

## ECOLOGY

# Quaternary climate changes as speciation drivers in the Amazon floodplains

Gregory Thom<sup>1,2\*</sup>, Alexander T. Xue<sup>3,4,5</sup>, André O. Sawakuchi<sup>6</sup>, Camila C. Ribas<sup>7</sup>, Michael J. Hickerson<sup>3,8,9</sup>, Alexandre Aleixo<sup>10,11</sup>, Cristina Miyaki<sup>1</sup>

The role of climate as a speciation driver in the Amazon has long been discussed. Phylogeographic studies have failed to recover synchronous demographic responses across taxa, although recent evidence supports the interaction between rivers and climate in promoting speciation. Most studies, however, are biased toward upland forest organisms, while other habitats are poorly explored and could hold valuable information about major historical processes. We conducted a comparative phylogenomic analysis of floodplain forest birds to explore the effects of historical environmental changes and current connectivity on population differentiation. Our findings support a similar demographic history among species complexes, indicating that the central portion of the Amazon River basin is a suture zone for taxa isolated across the main Amazonian sub-basins. Our results also suggest that changes in the fluvial landscape induced by climate variation during the Mid- and Late Pleistocene drove population isolation, leading to diversification with subsequent secondary contact.

## INTRODUCTION

The drivers of diversification of the Amazonian biota have been intriguing naturalists and researchers since the 19th century (1). The association between patterns of phenotypic variation and environmental history, as indicated by geological and paleoclimatic data, was the starting point for developing hypotheses that sought to understand the historical processes shaping biodiversity (2). Phylogenetic and comparative phylogeographic approaches have become a key way to test these diversification hypotheses, thus becoming the main focus of the current Amazonian biogeography (3). Among the most explored processes are the effects of past climatic changes promoting speciation in Amazonia (2). After decades of debate, comparative phylogeographic studies have failed to recover consistent demographic responses across taxa in both climatically more stable and unstable sectors of Amazonia, one of the main assumptions of the refugia hypothesis (3, 4). However, recent studies have suggested a dynamic interplay between rivers and climate promoting speciation, reporting progressively more recent diversification events toward the most climatically unstable areas of the Brazilian shield (4). Notwithstanding these studies, the vast majority of phylogeographic approaches focused exclusively on terrestrial organisms restricted to the upland forests, the most widespread and diverse environment in the biome (3–5). As a consequence, the debate about the role of climate change on speciation in Amazonia has to this day been restricted to the response of upland forests to shifts from wetter to

drier climates (2, 4). Amazonia, however, is an intricate mosaic with several distinct environments that harbor endemic fauna and flora, which could have been affected differently by past climatic oscillations. Hence, the detailed underpinnings of climatic oscillations as a driving force of the Amazonian biodiversity are still unclear.

Among the main forest environments in Amazonia, the floodplains stand out, housing the largest and most diverse flooded habitats in the world, covering more than 300,000 km<sup>2</sup> and harboring high levels of species endemism—circa 10% of tree species and 15% of nonaquatic bird species are endemic (6, 7). Nevertheless, the diversification processes that originated these endemic taxa are poorly characterized, and levels of diversity are largely unknown (8, 9). This is critical since the Amazonia floodplains are threatened by agriculture activities and hydroelectric dams, affecting several populations of endangered endemic species, eliciting conservation concerns and, thus, a need for further research (10). The first phylogeographic studies of floodplain species supported a scenario of widely distributed populations with low levels of genetic diversity and lack of population structure (8, 9, 11). This pattern reinforced the assumption that the linear connectivity of floodplain habitats enables high levels of gene flow over the entire distribution of floodplain taxa, which, in turn, are expected to have high dispersal capacity due to the highly dynamic environments (6). However, community composition turnover previously described for several taxonomic groups suggests a strong biological transition in central Amazonia floodplains that has not been properly explored with genetic data (7, 12).

Recent studies on antbird species associated with river edge forests revealed distinct genetic clusters restricted to the large Amazon River tributaries and a recent scenario of diversification potentially linked to climatic oscillations (13, 14). A flickering mechanism of the Amazonian floodplains over time could shape genetic diversity, genetic structure, and connectivity between populations, which may lead to speciation (15). In this case, rainfall variation and sea level changes associated to past climatic phases could change the erosion base level and sedimentary budget of rivers. This could then affect sediment erosion-storage balance and drive the retraction or expansion of floodplains, potentially hindering the distribution of river-created environments during specific periods in the past (14, 16, 17). From

<sup>1</sup>Departamento de Genética e Biologia Evolutiva, Universidade de São Paulo, Rua do Matão, 277, Cidade Universitária, São Paulo, SP 05508-090, Brazil. <sup>2</sup>Department of Ornithology, American Museum of Natural History, New York, NY 10024, USA. <sup>3</sup>Department of Biology, Marshak Science Building, City College of New York, New York, NY 10031, USA. <sup>4</sup>Human Genetics Institute of New Jersey and Department of Genetics, Rutgers University, 604 Allison Road, Piscataway, NJ 08854, USA. <sup>5</sup>Simons Center for Quantitative Biology, Cold Spring Harbor Laboratory, One Bungtown Road, Cold Spring Harbor, NY 11724, USA. <sup>6</sup>Instituto de Geociências, Universidade de São Paulo, Rua do Lago 562, São Paulo, SP 05508-080, Brazil. <sup>7</sup>Instituto Nacional de Pesquisas da Amazônia (INPA), Av. André Araújo 2936, Manaus, AM 69060-001, Brazil. <sup>8</sup>The Graduate Center, City University of New York, New York, NY 10016, USA. <sup>9</sup>New York Division of Invertebrate Zoology, American Museum of Natural History, New York, NY 10024, USA. <sup>10</sup>Museu Paraense Emílio Goeldi (MPEG), Caixa Postal 399, Belém, PA 66040-170, Brazil. <sup>11</sup>Finnish Museum of Natural History, University of Helsinki, Helsinki, Finland.

\*Corresponding author. Email: biogreggo@yahoo.com.br

a phylogeographic perspective, we expect to observe continuously or disrupted distribution ranges, with recently diverged [ $<0.8$  Ma; (16)] populations and high levels of gene flow due to secondary contact, given the relatively recent expansion of floodplains during the Holocene (18). In addition, given that a uniform continental-scale climatic change tends to affect populations simultaneously, we expect synchronic expansions of effective population sizes. Alternatively, the genetic diversity of floodplain populations could be subjected to the effects of other nonmutually exclusive processes that could have produced more pronounced effects than past climatic oscillations: (i) riverine isolation by distance (IBD), emerging from the river network architecture and extensive distribution of floodplain habitats that could generate gradients of genetic diversity correlated with geographic distance (19). In this case, we expect continuous distribution of the species along its entire geographic range and a gradual transition in patterns of genetic diversity and genetic structure matching an IBD model. (ii) Ecological gradients, which could emerge from spatially segregated habitats due to geomorphological and biogeochemical heterogeneities of rivers generating local adaptation of populations from distinct environments and leading to a reduction in gene flow. For example, geomorphological and hydrological conditions of the watersheds determine properties such as the water pH, concentration of suspended sediments, and amplitude of the flood pulse, resulting in floodplains with different environmental characteristics (20). In Amazonia, watersheds with headwaters in the Andes give origin to “várzea,” floodplains of sediment-rich whitewater rivers, while watersheds with headwaters in the shield areas produce “igapó,” floodplains of sediment-poor black/clearwater rivers. Therefore, abrupt environmental transitions can occur in floodplains of the Amazon River basin. In this scenario, we expect species continuously distributed over its entire geographic ranges with gradual or abrupt transition in patterns of genetic diversity and genetic structure due to reduced but continuous gene flow matching areas of ecological gradients, such as different types of floodplain forests (i.e., várzea versus igapó). (iii) Long-term physiographic changes within the Amazon River basin, which could cause rearrangement and structuring of the courses of the main Amazonian rivers since the Miocene and particularly during the Pleistocene, potentially allowing dispersal into new environments with subsequent diversification (21). Here, we could expect continuous or disrupted distribution ranges, divergence times overlapping with geologically more dynamic periods, and spatial congruence with known geological features, as past boundaries between watersheds and abandoned river channels as well as the absence or reduced levels of gene flow due to allopatry. Although these four described processes do not represent all potential mechanisms capable of affecting the diversification of floodplain species, they are fundamental to understand how the Amazonian floodplain forest biota emerged and diversify.

