global minor allele frequency (MAF) threshold of less than 1% removed from the dataset. In cases where multiple SNPs were found within the same read, only one locus was retained (chosen randomly per RAD tag) to avoid statistical bias from physical linkage (Lemay and Russello 2015; Zheng et al. 2012). Two samples from M14 had < 95% of loci genotyped and were removed, which resulted in 64 individuals from 13 populations genotyped at 10 275 SNPs (see Table S2 for summary of filtering steps). File types required for downstream analyses were created using the RADIATOR package (Gosselin 2017), PGDSpider v. 2.1.1.3 (Lischer and Excoffier 2012) and PLINK v. 1.9 (Chang et al. 2015).

*Phylogenomic relationships*

In order to evaluate the evolutionary relationships among *A. femoralis* possessing different femoral spot coloration we constructed a population tree by coalescence using SNAPP v. 1.4.1 (Bryant et al. 2012) implemented in BEAST v. 2.5 (Bouckaert et al. 2014). This analysis assumes a lack of gene flow among lineages which is inferred by phenotypic distinctiveness and further tested using clustering analyses. To reduce computational requirements and run times, we selected 2-3 representative individuals per population without signatures of between-population admixture (assessed though femoral spot color). We used our data set of 10 275 SNPs, and mutation rates (*u* and *v*) as estimated by SNAPP, with the birth rate (λ) of the Yule prior based on the number of samples used. The trial run for each dataset used a chain length of 1000 000 generations, sampling every 1 000 trees. We inspected final log files and created maximum clade credibility trees (median node heights) by combining three independent

runs in TreeAnnotator v. 2.5 implemented in BEAST after discarding 25% as burn-in.

*Detection of SNPs associated with selection*

We removed SNPs with evidence of being associated with selection because our population and landscape genetic inferences assume neutral loci (see e.g., Rellstab et al. 2015). Analyses to detect loci associated with selection were conducted on the full dataset using two different approaches. First, we detected SNPs under putatively positive or negative selection using $F_{ST}$ outlier analysis (OA) with BayeScan v.2.1 (Foll and Gaggiotti 2008), a Bayesian method based on a logistic regression model which is suited to detecting outliers in scenarios with low-admixtured samples while taking into account sample size and genetic structure (Villemereuil et al. 2014; Luu et al. 2017). We ran BayeScan using a prior model (prior odds parametrization) set to 100, thinning interval of 10-20 pilot runs of length 10 000, and burn-in of 50 000 steps. Second, we used Environmental Association Analysis (EAA) with Latent Factors Mixed Models (LFMM), implemented in the *R* package LEA v. 2.1.0 (Frichot and François 2015). LFMM uses a hierarchical Bayesian mixed model based on residuals from PCA that take population genetic structure into account (e.g. Benestan et al. 2016). We ran LFMMs for each of the four environmental variables which were previously identified as predictors of local abundance (Ferreira et al. 2018): land cover, silt content, temperature seasonality, and intensity of the dry season, separately using 10 000 iterations, a burn-in of 5 000 steps, and 5 repetitions. We set both BayeScan and LFMM with a false discovery rate of 0.05 (5%). We also investigated whether the SNPs identified as signaling selection could be attributed to a functional part of the genome in order to complement our tests of the influence of landscape variables on gene flow, as variables influencing connectivity may also impose selection (Armansin et al. 2020).

Consequently, gene annotations were sought for RAD tags that contained SNPs identified with both BayeScan and LFMM using the NCBI BLAST platform (Johnson et al. 2008). Sequences were annotated to genes classified as '*amphibians*' (taxid:8292), 'vertebrates' (taxid:7742) and aligned using the *Nanorana parkeri* (taxid:125878) reference genome (Sun et al. 2015), using BLAST with an E-value threshold of 0.0001.

All SNPs that provided evidence for selection were removed from the data set for all downstream analyses of genetic structure. Summary statistics were calculated for each of the modules and any remaining loci that deviated from Hardy-Weinberg Equilibrium at a Bonferroni-correction $\alpha = 0.004$ (1 000 simulations) were also excluded from the dataset. Estimates of observed ($H_O$) and expected ($H_E$) heterozygosity, inbreeding coefficients ($F_{IS}$) and private alleles were calculated using the R-package *diveRsity* v. 1.9.90 (Keenan et al. 2013) with 95% confidence interval calculated with 1 000 bootstraps.

*Genetic Structure*

Genetic structure was described with putatively neutral loci using the model-based clustering approaches implemented by ADMIXTURE (Alexander et al. 2009) and sNMF in the *R* package LEA v. 2.1.0 (Frichot et al. 2014). To ensure that the underlying genetic structure was not violating the assumptions of these models, we also carried out Discriminant Analysis of Principal Components (DAPC) calculated using the *R* package *adegenet* v. 2.1.1 (Jombart et al. 2010). Genetic partitioning was further described by calculating pairwise $F_{ST}$ between 11 sites in the R-package *adegenet* v. 1.3.1 (Jombart and Ahmed 2011).

sNMF is a method based on sparse non-negative Matrix Factorization algorithms (NMF) and least-squares optimization (Frichot et al. 2014). We tested the

number of genetic clusters (*K*) ranging from 1 to 11 (upper limit equal to the number of sampling localities) with 20 independent runs per test, alpha set at 100, a tolerance error of 0.00001, entropy set as true (where the cross-entropy criterion is calculated), a random seed of 50, and 10 000 interactions in the algorithm. The best-supported *K* was determined by the lowest error value of ancestry through the cross-entropy criterion. ADMIXTURE simultaneously estimates the probability of the observed genotypes using ancestry proportions and population allele frequencies (Alexander et al. 2009). Significance was defined at $p < 0.05$, above which individuals were considered pure. We ran ADMIXTURE using a cross-validation with a random seed as 43, the block relaxation algorithm as the point estimation method, QuasiNewton as the convergence acceleration algorithm, and a delta of < 0.0001 to terminate point estimations. The number of *K* was determined by the lowest cross-validation error value. DAPC is a multivariate method that performs discriminant functions to describe the relationships between clusters as well as membership probabilities of each individual for different groups, optimizing variance between groups while minimizing variance within groups (Jombart et al. 2010). We used cross-validation to define the number of principal components (PCs) retained in the analysis, identifying the optimal point in the trade-off between retaining too few and too many PCs in the model. We used the number of PCs associated with the lowest Root Mean Squared Error - RMSE as the optimum number for the PCA in the DAPC analysis. Eight PCs and two DAs were retained for the analyses, and explained 41% of the total variance. To test whether the number of sampled individuals in each module was sufficient for the inferences of genetic structure, we ran the above analyses with two alternative datasets: all individuals sampled, and three randomly chosen individuals for each module only.

*Construction of Environmental Resistance Surfaces*

To test the effects of landscape variables on genetic connectivity in *A. femoralis*, we used four environmental variables with ecological effects for the species as predictors of local abundance (see Ferreira et al. 2018): land cover, silt content, temperature seasonality (representing the annual range in temperatures) and the Walsh index, a measure of the intensity and duration of the dry season (Walsh 1996). Environmental data were obtained from the public repository Ambdata (www.dpi.inpe.br/Ambdata; Amaral et al. 2013), and converted to raster format using the *R* package *raster* v. 2.6.7 (Hijmans 2017) with a cell resolution of 30 arcsecond (1 $km^2$). To avoid model overparameterization, we tested for collinearity between variables through pairwise Pearson's correlations analyses based on values extracted of each sampling location. The four variables were not strongly correlated with each other ($r < 0.65$ in all cases) and were therefore retained. To facilitate comparisons among surfaces, we standardized all raster files to values between 1 and 100 (following Row et al. 2017, see Figure 2).

