Received: 22 October 2016 | Revised: 13 March 2016 | Accepted: 21 March 2017

DOI: 10.1111/mec.14142

ORIGINAL ARTICLE

WILEY MOLECULAR ECOLOGY

Wide but not impermeable: Testing the riverine barrier hypothesis for an Amazonian plant species

Alison G. Nazareno[1] | Christopher W. Dick[2] | Lúcia G. Lohmann[1]

[1]Departamento de Botânica, Universidade de São Paulo, São Paulo, SP, Brazil

[2]Department of Ecology and Evolutionary Biology, University of Michigan, Ann Arbor, MI, USA

Correspondence
Alison G. Nazareno, Departamento de Botânica, Universidade de São Paulo, São Paulo, SP, Brazil.
Email: alison_nazareno@yahoo.com.br
and
Lúcia G. Lohmann, Departamento de Botânica, Universidade de São Paulo, São Paulo, SP, Brazil.
Email: llohmann@usp.br

Funding information
Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP), Grant/Award Number: 2013/12633-8; 2015/07141-4, 2012/50260-6; Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), Grant/Award Number: 307781/2013-5; National Science Foundation, Grant/Award Number: FESD 1338694, DEB 1240869

Abstract

Wallace's riverine barrier hypothesis postulates that large rivers, such as the Amazon and its tributaries, reduce or prevent gene flow between populations on opposite banks, leading to allopatry and areas of species endemism occupying interfluvial regions. Several studies have shown that two major tributaries, Rio Branco and Rio Negro, are important barriers to gene flow for birds, amphibians and primates. No botanical studies have considered the potential role of the Rio Branco as a barrier, while a single botanical study has evaluated the Rio Negro as a barrier. We studied an Amazon shrub, *Amphirrhox longifolia* (A. St.-Hil.) Spreng (Violaceae), as a model to test the riverine barrier hypothesis. Twenty-six populations of *A. longifolia* were sampled on both banks of the Rio Branco and Rio Negro in the core Amazon Basin. Double-digest RADseq was used to identify 8,010 unlinked SNP markers from the nuclear genome of 156 individuals. Data relating to population structure support the hypothesis that the Rio Negro acted as a significant genetic barrier for *A. longifolia*. On the other hand, no genetic differentiation was detected among populations spanning the narrower Rio Branco, which is a tributary of the Rio Negro. This study shows that the strength of riverine barriers for Amazon plants is dependent on the width of the river separating populations and species-specific dispersal traits. Future studies of plants with contrasting life history traits will further improve our understanding of the landscape genetics and allopatric speciation history of Amazon plant diversity.

KEYWORDS
Amazon River, *Amphirrhox longifolia*, double-digest RADseq, population tree, single nucleotide polymorphism, Violaceae

1 | INTRODUCTION

The Amazon River and its ca. 10,000 tributaries have been the focus of several biogeographical studies due to their potential role in population isolation (Ayres & Clutton-Brock, 1992; Bates, Haffer, & Grismer, 2004; Boubli et al., 2015; Collevatti, Leoi, Leite, & Gribel, 2009; Fernandes, Wink, Sardelli, & Aleixo, 2014; Fernandes, Cohn-Haft, Hrbek, & Farias, 2014; Funk et al., 2007; Gascon et al., 2000; Hall & Harvey, 2002; Hayes & Sewlal, 2004; Link et al., 2015; Lougheed, Gascon, Jones, Bogart, & Boag, 1999; Maldonato-Coelho, Blake, Silveira, Batalha-Filho, & Ricklefs, 2013; Moraes, Pavan, Barros, & Ribas, 2016; Naka, Becchtoldt, Henriques, & Brumfield, 2012; Patton, da Silva, & Malcolm, 2000; Räsänen, Salo, Jungner, & Pittman, 1990; Ribas, Aleixo, Nogueira, Miyaki, & Cracraft, 2012; Salo et al., 1986; Smith et al., 2014; Solomon, Bacci, Martins, Vinha, & Mueller, 2008; Wallace, 1852). This complex river system is thought to have originated in the Miocene, some 11 million years ago, and took its present shape in the late Pliocene, approximately 2.4 million years ago (Baker et al., 2014). As such, the Amazon and its tributaries represent a potential and important cause of allopatric population differentiation for a plethora of taxa.

Wallace's (1852) riverine barrier hypothesis postulates that large rivers, such as the Amazon, reduce or prevent gene flow between populations on opposite banks, thereby explaining some biodiversity patterns currently found in this region. Several studies have shown that Rio Branco and Rio Negro tributaries are important barriers to gene flow for birds, amphibians and primates (Bates et al., 2004; Boubli et al., 2015; Naka et al. 2012; Räsänen et al., 1990; Ribas et al., 2012; Salo et al., 1986; Solomon et al., 2008). No botanical studies have considered the potential role of the Rio Branco as a barrier, and only one study to date has evaluated the Rio Negro as a barrier for plant dispersal (Collevatti et al., 2009). In this study, high rates of gene flow were found between populations of two species of *Caryocar* separated by the Rio Negro (Collevatti et al., 2009). The *Caryocar* species studied are enormous canopy-emergent trees pollinated by *Phyllostomus* bats and dispersed by strong swimming tapirs (*Tapirus terrestris*) (Collevatti et al., 2009). Extensive gene flow, even across rivers, may therefore be expected for *Caryocar*. On the other hand, plant species with more limited gene flow, such as insect-pollinated and terrestrially dispersed trees or shrubs, should exhibit very different response to riverine barriers. In addition to plant dispersal abilities, the physical traits (e.g., flow and dynamic) of the rivers are important factors in the establishment of putative genetic barriers.

The Neotropical plant species *Amphirrhox longifolia* (Violaceae) is a small, shrub-like treelet (up to 3 m in height) that is widely distributed in lowland forests from Costa Rica to eastern Brazil (Missouri Botanical Garden 2009). In the Amazon Basin, *A. longifolia* occurs at relatively high densities in blackwater floodplains (i.e., *igapó*, rich in humic substances, acidic and poor in nutrients), whitewater floodplains (i.e., *várzea*, rich in suspended matter and nutrients, with neutral pH), and nonflooding uplands (i.e., *terra firme*). *Amphirrhox longifolia* is least abundant in *terra firme*, where it occurs in the lowest stratum and is absent in the middle and upper strata (Parolin et al., 2004; Ribeiro et al., 1999). We selected *A. longifolia* as a model because of its abundance on both banks (left and right) of the Rio Negro and Rio Branco tributaries in the Amazon Basin. The species is self-incompatible and has hermaphroditic flowers that are largely actinomorphic, tubular and spurless (Braun, Dotter, Schlindwein, & Gottsberger, 2012), a floral syndrome associated with pollination by bees and butterflies (Braun et al., 2012). *Amphirrhox longifolia* has a telechoric ballistic seed dispersal mechanism. Because of its ballistic dispersal system, the small seeds of *A. longifolia* are likely to be dispersed short distances from the maternal tree. As the seeds do not float, a minimal role of water dispersal is expected. No previous evolutionary or genetic studies have been performed on this species.

Based on what is known about the ecology and life history of *A. longifolia*, we hypothesized that populations located on opposite margins of the Rio Negro and its tributary, the Rio Branco, would be more differentiated genetically than populations located on each side of these rivers, where gene flow is not impeded by any obvious barrier. We also expected that as the Rio Negro is generally wider (~1–24 km) than the Rio Branco (~0.5–3.5 km), genetic differentiation would be more pronounced in the Rio Negro than in the Rio Branco. In order to test these hypotheses, we used a high-throughput methodology (double-digest RADseq; Peterson, Weber, Kay, Fisher, & Hoekstra, 2012) to identify thousands of informative SNP markers from anonymous portions of the nuclear genome across all studied populations of *A. longifolia*. Every SNP identified through RADseq is likely to be independently inherited and thus represents an independent sample of the underlying evolutionary processes. With this study, we hoped to better understand the importance of Amazon tributaries as barriers for gene flow in a widespread rain forest plant species.

2 | MATERIAL AND METHODS

2.1 | Study area and sample collection

Located on the Guiana Shield ecoregion, the Rio Negro basin (754,925 km^2) and its largest tributary, the Rio Branco (235,073 km^2), were selected to test the riverine barrier hypothesis. The Rio Negro (Figure S1), a blackwater river spanning 1,700 km, is the fifth-largest river in the world and one of the largest tributaries of the Amazon Basin (Latrubesse, Stevaux, & Sinha, 2005). The geological history of this river system is complex, but geomorphological evidence suggests that the mouth of the Rio Negro moved 150 km eastward to its present location near the city of Manaus, Amazonas State, Brazil (Almeida-Filho & Miranda, 2007). The Rio Negro Basin is covered by lowland rain forest, submontane and montane forests and other vegetation types, including white sand-soil forests (Macedo & Prance, 1978). The highly diverse forests bordering the Rio Negro consist of several hundred species of trees (Ferreira, 2000; Keel & Prance, 1979). Deforestation along the Rio Negro has been minimal due in part to its nutrient-poor soils that are unsuitable for agriculture. The study area extends from above the mouth of the Rio Negro (Novo Airão, Amazonas State, Brazil) to 240 km up river, including the Rio Branco, the most important tributary of the Rio Negro. The Rio Branco (Figure S1) is 750 km long and is unique in the region because it is a sediment-rich "whitewater" river surrounded by clear blackwater rivers. Hence, the Rio Branco supports a different type of vegetation and plant community structure than is found along the shores of blackwater rivers (Worbes, 1997).