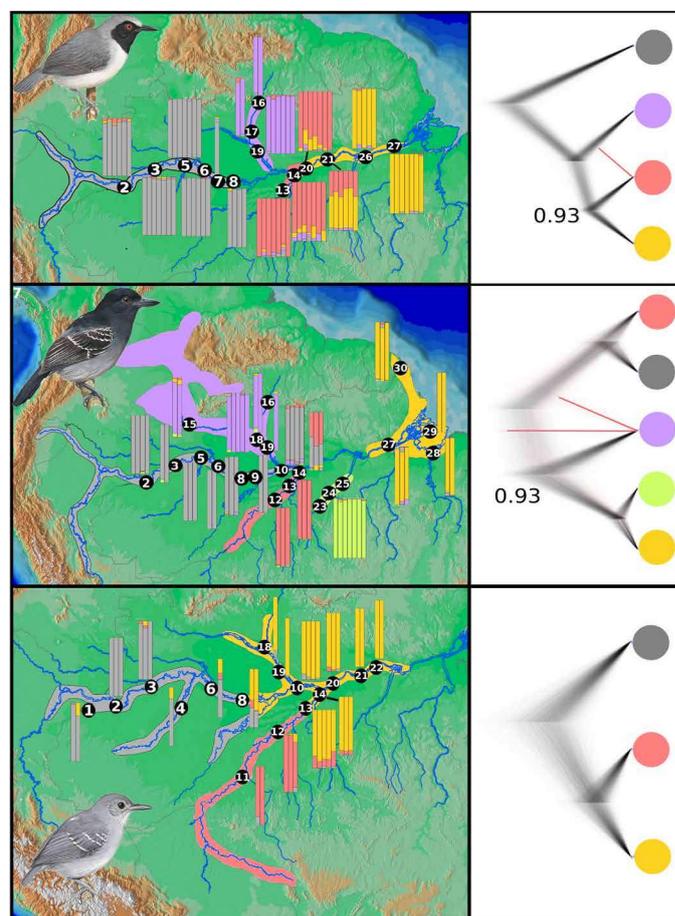
In the present study, we conducted a comparative phylogenomic analysis to elucidate if regional environmental changes affected floodplain forest birds. Specifically, we tested diversification hypotheses on widely distributed species complexes of antbirds, which comprise 12 subspecific taxa with overlapping geographic distributions. This is the first study to report the populational structure of various Amazonian floodplain species based on a comparative population genomic exploration of codiversification. The findings reported here support a similar pattern of diversification across species complexes, including area relationships, current gene flow among populations, and timing of divergence, as well as synchronous

population size changes. Our results suggest that Pleistocene climate change, which affected the fluvial sediment budget and floodplain habitat continuity, had a potentially important role in shaping the current genetic diversity of these endemic bird species. This mechanism possibly operates in an analogous way to the refugia hypothesis (2), with recurrent reductions in the availability of floodplain environments in the central portion of the basin, reinforcing the idea that paleoclimatic oscillations are intimately linked to speciation in Amazonia.

## RESULTS

### Summary of results, genetic structure, and phylogenetic relationships

Here, we selected three bird species complexes, namely, *Myrmoborus lugubris*, *Thamnophilus nigrocinereus*/*Thamnophilus cryptoleucus*, and *Myrmotherula assimilis*, distributed exclusively along floodplain forests of large Amazonian rivers (Fig. 1). Our sampling design includes most of the geographic distribution of the three species and



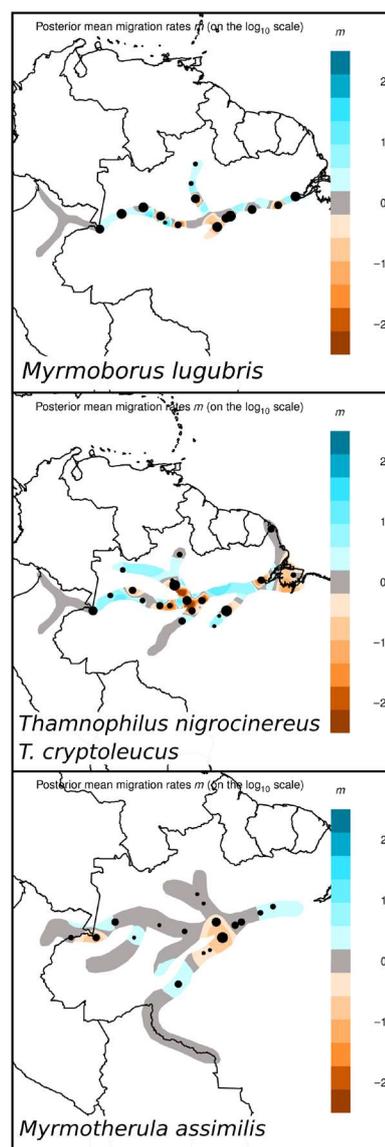
**Fig. 1. Geographic distribution, population structure, and species tree of studied lineages.** Numbers in the map are localities as in table S12. Barplots represent the coefficient of ancestry per individual obtained with sNMF. Numbers in the species tree are posterior probabilities  $<0.95$ , and all other nodes presented maximum posterior probabilities. The red lines in the species tree represent different topologies, sampled in the MCMC. Top: *M. lugubris*. Middle: *T. nigrocinereus*/*T. cryptoleucus*. Bottom: *M. assimilis*. Gray, Solimões population; pink, Madeira population; green, Tapajós population; purple, Negro population; yellow, Amazonas population (Amazonas + Negro for *M. assimilis*).

includes all named subspecific taxa. To obtain genetic variants, we deploy the sequence capture of thousands of ultraconserved elements (UCEs) (22). For a comprehensive description of the main results obtained with the implemented bioinformatics pipeline, see “Summary of results” section in the Supplementary Materials and table S1. To detect population structure across the Amazon River basin, we used the software sNMF and *k*-means clustering. The software sNMF can process large biallelic datasets efficiently with good accuracy, even in scenarios of departure from the Hardy Weinberg and linkage equilibria (23). To explore the spatial distribution of genetic diversity, testing for the effects of IBD, we used the estimated effective migration surface (EEMS), which maps genetic differentiation by estimating effective migration surface among demes based on a spatially explicit approach (24). All three analyses supported a scenario of population structure within sectors along the Amazon River mainstem and its main tributaries for the three species complexes studied (Figs. 1 and 2 and “Summary of results” in the Supplementary Materials). The number of ancestral populations (Fig. 1) was relatively concordant with the geographic distribution of described subspecies. The EEMS results supported considerable reductions in effective migration within the central portion of the Amazon River basin (Fig. 2). In the same region, differentiated populations are in contact, and individuals with mixed ancestry coefficients are predominantly located for all three species complexes. A scenario of secondary contact is indicated, as the effective genetic diversity in the central portion of the basin was higher than the average, indicating higher genetic dissimilarity among individuals (fig. S1). Although these patterns are broadly shared across the three species complexes, there were species idiosyncrasies. For example, the number and geographic placement of these phylogeographic breaks exhibited minor differences across species, and there is even further variation in the geographic location of individuals with mixed ancestry, as well as the ancestry sources and relative coefficient proportions within these individuals (Fig. 1).

Species trees estimated with SNAPP and TREEMIX revealed distinct splitting patterns of populations from the three species complexes with unsupported nodes for *M. lugubris* and *T. nigrocinereus/T. cryptoleucus* (Fig. 1). TREEMIX also indicated the presence of significant migration edges among populations in all three species complexes (Fig. 3 and “Summary of results” in the Supplementary Materials). A topology test performed using Fastsimcoal2 (FSC) supported a relationship different from that estimated with SNAPP only for *M. lugubris*, placing Negro and Madeira populations as sisters, as supported by previous studies (13) (table S2). The divergence pattern we observed indicated that the populations from the Solimões River were the first to split in *M. lugubris* and *M. assimilis*, while in *T. nigrocinereus/T. cryptoleucus*, the first divergence was between the ancestral populations of the Solimões and Madeira and the ancestral populations of the eastern Amazonia (tables S3 to S5).

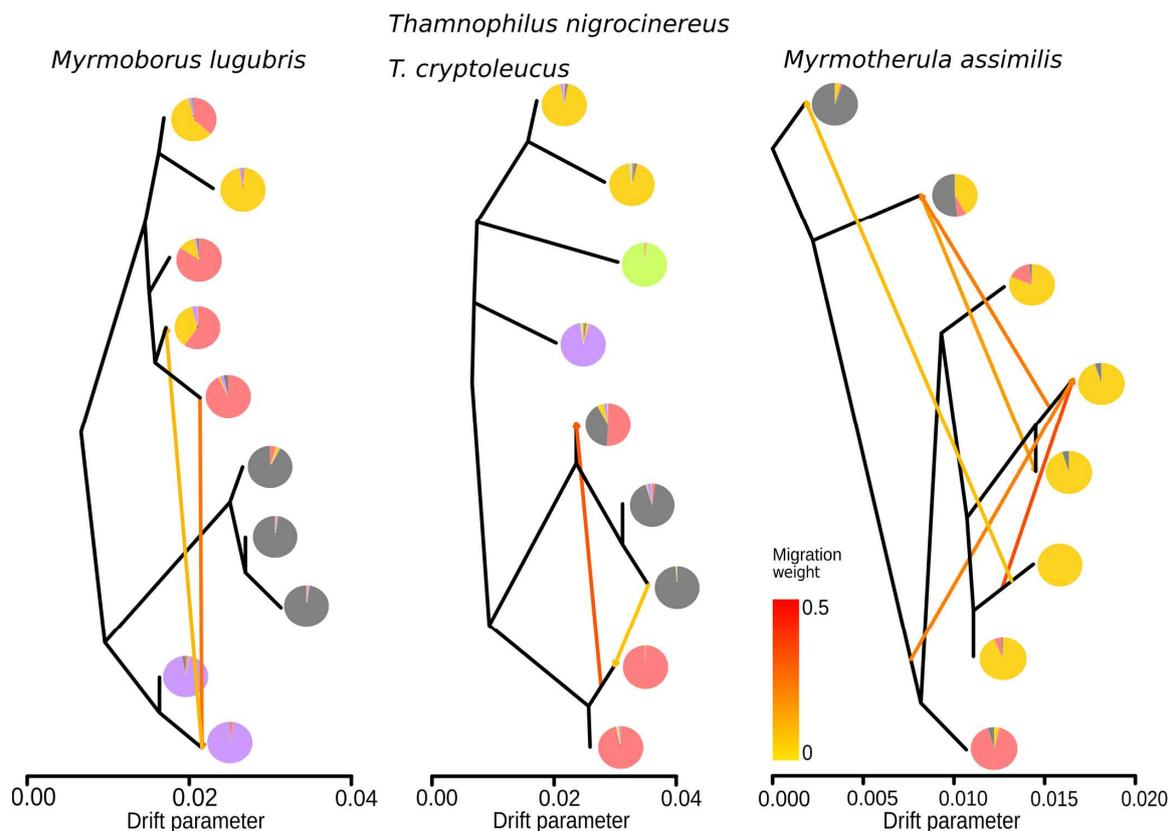
### Demographic parameter estimation

To estimate gene flow among populations, divergence time, and population size and to test demographic models associated with our hypotheses, we implemented three distinct coalescent-based approaches, Generalized Phylogenetic Coalescent Sampler (G-PhoCS) (25), FSC (26), and the Multiple Taxa Demographic Inference of Congruency in Events (Multi-DICE) (27). G-PhoCS is a full likelihood method that uses as input loci sequences containing all the variable sites (25). FSC approximates the likelihood of a demographic model, based on simulations, reducing the complexity of genomic data by using the site frequency spectrum as a summary statistic calculated from a randomly



