

# Genetic Consequences of the Transatlantic Slave Trade in the Americas

Steven J. Micheletti,<sup>1,\*</sup> Kasia Bryc,<sup>1</sup> Samantha G. Ancona Esselmann,<sup>1</sup> William A. Freyman,<sup>1</sup> Meghan E. Moreno,<sup>1</sup> G. David Poznik,<sup>1</sup> Anjali J. Shastri,<sup>1</sup> 23andMe Research Team,<sup>1</sup> Sandra Beleza,<sup>2</sup> and Joanna L. Mountain<sup>1</sup>

According to historical records of transatlantic slavery, traders forcibly deported an estimated 12.5 million people from ports along the Atlantic coastline of Africa between the 16th and 19th centuries, with global impacts reaching to the present day, more than a century and a half after slavery's abolition. Such records have fueled a broad understanding of the forced migration from Africa to the Americas yet remain underexplored in concert with genetic data. Here, we analyzed genotype array data from 50,281 research participants, which—combined with historical shipping documents—illustrate that the current genetic landscape of the Americas is largely concordant with expectations derived from documentation of slave voyages. For instance, genetic connections between people in slave trading regions of Africa and disembarkation regions of the Americas generally mirror the proportion of individuals forcibly moved between those regions. While some discordances can be explained by additional records of deportations *within* the Americas, other discordances yield insights into variable survival rates and timing of arrival of enslaved people from specific regions of Africa. Furthermore, the greater contribution of African women to the gene pool compared to African men varies across the Americas, consistent with literature documenting regional differences in slavery practices. This investigation of the transatlantic slave trade, which is broad in scope in terms of both datasets and analyses, establishes genetic links between individuals in the Americas and populations across Atlantic Africa, yielding a more comprehensive understanding of the African roots of peoples of the Americas.

## Introduction

The forced displacement of more than 12.5 million men, women, and children from Africa to the Americas between 1515 and 1865 has had significant social, cultural, health, and genetic impacts across the Americas.<sup>1–6</sup> The consequences of this displacement are myriad and complex, reflecting the varied economic goals and practices of the nations outfitting the ships that transported captive Africans.<sup>7,8</sup> The current narrative of this era of forced deportation has been compiled from thousands of historical shipping documents, as well as diaries, personal letters, sketches, and records of slave sales.<sup>9</sup> These data suggest that the number of enslaved people disembarking at each port in the Americas, the nation that shipped enslaved people to a given port, and the populations and geographic locations in Africa from which captives were taken differs from region to region within the Americas. For example, shipping manifests indicate that the number of enslaved people disembarking at Brazilian and Caribbean ports was far greater than the number disembarking at other ports in the Americas.<sup>5</sup> An estimated 10.1 million enslaved people, primarily male, disembarked in Central America, South America, and the Caribbean (accounting for more than 90% of all captives who were brought to the Americas), with fewer than half a million disembarking in mainland North America<sup>5</sup> (3% to 5% of the total). Concurrent with transatlantic voyages, slave traders also transferred nearly 500,000 enslaved people throughout the Americas with most documented intra-American voyages origi-

nating in the Caribbean.<sup>5,10</sup> Despite a majority of enslaved Africans disembarking into Latin America, estimates of the proportion of African ancestry for Latin Americans with African roots are lower on average than estimates for African Americans in the United States.<sup>11–15</sup> Further, patterns of embarkation of enslaved populations varied over time, whereby people primarily from 48 distinct ethnolinguistic groups were targeted across Atlantic Africa.<sup>16</sup> Prior to 1650, slave voyages most often originated in Senegambia and West Central Africa (in locations now known as the Democratic Republic of Congo and Angola). Although West Central Africa remained a primary embarkation region throughout the slave trading period (from 1650 to 1850), these voyages originated with increasing frequency at ports in the Bight of Benin and the Bight of Biafra, in locations now known as Benin, Nigeria, Cameroon, and Gabon.<sup>9</sup>

While records have shed light on the main trends of the transatlantic slave trade, the effects of under-documented practices, such as illegal slave trading and details of events after disembarkation in the Americas, remain less understood.<sup>17–20</sup> If under-documented events did not shift the overall paradigm of the slave trade and reproduction rates were equivalent among enslaved populations in the Americas, the degree of genetic similarity between the Americas and each slave-trading region of Africa would be expected to be correlated with the recorded number of enslaved people that disembarked into each region. Also, little is known about the extent to which enslaved people and their descendants continued to associate exclusively with those

