Supplementary Appendices for:

Genomic phylogeography of the White Crowned Manakin *Pseudopipra pipra* (Aves: Pipridae) illuminates a continental-scale radiation out of the Andes

Author Affiliations:

Jacob S. Berv[1,5*], Leonardo Campagna[1], Teresa J. Feo[2], Ivandy Castro Astor[3], Camila Ribas[4], Richard O. Prum[5], Irby J. Lovette[1]

1. Fuller Evolutionary Biology Program, Cornell Lab of Ornithology, 159 Sapsucker Woods Road, Ithaca, NY 14850, USA. Department of Ecology and Evolutionary Biology, Cornell University, 215 Tower Road, Ithaca, NY 14853, USA.

2. Department of Vertebrate Zoology, MRC-116, National Museum of Natural History, Smithsonian Institution, Washington, DC 20013, USA.

3. Department of Biology, City College of New York and CUNY Graduate Center, City University of New York, New York, NY 10031, USA

4. Coordenacão de Biodiversidade, Instituto Nacional de Pesquisas da Amazônia, Manaus, AM, Brazil

5. Department of Ecology and Evolutionary Biology, and Peabody Museum of Natural History, Yale University, New Haven, Connecticut, 06520, USA.

*Correspondence to: Jacob S. Berv (jsb439@cornell.edu)

Supplementary Appendix

Section 1

Expanded Taxonomic Summary

Pseudopipra coracina (Sclater 1856) **Andean White-crowned Manakin**

Distribution: Subtropical Andes from Venezuela south to Esmeraldas, Ecuador and San Martín, Peru.

Phylogenetic Position: Clade A1, plus multiple unsampled subspecies from the Colombian and Ecuadorian Andes (Fig. 3).

Comments: This apparently monophyletic group of northern Andean populations includes five currently recognized subspecies, each of which may be a distinct species. Three of these subspecies–*coracina, minima,* and *occulta*– have unique, highly differentiated vocal types, and diagnosable plumage differences. The vocal type of *bolivari* is unknown.

P. c. coracina (Sclater 1856)

Distribution: Subtropical forests of the eastern slope of the Andes from western Venezuela to Morona-Santiago, Ecuador.

Phylogenetic Position: Based on mtDNA sampled, a member of Clade A1.

Plumage: Males are moderately glossy on the back. White crown feathers are long with extensive black bases. Crowns are sometimes slightly grayish. Females are olive green with lighter yellow belly, and olive gray crown with more olive cheeks.

Lek Vocal Type: 8 (*errrwer*).

Call Vocal Type: Unknown

P. c. minima (Chapman 1914)

Distribution: Subtropical forests of western Cauca, Colombia south to Esmeraldas, Ecuador

Phylogenetic Position: Not Sampled, but a likely member of Clade A1.

Plumage: Males are moderately glossy; crown feathers are entirely white to their bases. No females were observed. Chapman (1914) reported that *minima* is smaller than *anthracina*, and that males lack prominent gray tips on undertails. Freile (2014) reported one specimen of a female from San Javier, Esmeraldas, Ecuador (100 meters) and provisionally identified it as *minima.* The specimen is bright olive above and below with a slightly grayish olive grown. However, this specimen is from a substantially lower altitude than Colombian records of *minima*, so it may represent an altitudinal migrant or a distinct population.

Lek Vocal Type: 9 (*reeee*)

Call Vocal Type: Unknown

P. c. bolivari (de Schauensee 1950)

Distribution: Subtropical forests of southern Córdoba, Colombia. (Not Sampled)

Phylogenetic Position: Not Sampled, but a likely member of Clade A1.

Plumage: None observed. Apparently known only from the type specimen from Cerro Murucucú, Córdoba, Colombia. de Schauensee (1950) described this male specimen as

having entirely white feathers in the forecrown, like *minima* and *unica,* but hindcrown feathers basally black like *coracina.*
Lek Vocal Type: Unknown
Call Vocal Type: unknown

P. c. unica (de Schauensee 1945)
Distribution: Subtropical forests of Magdalena Valley, Antioquia to Huila, Colombia.
Phylogenetic Position: Not Sampled, but a likely member of Clade A1.
Plumage: Males are moderate glossy, with long crown feathers that are white to their bases. Females are olive green above, and slightly gray on the crown; underparts uniform olive. de Schauensee (1945) described *unica* as glossier than coracina, with longer tail and very long crest.
Lek Vocal Type: 11a (*weer-dink*) and 11b (*shureeep*)
Call Vocal Type: unknown

P. c. occulta (Zimmer 1936)
Distribution: Eastern slope of the Andes from Zamora-Chinchipe, Ecuador (Freile 2014) south to San Martín, and Huánuco, Peru, west of the Rio Huallaga
Phylogenetic Position: Clade A1 (Fig. 3)
Plumage: Males are glossy with dark gray bases to crown feathers. Females are dark olive with dark gray crown and gray throat. Zimmer (1936) described occulta as similar to *comata* but adult males with the occipital feathers slightly shorter and with the crown and occipital feathers sooty at the base instead of entirely white.
Lek Vocal Type: 1 (*trill-dink*) and 10 (*bree*)
Call Vocal Type: unknown

Pseudopipra anthracina (Ridgway 1906) **Western White-crowned Manakin**
Distribution: Subtropical Costa Rica to Western Panama
Phylogenetic Position: Clade A2 (Fig. 3)
Plumage: Males less lustrous on back than all other *Pseudopipra* populations, white crown feathers gray or dark gray at base. Female are olive green with slaty crown and face. Ridgway (1906) considered *anthracina* to have shorter wings, smaller beak, less lustrous plumage than *pipra* with undertails tipped with gray.
Lek Vocal Type: 4 (*jureeee*)
Call Vocal Type: unknown

Pseudopipra comata (Berlepsch and Stolzmann 1894) **Junín White-crowned Manakin**
Distribution: Subtropical Andes of Peru from Cerro Azul, Loreto (east and south of the Rio Huallaga) to southern Huánuco, Pasco, Junín, and northern Cusco.
Phylogenetic Position: Clade B (Fig. 3).
Plumage: Males are glossy black above, crown feathers longer and entirely white to their bases. Females are bright olive green above, gray on crown and face, slightly gray on throat, dark olive below, and slightly dark gray on the belly.
Lek Vocal Type: One record, statistically similar to type 1 (*trill-dink*)

Call Vocal Type: unknown
Comment: *P. comata* is also composed of two well differentiated subclades. The northern clade (B1) is known from Cerro Azul in Loreto, Peru. The southern clade (B2) is known from extreme southern Huánuco (Cerros del Sira, 9˚30'S 74˚47'W; AMNH 820866, 820952), Pasco, Junín, and Cusco. The type locality of *comata* is Vitoc, Junín within the southern clade. Further investigation plumage and behavioral is necessary to determine whether the Cerro Azul populations should be recognized as a distinct, new taxon.

Pseudopipra pygmaea *(*Zimmer 1936) **Huallaga White-crowned Manakin**
Distribution: Tropical forest of Lower Rio Huallaga Valley, Peru
Phylogeneic Position: Sister to Clade F (mtDNA)
Plumage: Males: Glossy, with black bases to crown feathers. Females are olive above and gray below with a band of olive across the chest; crown and face only slightly darker than back, not gray. Zimmer (1936) described males as having long crest with gray bases, crown sometimes slightly ashy; females are much paler than *occulta*; throat and belly decidedly more whitish, breast paler duller green; lighter even than *microlopha*.
Lek Vocal Type: 2 (*deeeer*)
Call Vocal Type: 13
Comment: Lowland populations along the Rio Huallaga have been named *pygmaea* (Zimmer 1936). Our four samples of *pygmaea* from Jeberos, Peru did not yield sufficient quality DNA for RADseq, but all four had a phylogenetically distinct mtDNA haplotype which placed this lineage as the sister group to all other lowland populations of *Pseudopipra*. These populations have song type 2, which appears to be shared plesiomorphically with *P. discolor* and *P. microlopha separabilis* from Para, Brazil.

Pseudopipra discolor (Zimmer 1936) **Napo White-crowned Manakin**
Distribution: Tropical forest in Napo, Ecuador and northern Loreto, Peru south to the Rio Marañón.
Phylogenetic Position: Clade E (Fig. 3)
Distribution: Tropical forest in Napo, Ecuador south to the Rio Marañón
Plumage: Males are glossy black above, white crown feathers with black or dark gray bases. Females are dusky olive overall, slightly grayer on crown, and grayer belly. Zimmer (1936) described male *discolor* as glossier and bluer above than *pipra.*
Lek Vocal Type: 2 (*deeeer*)
Call Vocal Type: 13
Comment: This lineage was found to have both a distinct, unique history, with subsequent introgression with adjacent populations of the northern Amazonian clade. The nature of this introgression indicates this lineage may be best recognized as a distinct hybrid species.

Pseudopipra pipra (Linneaus 1758) **Northern White-crowned Manakin**
Distribution: Tropical forest of eastern Colombia, southern Venezuela, the Guianas, and Brazil north of the Amazon. West to the right (north) bank of the Rio Putumayo, Colombia.
Phylogenetic Position: Clade D (Fig. 3).

Plumage: Males are glossy black above, crown feathers longer with extensive black bases. Females are dark olive above, olive below, grayer on belly, and occasionally only slightly darker gray on crown.
Lek Vocal Type: 3 (*buzzzz*)
Call Vocal Type: 5 (*zeee*)

Pseudopipra microlopha (Zimmer 1929) **Southern White-crowned Manakin**
Distribution: Tropical forest of eastern Peru south of the Rio Marañón, and south of the Amazon east to Pará, Brazil, and subtropical forests between the Rio Huallaga and Rio Ucayali
Phylogenetic Position: Paraphyletic, including Clade C without Clade C7 (Fig. 3).
Comments: A paraphyletic group (with respect to *P. cephaleucos* from Brazilian Atlantic forest) which includes three, currently recognized subspecies, and four additional genetically well-supported monophyletic subgroups that may be recognized as new taxa. Furthermore, we identified a genetically distinct montane clade from the highlands between Rio Huallaga and Rio Ucayali that has not been previously described, and may have distinct plumage and vocal characters.

P. m. **undescribed subspecies**
Distribution: Subtropical forest from the highlands between Rio Huallaga and Rio Ucayali All samples are from a single locatlity: 77 km WNW Contamana, Loreto, Peru; 7.08333˚ S, 75.65˚ W).
Phylogenetic Position: Clade C2 (Fig. 3)
Plumage: Not examined.
Lek Vocal Type: Unknown
Call Vocal Type: 13

P. m. microlopha (Zimmer 1929)
Distribution: Eastern Peru south of the Rio Marañón and Rio Huallaga west to Rio Juruá and Rio Purus, Brazil.
Phylogenetic Position: Apparently paraphyletic, Clade C1 excluding C2 (Fig. 3)
Plumage: Males are glossy black above, with black or dark gray bases to white crown feathers. Females are dark olive above, occasionally with slightly gray crown, olive below, and graying on the belly.
Lek Vocal Type: 7 (*jeer*)
Call Vocal Type: 13

P. m. **undescribed subspecies**
Distribution: Right (east) bank of the Rio Purus to the left (west) bank Rio Madeira
Phylogenetic Position: Clade C3 (Fig. 3)
Plumage: Not examined.
Lek Vocal Type: Unknown
Call Vocal Type: 13

P. m. **undescribed subspecies**

Distribution: Right (east) bank of the Rio Madeira to the left (west) bank the Rio Tapajos.
Phylogenetic Position: (Clade C4, Fig. 3)
Plumage: Not examined.
Lek Vocal Type: Unknown
Call Vocal Type: 13

***P. m.* undescribed subspecies**
Distribution: Right (east) bank of the Rio Tapajos to the left (west) bank of the Rio Xingu
Phylogenetic Position: Clade C5 (Fig. 3)
Plumage: Not examined.
Lek Vocal Type: 6b
Call Vocal Type: 13

P. m. separabilis (Zimmer 1936)
Distribution: Right (east) bank of the Rio Xingu east to central and southern Pará.
Phylogenetic Position: Clade C6 (Fig. 3)
Plumage: Males are moderately glossy above, crown long with large, dark gray feather bases. Predefinitive male plumage light olive above, gray below, with medium gray crown.
Females are light olive above, light grayish below with olive wash on the breast. Zimmer (1939) Zimmer (1939) commented that adult males and females not distinguishable from *separabilis,* but he identified the distinct predefinitive male plumage
Lek Vocal Type: 2
Call Vocal Type: 13

Pseudopipra cephaleucos (Thunberg 1822) **Atlantic White-crowned Manakin**
Distribution: Tropical forest from Bahia south to northern Rio de Janeiro, Brazil.
Phylogenetic Position: Clade C7 (Fig. 3)
Plumage: Males are glossy black with a long and slightly gray crown. Crown feather have extensive dark gray bases. Predefinitive males have olive backs, pure white or grayish white crowns, and slate gray on the face, throat, and belly. Females have olive back, dusky gray on head, gray below, slightly olive on breast, lighter on belly.
Lek Vocal Type: 6a (*zeeeee-tonk*)
Call Vocal Type: 13

Section 2
Additional results and discussion

Additional phylogenetic results and comments on mutational spectra

At the level the concatenated alignments, the '20% missing' dataset had a total of 2,548 132bp loci (340,956 sites), 7.92% missing sites, 8,063 parsimony informative sites, and 5,709 variable parsimony-uninformative sites. The '50% missing' dataset had a total of 4,763 132 bp loci (626,868 sites), 20.48% missing sites, 15,365 parsimony informative sites, and 11,221 variable parsimony-uninformative sites. The '80%' missing dataset had a total of 7,907 132 bp loci (1,039,632 sites), 38.89% missing sites, 24,450 parsimony informative sites, and 17,839 variable parsimony-uninformative sites. Across these datasets, alpha from the GTR+G model was << 1 (0.069, SD=6.033×10^{-4}), indicating high among site rate heterogeneity for these ddRAD loci. Chi-square tests of base compositional heterogeneity rejected the hypothesis of compositional homogeneity (Chi-sq= 7822.93, df=702, P << 0.05), with slight bias observed on the AT-GC axis of compositional variation (A: 0.24968, C: 0.25225, G: 0.24301, T: 0.24301, on the largest 80% dataset). Maximum likelihood estimates of transition rates were ~8x transversion rates (A<-->G: 7.34 x G <--> T, C <--> T: 8.13 x G <--> T), as estimated in RAxML. Estimated rates among other nucleotide classes were ~1 relative to the fixed G <--> T rate, suggesting that the available GTR model in RAxML is likely over-parameterized for ddRAD data.

Reconstruction of mitochondrial ND2 gene tree

After obtaining mitochondrial DNA sequences for 168 individuals (Supplementary Table 1), we aligned these sequences using MAAFT (Standley and Katoh 2013). The alignment was visually inspected in Sequencher (Gene_Codes_Corporation 2010), and then analyzed in IQ-TREE 1.6.10 (Schmidt et al. 2014, Chernomor et al. 2016, Trifinopoulos et al. 2016, Hoang et al. 2017, Kalyaanamoorthy et al. 2017). We partitioned by codon position and generated a maximum likelihood tree using the MFP+MERGE model search and partitioning option, with 1000 ultrafast bootstrap replicates. MFP+MERGE detected that an optimal scheme comprised of three partition-models for each of the three codon positions (CP1: TIM2+F+I; CP2: TIM2+F+G4; CP3: TIM2+F+G4). Nodes recovered with ultrafast bootstrapped < 95 were collapsed. The recovered topology was entirely congruent with the topology presented in the main text as derived from ddRAD data, with a few exceptions (Supplementary Figure 14). Our mtDNA dataset included individuals from subspecies *coracina* and *pygmaea* which were derived from low quality tissue samples (and hence were not suitable for ddRAD sequencing). This enabled us to make a preliminary assessment of their phylogenetic affinities (main text), though nuclear genomic data should be collected in future studies. Notably, the introgressed western Napo lineage has mtDNA haplotypes which are members of the southern amazon clade (BS 98), which is consistent with the scenario of hybrid origin and introgression we develop in the main text. Because mtDNA is inherited matrilineally, a potential implication of this pattern is that the introgressed Napo lineage (S2a/S2 in Figure 6) was created when southern progenitor females were introgressed with northern males.

Reconstruction of ancestral elevational habit

We performed a Bayesian stochastic character mapping analysis (Huelsenbeck et al. 2003, Bollback 2006) to estimate the ancestral habit of *Pseudopipra*. In brief, we coded lineages as montane (>1000m) or lowland (< 1000m), applied a bi-directional Mk model

('ARD') and performed 100 simulations (see supplemental R code) using the RAxML topology. We used the SIMMAP implementation in phytools (Revell 2012). These analyses unambiguously reconstructed the ancestral habit of *Pseudopipra* to be montane.