**Fig. 2. EEMS for *M. lugubris*, *T. nigrocinereus/T. cryptoleucus*, and *M. assimilis*.** Color bar shows the effective migration rate on a log<sub>10</sub> scale relative to the average migration rate over the entire range of the corresponding species. The darker the blue, the higher the average effective migration rate. The darker the orange, the lower the average effective migration rate. The size of the black circles represents the number of sampled individuals in a given locality.

selected single nucleotide polymorphism (SNP) per locus (26). Multi-DICE tests for the degree of synchrony in size changes among populations through hierarchical codemographic modeling (27). The divergence times and migration rates [means and 95% highest posterior density (HPD)] estimated with the isolation-migration model of G-PhoCS (Fig. 4) were highly consistent among preliminary and final runs, despite the low effective sample size and high auto correlation time of some parameters (tables S3 to S5). Posterior mean divergence time for all three species ranged from 25,320 to 225,960 years ago (tables S3 to S5). The estimated means for the first divergence in each species complex were 107,880, 189,640 and 225,960 years ago for *M. lugubris*, *T. nigrocinereus/T. cryptoleucus*, and *M. assimilis*, respectively. Migration values varied considerably between pairs of populations



**Fig. 3. Population relationship graphs for *M. lugubris*, *T. nigrocinereus*/*T. cryptoleucus*, and *M. assimilis* obtained with TREEMIX. Colors represent populations as in Fig. 1. Pie charts represent the percentage of coefficient of ancestry obtained with sNMF.**

(tables S3 to S5), with higher migration bands related to populations occurring in whitewater rivers that are in contact in the central portion of the Amazon River basin (tables S3 to S5). The FSC results supported demographic models that include gene flow for all three species complexes (Tables 1 and 2 and tables S5 to S7). For *M. lugubris* and *T. nigrocinereus*/*T. cryptoleucus*, isolation with gene flow was the best-fit demographic model based on Akaike's relative weights, with all divergence times fixed on values estimated by G-PhoCS and instantaneous population size changes after divergence (Fig. 4, C and D, and Table 1). For *M. assimilis*, isolation with gene flow but without population size change was the highest supported demographic model with an Akaike's relative weight of 0.97 (Fig. 4E and Table 1). Parameter estimates based on the best-fit models for  $N_e$  were relatively congruent with the G-PhoCS results, as most of the estimates differed by no more than twofold between the two different methods (Table 2, tables S3 to S5, "Full versus composite likelihood" section in the Supplementary Materials).

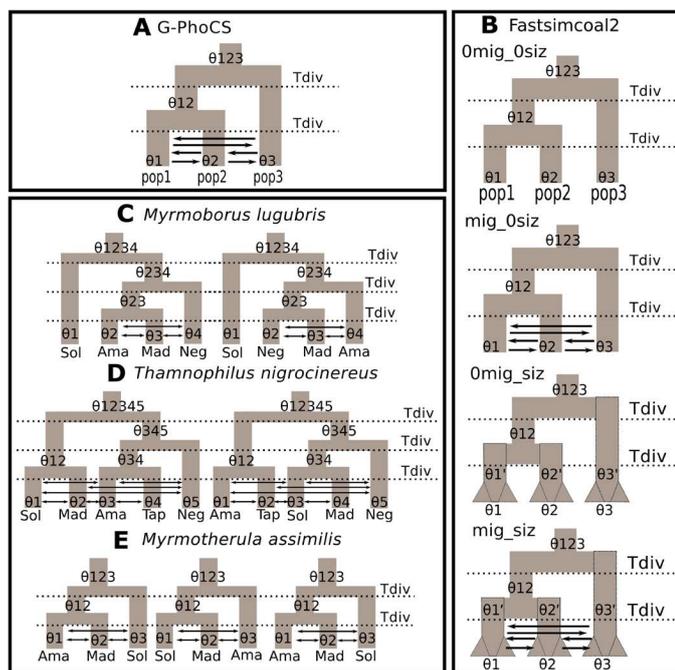
We explored the level of synchronicity of population size changes, as predicted by a scenario affected by climatic oscillation, with the R package Multi-DICE (27). The approximate Bayesian computation (ABC) model selection procedure supported 10 (of 15) populations with expansion as the best-fit model [posterior probability (PP) > 0.99] with intensities varying from ~5- to ~11-fold (based on mean PP), occurring from ~8000 to ~110,000 years ago (table S8 and figs. S2 and S3). When testing for the synchronicity of these expansions in the hierarchical codemographic approach of Multi-DICE, a high proportion (mean  $\zeta = 0.9301$ , with a mean cross-validation correla-

tion,  $r = 0.8356$ ; fig. S4 and tables S10 and S11) of the tested populations expanded at the same time, approximately ~0.2 Ma ( $r = 0.7325$ ; 95% HPD = 0.084 to 0.295 Ma).

## DISCUSSION

### Population structure along the Amazonian floodplains

The large Amazonian rivers have long been recognized as effective barriers impeding gene flow in upland forest birds and mammals and ultimately linked to the diversification of species complexes (3, 5). In contrast to this general prediction, we did not observe any sign of genetic differentiation of floodplain birds across opposite margins or between islands and margins but rather found a shared pattern of genetic structure associated with sectors along the Amazon River mainstem and its main tributaries. Our results suggest that the central portion of the Amazon River catchment, specifically in the Amazon River mainstem between the mouths of the Negro and Madeira Rivers, represents a suture zone across codistributed floodplain avian taxa, with a considerable number of phylogeographic breaks and hybridization events (12, 13, 28). Populations from the middle and upper courses of the Solimões, Madeira, Negro, and Tapajós Rivers, as well as from the lower course of the Amazon River, are completely or partially isolated from each other without a clear physical barrier across river channels (Fig. 1 and tables S2 to S4). This pattern conflicts with previous studies of floodplain bird species, which supported widely distributed populations without genetic structure over the entire Amazon River basin (9, 11).



**Fig. 4. Graphical representation of the explored demographic models.** Isolation-migration model implemented on G-PhoCS (A). Demographic models simulated in FSC (B). Note that for *M. lugubris* and *T. nigrocinereus*/*T. cryptoleucus*, these models included four and five populations, respectively. Names of the models are described in Table 1. The model mig\_siz\_ALL\_G-PhoCS is not depicted in the panel. Topologies tested with Fastsimcoal2 for *M. lugubris* (C), *T. nigrocinereus*/*T. cryptoleucus* (D), and *M. assimilis* (E).  $\theta$ , effective population size; Tdiv, divergence time; Sol, Solimões population; Mad, Madeira population; Neg, Negro population; Ama, Amazonas population; Tap, Tapajós population.

A recent study based on a large set of taxa suggested that floodplain species tend to present lower levels of genetic diversity and genetic structure than upland forest species (8). However, most of the studied taxa associated to floodplains also occupy other types of habitats, including disturbed areas. In contrast, our results support that habitat specificity affects connectivity among floodplain populations, causing river edge forest specialists to have pronounced genetic structure, in a similar fashion to upland forest organisms. The three species complexes studied here are part of a selected group of birds that show clear phenotypic diagnoses geographically structured over floodplains, while the vast majority of other floodplain species do not have clear intraspecific phenotypic variation across their entire distributions (12). However, as supported by our data (e.g., the Madeira population of *M. assimilis*), the Amazonian floodplains might harbor a high number of cryptic lineages, and future studies might support similar patterns to the one reported here, with genetic structure compartmentalized in different Amazonian rivers. Previous studies measuring the alpha diversity of floodplain forest sites between western (Solimões River) and eastern (Amazon River) Amazonia (7, 12), reported a strong community transition also in the central portion of the Amazon River basin, in the sector delimited by the mouths of the Negro and Madeira Rivers. Hence, despite the lack of intraspecific phenotypic variation in most species, these taxa are still restricted to specific sectors of the floodplains (e.g., Solimões River basin).

### Past climate change, fluvial sedimentation, and the diversification of Amazonian floodplain taxa

We found that despite potential intrinsic ecological differences between the three species complexes, their diversification patterns are similar, including topology, divergence time, gene flow, and historical demography (Fig. 1, Table 2, and tables S3 to S5). The diversification hypotheses we explored here are not mutually exclusive, given that the four proposed processes could have simultaneously affected the diversification of the species examined. However, our model-based approach and additional empirical analyses allowed us to explore which of the four scenarios had the most pronounced effect in driving the observed genetic diversity of the three species complexes. Despite the potential strong effects of IBD reported for taxa with linear distributions (19), the results obtained with the EEMS (24), the pattern of a genetic structure, and the best demographic model selected for each species complex cannot be explained solely by IBD. Another important factor that could affect the genetic diversity of floodplain populations is the heterogeneous distribution of habitats, which could generate local adaptation, thus reducing gene flow and, therefore, promoting ecological speciation. For example, the transition between sediment-rich whitewater versus sediment-poor black/clearwater rivers promotes the formation of várzea and igapó habitats, respectively, and affects allele frequency of fish species (20). However, the allopatric distribution of some populations, the synchronous demographic expansions in most of the populations occurring in whitewater rivers, and the variable levels of gene flow in each species complex (Fig. 1 and tables S2 to S4) indicate that ecological gradients solely cannot fully explain the observed pattern of genetic diversity. Last, the initial diversification of the three taxa studied here took place during the Mid- and Late Pleistocene, after the major reorganization of Amazonian tributaries and establishment of the current transcontinental Amazon River (29). This suggests that more recent and recurrent changes in riverine habitats produced the strongest signature in our data.

