Cultural Innovations Influence Patterns of Genetic Diversity in Northwestern Amazonia

Leonardo Arias, Roland Schröder, Alexander Hübner, Guillermo Barreto, Mark Stoneking, Brigitte Pakendorf

Cultural Innovations Influence Patterns of Genetic Diversity in Northwestern Amazonia

Leonardo Arias,*[,1,2] Roland Schröder,[1] Alexander Hübner,[1] Guillermo Barreto,[2] Mark Stoneking,*[,†,1] and Brigitte Pakendorf*[,†,3]

[1]Department of Evolutionary Genetics, Max Planck Institute for Evolutionary Anthropology, Leipzig, Germany

[2]Laboratorio de Genética Molecular Humana, Departamento de Biología, Universidad del Valle, Cali, Colombia

[3]Dynamique du Langage, UMR5596, CNRS & Université de Lyon, Lyon, France

[†]These authors contributed equally to this work.

***Corresponding authors:** E-mails: leoarias2@gmail.com; stonekg@eva.mpg.de; brigitte.pakendorf@cnrs.fr.

Associate editor: Connie Mulligan

All reads that aligned to the region of the NRY that was targeted by the capture-enrichment array were deposited in the European Nucleotide Archive (ENA) (accession no. PRJEB27777).

Abstract

Human populations often exhibit contrasting patterns of genetic diversity in the mtDNA and the nonrecombining portion of the Y-chromosome (NRY), which reflect sex-specific cultural behaviors and population histories. Here, we sequenced 2.3 Mb of the NRY from 284 individuals representing more than 30 Native American groups from Northwestern Amazonia (NWA) and compared these data to previously generated mtDNA genomes from the same groups, to investigate the impact of cultural practices on genetic diversity and gain new insights about NWA population history. Relevant cultural practices in NWA include postmarital residential rules and linguistic exogamy, a marital practice in which men are required to marry women speaking a different language. We identified 2,969 SNPs in the NRY sequences, only 925 of which were previously described. The NRY and mtDNA data showed different sex-specific demographic histories: female effective population size has been larger than that of males through time, which might reflect larger variance in male reproductive success. Both markers show an increase in lineage diversification beginning ~5,000 years ago, which may reflect the intensification of agriculture, technological innovations, and the expansion of regional trade networks documented in the archaeological evidence. Furthermore, we find similar excesses of NRY versus mtDNA between-population divergence at both the local and continental scale, suggesting long-term stability of female versus male migration. We also find evidence of the impact of sociocultural practices on diversity patterns. Finally, our study highlights the importance of analyzing high-resolution mtDNA and NRY sequences to reconstruct demographic history, since this can differ considerably between sexes.

Key words: **Y-chromosome, mtDNA, linguistic exogamy, population expansion, population history.**

Introduction

Uniparentally inherited mitochondrial DNA (mtDNA) and the nonrecombining portion of the Y-chromosome (NRY) have been used extensively to study human population history. Their sex-specific mode of inheritance is useful for contrasting the maternal versus paternal history of populations. Previous studies have found that human populations generally exhibit larger genetic differences for the NRY than for mtDNA (Seielstad et al. 1998; Kayser et al. 2003; Lippold et al. 2014), and several explanations revolving around cultural practices have been proposed to account for these differences. First, women migrate more often than men: around 70% of human societies are characterized by patrilocality (Burton et al. 1996), meaning that men generally remain in their birthplace, leading to an increase in the genetic differences in the NRY among populations. In contrast, the movement of women among populations results in a reduction of the genetic differences in the mtDNA (Seielstad et al. 1998; Oota et al. 2001; Jobling and Tyler-Smith 2003; Hamilton et al. 2005; Gunnarsdottir et al. 2011; Heyer et al. 2012; Verdu et al. 2013; Marchi et al. 2017). Second, some studies have claimed that these differences reflect a disparity between male and female effective population size (Ne), with the female Ne being larger than that of males. This might be a consequence of differential male reproductive success (i.e., many fewer males than females having offspring) (Tang et al. 2002; Wilder et al. 2004; Heyer et al. 2012). However, as shown by simulations, the excess of between-population differentiation for the NRY in comparison with the mtDNA decreases when comparing more geographically distant populations (Wilkins and Marlowe 2006). This might be due to the fact that the patterns of genetic diversity at a local scale reflect demographic and cultural practices over a relatively small number of generations, whereas at a larger geographic scale the

Article

genetic diversity might reflect old migration events and/or old common ancestry patterns. It has also been suggested that whereas short-distance migration between groups reflects patrilocality and hence is female-biased, long-distance migration might be male-biased, which would also weaken the excess NRY population differentiation over larger geographic scales (Marks et al. 2012).

Northwestern Amazonia (NWA) is particularly interesting in this regard, since human societies inhabiting this region exhibit great diversity in terms of languages, subsistence strategies, marital practices, and residential patterns, and humans have successfully adapted to its remarkable array of ecosystems since at least 10,000 years ago (Piperno 2011; Aceituno et al. 2013). A growing body of archaeological/anthropological evidence has changed the traditional view of NWA prehistory from one of small and isolated preagricultural societies to one where complex societies that were interconnected by extensive networks of exchange and trade developed (Vidal 1997, 2002; Santos-Granero 2002; Hornborg 2005; Heckenberger and Neves 2009). Furthermore, contrary to the view of Amazonia as a pristine forest with little potential for agriculture (Meggers 1954), the archaeological evidence also indicates that large regions across Amazonia were modified by intense human activity (Denevan 1992; Heckenberger et al. 2003; Erickson 2008; Clement et al. 2015; Bush et al. 2016; Levis et al. 2017). For example, the production of anthropogenic soils known as Amazonian Dark Earths, which range from a few to several hundred hectares in size, has been interpreted as an indication of increasing sedentism and intensification of agriculture (Erickson 2008; Arroyo-Kalin 2010).

Horticulturalist groups in Amazonia rely primarily on manioc (*Manihot esculenta*), but subsistence strategies are varied and extensive exchanges between horticulturalists and foragers have been documented (Jackson 1983; Milton 1984). Manioc production is exclusively women's work, and it includes planting, harvesting, processing, and the preparation of different foods and drinks (Hugh-Jones 1979; Jackson 1983; Heckler 2004). Women are also important in the exchange of manioc landraces among groups, since in several ethnolinguistic groups of NWA a newly married woman receives several manioc varieties from her mother and grandmother as part of her dowry before she leaves for the community of her husband (Peña-Venegas et al. 2014).

Although rich ethnographic descriptions exist for the region that highlight the diversity of Amazonian societies in terms of languages, subsistence strategies, marital practices, and residential patterns (Steward 1949; Sorensen 1967; Hugh-Jones 1979; Koch-Grünberg 1995; Chernela 2010), there is a general lack of studies that address the impact of these cultural practices on the patterns of genetic variation, as well as studies that reconcile the patterns of genetic variation at a local scale to the patterns of variation observed for the Americas in general. To address these issues, we have investigated the genetic diversity present in a comprehensive sample covering the extant ethnolinguistic diversity of NWA, including populations from different language families that have different subsistence strategies and different marital and other cultural practices. For example, some of the groups included in our study engage in a marital practice known as linguistic exogamy (Sorensen 1967; Stenzel 2005), in which men are required to marry women speaking a different language. We have employed next-generation DNA sequencing methods to assess the levels of diversity in the mtDNA and NRY, thereby avoiding the methodological differences that have characterized many previous studies. With these data, we aim to investigate how the patterns of variation in Native American populations from NWA are affected by different cultural practices. In particular, we address the following research questions: 1) Is there any evidence that the intensification of agriculture and development of extensive networks of exchange documented in the archaeological record had an impact on the genetic diversity in NWA? 2) To what extent have sociocultural patterns such as postmarital residence rules, linguistic exogamy, or the central role of women in manioc production affected patterns of genetic variation? 3) Is there a difference in patterns of local population differentiation in NWA versus those at a continental scale?

Results

We generated 284 NRY sequences from Native American individuals from NWA, using a hybridization capture method that covers a region of ~2.3 Mb. The average coverage per sample was 37.37 ± 15.55, which was reflected in the relatively low number of missing sites (mean = 3.99) (supplementary table 1, Supplementary Material online). We identified 2,969 SNPs in the sequences analyzed, of which only 925 have been described in the International Society of Genetic Genealogy (ISOGG) database (www.isogg.org accessed on 22.02.2018) and assigned to particular lineages. Thus, a total of 2,044 previously uncharacterized SNPs are reported here that allow a higher level of discrimination among sequences.

We compared the patterns of genetic variation between the NRY and mtDNA in 17 NWA ethnolinguistic groups, using a subset of the mtDNA genomes previously reported in Arias et al. (2018). All mtDNA lineages belonged to autochthonous Native American haplogroups, namely A2, B2, C1, or D1. In contrast, the NRY lineage diversity was dominated by haplogroup Q1, the main haplogroup observed among Native Americans, which in our data set reached a frequency of 91% (table 1). Within Q1 we identified a Cocama individual who exhibited a divergent Q1 sequence that marks the first split of the tree (supplementary fig. 1, Supplementary Material online) at 27.8 kya (95% highest posterior density [HPD] = 20.6–34.3 kya). The sequence was classified as haplogroup Q1b1a1 (renamed Q2 in the latest update of ISOGG accessed on 03.08.2018), since it carries derived alleles at the defining SNPs L275 and L612. Battaglia et al. (2013) had previously identified two individuals, one from Panama and the other from northern South America, who were positive for the SNP M378/Page100 that defines a sublineage within L275 in the haplogroup Q phylogeny. Q1b1a1 occurs at high frequency in Central, West, and South Asia (Balanovsky et al. 2017), and the estimates of the divergence between Q1a and Q1b are around 30–32 kya (Poznik et al. 2016; Kivisild 2017).

Table 1. NRY Haplogroup Frequencies for the 17 NWA Groups Included in the Population-Based Analyses.

Population	*n*	Q1	C2	R1	E1	I2	J1	J2	Language Family	Residence Pattern
Yucu-Matapi	24	1.00	0.00	0.00	0.00	0.00	0.00	0.00	Arawakan	Patrilocal
Curripaco	13	1.00	0.00	0.00	0.00	0.00	0.00	0.00	Arawakan	Patrilocal
Ach-Piapoco	22	0.91	0.00	0.05	0.00	0.05	0.00	0.00	Arawakan	Patrilocal
Carijona	6	0.83	0.00	0.17	0.00	0.00	0.00	0.00	Carib	Matrilocal
Desano	14	1.00	0.00	0.00	0.00	0.00	0.00	0.00	Eastern-Tukanoan	Patrilocal
Pira-Wanano	12	0.92	0.00	0.00	0.00	0.00	0.08	0.00	Eastern-Tukanoan	Patrilocal
Siona	10	1.00	0.00	0.00	0.00	0.00	0.00	0.00	Western-Tukanoan	Patrilocal
Coreguaje	12	0.92	0.00	0.08	0.00	0.00	0.00	0.00	Western-Tukanoan	Patrilocal
Sikuani	14	0.93	0.00	0.07	0.00	0.00	0.00	0.00	Guahiban	Matrilocal
Guayabero	18	0.83	0.11	0.06	0.00	0.00	0.00	0.00	Guahiban	Matrilocal
Saliba	11	0.73	0.00	0.09	0.09	0.00	0.00	0.09	Saliba-Piaroan	Ambi/neolocal
Mur-Uitoto	17	0.71	0.06	0.18	0.00	0.06	0.00	0.00	Huitotoan	Patrilocal
Puinave	18	0.89	0.00	0.06	0.06	0.00	0.00	0.00	Maku-Puinave	Patrilocal
Nukak	11	1.00	0.00	0.00	0.00	0.00	0.00	0.00	Maku-Puinave	Patrilocal
Tikuna	10	1.00	0.00	0.00	0.00	0.00	0.00	0.00	Tikuna	Patrilocal
Cocama	11	0.82	0.00	0.09	0.00	0.00	0.00	0.09	Tupian	Patrilocal
Yagua	10	1.00	0.00	0.00	0.00	0.00	0.00	0.00	Peba-Yaguan	Patrilocal

However, based on the presence of this sublineage in just three individuals it is difficult to determine if it can be considered autochthonous in Native Americans, or if instead it reflects postcontact admixture. In addition, haplogroup C2 (previously called C3), the other confirmed autochthonous lineage in the Americas (Roewer et al. 2013), was observed at very low frequency: only three individuals (1%) exhibited this haplogroup, two from the Guayabero and one from the Mur-Uitoto group. Furthermore, 18 individuals (8%) from the groups included in the population-based analyses carried lineages that are commonly regarded as the product of recent European or African admixture. Of these, haplogroup R1 was the most frequent, accounting for 5% of the total sequences. These observations are in agreement with previous studies that indicate that the patterns of admixture in American populations have been sex-biased (Mesa et al. 2000; Sans 2000; Rojas et al. 2010), particularly during colonial times, involving primarily European men and indigenous women.