We generated multiple resistance surfaces from our environmental variables to test multiple hypotheses about their effects on genetic distance following Yadav et al. (2019), evaluating each resistance surface model separately. We assumed that resistance in each raster cell was a function of environmental variables as follows:

$$ri = 1+\alpha \left(\frac{vi-1}{max-1}\right)^{\gamma},$$

where $ri$ is the resistance of raster cell $i$, $vi$ is the environmental variables value in cell $i$, and *max* is the maximum value of the raster surface (in our case 100, see above). Furthermore, $\alpha$ is a parameter that determines the maximum possible resistance value, and $\gamma$ is an exponent that determines the shape of the relationship (slope) between environmental variable values ($vi$) and resistance ($ri$), being linear when $\gamma = 1$ and nonlinear when $\gamma \neq 1$ (Shirk et al. 2010; Dudaniec et al. 2013, 2016). This approach has

been shown to effectively identify IBR including linear and non-linear relationships (Shirk et al. 2010; Dudaniec et al. 2013, 2016; Yadav et al. 2019). The equation expresses resistance as a function of the effect of landscape features. Based on previous information (Ferreira et al. 2018), we assume that the effects of land cover and temperature seasonality on resistance are negative and positive, respectively (Figure 3).

We used values of 0, 5, 10, 100, 1000 for intercept (α), and values of 0.01, 0.1, 0.5, 1, 5, 10, 100 for slope (γ) to create linear and non-linear resistance surfaces. Models where α is equal to zero (seven models for each landscape feature) are identical regardless of γ values, indicating no influence of resistance on genetic connectivity, which reduced the resistance surfaces for each dataset to 29 unique models. Values of $\gamma < 1$ represent resistance surfaces with increased sensitivity, $\gamma = 1$ represents a linear resistance relationship and $\gamma > 1$ are resistance surfaces with reduced sensitivity (Figure 2). We calculated pairwise resistance distance matrices for all landscape features using circuit theory (Hanks and Hooten 2013; McRae et al. 2008) as implemented in CIRCUITSCAPE v. 4.0.5 (McRae 2006). This approach identifies all possible pathways of movement between focal points across a given raster dataset and calculates average cumulative resistance between all pairwise sampling sites.

*Landscape genetic resistance modelling*

To evaluate the contribution of landscape features in explaining genetic differentiation, we fitted a Maximum-Likelihood Population-Effects (MLPE) mixed-effects model as implemented within the *mlpe_rga* function using the *R* package *ResistanceGA* v. 4.0-4 (Peterman 2018). This model uses individual pairwise metrics for genetic differentiation and landscape resistance, considering each pairwise data point as an observation. The lack of independence is incorporated as a population-level factor which distinguishes

between data points that share a common deme, and those that do not (Clarke et al. 2002; Row et al. 2017). Individual based pairwise genetic distance was measured as $F_{ST}/(1-F_{ST})$ and used as the dependent variable, resistance distance as the independent variable, and population as the random variable. We fitted the mixed-effects models using parameterization to account for the non-independence of values within pairwise distance matrices without restricting maximum-likelihood (Clarke et al. 2002; Van Strien et al. 2012). Next, to identify which model best described genetic distance among sites, we performed a model selection approach using Akaike Information Criteria (AICc). We then calculated the difference between the AIC of each model and the minimum AIC value found (Burnham and Anderson 2002; Diniz-Filho et al. 2008) with the lowest change in AICc score ($\Delta$AICc=0) and the largest AIC weight (wAICc) considered the most parsimonious model. These analyses were performed using the *R* package *ResistanceGA* v. 4.0-4 (Peterman 2018), with MLPE models fitted with *mlpe_rga* using the standard *lme4* v. 1.1-17 formula interface (Clarke et al. 2002; Bates et al. 2015), *magrittr* v. 1.5 (Bache and Wickham 2014), and *dplyr* v. 0.7.4 (Wickham et al. 2017).

*Effects of IBD and IBR on genetic differentiation*

We used a Mantel test (Mantel 1967) to estimate the significance of any relationship between pairwise $F_{ST}$ and geographic distance (km) using the function *mantel.randtest* implemented in the *ade4* v. 1.7-11 R-package (Dray and Dufour 2017), with 10 000 permutations. We also carried out an independent test for spatial autocorrelation between geographic and genetic distance using a Mantel correlogram (Oden and Sokal 1986), computed using the function *mantel.correlog* with 10 000 permutations. The number of geographic distance classes was selected by the Strurges equation, Pearson

correlation and correction of *p* values through FDR in the *R* package *vegan* v. 2.5.1 (Oksanen et al. 2018).

The effect of IBR decoupled from IBD was calculated using distance-based redundancy analysis (dbRDA) using *vegan* v. 2.5.1 (Oksanen et al. 2018). dbRDA is a direct extension of a multiple regression to model multivariate response data (Legendre and Gallagher 2001; Benestan et al. 2016), and was used to quantify the correlation between the best MLPE model for each landscape variable and $F_{ST}/(1-F_{ST})$, assuming models with genetic differentiation as the dependent variable and cost distances as independent variables, conditioned on IBD. We obtained statistical significance from each dbRDA model using Analyses of Variance (ANOVA; 1 000 permutations).

To verify that our limited sample size did not affected the MLPE and dbRDA inferences, we sub-sampled our data with three random individuals for each module, re-calculated $F_{ST}$ values, and correlated the complete and sub-sampled $F_{ST}$ matrices against each other. A correlation coefficient of 1.00 suggested that the sample sizes in the analyses provided reliable estimates.

## Results

### *$F_{ST}$ Outlier Analysis and Environmental Association Analysis*

Outlier analysis with BayeScan detected 174 SNPs with significantly high $F_{ST}$ (2.28%). The analysis with LFMM identified 1281, 912, 859 and 689 SNPs associated with land cover, the Walsh index, silt content and temperature seasonality, respectively. Of these, 43 SNPs were associated with each of the four environmental variables (Figure S1). Twenty-three outliers were in common for the BayeScan and LFMM analyses, none of which resulted in significant matches to either the *N. parkeri* genome or during BLAST searches using Genbank.

We removed the 23 loci in consensus between EAA and outlier approaches to produce an approximately neutral data set for population and landscape genetic analyses. Preliminary analyses indicated that inclusion or exclusion of these loci deviating from neutral expectations made no detectable difference to the results. Because of the strong genetic divergence of modules 1 and 2 from the remaining modules (see SNAPP analysis below), these two modules were excluded from the landscape genetic analyses to allow for subtle environmental influences on genetic structure to be detected. With the exclusion of the SNPs with signatures of selection and data from M1 and M2, a total of 7 609 SNPs were available for analysis. Summary statistics for modules M6-M14 are provided in Table 1.

*Population Tree*

The population tree constructed with SNAPP showed that individuals from the northern modules M1 and M2 (yellow femoral spot) belong to a strongly divergent lineage (Figure 4, Figure S2), consistent with the relatively high pairwise $F_{ST}$ values found between M1 or M2 and the other localities ($F_{ST}$ range 0.72-0.83). The remaining modules were split into three markedly shallower but distinct individual clades (posterior probability = 1.00 in all cases), with Cluster C formed by the most distal node (Figure 4).