<sup>1</sup>23andMe, Inc., Sunnyvale, CA 94043, USA; <sup>2</sup>University of Leicester, Leicester LE1 7RH, UK

\*Correspondence: [smicheletti@ucdavis.edu](mailto:smicheletti@ucdavis.edu)

<https://doi.org/10.1016/j.ajhg.2020.06.012>

© 2020 The Authors. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).



of similar ethnolinguistic origins after generations in the Americas.<sup>21</sup> If such association was common, Americans with African roots would be expected to have African ancestry from a single region of Africa in a pattern that is correlated with geography. Finally, previous studies indicate that individuals with mixed ancestry have higher levels of African ancestry in genomic regions inherited through the female line.<sup>22–24</sup> However, the extent of this sex bias in gene pool contributions may vary across the Americas due to regional differences in mortality of enslaved men,<sup>25</sup> rape of enslaved African women,<sup>26</sup> forced segregation,<sup>21</sup> and practices of reducing African representation by promoting reproduction with Europeans (racial whitening).<sup>27,28</sup>

Previous genetic studies focusing on people of African descent across the Americas found that most African Americans in the United States have more African ancestry from populations that lived near present-day Nigeria than from populations that lived elsewhere in Atlantic Africa.<sup>11,15</sup> However, multiple previous studies have lacked representation of West Central Africa, the largest source of enslaved people during the transatlantic slave trade, leading to an incomplete understanding of the dynamics of population interactions during and after the slave trade era. Here, we combined genetic data from individuals representing the major slave trading regions of Atlantic Africa, Europe, and the Americas and explored concordance with centuries of historical documents. We determined ancestral connections using identity by descent, local ancestry inference, and Bayesian dating techniques. In addition, we used ancestral admixture models to estimate the bias in gene pool contributions of African, European, and indigenous American men and women to current populations of the Americas. We used data from this large, diverse cohort of individuals to obtain unprecedented insights into the genetic impact of the transatlantic slave trade.

## Subjects and Methods

### Human Subjects

We included data from participants drawn from the millions of research participants of 23andMe, Inc., a consumer personal genetics company,<sup>29–31</sup> the 1000 Genomes Project,<sup>32</sup> the Human Genome Diversity Project,<sup>33</sup> Angola, and previous studies of the Democratic Republic of the Congo,<sup>34</sup> Sierra Leone,<sup>35</sup> and Khoe-San speaking people.<sup>36,37</sup> Angolan participant saliva samples were collected with research consent, which was approved by both the ethics committee of the University 11th of November (Universidade 11 de Novembro), Cabinda, Angola (REF: GD-FM/UoN/2016) and the University of Leicester ethics committee (REF: 11334-sdsb1-genetics). 23andMe research participants provided informed consent and participated in the research online, under a protocol approved by the external AAHRPP-accredited IRB, Ethical & Independent Review Services (E&I Review). 23andMe participants were given the option to fill out web-based questionnaires, including questions on ancestry, ethnicity, and grandparent birth locations. We included in our study data only of customers who signed IRB-approved consent documents,