During the winter of 2015, flowering individuals of *A. longifolia* were sampled from 26 populations located on both banks (left and right) of the Rio Negro and Rio Branco (Table 1, Figure 1). All populations of *A. longifolia* along the Rio Negro were located on blackwater floodplains and sampled within 44.1 km on the same riverbank. On the Rio Branco, all populations were sampled from whitewater floodplains within a 33.1 km range. For each population, we identified a corresponding population on the opposite riverbank. For the Rio Negro, corresponding populations located on opposite banks were separated by 4.2 km (Pop3 – Pop25, Figure 1) to 7.5 km (Pop5 – Pop24, Figure 1), while corresponding populations were separated by 1.0 km (Pop14 – Pop15, Figure 1) to 3.2 km (Pop8 – Pop21, Figure 1) on the Rio Branco. Leaves from six individuals of *A. longifolia* per population were sampled and desiccated in silica gel. Our population sample size was informed by an initial study with *A. longifolia*,

TABLE 1 Collection information for the populations of *Amphirrhox longifolia* (A. St.-Hil.) Spreng sampled along the Rio Negro and Rio Branco, Amazon Basin, Brazil

Locality	Population code	River bank	Latitude	Longitude	Voucher no.[a]
Rio Negro	P1	Right	02°04′25.4″S	61°30′32.8″W	AF113
Rio Negro	P2	Right	01°59′27.2″S	61°12′52.1″W	MB322
Rio Negro	P3	Right	01°54′21.0″S	61°20′08.9″W	AF116
Rio Negro	P4	Right	01°40′07.0″S	61°25′00.1″W	MB330
Rio Negro	P5	Right	01°33′14.2″S	61°30′27.8″W	VT407
Rio Negro	P6	Right	01°26′12.3″S	61°34′49.8″W	VT413
Rio Negro	P7	Right	01°22′05.2″S	61°45′55.3″W	AF125
Rio Negro	P8	Right	01°23 27.1″S	61°50′31.6″W	AF131
Rio Negro	P22	Left	01°23′47.5″S	61°47′58.0″W	VT449
Rio Negro	P23	Left	01°29′25.4″S	61°36′34.8″W	VT450
Rio Negro	P24	Left	01°35′40.8″S	61°33′43.7″W	VT455
Rio Negro	P25	Left	01°56′12.9″S	61°21′29.6″W	VT458
Rio Negro	P26	Left	02°01′36.4″S	61°15′25.1″W	VT460
Rio Negro	P27	Left	02°07′15.5″S	61°10′32.7″W	VT461
Rio Branco	P9	Right	01°14′40.0″S	61°49′52.0″W	AF132
Rio Branco	P11	Right	00°56′44.1″S	61°50′58.0″W	VT416
Rio Branco	P12	Right	00°50′55.7″S	61°51′33.1″W	VT417
Rio Branco	P13	Right	00°43′46.3″S	61°51′24.0″W	VT419
Rio Branco	P14	Right	00°35′07.0″S	61°48′15.3″W	VT425
Rio Branco	P15	Left	00°35′25.3″S	61°48′41.4″W	VT427
Rio Branco	P16	Left	00°43′34.5″S	61°52′05.7″W	VT429
Rio Branco	P17	Left	00°52′2.05″S	61°52′51.3″W	VT432
Rio Branco	P18	Left	00°56′46.4″S	61°52′32.1″W	VT434
Rio Branco	P19	Left	01°05′46.5″S	61°52′53.0″W	VT436
Rio Branco	p20	Left	01°14′42.7″S	61°50′56.2″W	VT442
Rio Branco	P21	Left	01°22′24.5″S	61°51′59.7″W	VT447

[a]All specimens are deposited at the University of São Paulo Herbarium (SPF), São Paulo, Brazil.

showing that large numbers of unlinked SNPs lead to results with high resolution even when just two individuals are sampled (Nazareno, Bemmels, Dick, & Lohmann, 2017). Individuals were sampled at intervals of at least 50 m to avoid sampling close relatives. One voucher specimen was sampled per population (Table 1). All vouchers were deposited at the Herbarium of the University of São Paulo (SPF), São Paulo, Brazil.

2.2 | Library preparation and sequencing

Genomic DNA was extracted from leaf samples of *A. longifolia* using the Macherey-Nagel kit (Macherey-Nagel GmbH & Co. KG), following the manufacturer's instructions. Four genomic libraries were created using a double-digest RADseq (ddRAD) protocol (Peterson et al., 2012), with modifications intended to minimize variance in the number of reads per individual within a pool. Specifically, PCRs were performed on each individual sample and PCR products were pooled for size selection, instead of pooling samples prior to size selection and PCR as recommended by Peterson et al. (2012). Before digestion reactions, double-stranded DNA concentrations were quantified using the Qubit dsDNA Assay Kit (Invitrogen) and all samples were adjusted to equal molar concentration. The initial amount of DNA, using the Qubit dsDNA Assay kit, varied from 350 to 500 ng for each sample. Each sample was digested with two high-fidelity restriction enzymes EcoRI and MseI (New England Biolabs). Digestion reactions were carried out in a total volume of 20 μl, using 17 μl of resuspended DNA, 5 units of EcoRI, 5 units of MseI and 1× CutSmart buffer (New England Biolabs) for 3 hr at 37°C, ending with a 20-min deactivation step at 65°C. Reactions were then purified with the Agencourt AMPure XP system (Beckman Coulter), following the manufacturer instructions, with elution in 40 μl TE buffer. In order to standardize the initial DNA mass to be added into an adapter ligation, the cleaned digests were quantified using Qubit. Adapter ligations were carried out in a total volume of 30 μl, combining 42 ng DNA, 0.22 μM of a non-sample-specific MseI adaptor (common for all samples), 0.33 μM of a sample-specific EcoRI double-strand adaptor for each DNA sample, 1U of T4 DNA ligase (New England BioLabs) and 1.3× T4 ligase buffer which were incubated at 23°C for 30 min. Reactions were then heat-killed at 65°C for 10 min following a slow cooling to room temperature (23°C). A total of 96 EcoRI double-stranded barcodes with a unique 10-base pair sequence were created using Python; for further details on the barcodes and the MseI oligos sequences, see Nazareno et al. (2017). After cleaning the reactions with the Agencourt AMPure XP system, ligation products were amplified in 20 μl PCRs, each containing 13.5 μl of the ligation product, 0.2 μM of each primer, 0.2 mM dNTPs, 1.0 mM $MgCl_2$, 0.5 U of iProof™ High-Fidelity DNA polymerase (Bio-Rad) and 2× of iProof buffer. The PCR protocol (98°C for 30s, 20 cycles of 98°C for 20s, 60°C for 30s and 72°C for 40s, followed by a final extension at 72°C for 10 min) was carried out in an Eppendorf PCR system. Before pooling samples at each library, each sample was purified using the Agencourt AMPure XP system and the DNA was quantified using Qubit. DNA concentration of each sample ranged from 2.36 to 3.54 ng/μl. Multiplexed libraries were prepared with approximately equal amounts of DNA among samples. Automated size selection was performed using a 2% agarose cartridge (Pippin Prep; Sage Science, Beverly, MA) to select genomic fragments at a target range size of 375–475 bp. Size, quantity and quality of each individual library were measured on the Agilent 2100 Bioanalyzer (Agilent Technologies) using the Agilent DNA 1000 Kit. Each library was sequenced (100-bp single-end reads) in a single lane of an Illumina HiSeq 2000 flow cell (Illumina Inc., San Diego, CA) at The Centre for Applied Genomics in Toronto, Canada (each lane was pooled with 34–65 *A. longifolia* samples).

2.3 | Identifying and genotyping SNPs

Files containing the raw sequence reads for all *A. longifolia* individuals were analysed in STACKS 1.35 (Catchen, Amores, Hohenlohe, Cresko, & Postlethwait, 2011; Catchen, Hohenlohe, Bassham, Amores, & Cresko, 2013) using *de novo* assembly. Initially, we used

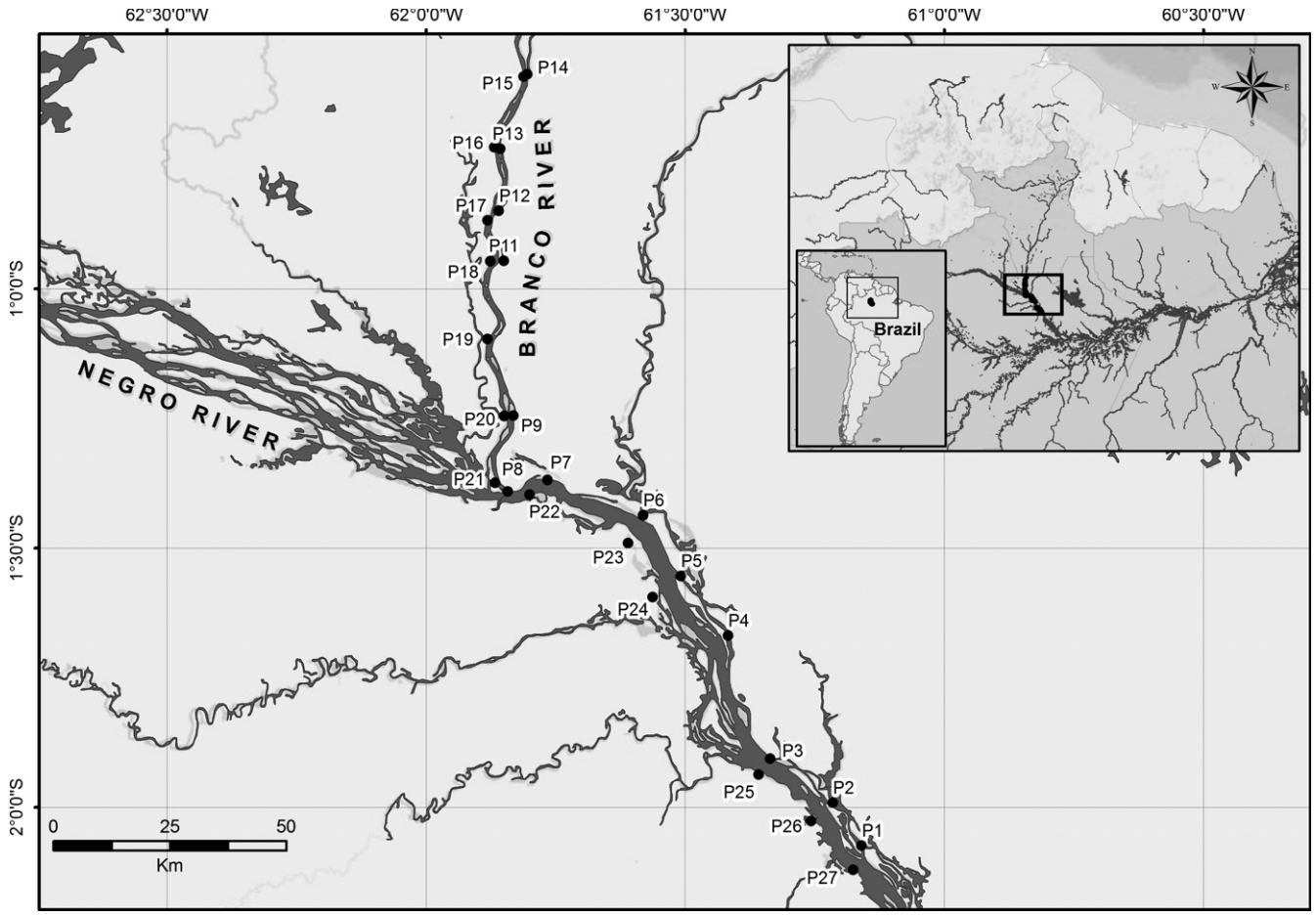

FIGURE 1 *Amphirrhox longifolia* (A. St.-Hil.) Spreng populations (P1 to P27) sampled in the wet season (March–April 2015) along the left and right banks of the Rio Negro and Rio Branco, northern Brazil