Corresponding with the genetic lineages identified using SNAPP, the population genetic inferences with ADMIXTURE, sNMF and DAPC produced a congruent result of three inferred genetic clusters from Module 6 onward (Figure 5, see also Figure S3). The first Cluster A comprised 14 individuals with red femoral spots across modules M6-M8 in dense forest. It was distinct from a second Cluster B, which comprised 24 individuals from five populations (BM8_M9 - M11) across dense and open forest. This

cluster largely comprised individuals with yellow femoral spots, with the exception of population BM8_9 with an intermediate (orange) coloration and evidence of genetic admixture (Figure 5). A third cluster (C, characterized by red femoral spots) was confined to 16 individuals from the eastern bank of the upper Madeira river (M12 to M14), an open forest area separated from the remainder of the transect by patches of savannah. Reducing the dataset to three individuals for all modules did not alter the genetic partitioning revealed by each of the three clustering methods, demonstrating that the sampling regime was sufficient to resolve genetic structure (Supplementary Figure S4).

*Isolation by geographic distance (IBD) and environmental resistance (IBR)*

Pairwise genetic distances ($F_{ST}$) across modules M6 to M14 ranged from 0.020 (M13 and M14) to 0.207 (M6 and M14; Table 2), with a strong association between genetic and geographic distances and therefore IBD (Mantel test: $p < 0.0001$, $r^2 = 0.96$, Figure 6). The Mantel correlograms calculated for seven classes of geographic distance revealed spatial autocorrelation in four cases: positively at geographic distances to 60 km ($r = 0.67$, $p < 0.001$) and 143 km ($r = 0.24$, $p = 0.02$), and negatively at distances of 476 km ($r = -0.61$, $p = 0.03$) and 560 km ($r = -0.61$, $p < 0.001$; Figure S5).

Our MLPE analysis showed that a land cover model with $\alpha = 5$ and $\gamma = 10$ explained 98% of the genetic variation (Table 3). The Walsh index explained 96% of the genetic variation at $\alpha = 100$ and $\gamma = 5$, and temperature seasonality and silt content explained 95% of the genetic variation each, at $\alpha = 10$ and 1000, and at $\gamma = 5$ and 1, respectively (Table 3). The $\alpha$ values determine the maximum resistance of the variables (e.g., in the case of Walsh index, $\alpha = 100$ suggests that landscape resistance to gene flow is 100 times greater than zero), and the $\gamma$ values indicate whether the variable

426 influenced genetic connectivity linearly or non-linearly. Silt presented a value of $\gamma = 1$,
427 suggesting a linear resistance relationship. All other confidence sets of resistance
428 surfaces presented values $\gamma > 1$, supporting resistance surfaces with reduced resistance
429 sensitivity. ΔAIC values were identical for the four landscape features (0.00),
430 supporting the maximum-likelihood models. In the dbRDA models, the Walsh index
431 captured 8.4% of the observed genetic variation ($F_{1,52} = 41.72$; $p = 0.001$), followed by
432 land cover (6.4%; $F_{1,52} = 26.85$, $p = 0.001$), temperature seasonality (5.3%; $F_{1,52} = 20.54$,
433 $p = 0.001$) and silt content (3.5%; $F_{1,52} = 11.79$, $p = 0.001$; Table 3; Figure S6).
434
435 **Discussion**
436 We characterized patterns of genetic structure and femoral spot coloration for the
437 brilliant-thighed poison frog *A. femoralis* that was sampled along an 880 km transect
438 through continuous rainforest in a major Amazonian interfluve. We revealed four
439 genetically distinct clusters, one derived from a deep lineage divergence, and each
440 corresponding with femoral spot coloration that differed between individuals from
441 adjacent clusters. Transitions between major forest types were not consistently
442 associated with the boundaries of genetic clusters. Genetic variation was characterized
443 by a pattern of IBD across hundreds of kilometers, and subtle but significant effects of
444 contemporary landscape features on the distribution of individual measures of genetic
445 variation.
446 Under a pronounced pattern of IBD, as is the case for our study system, genetic
447 clustering algorithms can overestimate the number of partitions or lead to misleading
448 admixture inferences (Frantz et al. 2009; Garcia-Erill and Albrechtsen 2020). We
449 nevertheless argue that the clusters identified along our *A. femoralis* transect represent
450 biologically meaningful entities, as they were identified through four independent

451 approaches and conform with a phenotypic trait (femoral spot coloration). While precise
452 time calibrations are beyond the scope of the present study, the latter also suggests that
453 the clusters have arisen from past rather than contemporary phenomena, addressing the
454 'time lag problem' of landscape genetic inferences (see e.g. Epps and Keyghobadi
455 2015). That the DAPC approach failed to identify the zone of admixture is not overly
456 surprising, as it does not assess differential ancestry proportions for each individual (see
457 also Miller et al. 2020).

458 Possible taxonomic implications of the deeply diverged population of *A.*
459 *femoralis* from the northeast of the PMI (localities M1 and M2) will require further
460 work. Timing the divergence is needed to evaluate the role of historical processes in
461 isolating these localities from the remainder of the PMI. The northeast of the PMI is
462 well drained, of young sedimentary origin (Late Pleistocene-Early Holocene, see e.g.
463 Sombroek 2001) and due to the proximity to the Amazon river subject to rapid changes
464 in topography and hydrology that might have resulted in periods of isolation (Hoorn et
465 al. 2010; Latrubesse et al. 2010; Pupim et al. 2019). At present, the populations from
466 M1 and M2 are also separated from the remainder of the transect by approximately 150
467 km of lowland dense forest unoccupied by *A. femoralis* (Ferreira et al. 2018). Isolation
468 by unsuitable habitat is also suggested for cluster C (M12-M14, red femoral spots),
469 which is separated from the remainder of the modules by secondary vegetation,
470 including intervening savannah over about 150 km along the transect, an ecological
471 barrier that is likely to have been further strengthened during the glacial periods in the
472 late Pleistocene (Cohen et al. 2014).

473 In contrast to the association of Clusters A and C, the area of contact between
474 Clusters A and B (M8-M9) does not occur at the location of a current ecotone. This
475 implies that the divergence of Clusters A and B might be linked to a barrier which is no

longer present. Our finding for *A. femoralis* contrasts with recent data on the genetic structure of a treefrog (Ortiz et al. 2018) and with plumage coloration in birds (De Abreu et al. 2018) along the same transect, which both reveal a zone of divergence spatially matching with the ecotone between open and closed forest (M10 and M11). For these species, it was concluded that present day environmental differences were responsible for the genetic partitioning.

Individuals in Cluster A possess different femoral spot coloration (red) from those in Cluster B (yellow), except in a relatively narrow (~100 km) zone of admixture where individuals possess orange femoral spots. This color transition mirrors a well-studied model hybrid zone system between the European red (fire)-bellied toad *Bombina bombina* and the yellow-bellied toad *B. variegate*, that form orange-bellied hybrids in parapatry (e.g. Szymura and Barton 1986, 1991). In this system, spatial separation through differential habitat preferences leads to a narrow zone of admixture despite the lack of pronounced postzygotic mating barriers (Vines et al. 2003). Another mechanism that can lead to narrow zones of admixture is sexual selection, and assortative mating in accordance with red or yellow femoral spot coloration has indeed been demonstrated with *A. femoralis* mate choice experiments (Ferreira et al. forthcoming). In addition, femoral spot coloration in *A. femoralis* also spatially varies in association with mimicry with syntopic toxic species (Amézquita et al. 2017), and evaluating locally co-occurring taxa to investigate such relationships may help shed light on the mechanisms underpinning the distribution of color variation at this locality.