Molecular Diversity

The patterns of NRY and mtDNA genetic diversity varied among NWA populations. The Eastern Tukanoan Desano and Pira-Wanano stand out in having much lower than average Y-chromosomal haplotype diversity in conjunction with higher than average values of mtDNA haplotype diversity (fig. 1*A*). In contrast, the Siona and the Ach-Piapoco had higher than average NRY and lower than average mtDNA haplotype diversity, whereas the Mur-Uitoto, Tikuna, and Cocama had higher than average haplotype diversity values for both markers; the hunter-gatherer groups Nukak and Sikuani had lower than average haplotype diversity values for both markers. In addition to low haplotype diversity values for the NRY, the Eastern Tukanoan groups also exhibited much lower than average values of the mean number of pairwise differences (MPD) for the NRY (6.6 in the Desano and 3.2 in the Pira-Wanano as opposed to 37.9 differences on average), whereas their mtDNA MPD values were in the average range. In contrast, the Guayabero and Mur-Uitototo showed the highest values of NRY MPD. However, when the C2 sequences were excluded, their values were within the range of the other populations (fig. 1*B* and supplementary table 2, Supplementary Material online).

The number of segregating sites varied considerably among groups for both mtDNA (range from 55 to 159) and the NRY (range from 14 to 412). In the NRY we observed five populations (Ach-Piapoco, Guayabero, Mur-Uitoto, Tikuna, and Cocama) with more than 200 segregating sites and three populations (Desano, Pira-Wanano, and Nukak) with less than 50 segregating sites. The high values in Guayabero and Mur-Uitoto are again explained by the presence of sequences belonging to haplogroup C2; when we excluded these sequences, the number of segregating sites dropped to 99 in Guayabero, although it remained high for Mur-Uitoto (166 segregating sites; fig. 1*C*). At the same time, the high numbers of segregating sites in the Ach-Piapoco, Tikuna, and Cocama might be an indication that these groups have a complex population history with heterogeneous origins.

Tajima's test of selective neutrality (fig. 1*D*) showed more negative *D* values for the NRY than for mtDNA, and these were significantly negative (P value < 0.05) in five populations, namely Desano, Mur-Uitoto, Puinave, Tikuna, and Cocama. In contrast, none of the *D* values estimated from the mtDNA were significant, not even the large positive values found for the Guayabero (2.3), Nukak (1.9), and Sikuani (1.3). Negative Tajima's *D* values are indicative of population expansion, whereas positive values suggest population contractions. As the Nukak and Sikuani are the only groups exhibiting positive Tajima's *D* values for both the mtDNA and NRY, they may have undergone a population contraction. In contrast, the lower *D* values observed in the NRY than in the mtDNA suggest that most of the NWA populations have undergone recent population expansions in the paternal line.

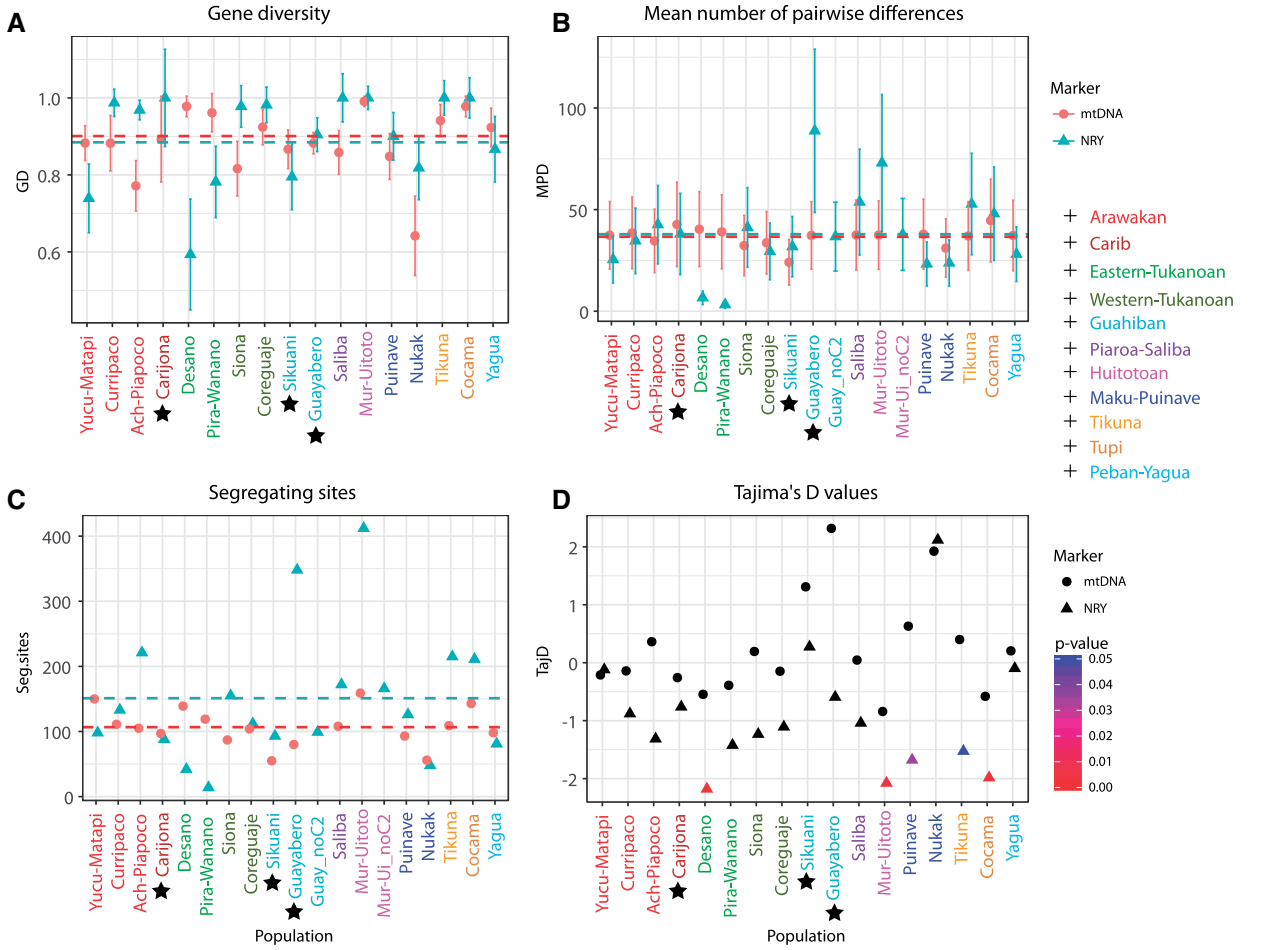

FIG. 1. Molecular diversity indices for the mtDNA and NRY in 17 groups of NWA. (*A*) Gene diversity. (*B*) Mean number of pairwise differences. (*C*) Number of segregating sites. (*D*) Tajima's *D* values. Dotted lines represent the average diversity values across all groups. Population names are color-coded by language family as indicated in the legend. Values for the NRY after removing sequences belonging to haplogroup C2 are shown next to Guayabero and Mur-Uitoto, the only groups carrying these sequences. Stars indicate the matrilocal groups.

The comparisons of the patterns of genetic variation between matrilocal and patrilocal populations are shown in figure 2. We obtained the distribution of the pairwise Φ_{ST}s, the haplotype diversity, and the MPD by sampling all possible combinations of triplets of patrilocal populations and estimating, for every triplet, the mean value of each statistic. After correcting for multiple comparisons, there were no significant differences between patrilocal and matrilocal groups in any of the values (based on One-sided tests). However, the distributions of the statistics were compatible with the expectations for patrilocality versus matrilocality (fig. 2): only 2% of the patrilocal triplets exhibited larger MPD values than the matrilocal groups for the NRY, whereas for mtDNA 90% of patrilocal triplets exhibited larger MPD values than matrilocal groups. Moreover, 74% of the patrilocal triplets exhibited smaller pairwise Φ_{ST} values for the mtDNA than matrilocal groups, whereas 95% of patrilocal groups exhibited larger Φ_{ST} values for the NRY than matrilocal groups.

Genetic Structure and Genetic Distances

The AMOVA results (table 2) showed that the percentage of between-population differences was higher for the NRY (27.24%) than for the mtDNA (12.79%), indicating that there is more genetic structure for the NRY and therefore larger genetic distances among populations, as we show below. In addition, the AMOVA results were consistent with our previous findings (Arias et al. 2018) that indicate that the distribution along rivers is a better predictor of the genetic structure in NWA than the general geographical location or the linguistic affiliation of groups (table 2). This is particularly true for the NRY, since the component of variation due to differences among groups defined by settlement along rivers was highly significant and larger than the component of variation due to differences between populations within groups.

We used the pairwise Φ_{ST} values as a measure of the genetic distances among populations. Pairwise distances were generally lower for mtDNA and fewer values were significantly different from zero (P value < 0.05) than for the NRY (supplementary fig. 2, Supplementary Material online). A visual representation of the matrices of pairwise Φ_{ST} distances are presented in the MDS plots (supplementary fig. 3, Supplementary Material online), where we observed different genetic relationships among populations for the mtDNA and the NRY. For instance, the Eastern Tukanoan groups Desano and Pira-Wanano were located in the center of the plot, together with several other groups, for the mtDNA, but were

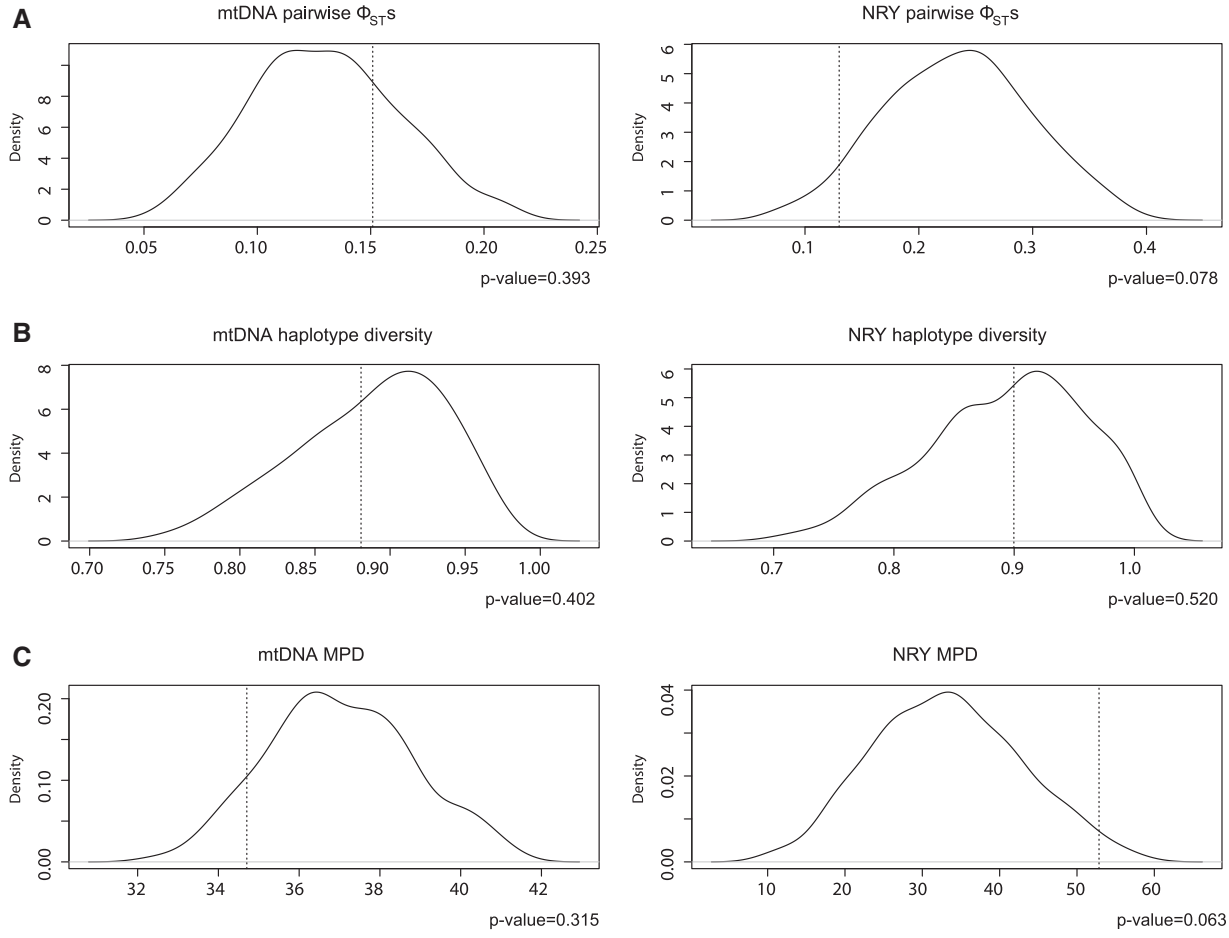

Fig. 2. Distribution of the mean values for: (*A*) Φ_{ST}s, (*B*) haplotype diversity, and (*C*) MPD for all possible triplets ($n = 286$) of patrilocal groups for the mtDNA (left) and NRY (right). The mean values for matrilocal groups are shown as dotted lines. One-side *P* values reported after Benjamin–Hochberg correction for multiple comparisons.

Table 2. AMOVA Results.

