Genome-wide Ancestry and Demographic History of African-Descendant Maroon Communities from French Guiana and Suriname

Cesar Fortes-Lima,[1,2] Antoine Gessain,[3] Andres Ruiz-Linares,[4,5,6] Maria-Cátira Bortolini,[7] Florence Migot-Nabias,[8] Gil Bellis,[9] J. Víctor Moreno-Mayar,[10] Berta Nelly Restrepo,[11] Winston Rojas,[12] Efren Avendaño-Tamayo,[12,13] Gabriel Bedoya,[12] Ludovic Orlando,[1,10] Antonio Salas,[14,15] Agnar Helgason,[16] M. Thomas P. Gilbert,[10,17] Martin Sikora,[10] Hannes Schroeder,[10,18,19,*] and Jean-Michel Dugoujon[1,19,*]

The transatlantic slave trade was the largest forced migration in world history. However, the origins of the enslaved Africans and their admixture dynamics remain unclear. To investigate the demographic history of African-descendant Marron populations, we generated genome-wide data (4.3 million markers) from 107 individuals from three African-descendant populations in South America, as well as 124 individuals from six west African populations. Throughout the Americas, thousands of enslaved Africans managed to escape captivity and establish lasting communities, such as the Noir Marron. We find that this population has the highest proportion of African ancestry (~98%) of any African-descendant population analyzed to date, presumably because of centuries of genetic isolation. By contrast, African-descendant populations in Brazil and Colombia harbor substantially more European and Native American ancestry as a result of their complex admixture histories. Using ancestry tract-length analysis, we detect different dates for the European admixture events in the African-Colombian (1749 CE; confidence interval [CI]: 1737–1764) and African-Brazilian (1796 CE; CI: 1789–1804) populations in our dataset, consistent with the historically attested earlier influx of Africans into Colombia. Furthermore, we find evidence for sex-specific admixture patterns, resulting from predominantly European paternal gene flow. Finally, we detect strong genetic links between the African-descendant populations and specific source populations in Africa on the basis of haplotype sharing patterns. Although the Noir Marron and African-Colombians show stronger affinities with African populations from the Bight of Benin and the Gold Coast, the African-Brazilian population from Rio de Janeiro has greater genetic affinity with Bantu-speaking populations from the Bight of Biafra and west central Africa.

Introduction

Between 1526 and 1875, an estimated 7 million Africans were transported to South America (see the Transatlantic Slave Trade Database [TASTD] in the Web Resources). According to documentary sources, they embarked from different historical coastal regions in Africa, including the Senegambia, the Windward Coast, the Gold Coast, the Bights of Benin and Biafra, west central Africa, and southeast Africa (see TASTD in Web Resources)[1] Of more than 100,000 Africans who arrived in the Spanish colony of Colombia between 1549 and 1800 (Table S1), 34% originated in west Africa (the Senegambia) and 45% in west central Africa (Angola, Cameroon, Gabon, and the Congo Basin) (see TASTD in Web Resources). By contrast, in Rio de Janeiro—the largest slaving port in the Portuguese colony of Brazil and the single most important disembarkation point in the Americas—84% of captives came from Angola in west central Africa (see TASTD in Web Resources).[1] In Suriname and French Guiana, the principal sources of enslaved Africans were the Gold Coast (23%), the Bight of Benin (also referred as the Slave Coast) (16%), and west central Africa (30%) (see TASTD in Web Resources). However, the region of embarkation does not necessarily reflect the slaves' geographical origins because most captives came from far inland rather than the documented port of

[1]Laboratoire d'Anthropologie Moléculaire et Imagerie de Synthèse, AMIS UMR5288, Centre National de la Recherche Scientifique (CNRS) -Université Paul Sabatier Toulouse III, Toulouse 31000, France; [2]Laboratory Eco-Anthropology and Ethno-Biology, UMR7206, CNRS-MNHN-University Paris Diderot, Musée de l'Homme, 17 Place du Trocadéro, 75016 Paris, France; [3]Oncogenic Virus Epidemiology and Pathophysiology Group, Department of Virology, CNRS UMR3569, Pasteur Institute, Paris 75015, France; [4]Department of Genetics, Evolution, and Environment, University College London, London WC1E 6BT, United Kingdom; [5]Ministry of Education Key Laboratory of Contemporary Anthropology and Collaborative Innovation Center of Genetics and Development, Fudan University, Shanghai 200438, China; [6]Laboratory of Biocultural Anthropology, Law, Ethics, and Health, CNRS/EFS ADES UMR7268, Aix-Marseille University, Marseille 13824, France; [7]Department of Genetics, Federal University of Rio Grande do Sul, Porto Alegre 91501-970, Brazil; [8]Mother and Child Facing Tropical Infections (MERIT), Research Institute for Development, Paris 5 University, Sorbonne Paris Cité, Paris 75006, France; [9]French Institute for Demographic Studies, Paris 75020, France; [10]Natural History Museum of Denmark, University of Copenhagen, Copenhagen 1350, Denmark; [11]Instituto Colombiano de Medicina Tropical, Universidad CES, Sabaneta, Antioquia 055450, Colombia; [12]Laboratory of Molecular Genetics, Institute of Biology, University of Antioquia, Medellín 050010, Colombia; [13]Grupo de Ciencias Básicas Aplicadas del Tecnológico de Antioquia, Tecnológico de Antioquia - Institución Universitaria, Medellín 050034, Colombia; [14]Unidade de Xenética, Departamento de Anatomía Patolóxica e Ciencias Forenses, Instituto de Ciencias Forenses, Facultade de Medicina, Universidade de Santiago de Compostela, Galicia 15782, Spain; [15]GenPoB Research Group, Instituto de Investigaciones Sanitarias, Hospital Clínico Universitario de Santiago, Galicia 15782, Spain; [16]DeCode Genetics, Reykjavik IS-101, Iceland; [17]Norwegian University of Science and Technology, University Museum, Trondheim 7491, Norway; [18]Faculty of Archaeology, Leiden University, Leiden 2333, the Netherlands
[19]These authors contributed equally to this work
*Correspondence: hschroeder@snm.ku.dk (H.S.), jean-michel.dugoujon@univ-tlse3.fr (J.-M.D.)
https://doi.org/10.1016/j.ajhg.2017.09.021.

embarkation.[1,2] Therefore, it has been extremely difficult to reconstruct the geographical or ancestral origins of enslaved Africans on the basis of documentary sources alone.[2,3]

Another source of information on the history of African-descendant populations has been the use of genetics.[4–6] For example, genetic studies have revealed significant regional differences in ancestry proportions within and between different African-descendant populations in the Americas,[7,8] different timing of admixture events,[9,10] and different patterns of sex-biased gene flow.[11–13] However, the challenge with reconstructing ancestral origins has been that African-descendant populations are generally highly admixed as a result of a long history of intermixing with Europeans, indigenous peoples, and other African groups.[11,14] These admixture dynamics have been associated with different migration patterns, drift effects, and the more or less pronounced geographic isolation of these populations.[8–12]

Maroon communities are descendants of Africans who escaped from slavery and formed independent settlements of free people; these communities persist to the present day in some parts of the Americas.[15] For example, the Noir Marron population of Suriname and French Guiana is one of the largest Maroon populations in South America and one of the most culturally, politically, and economically independent of all Maroon peoples in the Americas.[15,16] They are direct descendants of enslaved Africans who escaped from plantations in the Dutch colony of Suriname during the 16th and early 17th century.[17,18] Their relative isolation since the days of the transatlantic slave trade (TAST) presents a unique opportunity to study the ancestral origins of African-descendant populations;[19,20] however, these populations are underrepresented in genome-wide studies. To this day, each of the Noir Marron (or *Bushinengués*, meaning people of the forest) communities preserves a remarkable African heritage in their cultural traditions, languages, and social organization.[15,17] However, none of these traditions can be easily associated with a unique African source population because of inaccuracies and inconsistencies in some of the historical facts and dates and the names of African groups associated with their ancestral origin.[21,22]

To shed new light on the ancestral origins and demographic history of the Noir Marron, we generated a large genome-wide dataset of more than 4.3 million SNP markers for four Noir Marron communities (Figure S1), as well as two other African-descendant populations from southeastern Brazil and Colombia and six putative west African source populations. We used the data to evaluate whether there are differences in patterns of genetic ancestry between and within the Noir Marron communities and the other two non-Maroon populations. In addition, we applied model-based clustering and haplotype-based methods to infer genetic connections with putative founder source populations in Africa. Furthermore, we employed model-based approaches to infer the likely process and timing of admixture events. Lastly, we used X chromosome data to identify differences in patterns of sex-specific gene flow. Our study provides fundamentally new insights into the population history of African-descendant Maroon communities from French Guiana and Suriname.