Spatially structured transitions of coloration across an area of genetic admixture could serve as a mechanism to generate new phenotypes (Stelkens and Seehausen 2009; Sefc et al. 2017). In other poison dart frogs, hybridisation has indeed been shown to result in independent aposematic lineages and novel colors morphs (Medina et al. 2013;

Ebersbach et al. 2020). Examining the evolutionary history of admixed individuals with color variation across the wider distribution of *A. femoralis* in the Amazon basin will help establish the role of hybridization in generating this polymorphism. In addition, testing for assortative mating particularly for individuals possessing the orange phenotype and conditions allowing disassortative mating (e.g., low mate availability; Medina et al. 2013) will contribute towards a better understanding of the isolating processes involved.

Although contemporary environmental variation was not consistently associated with the four distinct genetic clusters we have described, genetic connectivity still varies with environmental conditions. Environmental variables have been shown to influence gene flow in other anurans. For example, IBR contributed an additional 10-20% in variation to models governed by IBD for the European common frog *Rana temporaria* (Van Buskirk and Jansen van Rensburg 2020). Given that this study was conducted in rugged, alpine terrain, such values are consistent with the environmental influence that we measured in a more gradually varying environment. We found that the influence of land cover was strongly supported by our MLPE models, confirming previous evidence that dense forest flooded by streams and overflowing rivers are not favorable habitats for *A. femoralis* (Ferreira et al. 2018). Our dbRDA analysis showed that the Walsh index was also associated with less connectivity. A possible explanation is that rainfall strongly determines the existence and persistence of water-filled ditches on the forest floor, a requirement for reproduction for many amphibians including *A. femoralis* (Menin et al. 2011; Ringler et al. 2015). Rainfall gradients and the two dominant forest phytophysiognomies in the PMI are autocorrelated, which likely explains the inconsistency with the highest ranking variable resolved with the MLPE and dbRDA analyses (forest cover versus Walsh Index, respectively). Open forests in the drier,

526 southwestern areas of the PMI are more seasonal and have lower stem densities and
527 higher tree mass compared to wetter, dense forest at northeastern parts (Sombroek 2001;
528 Cintra et al. 2013; Schietti et al. 2016).

529 Environmental variation also appears to impose different selective pressures
530 along the PMI, with environmental association analyses showing the largest number of
531 SNP loci associated with the Walsh index and forest cover. Further work with greater
532 SNP densities and a reference genome will contribute towards the identification of
533 genes under selection. Nonetheless, our existing results suggest that both the levels of
534 connectivity and differences in fitness that are associated with environmental variation
535 may contribute to the observed fine-scale patterns of genetic variation. We reduced the
536 risk of false positives in such inferences (see Hoban et al. 2016; Ahrens et al. 2018) by
537 considering only those loci which were identified by both BayeScan and LFMM.

538 Although strong IBD and environmental-based selection are conditions that may
539 lead to divergence in accordance with the gradient diversification hypothesis (Endler
540 1977), our data also suggest a role of historical processes in the generation of the
541 patterns of genetic divergence we describe for *A. femoralis*. In particular, the relatively
542 rapid restructuring of the Amazon region may give rise to conditions where historical
543 isolation and processes associated with secondary contact reduce the potential for
544 environmental gradients to strongly influence genetic and phenotypic variation. For
545 example, reinforcement by the development of reproductive character displacement
546 could potentially be a stronger influence on gene flow than the effects of environmental
547 gradients. Accumulating genetic data from additional species using the standardized
548 sampling system along the PMI provides a unique opportunity to look for traits (e.g.,
549 variation in mating cues) that predict whether current environmental transitions or

mechanisms associated with past landscapes generate diversity in areas of continuous habitat.

**Funding**

This study was funded by the Brazilian Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) and Fundação de Amparo à Pesquisa do Estado do Amazonas (FAPEAM) under grants conceded to AP Lima (CNPq: Programa Ciência sem Fronteiras process 401327/2012- 4; FAPEAM/CNPq/PRONEX process 586/10, Edital 003/2009 – number 137) and AS Ferreira received a PhD scholarship by CNPq (number 161883/2014-1) and by Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) through Programa de Doutorado-Sanduíche no Exterior (PDSE) (Finance Code 001, number 88881.133683/2016-01).

**Acknowledgments**

Specimens were collected under permit numbers 13.777 and 7836-1 (AP Lima) issued by the Instituto Chico Mendes de Conservação da Biodiversidade - ICMBio of the Ministry of Environment, Government of Brazil. We thank the Instituto Nacional de Pesquisas da Amazônia (INPA), the Programa de Pesquisa em Biodiversidade (PPBio/RAPELD), the Centro Integrado de Estudos da Biodiversidade Amazônica (INCT - CENBAM) and Santo Antônio Energia S.A. for logistical and institutional support. We thank Chris Barratt for providing R scripts, and Sonu Yadav and Alex Carey for helping with R scripts. We are grateful to all members of the Conservation Genetics Lab of Macquarie University for valuable discussion during early stages of this study. We are very grateful to the field assistants for their help in collecting

574 *Allobates femoralis* over the years, and to Bill Magnusson and Antoine Fouquet for
575 providing comments on the manuscript.

576

577 **Data Availability**

578 In accordance with the Journal of Heredity data archiving policy, we will have
579 submitted all the data and R scripts to Dryad.

580

581 **References**

582 Ahrens CW, Rymer PD, Stow A, Bragg J, Dollin S, Umbers KDL, Dudaniec RY. 2018.
583 The search for loci under selection: trends, biases and progress. *Mol Ecol*. 27:1342–
584 1356.

585

586 Alexander DH, Novembre J, Lange K. 2009. Fast model-based estimation of ancestry in
587 unrelated individuals. *Genet Res*. 19:1655–1664.

588

589 Alvares CA, Stape JL, Sentelhas PC, de Moraes Gonçalves JL, Sparovek G. 2013.
590 Köppen's climate classification map for Brazil. *Meteorol Z.* 22:711–728.

591

592 Amaral S, Costa CB, Arasato LS, Ximenes AC, Rennó CD. 2013. AMBDATA:
593 variáveis ambientais para modelos de distribuição de espécies (MDEs). *Anais do XVI*
594 *Simpósio Brasileiro de Sensoriamento Remoto* (SBSR). 16:6930–6937.