	Among Groups	Within Groups	Within Populations	# Groups
mtDNA				
Language	**−3.79**	**16.37****	**87.42****	**11**
Geography	**−0.83**	**13.53****	**87.30****	**6**
Rivers	**4.15***	**8.83****	**87.02****	**10**
Global Φ_{ST}	**12.79**			
Y-chromosome				
Language	**5.77**	**21.81****	**72.42****	**11**
Geography	**6.71***	**21.23****	**72.06****	**6**
Rivers	**14.45****	**13.35****	**72.2****	**10**
Global Φ_{ST}	**27.24**			

* *p*-value < 0.05
** *p*-value < 0.01

clearly differentiated based on the NRY (supplementary fig. 3, Supplementary Material online), and there was no obvious tendency for groups from the same language family to cluster together. Finally, a Mantel test revealed no significant correlations between genetic distances based on mtDNA and NRY sequences, nor between genetic and geographic distances (supplementary fig. 4, Supplementary Material online). Previously, adding rivers as an additional predictor variable resulted in a significant increase in the *R*-square value for the regression between mtDNA genetic distances and geographic distances (Arias et al. 2018). However, for the NRY adding rivers as an additional predictor resulted in just a slight (and nonsignificant) increase in the *R*-square value (*R* square = 0.037 *P* value= 0.30) in comparison with the simple regression between geographic and NRY genetic distances (*R* square = 0.013 *P* value= 0.47).

The shared haplotypes between pairs of populations provide information about their relationships, which can be due to common ancestry and/or gene flow between populations. Even though there was less sharing of haplotypes both between and within populations for the NRY in comparison with the mtDNA (fig. 3)—which can be explained by the differences in the amount of sequence and the mutation rate between the NRY and the mtDNA—the fact that there are shared NRY haplotypes between populations is noteworthy. According to the mutation rate used here of 7.6×10^{-10} subs/bp/year (Fu et al. 2014), for the sequenced region of ~2.3 Mb we expect to observe, on average, one substitution every 572 years, indicating that shared NRY haplotypes reflect recent contact or very recent population divergences; they thus further confirmed the contact events that we described previously based on mtDNA alone (Arias et al. 2018). For instance, the Guayaberos, Sikuani, Nukak, Puinave, Saliba,

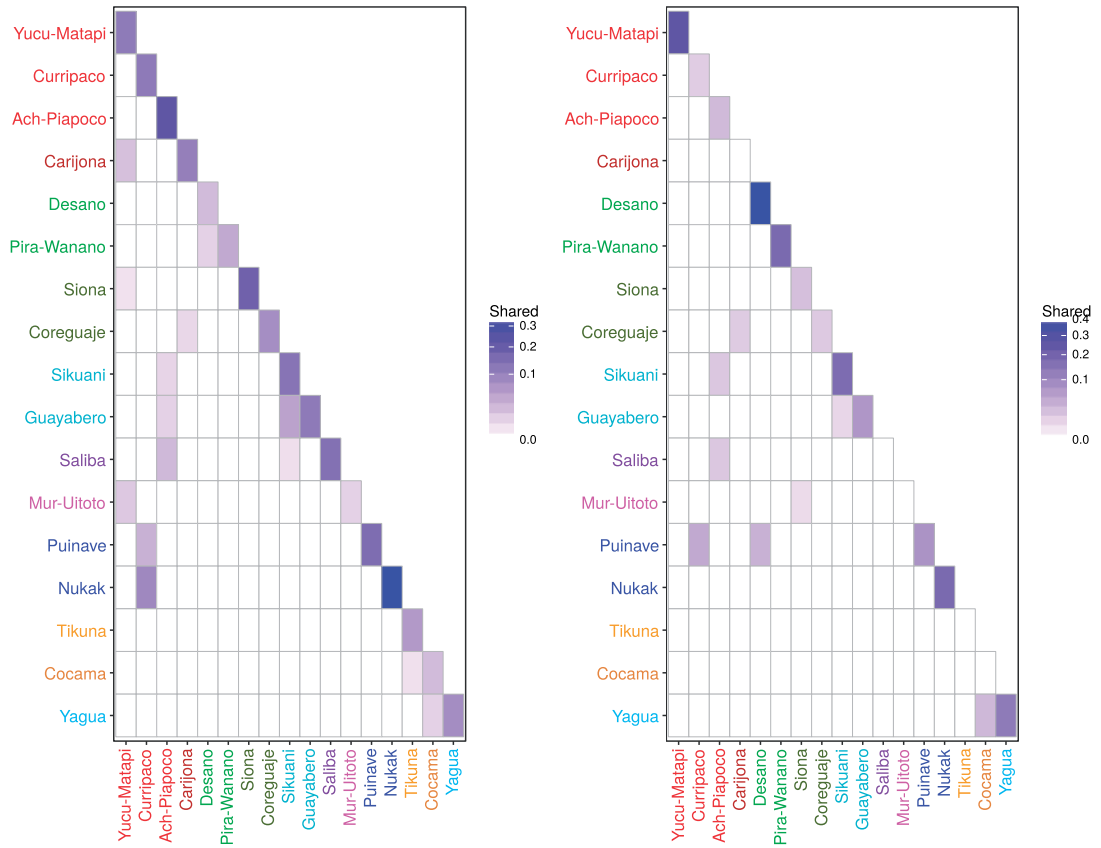

Fig. 3. Shared haplotypes between pairs of populations for the mtDNA (left) and NRY (right).

Ach-Piapoco, and Curripaco, who inhabit several tributaries of the Orinoco River, shared both identical and closely related mtDNA and NRY haplotypes (fig. 3 and supplementary fig. 5, Supplementary Material online—see circled clades), indicating that these groups have interacted for some time (Arias et al. 2018). This is additionally supported by ethnohistoric accounts of the existence of contact among the groups living along the Orinoco River and its tributaries (Rey Fajardo 1974; Morey and Morey 1980).

Phylogenetic and Demographic Reconstructions

The estimation of coalescence events through time and the phylogenies for the NRY haplogroup Q1 sequences and the mtDNA are depicted in figure 4 and supplementary figures S1 and S6, Supplementary Material online, respectively. The first split of the Q1 phylogeny separated a single individual belonging to subhaplogroup Q1b1a1 from the bulk of sequences belonging to haplogroup Q1a, which showed an initial event of diversification ~13 to 14 thousand years ago (kya). This diversification is evident in the number of coalescence events that happen around that time (fig. 4). An additional accumulation of coalescence events was observed after 5 kya, and 50% of all coalescence events happen during the last ~1 kya (fig. 4). The mtDNA also showed two regions of high accumulation of coalescence events (fig. 4), similar to what was found for the NRY. One such accumulation starts ~17.5 kya and the second one starts ~5 kya, with 50% of all coalescence events taking place in the last ~1 kya, as seen for the NRY.

The Bayesian skyline plot (BSP) demographic reconstructions for the mtDNA and the NRY are shown in figure 5. Both markers exhibited striking increases in the effective population size (Ne). However, there were differences in the estimates of Ne, the magnitude of increase, and the dates for the start of the population expansion. The mtDNA exhibits a single signal of population expansion starting ~17 kya, whereas in the NRY we observed an initial expansion that starts at ~13.5 kya and a subtle increase in Ne at ~3.5 kya, preceded by a reduction in Ne after 10 kya. Furthermore, the mtDNA exhibits larger Ne through time, being double the size of the NRY at the peak of both markers. The older signal of population expansion in the NRY was also observed in the network of haplotypes belonging to haplogroup Q1, which showed a star-like shape suggestive of a rapid diversification of lineages (supplementary fig. 5, Supplementary Material online). We estimated the date for this diversification around 13.6 ± 0.6 kya, based on the rho statistic as implemented in the software Network 4.6.1.5 (http://www.fluxus-engineering.com; last accessed August 27, 2018), using the substitution rate reported by Fu et al. (2014) (7.6×10^{-10} subs/bp/year).

We further investigated the differences between the BSPs for both markers. Although both the mtDNA and the NRY exhibit similar signals of recent coalescence events (fig. 4), the BSPs differ in that the NRY, but not the mtDNA, shows a recent increase in Ne. A possible explanation is violation of the panmixia assumption, as it has been shown that coalescence-based methods, including BSP, can then produce misleading results (Heller et al. 2013; Grant 2015). That is,

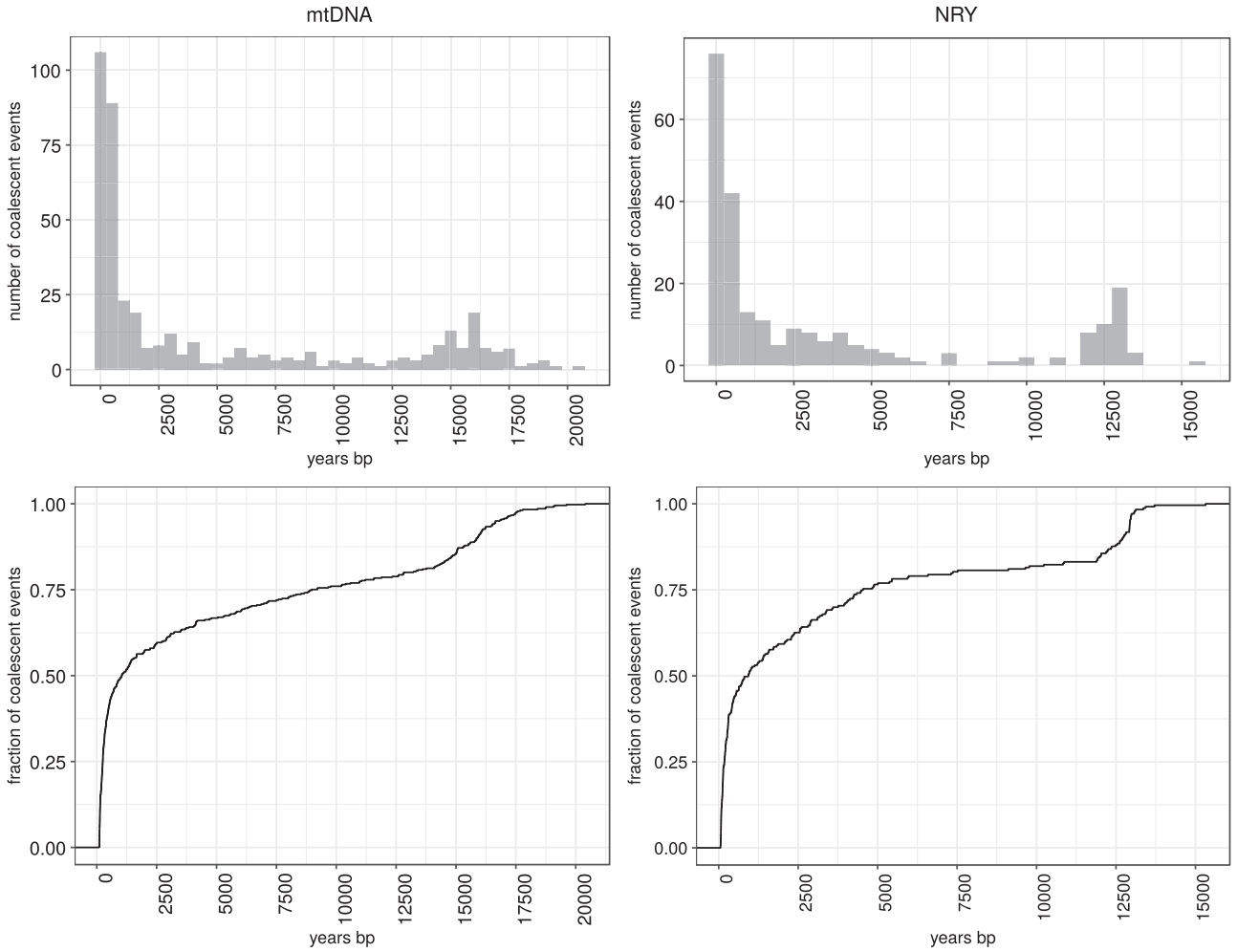

Fig. 4. Histograms and cumulative fraction of the number of coalescent events through time for the mtDNA (left) and NRY (right).

locally sampled populations that are interconnected by low levels of migration (a structured population) typically exhibit genealogies that resemble those of panmictic populations that have declined in size (Heller et al. 2013). One way to address this issue is to use different sampling strategies (see Materials and Methods section). The BSP results of the "pooled" and "scattered" strategies consistently identified the old signal of population expansion in both mtDNA and NRY, whereas some local populations showed a flatter trajectory of the Ne through time (supplementary fig. 7, Supplementary Material online). Moreover, we observed that the BSPs from the pooled and scattered sampling strategies of the NRY, but not of the mtDNA, showed an additional pulse of population size increase between 2.5 and 3.5 kya (supplementary fig. 7, Supplementary Material online). This suggests that this increase was not due to violation of the panmixia assumption, but rather indicates a true male-biased expansion.

Patterns of Genetic Diversity at the Local Versus Continental Scale in Native Americans

Our comparisons of the patterns of genetic diversity at different geographic scales are shown in figure 6. We observed differences in the levels of within-population diversity (i.e., nucleotide diversity) and between-population differentiation (i.e., Φ_{ST} values) between the mtDNA and the NRY. First, the nucleotide diversity in the mtDNA was larger than in the NRY, without significant differences between the local and continental levels. Second, the between-population differentiation for mtDNA was smaller than that for the NRY at both the local (mtDNA $\Phi_{ST} = 0.13$ vs. NRY $\Phi_{ST} = 0.27$) and continental level (mtDNA $\Phi_{ST} = 0.08$ vs. NRY $\Phi_{ST} = 0.17$). Whereas the between-population differentiation was larger at the local scale than at the continental scale for both mtDNA and the NRY, the relative amounts of NRY differentiation compared with mtDNA were the same (ratio of NRY Φ_{ST}/mtDNA $\Phi_{ST} = 2.1$ for both NWA and the continental sample).