Material and Methods

Population Samples and Genome-wide Genotyping Techniques

The DNA samples were collected from ten different locations in South America and west Africa. The samples from South America (n = 107) included 71 individuals from four Noir Marron communities in French Guiana and Suriname (23 Aluku, 23 Ndjuka, 19 Saramaka, and six Paramaka) (Figure S1), 16 African-Brazilians from Rio de Janeiro, and 20 African-Colombians from the Antioquia and Chocó departments in Colombia. The samples from west Africa (n = 124) included 19 Fon, 24 Bariba, and 24 Yoruba from Benin, 20 Ahizi and 17 Yacouba from the Ivory Coast, and 20 Bwa from Mali. Geographic locations and further information on these populations are provided in Table S2. The study was conducted with the informed consent of all voluntary participants and the authorization of the Comité Consultatif de Protection des Personnes dans la Recherche Biomédicale Outre-Mer III (Toulouse, France), Commission Nationale de l'Informatique et des Libertés (CNIL, France), l'Agence Française de Sécurité Sanitaire des Produits de Santé (AFSSAPS, France), the Ethic Committee of Université d'Abomey-Calavi (N° 07/T4/2015/CE/FSS/UAC) from Cotonou in Benin, the Research Ethics Committee of the Universidade Federal do Rio Grande do Sul (Resolution no. 98002/1998), the Brazilian Ethics Commission (CONEP number 1333/2002), and the Bioethics Committee of the Institute of Biology of the University of Antioquia in Colombia.

A total of 4,301,332 single-nucleotide polymorphisms (SNPs) for 231 individuals were genotyped on the Illumina HumanOmni5 Quad BeadChip (Illumina). Genotyping was performed with the Illumina Infinium Assay by the SNP&SEQ Technology platform in Uppsala, Sweden.[23] The results and genotyping quality were analyzed with GenomeStudio software v2011.1 (Illumina), and DNA strand positions were built according to the Genome Reference Consortium Human genome build 37 (GRCh37/hg19). So that the reproducibility of the genotyping could be assessed, three individuals were genotyped twice on separate arrays (Table S2), resulting in 99.92% genotype concordance, on average.

Assembled Datasets and Quality Control

We merged the genotype data with different reference panels to generate two genome-wide SNP datasets with different SNP densities (Table S3). Figure S2 depicts the methodological genome-wide framework that we applied to analyze each assembled dataset. Before merging, four South American and three west African individuals were removed on the basis of their intra-population probabilities of parent-offspring or full sibling relationships; this probability was estimated with KING[24] and PC-Relate.[25] The high-density SNP dataset consisted of the newly genotyped dataset of 224 individuals, 94 individuals from May et al.[26] and 1,812 individuals from the 1000 Genomes Project phase 3;[27] the latter included 157 African descendants from the Caribbean (ACB) and North America (ASW) (Table S3). We used PLINK v1.9[28] to perform the genotyping quality control (QC) (Figure S2), resulting in 1,782,673 SNPs for 2,130 individuals after QC and 244,529 SNPs after QC with pruning for linkage

disequilibrium (LD, $r^2 = 0.1$). For the low-density SNP dataset, we increased the African reference panel by adding previously published SNP-array data from 39 sub-Saharan African populations,[29–33] resulting in 229,586 SNPs for 3,127 individuals after QC and 83,366 SNPs after QC with LD-pruning ($r^2 = 0.1$) (Figure S2).

Estimating Continental Ancestries

First, we explored the genetic relationships among the African and African-descendant populations in the QC-filtered LD-pruned high-density SNP dataset by using FlashPCA v2.[34] We then performed both global (GAI) and local ancestry (LAI) analysis to estimate the four-way continental ancestry in each African-descendant population, and we evaluated the correlation between estimates from both approaches.[35,36] For GAI, we applied the unsupervised model-based clustering algorithm in ADMIXTURE v1.3[37] at K = 4 to the LD-pruned high-density SNP dataset. For LAI, we first used SHAPEIT2[38] to estimate haplotypic phased data across the chromosomes of each individual on the basis of the LD-unpruned high-density SNP dataset, and we then used RFMix v1.5.4[39] to estimate the continental haplotypic admixture along phased chromosomal segments (ancestry tracts). To maximize phasing accuracy, we phased the dataset by using the 1000 Genomes Project phase 3 dataset as a reference panel and the HapMap phase II b37 reference as a genetic map.[38,40] We performed RFMix analysis for a four-way ancestry model by using a random-forest algorithm based on two steps of expectation maximization (EM = 2).[39] We excluded the first and the last 2 Mb from the telomeres of each chromosome because of observed inaccurate LAI in those regions.[41] We ran RFMix in PopPhased mode to estimate local ancestry likelihoods in 0.2 cM genetic windows (~105 SNPs/window) and set the "generations after admixture" parameter to 14 generations.[42] Local ancestry assignments were determined at a 0.9 posterior probability threshold of a given ancestry estimated within each window with the random-forest algorithm.[10,39] For each local ancestry estimate, we calculated the average ancestry proportions for all ancestry tracts across autosomal chromosomes in each population, as well as for only short (>5 but ≤ 50 cM) and long (>50 cM) ancestry tracts.[10] To infer the genetic isolation within African-descendant populations, we estimated genomic inbreeding coefficients (FEstim) by using FSuite v1.0.4,[43,44] and genomic runs of homozygosity (ROH) across SNPs by using PLINK.[45] For both parameters, we performed a one-way ANOVA to test whether there are significant differences among African-descendent populations.

Estimating Sub-continental Ancestries

To disentangle both the African and European sub-continental ancestries, we used different approaches to compare genetic patterns between African-descendant and reference populations (Figure S2). First, on the basis of the LD-pruned low-density SNP dataset, we ran unsupervised ADMIXTURE analysis from K = 3 through K = 22 in 20 replicates for each K. We used the Greedy algorithm implemented in CLUMPP[46] for the optimal alignment of the replicates and the cross-validation (CV) test[47] to estimate the number of K's with the lowest value. We then used Surfer v12 (Golden Software) to visualize the African geographic distribution of the cluster with the highest proportions among African descendants. Second, on the basis of the LD-unpruned low-density SNP dataset, we used haplotype-based methods for masked and unmasked haploid genomes.

For masked haploid genomes, we first estimated ancestry tracts for individual haplotypes by using RFMix analysis at EM = 1.[10] We set the minimum number of reference haplotypes per tree node to 5[39] to avoid size biases resulting from unbalanced reference panel sizes in the application of the random forest algorithm (for 2,178 African, 503 European, 50 Native American, and 103 East Asian individuals). We then generated two datasets with masked haplotypes: one dataset restricted to reference haplotypes for African ancestry and masked African-descendant haploid genomes with at least 40% African ancestry; and the other dataset restricted to European reference haplotypes and masked African-descendant haploid genomes with at least 15% European ancestry.[10] For both datasets, we used VCFTools[48] to calculate Weir and Cockerham's weighted F_{ST} (ancestry-specific WC-F_{ST}) for pairwise populations between one admixed population and one reference population.[49] For the haploid genomes, we applied the multidimensional scaling approach described by Browning et al.[50] to perform ancestry-specific principal-components analysis (MDS-based ASPCA) for within-continental ancestry inference. We performed MDS with Euclidean distances by using the "dist" and "cmdscale" functions in R;[51] in the calculations, we treated masked positions as missing data and only used positions at which both haplotypes were non-missing.

For unmasked haploid genomes, we used GERMLINE[52] to estimate the distribution of genetic segments that are identical-by-descent (IBD) between pairs of individuals.[7] Each individual was analyzed as two distinct and separate haplotypes. We applied a filtering script to eliminate false-positive IBD matches by removing IBD segments smaller than 3 cM and regions with a large excess of IBD segments (> 20 cM).[7] We then calculated the sum of the total length of IBD sharing between each pair individuals within each population and normalized these sums by the sample size of each population.[10]

Reconstructing the Demographic History and Admixture Timing

To explore the recent admixture history of the African-descendant populations in our dataset, we applied two different haplotype-based admixture methods. First, we used GLOBETROTTER v2[53] to calculate the best-guess conclusion for admixture according to the goodness-of-fit (R^2 coefficient).[53] We identified haplotype sharing patterns among target and surrogate populations by using fineSTRUCTURE and ChromoPainter v2[54] with 10 steps of EM. We performed 100 bootstrap replicates to generate confidence intervals around the estimated date of admixture (95% CI).[53] Second, we used the model-testing approach implemented in TRACTS[55] to analyze the tract-length distributions of ancestry tracts estimated with RFMix (EM = 2). We then tested two- and three-pulse migration models with a founding admixture event between the African and Native American source populations and followed this by one (or two) subsequent pulse(s) of European admixture. To identify the best-fitting model, we first fit each model with 100 starting parameter randomizations and then evaluated the strength of the maximum-likelihood estimation (MLE) of each model.[55]

Detecting Sex-Biased Admixture Patterns

To detect sex-biased admixture patterns, we compared the average of autosomal and X chromosome continental ancestry proportions estimated for each population by evaluating data for males and females together as well as separately. To phase the