In the Amazon River basin, an abrupt increase in production and variability of sediment supply was recorded during the Pliocene and Pleistocene, which seems to be related to major climate shifts (30). Hence, enhanced availability and variability in sediment supply possibly had a severe impact on fluvial landscapes and favored the expansion of dynamic river-created environments such as floodplains along major Amazonian rivers. For example, variations in fluvial discharge due to climate changes in western Amazon allowed larger floodplains in the Solimões River during the Mid- to Late Pleistocene transition, and floodplains retreat during the final stages of the Late Pleistocene (18). This climate-induced environmental change coupled with the low dispersal capacity of antbirds and intimate relationship with the river edge forests (6) is congruent with the pattern of isolation and connectivity among populations. Historical expansion of floodplains along the Solimões-Amazon mainstem could have enabled dispersal and colonization events to new environments, mainly from the western, with long-term conditions for sediment accumulation (18), to the eastern Amazon River, with prevailing sediment bypassing conditions. Our results corroborate this scenario, indicating that the oldest splits in each complex are associated with the divergence of western Amazonia populations (17, 21), analogous to what has been found for upland terra-firme species as also a result of more stable humid conditions in western Amazonia (4).

After the expansion and establishment of a floodplain belt along the Solimões-Amazon mainstem, higher amplitude of orbital climate

**Table 1. Composite likelihood [max ln(L)], Akaike information criterion (AIC), and relative contribution (weights) for each of the demographic models tested for *M. lugubris*, *T. nigrocinereus*/*T. cryptoleucus*, and *M. assimilis*.** In bold: best model. 0mig\_0siz, demographic model assuming no migration between populations and no population size changes; 0mig\_siz, demographic model assuming no migration between populations but with population size changes; mig\_0siz, demographic model assuming migration between populations but no population size changes; mig\_siz, demographic model assuming migration between populations and population size changes; mig\_siz\_ALL\_G-PhoCS, demographic model assuming migration between populations and population size changes with all divergence times fixed on G-PhoCS mean estimations (tables S3 to S5); 2\*nparam, number of parameters in the models.

Taxon	Model	Max ln(L)	nparam	AIC	Relative weights
<i>M. lugubris</i>	0mig_0siz	-15,003.699	9	69,112.59	$2.72 \times 10^{-201}$
	0mig_siz	-14,986.31	17	69,048.51	$2.23 \times 10^{-187}$
	mig_0siz	-14,880.81	21	68,570.66	$1.29 \times 10^{-83}$
	mig-siz	-14,804.56	29	68,235.52	$7.72 \times 10^{-11}$
	<b>mig_siz_ALL_G-PhoCS</b>	<b>-14,795.316</b>	<b>27</b>	<b>68,188.95</b>	<b>1.00</b>
<i>T. nigrocinereus</i> / <i>T. cryptoleucus</i>	0mig_0siz	-17,880.913	12	82,368.65	$9.37 \times 10^{-239}$
	0mig_siz	-17,790.154	22	81,970.69	$2.44 \times 10^{-152}$
	mig_0siz	-17,777.802	26	81,921.8	$1.00 \times 10^{-141}$
	mig-siz	-17,672.851	36	81,458.49	$4.08 \times 10^{-41}$
	<b>mig_siz_ALL_G-PhoCS</b>	<b>-17,633.765</b>	<b>33</b>	<b>81,272.49</b>	<b>1.00</b>
<i>M. assimilis</i>	0mig_0siz	-3,415.642	6	15,741.61	$5.25 \times 10^{-66}$
	0mig_siz	-3,398.839	12	15,676.23	$8.27 \times 10^{-52}$
	<b>mig_0siz</b>	<b>-3,347.767</b>	<b>12</b>	<b>15,441.04</b>	<b><math>9.76 \times 10^{-01}</math></b>
	mig-siz	-3,346.776	18	15,448.47	$2.37 \times 10^{-02}$
	mig_siz_ALL_G-PhoCS	-3,352.539	16	15,471.01	$3.02 \times 10^{-07}$

cycles since 0.8 Ma would have intensified the variation of sediment production in the Andes, the main source of sediments to build floodplains, as well as of water discharge that drives inland erosion of base levels of Amazonian rivers. This climate-sedimentary dynamics modulates the sediment budget of the Amazon River basin and potentially produced several phases of floodplain expansion, through sediment aggradation, or contraction, due to channel incision (17, 18, 31). Furthermore, the sea level fall of more than 100 m recorded during the Last Glacial Maximum (LGM; ~21 ka) could have induced incision of the lower reach of the Amazon River mainstem (16), leading to the erosion and fragmentation of floodplains in the eastern Amazon River.

The scenario we propose here is a tentative explanation for the diversification of the three studied species, given that it is not possible to establish a direct link between divergence times and timing of specific landscape changes so far. Nevertheless, the time scale of divergence time estimates is compatible with orbital (17) and millennial rainfall variations (31, 32) that affected the sedimentary dynamics responsible for the expansion and contraction of Amazonian floodplains during the Mid- and Late Pleistocene (18). Abrupt rainfall increase occurred in the Amazon during Heinrich Stadials (HS) at ~17 ka (HS1), ~24 ka (HS2), ~31 ka (HS3), ~38 ka (HS3), ~48 ka (HS4), and ~60 ka (HS5) (31), with significant effect over the suspended sediments sustaining floodplains of the Solimões-Amazon mainstem (33). The combined effect of recurrent millennial events of increased precipitation such as the HS can result in significant changes in the distribution of flooded environments as the retreat of floodplains in the Solimões River during the Late Pleistocene (18). The higher water discharge induced by rainfall increase favors channel incision and induces sediment bypassing and floodplain erosion downstream the confluence of large rivers, potentially causing

fragmentation of floodplain environments and of populations associated with them. More recent conditions suitable for sediment trapping in the fluvial system and expansion of floodplains started during the Holocene (18), after the LGM and HS1. The Holocenic expansion (<11.7 ka) of floodplain environments could have allowed the dispersal of forest specialists across the Solimões-Amazon mainstem, leading to secondary contact among previously isolated populations, mainly through the lowermost parts of whitewater tributaries such as the Solimões and Madeira Rivers, where we detected higher levels of gene flow (Table 2 and tables S3 to S5), and recent evidence supports the presence of hybrid zones (13). Similarly, demographic changes were only detected in populations from sediment-rich whitewater rivers, mostly subtle expansions (table S9), suggesting that the higher sedimentary dynamics of whitewater rivers seems to have produced more unstable floodplain environments when compared with sediment-deficient clear/blackwater rivers (14). Similarly, the synchronous population size expansions detected here are congruent with continental-level oscillations, affecting the whole Amazonian basin at once. However, future studies are needed to improve the chronology of fragmentation and merging phases of Amazon floodplains and to allow more precise comparison between events of speciation and fluvial landscape changes.

In the modern Amazonian fluvial system, gaps in floodplain forest distribution occur in the downstream sectors of the Negro and Tapajós Rivers (16). In these rivers, the low amount of sediments derived from upstream sectors accumulate in the ria head, forming fluvial archipelagos, while the lower reaches become sediment deficient, impeding the development of floodplains. These areas lacking continuous floodplains can be up to 150 km long, matching the distribution limit of populations restricted to these rivers (except for

**Table 2. Parameter values of simulated models in FSC with the best composite likelihood for each of the three studied species and 95% confidence interval (in parentheses) obtained with parametric bootstraps.** Models are described in Fig. 4. mig\_0siz, demographic model assuming migration between populations but no population size changes; mig\_siz\_ALL\_G-PhoCS, demographic model assuming migration between populations and population size changes with all divergence times fixed on G-PhoCS mean estimations (tables S3 to S5); Ne, effective population size (number of diploid individuals). Ancestral population sizes follow the phylogenetic relationship between populations; n-m, ancestral population of Negro and Madeira; n-m-a, ancestral population of Amazonas, Negro, and Madeira; s-m, ancestral population of Solimões and Madeira; a-t, ancestral population of Amazonas and Tapajos; a-t-n, ancestral population of Amazonas, Tapajos, and Negro; m-a/n, ancestral population of Madeira and Amazonas/Negro; anc, ancestral population of the species complex; Mx → y, estimated migration band from x to y forward in time (number of individuals per generation); Tdiv, divergence time in years; s, Solimões population; m, Madeira population; n, Negro population; a, Amazonas population; t, Tapajos population; Epsilon, the ratio between the ancestral population size in relation to the current one when size change was allowed; tau, time of population size change.