Our estimates of the amount of between-population differences in the Americas contrast strikingly with the findings of Lippold et al. (2014), who reported much larger between-population differences for the mtDNA ($\Phi_{ST} = \sim 0.7$) than for the NRY ($\Phi_{ST} = \sim 0.2$) in a sample of 22 individuals from the CEPH Human Genome Diversity Panel. In contrast, we observed in our continental data set (composed of 77 indigenous individuals from four countries, see Materials and Methods section) a much smaller degree of differentiation

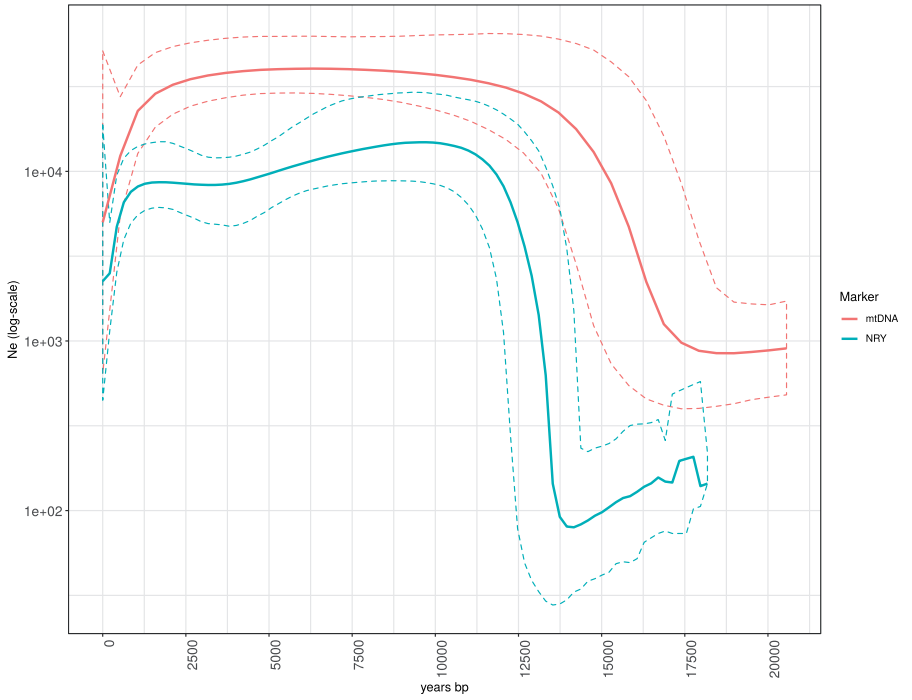

FIG. 5. BSPs for the mtDNA and NRY sequences from NWA. The dotted lines indicate the 95% HPD intervals. Ne was corrected for generation time according to (Fenner 2005), using 26 years for mtDNA and 31 years for NRY.

for the mtDNA ($\Phi_{ST} = 0.08$ averaged over five replicates) than that found by Lippold et al., whereas the NRY Φ_{ST} value (0.17) was comparable (supplementary table 3, Supplementary Material online).

Discussion

We have investigated the patterns of genetic variation in a region of ~2.3 Mb of the Y-chromosome and compared them to complete mitochondrial genomes in a large population-based sample of ethnolinguistic groups from NWA. This approach overcomes the drawbacks from previous studies regarding the quality and quantity of the data from which the levels of diversity in the mtDNA (typically sequences of the hypervariable segments of the control region) and in the Y-chromosome (typically biallelic SNPs and/or Y-STRs) have been assessed. We have identified 2,044 previously uncharacterized NRY SNPs, which provide new insights about NRY genetic diversity in Native Americans.

The Signals of an Old and a Recent Diversification in NWA

Recent genetic data from contemporary populations and ancient DNA have provided evidence that the initial peopling of the Americas by modern humans occurred between 15 and 20 kya (Fagundes et al. 2008; Raghavan et al. 2015; Llamas et al. 2016; Moreno-Mayar et al. 2018). Once in the Americas, populations quickly expanded and spread, reaching the southern extreme of South America in a few thousand years. The archaeological site of Monte Verde in Chile, dated to ~14.5 kya, is the result of this rapid movement across the Americas (Dillehay et al. 2015). Our phylogenetic and BSP reconstructions are consistent with the signals of an old population expansion, happening between 13.5 kya (NRY) and 17 kya (mtDNA), in agreement with the proposed time frame for the peopling of the Americas.

In addition, a recent event of lineage diversification is inferred from the accumulation of coalescence events for both mtDNA and the NRY (fig. 4). After ~5 kya there is evidence of a slight increase, and by ~1 kya a dramatic increase in the number of coalescence events, with around 50% of all coalescence events happening during the last 1,000 years. The expansion of widespread language families and/or archaeological cultures is often linked with the development of agriculture during the Neolithic (Renfrew 1999; Bellwood et al. 2002; Diamond and Bellwood 2003; Gignoux et al. 2011). Agriculture developed at different times in different parts of the world; for South America, the initial domestication of plants is dated to ~8 to 9 kya (Piperno 2011), considerably earlier than the lineage diversification we observe. However, it was only after ~5 kya that plant domesticates became a significant part of human diet (Piperno 2011; Arroyo-Kalin 2012; Clement et al. 2015; Levis et al. 2017). Paleoenvironmental information indicates that during the transition from the Middle to Late Holocene (~4.2 kya) there was a major climatic change in tropical South America from drier conditions to increased rainfall (Cross et al. 2000; Marchant and Hooghiemstra 2004; Wanner et al. 2008). The archaeological record indicates that this climate change induced vegetation changes and could have enabled a change in subsistence patterns among Amazonian societies (Iriarte et al. 2017), as seen by the fact that large areas of tropical rainforest were modified by human activity. This is indicated by changes in phytolith assemblages and the presence of

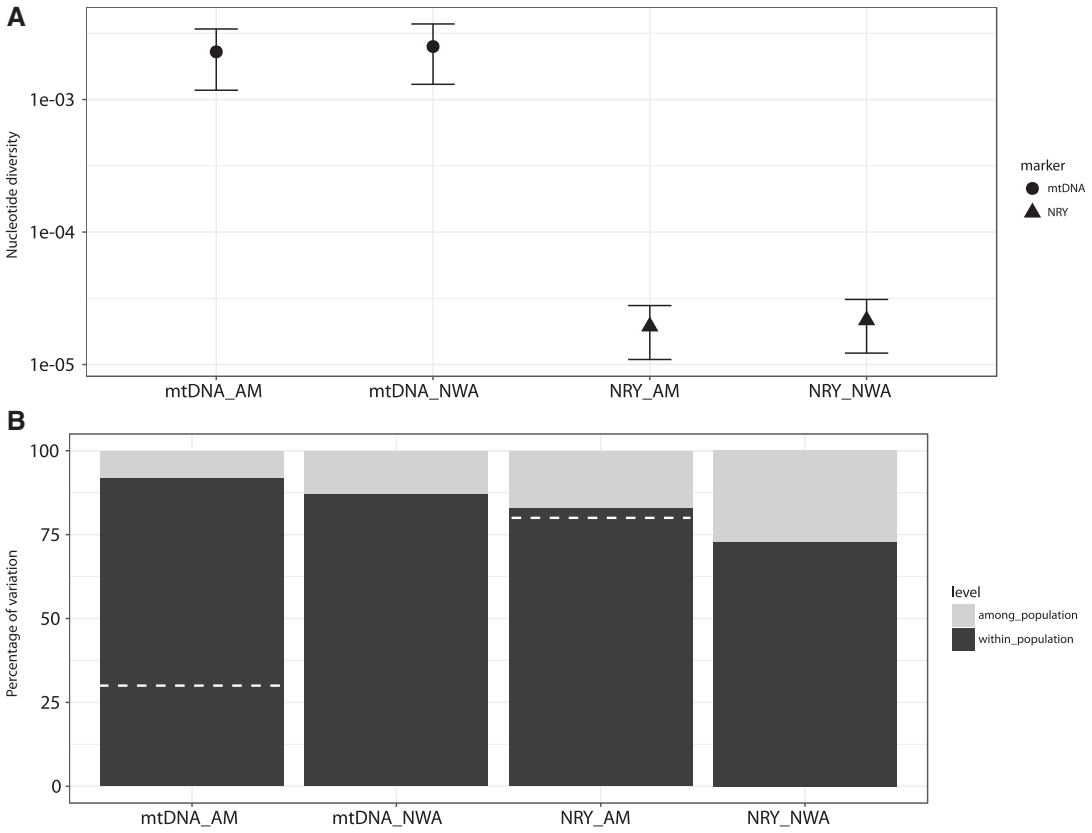

FIG. 6. (*A*) Nucleotide diversity and (*B*) AMOVA results in NWA and a sample of the Americas (AM) for the mtDNA and the NRY. Dotted lines represent the percentage of among population variation reported by Lippold et al. (2014) for the Americas.

anthropogenic soils, a result of both increasing sedentism and agricultural activities (Clement et al. 2015; Levis et al. 2017). Especially by one kya, the archaeological record shows a sudden increase in the number of sites with human occupation with the presence of dark earths (Eden et al. 1984; Neves 2008; Arroyo-Kalin 2010). Thus, the dates for the start of the lineage diversification of mtDNA and NRY lineages (fig. 4) agree with a phase of intensification of agriculture in South America. One crop that has been suggested as a candidate to allow this change in lifestyle is manioc (*M. esculenta*), especially the bitter variety, which grows well even in acidic and nutrient-poor soils, is more resistant to parasites, and produces larger and starchier tubers than the sweet manioc variety (Arroyo-Kalin 2010). However, bitter manioc tubers are highly toxic and require elaborate processing before consumption, including grating, washing, squeezing, and cooking. Therefore, technological innovations were necessary to facilitate the utilization of bitter manioc roots. These cultural innovations are likely to have been spread by migrating women: the low levels of population differentiation and the high amount of shared haplotypes in the mtDNA (supplementary figs. 2 and 3, Supplementary Material online and fig. 3) indicate that women were migrating extensively between groups. Manioc production is exclusively women's work and it includes: planting, harvesting, processing of roots, and food and drink preparation (Hugh-Jones 1979; Jackson 1983; Heckler 2004). Moreover, women are given manioc landraces as part of their dowry (Peña-Venegas et al. 2014). Thus, any technological and horticultural innovations developed in one ethnolinguistic group would have been quickly spread to other groups by migrating women. Similar observations on the role of women in the maintenance of manioc diversity through the exchange of seeds and varieties have been reported in Africa (Deletre et al. 2011).

These cultural changes went hand in hand with the establishment of complex trade networks, where multilingualism was a common trait, that connected distant parts of Amazonia, the Andes and the Caribbean (Lathrap 1973; Vidal 1997, 2002; Hornborg 2005; Heckenberger and Neves 2009). The last 4,000 years witnessed the geographic expansion of the major language families in different parts of South America, namely Arawakan, Tupi-Guarani, Carib, Ge, and Quechuan (Noelli 1998; Heckenberger 2002, 2013; Beresford-Jones and Heggarty 2011; Heggarty and Beresford-Jones 2012). Moreover, the analysis of summed calibrated radiocarbon dates for South America has shown the existence of a phase of population growth after ~5 kya (Goldberg et al. 2016). This period is associated with increased diversification of both mtDNA and NRY lineages in NWA, and hence may reflect processes that were happening in different parts of the continent, a cultural transition defined as the "Formative Period" (Ford 1969). The development of societies with increasing levels of complexity, that is, of societies living in permanent settlements and relying primarily on agriculture, producing sophisticated pottery and engaged in political, economic, and religious relationships with other societies, thus

appears to have had a notable impact on the diversification of both mtDNA and NRY lineages in NWA.

Patterns of Genetic Diversity at the Local Versus Continental Scale in Native Americans

Previous studies have shown that human populations generally exhibit larger genetic differences for the NRY than for mtDNA (Seielstad et al. 1998; Oota et al. 2001; Fagundes et al. 2002; Kayser et al. 2003; Nasidze et al. 2004; Pakendorf et al. 2007; Marchi et al. 2017). However, Lippold et al. (2014) found differences in the patterns of genetic variation among continental regions, with the Americas, Oceania, and Africa exhibiting larger between-population differences for the mtDNA than for the NRY. We further explored this discrepancy by comparing more individuals from the Americas ($n = 77$) in comparison with Lippold et al. (2014), who analyzed data for 22 individuals. In agreement with their results, we observed larger values of mtDNA nucleotide diversity than in the NRY. This could reflect differences in mutation rates between markers; however, previous studies (Wilson Sayres et al. 2014; Poznik et al. 2016) have shown that the differences in nucleotide diversity between the mtDNA and the NRY are maintained after correcting for differences in mutation rates.