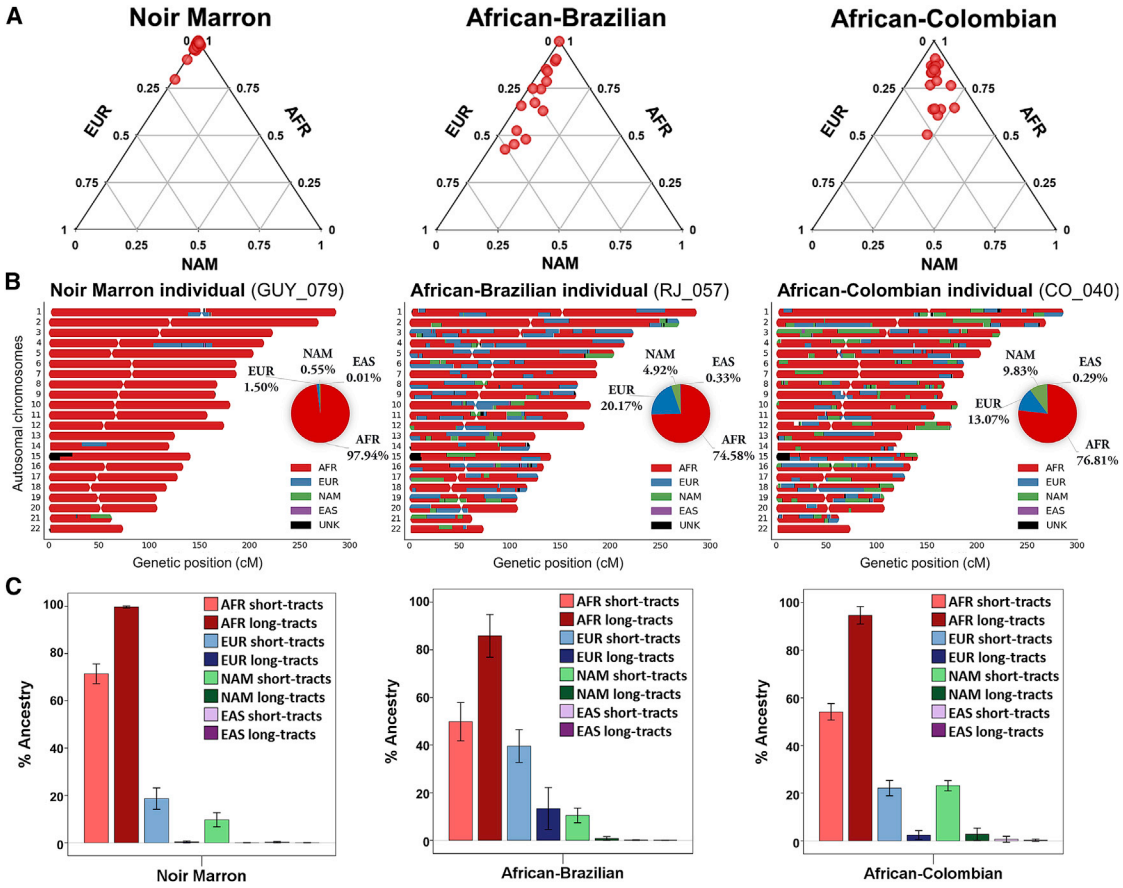

Figure 1. Genetic Diversity in the Three African-Descendant Populations from South America

(A) Ternary plots showing individual ancestry proportions in three African-descendant populations, based on ADMIXTURE analysis (AFR = African; EUR = European; and NAM = Native American). Each dot represents one individual.

(B) Local ancestry karyograms for three individuals with ancestry proportions similar to the estimated population averages. The African-Colombian and southeastern African-Brazilian haploid genomes present a complex mosaic of ancestry tracts across the genome, reflecting different demographic histories.

(C) Average ancestry proportions in short (>5 but ≤50 cM) and long (> 50 cM) ancestry tracts were estimated for each continental ancestry in African-descendant populations with RFMix (EM = 2). Bar plots were plotted with the indicated 95% confidence intervals.

X chromosome dataset corresponding to the non-pseudo-autosomal region (non-PAR), we first removed 157 haploid heterozygous SNPs and then phased the dataset by using the 1000 Genomes Project phase 3 haplotypes as a reference for the X chromosome non-PAR data. To estimate ancestry proportions for the phased X chromosome data on the basis of both GAI and LAI, we applied ADMIXTURE[37,56] and RFMix[39,50] analyses, respectively. We estimated means in the general population and for each sex separately, as well on the basis of the sex-biased admixture model.[13] To analyze differences between autosomal and X chromosome ancestry proportions in each population, we first calculated the admixture difference ratio (ΔAdmix ratio) for each continental ancestry[57] and then used the Wilcoxon signed-ranks (WSR) two-sided unpaired test to establish whether there are any statistically significant differences.[10]

Results

Genetic Diversity of African-Descendant Populations

We assessed the genetic ancestry of the Noir Marron and the two other African-descendant populations at a continental level by using both GAI and LAI approaches. Figure 1 shows the ancestry proportions estimated for each individual in the ADMIXTURE (Figure 1A and Figure S3) and RFMix (Figures 1B and 1C) analyses. In both approaches, the Noir Marron population has the highest proportion of African ancestry of all African-descendant populations in our dataset (see Table 1). The proportion of African ancestry is similar in all four Noir Marron communities (Tables S4 and S5), except for one Saramaka (90%) and one Ndjuka (79%) individual (Figure 1A). LAI further reveals that the Noir Marron have a low proportion (29%) of short European and Native American ancestry tracts, which is consistent with low-level gene flow from non-African groups after the formation of these communities (Figure 1C and Figure S4). In agreement, this population has the highest inbreeding coefficient (FEstim = 0.009, SD = ± 0.007) and total length of ROH (35,314 kb, SD = ±20,424 kb) of any of the five African-descendant populations in our dataset (Figure S5; one-way ANOVA test; p value < 0.001). Overall, these results suggest long-term genetic isolation in the

Table 1. Average Continental Ancestry Proportions Estimated in the Three African-Descendant Populations by Two Different Methods

	Global Ancestry Inference[a]				Local Ancestry Inference[b]			
	African	European	Native American	East Asian	African	European	Native American	East Asian
NM	97.4 ± 2.8	1.5 ± 2.8	0.9 ± 0.7	0.3 ± 0.3	98.2 ± 2.9	1.4 ± 2.8	0.4 ± 0.5	0.03 ± 0.1
AB	71.2 ± 17.8	23.3 ± 15.4	4.9 ± 3.5	0.6 ± 0.6	70.8 ± 17.8	24.0 ± 15.4	5.0 ± 3.5	0.2 ± 0.2
AC	75.6 ± 11.4	10.8 ± 5.9	12.6 ± 6.8	1.1 ± 1.8	76.8 ± 11.8	10.7 ± 6.2	11.9 ± 6.6	0.6 ± 1.7

We find a highly significant correlation between continental ancestry estimates obtained by both GAI and LAI approaches based on the autosomal data (Figure S7 and Table S6)—conversely, we found no correlation when we based our analysis on X-chromosomal data (Figure S8). Percents ± SD are shown. (NM = Noir Marron; AB = African-Brazilian; and AC = African-Colombian).
[a]ADMIXTURE analysis.
[b]RFMix analysis.

Noir Marron population.[43,58,59] By contrast, the African-Brazilian individuals in our dataset have a stronger European (~23%) and a smaller Native American component (~5%), whereas the African-Colombian individuals display similar proportions of Native American (~12%) and European (~11%) ancestry (Table 1). These results are mirrored in the FlashPCA2 analysis (Figure S6), reflecting the wide range of genetic variation within and between African-descendant populations in the Americas.

Sub-continental Ancestry Components in African-Descendant Populations

We applied three different methods to explore the sub-continental ancestry of African and European source populations. First, we used the clustering algorithm ADMIXTURE with a larger dataset of 51 African reference populations (Table S3). We identified 12 clusters in the CV test as the solution of K with the lowest error (Figure S9), and eight of those clusters are present in high proportions among African populations from specific regions across the continent (Figure S10). Figure 2A presents the geographical distribution of the dominant African ancestry component among African-descendant populations in Africa, revealing strong affinities with west African populations (Table S7). Figure 2B shows the average sub-continental admixture proportions in the three African-descendant populations, as well as their possible source populations across historical regions in Africa. Interestingly, the admixture profile of the Noir Marron closely matches that of present-day populations living in the historical Bight of Benin. Such populations include Fon, Bariba, and Yoruba from Benin and YRI and ESN from western Nigeria (Table S3).