the process_radtags program in STACKS to assign reads to individuals and eliminate poor quality reads and reads missing the expected EcoRI cut site (options –barcode_dist 4 -q -e ecoRI). All sequences were processed in ustacks to produce consensus sequences of RAD tags. The program ustacks takes a set of short-read sequences from a single sample as input and aligns them into exactly matching stacks. A maximum-likelihood framework (Hohenlohe et al., 2010) was then applied to estimate the diploid genotype for each individual of *A. longifolia* at each nucleotide position. The optimum minimum depth of coverage to create a stack was set at three sequences, the maximum distance allowed between stacks was two nucleotides, and the maximum number of stacks allowed per de novo locus was three. The stacks assembly enabled the Deleveraging algorithm (–d), which resolves overmerged tags, and the Removal algorithm (–*r*), which drops highly repetitive stacks and nearby errors from the algorithm. The alpha value for the SNP model was set at 0.05. Cstacks was used to build a catalogue of consensus loci containing all the loci from all the individuals and merging all alleles together. Then, each individual genotype was compared against the catalogue using sstacks. We then used rxstacks to exclude problematic loci with a log-likelihood <–100 and loci that matched a single catalogue locus (conf_limit = 0.25) or any nonbiological haplotypes (–prune_haplo) in more than 25% of individuals. Subsequently, cstacks and sstacks were performed again using the same parameters described above. The POPULATIONS program was run to obtain the loci that were present in at least 85% of the *A. longifolia* individuals and 100% of the sampling populations, with sequencing depth of 10×. Even though the use of multiple SNPs within loci has a strong effect on statistical power (Morin, Martien, & Taylor, 2009), we only included the first SNP per locus in the final analysis. All raw sequence reads are available from the National Center for Biotechnology Information Short Read Archive (Accession No. PRJNA381254).

2.4 | Quality control of genomic data

The numbers of raw sequence reads and unlinked SNPs were characterized for all populations of *A. longifolia*. We used the GENALEX 6.5 program (Peakall & Smouse, 2006) to remove SNP markers that were not at Hardy–Weinberg equilibrium (HW). We also used BayeScan 2.1 to remove the SNPs potentially under balancing and divergent selection; this software was run with 20 pilot runs of

10,000 iterations, a burn-in of 50,000 iterations and a final run of 100,000 iterations. In order to minimize false positives, prior odds of the neutral model were set to 10,000 (i.e., the neutral model is 10,000 times more likely than the model with selection; Foll & Gaggiotti, 2008).

2.5 | Population genomic analyses

After filtering SNP markers out of HWE and under selection, we assessed spatial patterns in the genetic structure between *A. longifolia* populations using several approaches. First, we calculated genetic distances among populations (DA: Nei, Tajima, & Tateno, 1983) and visualized the results by applying multidimensional scaling (MDS) in XL-STAT (Addinsoft), using the SMACOF method (Scaling by MAjorizing a COnvex Function), which minimizes the "normalized stress" (De Leeuw, 1977). The MDS is an ordination technique that plots populations as points in low-dimensional space while attempting to maintain the relative distances between populations as close as possible to the actual rank order of similarities between populations. Thus, *A. longifolia* populations with similar genetic structure are plotted closer together in ordination space established by a stress factor. MDS requires no assumptions regarding the cause of structure and does not assume Hardy–Weinberg or gametic equilibrium. In addition to the MDS analysis, a Bayesian model was developed using a Markov chain Monte Carlo (MCMC), as implemented in the R package GENELAND 4.0.2 (Guillot, 2012). This provided an alternative method of clustering populations as it incorporates spatial data in order to identify spatially explicit genetic discontinuities. This method operates by minimizing the Hardy–Weinberg and linkage disequilibrium that would result if individuals from different, randomly mating populations were incorrectly grouped into a population. We used a spatial model with correlated allele frequencies as proposed and implemented by Guillot, Santos, and Estoup (2008) and Guillot, Renaud, Ledevin, Michaux, and Claude (2012). Spatially explicit models take into account the spatial location of the individuals to improve the inference power of the substructure when differentiation occurs by limited gene flow driven by the presence of physical barriers. We conducted one hundred independent runs of 1,000,000 in length, discarding the first 500,000 iterations (burn-in) in postprocessing. The most likely number of k populations was unknown and hence treated as a simulated variable along with the MCMC simulations ($1 \leq k \leq 26$). The number of genetic clusters (K) was inferred as the modal number of genetic groups of the best run (based on posterior density values).

Second, we estimated pairwise genetic differentiation (F_{ST}) for *A. longifolia* populations using ANOVA following Weir and Cockerham (1984). We used the SPAGEDI program (Hardy & Vekemans, 2002) to calculate F_{ST}. Using the same program, we also estimated the significance of the deviation of F_{ST} values using a jackknife procedure over loci. In order to investigate the visual similarity between genetic and geographic distance from the MDS and GENELAND methods, we conducted a test for isolation by distance (IBD) to see whether this pattern meets the expectation of decreased genetic similarity with geographic distance using a Mantel test (Mantel, 1967). Furthermore, to test the riverine hypothesis for the Rio Negro and Rio Branco, the genetic structure was deconstructed by a multiple matrix regression in order to assess the relative contribution of long-term historical divergence and the effects attributed to IBD. For this, we used the model proposed by Legendre and Legendre (1998) which evaluates the relationship among three different matrices: (1) a matrix of pairwise genetic distances [$F_{ST}/(1 - F_{ST})$; Rousset, 1997;] between *A. longifolia* populations; (2) a matrix of Euclidian distances representing the geographic distance between pairwise *A. longifolia* populations; and (3) a pairwise binary matrix of isolation by the rivers expressing long-term historical divergence. The latter matrix was constructed by coding the position of each *A. longifolia* population pair relative to the river (populations on the same river bank were scored as "0" and those on different river banks as "1") (see Link et al., 2015; Manly, 1997; Telles & Diniz-Filho, 2005). Using 10,000 permutation tests of significance for the coefficient of correlation, multiple matrix regression and a single Mantel test were performed using R (R Development Core Team 2015).

Third, we used a nested hierarchical analysis of molecular variance (AMOVA—Excoffier, Smouse, & Quattro, 1992) to examine the effect of the rivers on the partitioning of genetic variation between populations. For each river, we defined two hierarchical levels at which we characterized population differentiation: between populations from opposite riverbanks and between *A. longifolia* populations along each bank. We used the software ARLEQUIN 3.5.2 (Excoffier & Lischer, 2010) to calculate population differentiation estimates and their statistical significance based on 20,000 random permutations.

Finally, historical relationships between *A. longifolia* populations in both rivers, inferred with a population graph analysis that allows both population splits and migration, were constructed with TreeMix 1.12 (Pickrell & Pritchard, 2012). Treemix incorporates a model to allow for population differentiation in the presence of postdivergence admixture/migration (m), which improves the likelihood fit of a bifurcating phylogeny. The resulting phylogeny is based on a composite maximum likelihood of the local optimum tree as proposed by Felsenstein (1981), with branch lengths proportional to the amount of genetic drift per branch. With this approach, inference is based on "shared genetic drift" between sets of populations, under the premise that shared drift implies a shared evolutionary history (Peter, 2016). We added stepwise migration edges and inspected the results for consistency between runs. The population graph and residuals were visualized using R (R Core Team, 2015).

3 | RESULTS

3.1 | Quality control of genomic data

The number of single-end raw reads of 101 bp produced for each lane of HiSeq 2000 Illumina varied from 51 million (library 4, with 34 *A. longifolia* samples) to 169 million (library 2, with 65 samples). Each read starts with a barcode sequence identifying a sample (up to

10 bp long) and the 6-bp restriction site, followed by 85 bp of usable data. The percentage of reads that passed the default quality filters, including a Phred quality score >33, and contained an identifiable barcode, varied from 68.2% (35,188,029 retained reads, library 4) to 90% (152,867,396 retained reads, library 2). The mean number of retained reads for *A. longifolia* samples was 2,192,050 ± 52,090 *SE*. At the population level, there were significant differences among the means of retained reads (Figure S2), which varied from 1,182,457 ± 147,076 *SE* (ranging from 721,570 to 1,639,058) for Population 21 to 3,225,298 ± 178,391 *SE* for Population 27 (ranging from 2,492,242 to 4,111,315).

Throughout the *A. longifolia* genome, we identified 8,085 polymorphic SNPs within the RAD tag sequences, for all populations, with 15% missing data and 10-fold coverage. Considering all significance levels (i.e., $\alpha = 0.05$, $\alpha = 0.01$ and $\alpha = 0.001$), 0.79% of SNPs deviated from HWE in at least 70% of the populations. SNPs that significantly deviated from HWE (a total of 65 SNPs) were discarded and not used in further analyses. We detected ten potential loci that were under diversifying selection for *A. longifolia* populations, with the false discovery rate (FDR) set to 0.05 (Figure S3). Thus, a total of 8,010 filtered SNPs were used in all genomic analyses.