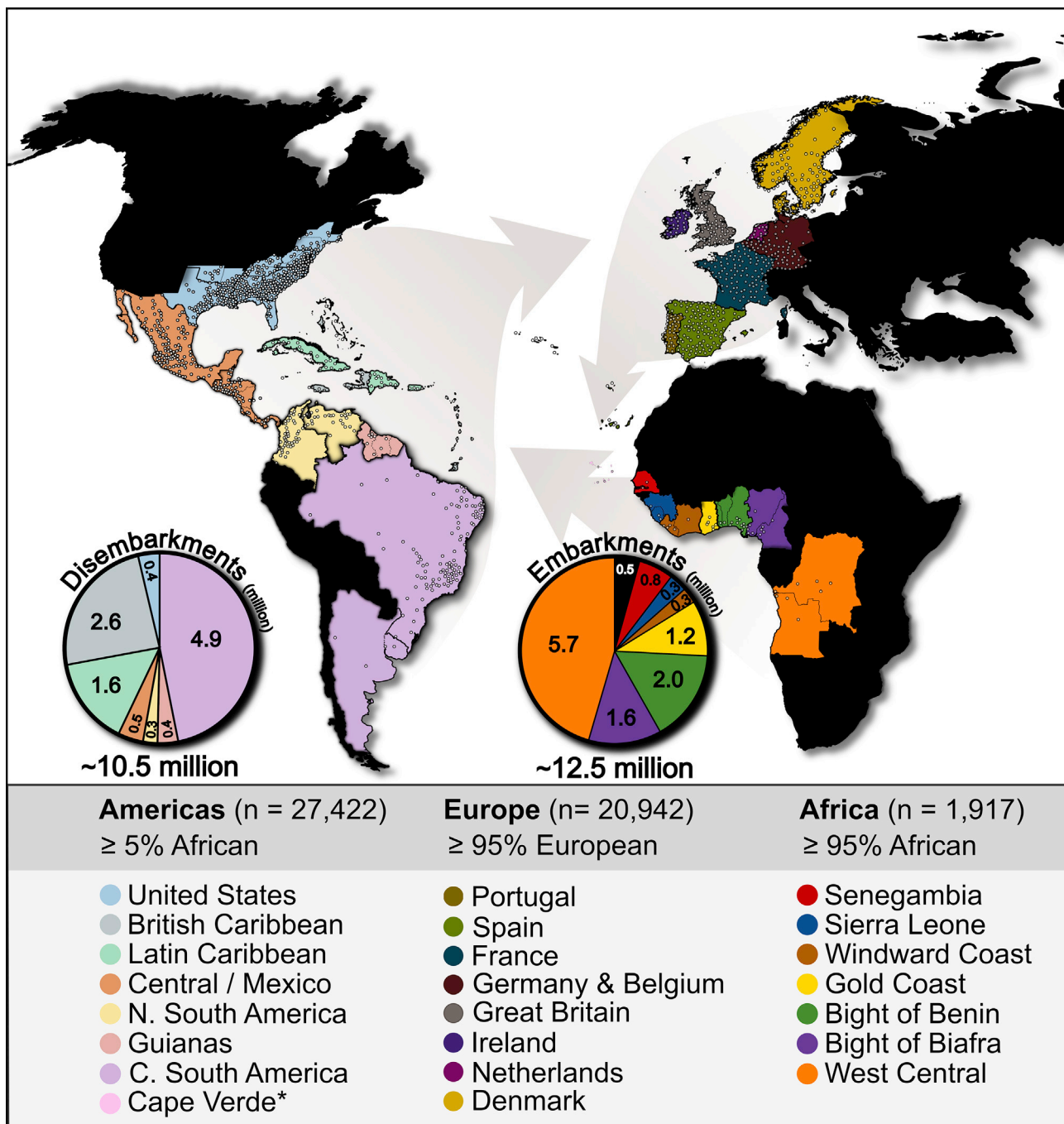
which are explicit about the use of customer data for ancestry research.

We grouped individuals across the Americas, Atlantic Africa, and Western Europe into cohorts designed to reflect historical regions (Tables S1–S3). Within the United States, we focused solely on coastal and surrounding states known to have been disembarkation locations for enslaved people.<sup>5,9</sup> We included the Cape Verde archipelago as control region since it is an Atlantic African country that was colonized by the Portuguese in the 15th century and has been inhabited primarily by the descendants of enslaved people from Senegambia.<sup>38</sup> In Africa, we merged contemporary administrative borders to match all major slave trading regions during the transatlantic slave trade except South East Africa, which constituted about 3.8% of all African embarkments, due to limited sample representation in Mozambique and Madagascar<sup>5,9</sup> (Figure 1; Table S1). Because records constitute number of captives deported by region of purchase, as opposed to ethnolinguistic affiliation, merging geographic boundaries in this manner allowed us to connect people of African descent directly to slave trading regions from which their African ancestors may have embarked.<sup>5</sup> Borders of Atlantic African countries generally aligned with those historical slave trading regions; however, we split Nigeria into western and eastern states along the Niger River because enslaved people from east of the Niger were deported from ports in the Bight of Biafra, whereas people west of the Niger were deported from ports in the Bight of Benin.<sup>9</sup> Categorizing Nigeria in this way also corresponds to the historical geographic occupancy of groups with genetic, ethnic, and linguistic differences: the Yoruba in the Bight of Benin and Esan and Igbo in the Bight of Biafra.<sup>4</sup> We considered current Atlantic African populations to be suitable proxies for past populations in Africa given limited migration and current population structure following a latitudinal gradient<sup>13,39,40</sup> (Figures S1–S6). We combined regions of Europe participating in the transatlantic slave trade by broad historical descriptors<sup>7</sup> and population structure (Figures S7–S9; Table S3), which are largely described by geographic proximity.