Taxon	Model	Parameters									
<i>M. lugubris</i>	mig_siz_ALL_G-PhoCS	Ne Solimoes	Ne Negro	Ne Madeira	Ne Amazonas	Ne n-m	Ne n-m-a	Ne anc	Ms→n	Mn→s	Ms→m
		577,831 (139,204–297,980)	187,753 (88,058–211,804)	41,753 (20,230–41,237)	64,241 (20,490–41,960)	592,914 (303,123–541,514)	402,991 (247,756–635,255)	88,469 (86,784–552,158)	0.04 (0.00–0.24)	0.50 (0.64–2.56)	0.02 (0.00–1.49)
		Mm→s	Ms→a	Ma→s	Mn→m	Mm→n	Mn→a	Ma→n	Mm→a	Ma→m	Epsilon s
		0.02 (0.01–0.23)	1.55 (0.00–3.70)	0.07 (0.01–0.10)	1.67 (0.03–4.68)	0.003 (0.00–0.04)	1.88 (0.00–4.35)	0.001 (0.002–0.08)	0.10 (0.09–6.39)	5.77 (0.00–10.49)	0.22 (0.22–1.38)
		Epsilon n	Epsilon m	Epsilon a	Tau s	Tau n	Tau m	Tau a			
		3.96 (0.65–8.74)	5.50 (3.65–8.49)	0.3 (0.28–5.18)	20,203 (18,539–105,746)	21,632 (2,545–59,913)	20,255 (11,054–31,161)	20,025 (2,489–23,263)			
<i>T. nigrocinereus</i>	mig_siz_ALL_G-PhoCS	Ne Solimoes	Ne Negro	Ne Madeira	Ne Amazonas	Ne Tapajos	Ne s-m	Ne a-t	Ne a-t-n	Ne anc	Ms→n
		231,432 (88,960–238,231)	196,767 (50,976–178,292)	127,012 (40,956–134,760)	100,777 (30,694–77,297)	116,840 (20,538–141,563)	311,298 (63,291–407,013)	447,604 (312,473–707,007)	211,583 (85,335–550,942)	356,013 (93,022–534,331)	1.36 (1.02–2.48)
		Mn→s	Ms→a	Ma→s	Ms→m	Mm→s	Mn→a	Ma→n	Mn→m	Mm→n	Ma→m
		0.52 (0.25–2.33)	0.064 (0.00–1.88)	1.15 (0.24–4.18)	0.30 (0.00–0.14)	1.04 (0.68–3.08)	0.03 (0.00–1.55)	0.57 (0.17–1.13)	0.22 (0.00–0.69)	0.08 (0.00–0.62)	0.01 (0.00–0.43)
		Mm→a	Ma→t	Mt→a	Epsilon s	Epsilon n	Epsilon m	Epsilon a	Epsilon t	Tau s	Tau n
		0.00 (0.00–5.46)	0.26 (0.06–0.51)	0.01 (0.00–1.52)	0.27 (0.38–5.88)	1.09 (2.01–7.15)	6.27 (3.62–15.70)	4.43 (1.43–18.93)	0.34 (0.22–5.06)	79,733 (34,392–138,714)	60,520 (7,006–113,621)
		Tau m	Tau a	Tau t							
		19,378 (3,871–18,503)	11,112 (1,429–12,285)	13,927 (9,354–46,256)							
<i>M. assimilis</i>	mig_0siz	Ne Solimoes	Ne Amazonas/Negro	Ne Madeira	Ne m-a/n	Ne anc	Mm→s	Ms→m	Mm→a/n	Ma/n → m	Ms→a/n
		358,432 (193,676–396,955)	525,318 (505,290–570,850)	765,488 (778,199–1,322,214)	195,303 (841,052–960,669)	249,877 (635,438–960,239)	0.34 (0.00–0.14)	0.20 (0.00–0.81)	3.05 (5.53–8.63)	3.94 (0.00–0.58)	2.37 (3.02–6.79)
		Ma/n→s	Tdiva/n-m								
		3.07 (3.93–5.74)	166,612 (24,182–173,088)								

*M. assimilis*; tables S3 to S5) (13, 14). Hence, we have robust evidence suggesting that current gaps in the distribution of floodplain environments, which were created during past climatic oscillations, are affecting the connectivity between some of the studied taxa, promoting allopatry.

Zones with narrow and unstable floodplains can also occur along the Amazon River mainstem. In this case, channel sectors combining higher water discharge and lower sediment supply induce erosion and sediment bypass, hindering the building of long-lasting substrates for floodplain forests. This condition results from the confluence of

large rivers with contrasting sediment loads such as the Negro and Solimões Rivers (34) and explains the narrow distribution of the floodplains in the Amazon mainstem between the Negro and Madeira River mouths. The high water supply and low sediment load from the Negro river promote erosion and sediment bypassing conditions in the Amazon mainstem until the sediment input from the Madeira River (35), restoring suitable conditions for the development of large and stable floodplains in the Amazon River mainstem. The current configuration of this region suggests that it could be a putative zone for interruptions on the distribution of river edge forests along the mainstem of the Amazon River basin during past climatic conditions. Our genomic results corroborate this idea, suggesting that this region seems to have faced alterations in the floodplain environment distribution, simultaneously isolating and contracting populations toward the middle and upper course of the Solimões, Negro, and Madeira Rivers from those of the lower course of the Amazon River. Landscape genetics studies focusing on species restricted to river edge forests that applied spatial explicit methods to estimate the directionality of geographic expansion could produce new evidence supporting this scenario.

The large temporal gap between stem and crown divergences of each lineage also corroborates the dynamism of floodplains driving diversification (8, 9, 13). As all three species complexes diverged from their sister taxa at least 5 Ma ago (36) and within each lineage, population divergence started only in the Mid- and Late Pleistocene. Therefore, there is a time window of several million years without diversification, indicating high levels of population extinctions or less opportunity for speciation when compared with the Amazonian upland species (8). However, the latter of these two hypotheses is less likely given the dynamic (expansion and retraction) history of the Amazonian floodplains (18, 21). Thus, more pronounced changes in the sedimentation dynamics of the rivers could recurrently extinguish populations, especially from the more unstable eastern Amazonia, which could therefore explain the long branches typically estimated for organisms restricted to the floodplains (8, 9, 13).

### Diversification in the Amazonian basin

The effects of paleoclimatic changes on Amazonian diversification have been long discussed (2). Recent evidence, based on upland forest organisms, has supported a complex scenario of diversification driven by interactions between climatic fluctuations and drainage evolution. Under this scenario, climate modulates environmental dynamics, leading to isolation and dispersion events, with rivers acting as semipermeable barriers to gene flow (4). Additional studies agree with the potential oscillations on the barrier effect of large rivers, which operate as a primary or secondary barrier for different species (5). Here, we propose a concordant scenario, with climate-driven river dynamics through time. We suggest that Pleistocene climatic variations potentially drove diversification in Amazonian floodplain taxa through a process analogous to the refugia hypothesis: the contraction or interruption of river edge forest resulting in isolated blocks restricted to the main Amazonian rivers. However, different from the original refugia hypothesis, fragmentation of river edge forests is controlled by rainfall influence on shifts of fluvial sedimentation, modulating the building and erosion of substrates colonized by Amazonian floodplain forests. Long-lasting changes in the landscape of floodplains, which are potential drivers of speciation, result from a combination of climate variation (rainfall shifts) and fluvial processes (e.g., meeting of rivers with contrasting sediment loads).

The combined interpretation of the diversification patterns of upland and floodplain forest (river-edge forests) organisms can bring new insights to understand the history of Amazonia as a whole. The Amazonian floodplains are highly dynamic and have been constantly affected by orbital ( $10^4$  to  $10^5$  years) and abrupt millennial ( $10^3$  years) climate variations during the Pleistocene (33). The region along the Amazon River mainstream between the mouths of the Negro and Madeira Rivers was probably very dynamic, not sustaining long-lasting floodplains, and was a potential periodic barrier for populations restricted to floodplain environments. This scenario also implies that periods of floodplain contraction could facilitate dispersal of upland forest taxa across large rivers, leading to colonization and fast diversification due to strong founder effects or allowing gene flow between partially isolated populations. In this scenario, comparative population genomic studies could test for potential signals of dispersal across rivers for upland forest lineages, coupled with disconnections among populations from floodplain environments, and explore whether this mechanism could fit a climate-driven species pump model (4).

## MATERIALS AND METHODS

### Sequence capture of UCEs and bioinformatics

We selected 80, 56, and 38 samples of *M. lugubris*, *T. nigrocinereus*/*T. cryptoleucus*, and *M. assimilis*, respectively. They represent all the 12 described subspecies, covering most of the geographic distribution of these taxa (table S12 and Fig. 1). For a detailed description of the study groups and river edge forests, see the “Study groups and floodplain environments” section in the Supplementary Materials. To obtain genome-wide data, we used sequence capture of UCEs (22). DNA extraction and quantification, capture and sequencing of UCEs, bioinformatic pipeline using Phyluce 1.4 (22), SNP calling, allele phasing, and the incorporation of phased alleles to the reference using the seccap\_pop (8) followed Thom *et al.* (13) (details in the “Sequence capture of UCEs and bioinformatics” section in the Supplementary Materials).

### Genetic structure

To explore the genetic structure of each species complex, we used sNMF (23) to test the best-fit number of ancestral populations ( $K$ ) and to cluster individuals to populations by applying sparse non-negative matrix factorization to compute least-square estimates of ancestry coefficients. We tested  $K$  values ranging from 1 to 8, running 100 replicates for each  $K$  value. Four values of alpha regularization parameter (1, 10, 100, and 1000) were used to test the robustness of the results. Similarly, we implemented the  $k$ -means (find.clusters) of ADEGENET 2.0 (37) to infer the number of genetic clusters. All components of the initial principal components analysis were retained. Last, we used discriminant analysis of principal components (DAPC) in ADEGENET 2.0 (37) to visualize the results. For the sNMF and  $k$ -means, we used as input our final SNP matrix with one randomly selected SNP per UCE loci.