3.2 | Population genomic structure and the genetic barrier hypothesis

Using the MDS and Bayesian clustering methods, we identified one potential barrier to gene flow in the Rio Negro. An examination of the stress value (Kruskal's stress = 0.131), which represents unexplained information, determined that two dimensions were sufficient to explain the genetic patterns for Rio Negro and Rio Branco. The genetic pattern that emerges from our data, as depicted in the MDS plot, separated samples from the Rio Negro into those on the left bank and those on the right bank, and grouped together all *A. longifolia* populations from both banks of the Rio Branco (Figure 2). The genetic structure pattern from the

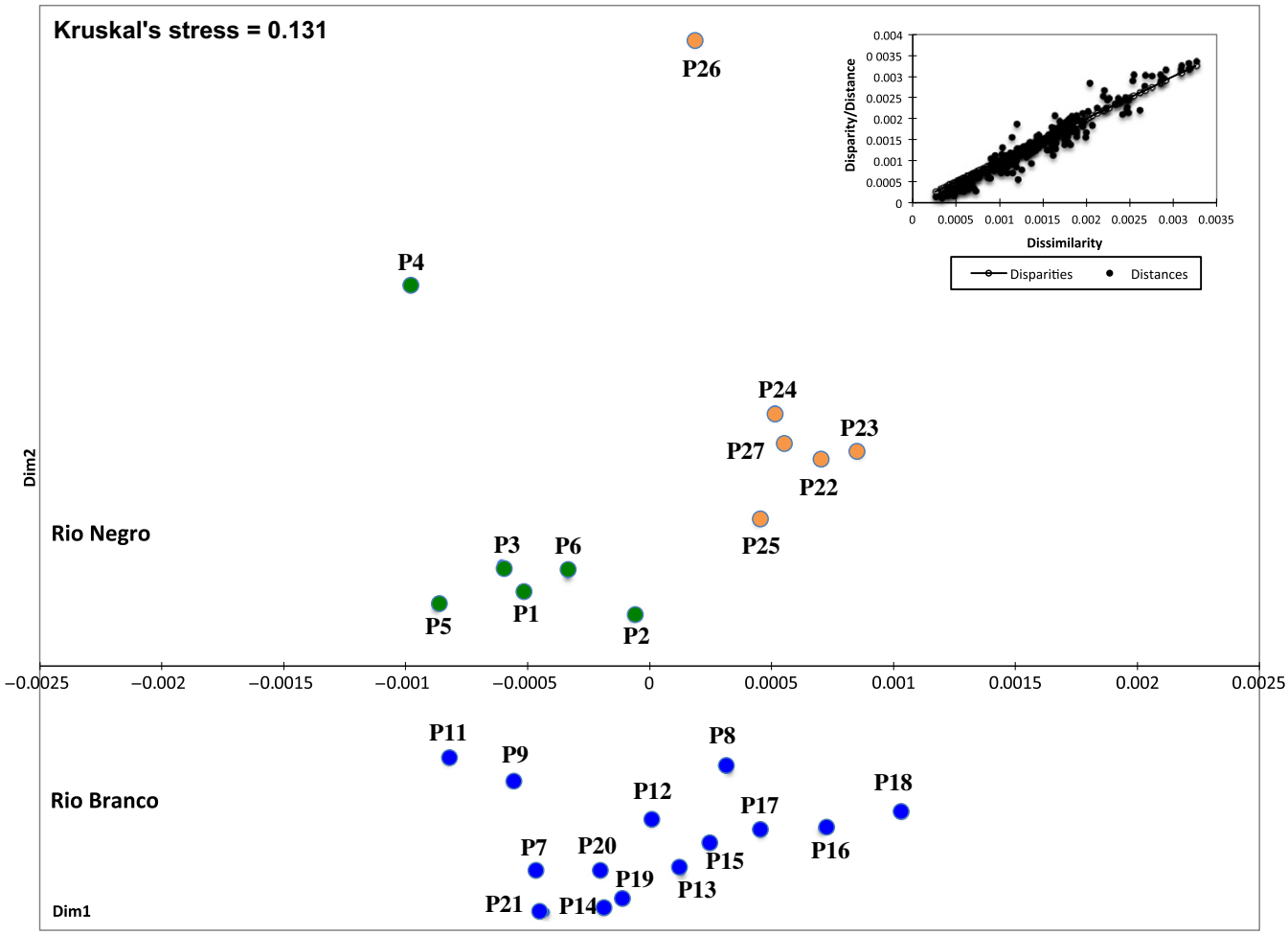

FIGURE 2 Patterns of geographic structure in *Amphirrhox longifolia* (A. St.-Hil.) Spreng as revealed by multidimensional scaling (MDS) of the matrix of genetic distances (DA; Nei et al., 1983; Kruskal's stress = 0.131). The first axis clearly separates populations from the Negro and Branco rivers. The Shepard diagram (inside) shows the quality of the MDS representation. The colours for the Amazonian rivers (and their banks) highlight the three clusters distinguished in the GENELAND analysis (Figure 3): Branco River (both banks) in blue; the banks of Negro River in orange and green [Colour figure can be viewed at wileyonlinelibrary.com]

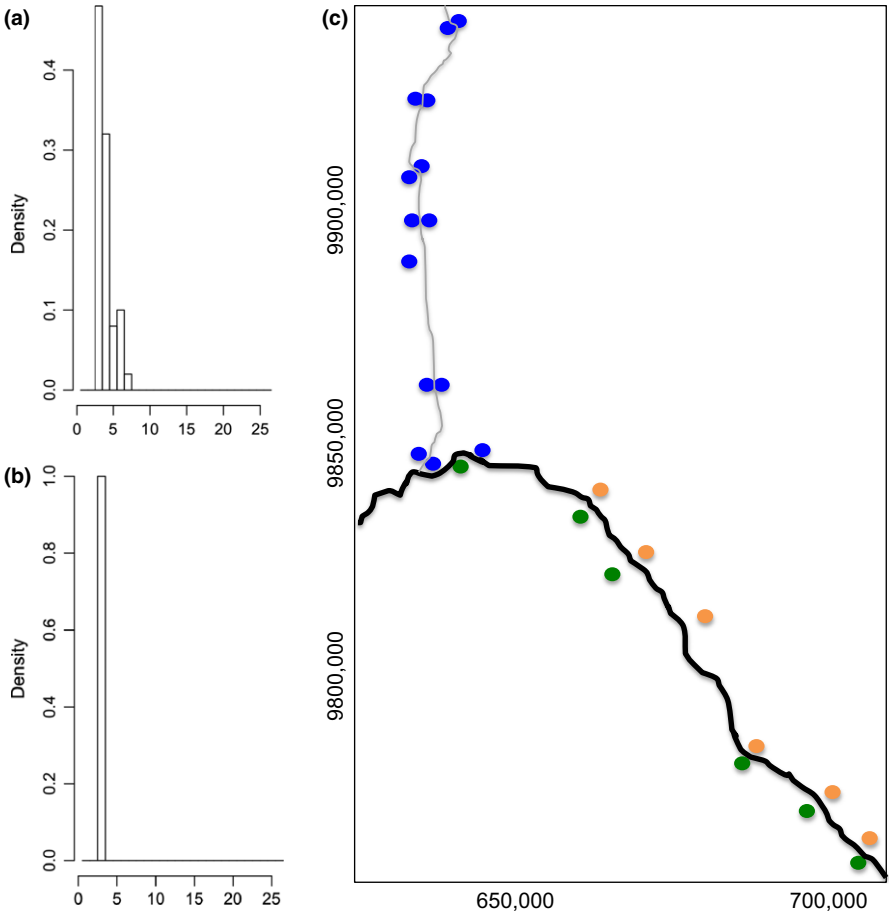

FIGURE 3 Population clustering analysis (a–c) for *Amphirrhox longifolia* (A. St.-Hil.) Spreng populations clearly delineated three groups of populations along the Rio Negro (green and orange circles) and Rio Branco (blue circle). Figures a and b show the density of the estimate of *k* along the Markov chain (after a burn-in of 1,000 × 100 iterations), when spatial data are used (a) or not used (b) in the analyses. Map C is based on population membership as calculated by GENELAND [Colour figure can be viewed at wileyonlinelibrary.com]

MDS analysis closely matched that obtained using Bayesian clustering analysis. Geneland results clearly delineated three groups with minimal variance in the posterior probabilities of population estimation over multiple runs using both spatial (Figure 3a) and nonspatial models (Figure 3b).

Pairwise estimates of F_{ST} varied from 0.0001 to 0.055 and all but five were statistically significant ($p < .05$), indicating limited differentiation between *A. longifolia* population pairs from the Rio Negro (see Table S1). Based on the results from the multiple matrix regression, a significant portion of the variation in pairwise genetic distances is explained either by the historic isolation caused by the river acting as a barrier to gene flow (53%) or isolation by distance (19%). Part of the genetic divergence remains unexplained (44%), and only a limited proportion (16%) of that variation can be explained by the combined effect of a river barrier plus isolation by distance (Figure 4). Results of simple matrix correlation between genetic and geographic distance were not significant when applied separately to both banks of the Rio Negro ($r = .1877$, $p = .1450$ for the right bank; $r = .0624$, $p = .3958$ for the left bank; Figure S4). However, simple matrix correlation between genetic and geographic distance led to highly significant results when applied between pairs of *A. longifolia* populations on

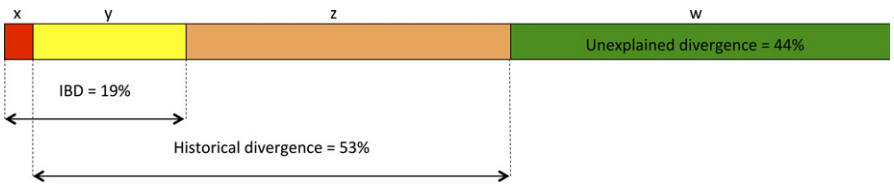

FIGURE 4 Analyses of multiple matrix regression using long-term historical divergence and isolation by distance (IBD) as predictors of genetic divergence among populations of *Amphirrhox longifolia* (A. St.-Hil.) Spreng along the left and right banks of the Rio Negro. The unexplained variation [$w = 1 - (x + y + z)$] was determined using both effects (IBD and long-term historical divergence) as predictors. The overlap between long-term historical divergence and IBD is equal to $y = (x + y) + (y + z) - (x + y + z)$, where $(x + y)$ is the coefficient of determination of the geographic distance (IBD), and $(y + z)$ is the coefficient of determination of the regression using history. Proportion of variation explained by history alone (z) is 37% and IBD alone (x) is 3% [Colour figure can be viewed at wileyonlinelibrary.com]

opposite banks exclusively (r = .6697, p = .007; Figure S4). Consequently, about 67% of the variation in genetic distance can be attributed to geographic distance between pairs of populations on opposite banks of the river.