We then investigated the sub-continental ancestry by using haplotype-based methods. Figure 2C shows the placement of masked African-descendant haploid genomes against a backdrop of African reference populations via MDS-based ASPCA. On the first principal component (PC1), the results first separate west African populations from southern and east African populations. They then separate populations throughout west Africa, from Gambia to Nigeria, on PC2. PC3 splits southern African populations, such as Namibia and South Africa, from eastern African populations, such as Kenya and Tanzania. In all PCs, the Noir Marron and African-Colombian individuals overlap with African populations residing in the Bight of Benin and, to a lesser extent, with Ghanaian populations (Akans, Kasem, and Nankam)[29] residing on the Gold Coast. By contrast, the African-Brazilians from Rio de Janeiro closely overlap with Bantu-speaking populations living in present-day Cameroon (Bantu sp., Semi-Bantu sp., and Nzime)[29,30] and Gabon (Nzebi)[30] in the Bight of Biafra and with present-day Angola (Ovimbundu, Kongo, and Kimbundu)[31] in west central Africa (Table S3). These results are in general agreement with historical sources[1,60] and are consistent with the lowest ancestry-specific WC-F_{ST} values detected between those pairwise populations (Table S8).

Furthermore, we evaluated the distribution of IBD segments shared between the African-descendant individuals and each African population by summing the cumulative amount of IBD shared between each pair of individuals within each population (Figure S11). We observed the highest values of summed IBD segments shared between each pair of African-descendant and African individuals who are representative of those historical regions. The Noir Marron have the highest amount of IBD sharing with populations from Benin (range 0.48–0.62), whereas the African-Colombians share most IBD segments with populations from Benin and Ghana (range 0.27–0.52), and the African-Brazilians share with populations from Angola (range 0.82–1.25) (Figure S11). Overall, the IBD results are in agreement with the MDS-based ASPCA analysis, highlighting strong genetic links between African-descendant populations in the Americas and specific historical regions in Africa.

With regard to European source populations, we find strong affinities between the European component of the African-Colombian and African-Brazilian populations in our dataset and Southern European populations in the ADMIXTURE analysis at K = 12 (Figure S10 and Table S7), as reported previously.[10,12] This result is mirrored in the MDS-based ASPCA for European ancestry tracts (Figure S12), which reveals substantial overlap between the masked haploid genomes from both admixed populations and haplotypes from Spain (IBS) and Italy (TSI).

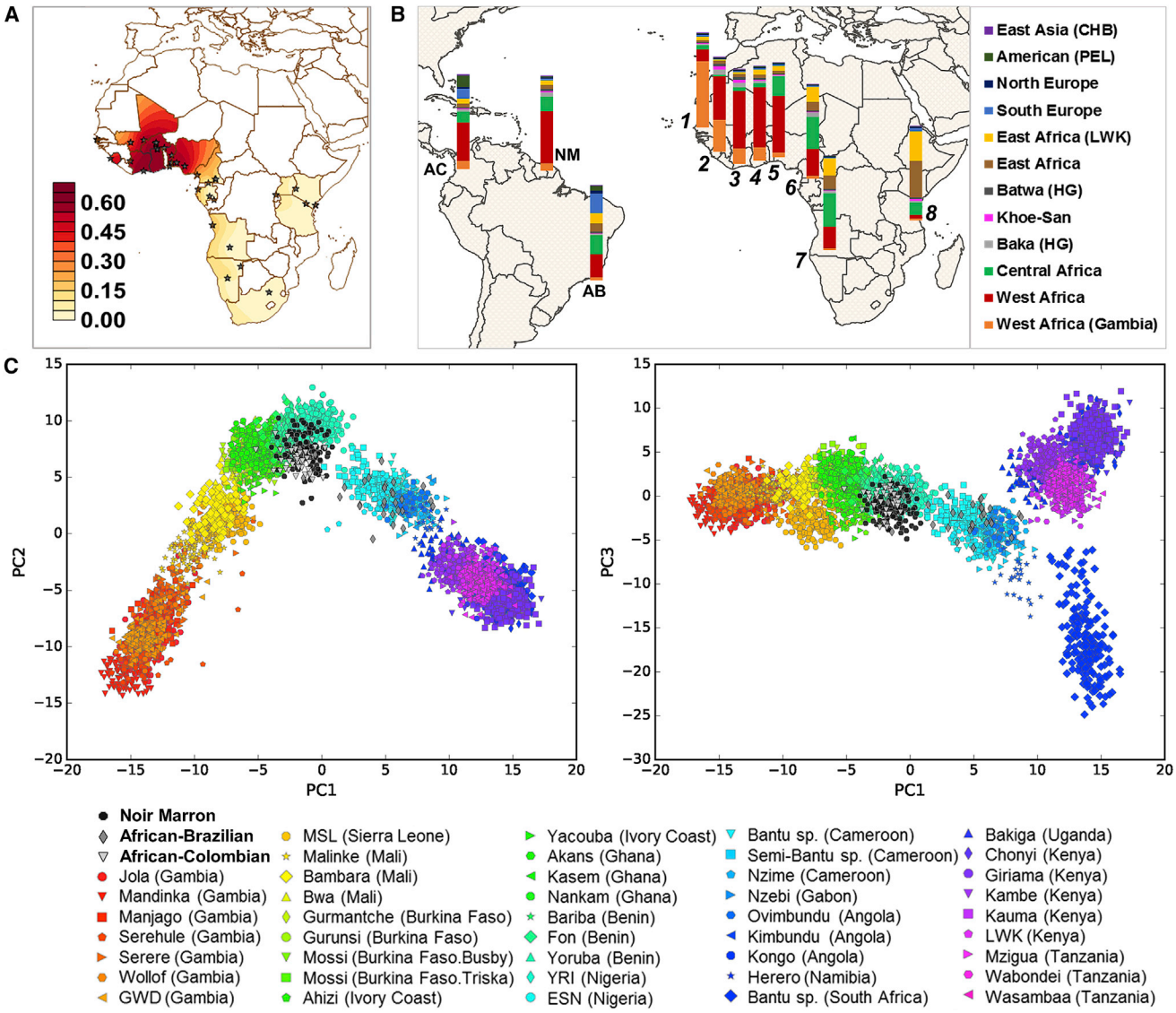

Figure 2. African Genetic Affinities of the Three African-Descendant Populations from South America

(A) Contour map showing the African geographical distribution of the predominant cluster among African-descendant populations, based on ADMIXTURE results (K = 12) (Table S7). Sampling locations for each African population are indicated by a star (Table S3).

(B) Average admixture profiles of African-descendant (NM = Noir Marron; AC = African-Colombian; and AB = southeastern African-Brazilian) populations and populations living in African historical coastal regions (1 = Senegambia; 2 = Sierra Leone; 3 = Windward Coast; 4 = Gold Coast; 5 = Bight of Benin; 6 = Bight of Biafra; 7= west central Africa; and 8 = southeast Africa). The admixture profile of the Noir Marron closely matches that of west African populations from the Bight of Benin, whereas the other two African-descendant populations show evidence of non-African gene flow.

(C) MDS-based ASPCA plot showing the placement of masked haploid genomes from African-descendant individuals against a backdrop of haploid genomes from African reference populations. Both plots reveal fine-scale genetic structure across sub-Saharan African populations and show different genetic affinities among Noir Marron, African-Colombian, and southeastern African-Brazilian individuals.

These findings are consistent with historical sources[1,3] and are also supported by ancestry-specific WC-F_{ST} values between those pairs of populations (Table S8).

Timing of Admixture Events in African-Descendant Populations

We applied two different methods to infer the demographic histories of the African-descendant populations in our dataset. We first used GLOBETROTTER analysis to determine the "best-guess" admixture scenarios on the basis of recombination distances between ancestry tracts.[53] We then tested multiple admixture models with regard to the ancestry tract-length distributions as revealed by TRACTS analysis.[10,55] The results are broadly consistent, in that both analyses identify recent admixture events associated with the period of the TAST (Table 2). From the GLOBETROTTER analysis, we infer multiple waves of admixture spanning the eighteenth century in the African-Colombian (1725–1743) and the African-Brazilian (1825–1835) individuals in our dataset (Table 2). We also

Table 2. Timing of Admixture Events according to GLOBETROTTER and TRACTS Analyses

	GLOBETROTTER			TRACTS		
	African Source	First Pulse	Second Pulse	African Source	First Pulse	Second Pulse
Noir Marron	Fon, Benin	PEL, 4.0%: 1750 CE (1725–1775)	CEU, 7.0%: 1775 CE (1750–1800)	YRI, Nigeria	PEL, 0.5%: 1783 CE (1755–1790)	CEU, 1.1%: 1791 CE (1767–1796)
African-Brazilian	Kimbundu, Angola	IBS, 25.0%: 1825 CE (1800–1850)	PEL, 6.3%: 1835 CE (1785–1864)	YRI, Nigeria	PEL, 5.8%: 1756 CE (1732–1772)	CEU, 20.6%: 1796 CE (1789–1804)
African-Colombian	Akans, Ghana	IBS, 16.0%: 1725 CE (1700–1750)	PEL, 21.0%: 1743 CE (1720–1766)	YRI, Nigeria	PEL, 11.8%: 1731 CE (1714–1744)	CEU, 9.8%: 1749 CE (1737–1764)

Table showing the best-guess conclusion for admixture in each African-descendant population, as well as ancestry proportions of each inferred source population. Admixture dates (95% CI) were estimated for a generation time of 25 years. We converted inferred generations since admixture (g) to the admixture year: 1950 – (g + 1) × 25.

detect two admixture events (at 1750 and 1775, respectively) in the Noir Marron population; however, because of its remarkably low proportions of non-African gene flow, those dates should be interpreted with caution.[53] Furthermore, we identify different European, Native American, and African ancestral source populations (Table 2). In the TRACTS analysis, a scenario with two migration pulses presents the model that best fits the observed data in all three South American populations (Figure 3A and Table 2). In the Noir Marron population, the inferred model predicts an initial pulse of Native American ancestry around 1750 (1725–1775 at 95% CI) followed by an admixture event involving Europeans around 1775 (1759–1800 at 95% CI) (Figure 3B). In the southeastern African-Brazilian population, the best-fitting model highlights one Native American pulse around 1756 (1732–1772 at 95% CI) and a European pulse around 1796 (1789–1804 at 95% CI). In the African-Colombian population, the Native American admixture was dated around 1731 (1714–1744 at 95% CI), and the European admixture was dated around 1749 (1737–1764 at 95% CI).