595

596 Amézquita A, Lima AP, Jehle R, Castellanos L, Ramos O, Crawford AJ, Gasser H,
597 Hödl W. 2009. Calls, colours, shapes, and genes: a multi-trait approach to the study of

598 geographic variation in the Amazonian frog *Allobates femoralis*. *Biol J Linn Soc.*
599 98:826–838.
600
601 Amézquita A, Ramos O, González MC, Rodríguez C, Medina I, Simões PI, Lima AP.
602 2017. Conspicuousness, color resemblance, and toxicity in geographically diverging
603 mimicry: The pan-Amazonian frog *Allobates femoralis*. *Evolution*. 71:1039–1050.
604
605 Anderson CD, Epperson BK, Fortin M-J, Holderegger R, James P, Rosenberg MS,
606 Scribner KT, Spear S. 2010. Considering spatial and temporal scale in landscape-
607 genetic studies of gene flow. *Mol Ecol*. 19:3565–3575.
608
609 Armansin NC, Stow AJ, Cantor M, Leu ST, Klarevas-Irby JA, Chariton AA, Farine DR.
610 2020. Social barriers in ecological landscapes: The social resistance hypothesis. *Trends*
611 *Ecol Evol.* 35:137–148.
612
613 Bache SM, Wickham H. 2014. magrittr: A Forward-Pipe Operator for R. R package
614 version 1.5. Available from: https://CRAN.R-project.org/package=magrittr
615
616 Balkenhol N, Dudaniec RY, Krutovsky KV, Johnson JS, Cairns DM, Segelbacher G,
617 Selkoe KA, von der Heyden S, Wang IJ, Selmoni O, Joost S. 2017. Landscape
618 genomics: Understanding relationships between environmental heterogeneity and
619 genomic characteristics of populations: Springer, Heidelberg.
620
621 Bates D, Maechler M, Bolker B, Walker S. 2015. Fitting linear mixed-effects models
622 using lme4. *J Stat Softw.* 67:1–48.

623

624 Beheregaray LB, Cooke GM, Chao NL, Landguth EL. 2015. Ecological speciation in
625 the tropics: insights from comparative genetic studies in Amazonia. *Front Genet*. 5:477.
626

627 Benestan L, Quinn BK, Maaroufi H, Laporte M, Clark FK, Greenwood SJ, Rochette R,
628 Bernatchez L. 2016. Seascape genomics provides evidence for thermal adaptation and
629 current-mediated population structure in American lobster (*Homarus americanus*). *Mol*
630 *Ecol*. 25:5073–5092.
631

632 Bouckaert R, Heled J, Kühnert D, Vaughan T, Wu CH, Xie D, Suchard MA, Rambaut
633 A, Drummond AJ. 2014. BEAST 2: A software platform for Bayesian evolutionary
634 analysis. *PLoS Comput Biol*. 10:e1003537.
635

636 Bryant D, Bouckaert R, Felsenstein J, Rosenberg NA, Roychoudhury A. 2012. Inferring
637 species trees directly from biallelic genetic markers: Bypassing gene trees in a full
638 coalescent analysis. *Mol Biol Evol*. 29:1917–1932.
639

640 Burnham K, Anderson D. 2002. Model selection and multi-model inference: A practical
641 information theoretic approach. 2nd edition, Springer-Verlag, New York, NY.
642

643 Chang CC, Chow CC, Tellier LCAM, Vattikuti S, Purcell SM, Lee JJ. 2015. Second-
644 generation PLINK: rising to the challenge of larger and richer datasets. *GigaScience*.
645 25:4–7.
646

647 Cintra BBL, Schietti J, Emillio T, Martins D, Moulatlet G, Souza P, Levis C, Quesada
648 CA, Schöngart J. 2013. Soil physical restrictions and hydrology regulate stand age and
649 wood biomass turnover rates of Purus–Madeira interfluvial wetlands in Amazonia.
650 *Biogeosciences*.10:7759–7774.
651
652 Clarke RT, Rothery P, Raybould AF. 2002. Confidence limits for regression
653 relationships between distance matrices: Estimating gene flow with distance. *J Agr Biol*
654 *Envir St*. 7:361–372.
655
656 Cohen MCL, Rossetti DF, Pessenda LCR, Friaes YS, Oliveira PE. 2014. Late
657 Pleistocene glacial forest of Humaitá–western Amazonia. *Palaeogeogr Palaeoclimatol*
658 *Palaeoecol*. 415:37–47.
659
660 Coyne JA, Orr HA. 2004. Speciation. Sinauer Associates, Inc., Sunderland.
661
662 De Abreu FHT, Schietti J, Anciães M. 2018. Spatial and environmental correlates of
663 intraspecific morphological variation in three species of passerine birds from the Purus–
664 Madeira interfluvium, Central Amazonia. *Evol Ecol*. 32:191–214.
665
666 Dean LL, Magalhaes IS, Foote A, D'Agostino D, McGowan S, MacColl ADC. 2019.
667 Admixture between ancient lineages, selection, and the formation of sympatric
668 stickleback species-pairs. *Mol Biol Evol*. 36:2481–2497.
669
670 Diniz-Filho JAF, Rangel TFLVB, Bini LM. 2008. Model selection and information
671 theory in geographical ecology. *Global Ecol Biogeogr*. 17:479–488.

Dray S, Dufour AB. 2007. The ade4 package: implementing the duality diagram for ecologists. *J Stat Softw*. 22:1–20.

Dudaniec RY, Rhodes JR, Wilmer JW, Lyons M, Lee KE, Mcalpine CA, Carrick FN. 2013. Using multilevel models to identify drivers of landscape-genetic structure among management areas. *Mol Ecol*. 22:3752–3765.

Dudaniec RY, Wilmer JW, Hanson JO, Warren M, Bell S, Rhodes JR. 2016. Dealing with uncertainty in landscape genetic resistance models: a case of three co-occurring marsupials. *Mol Ecol*, 25:470–486.

Ebersbach J, Posso-Terranova A, Bogdanowicz S, Gómez-Díaz M, García-González M X, Bolívar-García W, Andrés J. 2020. Complex patterns of differentiation and gene flow underly the divergence of aposematic phenotypes in *Oophaga* poison frogs. *Mol Ecol*. 29:1944–1956.

Elshire RJ, Glaubitz JC, Sun Q, Poland JA, Kawamoto K, Buckler ES, Mitchell SE. 2011. A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PLoS One*. 6(5):e19379.

Endler JA. 1977. Geographic variation, speciation, and clines: Princeton University Princeton.

696 Epps CW, Keyghobadi N. 2015. Landscape genetics in a changing world: disentangling
697 historical and contemporary influences and inferring change. *Mol Ecol.* 24:6021–6040.
698
699 Fan Y, Miguez-Macho G. 2010. Potential groundwater contribution to Amazon
700 evapotranspiration. *Hydrol Earth Syst Sci*. 14:2039–2056.
701
702 Ferreira AS, Jehle R, Stow AJ, Lima AP. 2018. Soil and forest structure predicts large-
703 scale patterns of occurrence and local abundance of a widespread Amazonian frog.
704 *PeerJ*. 6:e5424.
705
706 Fick SE, Hijmans RJ. 2017. Worldclim 2: New 1-km spatial resolution climate surfaces
707 for global land areas. *Int J Climatol*. 37:4302–4315.
708
709 Foll M, Gaggiotti O. 2008. A genome-scan method to identify selected loci appropriate
710 for both dominant and codominant markers: a Bayesian perspective. *Genetics*. 180:977–
711 993.
712
713 Fouquet A, Gilles A, Vences M, Marty C, Blanc M, Gemmell NJ. 2007.
714 Underestimation of species richness in Neotropical frogs revealed by mtDNA analyses.
715 *PLoS One*. 2(10):e1109.
716
717 Frantz AC, Cellina S, Krier A, Schley L, Burke T. 2009. Using spatial Bayesian
718 methods to determine the genetic structure of a continuously distributed population:
719 clusters or isolation by distance? *J Appl Ecol*. 46:493–505
720