In contrast to Lippold et al. (2014), our results showed smaller between-population differences for the mtDNA ($\Phi_{ST} = 0.08$) than for the NRY ($\Phi_{ST} = 0.17$) when comparing a larger number of sequences from the Americas (fig. 6). Thus, this region patterns as expected for human populations that are mostly patrilocal. The much larger between-population differences for the mtDNA than for the NRY detected by Lippold et al. in the Americas are likely a result of the small sample size used in their study: only 22 individuals from 5 populations. In addition, the Native American populations included in the HGDP are highly differentiated due to genetic drift (Reich et al. 2012), and Lippold et al. found that these populations were fixed for different mtDNA haplogroups (with the exception of the two Piapoco samples, whose sequences belong to different haplogroups). Thus, the much larger between-population differentiation observed by Lippold et al. for the mtDNA might be a result of pooling samples from geographically distant and isolated populations (Hammer et al. 2003). The impact of this sampling scheme may not be discernible with NRY data, since Native American Y-chromosomes are relatively homogenous, belonging to just one major haplogroup.

Interestingly, although we observed differences in the magnitude of the Φ_{ST} values between the local and continental comparisons, the ratio of NRY Φ_{ST}/mtDNA Φ_{ST} at each geographic scale is practically identical (~2.1). This is in contrast to simulation results that predict that the excess of between-population differentiation for the NRY, commonly observed in patrilocal groups, decreases when comparing more geographically distant populations (Wilkins and Marlowe 2006). The model used in the simulations assumes a shift from equal rates of female versus male migration to excess female migration as a consequence of a change to patrilocality at various times in the past (Wilkins and Marlowe 2006). Our results thus suggest that in the New World the migration rate has been larger for females than for males over considerable time periods, indicating that postmarital residence patterns in Native American populations may have been relatively stable over time. This is in contrast to the relatively recent shift from matrilocality toward patrilocality proposed for other regions of the world (Holden and Mace 2003; Wilkins and Marlowe 2006; Heyer et al. 2012). An ancestral tendency toward patrilocality in the New World would help explain the severely restricted NRY variation, with one major haplogroup predominating (Zegura et al. 2004; Dulik et al. 2012; Battaglia et al. 2013; Roewer et al. 2013; Jota et al. 2016). Alternatively, it could be that long-distance migration between populations in South America exhibits the same sex bias as short-distance migration, in contrast to other parts of the world (Marks et al. 2012). Whether this pattern of the same excess of NRY to mtDNA differentiation at both local and continental scales is indeed restricted to the New World, or might actually hold for other geographic regions as well, would require further empirical studies that contrast mtDNA and NRY variation in the same samples at different geographic scales within one continent.

Sex-Specific Patterns of Genetic Diversity

The BSPs and the diversity statistics (supplementary table 2, Supplementary Material online) indicate that overall the Ne of males has been smaller than that of females. One tentative explanation for this difference is that it reflects larger differences in reproductive success among males than among females. Some support for this explanation comes from the shape of the phylogenies (supplementary figs. 1 and 6, Supplementary Material online), since differences in reproductive success and the cultural transmission of fertility lead to unbalanced phylogenies (Blum et al. 2006; Heyer et al. 2015). We estimated a common index of tree imbalance (Colless index) and calculated whether the mtDNA and NRY trees were more unbalanced than 1,000 simulated trees generated under a Yule process (Bortolussi et al. 2006), that is, a simple pure birth process that assumes that the birth rate of new lineages is the same along the tree. We found that the NRY tree is more unbalanced than predicted by the Yule model (P value = 0.001), in accordance with the hypothesis of differences in reproductive success among males, whereas the mtDNA tree is not significantly different from trees generated by the Yule model (P value = 0.628). It has been suggested that highly mobile hunter-gatherer societies, such as those typical of most of human prehistory, were polygynous bands (Dupanloup et al. 2003); similarly, nomadic horticulturalist Amazonian societies exhibit strong differences in reproductive success due to the common practice of polygyny, especially among community chiefs, whose offspring also enjoy a high fertility (Neel 1970, 1980; Neel and Weiss 1975).

In addition, an expansion ~3.5 kya can be observed in the BSP based on the NRY, but not in the mtDNA BSP (fig. 5), suggesting an expansion specifically in the paternal line. This observation is supported by the larger number of negative Tajima's D values observed for the NRY than for the mtDNA (fig. 1D). Similar male-biased expansions have been observed in other studies using high-resolution NRY sequences

(Karmin et al. 2015; Batini et al. 2017; Oliveira et al. 2018), but the reasons for this are as yet unclear.

Further sex-specific patterns of genetic diversity may be the result of differences in sociocultural practices found among the populations in NWA. In this study, we have investigated in particular how the patterns of genetic variation are impacted by linguistic exogamy and postmarital residence patterns; these are discussed in the following sections.

Linguistic Exogamy

One of the most striking patterns among NWA populations is observed in the two Eastern Tukanoan groups, who show higher than average mtDNA haplotype diversity and lower than average NRY haplotype diversity (fig. 1A), as well as an extreme reduction of the MPD for the NRY (fig. 1B). In our previous study (Arias et al. 2018), which included more Eastern Tukanoan groups, we found that five out of the six groups showed higher than average mtDNA diversity. Eastern Tukanoan groups represent a prime example of the impact of cultural practices on the patterns of genetic diversity. Specifically, Eastern Tukanoans practice linguistic exogamy, a cultural norm in which marriages are required to occur between individuals speaking different languages. This system promotes the migration of women among ethnolinguistic groups, since it is reinforced by patrilocality and patrilineality, in which married men continue to live in their fathers' territory and the individual's identity is determined by the father's ethnolinguistic identity (Sorensen 1967; Stenzel 2005). For instance, Stenzel (2005) reports that in the Vaupes region 75–90% of the married men continue to reside in the subregion in which they lived before marriage, whereas this is true for only 50–58% of the women in the same region. This situation is likely responsible for the high levels of mtDNA and low levels of NRY genetic diversity found here for the Eastern Tukanoan Desano and Pira-Wanano.

Postmarital Residence Patterns

Patrilocal groups are expected to exhibit lower levels of within-population genetic diversity and higher levels of between-population divergence (i.e., high pairwise Φ_{ST}s) for the NRY, and higher levels of within-population diversity and lower levels of divergence for the mtDNA; matrilocal groups are expected to show the opposite pattern (Oota et al. 2001). However, as found previously in other regions of the world (Kumar et al. 2006; Ascunce et al. 2008; Gunnarsdottir et al. 2011), our data do not conform to the patterns expected for these postmarital residence practices. Only 3 populations out of 13 that are classified as patrilocal, namely, the two Eastern Tukanoan groups Desano and Pira-Wanano and the Arawakan group Yucu-Matapi, showed lower values of NRY haplotype diversity than two of the matrilocal groups, namely the Carijona and Guayabero, whereas only the Desano and Pira-Wanano exhibited larger mtDNA haplotype diversity values. The Sikuani, also classified as matrilocal, showed lower NRY haplotype diversity values, in contrast to what would be expected from their residence pattern. In general, there is considerable heterogeneity among patrilocal groups in the haplotype diversity values (fig. 1A). For instance, the patrilocal groups Ach-Piapoco and Siona showed a pattern that looks more like matrilocality, namely high NRY and low mtDNA haplotype diversity. Similarly, the three groups classified as matrilocal—Carijona, Sikuani, and Guayabero—did not significantly differ from the haplotype diversity values observed for many of the patrilocal groups (fig. 1A).

Since the patrilocal groups in our study are more heterogeneous and considerably outnumber the matrilocal groups, the resampling approach allowed for a more reliable comparison between residence patterns. This analysis detected that triplets of patrilocal groups do have lower MPD values for the NRY than the matrilocal groups (P value $= 0.063$). In the other comparisons, we also observed a tendency in the distribution of the evaluated statistics that is compatible with the general expectations for patrilocality versus matrilocality (fig. 2). Thus, patrilocal groups showed on average smaller pairwise Φ_{ST} and larger MPD for the mtDNA than matrilocal groups. Meanwhile, we observed the opposite pattern for the NRY, that is, larger pairwise Φ_{ST} values and smaller MPD in the patrilocal than in the matrilocal groups. In contrast, we did not find significant differences in the haplotype diversity values between matrilocal and patrilocal groups (fig. 2). These results indicate that NWA populations show considerable heterogeneity in the levels of genetic variation, which cannot be fully explained by a simple distinction between matrilocal and patrilocal groups. Indeed, previous studies that investigated the impact of postmarital residence practices on the patterns of genetic diversity have arrived at contrasting results, with some studies observing a good match between the marital practices and the levels of genetic diversity (Oota et al. 2001; Besaggio et al. 2007), whereas other studies failed to find such a match (Kumar et al. 2006; Ascunce et al. 2008; Gunnarsdottir et al. 2011; Ly et al. 2018). This suggests that although marital practices have an impact on patterns of diversity, this is just one of the factors at play in structuring the observed genetic variation (see Wilkins and Marlowe 2006; Heyer et al. 2012, and Marks et al. 2012 for further discussion). Social organization in humans is influenced by several interrelated aspects of human culture, the physical landscape, and subsistence strategies (Balaresque and Jobling 2007), and changes in each of these can affect the genetic diversity. For instance, Chaix et al. (2007) observed that a rapid change from nomadic pastoralism to sedentary farming ~20 generations ago among some Uzbek populations was followed by changes in the descent rules from patrilineal to extended or nuclear families, leading in turn to changes in the patterns of genetic diversity (Chaix et al. 2007). The mismatch observed here between postmarital residence rules and patterns of genetic diversity in the mtDNA and NRY among NWA populations could thus reflect changes in social organization due to range expansions of agriculturalist groups or as a consequence of European contact since the 16th century, which have impacted the expected signatures associated with marital rules. Furthermore, our results might be an indication that in practice there is more flexibility in postmarital residence rules than what is considered the cultural norm in theory (Hamilton et al. 2005; Walker 2015; Epps 2018; Ly et al. 2018).

To conclude, human prehistory is a complex phenomenon that can only be elucidated through the investigation of multiple lines of evidence from multiple disciplines. Our study has contributed to the understanding of the genetic history of the populations of NWA. By analyzing high-resolution mtDNA and NRY sequences, we have found that males and females have experienced different demographic histories. Cultural practices can account for some of the differences in the patterns of genetic diversity, for instance, by promoting differential reproductive success among individuals and determining the way the sexes disperse after marriage. We anticipate that the analysis of genome-wide data and ancient DNA studies of human remains will help fill in the remaining gaps in our knowledge.

Materials and Methods

Sample Collection

Saliva or blood samples were collected during several expeditions carried out by one of the authors (L.A.). Written informed consent was obtained from each participant and from the community leader and/or local/regional indigenous organizations after giving a full description of the aims of the study. All procedures were undertaken in accordance with the Declaration of Helsinki and the study was approved by the Institutional Review Committee on Human Ethics of the Universidad del Valle in Cali, Colombia, and the Ethics Commission of the University of Leipzig Medical Faculty. The total sample collection comprises 460 samples belonging to 40 different ethnolinguistic groups (see Arias et al. [2018] for details).

DNA Sequencing and Sequence Analysis

Double indexed DNA libraries (Kircher et al. 2012) were enriched for a region of 2.3 Mb of the nonrecombining region of the Y-chromosome (NRY) via in-solution capture (Kutanan et al. 2018), following the SureSelect protocol from Agilent with modifications described in Kircher et al. (2012). Three pools containing 90 samples each were prepared, and paired-end sequencing (read lengths 100 bp) was carried out on three lanes of the Illumina HiSeq 2500 platform; base-calling was performed with Bustard. Illumina adaptors were trimmed and reads starting from opposite directions were merged with leeHOM (Renaud et al. 2014) to completely overlap paired sequences. Finally, sequences were de-multiplexed with the program deML (Renaud et al. 2015) and aligned to the human reference genome *hg19* using BWA's *aln* algorithm. All sequence pairs that aligned to the NRY regions defined by Poznik et al (2013) were retained (Poznik et al. 2013). Duplicate reads were removed using PicardTools *MarkDuplicates*, indel realignment was performed using GATK *IndelRealigner* (McKenna et al. 2010), and base quality was re-calibrated using GATK *BaseRecalibrator*. We identified single nucleotide variants (SNVs) using GATK *UnifiedGenotyper* v3.3-0 across all samples simultaneously, setting the parameter *ploidy* to 1 and using dbSNP build 138 as a prior position list. The identified SNVs were further filtered as previously described by Barbieri et al (2016) and we obtained a set of 2,969 SNVs. We imputed all samples with missing genotype information at any of these variant sites using BEAGLE (Browning and Browning 2013) and assigned Y chromosome haplogroups using yhaplo (Poznik et al. 2016).

The complete mtDNA genome sequences were taken from a previous study (Arias et al. 2018).