Study participants from the 23andMe customer base indicated their four grandparents were born within the same country (Latin America and the Caribbean) or state (United States), whereas study participants from public databases and Africa-focused studies were included based on self-described historical ties to an African country relevant to the transatlantic slave trade.<sup>32–35</sup> Using 23andMe's ancestry inference algorithm,<sup>41</sup> we excluded data from individuals in the Americas with less than 5% African ancestry, Africans with less than 95% African ancestry, and Europeans with less than 95% European ancestry. We additionally pruned pairs of related individuals using refinedIBD<sup>42</sup> so that no two individuals in the study share more than 500 cM of DNA that is identical by descent. These filtering steps reduced a sample size of 183,915 individuals representing the Americas and Africa to 29,339, and 36,866 individuals representing Europe to 20,942. In remaining Atlantic African representatives, we compiled ethnic affiliations from self-reports in 23andMe customers and sample descriptions in other datasets, finding 37 unique ethnic affiliations (Table S1).

### Genotyping

DNA extraction and genotyping of 23andMe customers and Human Genome Diversity Project were performed on saliva samples by the National Genetics Institute (NGI), a CLIA-licensed clinical laboratory and subsidiary of Laboratory Corporation of America. Samples were genotyped on one of five genotyping platforms.



**Figure 1. Location of Individuals and Cohorts**

Arrows highlight the general direction of the triangular trade routes between continents during the transatlantic slave trade. Pie charts indicate the documented number of enslaved people embarking out of regions of Africa (~12.5 million total) and disembarking in regions of the Americas (~10.5 million total) between 1515 and 1865. Representatives of regions of the Americas and Europe indicated that they each have four grandparents born within the same country or US state. Representatives of Atlantic Africa either indicated four grandparents born within or historical ties to a country. Points indicate the ~16,000 unique grandparental geo-coordinates provided by participants. \*Cape Verde is an Atlantic African island country that, in the 15th century, was colonized by the Portuguese and inhabited primarily by enslaved people from Senegambia.

The V1 and V2 platforms were variants of the Illumina Human-Hap550+ BeadChip, with 560,000 SNPs of which ~25,000 custom SNPs were selected by 23andMe. The V3 platform was based on the Illumina OmniExpress+ BeadChip, with additional SNPs, totaling about 950,000 SNPs. The V4 and V5 platforms are custom arrays, including a lower-redundancy subset of V2 and V3 SNPs with

additional coverage of lower-frequency coding variation and about 570,000 SNPs. Whole-genome genotypes from 1000 Genomes Project<sup>32</sup> were subset to match those of V1-V5 platforms and Human Genome Diversity Panel individuals (HGDP)<sup>33</sup> were genotyped on both V1 and V5 platforms. Remaining African datasets were genotyped on V3 and V5 platforms. Within each

platform, we removed any sample with >5% genotype missingness, producing an average missingness rate of 0.48% across all samples.

### Population Structure

We used both ordination techniques and ADMIXTURE<sup>43</sup> to assess population structure within Atlantic Africa as well as in Atlantic Africa plus the Americas. For these techniques, we reduced linkage by pruning SNPs overlapping across all genotype platforms with  $r^2 < 0.5$  via PLINK,<sup>44</sup> resulting in 43,780 SNPs. We included additional Karitiana, Maya, Pima, and Surui individuals from HGDP<sup>23</sup> as Native American references. Ordination techniques consisted of principal component analysis (PCA) and uniform manifold approximation and projection<sup>43</sup> (UMAP) using 0.8 minimum distance, 20 nearest neighbors, and number of pcs based on the plateau of eigenvalues (8–15). We ran ADMIXTURE,<sup>45</sup> with K values from 2 to 20 in Atlantic African, European, and across all samples.