721 Frichot E, François O. 2015. LEA: an R package for landscape and ecological
722 association studies. *Methods Ecol Evol*. 6:925–929.
723
724 Frichot E, Mathieu F, Trouillon T, Bouchard G, François O. 2014. Fast and efficient
725 estimation of individual ancestry coefficients. *Genetics*. 196:973–983.
726
727 Garcia-Erill G, Albrechtsen A. 2020. Evaluation of model fit of inferred admixture
728 proportions. *Mol Ecol Resour*. 20:936–949.
729
730 Gosselin T. 2017. radiator: RADseq data exploration, manipulation and visualization
731 using R. R package version 0.0.5. Available from:
732 https://CRAN.Rproject.org/package=radiator
733
734 Grant T, Frost DR, Caldwell JP, Gagliardo R, Haddad CFB, Kok PJR, Means DB,
735 Noonan BP, Schargel WE, Wheeler W. 2006. Phylogenetic systematics of dart poison
736 frogs and their relatives (Anura: Athesphatanura: Dendrobatidae). *AMNH Res Library*.
737 299:1–262.
738
739 Grant T, Rada M, Anganoy-Criollo M, Batista A, Dias PH, Jeckel AM, Machado DJ,
740 Rueda-Almonacid JV. 2017. Phylogenetic systematics of dart-poison frogs and their
741 relatives revisited (Anura: Dendrobatoidea). *S Am J Herpetol*. 12:1–90.
742
743 Hanks EM, Hooten MB. 2013. Circuit theory and model-based inference for landscape
744 connectivity. *J Am Stat Assoc*. 108:22–33.
745

746 Hoban S, Kelley JL, Lotterhos KE, Antolin MF, Bradburd G, Lowry DB, Poss ML,
747 Reed LK, Storfer A, Whitlock MC. 2016. Finding the genomic basis of local
748 adaptation: Pitfalls, practical solutions, and future directions. *Am Nat*. 188:379–397.
749
750 Hijmans RJ. 2017. raster: Geographic Data Analysis and Modeling. R package version
751 2.6-7. Available from: https://CRAN.R-project.org/package=raster
752
753 Hoorn C, Wesselingh FP, Steege Hter, Bermudez MA, Mora A, Sevink J, Sanmartín I,
754 Sanchez-Meseguer A, Anderson CL, Figueiredo JP, *et al*. 2010. Amazonia through
755 time: Andean uplift, climate change, landscape evolution, and biodiversity. *Science*.
756 330:927–931.
757
758 IBGE (1997). Recursos naturais e meio ambiente: uma visão do Brasil. Second Edition.
759 Rio de Janeiro: Instituto Brasileiro de Geografia e Estatística (IBGE).
760
761 Jenkins DG, Carey M, Czerniewska J, Fletcher J, Hether T, Jones A, Knight S, Knox J,
762 Long T, Mannino M, *et al*. 2010. A meta-analysis of isolation by distance: relic or
763 reference standard for landscape genetics? *Ecography*. 33:315–320.
764
765 Johnson M, Zaretskaya I, Raytselis Y, Merezhuk Y, McGinnis S, Madden TL. 2008.
766 NCBI BLAST: a better web interface. *Nucleic Acids Res*, 36:5–9.
767
768 Jombart T, Ahmed I. 2011. adegenet 1.3-1: new tools for the analysis of genome-wide
769 SNP data. *Bioinformatics*. 27:3070–3071.
770

771 Jombart T, Devillard S, Balloux F. 2010. Discriminant analysis of principal
772 components: a new method for the analysis of genetically structured populations. *BMC*
773 *Genetics*. 11:94.
774
775 Kaefer IL, Montanarin A, Costa RS, Lima AP. 2012. Temporal patterns of reproductive
776 activity and site attachment of the brilliant-thighed frog *Allobates femoralis* from
777 Central Amazonia. *J Herpetol*. 46:549–554.
778
779 Keenan K, McGinnity P, Cross TF, Crozier WW, Prodöhl PA. 2013. diveRsity: An R
780 package for the estimation of population genetics parameters and their associated errors.
781 *Methods Ecol Evol*. 4:782–788.
782
783 Kilian A, Wenzl P, Huttner E, Carling J, Xia L, Blois H, Caig V, Heller-Uszynska K,
784 Jaccoud D, Hopper C, *et al.* 2012. Diversity Arrays Technology: A generic genome
785 profiling technology on open platforms. In: Pompanon F, Bonin A (eds). Data
786 production and analysis in population genomics. Methods in molecular biology
787 (Methods and Protocols), vol. 888. Humana Press, Totowa, NJ.
788
789 Latrubesse EM, Cozzuol M, Silva-Caminha SAF, Rigsby CA, Absy MA, Jaramillo C.
790 2010. The late Miocene paleogeography of the Amazon basin and the evolution of the
791 Amazon River system. *Earth-Sci Rev*. 99:99e124.
792
793 Legendre P, Gallagher ED. 2001. Ecologically meaningful transformations for
794 ordination of species data. *Oecologia*. 129:271–280.
795

796 Leite RN, Rogers DS. 2013. Revisiting Amazonian phylogeography: insights into
797 diversification hypotheses and novel perspectives. *Org Divers Evol*. 13:639–664.
798
799 Lemay MA, Russello MA. 2015. Genetic evidence for ecological divergence in kokanee
800 salmon. *Mol Ecol*. 24:798–811.
801
802 Lischer HEL, Excoffier L. 2012. PGDSpider: An automated data conversion tool for
803 connecting population genetics and genomics programs. *Bioinformatics*. 28:298–299.
804
805 Luu K, Bazin E, Blum MGB. 2017. pcadapt: an R package for performing genome
806 scans for selection based on principal component analysis. *Mol Ecol Resour*. 17:67–77.
807
808 Magnusson WE, Braga-Neto R, Pezzini F, Baccaro F, Bergallo H, Penha J, Rodrigues
809 D, Verdade LM, Lima A, Albernaz AL, *et al*. 2013. *Biodiversity and integrated*
810 *environmental monitoring*. Manaus: Áttema. p. 356.
811
812 Manel S, Schwartz MK, Luikart G, Taberlet P. 2003. Landscape genetics: combining
813 landscape ecology and population genetics. *Trends Ecol Evol*. 18:189–197.
814
815 Mantel N. 1967. The detection of disease clustering and a generalized regression
816 approach. *Cancer Res*. 27:209–220.
817
818 Marques DA, Lucek K, Sousa VC, Excoffier L, Seehausen O. 2019. Admixture
819 between old lineages facilitated contemporary ecological speciation in Lake Constance
820 stickleback. *Nat Commun*. 10:4240.

Mayr E. 1963. Animal Species and Evolution. Harvard University Press, London.

Medina I, Wang IJ, Salazar C, Amézquita A. 2013. Hybridization promotes color polymorphism in the aposematic harlequin poison frog, *Oophaga histrionica*. *Ecol Evol*. 3:4388–4400.

Menin M, Waldez F, Lima AP. 2011. Effects of environmental and spatial factors on the distribution of anuran species with aquatic reproduction in central Amazonia. *Herpetol J*. 21:255–261.

McRae BH, Dickson BG, Keitt TH, Shah VB. 2008. Using circuit theory to model connectivity in ecology, evolution, and conservation. *Ecology*. 89:2712–2724.

McRae BH. 2006. Isolation by resistance. *Evolution*. 60:1551–1561.

Miller JM, Cullingham CI, Peery RM. 2020. The influence of a priori grouping on inference of genetic clusters: simulation study and literature review of the DAPC method. *Heredity.* early online.