Sex-Specific Admixture Histories in African-Descendant Populations

We detect patterns of sex-biased gene flow in all studied African-descendant populations (African-Brazilian, African-Colombian, ACB, and ASW), except for the Noir Marron because of their low levels of admixture. When comparing ancestry proportions between X chromosome and autosomal data via both LAI and GAI, we observe a significantly higher African component in the X chromosome data than in the autosomal data of each admixed population (Figure 3C and Table S9), as well as for females and males analyzed separately (Figure S13). Conversely, European paternal gene flow was evidenced by a higher proportion of European ancestry in the autosomal data than in the X chromosome data. This is consistent with male-biased European gene flow at the time of admixture.[13] In accordance with other studies,[11,13] we find that the patterns of sex-biased gene flow differ significantly between different admixed African-descendant populations in our dataset. This is clearly shown in Figure 3D, which reveals significant differences in ΔAdmix ratios between different African-descendant populations. For example, the African-Brazilians in our dataset show higher levels of European male-biased gene flow than any other African-descendant population (Figure 3D and Table S5). Surprisingly, the African-Colombians also show significantly higher levels of Native American male-biased gene flow (Table S9; WSR test; p value 0.041 for GAI and 0.037 for LAI estimated from data from females and males together).[57,61] Overall, these results highlight the complex history of sex-biased gene flow between African-descendant and non-African populations in the Americas.

Discussion

African-descendant populations in the Americas have varying degrees of European and Native American admixture depending on their population history.[8,11] For example, we find that the African-Colombians and southeastern African-Brazilians in our dataset have on average around 29% and 24% of non-African ancestry, respectively (Table 1). Other African-descendant populations in the Americas, such as the African-Bolivians from Yungas in Bolivia, ACB from Barbados, and Jamaicans, have similar amounts of non-African ancestry, ranging from 11%–20%.[9,11,62] By contrast, the Noir Marron have relatively little non-African ancestry (less than 3% on average). In fact, despite living surrounded by European settlers and indigenous groups for more than four centuries,[17,18] the Noir Marron have the highest proportion (~98%) of genome-wide African ancestry of all studied African-descendant populations in the Americas (Table S5).[19,20] This result can be explained by the relative genetic isolation since these communities were founded. This is also reflected in the FEstim and ROH estimates, which support the long-term genetic isolation of the Noir Marron, who do not promote consanguineous marriage (Figure S5).[17,63,64]

The patterns of genetic ancestry reflect the different population histories and admixture dynamics of the three South American populations in our dataset.[14] The African-Brazilian individuals in our dataset have higher proportions of European ancestry (24%),[12,65] reflecting both their

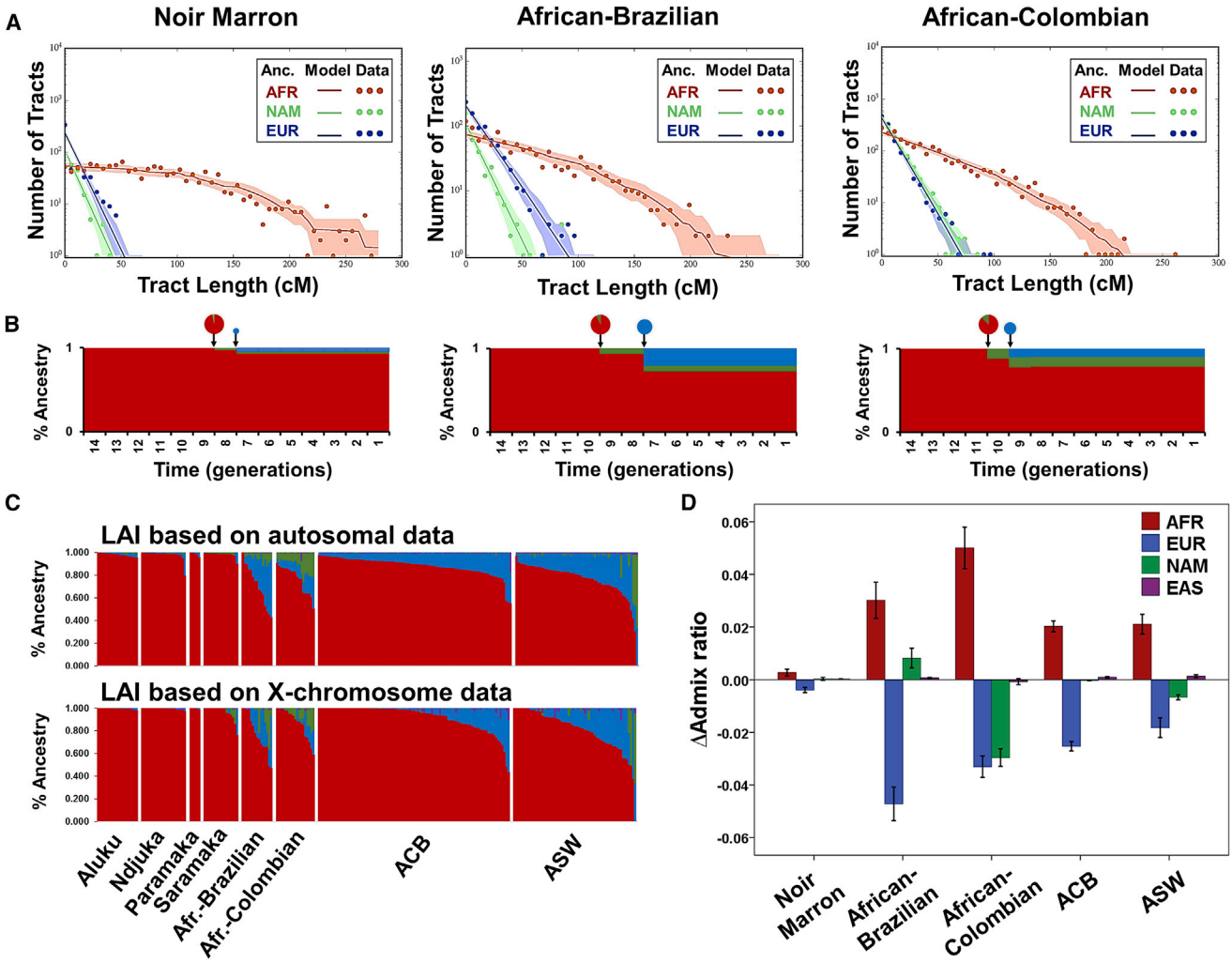

Figure 3. Timing of Admixture Events and Sex-Biased Gene Flow in the Three African-Descendant Populations from South America
(A) Length distributions of ancestry tracts within each study population; the best-fitting model, allowing for two pulse migration events, is shown. The data points represent the observed distribution of ancestry tracts estimated with RFMix (EM = 2); solid-colored lines represent the distribution from the model, and shaded areas indicate 68.3% confidence intervals of the predicted models.[55]
(B) Admixture time estimates (in generations), migration events, volume of migrants, and ancestry proportions over time for each population under the best-fitting model.[55]
(C) Bar plots showing differences in sex-specific ancestry contributions between different African-descendant populations as estimated with RFMix (EM = 2). The plots show ancestry proportions estimated for males and females together on the basis of LAI.
(D) ΔAdmix ratios showing strong African-female contribution (positive values) and European-male contribution (negative values) in admixed African-descendant populations. Bar plots were plotted with the indicated 95% confidence intervals.

history of admixture and current policies aimed at sociopolitical integration of multi-ethnic groups, whereas their Native American ancestry averages only around 5%, indicating limited gene flow from indigenous groups. By contrast, the African-Colombians in our dataset display significantly higher proportions of Native American ancestry (12%), highlighting their history of admixture with indigenous peoples.[14,66]