STRUCTURE - additional results

The Evanno method applied to the whole dataset detected a significant shift in the rate of change of the log probability of the data between K1 and K2, indicating a deep hierarchical split in the data. As STRUCTURE infers the degree of admixture among individuals, this assignment is not directly comparable to K-means phenetic cluster solutions, which lump individuals categorically based on overall genetic similarity. That said, there was broad overlap in cluster assignment.

Descriptive Population Genetic Statistics – methods

For population genetic statistics, we considered eighteen population-areas (Figure 3, 4). Most of these populations are delimited by clear geographic barriers (e.g., rivers in the cases of previously identified areas of endemism, the Andes, or the Cerrado belt) and have strong phylogenetic support. Two subgroups within the broad Northern Amazonian + Guiana Shield lowland clade were defined on the basis of low-support monophyly in the RAxML analysis *and* coincidence with geographic features. One of these comprised individuals unambiguously assigned to northern Amazonian clade in phylogenetic analysis, but which were also restricted to the eastern Napo area of endemism, east of the Rio Putumayo (brown dots in Figures 3, 4, 'weakly resolved eastern Napo' – abbreviated in R code and Supplementary Figures as 'GSNapo'). The second comprised individuals found near the coasts in Suriname and the Brazilian state of Amapá, east of the Essequibo river (pale blue dots in Figures 3, 4: 'Suriname + Amapá' – abbreviated in R code and Supplementary Figures as 'GSSR'). Another subgroup was defined on the basis of restriction to the Jaú area of endemism (pale yellow dots in Figures 3, 4: 'unresolved Jaú' – abbreviated in R code and Supplementary Figures as 'GSImeri'). Lastly, a fourth group of individuals included all other individuals in the lowland northern Amazon clade, restricted to the Guiana Shield (green dots in Figure 3, 4: 'weakly resolved Guiana Shield' – abbreviated in R code and Supplementary Figures as 'GS'), comprising individuals east of the Jaú group (above), and west of those in the Suriname + Amapá group. The primary geographic barriers in this region separating western and eastern Guiana Shield populations seems to be the Guiana Highlands, which is where tepuis are found, as well as the Essequibo river.

For these descriptive analyses, we focus on the aforementioned eighteen areas as units of comparison because focusing on broader populations delimited by cluster analyses would likely generate statistics biased by population sub-structure—i.e, lower than expected heterozygosities (Wahlund 1928). Further, groups delimited by broader cluster assignments may be more reflective of ancestral populations, and therefore not indicative of presently restricted groups (ie, inappropriately moving migrants back to their source populations, Kuhner 2006). Statistics were calculated using dataset 2, as this includes that largest number of putatively unlinked markers (the first SNP from each of 2,581 ddRAD loci), unless otherwise indicated.

To estimate a measure of genetic diversity across these sampling regions, we calculated the rarefied allelic richness per population (restricted to populations comprising > 5 individuals) using the allelic.richness function in the hierfstat R package (Goudet 2005), after removing all sites with missing genotypes (Supplemental R Script).

We also calculated the inbreeding coefficient F_{IS}, defined as $(H_S - H_I)/H_S$, where H_I is the mean expected heterozygosity per individual within subpopulations, and H_S is the mean expected heterozygosity within random mating populations (Goudet 2005). We generated 100,000 bootstrapped estimates of F_{IS}, sampling over loci per population, using the boot.ppfis function in hierfstat (Goudet 2005). For recently hybrid individuals, F1s should be more outbred (relative heterozygosity) than their parental genotypes. We tested the hypothesis that the introgressed western Napo population is composed of recently introgressed individuals by estimating the inbreeding coefficient for a simulated F1 population, comprised of the progenitor lineages discussed in the main text. We generated a simulated F1 population using the hybridize function in adegenet (Jombart 2008), and then estimated it's inbreeding coefficient as described above to compare to empirical estimates from source populations.

To perform a preliminary assessment of the potential for evolutionary processes deviating from the assumptions of Hardy-Weinberg equilibrium, we applied the hw.test function in the pegas R package (Paradis 2010) with 1000 Monte Carlo permutations of alleles to compute an exact p value for each locus within each population. To assess the assumption of linkage intrinsic to most model-based analyses in this study, we computed the Standardized Index of Association $\bar{r}_d$ (Brown et al. 1980, Agapow and Burt 2001) within populations using the poppr summary function in the poppr R package (Kamvar et al. 2014), and estimated p values with 1000 permutations. We estimated pairwise Weir and Cockerham's (Weir and Cockerham 1984) F_{st} among all 18 areas, and evaluated significance using 1000 bootstrapped datasets to estimate 95% confidence intervals using the 'assigner' R package (Gosselin et al. 2016).

Lastly, we quantified differentiation among two hierarchical strata recapitulating 1) deep coalescent structure (6 groups as identified by SVDquartets (~K5 from STRUCTURE + putative introgressed Napo hybrids as a separate group), and 2) populations identified in phylogenetic analyses which coincide with geographic barriers (18 groups), with analysis of molecular variance (AMOVA) (Excoffier et al. 1992). We used the poppr.amova wrapper function in the poppr R package (Kamvar et al. 2014) to perform AMOVA on adegenet genind objects, set to use the ade4 implementation of AMOVA with 1000 permutations to assess significance. For AMOVA calculations we used dataset 1, to minimize within individual variance.

Descriptive population genetic statistics – results

Missing data (dataset2) was quite low across areas (mean: ~9.5%, SD: 5.3% and ranged from ~2.4% (Jaú subgroup of the Guiana Shield clade) to a maximum of 20.5% (Costa Rica, though this was somewhat of an outlier – 75% of these areas had less than 13% missing data overall). Despite the fact that our sampling scheme among areas delimited by geographic boundaries had high variance relative to the mean (mean: 12.94 [1 - 70], SD: 16.95284, CoV: 1.31), the sum rarefied estimates of allele counts in each of 13 areas (with > 5 individuals, and after filtering out all sites with missing genotypes) were

similar. For dataset 2 (2581 SNPs): mean number of alleles: 297.76, SD: 6.44, CoV: 0.022. The greatest allelic diversity was found in the Jaú (n=14, 305.75 alleles) and introgressed western Napo population (n=10, 303.8 alleles). The lowest allelic richness was found to be in the Rio (n=10, 287.15) and Bahia (n=6, 285.48 alleles) Atlantic Forest populations, followed closely by Panamanian populations (n=5, 292 alleles). These results are generally consistent with our EEMS analysis (Supplementary Figure 12). Summary table below:

Population	**Alleles**
Atlantic Forest (Bahia)	285.4848
Atlantic Forest (Rio)	287.1516
Central America (Panama)	292.0000
South Andean Peru (South)	294.0000
Eastern Inambari endemic	297.4178
Xingu endemic	297.5155
Weakly resolved Guiana Shield (western)	299.4562
Weakly resolved Suriname + Amapá	299.5588
Inambari endemic (western)	302.4655
Weakly resolved eastern Napo	302.4864
Tapajós endemic	303.7529
Western Napo introgressed lineage	303.7964
Unresolved Jaú	305.7571

Most populations were detected to be significantly inbred ($F_{IS} > 1$**,** Supplementary Figure 9), with lower 95% confidence intervals > 0. Panamanian, Costa Rican, South Andean (North clade), Rondônia, and Espírito Santo clades had 95% confidence intervals which overlapped zero, and thus cannot be confidently inferred to have positive or negative F_{IS}. The simulated F1 population, however, did have significantly negative F_{IS}, as predicted. This pattern implies that the introgressed western Napo population, which was detected to have a significantly positive F_{IS}, is not likely to include recently introgressed individuals. Indeed, the confidence intervals for eastern Napo, Jaú, Inambari and western Napo popularions, are generally overlapping, with similar means (mean of mean estimates ~0.17, SD of mean estimates ~0.02, Supplementary Figure 9).

Pairwise population estimates of Weir and Cockerham's F_{st} ranged from essentially undifferentiated, to almost entirely distinct. At the most extreme: comparing the geographically proximate eastern Napo and Jaú populations (both weakly resolved in phylogenetic analyses, but likely sister) --F_{st}: 0.0045. By contrast, comparing an Atlantic forest Espírito Santo population to a population in Panama indicates an F_{st} of 0.81, or almost entirely differentiated. Overall, population average F_{st} was very high: 0.196 [0.188-0.204] (Supplementary Figure 10).

After correcting for multiple tests with the Benjamin & Hochberg correction, exact tests of Hardy-Weinberg equilibrium suggested most loci in most populations were in equilibrium. However, a small number of loci in the western Guiana group (123 loci), Suriname+Amapá (53 loci) and Tapajós (14 loci) areas were identified as being out of Hardy-Weinberg equilibrium. Estimates of $\bar{r}_d$ within these populations indicated that there

was no strong evidence of linkage among loci within populations, except for the Tapajós area, in which weak linage was detected ($\bar{r}_d$: 0.005956, p= 0.000999).

Lastly, an AMOVA detected significant population differentiation at all evaluated levels, including between coalescent units (well supported clades form SVDquartets) (~32%) and between samples within coalescent units (~5%) ($p < 0.001$ for all). Re-running the same AMOVA with evolutionary distances estimated with RAxML branch lengths (instead of the default allelic distance) indicated the same pattern, but with more of the variance explained by coalescent and population level strata (41.3% and 12.6% respectively). Both AMOVA analyses detected a significant proportion of the variance attributable to within sample variance (62% and 46% respectively).

Isolation by distance and the effect of geography

The evolutionary history of *Pseudopipra* within the Amazon basin appears to be deeply connected to the South American landscape, adding additional support to a rich body of literature endorsing this hypothesis (Cracraft and Prum 1988, Brumfield 2012). For virtually all evaluated cases, we find significant effects of geographic barriers on structuring genetic variation within this species complex, including the Amazon River and most associated tributaries (Supplementary Table 2b and Figure 7). Further afield, the 'dry-diagonal' Cerrado belt appears to have strongly isolated Atlantic Forest lineages from their southeastern Amazonian Xingu relatives, as do the Andes exhibit a disproportionate effect on divergence between Peruvian foothills populations and Central American lineages (with the caveat that our sampling in that area is sparse, so our power to infer spatial patterns is necessarily limited).

The establishment of the Amazonian river system has recently been questioned as a driver of species—level variation across key areas in the Neotropics (Oliveira et al. 2017, Santorelli et al. 2018). These recent studies used distributional data to infer the effects of key proposed barriers and concluded that while large rivers clearly limit some Amazonian species—the large number of exceptions to this 'rule' point towards alternative speciation mechanisms as the norm, rather than as the exception. Indeed, rivers can plausibly function as contemporary species limits without being the source of such limits (Santorelli et al. 2018). In the case of *Pseudopipra,* river barriers have clearly contributed to contemporary patterns of genetic diversity, regardless of whether or not the formation of the Amazonian drainage system was the primary driver of generating that diversity. Importantly, studies which rely on distributional data alone are limited in that their statistical power is entirely contingent on the accuracy of species and subspecies delimitation. In the biogeographic context of the Amazon, this is likely to be enormously underestimated for birds (Brumfield 2012, Smith et al. 2014). This fundamental limitation in our knowledge of cryptic avian diversity is therefore likely to bias inferences derived from distributional data, which is based on mostly untested species limits. Indeed, most studies that use genetic data to investigate the effect of river or other physical barriers in structuring Neotropical avian diversity have inferred strong, though varying effects (e.g. Moore et al. 2008, Harvey and Brumfield 2015, Naka and Brumfield 2018).

A number of authors have also noted that the practice of identifying genetic clusters with model based approaches often fail to appropriately account for the effects of isolation by distance (Guillot et al. 2013), and various methods are in development to improve our

ability to model such correlated phenomena (Bradburd et al. 2013, Botta et al. 2015, Petkova et al. 2015, Bradburd et al. 2017). STRUCTURE in particular has been highlighted as potentially suffering from over-estimating K as a consequence of spatial autocorrelation in widely distributed genetic data (Bradburd et al. 2017). Our STRUCTURE analysis appears to exhibit this behavior for the southern Amazon, with a genetic cline of admixture that falls on a longitudinal gradient across the southern Amazon and ends in the well differentiated Atlantic Forest Rio population. While it is plausible that isolation by distance, combined with physical barriers to gene flow, could generate a similar pattern (as implied by our phylogenetic analyses), it is important to keep this caveat in mind when interpreting STRUCTURE results. For example, STRUCTURE may suggest that a scenario of K2 with an admixture gradient between two populations is preferred, when K1 with an isolation by distance effect may be a better description and more biologically plausible model for the data (Bradburd et al. 2017). The degree to which this kind of spatial autocorrelation confounds STRUCTURE-like analyses at large remains an open and important area of inquiry. Our EEMS analysis attempts to circumvent this issue entirely, assuming a more biologically realistic process of continuous differentiation across a heterogeneous landscape, however it does not provide unambiguous insight into hypotheses of species delimitation.

Notes on congruent patterns with Ceratopipra

Within the manakins, a recent molecular phylogeny (Ohlson et al. 2013) placed *Pseudopipra* as sister to the genus *Ceratopipra*, which includes five well-recognized species that are extensively codistribtuted with *Pseudopipra.* The breakpoints among these *Ceratopipra* species are highly concordant with the breakpoints among the genetic clusters within the *Pseudopipra* complex that we have presented here, implying that these taxa have many components of their phylogeographic history in common. *Pesudopipra* is the Andean sister group to the lowland Ceratopipra, which has itself expanded into montane habitats twice (*corunta* and *chloromeros*). By contrast, *Pseudopipra* expanded from the Andes into the lowlands.

Ceratopipra erythrocephala is distributed in the northern Amazon, and *C. rubrocapilla* has a range encompassing the southern Amazon and the Atlantic Forest. *C. mentalis* is distributed in Central America and south-ward into the Chocó and the western edges of Columbia and Ecuador *C. chloromeros* has a narrow distribution in the lower montane forests of the southern Peruvian and northern Bolivia Andes. The distributions of *C. erythrocephala* and *rubrocapilla* are extensively with the Guianan Shield and Southern Amazonian clades of *Pseudopipra*. However, the *Pseudopipra* radiation also has some important differences from *Ceratopipra. C. cornuta* is distributed in montane forests of tepuis in Venezuela and western Guyuna, at altitudes where *Pseudopipra* does not occur. In contrast, *Pseudopipra* has extensive montane populations in the Andes from Peru to Colombia, and *C. chloromeros* is only distributed in the easterns slope of the Andes in Peru and Bolivia. *C. mentalis* is found in lowland tropical forest at lower altitudes than the lower montane populations of *Pseudopipra* in Central America. Furthermore, the Chocó population of *Pseudopipra* is also lower montane in distribution, and not continuous with Central America. Lastly, the phylogenetic relationships among the differentiated lineages of *Ceratopipra* and *Pseudopipra* are not congruent. In *Ceratopipra,* the northern and southern

Amazonian lineages are not sister taxa. Rather, the northern Amazonian *erythrocephala* is sister to the Central American and Chocó *mentalis,* and southern Amazonian *rubrocapilla* is sister to the Andean *chloromeros* (Ohlson 2013).

Vocal variation

Pseudopipra vocalizations have 1-3 buzzy or tonal notes. We measured: 1) starting frequency, 2) ending frequency, 3) minimum frequency, 4) maximum frequency, 5) number of notes, and 6) duration of the entire vocalization (see Supplementary Figure 13). To obtain a conservative estimate of the number of individuals sampled, we took measurements of one vocalization from each recording. When there were multiple recordings by the same recordist on the same day and location, only one of the recordings was measured. Some recordists raised the possibility that the tonal notes, particularly the 'tonk' in vocal type 1, may be a mechanical sound, but further research is required to determine which sounds are vocalizations and which are mechanical sonations. We performed principal components analysis (PCA) and logistic regression on the vocal measurements to test for significant differences between the vocal types and to reduce the dimensionality of the data for comparison to results from analysis of genetic data. The PCA analysis was performed using the princomp function and the logistic regression was performed using the glm function, both in the stats R package (R_Core_Team 2018). The geographic distribution of each vocal type was assessed using latitude and longitude coordinates included in the metadata of each recording. When no coordinates were available, we determined latitude and longitude based on the description of the locality. Because no sound records were directly associated with genetic samples in this study, we used geographic proximity to vocalization recordings and localization to areas of endemism or areas bounded by clear physical barriers to associate vocal types to genetic samples. This approach assumes that genetically and geographically proximate individuals are likely to share the same vocal type and enabled us to perform a preliminary assessment of how variation in vocalizations maps onto existing genetic variation. Testing the fine-scale association of genetic and vocalization boundaries will require extensive field sampling of both traits from individual manakins.