Y-Chromosome Sequences

We generated NRY sequences from 284 individuals. From these we excluded seven "Mestizo" individuals (i.e., individuals with paternal origin from outside NWA) as well as 30 sequences belonging to nonautochthonous haplogroups. For the population-based analyses we excluded a further 32 sequences, since we restricted these analyses to populations with a sample size of at least 10 individuals for which we have both mtDNA and NRY sequence data. The sole exception was the Carijona ($n = 6$), since it is the only Carib-speaking group living in NWA. Some ethnolinguistic groups were merged into single populations based on linguistic criteria when their population sizes were smaller than 10 individuals, as previously described in Arias et al. (2018). Thus, 215 NRY sequences and 330 complete mtDNA genome sequences from 17 ethnolinguistic groups were included in the population-based analyses. In contrast, for the phylogenetic and demographic reconstructions we included all the autochthonous sequences with the exception of three sequences belonging to haplogroup C2; these reconstructions were therefore based on 244 sequences belonging to haplogroup Q1. Seven populations included in Arias et al. (2018) were excluded due to insufficient numbers of NRY sequences. These are: the Eastern Tukanoan groups Tuka-Tatuyo, Siriano, Other-ET, and Tanimuka, as well as the three groups from the Andean foothills: Pasto, Kamentsa, and Inga. The population-based analyses included: the estimation of haplogroup frequencies (by counting); diversity values (i.e., haplotype diversity, mean number of pairwise differences [MPD], and nucleotide diversity), AMOVA, and pairwise Φ_{ST} genetic distances, which were computed in Arlequin 3.5. (Excoffier and Lischer 2010); multidimensional scaling (MDS) analyses based on pairwise genetic distance matrices, which were performed with the R package *MASS* (Venables and Ripley 2002); and the analysis of shared haplotypes, which was performed with in-house R scripts.

Furthermore, we investigated the impact of patrilocality and matrilocality on the patterns of genetic diversity. Our data set contains 13 populations classified as patrilocal, three as matrilocal, and one as ambi/neolocal (table 1), as reported in different sources (Morey et al. 1973; Morey and Morey 1980; Walker et al. 2010; Kirby et al. 2016). To control for the difference in the number of groups in each category and the heterogeneity among patrilocal groups, we devised a resampling strategy in which we sampled all possible triplets of patrilocal populations (286 combinations in total). For each triplet we computed the average of the observed values of each of the following statistics: pairwise Φ_{ST} values, haplotype diversity, and the MPD for the mtDNA and the NRY. We then compared the distribution of these values among all

triplets of patrilocal groups to the average value estimated for the three matrilocal groups.

Genetic Structure and Population Relationships

For the AMOVA we defined clusters based on linguistic criteria, geographic proximity, and the distribution of populations along rivers; the latter is an important factor in structuring the mtDNA variation among NWA populations (Arias et al. 2018). For inferring the relationships among populations, we used pairwise Φ_{ST} values, as well as the proportion of shared haplotypes between populations. In addition, we performed Mantel tests with the R package *ade4* (Dray and Dufour 2007) to evaluate whether there are significant correlations between pairwise genetic distances among populations for the NRY and the mtDNA, as well as between genetic distances and geographic distances. For the matrix of geographic distances we used the geographic coordinates of the locality that contained the majority of individuals for each ethnolinguistic group, and we calculated great circle distances between locations with the R package *geosphere* (Hijmans 2016). Furthermore, to account for the role of rivers as facilitators of contact among groups, we performed a multiple regression analysis on distance matrices (Goslee and Urban 2007) as previously described in Arias et al. (2018). For this analysis, we consider the matrix of NRY Φ_{ST} values, a matrix of geographic distances, and a matrix of river distances, in which populations living along the same river or on rivers closely connected take a distance value of zero and populations living on different rivers a value of one.

Phylogenetic Inferences and Demographic Reconstructions

We reconstructed the phylogenies of the mtDNA and NRY sequences, using 428 complete mtDNA genome sequences reported by Arias et al. (2018), and 244 NRY sequences belonging to haplogroup Q1. We estimated maximum clade credibility trees for both mtDNA and NRY sequences with BEAST 1.8.2 (Drummond et al. 2012). The best nucleotide substitution model was estimated with jModeltest 2.1.7 (Darriba et al. 2012), and we tested if the data best fit a strict clock or an uncorrelated log normal relaxed clock model using stepping-stone sampling (Baele et al. 2013) and Bayes factor analysis (Kass and Raftery 1995). Finally, we used the substitution rate of 7.6×10^{-10} per base pair per year reported by Fu et al. (2014) for the NRY, and substitution rates of 1.708×10^{-8} and 9.883×10^{-8} per base pair per year reported by Soares et al. (2009) for the coding and noncoding regions, respectively, of the mtDNA genome. In addition, a median-joining network of haplotypes for the NRY sequences was generated with the software Network 4.6.1.5 and visualized with Network Publisher 2.0.0.1 (http://www.fluxus-engineering.com).

We made inferences about changes in population size through time with Bayesian skyline plots (BSPs) as implemented in BEAST 1.8.2 (Drummond et al. 2012). For this we used the same settings defined for the phylogenetic reconstructions (substitution model, substitution rate, clock model, etc.), but using the coalescent tree prior BSP. We performed this analysis by haplogroup and by population. Furthermore, since violations of the panmixia assumption are known to lead to misleading results in coalescent-based methods (Heller et al. 2013; Grant 2015), we implemented the sampling strategies suggested by Städler et al. (2009), namely, "pooled" and "scattered", and compared them with the results from the groups included in the population-based analyses, which corresponds to the "local" sampling scheme in (Städler et al. 2009). The pooled sample strategy consisted of randomly sampling four individuals from each population, whereas the scattered sample consisted of randomly sampling one individual from each population; each strategy was replicated ten times. This was performed for both the mtDNA and the NRY, and BSPs were obtained for each replicate.

Patterns of Genetic Diversity at a Local Versus Continental Scale in Native Americans

To investigate the effect of geographic scale on the patterns of genetic diversity, we compared several summary statistics at two levels: first, our local populations from NWA; and second, a comparative data set of published indigenous mtDNA and NRY sequences from the Americas. The NRY comparative data set was composed of 52 sequences from the Americas (Karmin et al. 2015; Mallick et al. 2016; Poznik et al. 2016), grouped by country of origin (i.e., Argentina $n = 12$, Mexico $n = 15$, and Peru $n = 25$) and a random sample ($n = 25$) of NRY sequences from NWA as representative of Colombia. We produced five NRY data set replicates, by randomly sampling 25 NRY sequences from Colombia five independent times and merging with our 52 sequences from the literature. The mtDNA comparative data set contained the same number of sequences by country as the NRY data set (Argentina = 12, Mexico = 15, and Peru = 25, Colombia = 25). Five data set replicates were generated by randomly sampling mtDNA sequences from a large collection of sequences from the literature (Tamm et al. 2007; Fagundes et al. 2008; Perego et al. 2009, 2010; Kumar et al. 2011; Bodner et al. 2012; Cardoso et al. 2012; de Saint Pierre et al. 2012; Gomez-Carballa et al. 2012; Achilli et al. 2013; Lippold et al. 2014; Mizuno et al. 2014; Lee and Merriwether 2015; Barbieri et al. 2017; Arias et al. 2018). We estimated the nucleotide diversity and performed an AMOVA considering two hierarchical levels, namely differences among populations and within populations. We compared the average values over the five replicates to the values observed in the NWA data set.

Supplementary Material

Supplementary data are available at *Molecular Biology and Evolution* online.

Acknowledgments

We greatly acknowledge the contribution of all sample donors, communities, community leaders, and regional indigenous organizations, without whose contribution this study would not be possible. We dedicate this article to Rafael Rodríguez, who assisted greatly with the fieldwork and passed away in September 2017. We also acknowledge Enrico

Macholdt, Sandra Oliveira, Michael Dannemann, and Benjamin Peter for advice with data analysis. We thank Anna Paschall for her assistance with sample preparation, and Thiago Chacon, Zachary O'Hagan, and Jorge Rosés Labrada for their insights into the ethnohistory of Amazonia. We thank David Beresford-Jones, Paul Heggarty, and Frank Seifart for comments on an earlier version of this manuscript. B.P. acknowledges the LABEX ASLAN (ANR-10-LABX-0081) of Université de Lyon for its financial support within the program "Investissements d'Avenir" (ANR-11-IDEX-0007) of the French government operated by the National Research Agency (ANR). L.A. was supported by a graduate grant from COLCIENCIAS. This research was supported by funds from the Max Planck Society. Grant sponsorship: Max Planck Society + Colciencas.

References

Aceituno FJ, Loaiza N, Delgado-Burbano ME, Barrientos G. 2013. The initial human settlement of Northwest South America during the Pleistocene/Holocene transition: synthesis and perspectives. *Quatern Int*. 301:23–33.

Achilli A, Perego UA, Lancioni H, Olivieri A, Gandini F, Hooshiar Kashani B, Battaglia V, Grugni V, Angerhofer N, Rogers MP, et al. 2013. Reconciling migration models to the Americas with the variation of North American native mitogenomes. *Proc Natl Acad Sci U S A*. 110(35):14308–14313.

Arias L, Barbieri C, Barreto G, Stoneking M, Pakendorf B. 2018. High-resolution mitochondrial DNA analysis sheds light on human diversity, cultural interactions, and population mobility in Northwestern Amazonia. *Am J Phys Anthropol*. 165(2):238–255.

Arroyo-Kalin M. 2010. The Amazonian formative: crop domestication and anthropogenic soils. *Diversity* 2(4): 473–504.

Arroyo-Kalin M. 2012. Slash-burn-and-churn: landscape history and crop cultivation in pre-Columbian Amazonia. *Quatern Int*. 249:4–18.

Ascunce MS, Gonzalez-Oliver A, Mulligan CJ. 2008. Y-chromosome variability in four Native American populations from Panama. *Hum Biol*. 80(3): 287–302.

Baele G, Li WL, Drummond AJ, Suchard MA, Lemey P. 2013. Accurate model selection of relaxed molecular clocks in Bayesian phylogenetics. *Mol Biol Evol*. 30(2):239–243.

Balanovsky O, Gurianov V, Zaporozhchenko V, Balaganskaya O, Urasin V, Zhabagin M, Grugni V, Canada R, Al-Zahery N, Raveane A, et al. 2017. Phylogeography of human Y-chromosome haplogroup q3-l275 from an academic/citizen science collaboration. *BMC Evol Biol*. 17(S1):18.

Balaresque P, Jobling MA. 2007. Human populations: houses for spouses. *Curr Biol*. 17(1):R14–R16.

Barbieri C, Hubner A, Macholdt E, Ni S, Lippold S, Schroder R, Mpoloka SW, Purps J, Roewer L, Stoneking M, et al. 2016. Refining the Y chromosome phylogeny with southern African sequences. *Hum Genet*. 135(5):541–553.

Barbieri C, Sandoval JR, Valqui J, Shimelman A, Ziemendorff S, Schroder R, Geppert M, Roewer L, Gray R, Stoneking M, et al. 2017. Enclaves of genetic diversity resisted Inca impacts on population history. *Sci Rep*. 7(1):17411.

Batini C, Hallast P, Vagene AJ, Zadik D, Eriksen HA, Pamjav H, Sajantila A, Wetton JH, Jobling MA. 2017. Population resequencing of European mitochondrial genomes highlights sex-bias in Bronze Age demographic expansions. *Sci Rep*. 7(1):12086.

Battaglia V, Grugni V, Perego UA, Angerhofer N, Gomez-Palmieri JE, Woodward SR, Achilli A, Myres N, Torroni A, Semino O. 2013. The first peopling of South America: new evidence from Y-chromosome haplogroup Q. *Plos One* 8(8):e71390.

Bellwood, P. and C. Renfrew, editors. Examining the farming/language dispersal hypothesis. Cambridge: McDonald Institute for Archaeological Research.

Beresford-Jones D, Heggarty P. 2011. What role for language prehistory in redefining archaeological "culture"? A case study on new horizons in the Andes. In: Roberts BW, Vander Linden M, editors. Investigating archaeological cultures: material culture, variability, and transmission. New York: Springer New York. p. 355–386.

Besaggio D, Fuselli S, Srikummool M, Kampuansai J, Castri L, Tyler-Smith C, Seielstad M, Kangwanpong D, Bertorelle G. 2007. Genetic variation in Northern Thailand Hill Tribes: origins and relationships with social structure and linguistic differences. *BMC Evol Biol*. 7(2 Suppl):S12.

Blum MG, Heyer E, Francois O, Austerlitz F. 2006. Matrilineal fertility inheritance detected in hunter-gatherer populations using the imbalance of gene genealogies. *PLoS Genet*. 2(8):e122.

Bodner M, Perego UA, Huber G, Fendt L, Rock AW, Zimmermann B, Olivieri A, Gomez-Carballa A, Lancioni H, Angerhofer N, et al. 2012. Rapid coastal spread of First Americans: novel insights from South America's Southern Cone mitochondrial genomes. *Genome Res*. 22(5):811–820.

Bortolussi N, Durand E, Blum M, Francois O. 2006. Aptreeshape: statistical analysis of phylogenetic tree shape. *Bioinformatics* 22(3):363–364.

Browning BL, Browning SR. 2013. Improving the accuracy and efficiency of identity-by-descent detection in population data. *Genetics* 194(2):459–471.

Burton ML, Moore CC, Romney AK, Aberle DF, Barcelo JA, Dow MM, Guyer JI, Kronenfeld DB, Levy JE, Linnekin J. 1996. Regions based on social structure. *Curr Anthropol*. 37(1):87–123.

Bush MB, Correa-Metrio A, McMichael CH, Sully S, Shadik CR, Valencia BG, Guilderson T, Steinitz-Kannan M, Overpeck JT. 2016. A 6900-year history of landscape modification by humans in lowland Amazonia. *Quatern Sci Rev*. 141:52–64.

Cardoso S, Alfonso-Sanchez MA, Valverde L, Sanchez D, Zarrabeitia MT, Odriozola A, Martinez-Jarreta B, de Pancorbo MM. 2012. Genetic uniqueness of the Waorani tribe from the Ecuadorian Amazon. *Heredity (Edinb)* 108(6):609–615.