To test whether the observed genetic diversity among populations of each species complex deviates from the IBD scenario and identify corridors or barriers to gene flow, we used EEMS (24). Euclidian genetic dissimilarity matrices between individuals for each species complex were generated using the SNP datasets in ADEGENET 2.0 (37) with the function “dist.genpop” using method 4 (Rogers’ distance). Habitat polygons were produced on the basis

of the geographic distribution of each species complex, and 300 demes were distributed over the habitat area. A single Markov chain Monte Carlo (MCMC) run was performed for  $30 \times 10^6$  generations for each species with the first  $5 \times 10^6$  generations excluded as burn-in. The convergence of the MCMC was visually assessed by plotting the log posterior probabilities of each run.

### Phylogenetic relationships, species tree, and gene flow

We inferred the phylogenetic relationship between populations inferred with sNMF/DAPC analyses for each species complex using the coalescent-based method implemented in SNAPP (38). SNAPP infers the likelihood of a given species tree using allele frequency of unlinked SNPs bypassing the need to integrate the probabilities of gene trees in the function of a given species tree. We used gamma rate priors for alpha and beta parameters, with all other priors with default values. Two replicates of 2.5 million MCMC generations were run with 100,000 burn-in iterations for each species complex. Estimated parameters were sampled every 500 generations. Burn-in values for the MCMC chains were assessed in Tracer 1.4 (<http://tree.bio.ed.ac.uk/software/tracer/>). To estimate the graph of relationships among populations of each species complex accounting for admixture, we used TREEMIX 1.12 (39). We ran TREEMIX varying from zero to six migration edges for each species complex. The best model was selected observing the significance of migration edges and the residue covariance matrix. Given the high proportions of the shared coefficient of ancestry estimated with sNMF for some localities and to improve our resolution in detecting introgressed populations, we subdivided populations by clustering individuals from neighborhood localities and with similar proportions of ancestry. In addition, we implemented the *f*<sub>3</sub> statistics using the three-pop test for admixture (40). All possible clusters of populations were tested. We assumed as groups the genetic clusters inferred by sNMF. Localities with individuals with a shared coefficient of ancestry close to 0.5 (potentially introgressed) were assumed to belong to distinct groups (intermediary).

Given the gene flow pattern inferred by TREEMIX for *M. lugubris* and *M. assimilis* and the low node supports in *T. nigrocinereus*/*T. cryptoleucus* inferred by SNAPP, we performed a topology test using FSC (26), assuming the presence of gene flow before testing the demographic models and estimating parameters. For *M. lugubris*, we tested two models assuming constant gene flow since divergence but varying the topology (Fig. 4C): (i) consensus tree obtained with SNAPP (Solimões,(Negro,(Madeira,Amazonas))), and (ii) alternative topology obtained with SNAPP clustering the populations that exchange alleles as identified by TREEMIX (Solimões,(Amazonas,(Negro, Madeira))). For *T. nigrocinereus*/*T. cryptoleucus*, we tested two models that varied the relationship of the Negro River population with the other groups (Fig. 4D): (i) consensus tree obtained with SNAPP ((Solimões, Madeira),(Negro,(Tapajós,Amazonas))), and (ii) alternative topology obtained with SNAPP (((Solimões, Madeira), Negro),(Tapajós,Amazonas)). In *M. assimilis*, we tested three possible topologies (Fig. 4E): (i) consensus tree obtained with SNAPP (Solimões, (Madeira, Amazonas/Negro)), (ii) (Amazonas/Negro, (Solimões, Madeira)), and (iii) (Solimões,(Madeira, Amazonas/Negro)).

### Species complex demography using multipopulation models

Since demographic analyses can be biased by the particular intrinsic characteristics of the methods used, we estimated demographic

parameters using two distinct and complementary approaches. First, we used the G-PhoCS (25) to estimate mutational-scaled ( $\mu$ ) divergence times given in  $\tau$  ( $\tau = T\mu/g$ , where  $T$  is the absolute divergence time in years,  $g$  is the average generation time), effective population sizes based on  $\theta$  ( $\theta = 4N_e\mu$ , where  $N_e$  is the absolute effective population size in a number of individuals), and gene flow between populations measured as migration bands (Fig. 4). G-PhoCS uses unphased loci sequences as input, integrating the likelihood computation over all possible gametic phases. A single preliminary MCMC run was conducted for each of the three species complexes using default priors. Each preliminary run included 300,000 iterations, sampling every five iterations, and the first 100,000 iterations were conservatively excluded as burn-in. A final run for each species complex was then performed given  $\alpha = 1.0$  and  $\beta = 5000$  for  $\tau$  as well as  $\theta$ ,  $\alpha = 1.2$  and  $\beta = 0.01$  for mutation-scaled migration rates, 500,000 iterations sampling every five, and fine-tune parameters automatically estimated during the first 10,000 burn-in iterations. We assumed an average mutation rate of  $2.5 \times 10^{-9}$  (41) and a generation time of 1 year (see Discussion). To reduce the computational time, we subsampled between 5 and 10 individuals per population. Migration bands were considered to have significant levels of gene flow if the 95% Bayesian confidence interval of the migration rate did not overlap zero.

To explicitly test distinct demographic scenarios and estimate parameters for comparison with the G-PhoCS results, we implemented a model-based approach in FSC v.2.5.2 (26). FSC calculates the composite likelihood of a simulated joint site frequency spectrum (jSFS) given a specific scenario of diversification, which is optimized against the observed data to estimate population genetic parameters such as divergence time, effective population size and gene flow. Final variant call format (VCF) files were converted in jSFS using *daI* 1.7 (42). We tested five demographic models, which combined with our empirical analyses enabled us to test the assumptions of our diversification hypotheses. The five models we tested varied the presence and search interval for divergence time, gene flow, and postdivergence population size changes (Fig. 4): 0mig\_0siz, demographic model assuming no migration between populations and no population size changes after divergence; 0mig\_siz, demographic model assuming no migration between populations but with population size changes after divergence (expansions and/or bottlenecks); mig\_0siz, demographic model assuming migration between populations but no population size changes; mig\_siz, demographic model assuming migration between populations and population size changes; mig\_siz\_ALL\_G-PhoCS, demographic model assuming migration between populations and population size changes with all divergence times fixed using G-PhoCS mean estimations. To calibrate parameter estimates to absolute values that are biologically meaningful, we fixed the most ancient divergence event for each species complex using the G-PhoCS posterior distribution mean estimate, since G-PhoCS is known to produce robust estimates of divergence times even under gene flow (25). This procedure allowed us to omit monomorphic sites, thus using only randomly selected SNPs for the analyses. We performed the composite likelihood search using 50 independent runs per model, given 40 to 50 cycles of Brent algorithm for each independent run and 100,000 genealogical simulations per generation of a single jSFS. To select the best-fit model to the observed data, we implemented Akaike's weight of evidence, which is based on the Akaike information criterion [ $AIC = 2k - 2\ln(L)$ , where  $k$  is the number of parameters estimated in the model, and  $L$  is the composite likelihood value] for the single run with the highest composite likelihood

per model. Parameter estimates were then derived from the highest likelihood run of the best-fit model. To obtain 95% confidence intervals, we performed 50 parametric bootstrap estimates simulating jSFSs based on the parameter values of the best model and reestimating parameters using these simulated datasets.

### Synchronicity of size changes among populations

In addition to the population size changes estimated in the FSC multipopulation models, we tested for the degree of synchrony in size changes among populations through hierarchical codemographic modeling with the Multi-DICE R package (27). To clarify, this method treats each extant lineage throughout the entire dataset as an independent single population and requires the same number of samples per population. Thus, we considered 15 populations projected to six diploid individuals each. To better inform our hierarchical codemographic model, we first performed a preliminary analysis that involved independent inferences for every single population. We used FSC to generate 400,000 coalescent simulations per each of three demographic syndromes: (i) instantaneous population expansion, (ii) constant size, and (iii) instantaneous population contraction. We then applied an ABC procedure to iteratively compare every 1 of the 15 observed folded SFS with this table of 1.2 million simulations. We performed ABC inference with the ABC R package (43) by accepting the 0.125% of simulations (1500 simulations) with the shortest Euclidean distance to the observed SFS.

Afterward, we grouped SFS belonging to populations with the same best-fit syndrome and converted these into the aggregate SFS (aSFS) to estimate the coefficient of synchrony in population size changes,  $\zeta$  (27). We then implemented a hierarchical ABC (hABC) approach, comparing this observed data vector with simulations generated under a hierarchical codemographic model across different degrees of temporal synchrony in instantaneous population size change. (Details on this procedure including priors for the single population and codemography using Multi-DICE are described in the “Synchronicity of size changes among populations” section in the Supplementary Materials.

### SUPPLEMENTARY MATERIALS

Supplementary material for this article is available at <http://advances.sciencemag.org/cgi/content/full/6/11/eaax4718/DC1>

Supplementary Material and Methods

Study groups and floodplain environments

Sequence capture of UCEs and bioinformatics

Synchronicity of size changes among populations

Summary of results

Full versus composite likelihood

Table S1. Summary of bioinformatics results.

Table S2. Topologies tested for *M. lugubris*, *T. nigrocinereus*/*T. cryptoleucus*, and *M. assimilis* in FSC.

Table S3. Demographic parameters estimated for *M. lugubris* obtained with G-PhoCS.

Table S4. Demographic parameters estimated for *T. nigrocinereus*/*T. cryptoleucus* obtained with G-PhoCS.

Table S5. Demographic parameters estimated for *M. assimilis* obtained with G-PhoCS.

Table S6. Parameter values of simulated models in FSC with the best composite likelihood for *M. lugubris*.

Table S7. Parameter values of simulated models in FSC with the best composite likelihood for *T. nigrocinereus*/*T. cryptoleucus*.

Table S8. Parameter values of simulated models in FSC with the best composite likelihood for *M. assimilis*.