Literature Cited

Agapow PM, Burt A. 2001. Indices of multilocus linkage disequilibrium. Molecular Ecology Notes, 1:101-102.

Bollback JP. 2006. SIMMAP: Stochastic character mapping of discrete traits on phylogenies. BMC Bioinformatics, 7.

Botta F, Eriksen C, Fontaine MC, Guillot G, Yu D. 2015. Enhanced computational methods for quantifying the effect of geographic and environmental isolation on genetic differentiation. Methods in Ecology and Evolution, 6:1270-1277.

Bradburd G, Coop G, Ralph P. 2017. Inferring Continuous and Discrete Population Genetic Structure Across Space. bioRxiv.

Bradburd GS, Ralph PL, Coop GM. 2013. DISENTANGLING THE EFFECTS OF GEOGRAPHIC AND ECOLOGICAL ISOLATION ON GENETIC DIFFERENTIATION. Evolution, 67:3258-3273.

Brown AHD, Feldman MW, Nevo E. 1980. Multilocus Structure of Natal Populations of Hordeum Spontaneum. Genetics, 96:523-536.

Brumfield RT. 2012. Inferring the origins of lowland Neotropical birds. The Auk, 129:367-376.

Chernomor O, Minh BQ, von Haeseler A. 2016. Terrace Aware Data Structure for Phylogenomic Inference from Supermatrices. Systematic Biology, 65:997-1008.

Cracraft J, Prum RO. 1988. PATTERNS AND PROCESSES OF DIVERSIFICATION: SPECIATION AND HISTORICAL CONGRUENCE IN SOME NEOTROPICAL BIRDS. Evolution, 42:603-620.

Excoffier L, Smouse PE, Quattro JM. 1992. Analysis of Molecular Variance Inferred from Metric Distances among DNA Haplotypes: Application to Human Mitochondrial DNA Restriction Data. Genetics, 131:479-491.

Gene_Codes_Corporation. 2010. Sequencher version 5.0 sequence analysis software. Ann Arbor Michigan.

Gosselin T, Anderson EC, Ferchaud A-L. 2016. thierrygosselin/assigner: v.0.4.0 (Version 0.4.0). Zenodo. http://doi.org/10.5281/zenodo.197418.

Goudet J. 2005. hierfstat, a package for r to compute and test hierarchical F-statistics. Molecular Ecology Notes, 5:184-186.

Guillot G, Schilling R, Porcu E, Bevilacqua M. 2013. Validity of covariance models for the analysis of geographical variation. arXiv:1311.4136.

Harvey MG, Brumfield RT. 2015. Genomic variation in a widespread Neotropical bird (Xenops minutus) reveals divergence, population expansion, and gene flow. Molecular Phylogenetics and Evolution, 83:305-316.

Hoang DT, Vinh LS, Chernomor O, Minh BQ, von Haeseler A. 2017. UFBoot2: Improving the Ultrafast Bootstrap Approximation. Molecular Biology and Evolution, 35:518-522.

Huelsenbeck JP, Nielsen R, Bollback JP. 2003. Stochastic Mapping of Morphological Characters. Systematic Biology, 52:131-158.

Jombart T. 2008. adegenet: a R package for the multivariate analysis of genetic markers. Bioinformatics, 24:1403-1405.

Kalyaanamoorthy S, Minh BQ, Wong TKF, von Haeseler A, Jermiin LS. 2017. ModelFinder: fast model selection for accurate phylogenetic estimates. Nature Methods, 14:587.

Kamvar ZN, Tabima JF, Grünwald NJ. 2014. Poppr: an R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction. PeerJ, 2:e281.

627 Kuhner MK. 2006. LAMARC 2.0: maximum likelihood and Bayesian estimation of population
628 parameters. Bioinformatics, 22:768-770.
629 Moore RP, Robinson WD, Lovette IJ, Robinson TR. 2008. Experimental evidence for extreme
630 dispersal limitation in tropical forest birds. Ecology Letters, 11:960-968.
631 Naka LN, Brumfield RT. 2018. The dual role of Amazonian rivers in the generation and
632 maintenance of avian diversity. Science Advances, 4.
633 Ohlson JI, Fjeldså J, Ericson PGP. 2013. Molecular phylogeny of the manakins (Aves:
634 Passeriformes: Pipridae), with a new classification and the description of a new genus.
635 Molecular Phylogenetics and Evolution, 69:796-804.
636 Oliveira U, Vasconcelos MF, Santos AJ. 2017. Biogeography of Amazon birds: rivers limit
637 species composition, but not areas of endemism. Scientific Reports, 7:2992.
638 Paradis E. 2010. pegas: an R package for population genetics with an integrated–modular
639 approach. Bioinformatics, 26:419-420.
640 Petkova D, Novembre J, Stephens M. 2015. Visualizing spatial population structure with
641 estimated effective migration surfaces. Nature Genetics, 48:94.
642 R_Core_Team. 2018. R: A language and environment for statistical computing.
643 Revell LJ. 2012. Phytools: An R package for phylogenetic comparative biology (and other
644 things). Methods in Ecology and Evolution, 3:217-223.
645 Santorelli S, Magnusson WE, Deus CP. 2018. Most species are not limited by an Amazonian
646 river postulated to be a border between endemism areas. Scientific Reports, 8:2294.
647 Schmidt HA, Minh BQ, von Haeseler A, Nguyen L-T. 2014. IQ-TREE: A Fast and Effective
648 Stochastic Algorithm for Estimating Maximum-Likelihood Phylogenies. Molecular
649 Biology and Evolution, 32:268-274.
650 Smith BT, McCormack JE, Cuervo AM, Hickerson MJ, Aleixo A, Cadena CD, Perez-Eman J,
651 Burney CW, Xie X, Harvey MG, Faircloth BC, Glenn TC, Derryberry EP, Prejean J,
652 Fields S, Brumfield RT. 2014. The drivers of tropical speciation. Nature, 515:406-409.
653 Standley DM, Katoh K. 2013. MAFFT Multiple Sequence Alignment Software Version 7:
654 Improvements in Performance and Usability. Molecular Biology and Evolution, 30:772-
655 780.
656 Trifinopoulos J, Nguyen L-T, Minh BQ, von Haeseler A. 2016. W-IQ-TREE: a fast online
657 phylogenetic tool for maximum likelihood analysis. Nucleic Acids Research, 44:W232-
658 W235.
659 Wahlund S. 1928. ZUSAMMENSETZUNG VON POPULATIONEN UND
660 KORRELATIONSERSCHEINUNGEN VOM STANDPUNKT DER
661 VERERBUNGSLEHRE AUS BETRACHTET. Hereditas, 11:65-106.
662 Weir BS, Cockerham CC. 1984. ESTIMATING F-STATISTICS FOR THE ANALYSIS OF
663 POPULATION STRUCTURE. Evolution, 38:1358-1370.
664

Supplementary Figures and Tables

Supplementary Table 1. Specimen data table (separate file)

Supplementary Table 2a. Results from Mantel tests

Supplementary Table 2a.

Region	locality code	mantel r	two-tailed p	lower 2.5% limit	upper 97.5% limit	log.d	perm	significant (p < 0.001)
Full dataset		0.748	0.0001	0.729	0.763	F	10000	*
Central America - Costa Rica	CACR	0.459	0.2551	0.017	1.000	F	10000	-
Central America - Panama	CAPA	0.694	0.1019	0.021	0.921	F	10000	-
North Andean – Marañón	CAMA	NA	NA	NA	NA	F	10000	-
South Andean Peru (North)	CPN	-0.261	1	-0.261	-0.261	F	10000	-
South Andean Peru (South)	CPS	0.816	0.0993	0.618	0.998	F	10000	-
weakly resolved Guiana Shield	GS	0.099	0.0947	0.058	0.134	F	10000	-
unresolved Jaú	GSIMERI	0.096	0.243	0.043	0.166	F	10000	-
weakly resolved eastern Napo	GSNAPO	0.313	0.0513	0.177	0.560	F	10000	-
weakly resolved Suriname + Amapá	GSSR	0.069	0.2483	0.032	0.121	F	10000	-
Western Napo introgressed lineage	PH	0.508	0.017	0.340	0.754	F	10000	-
Western Inambari endemic	INAMBARI	0.608	0.001	0.382	0.848	F	10000	*
Eastern Inambari endemic	INAMBARIE	0.004	0.981	-0.196	0.244	F	10000	-
Rondônia endemic	RONDONIA	NA	NA	NA	NA	F	10000	-
Tapajós endemic	TAPAJOS	0.250	0.0148	0.182	0.313	F	10000	-
Xingu endemic	XINGU	0.079	0.8572	-0.170	0.275	F	10000	-
Atlantic Forest – Bahia	AFBAHIA	0.484	0.0173	0.173	0.804	F	10000	-
Atlantic Forest – Espírito Santo	AFES	NA	NA	NA	NA	F	10000	-
Atlantic Forest – Rio	AFRIO	0.701	0.0007	0.615	0.782	F	10000	*

Region	locality code	mantel r	two-tailed p	lower 2.5% limit	upper 97.5% limit	log.d	perm	significant (p < 0.001)
Full dataset		0.391	0.0001	0.379	0.401	T	10000	*
Central America - Costa Rica	CACR	0.459	0.253	0.017	1.000	T	10000	-
Central America - Panama	CAPA	0.519	0.2532	0.021	0.919	T	10000	-
North Andean – Marañón	CAMA	NA	NA	NA	NA	T	10000	-
South Andean Peru (North)	CPN	-0.261	1	-0.261	-0.261	T	10000	-
South Andean Peru (South)	CPS	0.810	0.032	0.669	0.998	T	10000	-
weakly resolved Guiana Shield	GS	0.091	0.0022	0.063	0.119	T	10000	-
unresolved Jaú	GSIMERI	0.536	0.0001	0.319	0.684	T	10000	*
weakly resolved eastern Napo	GSNAPO	0.313	0.0479	0.161	0.574	T	10000	-
weakly resolved Suriname + Amapá	GSSR	0.065	0.1054	0.026	0.117	T	10000	-
Western Napo introgressed lineage	PH	0.248	0.091	0.102	0.407	T	10000	-
Western Inambari endemic	INAMBARI	0.714	0.0008	0.499	0.952	T	10000	*
Eastern Inambari endemic	INAMBARIE	-0.004	0.9801	-0.304	0.213	T	10000	-
Rondônia endemic	RONDONIA	NA	NA	NA	NA	T	10000	-
Tapajós endemic	TAPAJOS	0.247	0.0011	0.115	0.353	T	10000	-
Xingu endemic	XINGU	0.079	0.854	-0.112	0.275	T	10000	-
Atlantic Forest – Bahia	AFBAHIA	0.481	0.0693	0.298	0.885	T	10000	-
Atlantic Forest – Espírito Santo	AFES	NA	NA	NA	NA	T	10000	-
Atlantic Forest – Rio	AFRIO	0.747	0.0012	0.676	0.912	T	10000	-

Supplementary Table 2b. Results from partial Mantel tests

Supplementary Table 2b.

Approximate Barrier	Comparison (populations)	partial mantel r	two-tailed p	lower 2.5% limit	upper 97.5% limit	log.d	perm	significant (p < 0.001)
Cordillera de Talamanca	Costa Rica vs Panama	-0.159	0.2972	-0.267	0.155	F	10000	-
Andes (1)	Central America vs Marañón	-0.591	0.0967	-0.729	-0.434	F	10000	-
Andes (2)	Central America vs (Marañón + South Andean Peru)	0.041	0.4030	-0.248	0.137	F	10000	-
Andes (3)	Central America vs (Everything else)	-0.267	0.0001	-0.295	-0.229	F	10000	*
Rio Ucayali	South Andean Peru vs Inambari	-0.906	0.0001	-0.940	-0.872	F	10000	*
Eastern Marañón + Hauallaga Rivers	Introgressed western Napo vs Inambari	-0.896	0.0001	-0.917	-0.878	F	10000	*
Rio Putumayo	Introgressed western Napo vs eastern Napo	-0.744	0.0001	-0.826	-0.702	F	10000	*
Rio Purus	Western Inambari vs eastern Inambari	-0.603	0.0001	-0.720	-0.044	F	10000	*
Rio Madeira	Eastern Inambari vs Rondônia	-0.550	0.0001	-0.697	-0.129	F	10000	*
Rio Tapajós	Rondônia vs Tapajós	-0.175	0.0441	-0.233	-0.114	F	10000	-
Rio Xingu	Tapajós vs Xingu	-0.318	0.0004	-0.383	-0.240	F	10000	*
Cerrado (1)	All pooled pops vs pooled Atlantic Forest	-0.183	0.0008	-0.229	-0.112	F	10000	*
Cerrado (2)	Xingu vs Bahia	-0.565	0.0001	-0.609	-0.434	F	10000	*
Rio Japurá	Eastern Napo vs Jaú	-0.151	0.0146	-0.297	-0.109	F	10000	-
Rio Negro	Jaú vs central Guiana Shield	-0.053	0.3107	-0.088	-0.012	F	10000	-
Rio Essequibo	central Guiana Shield vs eastern Guiana shield	-0.350	0.0001	-0.391	0.060	F	10000	*
Rio Amazonas	All pooled lowland N vs all pooled lowland S	-0.912	0.0001	-0.918	-0.907	F	10000	*

Approximate Barrier	Comparison (populations)	partial mantel r	two-tailed p	lower 2.5% limit	upper 97.5% limit	log.d	perm	significant (p < 0.001)
Cordillera de Talamanca	Costa Rica vs Panama	-0.744	0.0001	-0.802	-0.693	T	10000	*
Andes (1)	Central America vs Marañón	-0.996	0.0001	-0.997	-0.963	T	10000	*
Andes (2)	Central America vs (Marañón + South Andean Peru)	-0.829	0.0001	-0.942	-0.781	T	10000	*
Andes (3)	Central America vs (Everything else)	-0.464	0.0001	-0.495	-0.420	T	10000	*
Rio Ucayali	South Andean Peru vs Inambari	-0.910	0.0001	-0.938	-0.883	T	10000	*
Rio Marañón + Solimões	Introgressed western Napo vs Inambari	-0.880	0.0001	-0.900	-0.862	T	10000	*
Rio Putumayo	Introgressed western Napo vs eastern Napo	-0.900	0.0001	-0.941	-0.884	T	10000	*
Rio Purus	Western Inambari vs eastern Inambari	-0.833	0.0001	-0.863	-0.811	T	10000	*
Rio Madeira	Eastern Inambari vs Rondônia	-0.751	0.0001	-0.814	-0.650	T	10000	*
Rio Tapajós	Rondônia vs Tapajós	-0.316	0.0001	-0.388	-0.211	T	10000	*
Rio Xingu	Tapajós vs Xingu	-0.703	0.0001	-0.754	-0.655	T	10000	*
Cerrado (1)	All pooled pops vs pooled Atlantic Forest	-0.721	0.0001	-0.745	-0.702	T	10000	*
Cerrado (2)	Xingu vs Bahia	-0.934	0.0001	-0.952	0.070	T	10000	*
Rio Japurá	Eastern Napo vs Jaú	-0.399	0.0001	-0.490	-0.329	T	10000	*
Rio Negro	Jaú vs central Guiana Shield	-0.036	0.4825	-0.069	0.001	T	10000	-
Rio Essequibo	central Guiana Shield vs eastern Guiana shield	-0.470	0.0001	-0.503	-0.444	T	10000	*
Rio Amazonas	All pooled lowland N vs all pooled lowland S	-0.915	0.0001	-0.921	-0.911	T	10000	*

Supplementary Table 3: Song measures (separate file)

Supplementary Table 4: Song recording metadata (separate file)

Supplementary Table 5: G-PhoCS parameters

Supplementary Table 5a, θ: effective population size

	Western Napo	Eastern Napo	Inambari	MRCA Western Napo, Inambari	MRCA Inambari, Eastern Napo
median	4950840	2807467.5	2797062.5	5356300	1429540
95% HPD Interval LOW	4527457.5	2498760	2566870	4658055	1352200
95% HPD Interval HIGH	5378937.5	3132990	3044770	6047630	1512647.5