African Source Populations

Our findings suggest different source populations for the three African-descendant populations in our dataset. The Noir Marron, with their high proportion of African ancestry (~98%), have strong genetic affinities with African populations residing in the historical Bight of Benin (Figures 2B and 2C), highlighting this region as an important source population. Interestingly, this result is consistent with linguistic studies based on hundreds of African words that are commonly used by Noir Marron communities today; these studies reveal striking parallels and functional similarities with the eastern Gbe language sub-family that is spoken today by the Fon and other populations residing in the Bight of Benin.[67–72] Together, the genetic and linguistic data point to the Bight of Benin as the most likely source area for the founding population of the Noir Marron. African-descendants from Colombia show affinities similar to those shown by the MDS-based ASPCA and to ancestry-specific WC-F_{ST} results (Figure 2C and Table S8), highlighting the importance of the Gold Coast and the Bight of Benin as a source for enslaved

Africans during the period of the TAST. By contrast, the African descendants from southeastern Brazil show stronger affinities with Bantu-speaking populations from west central Africa, including Cameroon, Gabon, and Angola (Figure 2C and Figure S11), as suggested by the MDS-based ASPCA, ADMIXTURE results, and IBD sharing patterns (Figure 2C and Figures S10 and S11). These results are broadly consistent with historical sources (see TASTD in Web Resources),[1,60] which suggest that the clear majority (~84%) of enslaved Africans who disembarked in Rio de Janeiro between 1651 and 1856 came mainly from Luanda (Angola) and other areas in west central Africa.

Admixture Timing

Historical sources indicate that the number of imports during the TAST fluctuated widely between different European colonies and over time (Table S1) (see TASTD in Web Resources).[1] For example, the slave trade to Spanish America began much earlier than to other European colonies in South America. Thus, ca. 85% of enslaved Africans who were transported to Cartagena in Colombia during the entire period of the TAST arrived between 1549 and 1650 (Table S1), to replace the rapidly dwindling indigenous workforce.[3] By contrast, the slave trade to Brazil peaked around 200 years later, during the early 19th century (Table S1) (see TASTD in Web Resources).[1] Interestingly, these historically attested differences in slave imports are also reflected in the genetic data in the sense that the timing of inferred admixture events differs markedly between the three African-descendant populations in our dataset (although, admittedly, not as much as expected on the basis of historical sources;[1] see Table 2 and TASTD in Web Resources). For example, we infer earlier admixture dates for African-Colombians than for the African-Brazilians and the Noir Marron in our dataset (Figure 3B and Table 2); these dates reflect the earlier influx of enslaved Africans to Colombia. In the African-Brazilian population from Rio de Janeiro, the inferred admixture dates overlap with the increased number of slave imports to Brazil during the late 18th and early 19th centuries (Table S1), whereas for the Noir Marron, the inferred dates for the admixture with indigenous groups fall close to the formation of the first two Noir Marron communities (Ndjuka and Saramaka) in the 1760s.[17,18]

In addition, the GLOBETROTTER results add further support to the sub-continental ancestry results described above, in that the Fon from Benin were the best proxy for the African ancestral source population of the Noir Marron, whereas for the southeastern African-Brazilians and the African-Colombians it was the Kimbundu from Angola and the Akans from Ghana, respectively (Table 2). These findings are also broadly consistent with historically attested patterns of population movements during the TAST (see TASTD in Web Resources).[31,60]

Evidence of Sex-Biased Gene Flow

Historical data indicate that before 1820 four enslaved Africans crossed the Atlantic Ocean for every European and that, overall, more African men than women were enslaved and transported to the Americas (see TASTD in Web Resources).[1] For example, of more than 1.1 million enslaved Africans who disembarked in Rio de Janeiro between 1576 to 1856, 71% were males and 29% were females, whereas in Cartagena 67% were males and 33% females (Table S1 and TASTD in Web Resources). Notwithstanding this, we found signatures of European male-biased gene flow in all four admixed African-descendant populations in our dataset (Table S9 and Figure S13). This signal is particularly strong in the African-Brazilian population from Rio de Janeiro, which has the highest proportion of European male ancestry of all African-descendant populations in our dataset (Table S5 and Figure 3D). This result follows the same pattern as other African-descendant populations in the Americas and reflects the predominant social dynamics of colonial slave societies.[8,11] Interestingly, we also detect a higher male Native American contribution in the African-Colombian individuals in our dataset. Although this goes against some earlier studies,[8,66,73] which suggest higher female Native American contributions in African-descendant and Latino populations, it does underline the notable heterogeneity in admixture patterns among African-descendant populations.

In summary, this study presents the first genome-wide characterization of the Noir Marron from French Guiana and Suriname, who present some of the highest levels of African ancestry of any African-descendant population in the Americas. Using both genotype- and haplotype-based approaches, we were able to identify putative ancestral source populations of the Noir Marron and neighboring African-descendant populations. Although the Noir Marron and African-Colombian individuals in our dataset showed strong genetic affinities with present-day populations from the Bight of Benin and the Gold Coast, the African-Brazilians from Rio de Janeiro showed stronger affinities with present-day populations from west central Africa. These results reflect historically attested patterns of population movements during the TAST. Furthermore, we find evidence for different admixture dates and distinct patterns of sex-biased gene flow, which reflect the contrasting population histories and admixture dynamics of African-descendant populations in the Americas. Taken together, our results shed new light on the demographic consequences of the TAST and illustrate how genetic data can complement historical sources in assessments of historical population movements.

Accession Numbers

The European Genome-phenome Archive (EGA) accession number for the genome-wide SNP data of 107 African-descendant and 124 west African individuals on whom we report in this paper is EGAD00010001283.

Supplemental Data

Supplemental Data include 13 Supplemental Figures, nine Supplemental Tables, and Supplemental References.

Acknowledgments

The authors would like to acknowledge the volunteers who donated DNA samples and whose genetic data are reported in this study for the first time. We thank Bernard Carme, Mathieu Nacher, Gabriel Carles, Nicolas Pouillot, and Nicolle Joly for the new sampling performed in the French Guiana; Loic Niel for the logistical organization at "Etablissement Français du Sang" (EFS) in French Guiana; and Jacques Chiaroni and Stéphane Mazieres for the DNA extraction. We thank all the members of EUROTAST network for their helpful comments and feedback. We thank Nicolas Brucato, Hákon Jónsson, Petr Triska, Alicia Martin, Simon Gravel, Soheil Baharian, Nicolas Valdeyron, Francis Roubinet, Esteban Parra, and Paul Verdu for helpful discussions on the analyses. We thank Francis Dupuy for his comments regarding to the Noir Marron communities' histories. We thank Garrett Hellenthal for providing us with GLOBETROTTER v2. We are grateful to the Genotoul bioinformatics platform Toulouse Midi-Pyrénées (Bioinfo Genotoul) for providing help and computing resources. Genotyping was performed at the SNP&SEQ Technology Platform in Uppsala, Sweden (www.genotyping.se). The research leading to these results received funding from the European Union through the Marie Curie Actions (FP7/2007-2013, grant no. 290344, EUROTAST). C.F.-L. was supported by the EUROTAST Marie Curie Initial Training Network (FP7/2007-2013, grant no. 290344). H.S. was funded by an ERC Synergy Grant (grant no. 319209, NEXUS1492) and a grant from HERA and the European Union (grant no. 649307, CITIGEN). This work is dedicated to the memory of Professor Georges Larrouy.

Received: May 3, 2017
Accepted: September 22, 2017
Published: November 2, 2017

Web Resources

1000 Genomes Project Phase 3, ftp://ftp.1000genomes.ebi.ac.uk/vol1/ftp//release/20130502/
Ancestry Pipeline, https://github.com/armartin/ancestry_pipeline/
European Genome-phenome Archive, https://ega-archive.org/studies/EGAS00001002535
EUROTAST Genotype Data, http://eurotast.eu/eurotast-genotype-data/
fineSTRUCTURE/ChromoPainter v2 & GLOBETROTTER v2, https://people.maths.bris.ac.uk/~madjl/finestructure/finestructure.html
Local Ancestry Inference Pipeline, http://faculty.washington.edu/sguy/local_ancestry_pipeline/
TRACTS, https://github.com/sgravel/tracts
Trans-Atlantic Slave Trade Database, http://www.slavevoyages.org
RFMix, https://sites.google.com/site/rfmixlocalancestryinference/