Chaix R, Quintana-Murci L, Hegay T, Hammer MF, Mobasher Z, Austerlitz F, Heyer E. 2007. From social to genetic structures in central Asia. *Curr Biol*. 17(1): 43–48.

Chernela JM. 2010. The Wanano Indians of the Brazilian Amazon: a sense of space. Austin (TX): University of Texas Press.

Clement CR, Denevan WM, Heckenberger MJ, Junqueira AB, Neves EG, Teixeira WG, Woods WI. 2015. The domestication of Amazonia before European conquest. *Roy Soc B Biol Sci*. 282(1812): 20150813–20150840.

Cross SL, Baker PA, Seltzer GO, Fritz SC, Dunbar RB. 2000. A new estimate of the Holocene Lowstand level of Lake Titicaca, central Andes, and implications for tropical palaeohydrology. *Holocene* 10(1):21–32.

Darriba D, Taboada GL, Doallo R, Posada D. 2012. Jmodeltest 2: more models, new heuristics and parallel computing. *Nat Methods* 9(8):772.

Deletre M, McKey DB, Hodkinson TR. 2011. Marriage exchanges, seed exchanges, and the dynamics of manioc diversity. *Proc Natl Acad Sci U S A*. 108(45):18249–18254.

Denevan WM. 1992. The pristine myth: the landscape of the Americas in 1492. *Ann Assoc Am Geogr*. 82(3):369–385.

de Saint Pierre M, Gandini F, Perego UA, Bodner M, Gómez-Carballa A, Corach D, Angerhofer N, Woodward SR, Semino O, Salas A, et al. 2012. Arrival of Paleo-Indians to the Southern Cone of South America: new clues from mitogenomes. *PLoS One* 7(12):e51311.

Diamond J, Bellwood P. 2003. Farmers and their languages: the first expansions. *Science* 300(5619):597–603.

Dillehay TD, Ocampo C, Saavedra J, Sawakuchi AO, Vega RM, Pino M, Collins MB, Scott Cummings L, Arregui I, Villagran XS, et al. 2015. New archaeological evidence for an early human presence at monte verde, chile. *PLoS One* 10(11):e0141923.

Dray S, Dufour AB. 2007. The ade4 package: implementing the duality diagram for ecologists. *J Stat Softw*. 22(4):1–20.

Drummond AJ, Suchard MA, Xie D, Rambaut A. 2012. Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Mol Biol Evol*. 29(8):1969–1973.

Dulik MC, Owings AC, Gaieski JB, Vilar MG, Andre A, Lennie C, Mackenzie MA, Kritsch I, Snowshoe S, Wright R, et al. 2012. Y-chromosome analysis reveals genetic divergence and new founding native lineages in athapaskan- and eskimoan-speaking populations. *Proc Natl Acad Sci U S A*. 109(22):8471–8476.

Dupanloup I, Pereira L, Bertorelle G, Calafell F, Prata MJ, Amorim A, Barbujani G. 2003. A recent shift from polygyny to monogamy in humans is suggested by the analysis of worldwide Y-chromosome diversity. *J Mol Evol*. 57(1):85–97.

Eden MJ, Bray W, Herrera L, Mcewan C. 1984. Terra-preta soils and their archaeological context in the caqueta basin of southeast Colombia. *Am Antiquity* 49(01):125–140.

Epps P. 2018. Contrasting linguistic ecologies: indigenous and colonially mediated language contact in northwest Amazonia. *Lang Commun*.

Erickson CL. 2008. Amazonia: the historical ecology of a domesticated landscape. In: Silverman H, William HI, editors. Handbook of South American archaeology. New York: Springer-Verlag New York. p. 157–183.

Excoffier L, Lischer HE. 2010. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Mol Ecol Resour*. 10(3):564–567.

Fagundes NJ, Bonatto SL, Callegari-Jacques SM, Salzano FM. 2002. Genetic, geographic, and linguistic variation among South American Indians: possible sex influence. *Am J Phys Anthropol*. 117(1):68–78.

Fagundes NJ, Kanitz R, Eckert R, Valls AC, Bogo MR, Salzano FM, Smith DG, Silva WA Jr, Zago MA, Ribeiro-dos-Santos AK, et al. 2008. Mitochondrial population genomics supports a single pre-Clovis origin with a coastal route for the peopling of the Americas. *Am J Hum Genet*. 82(3):583–592.

Fenner JN. 2005. Cross-cultural estimation of the human generation interval for use in genetics-based population divergence studies. *Am J Phys Anthropol*. 128(2):415–423.

Ford, JA. 1969. A comparison of formative cultures in the Americas: Diffusion or the psychic unity of man. Washington (DC): Smithsonian Institution Press.

Fu Q, Li H, Moorjani P, Jay F, Slepchenko SM, Bondarev AA, Johnson PL, Aximu-Petri A, Prufer K, de Filippo C, et al. 2014. Genome sequence of a 45,000-year-old modern human from western Siberia. *Nature* 514(7523):445–449.

Gignoux CR, Henn BM, Mountain JL. 2011. Rapid, global demographic expansions after the origins of agriculture. *Proc Natl Acad Sci U S A*. 108(15):6044–6049.

Goldberg A, Mychajliw AM, Hadly EA. 2016. Post-invasion demography of prehistoric humans in South America. *Nature* 532(7598):232–235.

Gomez-Carballa A, Ignacio-Veiga A, Alvarez-Iglesias V, Pastoriza-Mourelle A, Ruiz Y, Pineda L, Carracedo A, Salas A. 2012. A melting pot of multicontinental mtDNA lineages in admixed Venezuelans. *Am J Phys Anthropol*. 147(1):78–87.

Goslee SC, Urban DL. 2007. The ecodist package for dissimilarity-based analysis of ecological data. *J Stat Softw*. 22(7):1–19.

Grant WS. 2015. Problems and cautions with sequence mismatch analysis and Bayesian skyline plots to infer historical demography. *J Hered*. 106(4):333–346.

Gunnarsdottir ED, Nandineni MR, Li M, Myles S, Gil D, Pakendorf B, Stoneking M. 2011. Larger mitochondrial DNA than Y-chromosome differences between matrilocal and patrilocal groups from Sumatra. *Nat Commun*. 2:228.

Hamilton G, Stoneking M, Excoffier L. 2005. Molecular analysis reveals tighter social regulation of immigration in patrilocal populations than in matrilocal populations. *Proc Natl Acad Sci U S A*. 102(21):7476–7480.

Hammer MF, Blackmer F, Garrigan D, Nachman MW, Wilder JA. 2003. Human population structure and its effects on sampling Y chromosome sequence variation. *Genetics* 164(4):1495–1509.

Heckenberger MJ. 2002. Rethinking the Arawakan diaspora: hierarchy, regionality, and the Amazonian formative. In: Hill JD, Santos-Granero F, editors. Comparative Arawakan histories: rethinking language family and culture area in Amazonia. Urbana (IL): University of Illinois Press. p. 99–122.

Heckenberger MJ. 2013. 51 Amazonia: archaeology. In: Ness I, editor. The encyclopedia of global human migration. Hoboken (NJ): Blackwell Publishing Ltd.

Heckenberger MJ, Kuikuro A, Kuikuro UT, Russell JC, Schmidt M, Fausto C, Franchetto B. 2003. Amazonia 1492: pristine forest or cultural parkland? *Science* 301(5640):1710–1714.

Heckenberger MJ, Neves EG. 2009. Amazonian archaeology. *Annu Rev Anthropol*. 38(1):251–266.

Heckler SL. 2004. Tedium and creativity: the valorization of manioc cultivation and Piaroa women. *J R Anthropol Inst*. 10(2):241–259.

Heggarty P, Beresford-Jones D. 2012. Conclusion: a cross-disciplinary prehistory for the Andes? Surveying the state of the art. Archaeology and language in the Andes a cross-disciplinary exploration of prehistory. New York: Oxford New York Published for the British Academy by Oxford University Press. p. 407–434.

Heller R, Chikhi L, Siegismund HR. 2013. The confounding effect of population structure on Bayesian skyline plot inferences of demographic history. *PLoS One* 8(5):e62992.

Heyer E, Brandenburg JT, Leonardi M, Toupance B, Balaresque P, Hegay T, Aldashev A, Austerlitz F. 2015. Patrilineal populations show more male transmission of reproductive success than cognatic populations in central Asia, which reduces their genetic diversity. *Am J Phys Anthropol*. 157(4):537–543.

Heyer E, Chaix R, Pavard S, Austerlitz F. 2012. Sex-specific demographic behaviours that shape human genomic variation. *Mol Ecol*. 21(3):597–612.

Hijmans RJ. 2016. Geosphere: spherical trigonometry. R package version 1.5-5 ed. Available from: https://CRAN.R-project.org/package=geosphere.

Holden CJ, Mace R. 2003. Spread of cattle led to the loss of matrilineal descent in Africa: a coevolutionary analysis. *Proc Biol Sci*. 270(1532):2425–2433.

Hornborg A. 2005. Ethnogenesis, regional integration, and ecology in prehistoric Amazonia—toward a system perspective. *Curr Anthropol*. 46(4):589–620.

Hugh-Jones S. 1979. The palm and the Pleiades: initiation and cosmology in northwest Amazonia. Cambridge; New York: Cambridge University Press.

Iriarte J, Smith RJ, de Souza JG, Mayle FE, Whitney BS, Cardenas ML, Singarayer J, Carson JF, Roy S, Valdes P. 2017. Out of Amazonia: Late-Holocene climate change and the Tupi-Guarani trans-continental expansion. *Holocene* 27(7):967–975.

Jackson JE. 1983. The fish people: linguistic exogamy and Tukanoan identity in northwest Amazonia. Cambridge: Cambridge University Press.

Jobling MA, Tyler-Smith C. 2003. The human Y chromosome: an evolutionary marker comes of age. *Nat Rev Genet*. 4(8):598–612.

Jota MS, Lacerda DR, Sandoval JR, Vieira PP, Ohasi D, Santos-Junior JE, Acosta O, Cuellar C, Revollo S, Paz YMC, et al. 2016. New native South American Y chromosome lineages. *J Hum Genet*. 61(7):593–603.

Karmin M, Saag L, Vicente M, Wilson Sayres MA, Jarve M, Talas UG, Rootsi S, Ilumae AM, Magi R, Mitt M, et al. 2015. A recent bottleneck of Y chromosome diversity coincides with a global change in culture. *Genome Res*. 25(4): 459–466.

Kass RE, Raftery AE. 1995. Bayes factors. *J Am Stat Assoc*. 90(430):773–795.

Kayser M, Brauer S, Weiss G, Schiefenhovel W, Underhill P, Shen P, Oefner P, Tommaseo-Ponzetta M, Stoneking M. 2003. Reduced Y-chromosome, but not mitochondrial DNA, diversity in human populations from West New Guinea. *Am J Hum Genet*. 72(2): 281–302.

Kirby KR, Gray RD, Greenhill SJ, Jordan FM, Gomes-Ng S, Bibiko HJ, Blasi DE, Botero CA, Bowern C, Ember CR, et al. 2016. D-place: a global database of cultural, linguistic and environmental diversity. *PLoS One* 11(7):e0158391.