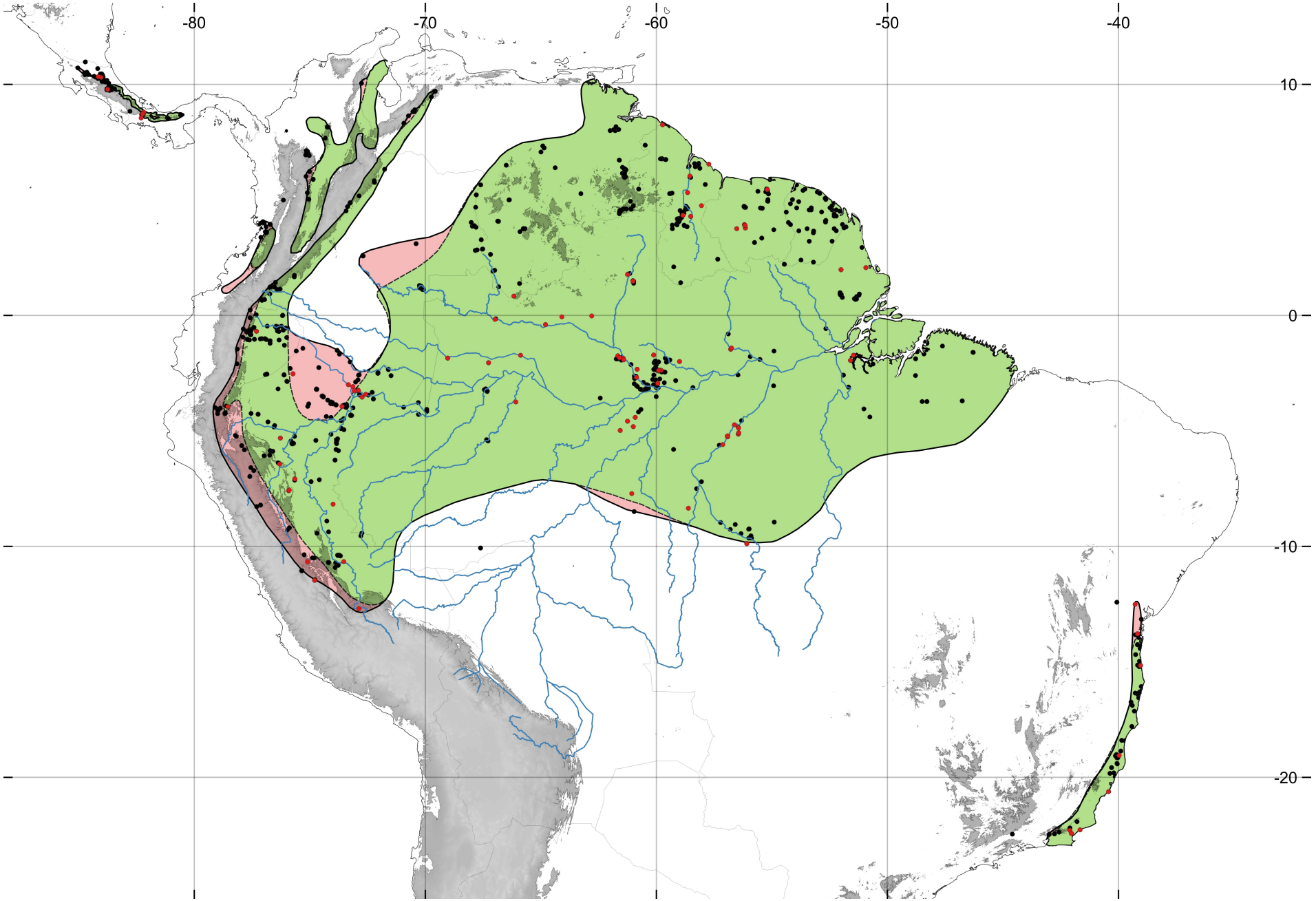
Supplementary Table 5b, τ: splitting time in generations

	MRCA Western Napo, Inambari, Eastern Napo	
median	484600	1025420
95% HPD Interval LOW	442610	954030
95% HPD Interval HIGH	532120	1099450

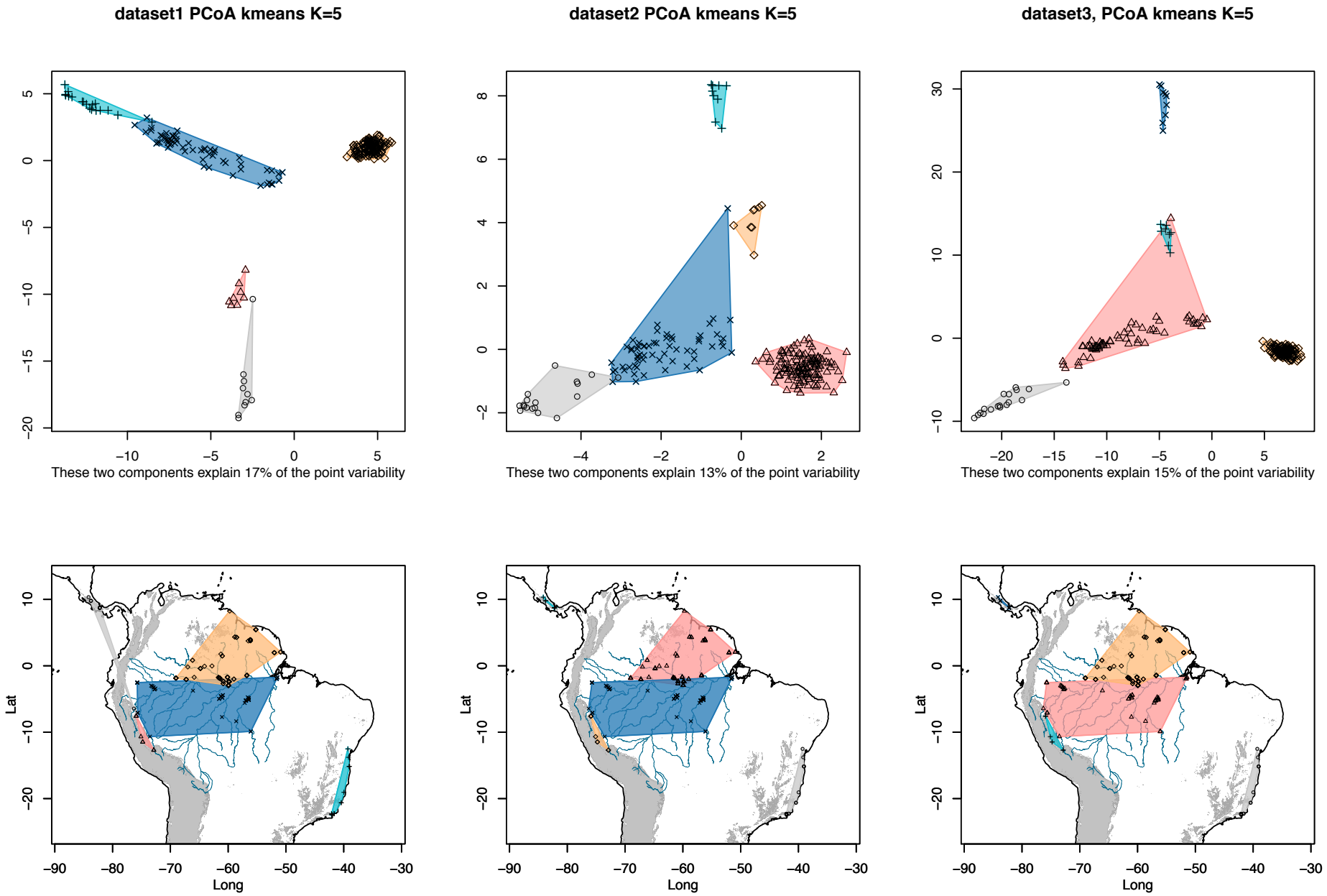
Supplementary Table 5c, m : migration rate (migrants per generation)

	Eastern Napo to	Inambari to Eastern	apo to West	Western Napo to Eastern Napo	Inambari to Western Napo	Western Napo to Inambari
median	0.0663	0.4196	1.7508	0.3671	0.0000	0.0000
95% HPD Interval LOW	0.0000	0.2670	1.0470	0.0000	0.0000	0.0000
95% HPD Interval HIGH	0.1763	0.6109	2.5872	0.8184	0.0056	0.0028

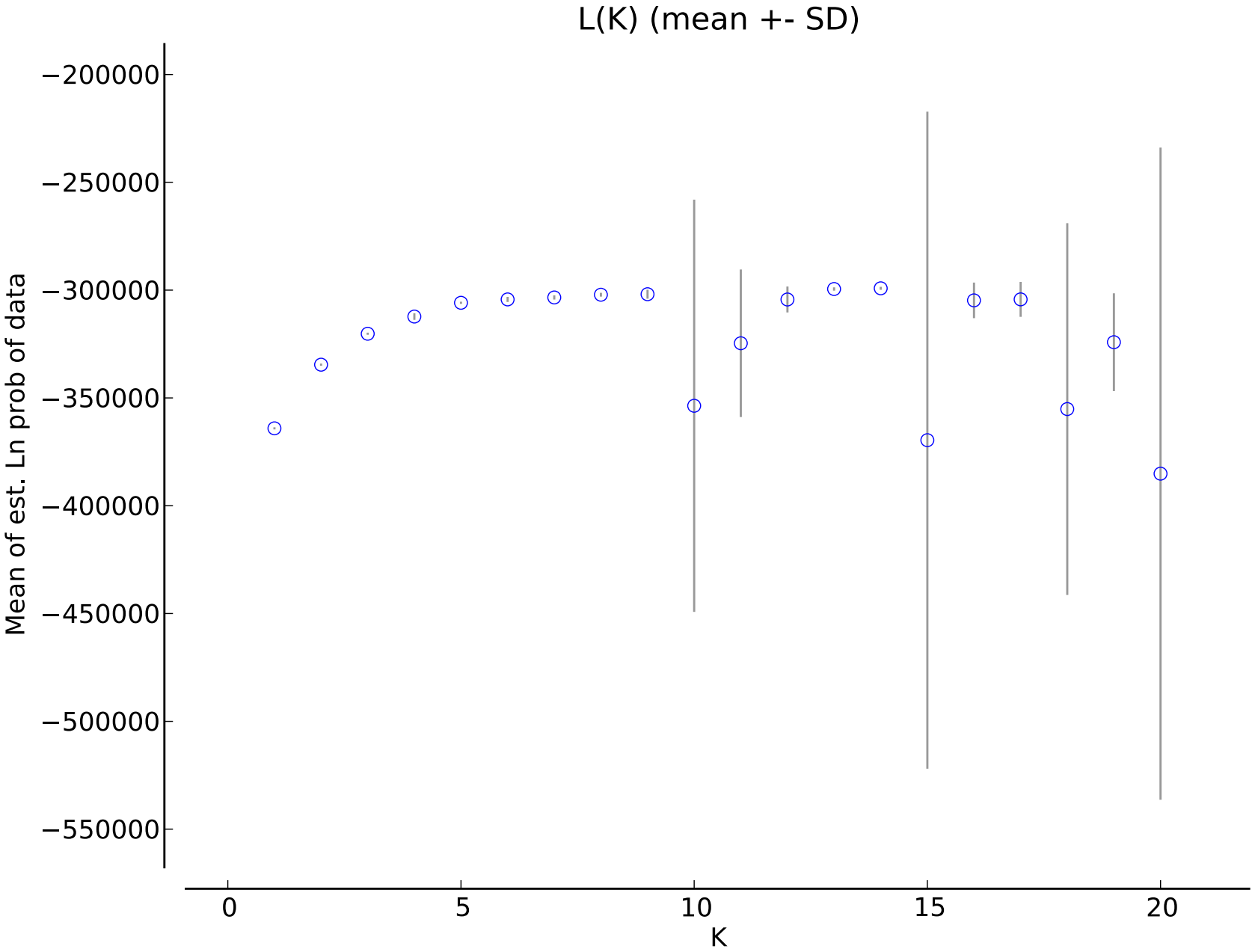
Supplementary Figure 1. GBIF occurrence records.
Here, we plot all GBIF occurrence records at the time of writing (in black) with our sampling localities (in red). The BirdLife approximate range map is shown in light green, and our modifications to this map are shown in pink to account for major inaccuracies in the available genus range map. This figure is provided primarily to illustrate that the BirdLife range map is inaccurate in the western Amazon, in Loreto, Peru, where our analyses detect an introgressed hybrid lineage.


-80
-70
-60
-50
-40
10
0
-10
-20

Supplementary Figure 2. K-means clustering of SNP data
In this figure, the PCoA projections of the SNP data are indicated on the top row, with minimum convex hulls (minimum implied range) and plotting symbols indicating the optimal K-means K5 clustering solution. PCoA explained 13-17% of the variance in the SNP data on the first two axes, and K-means clustering assignments derived from each dataset recovers nearly identical population assignments. Clustering of dataset 1 (1960 SNPs, 0.05 MAF) was identical to the clustering solutions for datasets 2 and 3, except for the assignment of one important individual (5444.PE.MAR), which links Central American lineages to our San Martín specimen in North Andean, Peru. Plotting symbols and colored convex hulls reflect cluster assignment (Hull colors are synonymous only across plotted columns, see supplementary R script). For datasets 2 (2581 SNPs) and 3 (5099 SNPs), K-means clustering detected the following groups: all Guiana Shield (Clade D in Figure 3), Atlantic Forest (Clade C7), South Andean Peru (Clade B), Southern Amazon including the western Napo population (Clade C + Clade E in Figure 3), and Central America (Clade A1 in Figure 3).

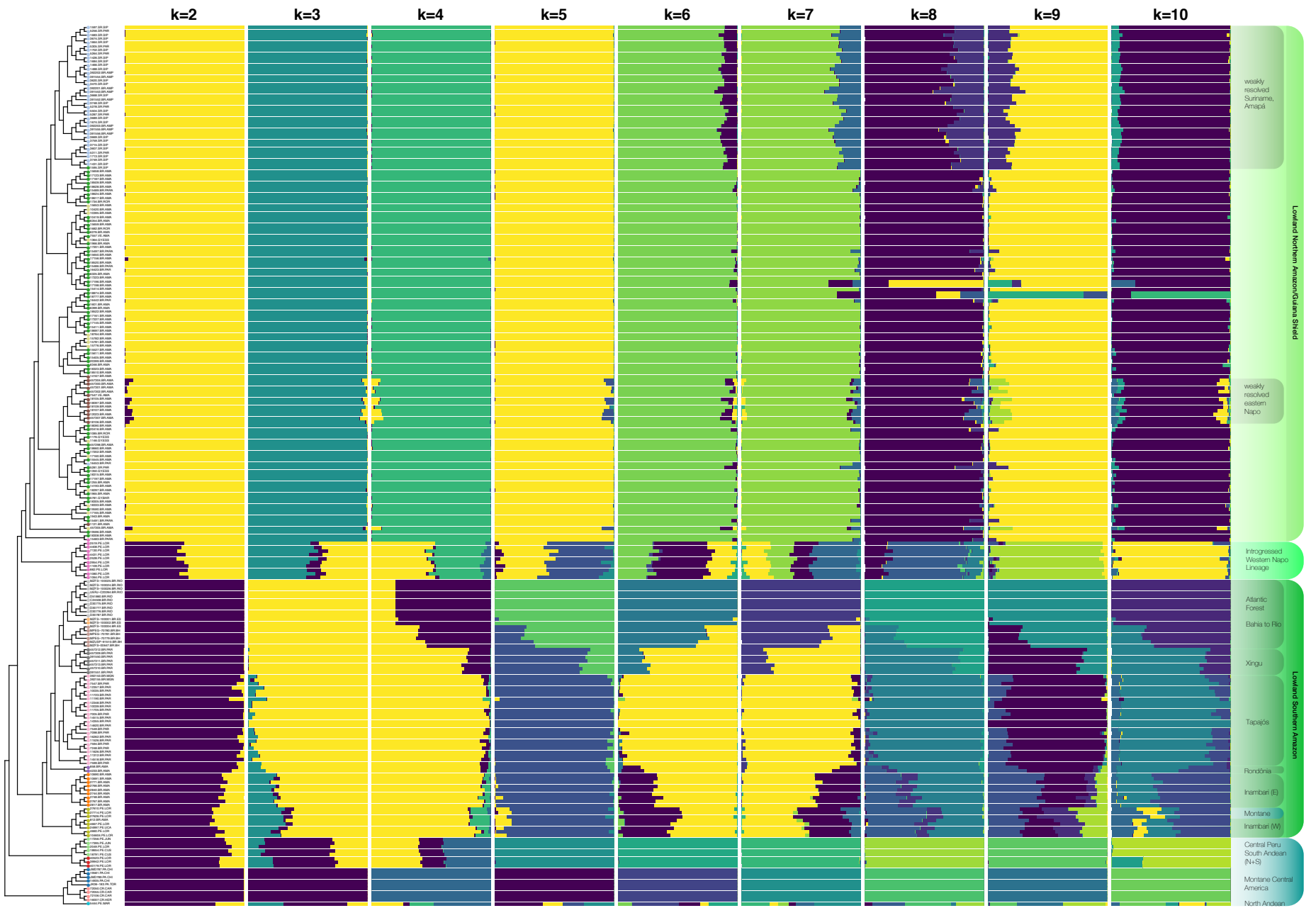

dataset1 PCoA kmeans K=5
These two components explain 17% of the point variability
dataset2 PCoA kmeans K=5
These two components explain 13% of the point variability
dataset3, PCoA kmeans K=5
These two components explain 15% of the point variability
Lat
Long

Supplementary Figure 3 – log likelihoods of STRUCTURE runs
Summarized log likelihood values across STRUCTURE runs for each value of K, with a plateau starting at K~5, and variance across runs increasing dramatically after K10.