Montanarin A, Kaefer IL, Lima AP. 2011. Courtship and mating behaviour of the brilliant-thighed frog *Allobates femoralis* from Central Amazonia: implications for the study of a species complex. *Ethol Ecol Evol*. 23:141–150.

Moritz C, Patton JL, Schneider CJ, Smith TB. 2000. Diversification of rainforest

846 faunas: an integrated molecular approach. *Ann Rev Ecol S.* 31:533–563.
847
848 Naka LN, Bechtoldt CL, Henriques LMP, Brumfield RT, Heard AESB, McPeek EMA.
849 2012. The role of physical barriers in the location of avian suture zones in the Guiana
850 Shield, northern Amazonia. *Am Nat*. 179:E115–E132.
851
852 Nazareno AG, Dick CW, Lohmann LG. 2017. Wide but not impermeable: testing the
853 riverine barrier hypothesis for an Amazonian plant species. *Mol Ecol*. 26:3636–3648.
854
855 Nosil P. 2012. Ecological speciation: Oxford University Press, Oxford.
856
857 Oden NL, Sokal RR. 1986. Directional autocorrelation: an extension of spatial
858 correlograms to two dimensions. *Syst Zool*. 35:608–617.
859
860 Oksanen J, Blanchet FG, Friendly M, Kindt R, Legendre P, McGlinn D, Minchin PR,
861 O'Har RB, Simpson GL, Solymos P, *et al*. 2018. vegan: Community Ecology Package.
862 R package version 2.5-1. Available from: https://CRAN.R-project.org/package=vegan
863
864 Ortiz DA, Lima AP, Werneck FP. 2018. Environmental transition zone and rivers shape
865 intraspecific population structure and genetic diversity of an Amazonian rain forest tree
866 frog. *Evol Ecol*. 32:359–378.
867
868 Pabijan M, Palomar G, Antunes B, Antoł W, Zieliński P, Babik W. 2020. Evolutionary
869 principles guiding amphibian conservation. *Evol Appl*. 13:857–878.
870

871 Pašukonis A, Trenkwalder K, Ringler M, Ringler E, Mangione R, Steininger J,
872 Warrington I, Hödl W. 2016. The significance of spatial memory for water finding in a
873 tadpole-transporting frog. *Anim Behav*. 116:89–98.
874
875 Peterman WE. 2018. ResistanceGA: An R package for the optimization of resistance
876 surfaces using genetic algorithms. *Methods Ecol Evol*. 9:1638–1647.
877
878 Pupim FN, Sawakuchi AO, Almeida RP, Ribas CC, Kern AK, Hartmann GA, Chiessi
879 CM, Tamura LN, Mineli TD, Savian JF, *et al*. 2019. Chronology of Terra Firme
880 formation in Amazonian lowlands reveals a dynamic Quaternary landscape. *Quat Sci*
881 *Rev*. 210:154–163.
882
883 Rellstab C, Gugerli F, Eckert AJ, Hancock AM, Holderegger R. 2015. A practical guide
884 to environmental association analysis in landscape genomics. *Mol Ecol*. 24:4348–4370.
885
886 Ribas CC, Aleixo A, Nogueira ACR, Miyaki CY, Cracraft J. 2012. A
887 palaeobiogeographic model for biotic diversification within Amazonia over the past
888 three million years. *P Roy Soc B-Biol Sci*. 279:681–689.
889
890 Ribas CC, Aleixo A, Gubili C, d'Horta F, Brumfield RT, Cracraft J. 2018.
891 Biogeography and diversification of Rhegmatorhina (Aves: Thamnophilidae):
892 implications for the evolution of Amazonian landscapes during the Quaternary. *J*
893 *Biogeogr*. 45:917–928.
894

895 Ringler E, Pašukonis A, Hödl W, Ringler M. 2013. Tadpole transport logistics in a
896 Neotropical poison frog: indications for strategic planning and adaptive plasticity in
897 anuran parental care. *Front Zool*. 10:67.
898
899 Ringler M, Ursprung E, Hödl W. 2009. Site fidelity and patterns of short-and long-term
900 movement in the brilliant-thighed poison frog *Allobates femoralis* (Aromobatidae).
901 *Behav Ecol Sociobiol*. 3:1281–1293.
902
903 Ringler M, Hödl W, Ringler E. 2015. Populations, pools, and peccaries: simulating the
904 impact of ecosystem engineers on rainforest frogs. *Behav Ecol*. 26:340–349.
905
906 Roithmair ME. 1994. Field studies on reproductive behaviour in two dart-poison frog
907 species (*Epipedobates femoralis*, *Epipedobates trivittatus*) in Amazonian Peru. *Herpetol*
908 *J*. 4:77–85.
909
910 Row JR, Knick ST, Oyler-McCance SJ, Lougheed SC, Fedy BC. 2017. Developing
911 approaches for linear mixed models in genetics through landscape-directed dispersal
912 simulations. *Ecol Evol*. 7:3751–3761.
913
914 Santos JC, Coloma LA, Summers K, Caldwell JP, Ree R, Cannatella DC. 2009.
915 Amazonian amphibian diversity is primarily derived from Late Miocene Andean
916 lineages. *PLoS Biol*. 7:e1000056.
917
918 Schietti J, Martins D, Emilio T, Souza PF, Levis C, Baccaro FB, Pinto JLPV, Moulatlet
919 GM, Stark SC, Sarmento, K, *et al*. 2016. Forest structure along a 600 km transect of

920 natural disturbances and seasonality gradients in central-southern Amazonia. *J Ecol.*
921 104:1335–1346.
922
923 Sefc KM, Mattersdorfer K, Ziegelbecker A, Neuhüttler N, Steiner O, Goessler W,
924 Koblmüller S. 2017. Shifting barriers and phenotypic diversification by hybridization.
925 *Ecol Lett.* 20:651–662.
926
927 Sexton JP, Hangartner SB, Hoffmann AA. 2014. Genetic isolation by
928 environment or distance: which pattern of gene flow is most common? *Evolution*. 68:1–
929 15.
930
931 Shafer ABA, Wolf JBW. 2013. Widespread evidence for incipient ecological
932 speciation: a meta-analysis of isolation-by-ecology. *Ecol Lett.* 16:940–950.
933
934 Shirk AJ, Wallin DO, Cushman SA, Rice CG, Warheit KI. 2010. Inferring landscape
935 effects on gene flow: a new model selection framework. *Mol Ecol*. 19:3603–3619.
936
937 Silverstone P. 1975. A revision of the poison-arrow frogs of the genus *Dendrobates*
938 Wagler. *Nat Hist Mus Los Ang Cty Sci Bull*. 21:1–55.
939
940 Simões PI, Lima AP, Farias IP. 2010. The description of a cryptic species related to the
941 pan Amazonian frog *Allobates femoralis* (Boulenger 1883) (Anura: Aromobatidae).
942 *Zootaxa*. 2406:1–28.
943