Kircher M, Sawyer S, Meyer M. 2012. Double indexing overcomes inaccuracies in multiplex sequencing on the Illumina platform. *Nucleic Acids Res*. 40(1):e3.
Kivisild T. 2017. The study of human Y chromosome variation through ancient DNA. *Hum Genet*. 136(5):529–546.
Koch-Grünberg T. 1995. Dos años entre los indios: viajes por el noroeste brasileño, 1903–1905. Bogotá (Colombia): Editorial Universidad Nacional.
Kumar S, Bellis C, Zlojutro M, Melton PE, Blangero J, Curran JE. 2011. Large scale mitochondrial sequencing in Mexican Americans suggests a reappraisal of Native American origins. *BMC Evol Biol*. 11:293.
Kumar V, Langstieh BT, Madhavi KV, Naidu VM, Singh HP, Biswas S, Thangaraj K, Singh L, Reddy BM. 2006. Global patterns in human mitochondrial DNA and Y-chromosome variation caused by spatial instability of the local cultural processes. *PLoS Genet*. 2(4):e53.
Kutanan W, Kampuansai J, Changmai P, Flegontov P, Schroder R, Macholdt E, Hubner A, Kangwanpong D, Stoneking M. 2018. Contrasting maternal and paternal genetic variation of hunter-gatherer groups in thailand. *Sci Rep*. 8(1):1536.
Lathrap DW. 1973. The antiquity and importance of long-distance trade relationships in the moist tropics of pre-Columbian South America. *World Archaeol*. 5(2):170–186.
Lee EJ, Merriwether DA. 2015. Identification of whole mitochondrial genomes from venezuela and implications on regional phylogenies in South America. *Hum Biol*. 87(1):29–38.
Levis C, Costa FRC, Bongers F, Pena-Claros M, Clement CR, Junqueira AB, Neves EG, Tamanaha EK, Figueiredo FOG, Salomao RP, et al. 2017. Persistent effects of pre-Columbian plant domestication on Amazonian forest composition. *Science* 355(6328):925.
Lippold S, Xu H, Ko A, Li M, Renaud G, Butthof A, Schroder R, Stoneking M. 2014. Human paternal and maternal demographic histories: insights from high-resolution Y chromosome and mtDNA sequences. *Investig Genet*. 5:13.
Llamas B, Fehren-Schmitz L, Valverde G, Soubrier J, Mallick S, Rohland N, Nordenfelt S, Valdiosera C, Richards SM, Rohrlach A, et al. 2016. Ancient mitochondrial DNA provides high-resolution time scale of the peopling of the Americas. *Sci Adv*. 2(4):e1501385.
Ly G, Alard B, Laurent R, Lafosse S, Toupance B, Monidarin C, Diffloth G, Bourdier F, Evrard O, Pavard S, et al. 2018. Residence rule flexibility and descent groups dynamics shape uniparental genetic diversities in South East Asia. *Am J Phys Anthropol*. 165(3):480–491.
Mallick S, Li H, Lipson M, Mathieson I, Gymrek M, Racimo F, Zhao M, Chennagiri N, Nordenfelt S, Tandon A, et al. 2016. The Simons Genome Diversity Project: 300 genomes from 142 diverse populations. *Nature* 538(7624):201–206.
Marchant R, Hooghiemstra H. 2004. Rapid environmental change in African and South American tropics around 4000 years before present: a review. *Earth Sci Rev*. 66(3–4):217–260.
Marchi N, Hegay T, Mennecier P, Georges M, Laurent R, Whitten M, Endicott P, Aldashev A, Dorzhu C, Nasyrova F, et al. 2017. Sex-specific genetic diversity is shaped by cultural factors in inner Asian human populations. *Am J Phys Anthropol*. 162(4):627–640.
Marks SJ, Levy H, Martinez-Cadenas C, Montinaro F, Capelli C. 2012. Migration distance rather than migration rate explains genetic diversity in human patrilocal groups. *Mol Ecol*. 21(20): 4958–4969.
McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, Garimella K, Altshuler D, Gabriel S, Daly M, et al. 2010. The genome analysis toolkit: a mapreduce framework for analyzing next-generation DNA sequencing data. *Genome Res*. 20(9):1297–1303.
Meggers BJ. 1954. Environmental limitation on the development of culture. *Am Anthropol*. 56(5):801–824.
Mesa NR, Mondragon MC, Soto ID, Parra MV, Duque C, Ortiz-Barrientos D, Garcia LF, Velez ID, Bravo ML, Munera JG, et al. 2000. Autosomal, mtDNA, and Y-chromosome diversity in amerinds: pre- and post-Columbian patterns of gene flow in South America. *Am J Hum Genet*. 67(5):1277–1286.
Milton K. 1984. Protein and carbohydrate resources of the Maku Indians of Northwestern Amazonia. *Am Anthropol*. 86(1):7–27.
Mizuno F, Gojobori J, Wang L, Onishi K, Sugiyama S, Granados J, Gomez-Trejo C, Acuna-Alonzo V, Ueda S. 2014. Complete mitogenome analysis of indigenous populations in Mexico: its relevance for the origin of Mesoamericans. *J Hum Genet*. 59(7):359–367.
Moreno-Mayar JV, Potter BA, Vinner L, Steinrucken M, Rasmussen S, Terhorst J, Kamm JA, Albrechtsen A, Malaspinas AS, Sikora M, et al. 2018. Terminal Pleistocene Alaskan genome reveals first founding population of Native Americans. *Nature* 553(7687):203–207.
Morey NC, Morey RV. 1980. Los sáliva. In: Coppens W, editor. Los aborígenes de Venezuela i. Caracas (Venezuela): La Salle. p. 241–306.
Morey NC, Morey RV, Metzger DJ. 1973. Guahibo band organization. *Antropologica* 36:83–95.
Nasidze I, Ling EY, Quinque D, Dupanloup I, Cordaux R, Rychkov S, Naumova O, Zhukova O, Sarraf-Zadegan N, Naderi GA, et al. 2004. Mitochondrial DNA and Y-chromosome variation in the Caucasus. *Ann Hum Genet*. 68(Pt 3):205–221.
Neel JV. 1970. Lessons from a "primitive" people. *Science* 170(3960):815–822.
Neel JV. 1980. On being headman. *Perspect Biol Med*. 23(2–1):277–294.
Neel JV, Weiss KM. 1975. The genetic structure of a tribal population, the yanomama indians. Xii. Biodemographic studies. *Am J Phys Anthropol*. 42(1):25–51.
Neves EG. 2008. Ecology, ceramic chronology and distribution, long-term history, and political change in the Amazonian floodplain. Handbook of South American archaeology. New York: Springer. p. 359–379.
Noelli FS. 1998. The Tupi: explaining origin and expansions in terms of archaeology and of historical linguistics. *Antiquity* 72(277):648–663.
Oliveira S, Hübner A, Fehn A-M, Aço T, Lages F, Pakendorf B, Stoneking M, Rocha J. 2018. The role of matrilineality in shaping patterns of Y chromosome and mtDNA sequence variation in southwestern Angola. *bioRxiv*. doi:https://doi.org/10.1101/349878.
Oota H, Settheetham-Ishida W, Tiwawech D, Ishida T, Stoneking M. 2001. Human mtDNA and Y-chromosome variation is correlated with matrilocal versus patrilocal residence. *Nat Genet*. 29(1):20–21.
Pakendorf B, Novgorodov IN, Osakovskij VL, Stoneking M. 2007. Mating patterns amongst Siberian reindeer herders: inferences from mtDNA and y-chromosomal analyses. *Am J Phys Anthropol*. 133(3):1013–1027.
Peña-Venegas PC, Stomph JT, Verschoor G, Lopez-Lavalle AL, Struik CP. 2014. Differences in manioc diversity among five ethnic groups of the Colombian Amazon. *Diversity* 6(4):792.
Perego UA, Achilli A, Angerhofer N, Accetturo M, Pala M, Olivieri A, Hooshiar Kashani B, Ritchie KH, Scozzari R, Kong QP, et al. 2009. Distinctive Paleo-Indian migration routes from Beringia marked by two rare mtDNA haplogroups. *Curr Biol*. 19(1):1–8.
Perego UA, Angerhofer N, Pala M, Olivieri A, Lancioni H, Hooshiar Kashani B, Carossa V, Ekins JE, Gomez-Carballa A, Huber G, et al. 2010. The initial peopling of the Americas: a growing number of founding mitochondrial genomes from Beringia. *Genome Res*. 20(9):1174–1179.
Piperno DR. 2011. The origins of plant cultivation and domestication in the new world tropics: patterns, process, and new developments. *Curr Anthropol*. 52(S4):S453–S470.
Poznik GD, Henn BM, Yee MC, Sliwerska E, Euskirchen GM, Lin AA, Snyder M, Quintana-Murci L, Kidd JM, Underhill PA, et al. 2013. Sequencing Y chromosomes resolves discrepancy in time to common ancestor of males versus females. *Science* 341(6145): 562–565.
Poznik GD, Xue Y, Mendez FL, Willems TF, Massaia A, Wilson Sayres MA, Ayub Q, McCarthy SA, Narechania A, Kashin S, et al. 2016. Punctuated bursts in human male demography inferred from 1,244 worldwide Y-chromosome sequences. *Nat Genet*. 48(6):593–599.
Raghavan M, Steinrücken M, Harris K, Schiffels S, Rasmussen S, DeGiorgio M, Albrechtsen A, Valdiosera C, Ávila-Arcos MC, Malaspinas A-S, et al. 2015. Genomic evidence for the Pleistocene and recent population history of Native Americans. *Science* 349(6250):aab3884.

Reich D, Patterson N, Campbell D, Tandon A, Mazieres S, Ray N, Parra MV, Rojas W, Duque C, Mesa N, et al. 2012. Reconstructing Native American population history. *Nature* 488(7411):370–374.

Renaud G, Stenzel U, Kelso J. 2014. Leehom: adaptor trimming and merging for illumina sequencing reads. *Nucleic Acids Res*. 42(18):e141.

Renaud G, Stenzel U, Maricic T, Wiebe V, Kelso J. 2015. Deml: robust demultiplexing of illumina sequences using a likelihood-based approach. *Bioinformatics* 31(5):770–772.

Renfrew C. 1999. Archaeology and language: the puzzle of Indo-European origins. New York: Cambridge University Press.

Rey Fajardo J. 1974. Documentos jesuiticos relativos a la historia de la Compania de Jesus en Venezuela. Caracas (Venezuela): Academia Nacional de la Historia de Venezuela.

Roewer L, Nothnagel M, Gusmao L, Gomes V, Gonzalez M, Corach D, Sala A, Alechine E, Palha T, Santos N, et al. 2013. Continent-wide decoupling of y-chromosomal genetic variation from language and geography in native South Americans. *PLoS Genet*. 9(4):e1003460.

Rojas W, Parra MV, Campo O, Caro MA, Lopera JG, Arias W, Duque C, Naranjo A, Garcia J, Vergara C, et al. 2010. Genetic make up and structure of Colombian populations by means of uniparental and biparental DNA markers. *Am J Phys Anthropol*. 143(1):13–20.

Sans M. 2000. Admixture studies in Latin America: from the 20th to the 21st century. *Hum Biol*. 72(1):155–177.

Santos-Granero F. 2002. The Arawakan matrix: ethos, language, and history in Native South America. In: Hill JD, Santos-Granero F, editors. Comparative Arawakan histories: rethinking language family and culture area in Amazonia. Urbana (IL): University of Illinois Press. p. 25–50.

Seielstad MT, Minch E, Cavalli-Sforza LL. 1998. Genetic evidence for a higher female migration rate in humans. *Nat Genet*. 20(3):278–280.

Soares P, Ermini L, Thomson N, Mormina M, Rito T, Rohl A, Salas A, Oppenheimer S, Macaulay V, Richards MB. 2009. Correcting for purifying selection: an improved human mitochondrial molecular clock. *Am J Hum Genet*. 84(6):740–759.

Sorensen AP. 1967. Multilingualism in the Northwest Amazon. *Am Anthropol*. 69(6):670–684.

Städler T, Haubold B, Merino C, Stephan W, Pfaffelhuber P. 2009. The impact of sampling schemes on the site frequency spectrum in nonequilibrium subdivided populations. *Genetics* 182(1):205–216.

Stenzel K. 2005. Multilingualism in the northwest Amazon, revisited. Memorias del Congreso de Idiomas Indígenas de Latinoamérica-II; 2005; Austin (TX). University of Texas.

Steward JH. 1949. Handbook of South American Indians. Vol. 5. Washington (DC): Smithsonian Institution.

Tamm E, Kivisild T, Reidla M, Metspalu M, Smith DG, Mulligan CJ, Bravi CM, Rickards O, Martinez-Labarga C, Khusnutdinova EK, et al. 2007. Beringian standstill and spread of Native American founders. *PLoS One* 2(9):e829.

Tang H, Siegmund DO, Shen P, Oefner PJ, Feldman MW. 2002. Frequentist estimation of coalescence times from nucleotide sequence data using a tree-based partition. *Genetics* 161(1):447–459.

Venables WN, Ripley BD. 2002. Modern applied statistics with S. New York: Springer.

Verdu P, Becker NS, Froment A, Georges M, Grugni V, Quintana-Murci L, Hombert JM, Van der Veen L, Le Bomin S, Bahuchet S, et al. 2013. Sociocultural behavior, sex-biased admixture, and effective population sizes in Central African Pygmies and non-Pygmies. *Mol Biol Evol*. 30(4):918–937.

Vidal SM. 1997. Liderazgo y confederaciones multiétnicas amerindias en la Amazonia luso-hispana del siglo xviii. *Antropologica*. 87:19–46.

Vidal SM. 2002. Secret religious cults and political leadership: multiethnic confederacies from Northwestern Amazonia. Comparative Arawakan histories. Urbana (IL): University of Illinois Press. p. 248–268.

Walker RS. 2015. Human residence patterns. In: Scott RA, Kosslyn SM, editors. Emerging trends in the social and behavioral sciences. Hoboken (NJ): John Wiley & Sons, Inc.

Walker RS, Flinn MV, Hill KR. 2010. Evolutionary history of partible paternity in lowland South America. *Proc Natl Acad Sci U S A*. 107(45):19195–19200.

Wanner H, Beer J, Butikofer J, Crowley TJ, Cubasch U, Fluckiger J, Goosse H, Grosjean M, Joos F, Kaplan JO, et al. 2008. Mid- to Late Holocene climate change: an overview. *Quatern Sci Rev*. 27(19–20): 1791–1828.

Wilder JA, Mobasher Z, Hammer MF. 2004. Genetic evidence for unequal effective population sizes of human females and males. *Mol Biol Evol*. 21(11):2047–2057.

Wilkins JF, Marlowe FW. 2006. Sex-biased migration in humans: what should we expect from genetic data? *Bioessays* 28(3):290–300.

Wilson Sayres MA, Lohmueller KE, Nielsen R. 2014. Natural selection reduced diversity on human Y chromosomes. *PLoS Genet*. 10(1):e1004064.

Zegura SL, Karafet TM, Zhivotovsky LA, Hammer MF. 2004. High-resolution SNPs and microsatellite haplotypes point to a single, recent entry of Native American Y chromosomes into the Americas. *Mol Biol Evol*. 21(1):164–175.