On the other hand, for populations along the Rio Branco, the coefficient of determination of the multiple matrix regression was not significant (R^2 = .06424, p = .40) and no correlation between genetic distance and geographic distance (r = .176, p = .2211) was found. In addition, the binary model matrix designed to separate right and left populations on the Rio Branco was not significantly correlated with genetic distance (r = .1856, p = .1569). These results indicate that Rio Branco has not acted as a barrier to gene flow for *A. longifolia*. Results of simple matrix correlation between genetic and geographic distance were also not significant when applied separately to both banks of the river (r = −.0528, p = .6243 for the right bank; r = −.1232, p = .7026 for the left bank), nor when applied only between pairs of *A. longifolia* populations on opposite banks of the river (r = .0786, p = .50). The matrix of geographic distance and the pairwise F_{ST} values quantifying genetic differentiation among *A. longifolia* populations on the Rio Branco are shown in Table S2. Pairwise estimates of F_{ST} varied from 0.0045 to 0.0420, and all but two were statistically significant (p < .05), indicating low levels of differentiation between *A. longifolia* population pairs from the Rio Branco (Table S2).

The hierarchical multilocus evaluation of genetic differentiation performed using an AMOVA revealed that a greater proportion of the overall genetic variation exists within populations (97.52% and 98.73%) than either between *A. longifolia* populations from opposite banks (1.43% and 0.01%) or among populations from the same riverbank (1.05% and 1.26%) for the Rio Negro and Rio Branco, respectively (Table 2). In accordance with our previous results, which indicate that Rio Negro can be a barrier to gene flow, most of the genetic variation was weak but significant and attributable to differences observed between banks (ϕ_{CT} = 0.014, p < .001) rather than among populations on the same riverbanks (ϕ_{SC} = 0.010, p < .001). We did not find this pattern for the Rio Branco (Table 2), strengthening our findings that this tributary is not a genetic barrier for *A. longifolia*.

TABLE 2 Analyses of molecular variance (AMOVA) of *Amphirrhox longifolia* (A. St.-Hil.) Spreng for (A) Rio Negro and (B) Rio Branco, Amazon Basin, Brazil

	Sum of squares	Variance components	Percentage of variation	*p*-Value
(A) Rio Negro				
Between banks	156.23	1.11	1.43	<.001
Among populations within banks	1066.44	0.82	1.05	<.001
Within populations	11252.33	75.78	97.52	<.001
Total	12475.00	77.71		
(B) Rio Branco				
Between banks	77.71	0.01	0.01	.365
Among populations within banks	842.41	0.86	1.26	<.001
Within populations	8135.12	67.32	98.73	<.001
Total	9055.24	68.19		

3.3 | Population trees and historical migration events

The best maximum-likelihood tree, considering both rivers, largely reiterates the results for the relationships among *A. longifolia* populations, indicating that populations on each river are more closely related to each other (Figure 5a). Furthermore, the length of the internal branches on the tree—a measure of dissimilarity between populations—implies that *A. longifolia* populations from Rio Negro show greater differentiation than those from the Rio Branco (Figure 5a). The graph model tree for both rivers explained 98.81% of the variance in relatedness between *A. longifolia* populations. In addition, the populations on each side of the Rio Negro clearly formed two distinct clades, corresponding to the left and right banks (Figure 5a). We examined the residuals of the model's fit to identify aspects of ancestry not captured by the tree. The residuals show that there is high standard error between some population pairs, suggesting that they may be candidates for migration events (Figure 5b). For instance, the residual between Pop2 (right bank of the Rio Negro) and Pop23 (left bank) is greater than zero, although the standard error is low (Figure 5b). For the Rio Branco, where populations from both banks of the river were grouped together, a similar pattern was observed between Pop14 and Pop19, which are on opposite riverbanks. We then sequentially added historical migration events (up to five events) to the tree. This graph model (Figure 6) explained 99.81% of the variance in relatedness between populations. The migrations could be observed occurring between populations from the same bank, between populations on opposite banks and between populations from different rivers. This result suggests that gene flow was feasible between *A. longifolia* populations located on opposite riverbanks.

When we analysed both rivers separately, the best maximum-likelihood tree explained 96.95% (Figure 7a) of the variance in relatedness between *A. longifolia* populations from the Rio Negro. The residuals for this tree showed that there is high standard error between some population pairs (e.g., Pop2-Pop23, Pop4-Pop27, Pop4-Pop22, Pop6-Pop27, Pop22-Pop26), mainly for populations on opposite banks of the river (Figure S5A), indicating these may be candidates for migration events. Historical migration events (up to nine events) were added sequentially to the tree. The graph model (Figure 7b) explained 99.94% of the variance in relatedness between *A. longifolia* populations. The first added migration edge goes from Pop23 to the node between Pop2 and Pop3 with a weight of 0.14 (Figure 7b) as indicated by the residuals from the initial model graph tree (Figure S5A). Although there are migrations between *A. longifolia* populations within and among banks of the Rio Negro (Figure 7b), migration events with high weights were observed for population

pairs on the same bank of the Rio Negro (e.g., Pop22-Pop26 and Pop22-Pop25; Figure 7b), with population 22 being the most important source of migrants.

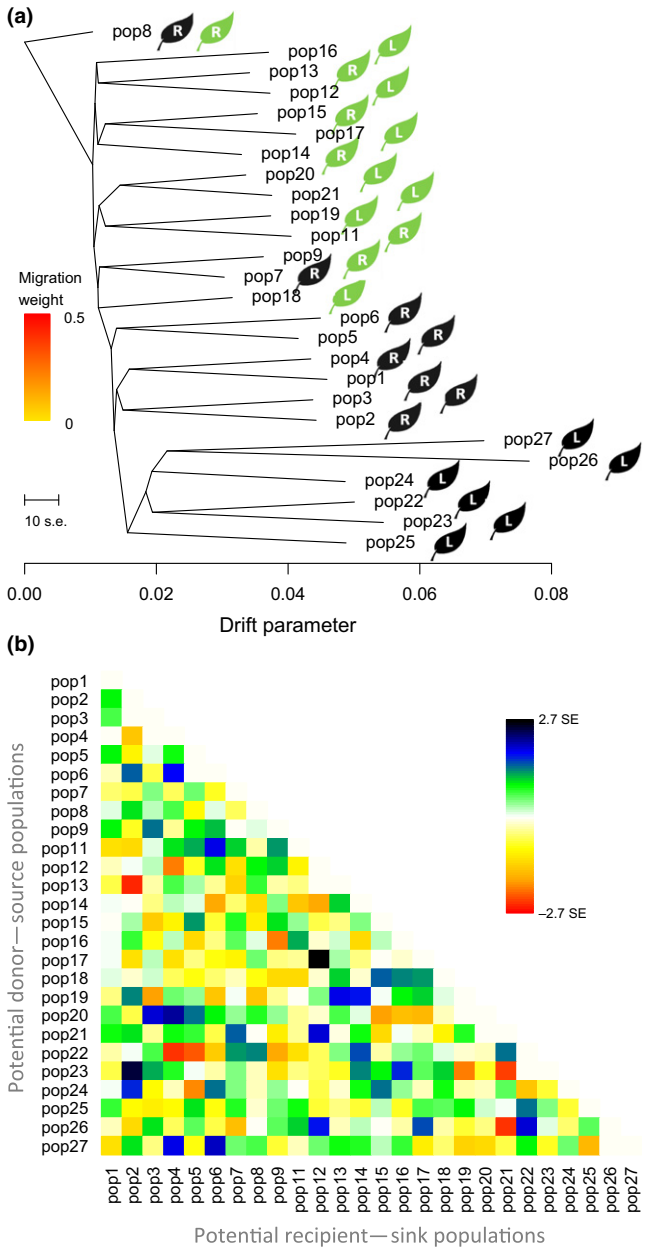

FIGURE 5 (a) Structure of the graph inferred by TreeMix for *Amphirrhox longifolia* (A. St.-Hil.) Spreng populations along the left (L) and right (R) banks of the Rio Negro (black leaf) and Rio Branco (green leaf). Pop8, located within the Rio Negro/Rio Branco interfluvium, was fixed as the root. The scale bar shows ten times the average standard error of the entries in the sample covariance matrix. The drift parameter reflects the amount of genetic drift that has occurred between *A. longifolia* populations. (b) Plotted is the residual fit from the maximum-likelihood tree for both the Rio Negro and Rio Branco. Colours are described in the palette on the right. Residuals above zero represent populations that are more closely related to each other in the data than in the best-fit tree, and are thus candidates for migration events [Colour figure can be viewed at wileyonlinelibrary.com]

For the Rio Branco, the best maximum-likelihood tree explained 95.97% (Figure 8a) of the variance in relatedness between *A. longifolia* populations. The residuals for this tree showed that there is high standard error between some population pairs (e.g., Pop8-Pop19, Pop9-Pop16, Pop11-Pop16, Pop12-Pop17, Pop14-Pop15), mainly for populations on opposite banks (Figure S5B), suggesting candidate populations for migration events. Historical migration events (up to ten events) were added sequentially to the tree. The graph model tree (Figure 8b) explained 99.81% of the variance in relatedness between populations. The first added migration edge goes from Pop21 to the node between Pop19 and Pop11 with a weight of 0.64 (Figure 8b). Although we found migration events between *A. longifolia* populations within and among banks of the Rio Branco (Figure 8b), the greatest migration weight (0.80) was observed for a population pair on opposite banks of the river (Pop12-Pop17; Figure 8b). Population 9 (Pop9) was the greatest source of migrants as migrations were observed between population 9 and populations 11, 12, 15 and 17. Overall, the magnitude of the weight of migration events was higher between *A. longifolia* populations on the Rio Branco (maximum of 1.0; Figure 8b) than in the Rio Negro (maximum of 0.5; Figure 7b).