Table S9. Model selection and parameter estimation of Multi-DICE single-population models for *M. assimilis*, *M. lugubris*, and *T. nigrocinereus*/*T. cryptoleucus*.

Table S10. Posterior estimates obtained with the hABC procedure implemented in Multi-DICE using aSFS.

Table S11. Results of the “leave one out” cross-validations obtained with the hABC procedure implemented in Multi-DICE using the aSFS.

Table S12. Individuals sampled of *M. lugubris*, *M. assimilis*, and *T. nigrocinereus*/*T. cryptoleucus*.  
Fig. S1. Estimated effective diversity surface obtained in EEMS for *M. lugubris*, *T. nigrocinereus*/*T. cryptoleucus*, and *M. assimilis*.

Fig. S2. Principal components analyses of the simulated single-population models with Multi-DICE.

Fig. S3. Principal components analyses on the retained simulated SFS of each population.

Fig. S4. Principal components analysis of simulated and observed aSFSs performed in Multi-DICE.

References (44–50)

[View/request a protocol for this paper from Bio-protocol.](#)

### REFERENCES AND NOTES

1. A. R. Wallace, On the monkeys of the Amazon. *Ann. Mag. Nat. Hist.* **14**, 451–454 (1854).
2. J. Haffer, Speciation in amazonian forest birds. *Science* **165**, 131–137 (1969).
3. G. Thom, A. Aleixo, Cryptic speciation in the white-shouldered antshrike (*Thamnophilus aethiops*, Aves – Thamnophilidae): The tale of a transcontinental radiation across rivers in lowland Amazonia and the northeastern Atlantic Forest. *Mol. Phylogenet. Evol.* **82**, 95–110 (2015).
4. S. M. Silva, A. T. Peterson, L. Carneiro, T. C. T. Burlamaqui, C. C. Ribas, T. Sousa-Neves, L. S. Miranda, A. M. Fernandes, F. M. d’Horta, L. E. Araújo-Silva, R. Batista, C. H. M. M. Bandeira, S. M. Dantas, M. Ferreira, D. M. Martins, J. Oliveira, T. C. Rocha, C. H. Sardelli, G. Thom, P. S. Rêgo, M. P. Santos, F. Sequeira, M. Vallinoto, A. Aleixo, A dynamic continental moisture gradient drove Amazonian bird diversification. *Sci. Adv.* **5**, eaat5752 (2019).
5. L. N. Naka, R. T. Brumfield, The dual role of Amazonian rivers in the generation and maintenance of avian diversity. *Sci. Adv.* **4**, eaar8575 (2018).
6. J. V. Remsen Jr., T. A. Parker III, Contribution of river-created habitats to bird species richness in Amazonia. *Biotropica* **15**, 223–231 (1983).
7. A. L. Albernaz, R. L. Pressey, L. R. F. Costa, M. P. Moreira, J. F. Ramos, P. A. Assunção, C. H. Francison, Tree species compositional change and conservation implications in the white-water flooded forests of the Brazilian Amazon. *J. Biogeogr.* **39**, 869–883 (2012).
8. M. G. Harvey, A. Aleixo, C. C. Ribas, R. T. Brumfield, Habitat association predicts genetic diversity and population divergence in amazonian birds. *Am. Nat.* **190**, 631–648 (2017).
9. A. Aleixo, Historical diversification of floodplain forest specialist species in the Amazon: A case study with two species of the avian genus *Xiphorhynchus* (Aves: Dendrocolaptidae). *Biol. J. Linn. Soc. Lond.* **89**, 383–395 (2006).
10. E. M. Latrubesse, E. Y. Arima, T. Dunne, E. Park, V. R. Baker, F. M. d’Horta, C. Wight, F. Wittmann, J. Zuanon, P. A. Baker, C. C. Ribas, R. B. Norgaard, N. Filizola, A. Ansar, B. Flyvbjerg, J. C. Stevaux, Damming the rivers of the Amazon basin. *Nature* **546**, 363–369 (2017).
11. C. D. Cadena, N. Gutiérrez-Pinto, N. Dávila, R. T. Chesser, No population genetic structure in a widespread aquatic songbird from the Neotropics. *Mol. Phylogenet. Evol.* **58**, 540–545 (2011).
12. M. Cohn-Haft, L. N. Naka, A. M. Fernandes, Padrões de distribuição da avifauna da várzea dos rios Solimões-Amazonas. Pp. 287–324. In: A. L. Albernaz (org), Conservação da Várzea, Identificação e Caracterização de Regiões Biogeográficas. IBAMA/ ProVárzea/ INPA, 2007).
13. G. Thom, F. R. Do Amaral, M. J. Hickerson, A. Aleixo, L. E. Araújo-Silva, C. C. Ribas, E. Choueri, C. Y. Miyaki, Phenotypic and genetic structure support gene flow generating gene tree discordances in an amazonian floodplain endemic species. *Syst. Biol.* **67**, 700–718 (2018).
14. É. L. Choueri, C. Gubili, S. H. Borges, G. Thom, A. O. Sawakuchi, E. A. A. Soares, C. C. Ribas, Phylogeography and population dynamics of Antbirds (Thamnophilidae) from Amazonian fluvial islands. *J. Biogeogr.* **44**, 2284–2294 (2017).
15. J. C. Avice, *Phylogeography: The History and Formation of Species* (Harvard Univ. Press, 2000).
16. G. Irion, J. Müller, J. O. Morais, G. Keim, J. N. de Mello, W. J. Junk, The impact of Quaternary sea level changes on the evolution of the Amazonian lowland. *Hydrol. Process.* **23**, 3168–3172 (2009).
17. A. Govin, C. M. Chiessi, M. Zabel, A. O. Sawakuchi, D. Heslop, T. Hörner, Y. Zhang, S. Multiza, Terrigenous input off northern South America driven by changes in Amazonian climate and the North Brazil Current retroflection during the last 250 ka. *Clim. Past Discuss.* **9**, 5855–5898 (2013).
18. F. N. Pupim, A. O. Sawakuchi, R. P. Almeida, C. C. Ribas, A. K. Kern, G. A. Hartmann, C. M. Chiessi, L. N. Tamura, T. D. Mineli, J. F. Savian, C. H. Grohmann, D. J. Bertassoli Jr., A. G. Stern, F. W. Cruz, J. Cracraft, Chronology of Terra Firme formation in Amazonian lowlands reveals a dynamic Quaternary landscape. *Quat. Sci. Rev.* **210**, 154–163 (2019).
19. A. T. Thomaz, M. R. Christie, L. L. Knowles, The architecture of river networks can drive the evolutionary dynamics of aquatic populations. *Evolution* **70**, 731–739 (2016).