### Voyage Records

We accessed transatlantic historical records from the Slave Voyages database.<sup>5</sup> The Slave Voyages database contains a collection of records from a wide range of published and archival information on more than 36,000 voyages across the Atlantic between 1492 and the early nineteenth century.<sup>5</sup> The database contains more than 100 possible variables for each voyage, including voyage dates, disembarkation/embarkation ports, crew information, and voyage outcomes. Some records that lacked pertinent data were imputed using information such as carrying capacity of vessels and characteristics of similar documented voyages.<sup>5</sup> Using the Slave Voyages database, we compiled estimates of enslaved African embarkation (“Imputed total num slaves purchased exact”), disembarkation (“Imputed total slaves disembarked exact”), proportions of males (“Imputed percentage male exact”), and proportion of children (“Imputed percentage child exact”) for each Atlantic African (“Imputed principal region of slave purchase lang en exact”) and American region (“Imputed principal port slave dis lang en exact;” “Imputed principal region slave dis lang en exact”), by voyage year (“Imputed arrival at port of dis exact”). For disembarkation estimates of enslaved Africans into the Americas, we mapped the principal port of disembarkation and/or principal region of disembarkation to American region and grouped by Atlantic African region. We estimated shipment sex bias using the “Voyage - percent males on” field and shipment age bias using the “Voyage - percent children on” field. We similarly utilized the intra-American slave trade database, which contains records of voyages within regions of the Americas that occurred in tandem with the transatlantic slave trade.<sup>5,46</sup> We aggregated estimates of the number of enslaved people forcefully deported between regions of the Americas using the “Imputed total disembarked,” “Imputed principal place of purchase,” and “Imputed principal place of slave landing” fields.<sup>5,46</sup>

### Ancestry Inference

We used 23andMe’s Ancestry Composition pipeline to identify the ancestral origins of chromosomal segments in individuals using all available genotype data.<sup>41,47</sup> Thus, the number of SNPs used to predict ancestry varied between 560,000 and 950,000 based on the platform an individual was genotyped on. Ancestry Composition uses support vector machine classifiers to assign ancestry labels to locally phased 300-SNP windows, whose ~1

Mb size increases the power for detecting ancestry origins within a window while still accurately predicting fine-scale local ancestry.<sup>48</sup> These initial labels are then processed with an autoregressive pair hidden Markov model to smooth and correct phasing errors. The resulting posterior probabilities are recalibrated with an isotonic regression model. Ancestry Composition assigns 33 local populations, four of which correspond to Atlantic African countries directly impacted by the slave trade, and four of which correspond to western European countries that conducted the slave trade.<sup>3</sup> The limited number of populations Ancestry Composition predicts reflects the inherent difficulty to distinguish populations with complex histories, and thus is a balance between accurate assignment of genetic segments and the number of unique populations. In Africa, potential local populations are as follows: Nigerian (Nigeria), Senegambian (The Gambia, Guinea, Guinea Bissau, Senegal), Coastal West African (Sierra Leone, Ghana, Côte d’Ivoire, Liberia), and Congolese (Angola, the Democratic Republic of the Congo). Local European populations are as follows: British and Irish (the United Kingdom and the Republic of Ireland), French and German (Austria, Belgium, France, Germany, the Netherlands, Luxembourg, Switzerland), Iberian (Spain and Portugal), and Scandinavian (Denmark, Iceland, Norway, and Sweden).

### Identity by Descent

We estimated identity by descent (IBD) within and between people representing regions in the Americas, Atlantic Africa, and western Europe using the software package refinedIBD.<sup>42</sup> We first removed all missing positions and subset markers to the intersection between all platforms, resulting in 87,847 unique positions used in all IBD analyses. While lowering SNP density may reduce the ability to detect smaller segments, removing missing positions minimizes the detection of false-positive segments.<sup>49</sup> With the preference of reducing false-positive calls, we additionally specified a minimum LOD score of 3.3 and a minimum reported length of 3 cM. All other parameters were set to default. We used mean pairwise IBD ( $_{ms}IBD$ ) to assess the overall IBD being shared between (or within) regions whereby  $_{ms}IBD$  is the sum of IBD shared between populations divided by the total pairwise comparisons of individuals in populations. We used this metric because it accounts for imbalances in sample sizes—although other metrics such as the proportion of individuals with IBD match, skewness of IBD segments, and number of IBD segments shared yield identical patterns (Table S4). We also determined the mean number of regions within Europe or Africa with which individuals from the Americas share a  $\geq 5$  cM segment ( $_{mc}IBD_{region}$ ). That is, an individual from a region in the Americas must share at least 5 cM with a single individual from a region in Europe or Africa to be considered to have an IBD connection to that region.