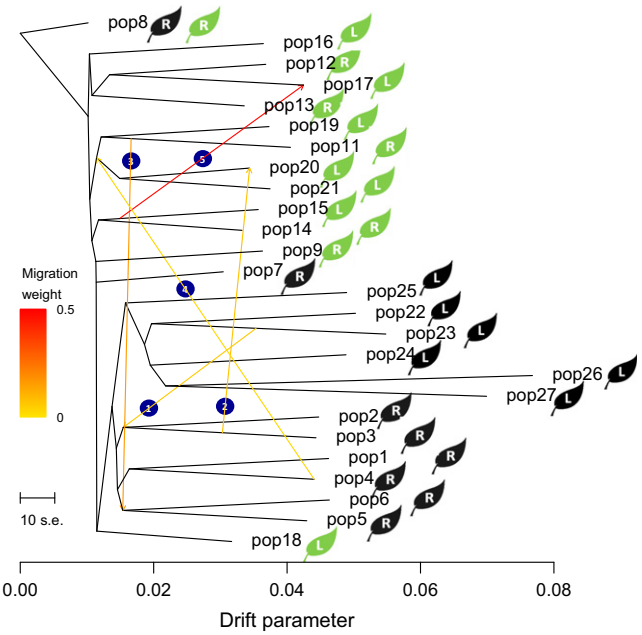

FIGURE 6 Structure of the graph inferred by TreeMix, with five migration edges (m = 1 to 5 events), for *Amphirrhox longifolia* (A. St.-Hil.) Spreng populations along the left and right banks of the Rio Negro (black leaf) and Rio Branco (green leaf). Pop8, located within the Rio Negro/Rio Branco interfluvium, was fixed as the root. Migration arrows, coloured according to their weight, relates to the fraction of alleles in the descendant population that originated in the ancestral population and suggests multiple plausible instances of migration. The scale bar shows ten times the average standard error of the entries in the sample covariance matrix. The drift parameter reflects the amount of genetic drift that has occurred between *A. longifolia* populations [Colour figure can be viewed at wileyonlinelibrary.com]

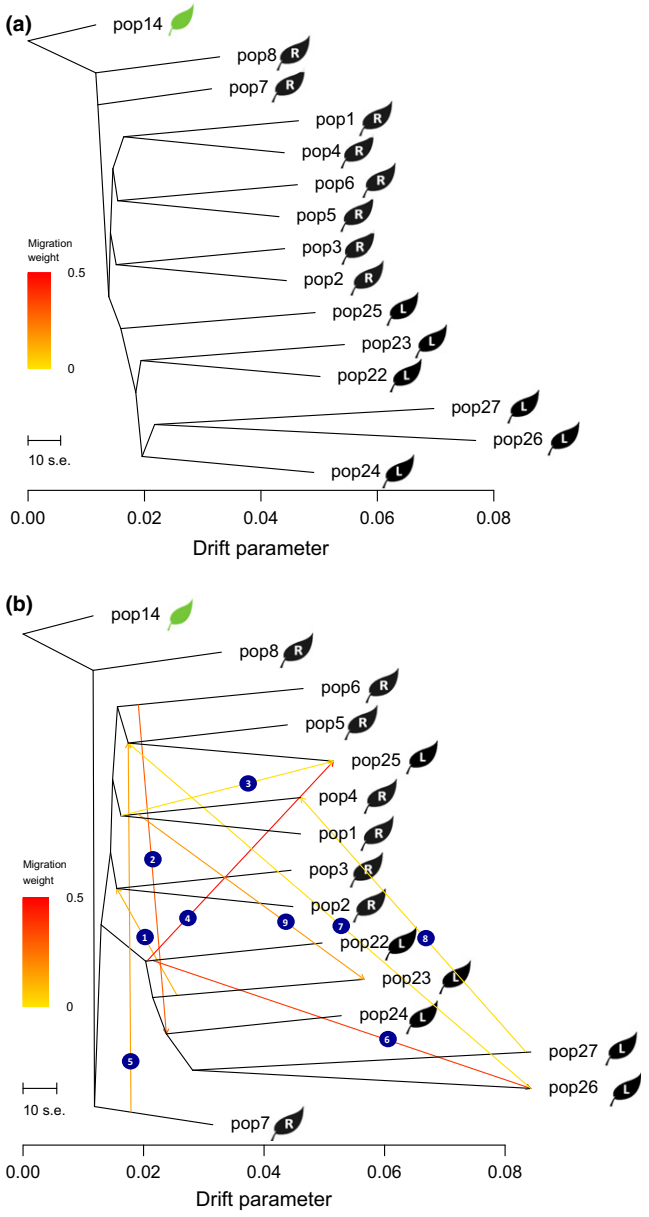

FIGURE 7 (a) Structure of the graph inferred by TreeMix for *Amphirrhox longifolia* (A. St.-Hil.) Spreng populations along the left (black leaf—L) and right (black leaf—R) banks of the Rio Negro. (b) Population tree with branch lengths scaled to the amount of genetic drift between banks of the Rio Negro and inferred proportion of migration (m = 1 to 9 events) are shown by arrows. Pop14 from the Rio Branco (green leaf) was fixed as the root. The scale bar shows ten times the average standard error of the entries in the sample covariance matrix. The drift parameter reflects the amount of genetic drift that has occurred between *A. longifolia* populations. The residual fit from the graph A is presented in Figure S5 [Colour figure can be viewed at wileyonlinelibrary.com]

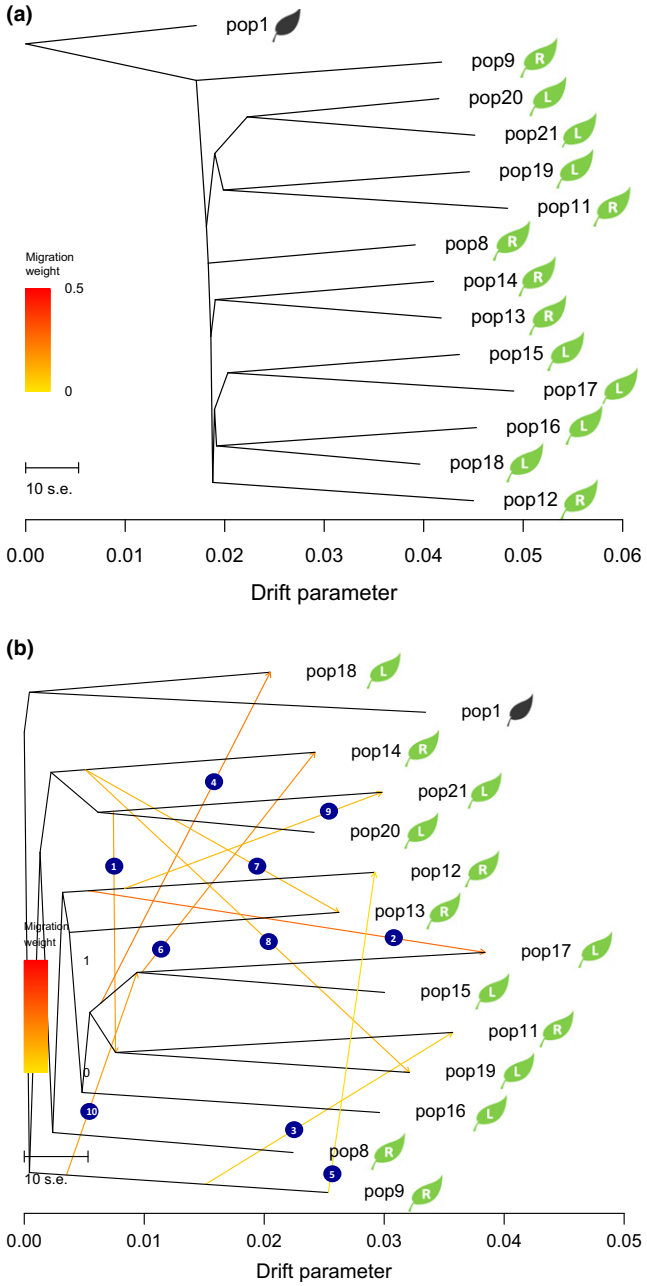

FIGURE 8 (a) Structure of the graph inferred by TreeMix for *Amphirrhox longifolia* (A. St.-Hil.) Spreng populations along the left (green leaf—L) and right (green leaf—R) banks of the Rio Branco. (b) Population tree with branch lengths scaled to the amount of genetic drift between banks of the Rio Branco and inferred proportion of migration (m = 1 to 10 events) are shown by arrows. Pop1 from the Rio Negro (black leaf) was fixed as the root. The scale bar shows ten times the average standard error of the entries in the sample covariance matrix. The drift parameter reflects the amount of genetic drift that has occurred between *A. longifolia* populations. The residual fit from the graph A is presented in Figure S5 [Colour figure can be viewed at wileyonlinelibrary.com]

4 | DISCUSSION

The Amazon Basin is reknowned for its rich biological diversity, being one of the most diverse regions on the planet. Biologists have long sought to understand evolutionary processes governing its diversification history. Biogeographical studies have invoked marine incursions (Lovejoy, Albert, & Crampton, 2006; Webb, 1995), Pleistocene forest refugia (Haffer, 1969, 1997; Haffer & Prance, 2001) and riverine barriers (Ayres & Clutton-Brock, 1992; Gascon et al., 2000; Wallace, 1852) as allopatric drivers of speciation. The

evolutionary history of Amazon diversity is complex, and even a basic understanding of Amazon landscape history remains elusive (Baker et al., 2014). Identifying and explaining patterns of population genetic structure in Amazon plants can help to advance understanding of the region's physical and biotic history. Relative to studies of animal groups, there are few botanical studies of Amazon biogeography (Fiaschi & Pirani, 2009).

In this study, we evaluated Wallace's (1852) riverine barrier hypothesis as it applies to plants along two major tributaries of the Amazon River, the Rio Negro and Rio Branco. Both rivers are allopatric barriers for primates and birds (Bonvicino et al., 2003; Boubli et al., 2008; Boubli et al., 2015; Naka et al. 2012). The Rio Negro has been studied more frequently (Bonvicino et al., 2003; Boubli et al., 2008; Boubli et al., 2015; Fernandes, Wink, & Aleixo, 2012; Fernandes, Wink, et al., 2014; Horta, Cuervo, Ribas, Brumfield, & Miyaki, 2013; Moraes et al., 2016; Naka et al. 2012; Ribas et al., 2012), and numerous zoological studies focussed on the Rio Negro generally find support for the riverine barrier hypothesis (e.g., Avila-Pires, 1995; Peres, Patton, & da Silva, 1996; Hayes & Sewlal, 2004; Funk et al., 2007; Fernandes, Gonzales, Wink, & Aleixo, 2013; Fernandes, Cohn-Haft, et al., 2014; —but see Patton, Dasilva, & Malcolm, 1994; Lougheed et al., 1999; Gascon et al., 2000; Patton et al., 2000).

Commensurate with its high species richness, most Amazon tree populations occur in low densities or have scattered distributions (Dick, Hardy, Jones, & Petit, 2008). We selected *Amphirrrhox longifolia* to examine the riverine barrier hypothesis in part because it was one of the few plant species common enough to be sampled along both the Rio Branco and the Rio Negro. In order to perform a robust genetic analysis using relatively small population samples, we utilized the extensive genomic coverage available through RADseq (Nazareno et al., 2017). There was no prior population genetic data available for *Amphirrhox longifolia*. Although the specific pollinator of *A. longifolia* is unknown, its floral morphology suggests pollination by short-flying small insects. The fruits and seeds of this widespread species do not have specific adaptations for dispersal by water or by fish.