20. L. B. Beheregaray, G. M. Cooke, N. L. Chao, E. L. Landguth, Ecological speciation in the tropics: Insights from comparative genetic studies in Amazonia. *Front. Genet.* **5**, 477 (2015).
21. D. F. Rossetti, M. C. L. Cohen, S. H. Tatum, A. O. Sawakuchi, É. H. Cremon, J. C. R. Mittani, T. C. Bertani, C. J. A. S. Munita, D. R. G. Tudela, M. Yee, G. Moya, Mid-Late Pleistocene OSL chronology in western Amazonia and implications for the transcontinental Amazon pathway. *Sediment. Geol.* **330**, 1–15 (2015).
22. B. C. Faircloth, J. E. McCormack, N. G. Crawford, M. G. Harvey, R. T. Brumfield, T. C. Glenn, Ultraconserved elements anchor thousands of genetic markers spanning multiple evolutionary timescales. *Syst. Biol.* **61**, 717–726 (2012).
23. E. Frichot, F. Mathieu, T. Trouillon, G. Bouchard, O. François, Fast and efficient estimation of individual ancestry coefficients. *Genetics* **196**, 973–983 (2014).
24. D. Petkova, J. Novembre, M. Stephens, Visualizing spatial population structure with estimated effective migration surfaces. *Nat. Genet.* **48**, 94–100 (2016).
25. I. Gronau, M. J. Hubisz, B. Gulko, C. G. Danko, A. Siepel, Bayesian inference of ancient human demography from individual genome sequences. *Nat. Genet.* **43**, 1031–1034 (2011).
26. L. Excoffier, I. Dupanloup, E. Huerta-Sánchez, V. C. Sousa, M. Foll, Robust demographic inference from genomic and SNP data. *PLoS Genet.* **9**, e1003905 (2013).
27. A. T. Xue, M. J. Hickerson, Multi-dice: R package for comparative population genomic inference under hierarchical co-demographic models of independent single-population size changes. *Mol. Ecol. Resour.* **17**, e212–e224 (2017).
28. N. G. Swenson, D. J. Howard, Do suture zones exist? *Evolution* **58**, 2391–2397 (2004).
29. C. Hoorn, G. R. Bogotá-A, M. Romero-Baez, E. I. Lammertsma, S. G. A. Flantua, E. L. Dantas, R. Dino, D. A. do Carmo, F. Chemale Jr., The Amazon at sea: Onset and stages of the Amazon River from a marine record, with special reference to Neogene plant turnover in the drainage basin. *Glob. Planet. Chang.* **153**, 51–65 (2017).
30. E. I. Lammertsma, S. R. Troelstra, J.-A. Flores, F. Sangiorgi, F. Chemale Jr., D. A. do Carmo, C. Hoorn, Primary productivity in the western tropical Atlantic follows Neogene Amazon River evolution. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **506**, 12–21 (2018).
31. H. Cheng, A. Sinha, F. W. Cruz, X. Wang, R. L. Edwards, F. M. d'Horta, C. C. Ribas, M. Vuille, L. D. Stott, A. S. Auler, Climate change patterns in Amazonia and biodiversity. *Nat. Commun.* **4**, 1411 (2013).
32. X. Wang, R. L. Edwards, A. S. Auler, H. Cheng, X. Kong, Y. Wang, F. W. Cruz, J. A. Dorale, H.-W. Chiang, Hydroclimate changes across the Amazon lowlands over the past 45,000 years. *Nature* **541**, 204–207 (2017).
33. C. Häggi, E. Scheuß, A. O. Sawakuchi, C. M. Chiessi, S. Mulitza, D. J. Bertassoli, J. Hefter, M. Zabel, P. A. Baker, S. Schouten, Modern and late Pleistocene particulate organic carbon transport by the Amazon River: Insights from long-chain alkyl diols. *Geochim. Cosmochim. Acta* **262**, 1–19 (2019).
34. C. Gualtieri, N. Filizola, M. de Oliveira, A. M. Santos, M. Ianniruberto, A field study of the confluence between Negro and Solimões Rivers. Part 1: Hydrodynamics and sediment transport. *Compt. Rendus Geosci.* **350**, 31–42 (2018).
35. N. Filizola, J. L. Guyot, Suspended sediment yields in the Amazon basin: An assessment using the Brazilian national data set. *Hydrol. Process.* **23**, 3207–3215 (2009).
36. G. A. Bravo, Phenotypic and Niche Evolution in the Antbirds (Aves: Thamnophilidae) (2012); [http://digitalcommons.lsu.edu/cgi/viewcontent.cgi?article=5009&context=gradschool\\_dissertations](http://digitalcommons.lsu.edu/cgi/viewcontent.cgi?article=5009&context=gradschool_dissertations).
37. T. Jombart, I. Ahmed, adegenet 1.3-1: New tools for the analysis of genome-wide SNP data. *Bioinformatics* **27**, 3070–3071 (2011).
38. D. Bryant, R. Bouckaert, J. Felsenstein, N. A. Rosenberg, A. RoyChoudhury, Inferring species trees directly from diallelic genetic markers: Bypassing gene trees in a full coalescent analysis. *Mol. Biol. Evol.* **29**, 1917–1932 (2012).
39. J. K. Pickrell, J. K. Pritchard, Inference of population splits and mixtures from genome-wide allele frequency data. *PLoS Genet.* **8**, e1002967 (2012).
40. D. Reich, K. Thangaraj, N. Patterson, A. L. Price, L. Singh, Reconstructing Indian population history. *Nature* **461**, 489–494 (2009).
41. K. Nadachowska-Brzyska, C. Li, L. Smeds, G. Zhang, H. Ellegren, Temporal dynamics of avian populations during pleistocene revealed by whole-genome sequences. *Curr. Biol.* **25**, 1375–1380 (2015).
42. R. N. Gutenkunst, R. D. Hernandez, S. H. Williamson, C. D. Bustamante, Inferring the joint demographic history of multiple populations from multidimensional SNP frequency data. *PLoS Genet.* **5**, e1000695 (2009).
43. K. Csilléry, O. François, M. G. B. Blum, abc: An R package for approximate Bayesian computation (ABC). *Methods Ecol. Evol.* **3**, 475–479 (2012).
44. K. Zimmer, M. L. Isler, Typical antbirds (Thamnophilidae), in *Handbook of the Birds of the World Alive*, J. del Hoyo, A. Elliott, J. Sargatal, D. A. Christie, E. de Juana, Eds. (Lynx Edicions: Barcelona, 2016); [www.hbw.com/node/52291](http://www.hbw.com/node/52291) [Verified 29 July 2016].
45. M. G. Grabherr, B. J. Haas, M. Yassour, J. Z. Levin, D. A. Thompson, I. Amit, X. Adiconis, L. Fan, R. Raychowdhury, Q. Zeng, Z. Chen, E. Mauceli, N. Hacohen, A. Gnirke, N. Rhind, F. di Palma, B. W. Birren, C. Nusbaum, K. Lindblad-Toh, N. Friedman, A. Regev, Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat. Biotechnol.* **29**, 644–652 (2011).
46. H. Li, R. Durbin, Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).
47. H. Li, B. Handsaker, A. Wysoker, T. Fennell, J. Ruan, N. Homer, G. Marth, G. Abecasis, R. Durbin; 1000 Genome Project Data Processing Subgroup, The sequence alignment/map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
48. A. McKenna, M. Hanna, E. Banks, A. Sivachenko, K. Cibulskis, A. Kernytsky, K. Garimella, D. Altshuler, S. Gabriel, M. Daly, M. A. DePristo, The genome analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* **20**, 1297–1303 (2010).
49. M. G. Harvey, B. T. Smith, T. C. Glenn, B. C. Faircloth, R. T. Brumfield, Sequence capture versus restriction site associated DNA sequencing for shallow systematics. *Syst. Biol.* **65**, 910–924 (2016).
50. M. Maldonado-Coelho, Climatic oscillations shape the phylogeographical structure of Atlantic Forest fire-eye antbirds (Aves: Thamnophilidae). *Biol. J. Linn. Soc. Lond.* **105**, 900–924 (2012).

**Acknowledgments:** We thank the curator and curatorial assistants of the Instituto Nacional de Pesquisas da Amazônia (INPA), Laboratório de Genética e Evolução Molecular de Aves (LGEMA), and Museu Paraense Emílio Goeldi (MPEG) for allowing us to borrow tissue samples and study specimens under their care. We thank members of the Hickerson lab, more specifically I. Overcast, for the discussion related to the use of model-based approaches. We thank L. E. Araujo-Silva for the partnership during field expeditions and to F. Raposo do Amaral for the extensive revision on the manuscript. This work was developed in the Research Center on Biodiversity and Computing (BioComp) of the Universidade de São Paulo (USP), supported by the USP Provost's Office for Research. Specimens were collected under the ICBio/SISBIO permanent collecting permits #72232 and #10420 and approved by the Animal Ethics Committee of the University of São Paulo, protocol #188/2013. **Funding:** This study was co-funded by FAPESP (BIOTA, 2012/50260-6 and 2013/50297-0), NSF (DOB 1343578), NASA, CNPq (310593/2009-3, 574008/2008-0, 563236/2010-8, and 471342/2011-4), PEER/USAID (AID-OAA-A-11-00012), and FAPESPA (ICAAF 023/2011). G.T. was granted by CAPES and then FAPESP scholarships (2014/00113-2, 2015/12551-7, 2018/17869-3, and 2017/25720-7). A.A., C.C.R., and C.M. are supported by CNPq research productivity fellowships (310880/2012-2, 308927/2016-8, and 303713/2015-1). M.J.H. was funded by FAPESP (BIOTA, 2013/50297-0), NASA through the Dimensions of Biodiversity Program (DOB 1343578), and the NSF (DEB-1253710). **Author contributions:** G.T. and C.M. conceived the study. G.T. and other collectors obtained the samples in the field. G.T. performed bioinformatics and phylogeographic analyses with the assistance of A.T.X. and M.J.H. G.T., A.T.X., A.O.S., and C.M. wrote the paper with input from A.A., C.C.R., and M.J.H. All authors reviewed and approved the final version of the manuscript before submission. **Competing interest:** The authors declare that they have no competing interests. **Data and materials availability:** All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials. Additional data related to this paper are available at <https://github.com/GregoryThom/Quaternary-climate-changes-as-speciation-drivers-in-the-Amazon-Floodplains>, or may be requested from the authors. All genetic data are available at the Sequence Read Archive (SRA); BioProject ID: PRJNA595086; SubmissionID: SUB6206537; [www.ncbi.nlm.nih.gov/bioproject/595086](http://www.ncbi.nlm.nih.gov/bioproject/595086).

Submitted 26 March 2019  
Accepted 13 December 2019  
Published 11 March 2020  
10.1126/sciadv.aax4718

**Citation:** G. Thom, A. T. Xue, A. O. Sawakuchi, C. C. Ribas, M. J. Hickerson, A. Aleixo, C. Miyaki, Quaternary climate changes as speciation drivers in the Amazon floodplains. *Sci. Adv.* **6**, eaax4718 (2020).

## Quaternary climate changes as speciation drivers in the Amazon floodplains

Gregory Thom, Alexander T. Xue, André O. Sawakuchi, Camila C. Ribas, Michael J. Hickerson, Alexandre Aleixo and Cristina Miyaki

*Sci Adv* 6 (11), eaax4718.  
DOI: 10.1126/sciadv.aax4718

ARTICLE TOOLS	<a href="http://advances.sciencemag.org/content/6/11/eaax4718">http://advances.sciencemag.org/content/6/11/eaax4718</a>
SUPPLEMENTARY MATERIALS	<a href="http://advances.sciencemag.org/content/suppl/2020/03/09/6.11.eaax4718.DC1">http://advances.sciencemag.org/content/suppl/2020/03/09/6.11.eaax4718.DC1</a>
REFERENCES	This article cites 46 articles, 5 of which you can access for free <a href="http://advances.sciencemag.org/content/6/11/eaax4718#BIBL">http://advances.sciencemag.org/content/6/11/eaax4718#BIBL</a>
PERMISSIONS	<a href="http://www.sciencemag.org/help/reprints-and-permissions">http://www.sciencemag.org/help/reprints-and-permissions</a>

Use of this article is subject to the [Terms of Service](#)

---

*Science Advances* (ISSN 2375-2548) is published by the American Association for the Advancement of Science, 1200 New York Avenue NW, Washington, DC 20005. The title *Science Advances* is a registered trademark of AAAS.

Copyright © 2020 The Authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original U.S. Government Works. Distributed under a Creative Commons Attribution NonCommercial License 4.0 (CC BY-NC).