L(K) (mean +- SD)
Mean of est. Ln prob of data
−200000
−250000
−300000
−350000
−400000
−450000
−500000
−550000
0
5
10
15
20
K

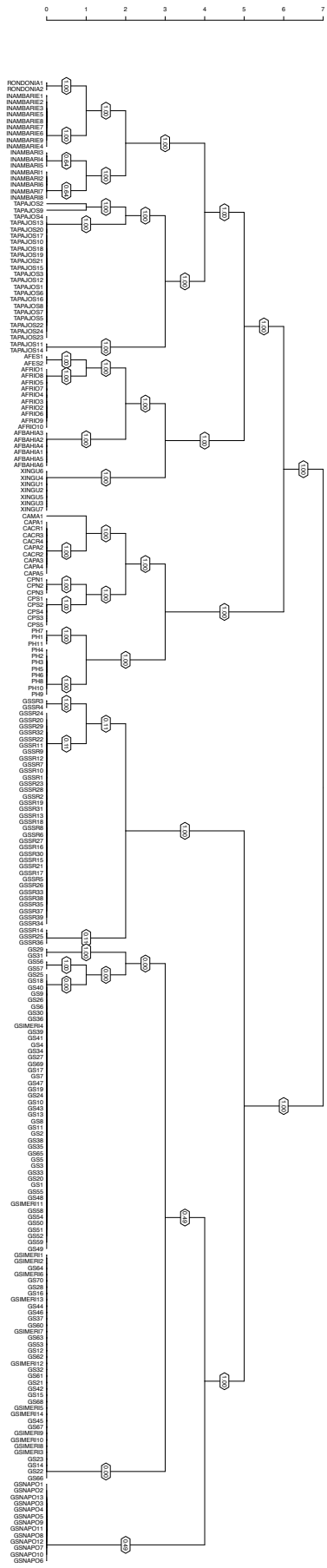
Supplementary Figure 4. STRUCTURE output for k2-10

Full STRUCTURE output for dataset 1, indicating population assignments and admixture for K2-10. The likelihood of each evaluated number of K clusters from 1:20 plateaued at K = 5, with the standard deviation across runs increasing rapidly after this point (Supplementary Figure 3). See results text for descriptions of these analyses. At K2, the first partition divides the dataset into broad northern and southern Amazonian groups, with all Andean and Central American samples assigned to predominantly southern Amazonian genetic provenance, with some northern admixture. Western Napo individuals are detected as an approximately even mixture of northern and southern Amazonian genomes. At K5, the five identified clusters broadly correspond to wide biogeographic Amazonian regions which encompass multiple areas of endemism (see text). For each barplot, colors are sampled randomly from a 20 color viridis color palette for each run (i.e., they are not synonymous across values of K, see supplementary R script). The tree below corresponds to the RAxML result using the 50% haplotype dataset, with tip labels and colors indicating group membership to one of eighteen population-areas. Colored tip labels correspond to clade label colors in Figures 3 and 4. At higher K, the broad-scale population assignments inferred at K5 are similar, however additional admixture components are inferred for most groups. The introgressed western Napo clade is eventually placed into its own cluster at K9-10.

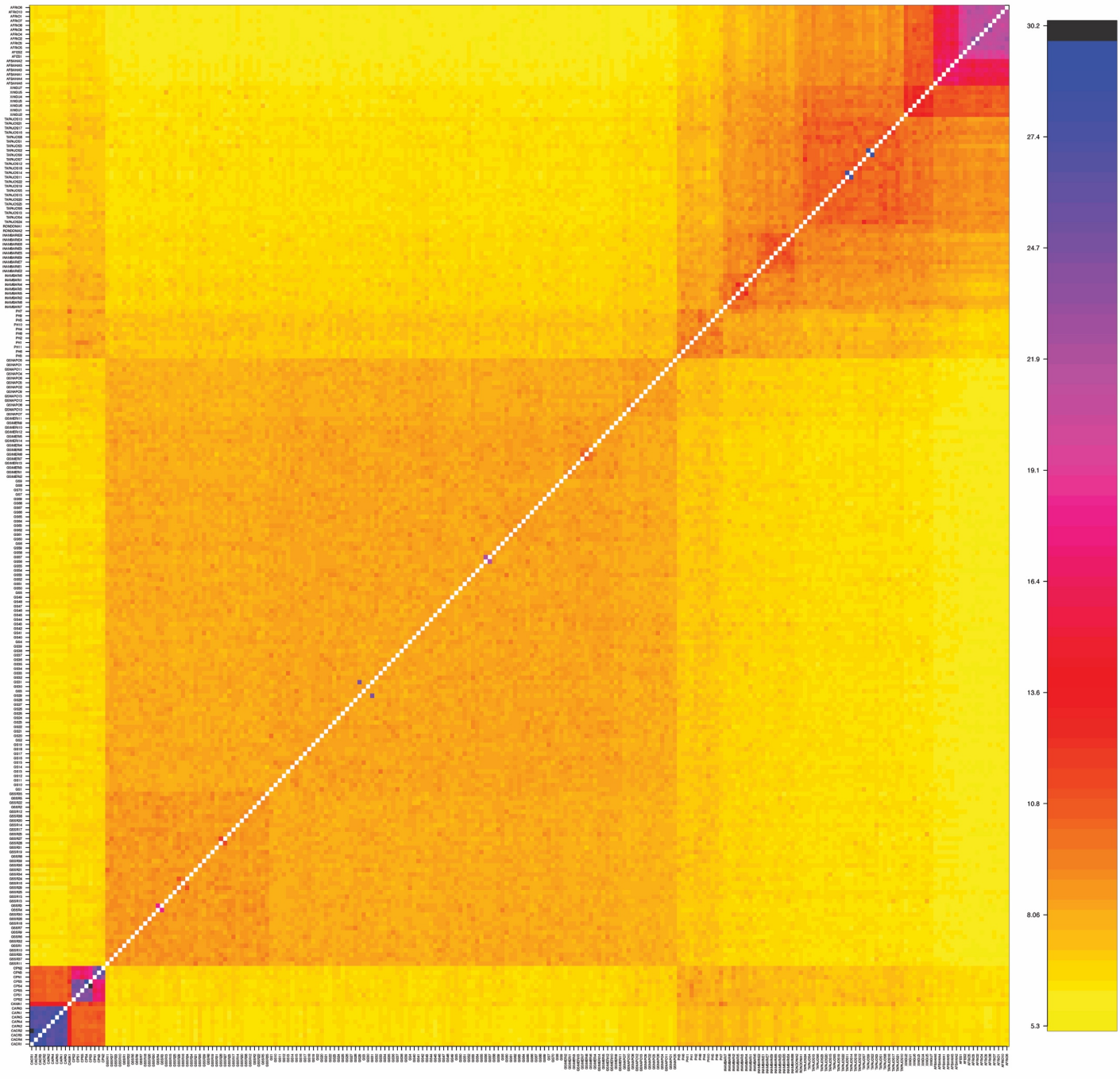

k=2
k=3
k=4
k=5
k=6
k=7
k=8
k=9
k=10
weakly resolved Suriname, Amapá
weakly resolved eastern Napo
Introgressed Western Napo Lineage
Lowland Northern Amazon/Guiana Shield
Atlantic Forest
Bahia to Rio
Xingu
Tapajós
Rondônia
Inambari (E)
Montane
Inambari (W)
Lowland Southern Amazon
Central Peru South Andean (N+S)
Montane Central America
North Andean

Supplementary Figure 5. fineRADstructure population assignment dendrogram
Clustering dendrogram generated from fineRADstructure population assignment. Note: this is not a phylogenetic hypothesis, but rather, a clustering based on genomic similarity which considers data from the full co-ancestry matrix. Tip labels correspond to population codes used internally for R scripts and other analyses. Each of these codes has a 1:1 correspondence with the labeled localities in Figures 3 and 4:

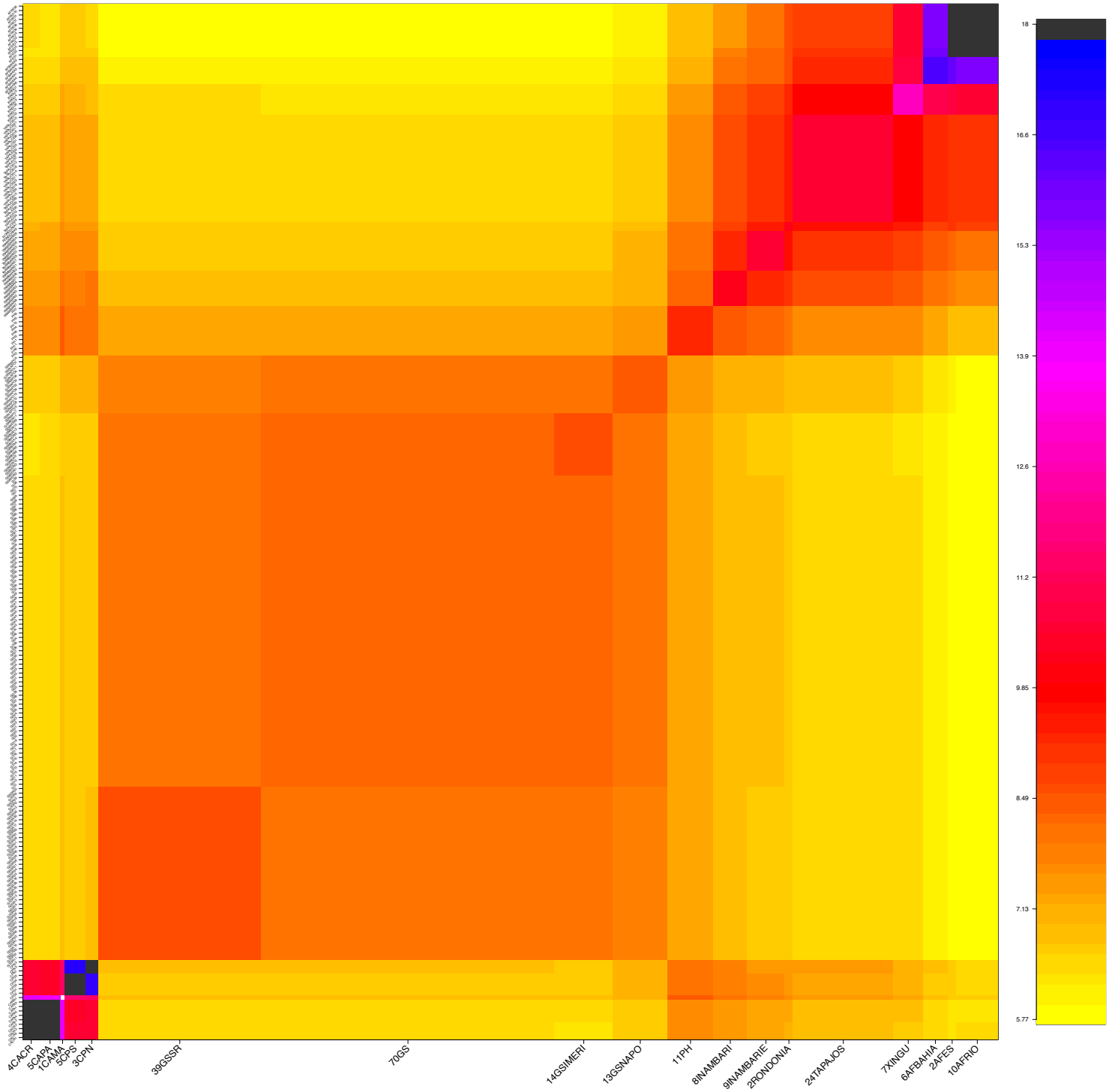
CAMA: North Andean – San Martín, Peru
CACR: Central America - Costa Rica
CAPA: Central America - Panama
CPS: South Andean Peru (South)
CPN: South Andean Peru (North)
INAMBARI: Western Inambari endemic
INAMBARIE: Eastern Inambari endemic
RONDONIA: Rondônia endemic
TAPAJOS: Tapajós endemic
XINGU: Xingu endemic
AFBAHIA: Atlantic Forest – Bahia
AFES: Atlantic Forest – Espírito Santo
AFRIO: Atlantic Forest – Rio
PH (putative hybrid): Western Napo introgressed lineage
GSIMERI: unresolved Jaú
GSNAPO: weakly resolved eastern Napo
GS: weakly resolved Guiana Shield
GSSR: weakly resolved Suriname + Amapá


0
1
2
3
4
5
6
7
RONDONIA1
RONDONIA2
INAMBARIE1
INAMBARIE2
INAMBARIE3
INAMBARIE5
INAMBARIE8
INAMBARIE7
INAMBARIE6
INAMBARIE9
INAMBARIE4
INAMBARI3
INAMBARI4
INAMBARI5
INAMBARI1
INAMBARI2
INAMBARI8
INAMBARI7
INAMBARI6
TAPAJOS2
TAPAJOS9
TAPAJOS4
TAPAJOS13
TAPAJOS20
TAPAJOS17
TAPAJOS10
TAPAJOS18
TAPAJOS19
TAPAJOS21
TAPAJOS15
TAPAJOS3
TAPAJOS12
TAPAJOS1
TAPAJOS6
TAPAJOS16
TAPAJOS8
TAPAJOS7
TAPAJOS5
TAPAJOS22
TAPAJOS24
TAPAJOS23
TAPAJOS11
TAPAJOS14
AFES1
AFES2
AFRIC1
AFRIC8
AFRIC5
AFRIC7
AFRIC4
AFRIC3
AFRIC2
AFRIC6
AFRIC9
AFRIC10
AFBAHIA3
AFBAHIA2
AFBAHIA4
AFBAHIA1
AFBAHIA5
AFBAHIA6
XINGU6
XINGU4
XINGU1
XINGU2
XINGU5
XINGU3
XINGU7
CAMA1
CAPA1
CACR1
CACR3
CACR4
CAPA2
CACR2
CAPA3
CAPA4
CAPA5
CPN1
CPN2
CPN3
CPS1
CPS2
CPS4
CPS3
CPS5
PH7
PH1
PH11
PH4
PH2
PH3
PH5
PH6
PH8
PH10
PH9
GSSR3
GSSR4
GSSR24
GSSR20
GSSR29
GSSR32
GSSR22
GSSR11
GSSR9
GSSR12
GSSR7
GSSR10
GSSR1
GSSR23
GSSR28
GSSR2
GSSR19
GSSR31
GSSR13
GSSR18
GSSR8
GSSR6
GSSR27
GSSR16
GSSR30
GSSR15
GSSR21
GSSR17
GSSR5
GSSR26
GSSR33
GSSR38
GSSR35
GSSR37
GSSR39
GSSR34
GSSR14
GSSR25
GSSR36
GS29
GS31
GS56
GS57
GS25
GS18
GS40
GS9
GS26
GS6
GS30
GS36
GSIMER84
GS39
GS41
GS4
GS34
GS27
GS69
GS17
GS7
GS47
GS19
GS24
GS10
GS43
GS13
GS8
GS11
GS2
GS38
GS35
GS65
GS5
GS3
GS33
GS20
GS1
GS55
GS48
GSIMER11
GS58
GS54
GS50
GS51
GS52
GS59
GS49
GSIMER1
GSIMER2
GS64
GSIMER6
GS70
GS28
GS16
GSIMER13
GS44
GS46
GS37
GS60
GSIMER7
GS63
GS53
GS12
GS62
GSIMER12
GS32
GS61
GS21
GS42
GS15
GS68
GSIMER5
GSIMER14
GS45
GS67
GSIMER9
GSIMER10
GSIMER8
GSIMER3
GS23
GS14
GS22
GS66
GSNAPO1
GSNAPO2
GSNAPO13
GSNAPO3
GSNAPO4
GSNAPO5
GSNAPO9
GSNAPO11
GSNAPO8
GSNAPO12
GSNAPO7
GSNAPO10
GSNAPO6

Supplementary Figure 6. chromoPainter co-ancestry matrix
The raw co-ancestry matrix from the full haplotyle dataset, output from the fineRADstructure program. See Supplementary Figure 5 caption for descriptions of localities, matching those in Figure 3 and 4.