944 Slatkin M. 1987. Gene flow and the geographic structure of natural populations.
945 Science. 236:787.
946
947 Sombroek W. 2001. Spatial and temporal patterns of Amazon rainfall. AMBIO: *J Hum*
948 *Environ*. 30:388–396.
949
950 Stelkens RB, Seehausen O. 2009. Genetic distance between species predicts novel trait
951 expression in their hybrids. *Evolution*. 63:884–897.
952
953 Storfer A, Murphy MA, Spear SF, Holderegger R, Lisette P, Waits LP. 2010.
954 Landscape genetics: where are we now? *Mol Ecol.* 19:3496–3514.
955
956 Stow AJ, Sunnucks P, Briscoe DA, Gardner MG. 2001. The impact of habitat
957 fragmentation on dispersal of Cunningham's skink (*Egernia cunninghami*): evidence
958 from allelic and genotypic analyses of microsatellites. *Mol Ecol*. 10:867–878
959
960 Sun Y-B, Xiong Z-J, Xiang X-Y, Liu S-P, Zhou W-W, Tu X-L, Zhong L, Wang L, Wu
961 D-D, Zhang B-L, *et al*. 2015. Whole-genome sequence of the Tibetan frog *Nanorana*
962 *parkeri* and the comparative evolution of tetrapod genomes. *PNAS USA*. 112:E1257–
963 E1262.
964
965 Szymura JM, Barton N. 1986. Genetic analyses of a hybrid zone between the fire-
966 bellied toads *Bombina bombina* and *B. variegate*, near Cracow in southern Poland.
967 *Evolution*. 40:1141–1159.
968

969 Szymura JM, Barton N. 1991. The genetic structure of the hybrid zone between the fire-
970 bellied toads *Bombina bombina* and *B. variegata*: comparisons between transects and
971 between loci. *Evolution*. 45:237–261.
972
973 Thom G, Xue AT, Sawakuchi AO, Ribas CC, Hickerson MJ, Aleixo A, Miyaki C.
974 2020. Quarternary climate changes as speciation drivers in the Amazonian floodplains.
975 *Sci Adv*. 6:eaax4718.
976
977 Ursprung E, Ringler M, Jehle R, Hödl W. 2011. Strong male/male competition allows
978 for nonchoosy females: high levels of polygynandry in a territorial frog with paternal
979 care. *Mol Ecol*. 20:1759–71.
980
981 Van Buskirk J, Jansen van Rensburg A. 2020. Relative importance of isolation-by-
982 environment and other determinants of gene flow in an alpine amphibian. *Evolution*.
983 74:962–978.
984
985 Van Strien MJ, Keller D, Holderegger R. 2012. A new analytical approach to landscape
986 genetic modelling: least-cost transect analysis and linear mixed models. *Mol Ecol*.
987 21:4010–4023.
988
989 Villemereuil P, Frichot É, Bazin É, Olivier F, Gaggiotti OE. 2014. Genome scan
990 methods against more complex models: when and how much should we trust them? *Mol*
991 *Ecol*. 23:2006–2019.
992

993 Vines TH, Kohler SC, Thiel M, Ghira I, Sands TR, MacCallum CJ, Barton NH,
994 Nürnberger B. 2003. The maintenance of reproductive isolation in a mosaic hybrid zone
995 between the fire-bellied toads *Bombina bombina* and *B. variegata*. *Evolution.* 57:1876–
996 1888.
997
998 Wallace AR. 1852. On the monkeys of the Amazon. *Proc Zool Soc Lond*. 20:107–110.
999
1000 Walsh RPD. 1996. The climate. In: Richards PW. (ed). The Tropical Rain Forest: an
1001 ecological study. Cambridge University Press.
1002
1003 Wickham H, Francois R, Henry L, Müller K. 2017. dplyr: A Grammar of Data
1004 Manipulation. R package version 0.7.4. Available from: https://CRAN.R-
1005 project.org/package=dplyr
1006
1007 Wright S. 1943. Isolation by distance. *Genetics*. 28:114.
1008
1009 Yadav S, Stow AJ, Dudaniec RY. 2019. Detection of environmental and morphological
1010 adaptation despite high landscape genetic connectivity in a pest grasshopper
1011 (*Phaulacridium vittatum*). *Mol Ecol*. 28:3395–3412.
1012
1013 Zeisset I, Beebee TJC. 2008. Amphibian phylogeography: a model for understanding
1014 historical aspects of species distributions. *Heredity*. 101:109–119.
1015

Zheng X, Levine D, Shen J, Gogarten SM, Laurie C, Weir BS. 2012. A high-performance computing toolset for relatedness and principal component analysis of SNP data. *Bioinformatics*. 28:3326–3328.

**Table and figure captions**

**Table 1**. Number of sampled individuals ($N_{TOTAL}$) and summary genetic data at each sampling site for *Allobates femoralis* along the Purus-Madeira interfluve in central-southern Amazonia. Heterozygosity ($H_O$), expected heterozygosity ($H_E$), inbreeding coefficient ($F_{IS}$) and their low and high values (95%), number of private alleles (PA) and probability of deviating from Hardy–Weinberg equilibrium (HWE) are provided.

**Table 2**. Pairwise genetic distances $F_{ST}$ (below diagonal) and geographic distance (in Km) between *Allobates femoralis* sampling locations (above diagonal) within the Purus-Madeira interfluve.

**Table 3**. Summary of model selection using MLPE and dbRDA that evaluated the effects of isolation by resistance (IBR) on genetic distance ($\log(F_{ST}/1-F_{ST})$). For MLPE, the Akaike Information Criteria (AIC), $r^2$ value, standard error (SE) and the parameter combination ($\alpha$ and $\gamma$) is given for the best models for each landscape variable. For dbRDA the magnitude of difference is given by the *t*-value and the F and *p* values were obtained by ANOVA. Bolded *p* values show significant effects of IBR on genetic distance.

**Figure 1**. The distribution of modules from which samples of *Allobates femoralis* were collected in the Purus-Madeira interfluve, central-southern Amazonia, Brazil. White circles indicate absence of *A. femoralis*. For sample sizes at each module see Table 1. See online version for full colors.

**Figure 2**. Rasters capturing each of the four environmental variables used in CIRCUITSCAPE to generate resistance distance matrices between each pair of

sampling locations a) land cover, b) silt content, c) temperature seasonality - Bio4 and d) Walsh index. See online version for full colors.

**Figure 3.** The isolation-by-resistance (IBR) relationships tested for the effect of land cover and temperature seasonality on genetic distance $F_{ST}/(1-F_{ST})$ using seven values of $\gamma$ (0.01, 0.1, 0.5, 1, 5, 10, 100). The different slopes are not shown ($\alpha$ values) and are displayed here for $\alpha = 5$ here for simplicity. The curves show decreasing landscape resistances from right to left for land cover (A) and left to right for temperature seasonality (B).

**Figure 4**. A population tree generated using SNAPP, and a histogram showing individual ancestry proportions color coded to correspond to each genetic cluster, estimated using ADMIXTURE. The location of the collection modules are color coded to reflect the color assigned to each genetic cluster in the ADMIXTURE plot (the white circles for M3-M5 indicate the absence of *A. femoralis*). Posterior probabilities obtained at each node are shown on the tree. Cluster 1 corresponds to individuals with yellow femoral spots, Cluster A corresponds to individuals with red femoral spots, Cluster B corresponds to individuals with yellow femoral spots, with a zone of admixture between Cluster A-B (BM8-9) with an intermediate color phenotype (orange), and Cluster C corresponds to individuals with red femoral spot. See online version for full colors.

**Figure 5**. Histograms for individual *A. femoralis* sampled along the Purus-Madeira interfluve using three different clustering approaches a) ADMIXURE, b) sNMF and c) DAPC. Each individual is represented by a bar partitioned into different colors to represent individual ancestry proportions. *K* represents the most likely number of genetic clusters. See online version for full colors.

**Figure 6**. Relationship between genetic and geographic distance in *A. femoralis* across the Purus-Madeira interfluve.