References

1. Eltis, D., and Richardson, D. (2013). Routes to Slavery: Direction, Ethnicity and Mortality in the Transatlantic Slave Trade (Routledge).
2. Morgan, P.D. (1997). The cultural implications of the Atlantic slave trade: African regional origins, American destinations and new world developments. Slavery Abol. *18*, 122–145.
3. Klein, H.S., and Vinson, B., III. (2007). African slavery in Latin America and the Caribbean (Oxford University Press).
4. Tishkoff, S.A., Reed, F.A., Friedlaender, F.R., Ehret, C., Ranciaro, A., Froment, A., Hirbo, J.B., Awomoyi, A.A., Bodo, J.-M., Doumbo, O., et al. (2009). The genetic structure and history of Africans and African Americans. Science *324*, 1035–1044.
5. Bryc, K., Auton, A., Nelson, M.R., Oksenberg, J.R., Hauser, S.L., Williams, S., Froment, A., Bodo, J.M., Wambebe, C., Tishkoff, S.A., and Bustamante, C.D. (2010). Genome-wide patterns of population structure and admixture in West Africans and African Americans. Proc. Natl. Acad. Sci. USA *107*, 786–791.
6. Schroeder, H., Ávila-Arcos, M.C., Malaspinas, A.-S., Poznik, G.D., Sandoval-Velasco, M., Carpenter, M.L., Moreno-Mayar, J.V., Sikora, M., Johnson, P.L.F., Allentoft, M.E., et al. (2015). Genome-wide ancestry of 17th-century enslaved Africans from the Caribbean. Proc. Natl. Acad. Sci. USA *112*, 3669–3673.
7. Baharian, S., Barakatt, M., Gignoux, C.R., Shringarpure, S., Errington, J., Blot, W.J., Bustamante, C.D., Kenny, E.E., Williams, S.M., Aldrich, M.C., and Gravel, S. (2016). The Great Migration and African-American Genomic Diversity. PLoS Genet. *12*, e1006059.
8. Bryc, K., Durand, E.Y., Macpherson, J.M., Reich, D., and Mountain, J.L. (2015). The genetic ancestry of African Americans, Latinos, and European Americans across the United States. Am. J. Hum. Genet. *96*, 37–53.
9. Martin, A.R., Gignoux, C.R., Walters, R.K., Wojcik, G.L., Neale, B.M., Gravel, S., Daly, M.J., Bustamante, C.D., and Kenny, E.E. (2017). Human Demographic History Impacts Genetic Risk Prediction across Diverse Populations. Am. J. Hum. Genet. *100*, 635–649.
10. Moreno-Estrada, A., Gravel, S., Zakharia, F., McCauley, J.L., Byrnes, J.K., Gignoux, C.R., Ortiz-Tello, P.A., Martínez, R.J., Hedges, D.J., Morris, R.W., et al. (2013). Reconstructing the population genetic history of the Caribbean. PLoS Genet. *9*, e1003925.
11. Mathias, R.A., Taub, M.A., Gignoux, C.R., Fu, W., Musharoff, S., O'Connor, T.D., Vergara, C., Torgerson, D.G., Pino-Yanes, M., Shringarpure, S.S., et al.; CAAPA (2016). A continuum of admixture in the Western Hemisphere revealed by the African Diaspora genome. Nat. Commun. *7*, 12522.
12. Kehdy, F.S.G., Gouveia, M.H., Machado, M., Magalhães, W.C.S., Horimoto, A.R., Horta, B.L., Moreira, R.G., Leal, T.P., Scliar, M.O., Soares-Souza, G.B., et al.; Brazilian EPIGEN Project Consortium (2015). Origin and dynamics of admixture in Brazilians and its effect on the pattern of deleterious mutations. Proc. Natl. Acad. Sci. USA *112*, 8696–8701.
13. Goldberg, A., and Rosenberg, N.A. (2015). Beyond 2/3 and 1/3: The Complex Signatures of Sex-Biased Admixture on the X Chromosome. Genetics *201*, 263–279.
14. Ruiz-Linares, A., Adhikari, K., Acuña-Alonzo, V., Quinto-Sanchez, M., Jaramillo, C., Arias, W., Fuentes, M., Pizarro, M., Everardo, P., de Avila, F., et al. (2014). Admixture in Latin America: geographic structure, phenotypic diversity and self-perception of ancestry based on 7,342 individuals. PLoS Genet. *10*, e1004572.
15. Price, R. (2013). Maroon Societies: Rebel Slave Communities in the America (Knopf Doubleday Publishing Group).
16. Price, R. (2002). Maroons in Suriname and Guyane: How many and where. New West Indian Guide *76*, 81–88.
17. Price, R., and Price, S. (2003). Les Marrons (Vents d'Ailleurs).
18. Price, R. (2013). Les Premiers Temps. La conception de l'histoire des Marrons Saamaka (Vents d'Ailleurs).
19. Brucato, N., Tortevoye, P., Plancoulaine, S., Guitard, E., Sanchez-Mazas, A., Larrouy, G., Gessain, A., and Dugoujon,

J.-M. (2009). The genetic diversity of three peculiar populations descending from the slave trade: Gm study of Noir Marron from French Guiana. C. R. Biol. *332*, 917–926.
20. Brucato, N., Cassar, O., Tonasso, L., Tortevoye, P., Migot-Nabias, F., Plancoulaine, S., Guitard, E., Larrouy, G., Gessain, A., and Dugoujon, J.-M. (2010). The imprint of the Slave Trade in an African American population: mitochondrial DNA, Y chromosome and HTLV-1 analysis in the Noir Marron of French Guiana. BMC Evol. Biol. *10*, 314.
21. Bilby, K.M., and Bilby, K.M. (2001). Aleke: New music and new identities in the Guianas. Latin American Music Review *22*, 31–47.
22. Price, R. (2010). Travels with Tooy: History, memory, and the African American imagination (University of Chicago Press).
23. Gunderson, K.L., Steemers, F.J., Lee, G., Mendoza, L.G., and Chee, M.S. (2005). A genome-wide scalable SNP genotyping assay using microarray technology. Nat. Genet. *37*, 549–554.
24. Manichaikul, A., Mychaleckyj, J.C., Rich, S.S., Daly, K., Sale, M., and Chen, W.-M. (2010). Robust relationship inference in genome-wide association studies. Bioinformatics *26*, 2867–2873.
25. Conomos, M.P., Reiner, A.P., Weir, B.S., and Thornton, T.A. (2016). Model-free Estimation of Recent Genetic Relatedness. Am. J. Hum. Genet. *98*, 127–148.
26. May, A., Hazelhurst, S., Li, Y., Norris, S.A., Govind, N., Tikly, M., Hon, C., Johnson, K.J., Hartmann, N., Staedtler, F., and Ramsay, M. (2013). Genetic diversity in black South Africans from Soweto. BMC Genomics *14*, 644.
27. 1000 Genomes Project Consortium, Auton, A., Brooks, L.D., Durbin, R.M., Garrison, E.P., Kang, H.M., Korbel, J.O., Marchini, J.L., McCarthy, S., McVean, G.A., and Abecasis, G.R. (2015). A global reference for human genetic variation. Nature *526*, 68–74.
28. Chang, C.C., Chow, C.C., Tellier, L.C., Vattikuti, S., Purcell, S.M., and Lee, J.J. (2015). Second-generation PLINK: rising to the challenge of larger and richer datasets. Gigascience *4*, 7.
29. Busby, G.B., Band, G., Si Le, Q., Jallow, M., Bougama, E., Mangano, V.D., Amenga-Etego, L.N., Enimil, A., Apinjoh, T., Ndila, C.M., et al.; Malaria Genomic Epidemiology Network (2016). Admixture into and within sub-Saharan Africa. eLife *5*, e15266.
30. Patin, E., Siddle, K.J., Laval, G., Quach, H., Harmant, C., Becker, N., Froment, A., Régnault, B., Lemée, L., Gravel, S., et al. (2014). The impact of agricultural emergence on the genetic history of African rainforest hunter-gatherers and agriculturalists. Nat. Commun. *5*, 3163.
31. Patin, E., Lopez, M., Grollemund, R., Verdu, P., Harmant, C., Quach, H., Laval, G., Perry, G.H., Barreiro, L.B., Froment, A., et al. (2017). Dispersals and genetic adaptation of Bantu-speaking populations in Africa and North America. Science *356*, 543–546.
32. Triska, P., Soares, P., Patin, E., Fernandes, V., Cerny, V., and Pereira, L. (2015). Extensive Admixture and Selective Pressure Across the Sahel Belt. Genome Biol. Evol. *7*, 3484–3495.
33. Schlebusch, C.M., Skoglund, P., Sjödin, P., Gattepaille, L.M., Hernandez, D., Jay, F., Li, S., De Jongh, M., Singleton, A., Blum, M.G.B., et al. (2012). Genomic variation in seven Khoe-San groups reveals adaptation and complex African history. Science *338*, 374–379.
34. Abraham, G., Qiu, Y., and Inouye, M. (2017). FlashPCA2: principal component analysis of Biobank-scale genotype datasets. Bioinformatics *33*, 2776–2778.
35. Kidd, J.M., Gravel, S., Byrnes, J., Moreno-Estrada, A., Musharoff, S., Bryc, K., Degenhardt, J.D., Brisbin, A., Sheth, V., Chen, R., et al. (2012). Population genetic inference from personal genome data: impact of ancestry and admixture on human genomic variation. Am. J. Hum. Genet. *91*, 660–671.
36. Sankararaman, S., Sridhar, S., Kimmel, G., and Halperin, E. (2008). Estimating local ancestry in admixed populations. Am. J. Hum. Genet. *82*, 290–303.
37. Alexander, D.H., Novembre, J., and Lange, K. (2009). Fast model-based estimation of ancestry in unrelated individuals. Genome Res. *19*, 1655–1664.
38. Delaneau, O., Howie, B., Cox, A.J., Zagury, J.-F., and Marchini, J. (2013). Haplotype estimation using sequencing reads. Am. J. Hum. Genet. *93*, 687–696.
39. Maples, B.K., Gravel, S., Kenny, E.E., and Bustamante, C.D. (2013). RFMix: a discriminative modeling approach for rapid and robust local-ancestry inference. Am. J. Hum. Genet. *93*, 278–288.
40. Delaneau, O., Marchini, J.; 1000 Genomes Project Consortium; and 1000 Genomes Project Consortium (2014). Integrating sequence and array data to create an improved 1000 Genomes Project haplotype reference panel. Nat. Commun. *5*, 3934.
41. Bhatia, G., Tandon, A., Patterson, N., Aldrich, M.C., Ambrosone, C.B., Amos, C., Bandera, E.V., Berndt, S.I., Bernstein, L., Blot, W.J., et al. (2014). Genome-wide scan of 29,141 African Americans finds no evidence of directional selection since admixture. Am. J. Hum. Genet. *95*, 437–444.
42. Jin, W., Wang, S., Wang, H., Jin, L., and Xu, S. (2012). Exploring population admixture dynamics via empirical and simulated genome-wide distribution of ancestral chromosomal segments. Am. J. Hum. Genet. *91*, 849–862.
43. Leutenegger, A.-L., Prum, B., Génin, E., Verny, C., Lemainque, A., Clerget-Darpoux, F., and Thompson, E.A. (2003). Estimation of the inbreeding coefficient through use of genomic data. Am. J. Hum. Genet. *73*, 516–523.
44. Gazal, S., Sahbatou, M., Babron, M.-C., Génin, E., and Leutenegger, A.-L. (2014). FSuite: exploiting inbreeding in dense SNP chip and exome data. Bioinformatics *30*, 1940–1941.
45. Cassidy, L.M., Martiniano, R., Murphy, E.M., Teasdale, M.D., Mallory, J., Hartwell, B., and Bradley, D.G. (2016). Neolithic and Bronze Age migration to Ireland and establishment of the insular Atlantic genome. Proc. Natl. Acad. Sci. USA *113*, 368–373.
46. Jakobsson, M., and Rosenberg, N.A. (2007). CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. Bioinformatics *23*, 1801–1806.
47. Alexander, D.H., and Lange, K. (2011). Enhancements to the ADMIXTURE algorithm for individual ancestry estimation. BMC Bioinformatics *12*, 246.
48. Danecek, P., Auton, A., Abecasis, G., Albers, C.A., Banks, E., DePristo, M.A., Handsaker, R.E., Lunter, G., Marth, G.T., Sherry, S.T., et al.; 1000 Genomes Project Analysis Group (2011). The variant call format and VCFtools. Bioinformatics *27*, 2156–2158.
49. Weir, B.S., and Cockerham, C.C. (1984). Estimating F-Statistics for the Analysis of Population Structure. Evolution *38*, 1358–1370.
50. Browning, S.R., Grinde, K., Plantinga, A., Gogarten, S.M., Stilp, A.M., Kaplan, R.C., Avilés-Santa, M.L., Browning, B.L., and Laurie, C.C. (2016). Local Ancestry Inference in a Large