30.2
27.4
24.7
21.9
19.1
16.4
13.6
10.8
8.06
5.3

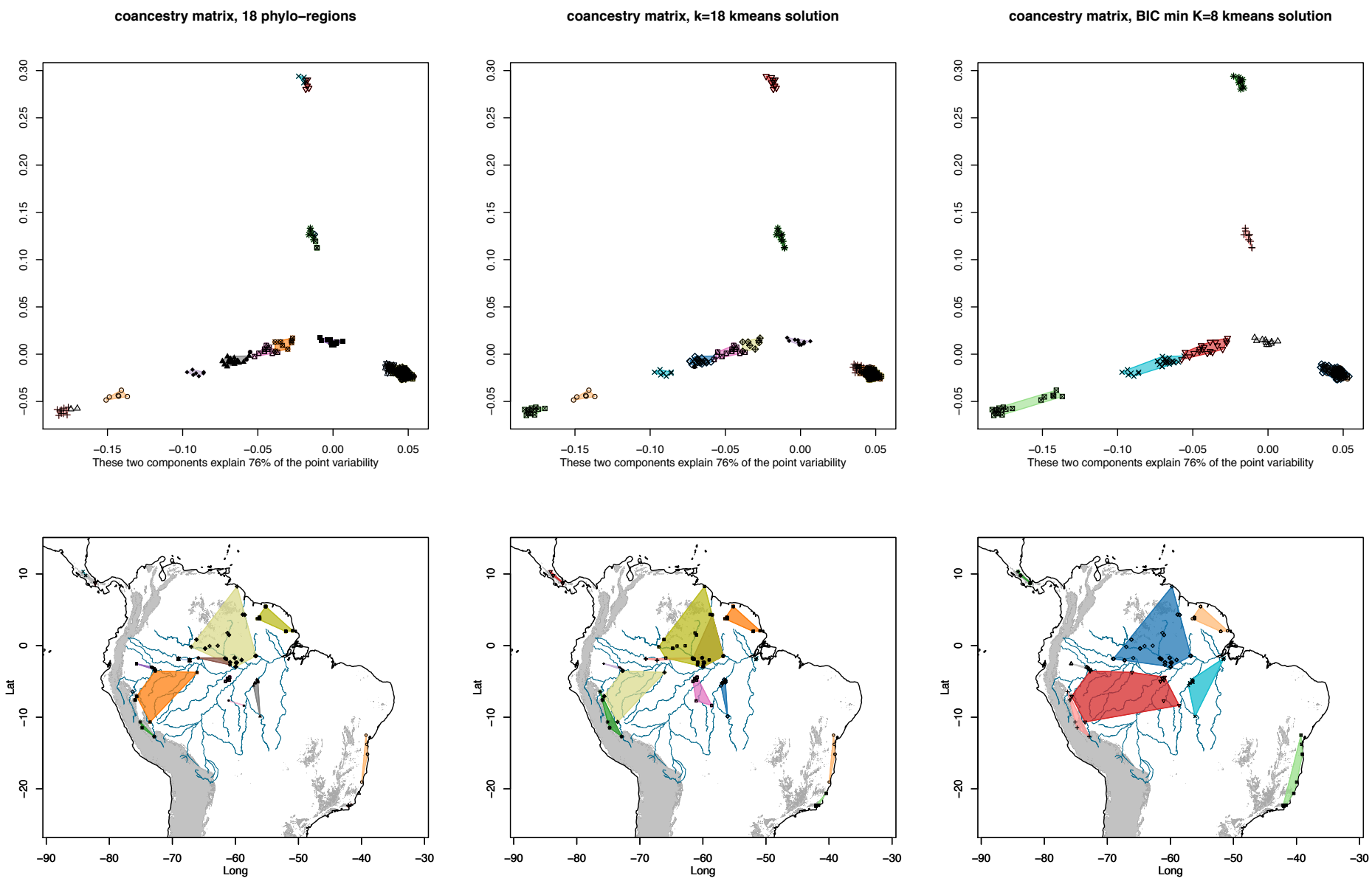
Supplementary Figure 7. chromoPainter co-ancestry matrix
The co-ancestry matrix from the full haplotyle dataset, with values averaged across 18 focal population-areas. See Supplementary Figure 5 caption for descriptions of localities, matching those in Figure 3 and 4.


18
16.6
15.3
13.9
12.6
11.2
9.85
8.49
7.13
5.77
4CACR
5CAPA
1CAMA
5CPS
3CPN
39GSSR
70GS
14GSIMERI
13GSNAPO
11PH
8INAMBARI
9INAMBARIE
2RONDONIA
24TAPAJOS
7XINGU
6AFBAHIA
2AFES
10AFRIO

Supplementary Figure 8. K-means clustering of the full dataset co-ancestry matrix
The Lawson et al. (2012) 'normalized PCA' approach provided with the fineRADstructure software captured 89% of the variance in the genetic data on the first four component axes (axis 1: 51.4%, axis 2, 24.9%, axis 3: 7.77%, axis 4: 4.94%). Thus, the co-ancestry matrix reflects substantially more information than standard PCoA/PCA of SNP data (Supplementary Figure 2). K-means phenetic clustering of the co-ancestry matrix more finely partitions the genetic data and explains a much greater proportion of the overall genetic variance than K-means clustering of the raw SNP data. Top row: normalized PCA projection of all individuals on the first two component axes, which capture ~76% of the point variability. Plotting symbols and colored convex hulls reflect cluster assignment. Hull colors are sampled randomly from a 20-color palette for each dataset (i.e., they are synonymous only across plot columns, see supplementary R script).

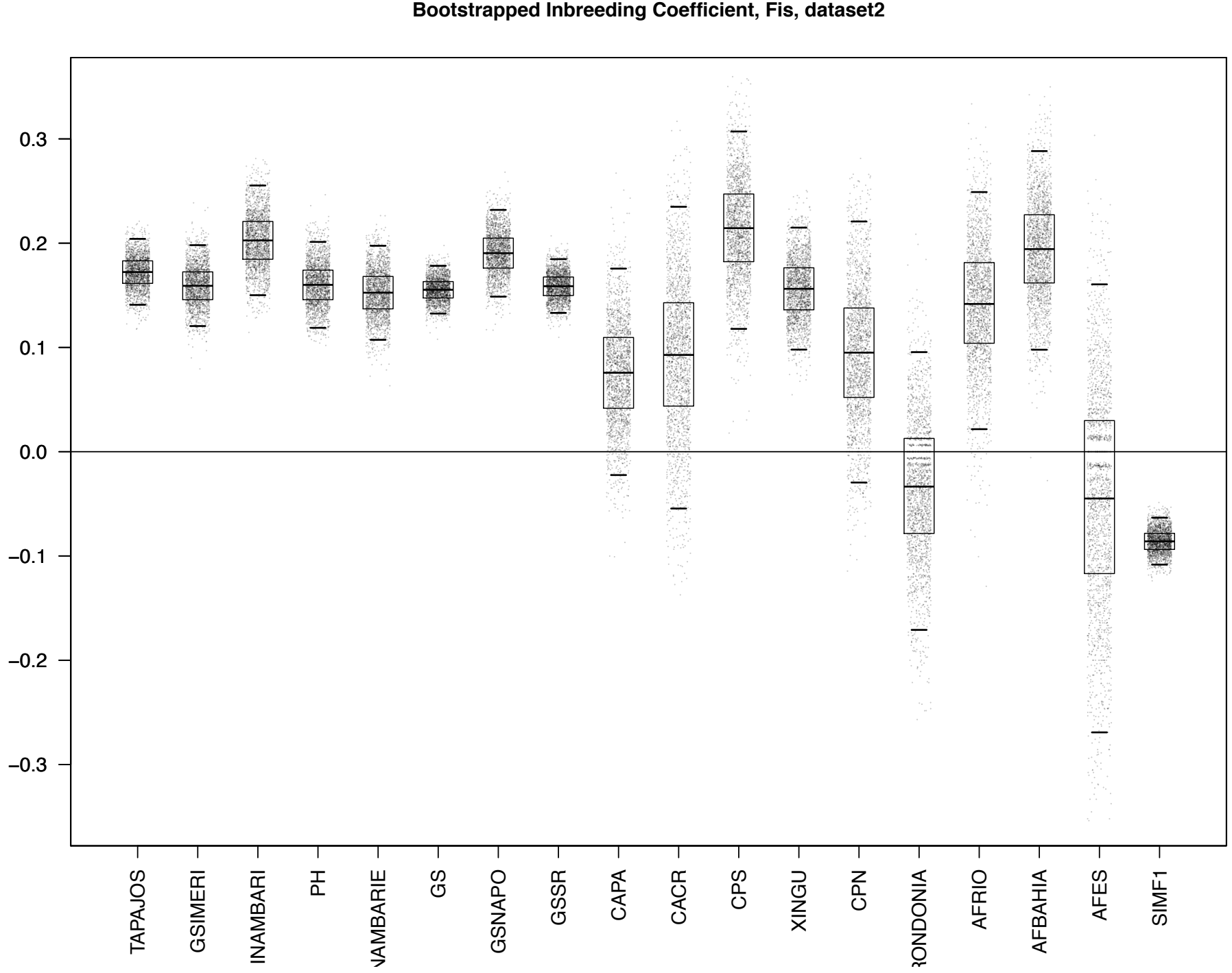
The leftmost pair of plots indicate membership to one of eighteen focal population-areas (i.e., not K-means assignments, see Supplemental Appendix text for justification) and are shown as minimum convex hulls in co-ancestry PC space (top) as well as projected onto a map (bottom, also shown in Figure 4). The center pair of plots shows the K-means clustering solution of the co-ancestry matrix when K is fixed to 18 (i.e., not based on BIC scores). Intriguingly, this produces a similar set of groups as shown in the leftmost pair.

In the rightmost two plots, we show the K-means optimum clustering solution of the co-ancestry matrix, with a BIC minimum plateau of ~8. This set of groups is generally concordant with hierarchical strata determined in earlier analyses, but also further partitioned relative to standard PC analyses on our SNP data. This clustering solution identified 1) Central America (Clade A2 in Figure 3), 2) South Andean Peru + San Martín (North Andean Peru); (Clade B + Clade A1 in Figure 3), 3) western Napo (Clade E in Figure 3), 4) Inambari + Rondônia (Clades C1, C3, and C4 in Figure 3), 5) Tapajós + Xingu (Clades C5 + C6 in Figure 3), 6) eastern Napo, Jaú, western Guiana shield, 7) Suriname + Amapá (Clade D in Figure 3) and 8) the Atlantic Forest (Clade C7), as separate groups which explain a majority of the variance in the data. Notably, this solution is entirely compatible with our phylogenetic hypothesis, except for the clustering of our single San Martín sample with geographically proximate Peruvian populations, rather than Central American populations (see discussion). This solution generally recapitulates subspecies boundaries (Figure 2).


coancestry matrix, 18 phylo–regions
coancestry matrix, k=18 kmeans solution
coancestry matrix, BIC min K=8 kmeans solution
These two components explain 76% of the point variability
Lat
Long

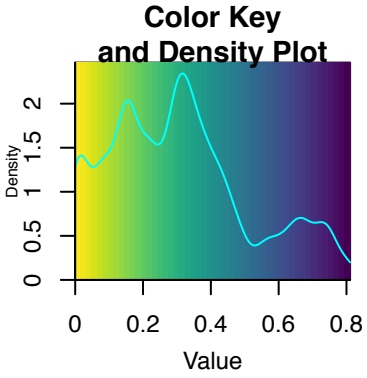
Supplementary Figure 9. Estimates of Inbreeding coefficients F_{IS}
Most populations were detected to be significantly inbred ($F_{IS} > 1$), with lower 95% confidence intervals > 0. Panamanian, Costa Rican, South Andean (North clade), Rondônia, and Espírito Santo clades had 95% confidence intervals which overlapped zero, and thus cannot be confidently inferred to have positive or negative F_{IS}. The simulated F1 (SIMF1) population however, did have significantly negative F_{IS}, as predicted. This pattern implies that the introgressed western Napo population, which was detected to have a significantly positive F_{IS}, is not likely to include recently introgressed individuals. The confidence intervals for eastern Napo, Jaú, Inambari and western Napo popularions are generally overlapping, with similar means (mean of mean estimates ~0.17, SD of mean estimates ~0.02). Locality codes below:

CAMA: North Andean – San Martín (North Andean Peru)
CACR: Central America - Costa Rica
CAPA: Central America - Panama
CPS: South Andean Peru (South)
CPN: South Andean Peru (North)
INAMBARI: Western Inambari endemic
INAMBARIE: Eastern Inambari endemic
RONDONIA: Rondônia endemic
TAPAJOS: Tapajós endemic
XINGU: Xingu endemic
AFBAHIA: Atlantic Forest – Bahia
AFES: Atlantic Forest – Espírito Santo
AFRIO: Atlantic Forest – Rio
PH (putative hybrid): Western Napo introgressed lineage
GSIMERI: unresolved Jaú
GSNAPO: weakly resolved eastern Napo
GS: weakly resolved Guiana Shield
GSSR: weakly resolved Suriname + Amapá

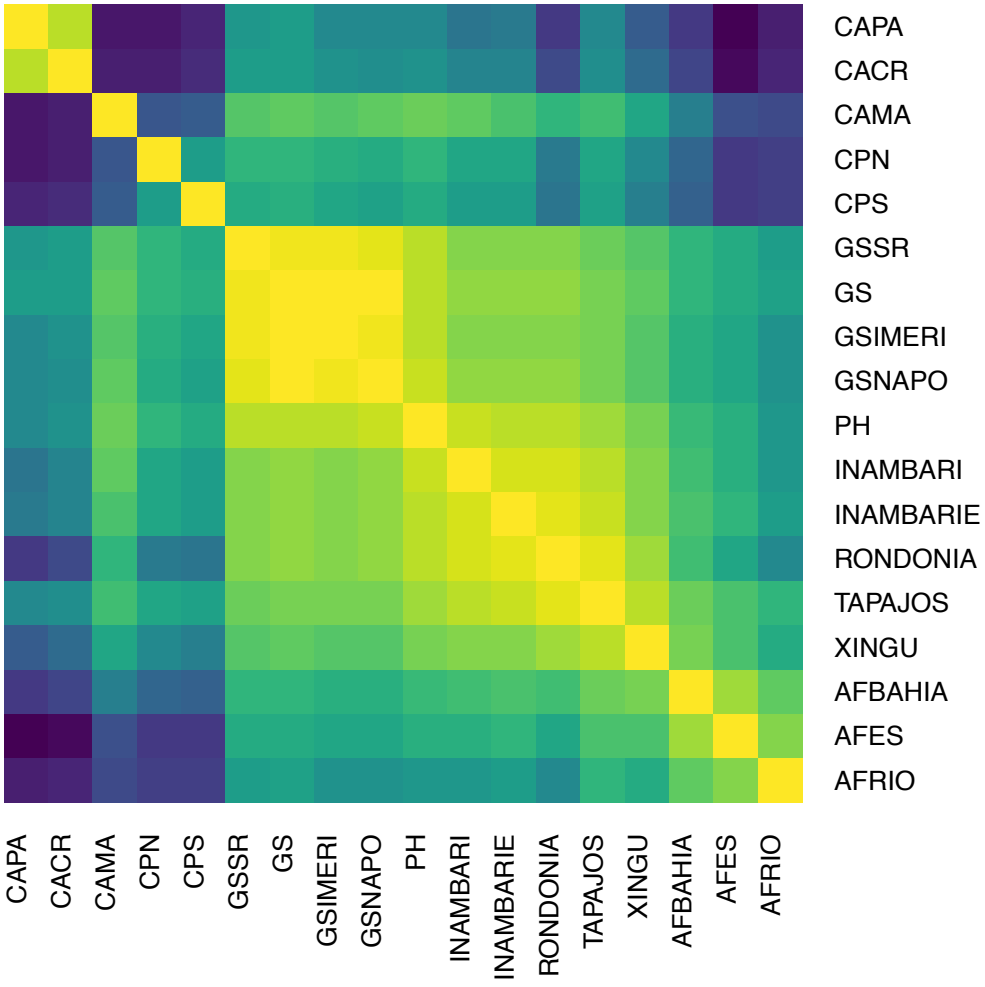

Bootstrapped Inbreeding Coefficient, Fis, dataset2
0.3
0.2
0.1
0.0
−0.1
−0.2
−0.3
TAPAJOS
GSIMERI
INAMBARI
PH
NAMBARIE
GS
GSNAPO
GSSR
CAPA
CACR
CPS
XINGU
CPN
RONDONIA
AFRIO
AFBAHIA
AFES
SIMF1

Supplementary Figure 10. Population pairwise F_{st}

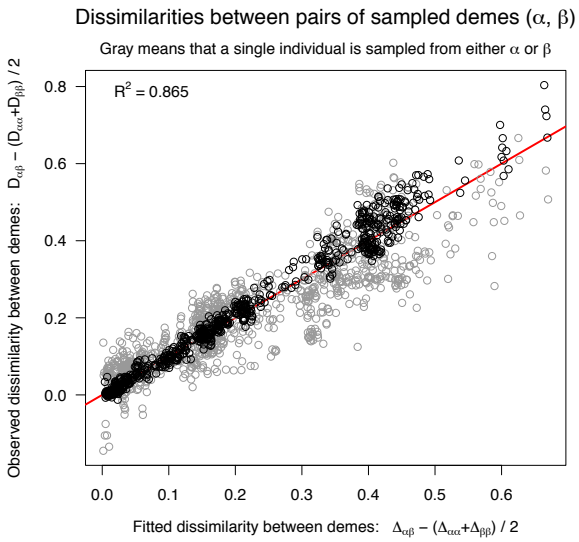
We estimated pairwise Weir and Cockerham's (Weir and Cockerham 1984) F_{st} among all 18 focal areas, and evaluated significance using 1000 bootstrapped datasets to estimate 95% confidence intervals using the 'assigner' R package (Gosselin et al. 2016). Here, we show these results plotted as a pairwise distance heatmap. Average pairwise F_{st} ranged from essentially undifferentiated (F_{st}: 0.0045, comparing GSNAPO, and GSIMERI (comparing eastern Napo to Jaú)), to almost entirely differentiated (F_{st}: 0.81, comparing AFES to CAPA (comparing Espírito Santo to Panama)). Overall population F_{st} was very high ~0.196 [0.188-0.204], indicating substantial population level differentiation among focal areas for *Pseudopipra*. Locality codes are the same as those in Supplementary Figure 9.