Our population genetic data indicated significant population genetic structure for *A. longifolia* throughout the Rio Negro. Although a moderate proportion of the genetic variation (56%) can be explained by the river acting as a barrier, population genetic analyses revealed limited gene flow between riverbanks separated by 4.2 to 6.7 km. In contrast to our findings along the Rio Negro, our analysis did not recover genetic structuring across the Rio Branco. This is the first genetic study (to our knowledge) of plants along the Rio Branco. This absence of genetic structure has been noted for primates (Boubli et al., 2015), although for studied birds the Rio Branco is apparently a significant gene-flow barrier (Naka et al. 2012).

Both the age and changing magnitude of the Amazon and its tributaries have likely impacted plant genetic structure and speciation. For example, the Rio Negro responded significantly to environmental changes in the Late Pleistocene and Holocene (Latrubesse & Franzinelli 2005). As such, rivers that seem to be dispersal barriers today may not have been barriers in the past. Hydrological connectivity over historical and geological timescales in the Rio Negro may have occurred due to the translocation of islands between riverbanks (Silva et al., 2007). Furthermore, molecular studies in multiple animal taxa show younger divergences across the Rio Branco than the Rio Negro (Boubli et al., 2015; Ribas et al., 2012). Although the full geological history of the Amazon and its tributaries may never be well understood, a fuller understanding of regional landscape history may enable us to better predict genetic patterns in widely distributed plant species.

Collevatti et al. (2009) performed the only botanical study of the riverine barrier hypothesis focused the Rio Negro. Their study focused on a low-density and widely distributed canopy-emergent tree species (*Caryocar villosum*) that grows in the upland forests, and a habitat-specific tree (*C. microcarpum*) that grows in seasonally flooded blackwater forests (Collevatti et al., 2009). The authors did not find evidence of genetic differentiation. This result was perhaps not unexpected given the long distances of gene flow associated with bat pollination and seed dispersal by strong swimming tapirs and fish (Collevatti et al., 2009). Clearly, population-level and comparative studies for multiple plant species are needed for a more comprehensive assessment of the role of the Amazonian waterways on Amazonian plants. We think that the use of genomic approaches like RADseq will open Amazon forests to a broad array of landscape genetic and biogeographical analyses.

ACKNOWLEDGEMENTS

The authors thank Jordan Bemmels for valuable input on an earlier version of this manuscript. We thank Verônica Thode, Maila Beyer, Beatriz Gomez, Annelise Frazão and Osmar Pereira for their great help during fieldwork. We also thank the Core Facility for Scientific Research (CEFAP) of the Universidade de São Paulo for computational support. Dr. Mark Ungerer and three anonymous reviewers provided comments and suggestions that greatly improved this manuscript.

DATA ACCESSIBILITY

All input files are available from Dryad (https://doi.org/10.5061/dryad.p8n59). Raw sequence data files for all the *A. longifolia* populations are available from the National Center for Biotechnology Information Short Read Archive (NCBI-SRA).

AUTHOR CONTRIBUTIONS

A.G.N. and L.G.L. designed the study and coordinated sample collection. A.G.N. conducted molecular work, performed analyses and led the writing of the manuscript with input from all co-authors. C.W.D. provided laboratory assistance, analytical input and troubleshooting.

REFERENCES

Almeida-Filho, R., & Miranda, F. P. (2007). Mega capture of the Rio Negro and formation of the Anavilhanas Archipelago, Central

Amazonia, Brazil: Evidences in an SRTM digital elevation model. *Remote Sensing of Environment*, *110*, 387–392.

Avila-Pires, T. C. S. (1995). Lizards of Brazilian Amazonia (Reptilia: Squamata). *Zoologische Verhandelingen*, *299*, 1–706.

Ayres, J. M., & Clutton-Brock, T. H. (1992). River boundaries and species range size in Amazonian primates. *The American Naturalist*, *140*, 531–537.

Baker, P. A., Fritz, S. C., Dick, C. W., Eckert, A. J., Horton, B. K., Manzoni, S., … Battisti, D. (2014). The emerging field of geogenomics: Constraining geological problems with genetic data. *Earth Science Reviews*, *135*, 38–47.

Bates, J. M., Haffer, J., & Grismer, E. (2004). Avian mitochondrial DNA sequence divergence across a headwater stream of the Rio Tapajos, a major Amazonian river. *Journal of Ornithology*, *145*, 199–205.

Bonvicino, C. R., Boubli, J. P., Otazu, I. B., Almeida, F. C., Nascimento, F. F., Coura, J. R., & Seuanez, H. N. (2003). Morphologic, karyotypic and molecular evidence of a new form of Chiropotes (Primates, Pitheciinae). *American Journal of Primatology*, *61*, 123–133.

Boubli, J. P., da Silva, M. N. F., Amado, M. V., Hrbek, T., Pontual, F. B., & Farias, I. P. (2008). A taxonomic reassessment of *Cacajao melanocephalus* Humboldt (1811), with the description of two new species. *International Journal of Primatology*, *29*, 723–741.

Boubli, J. P., Ribas, C., Alfaro, J. W. L., Alfaro, M. E., da Silva, M. N., Pinho, G. M., & Farias, I. P. (2015). Spatial and temporal patterns of diversification on the Amazon: A test of the riverine hypothesis for all diurnal primates of Rio Negro and Rio Branco in Brazil. *Molecular Phylogenetics and Evolution*, *82*, 400–412.

Braun, M., Dotter, S., Schlindwein, C., & Gottsberger, G. (2012). Can nectar be a disadvantage? Contrasting pollination natural histories of two woody Violaceae from the Neotropics. *International Journal of Plant Sciences*, *173*, 161–171.

Catchen, J. M., Amores, A., Hohenlohe, P., Cresko, W., & Postlethwait, J. H. (2011). Stacks: Building and genotyping loci de novo from short-read sequences. *G3 (Bethesda)*, *1*, 171–182.

Catchen, J. M., Hohenlohe, P. A., Bassham, S., Amores, A., & Cresko, W. A. (2013). Stacks: An analysis tool set for population genomics. *Molecular Ecology*, *22*, 3124–3140.

Collevatti, R. G., Leoi, L. C. T., Leite, S. A., & Gribel, R. (2009). Contrasting patterns of genetic structure in *Caryocar* (Caryocaraceae) congeners from flooded and upland Amazonian forests. *Biological Journal of the Linnean Society*, *98*, 278–290.

De Leeuw, J. (1977). Correctness of Kruskal's algorithms for monotone regression with ties. *Psychometrika*, *42*, 141–144.

Dick, C. W., Hardy, O. J., Jones, F. A., & Petit, R. J. (2008). Spatial scales of pollen and seed-mediated gene flow in tropical rain forest trees. *Tropical Plant Biology*, *1*, 20–33.

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

Excoffier, L., Smouse, P., & Quattro, J. (1992). Analysis of molecular variance inferred from metric distances among DNA haplotypes: Application to human mitochondrial DNA restriction data. *Genetics*, *131*, 479–491.

Felsenstein, J. (1981). Evolutionary trees from gene frequencies and quantitative characters: Finding maximum likelihood estimates. *Evolution*, *35*, 1229–1242.

Fernandes, A. M., Cohn-Haft, M., Hrbek, T., & Farias, I. P. (2014). Rivers acting as barriers for bird dispersal in the Amazon. *Revista Brasileira de Ornitologia*, *22*, 363–373.

Fernandes, A. M., Gonzales, J., Wink, M., & Aleixo, A. (2013). Multilocus phylogeography of the Wedge-billed Woodcreeper *Glyphorynchus spirurus* (Aves, Furnariidae) in lowland Amazonia: Widespread cryptic diversity and paraphyly reveal a complex diversification pattern. *Molecular Phylogenetics and Evolution*, *66*, 270–282.

Fernandes, A. M., Wink, M., & Aleixo, A. (2012). Phylogeography of the chestnut-tailed antbird (*Myrmeciza hemimelaena*) clarifies the role of rivers in Amazonian biogeography. *Journal of Biogeography*, *39*, 1524–1535.

Fernandes, A. M., Wink, M., Sardelli, C., & Aleixo, A. (2014). Multiple speciation across the Andes and throughout Amazonia: The case of the spot-backed antbird species complex (*Hylophylax naevius/Hylophylax naevioides*). *Journal of Biogeography*, *41*, 1094–1104.

Ferreira, L. V. (2000). Effect of flooding duration on species richness, floristic composition and forest structure in river margin habitats in Amazonian blackwater floodplain forests: Implications for future design of protected areas. *Biodiversity Conservation*, *9*, 1–14.

Fiaschi, P., & Pirani, J. R. (2009). Review of plant biogeographic studies in Brazil. *Journal of Systematics and Evolution*, *47*(5), 477–496.

Foll, M., & Gaggiotti, O. E. (2008). A genome scan method to identify selected loci appropriate for both dominant and codominant markers: A Bayesian perspective. *Genetics*, *180*, 977–993.

Funk, W. C., Caldwell, J. P., Peden, C. E., Padial, J. M., De la Riva, I., & Cannatella, D. C. (2007). Tests of biogeographic hypotheses for diversification in the Amazonian forest frog, *Physalaemus petersi*. *Molecular Phylogenetics and Evolution*, *44*, 825–837.

Gascon, C., Malcolm, J. R., Patton, J. L., da Silva, M. N. F., Bogart, J. P., Lougheed, S. C., … Boag, P. T. (2000). Riverine barriers and the geographic distribution of Amazonian species. *Proceedings of the National Academy of Sciences of the United States of America*, *97*, 13672–13677.

Guillot, G. (2012). *Population genetic and morphometric data analysis using R and the Geneland program*. Retrieved from http://www2.imm.dtu.dk/~gigu/Geneland/Geneland-Doc.pdf (accessed 20 July 2016)

Guillot, G., Renaud, S., Ledevin, R., Michaux, J., & Claude, J. (2012). A unifying model for the analysis of phenotypic, genetic and geographic data. *Systematic Biology*, *61*, 897–911.