US-Based Hispanic/Latino Study: Hispanic Community Health Study/Study of Latinos (HCHS/SOL). G3 (Bethesda) *6*, 1525–1534.
51. R Core Team (2014). R: A language and environment for statistical computing (Version 3.0. 2) (Vienna, Austria: R Foundation for Statistical Computing).
52. Gusev, A., Lowe, J.K., Stoffel, M., Daly, M.J., Altshuler, D., Breslow, J.L., Friedman, J.M., and Pe'er, I. (2009). Whole population, genome-wide mapping of hidden relatedness. Genome Res. *19*, 318–326.
53. Hellenthal, G., Busby, G.B.J., Band, G., Wilson, J.F., Capelli, C., Falush, D., and Myers, S. (2014). A genetic atlas of human admixture history. Science *343*, 747–751.
54. Lawson, D.J., Hellenthal, G., Myers, S., and Falush, D. (2012). Inference of population structure using dense haplotype data. PLoS Genet. *8*, e1002453.
55. Gravel, S. (2012). Population genetics models of local ancestry. Genetics *191*, 607–619.
56. Shringarpure, S.S., Bustamante, C.D., Lange, K., and Alexander, D.H. (2016). Efficient analysis of large datasets and sex bias with ADMIXTURE. BMC Bioinformatics *17*, 218.
57. Rishishwar, L., Conley, A.B., Wigington, C.H., Wang, L., Valderrama-Aguirre, A., and Jordan, I.K. (2015). Ancestry, admixture and fitness in Colombian genomes. Sci. Rep. *5*, 12376.
58. Pemberton, T.J., Absher, D., Feldman, M.W., Myers, R.M., Rosenberg, N.A., and Li, J.Z. (2012). Genomic patterns of homozygosity in worldwide human populations. Am. J. Hum. Genet. *91*, 275–292.
59. Pemberton, T.J., and Rosenberg, N.A. (2014). Population-genetic influences on genomic estimates of the inbreeding coefficient: a global perspective. Hum. Hered. *77*, 37–48.
60. Gomes, F. (2012). The Atlantic demographics of Africans in Rio de Janeiro in the seventeenth, eighteenth, and nineteenth centuries: some patterns based on parish registers. Hist. Cienc. Saude Manguinhos *19* (*Suppl 1*), 81–106.
61. Bedoya, G., Montoya, P., García, J., Soto, I., Bourgeois, S., Carvajal, L., Labuda, D., Alvarez, V., Ospina, J., Hedrick, P.W., and Ruiz-Linares, A. (2006). Admixture dynamics in Hispanics: a shift in the nuclear genetic ancestry of a South American population isolate. Proc. Natl. Acad. Sci. USA *103*, 7234–7239.
62. Pardo-Seco, J., Heinz, T., Taboada-Echalar, P., Martinón-Torres, F., and Salas, A. (2016). Mapping the genomic mosaic of two 'Afro-Bolivians' from the isolated Yungas valleys. BMC Genomics *17*, 207.
63. Kirin, M., McQuillan, R., Franklin, C.S., Campbell, H., McKeigue, P.M., and Wilson, J.F. (2010). Genomic runs of homozygosity record population history and consanguinity. PLoS ONE *5*, e13996.
64. Szpiech, Z.A., Xu, J., Pemberton, T.J., Peng, W., Zöllner, S., Rosenberg, N.A., and Li, J.Z. (2013). Long runs of homozygosity are enriched for deleterious variation. Am. J. Hum. Genet. *93*, 90–102.
65. Manta, F.S.N., Pereira, R., Caiafa, A., Silva, D.A., Gusmão, L., and Carvalho, E.F. (2013). Analysis of genetic ancestry in the admixed Brazilian population from Rio de Janeiro using 46 autosomal ancestry-informative indel markers. Ann. Hum. Biol. *40*, 94–98.
66. Salas, A., Acosta, A., Alvarez-Iglesias, V., Cerezo, M., Phillips, C., Lareu, M.V., and Carracedo, A. (2008). The mtDNA ancestry of admixed Colombian populations. Am. J. Hum. Biol. *20*, 584–591.
67. Essegbey, J., James, E., van den Berg, M., and van de Vate, M. (2013). Possibility and necessity modals in Gbe and Surinamese creoles. Lingua *129*, 67–95.
68. Essegbey, J., James, E., Bettina, M., and Donald, W. (2013). Cross-linguistic influence in language creation: Assessing the role of the Gbe languages in the formation of the Creoles of Suriname. Lingua *129*, 1–8.
69. Migge, B., Bettina, M., and Donald, W. (2013). Fact-type complements in Gbe and the Surinamese Creoles. Lingua *129*, 9–31.
70. Huttar, G.L., Aboh, E.O., and Ameka, F.K. (2013). Relative clauses in Suriname creoles and Gbe languages. Lingua *129*, 96–123.
71. Muysken, P.C., and Smith, N. (2015). Surviving the Middle Passage: The West Africa-Surinam Sprachbund (Walter de Gruyter GmbH & Co KG).
72. Daval-Markussen, A., and Bakker, P. (2011). A phylogenetic networks approach to the classification of English-based Atlantic creoles. English World-Wide *32*, 115–146.
73. Bryc, K., Velez, C., Karafet, T., Moreno-Estrada, A., Reynolds, A., Auton, A., Hammer, M., Bustamante, C.D., and Ostrer, H. (2010). Colloquium paper: genome-wide patterns of population structure and admixture among Hispanic/Latino populations. Proc. Natl. Acad. Sci. USA *107* (*Suppl 2*), 8954–8961.