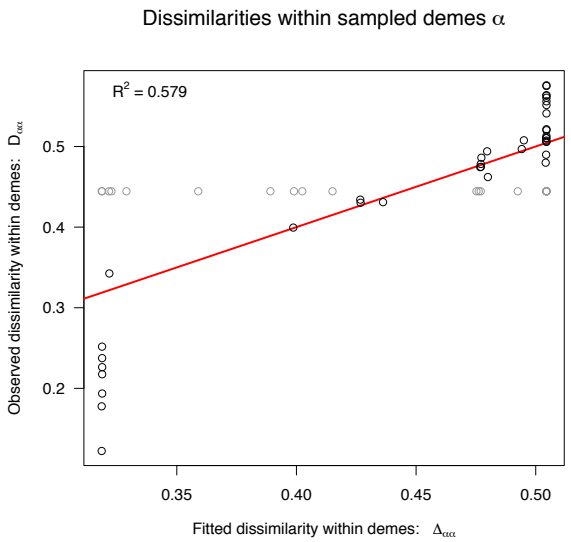

Color Key
and Density Plot
Density
0
0.5
1
1.5
2
0
0.2
0.4
0.6
0.8
Value

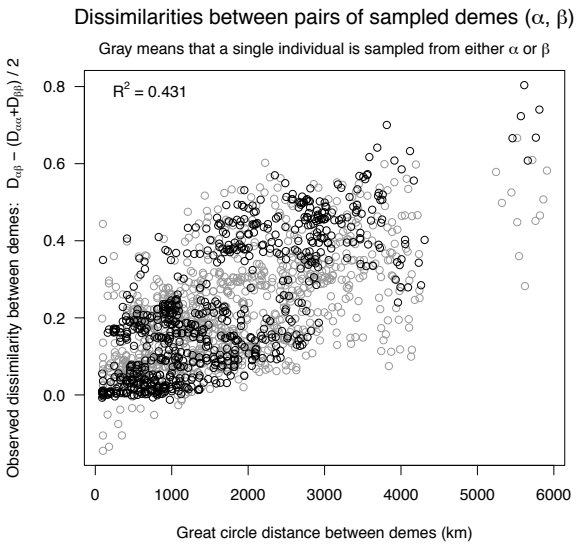
Fst heatmap


CAPA
CACR
CAMA
CPN
CPS
GSSR
GS
GSIMERI
GSNAPO
PH
INAMBARI
INAMBARIE
RONDONIA
TAPAJOS
XINGU
AFBAHIA
AFES
AFRIO

Supplementary Figure 11. EEMS model fit
Regressing the observed dissimilarity between demes against the fitted dissimilarity between demes provides an indication of model fit (Petkova et al. 2015). For the present dataset, model fit (leftmost plot) was very high ($R^2 = 0.865$). Within demes (central plot, within demes that represent more than a single individual), model fit was somewhat less, but still high ($R^2 = 0.579$). Lastly, comparing observed dissimilarity between demes against great circle distance between demes suggested a strong signal of isolation by distance operating at the scale of the entire dataset ($R^2 = 0.431$). Overall, the EEMS model does a very good job of describing spatially structured variation in this dataset.

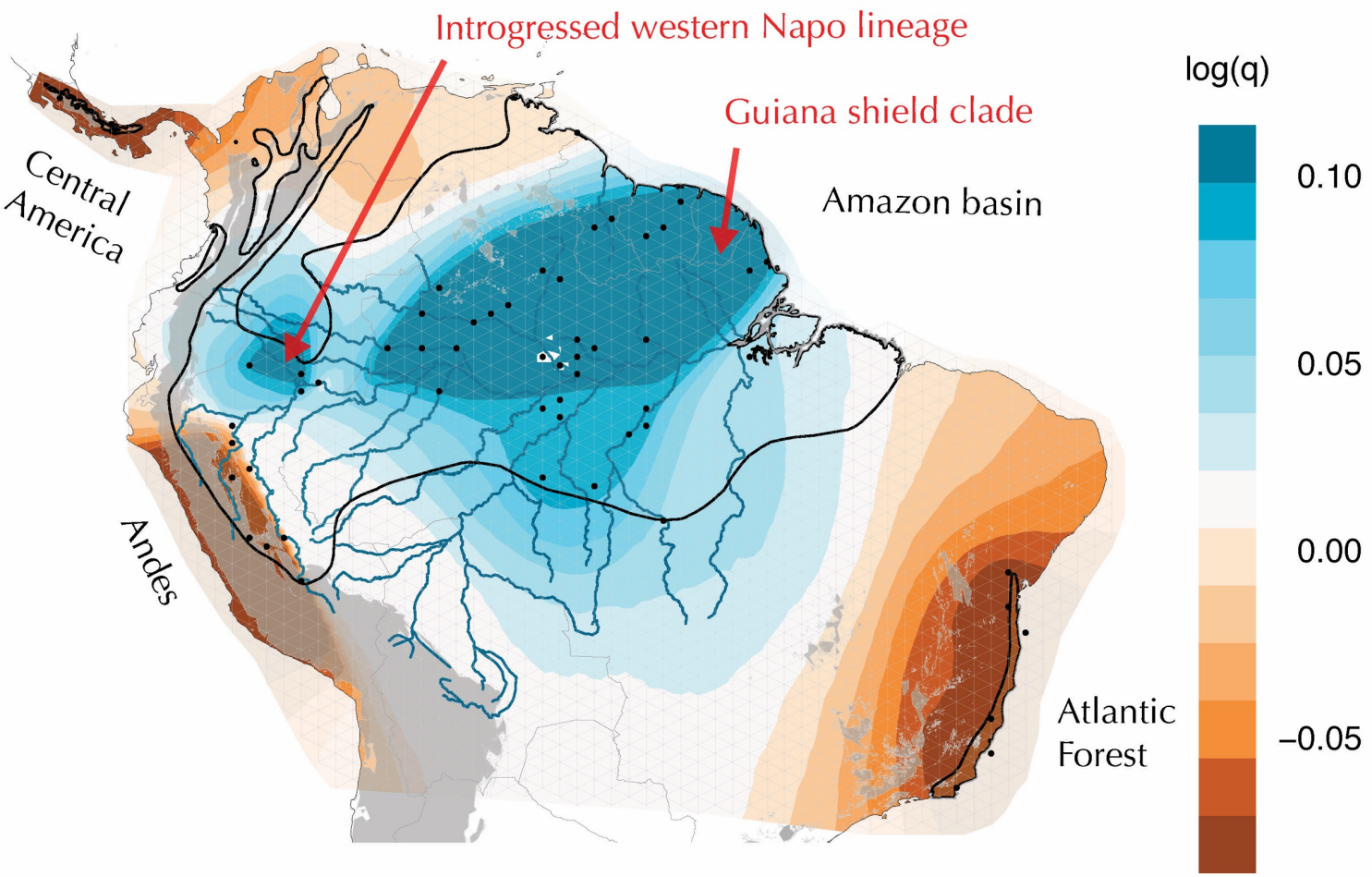

Dissimilarities between pairs of sampled demes (α, β)
Gray means that a single individual is sampled from either α or β
R² = 0.865
Observed dissimilarity between demes: Dαβ − (Dαα+Dββ) / 2
Fitted dissimilarity between demes: Δαβ − (Δαα+Δββ) / 2
0.0
0.2
0.4
0.6
0.8
0.0
0.1
0.2
0.3
0.4
0.5
0.6


Dissimilarities within sampled demes α
R² = 0.579
Observed dissimilarity within demes: Dαα
Fitted dissimilarity within demes: Δαα
0.2
0.3
0.4
0.5
0.35
0.40
0.45
0.50

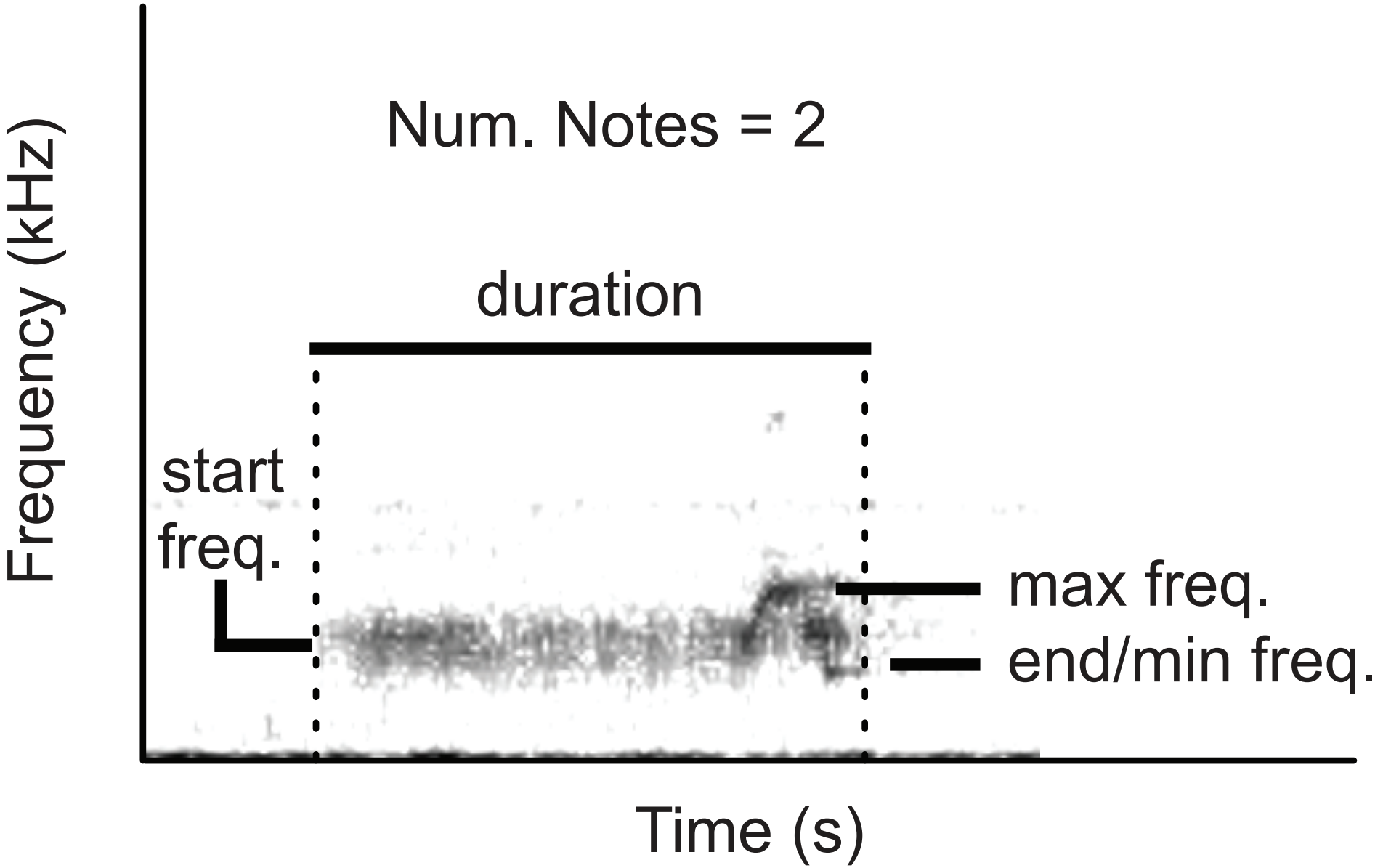

Dissimilarities between pairs of sampled demes (α, β)
Gray means that a single individual is sampled from either α or β
R² = 0.431
Observed dissimilarity between demes: Dαβ − (Dαα+Dββ) / 2
Great circle distance between demes (km)
0.0
0.2
0.4
0.6
0.8
0
1000
2000
3000
4000
5000
6000

Supplementary Figure 12. EEMS estimated genetic diversity
Two broad clusters of relatively high genetic diversity (greatest heterozygosity) were detected in EEMS. The first is centered along the Amazon river and was relatively uniform within the northern Amazonian basin, reflecting relatively high diversity in the Guiana shield. The second relatively high diversity group reflected the introgressed western Napo population. Areas of relatively low genetic diversity included the Atlantic Forest, the Peruvian Andes, and Central American lineages. These results are generally consistent with our estimates of allelic richness (Supplementary Appendix for details).

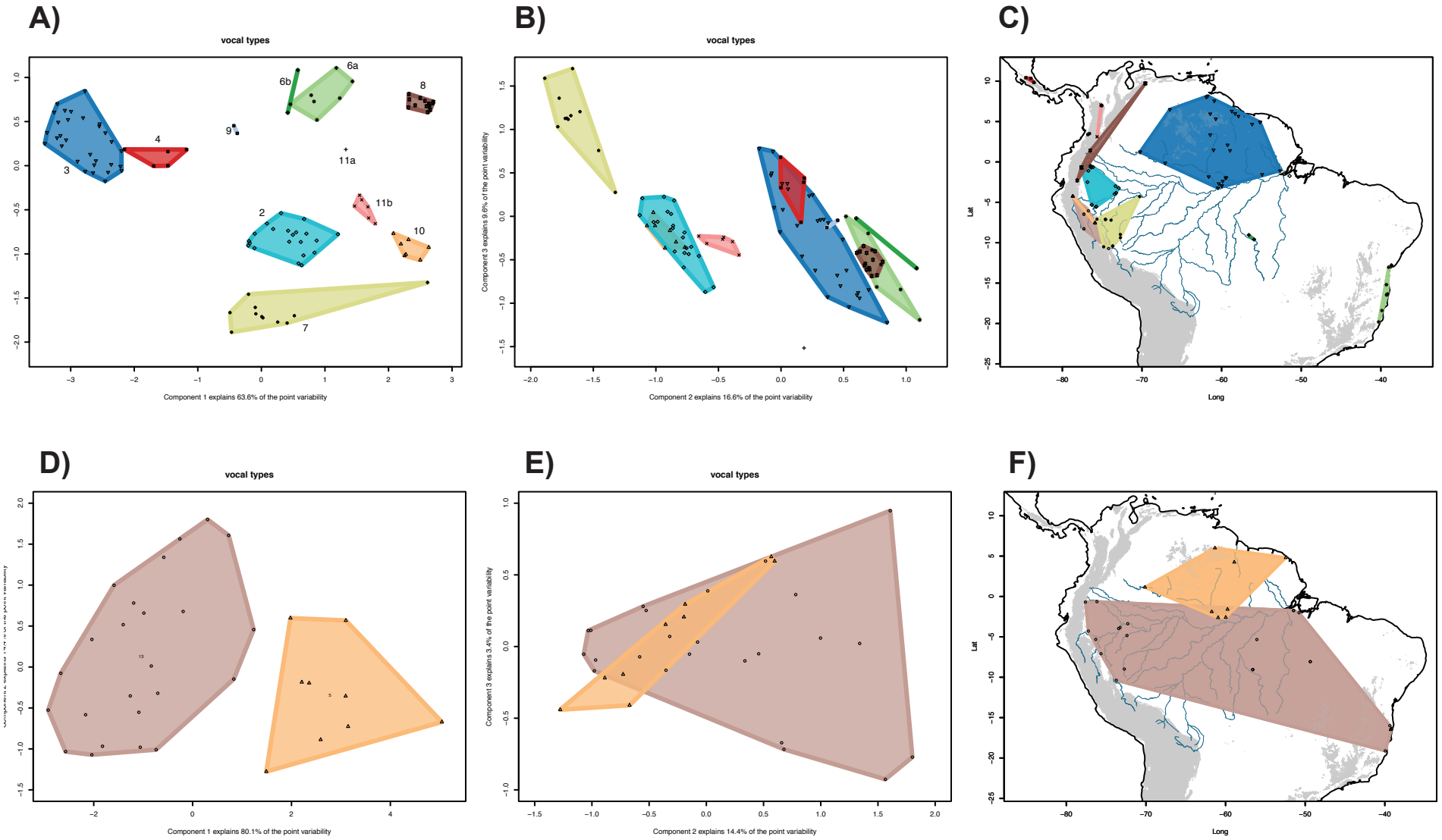
Posterior mean diversity rates q (on the log10 scale)