Guillot, G., Santos, F., & Estoup, A. (2008). Analyzing georeferenced population genetics data with Geneland: A new algorithm to deal with null alleles and a friendly graphical user interface. *Bioinformatics*, *24*, 1406–1407.

Haffer, J. (1969). Speciation in Amazonian forest birds. *Science*, *165*, 131–137.

Haffer, J. (1997). Alternative models of vertebrate speciation in Amazonia: An overview. *Biodiversity and Conservation*, *6*, 451–476.

Haffer, J., & Prance, G. T. (2001). Climatic forcing of evolution in Amazonia during the Cenozoic: On the refuge theory of biotic differentiation. *Amazoniana- Limnologia et Oecologia Regionalis Systemae Fluminis Amazonas*, *16*, 579–605.

Hall, J. P. W., & Harvey, D. J. (2002). The phylogeography of Amazonia revisited: New evidence from riodinid butterflies. *Evolution*, *56*, 1489–1497.

Hardy, O. J., & Vekemans, X. (2002). SPAGeDi: A versatile computer program to analyze spatial genetic structure at the individual or population levels. *Molecular Ecology Notes*, *2*, 618–620.

Hayes, F. E., & Sewlal, J. N. (2004). The Amazon river as a dispersal barrier to passerine birds: Effects of river width, habitat and taxonomy. *Journal of Biogeography*, *31*, 1809–1818.

Hohenlohe, P. A., Bassham, S., Etter, P. D., Stiffler, N., Johnson, E. A., & Cresko, W. A. (2010). Population genomics of parallel adaptation in three-spine stickleback using sequenced RAD tags. *PLoS Genetics*, *6*, e1000862.

Horta, F. M., Cuervo, A. M., Ribas, C. C., Brumfield, R. T., & Miyaki, C. Y. (2013). Phylogeny and comparative phylogeography of *Sclerurus* (Aves: Furnariidae) reveal constant and cryptic diversification in an old radiation of rain forest understory specialists. *Journal of Biogeography*, *40*, 37–49.

Keel, S. H., & Prance, G. T. (1979). Studies of the vegetation of a black water igapó (Rio Negro-Brazil). *Acta Amazonia*, *9*, 645–655.

Latrubesse, E. M., & Franzinelli, E. (2005). The late Quaternary evolution of the Negro River Amazon, Brazil: Implications for island and floodplain formation in large anabranching tropical systems. *Geomorphology*, *70*, 372–397.

Latrubesse, E. M., Stevaux, J. C., & Sinha, R. (2005). Tropical rivers. *Geomorphology*, *70*, 187–206.

Legendre, P., & Legendre, L. (1998). *Numerical ecology*, 2nd ed. Amsterdam: Elsevier Science.

Link, A., Valencia, L. M., Céspedes, L. N., Duque, L. D., Cadena, C. D., & Di Fiore, A. (2015). Phylogeography of the critically endangered brown spider monkey (*Ateles hybridus*): Testing the riverine barrier hypothesis. *International Journal of Primatology*, *36*, 530–547.

Lougheed, S. C., Gascon, C., Jones, D. A., Bogart, J. P., & Boag, P. T. (1999). Ridges and rivers: A test of competing hypotheses of Amazonian diversification using a dart-poison frog (*Epipedobates femoralis*). *Proceedings of the Royal Society of London Series B-Biological Sciences*, *266*, 1829–1835.

Lovejoy, N. R., Albert, J. S., & Crampton, W. G. R. (2006). Miocene marine incursions and marine/freshwater transitions: Evidence from Neotropical fishes. *Journal of South American Earth Sciences*, *21*, 5–13.

Macedo, M., & Prance, G. T. (1978). Notes on the vegetation of Amazonia II. The dispersal of plants in Amazonian white sand campinas: The campinas as functional islands. *Brittonia*, *30*, 203–215.

Maldonato-Coelho, M., Blake, J. G., Silveira, L. F., Batalha-Filho, H., & Ricklefs, R. E. (2013). Rivers, refuges and population divergence of fire-eye antbirds (*Pyriglena*) in the Amazon Basin. *Journal of Evolutionary Biology*, *26*, 1090–1107.

Manly, B. F. J. (1997). *Randomization, bootstrap and Monte Carlo Methods in Biology*. London, UK: Chapman & Hall.

Mantel, N. (1967). The detection of disease clustering and a generalized regression approach. *Cancer Research*, *27*, 209–220.

Missouri Botanical Garden (2009). *Amphirrhox longifolia (A. St.-Hil.) Spreng*. Retrieved from http://www.tropicos.org/Name/33800557 (accessed 10 June 2016).

Moraes, L. J. C. L., Pavan, D., Barros, M. C., & Ribas, C. C. (2016). The combined influence of riverine barriers and flooding gradients on biogeographical patterns for amphibians and squamates in south-eastern Amazonia. *Journal of Biogeography*, *43*, 2113–2124.

Morin, P. A., Martien, K. K., & Taylor, B. L. (2009). Assessing statistical power of SNPs for population structure and conservation studies. *Molecular Ecology Resources*, *9*, 66–73.

Naka, L. N., Becchtoldt, C. L., Henriques, L. M. P., & Brumfield, R. T. (2012). The role of physical barriers in the location of avian suture zones in the Guiana Shield, northern Amazonia. *The American Naturalist*, *179*, E115–E132.

Nazareno, A. G., Bemmels, J. B., Dick, C. W., & Lohmann, L. G. (2017). Minimum sample sizes for population genomics: An empirical study from an Amazonian plant species. *Molecular Ecology Resources*. https://doi.org/10.1111/1755-0998.12654

Nei, M., Tajima, F., & Tateno, Y. (1983). Accuracy of estimated phylo- genetic trees from molecular data. *Journal of Molecular Evolution*, *19*, 153–170.

Parolin, P., De Simone, O., Haase, K., Waldhoff, D., Rottenberger, S., & Kuhn, U. (2004). Central Amazon floodplain forests: Tree survival in a pulsing system. *Botanical Review*, *70*, 357–380.

Patton, J. L., da Silva, M. N., & Malcolm, J. R. (2000). Mammals of the Rio Jurua and the evolutionary and ecological diversification of Amazonia. *Bulletin of the American Museum of Natural History*, *244*, 1–306.

Patton, J. L., Dasilva, M. N. F., & Malcolm, J. R. (1994). Gene genealogy and differentiation among arboreal spiny rats (Rodentia, Echimyidae) of the Amazon Basin - a test of the riverine barrier hypothesis. *Evolution*, *48*, 1314–1323.

Peakall, R., & Smouse, P. E. (2006). GenAlEx 6: Genetic analysis in Excel. Population genetics software for teaching and research. *Molecular Ecology Notes*, *6*, 288–295.

Peres, C. A., Patton, J. L., & da Silva, M. N. F. (1996). Riverine barriers and gene flow in Amazonian Saddle-Back Tamarins. *Folia Primatology*, *67*, 113–124.

Peter, B. M. (2016). Admixture, population structure, and F-statistics. *Genetics*, *202*, 1485–1501.

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. *PLoS ONE*, *7*, e37135.

Pickrell, J. K., & Pritchard, J. K. (2012). Inference of population splits and mixtures from genome-wide allele frequency data. *PLoS Genetics*, *8* (11), e1002967.

R Core Team (2015). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from http://www.R-project.org/

Räsänen, M., Salo, J. S., Jungner, H., & Pittman, L. R. (1990). Evolution of the western Amazon lowland relief: Impact of Andean foreland dynamics. *Terra Nova*, *2*, 320–332.

Ribas, C. C., Aleixo, A., Nogueira, A. C. R., Miyaki, C. Y., & Cracraft, J. (2012). A palaeobiogeographic model for biotic diversification within Amazonia over the past three million years. *Proceedings of the Royal Society of London Series B-Biological Sciences*, *279*(1729), 681–689.

Ribeiro, J. E. L. S., Hopkins, M. J. G., Vicentini, A., Sothers, C. A., Costa, M. A. S., Brito, J. M., … Procópio, L. C. (1999). *Flora da Reserva Ducke, guia de identificação das plantas vasculares de uma floresta de terra firme na Amazônia Central*. Manaus: INPA-DFID.

Rousset, F. (1997). Genetic differentiation and estimation of gene flow from F-statistics under isolation by distance. *Genetics*, *145*, 1219–1228.

Salo, J., Kalliola, R., Hakkinen, I., Makinen, Y., Puhalla, M., & Coley, P. (1986). River dynamics and the diversity of Amazon lowland forest. *Nature*, *322*, 254–258.

Silva, C. L., Morales, N., Crosta, A. P., Solange, S., Costa, S. S., & Jiménez-Rueda, J. R. (2007). Analysis of tectonic-controlled fluvial morphology and sedimentary processes of the western Amazon Basin: An approach using satellite images and digital elevation model. *Anais da Academia Brasileira de Ciências*, *79*, 693–711.

Smith, B. T., McCormack, J. E., Cuervo, A. M., Hickerson, M. J., Aleixo, A., Cadena, C. D., … Brumfield, R. T. (2014). The drivers of tropical speciation. *Nature*, *515*, 406–409.

Solomon, S. E., Bacci, M. Jr, Martins, J. Jr, Vinha, G. G., & Mueller, U. G. (2008). Paleodistributions and comparative molecular phylogeography of leafcutter ants (*Atta* spp.) provide new insight into the origins of Amazonian diversity. *PLoS ONE*, *3*(7), e2738.

Telles, M. P. C., & Diniz-Filho, J. A. F. (2005). Multiple Mantel tests and isolation-by-distance, taking into account long-term historical divergence. *Genetics and Molecular Research*, *4*, 742–748.

Wallace, A. R. (1852). On the monkeys of the Amazon. *Proceedings of the Zoological Society of London*, *20*, 107–110.

Webb, S. D. (1995). Biological implications of the middle Miocene Amazon seaway. *Science*, *269*, 361–362.

Weir, B. S., & Cockerham, C. C. (1984). Estimating F-statistics for the analysis of population structure. *Evolution*, *38*, 1358–1370.

Worbes, M. (1997). The forest ecosystem of the floodplains. In: W. J. Junk (Ed.), *The Central Amazon floodplain* (pp. 223–265). Berlin: Springer.

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article.

How to cite this article: Nazareno AG, Dick CW, Lohmann LG. Wide but not impermeable: Testing the riverine barrier hypothesis for an Amazonian plant species. *Mol Ecol*. 2017;26:3636–3648. https://doi.org/10.1111/mec.14142