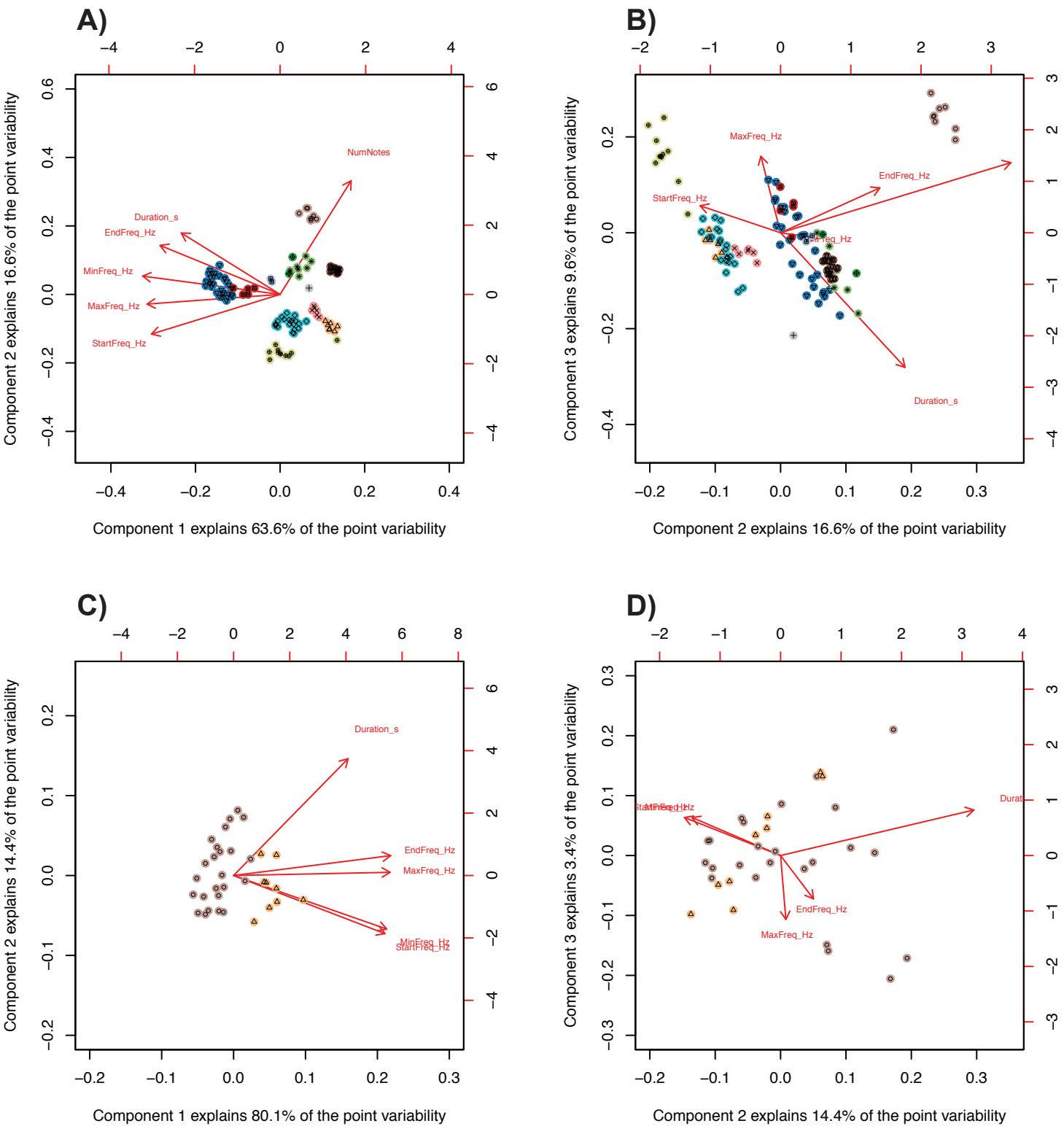
Supplementary Figure 13. Quantification of song variation
All *Pseudopipra* songs start with a single broad frequency, buzzy note. In three of the song types (Types 1, 6, and 8), the initial buzzy note is also followed by one or two shorter tonal notes. We measured: 1) starting frequency, 2) ending frequency, 3) minimum frequency, 4) maximum frequency, 5) number of notes, and 6) duration of the entire song, the buzzy note, and the tonal notes when present. To obtain a conservative estimate of the number of individuals sampled, we took measurements of one song from each recording.


Num. Notes = 2
duration
start
freq.
max freq.
end/min freq.
Frequency (kHz)
Time (s)

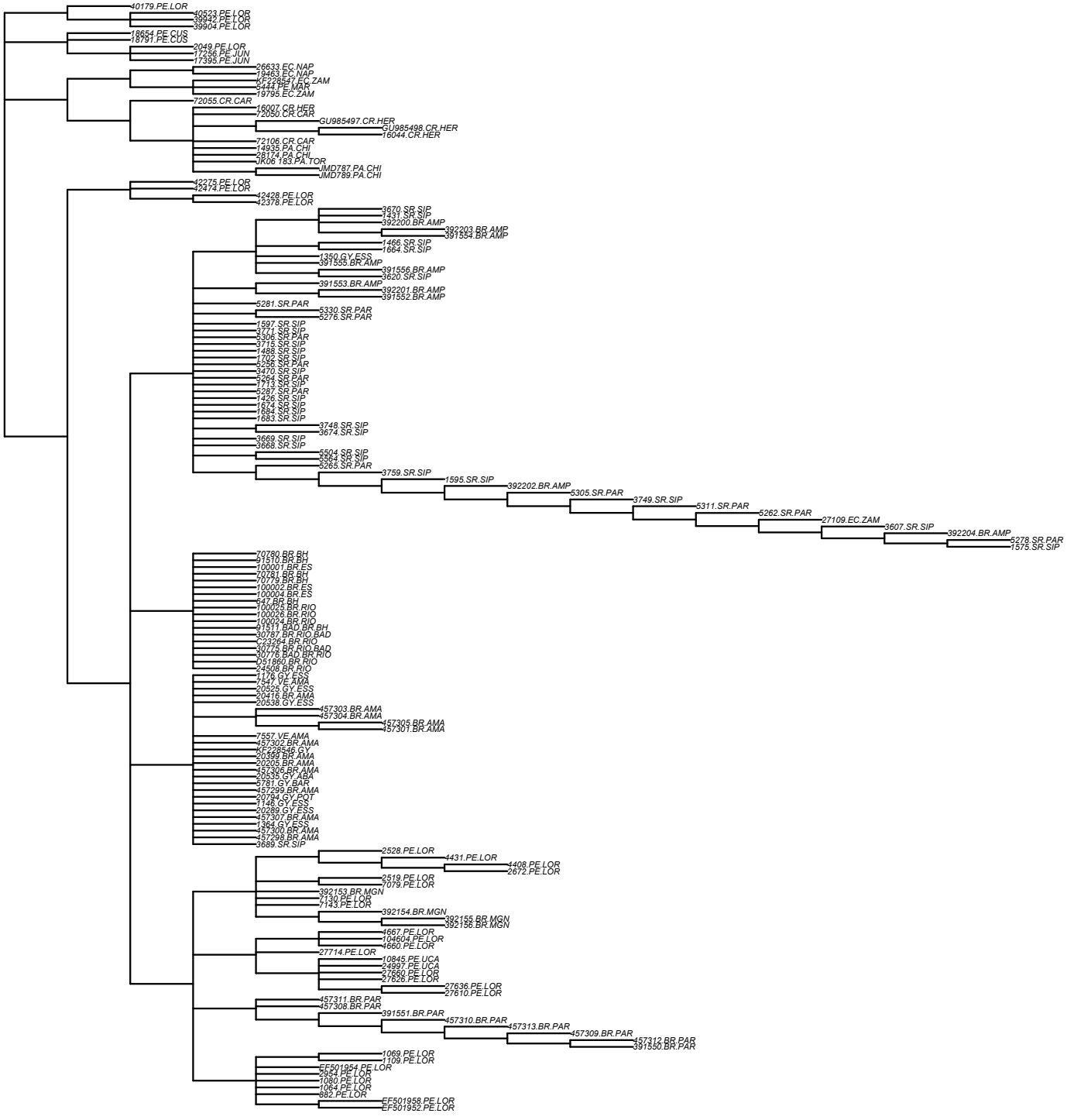
Supplementary Figure 14. Expanded summary of lekking vocalization phenotypes
PCA and logistic regression on lekking vocalizations (n=114, Supplementary Table 3) found significant differences among types ($p < 0.001$). The first three axes of a PCA explained ~90% of the variation in lekking vocalization characters, with PC1 (~64%) primarily explaining variation in note number and frequency. Panels A and D show vocalization records plotted into the first and second components of a principle components analysis. Panels B and E show vocalization data points plotted into the second and third principle components. Lastly, Panels C and F are reproduced from Figure 8 and 9 in the main text. All vocal types can be quantitatively discriminated.


A)
vocal types
6a
6b
8
9
4
3
11a
11b
2
10
7
Component 1 explains 63.6% of the point variability
B)
vocal types
Component 3 explains 9.6% of the point variability
Component 2 explains 16.6% of the point variability
C)
Lat
Long
D)
vocal types
Component 1 explains 80.1% of the point variability
E)
vocal types
Component 3 explains 3.4% of the point variability
Component 2 explains 14.4% of the point variability
F)
Lat
Long

Supplementary Figure 15 – Vocalization Loading Plots
PCA plots from Figure 14, shown here with loading vectors projected in to principle component space. The direction and length of the vectors indicate the direction and strength of the variation in a particular direction. Panels A and B show data for lek vocalizations, while C and D show data for call vocalizations.


A)
NumNotes
Duration_s
EndFreq_Hz
MinFreq_Hz
MaxFreq_Hz
StartFreq_Hz
Component 2 explains 16.6% of the point variability
Component 1 explains 63.6% of the point variability
B)
MaxFreq_Hz
EndFreq_Hz
StartFreq_Hz
Duration_s
Component 3 explains 9.6% of the point variability
Component 2 explains 16.6% of the point variability
C)
Duration_s
EndFreq_Hz
MaxFreq_Hz
Component 2 explains 14.4% of the point variability
Component 1 explains 80.1% of the point variability
D)
EndFreq_Hz
MaxFreq_Hz
Component 3 explains 3.4% of the point variability
Component 2 explains 14.4% of the point variability

Supplementary Figure 16 – Mitochondrial ND2 gene tree.
Nodes with ultrafast bootstrap scores of lower than 95 are collapsed. The inferred topology is congruent with the topology presented in the main text derived from ddRAD sequencing data, with one exception: Western Napo haplotypes cluster with southern Amazonian lowland haplotypes (see results and discussion). Otherwise, there are no strongly supported conflicts (see discussion above) with our signal from nuclear genomic DNA.


40179.PE.LOR
40523.PE.LOR
39542.PE.LOR
39904.PE.LOR
18654.PE.CUS
18791.PE.CUS
2049.PE.LOR
17256.PE.JUN
17395.PE.JUN
26633.EC.NAP
19463.EC.NAP
KF228547.EC.ZAM
5441.PE.MAR
19795.EC.ZAM
72055.CR.CAR
16007.CR.HER
72050.CR.CAR
GU985497.CR.HER
GU985498.CR.HER
16044.CR.HER
72106.CR.CAR
14935.PA.CHI
28174.PA.CHI
JK06 183.PA.TOR
JMD787.PA.CHI
JMD789.PA.CHI
42275.PE.LOR
42474.PE.LOR
42428.PE.LOR
42378.PE.LOR
3670.SR.SIP
1431.SR.SIP
392200.BR.AMP
392203.BR.AMP
391554.BR.AMP
1466.SR.SIP
1664.SR.SIP
1350.GY.ESS
391555.BR.AMP
391556.BR.AMP
3620.SR.SIP
391553.BR.AMP
392201.BR.AMP
391552.BR.AMP
5281.SR.PAR
5330.SR.PAR
5276.SR.PAR
1597.SR.SIP
1771.SR.SIP
5306.SR.PAR
3715.SR.SIP
1488.SR.SIP
1702.SR.SIP
5296.SR.PAR
3470.SR.SIP
5264.SR.PAR
1713.SR.SIP
5287.SR.PAR
1426.SR.SIP
1674.SR.SIP
1864.SR.SIP
1683.SR.SIP
3748.SR.SIP
3674.SR.SIP
3669.SR.SIP
3668.SR.SIP
5504.SR.SIP
5564.SR.SIP
5265.SR.PAR
3759.SR.SIP
1595.SR.SIP
392202.BR.AMP
5305.SR.PAR
3749.SR.SIP
5311.SR.PAR
5262.SR.PAR
27109.EC.ZAM
3607.SR.SIP
392204.BR.AMP
5278.SR.PAR
1575.SR.SIP
70780.BR.BH
71510.BR.BH
100001.BR.ES
70781.BR.BH
70179.BR.BH
100002.BR.ES
100004.BR.ES
647.BR.BH
100023.BR.RIO
100026.BR.RIO
100024.BR.RIO
41511.BAD.BR.BH
30787.BR.RIO.BAD
C23264.BR.RIO
30775.BR.RIO.BAD
30776.BAD.BR.RIO
051860.BR.RIO
24506.BR.RIO
1176.GY.ESS
7547.VE.AMA
20525.GY.ESS
20416.BR.AMA
20538.GY.ESS
457303.BR.AMA
457304.BR.AMA
457305.BR.AMA
457301.BR.AMA
7557.VE.AMA
457302.BR.AMA
KF228546.GY
20399.BR.AMA
20205.BR.AMA
457306.BR.AMA
20535.GY.ABA
5761.GY.BAR
457299.BR.AMA
20794.GY.POT
1146.GY.ESS
20289.GY.ESS
457307.BR.AMA
1364.GY.ESS
457300.BR.AMA
457298.BR.AMA
3689.SR.SIP
2528.PE.LOR
4431.PE.LOR
4408.PE.LOR
2672.PE.LOR
2519.PE.LOR
7079.PE.LOR
392153.BR.MGN
7130.PE.LOR
7143.PE.LOR
392154.BR.MGN
392155.BR.MGN
392156.BR.MGN
4667.PE.LOR
104604.PE.LOR
4660.PE.LOR
27714.PE.LOR
10845.PE.UCA
24997.PE.UCA
27660.PE.LOR
27626.PE.LOR
27636.PE.LOR
27610.PE.LOR
457311.BR.PAR
457308.BR.PAR
391551.BR.PAR
457310.BR.PAR
457313.BR.PAR
457309.BR.PAR
457312.BR.PAR
391550.BR.PAR
1069.PE.LOR
1109.PE.LOR
EF501954.PE.LOR
2624.PE.LOR
1080.PE.LOR
1064.PE.LOR
882.PE.LOR
EF501958.PE.LOR
EF501952.PE.LOR

Literature Cited

Gosselin T, Anderson EC, Ferchaud A-L. 2016. thierrygosselin/assigner: v.0.4.0 (Version 0.4.0). Zenodo. http://doi.org/10.5281/zenodo.197418.

Lawson DJ, Hellenthal G, Myers S, Falush D. 2012. Inference of Population Structure using Dense Haplotype Data. PLOS Genetics, 8:e1002453.

Petkova D, Novembre J, Stephens M. 2015. Visualizing spatial population structure with estimated effective migration surfaces. Nature Genetics, 48:94.

Weir BS, Cockerham CC. 1984. ESTIMATING F-STATISTICS FOR THE ANALYSIS OF POPULATION STRUCTURE. Evolution, 38:1358-1370.