### Time to Most Recent Common Ancestor

To compare the timing of genetic connections between African populations and individuals from the Americas with historical records of the transatlantic shipping of enslaved Africans, we estimated the age of IBD segments using a Bayesian method.<sup>50</sup> The length of an IBD segment shared between two individuals given the time to their most recent common ancestor (MRCA) is Erlang-2 distributed with  $\lambda = g/50$ , where  $g$  is the number of generations back in time to the MRCA.<sup>50</sup> If we let  $l$  represent an IBD segment length observed in cM, and we allow the discrete number of generations  $g$  to be continuous  $t$ , this distribution is:



$$\mathcal{T}(l | t) = l \left( \frac{t}{50} \right)^2 e^{-\frac{t}{50} l}.$$

However, in practice, we only observe IBD segments larger than some length predetermined by our method to estimate IBD. Thus, we must condition the distribution of IBD segment lengths observed given the TMRCA on the minimum possible segment size

$$\mathcal{T}(l | t, l_{\min}) = \frac{l \left( \frac{t}{50} \right)^2 e^{-\frac{t}{50} l}}{1 - \int_0^{l_{\min}} s e^{-\frac{t}{50} s} ds},$$

where  $l_{\min}$  is the predetermined minimum IBD segment length in cM.

From the set of all IBD segments observed between sampled individuals in two populations we are interested in estimating the time of a MRCA that may have occurred during the transatlantic slave trade. However, not all of the MRCAs between each pair of individuals in the two populations lived during this time period. Some of the IBD segments may come from recent, post-slave trade immigration, and some of the IBD segments may be from more distant shared ancestors that predate the slave trade. We modeled the probability that an observed IBD segment comes from each of the three time periods post-slave trade, slave trade, and pre-slave trade as  $\omega_1$ ,  $\omega_2$ , and  $\omega_3$ . Given the possibility of ancestors from each of those time periods, the distribution of IBD segment lengths becomes:

$$\mathcal{T}(\mathcal{L} | t_1, t_2, t_3, \omega_1, \omega_2, \omega_3, l_{\min}) = \prod_{i=1}^n [\omega_1 \mathcal{T}(l_i | t_1, l_{\min}) + \omega_2 \mathcal{T}(l_i | t_2, l_{\min}) + \omega_3 \mathcal{T}(l_i | t_3, l_{\min})],$$

where  $t_1$ ,  $t_2$ , and  $t_3$  are the TMRCA for post-slave trade, slave trade, and pre-slave trade, respectively, and  $\mathcal{L} = \{l_1, l_2, \dots, l_n\}$  is the set of all IBD segment lengths observed in cM.

We placed uniform prior distributions on  $t_1$ ,  $t_2$ , and  $t_3$ . These distributions assume we know the time (in generations) of the start and end of the transatlantic slave trade. Assuming a generation time of 20 to 27 years, the transatlantic slave trade began roughly 14 to 20 generations ago.<sup>9,22</sup> We set our priors accordingly:

$t_1 \sim \text{Uniform}(0, 4.5)$ ,  $t_2 \sim \text{Uniform}(4.5, 20)$ ,  $t_3 \sim \text{Uniform}(20, 100)$ , and  $\omega_1, \omega_2, \omega_3 \sim \text{Dirichlet}(1, 1, 1)$ . We found that the relative timings estimated by our results were robust to a range of different priors, for example assigning the slave trade start and end times as 25 and 4 generations ago. In order to estimate posterior distributions of  $t_1$ ,  $t_2$ ,  $t_3$ ,  $\omega_1$ ,  $\omega_2$ , and  $\omega_3$ , we ran an MCMC analysis for 200,000 iterations, sampling parameter values every 40 iterations, and discarding the first 25% of samples as burnin. We assessed convergence of the MCMC by checking that trace plots exhibited apparent stationarity and ensuring that the effective sample size for each estimated parameter was more than 600.

## Ancestral Sex Bias

QA We determined unequal sex contributions to the gene pool using an admixture model with X chromosome ancestry estimates, autosome ancestry estimates,<sup>22,51</sup> and haplogroup imbalance between Y and mitochondrial haplogroups.<sup>52</sup> We estimated sex-specific contributions, whereby multiple source populations contribute to an admixed population, assuming a single admixture event (for generations 1 to 15).<sup>51</sup> We used this mechanistic admixture model to explicitly compare autosome to X chromosome ancestry to determine female ( $s_f$ ) and male contributions ( $s_m$ ) to the gene pools of populations with African ancestry. This discrete-time model specifically derives the moment of the distribution of the autosomal

admixture fraction from a specific source population and allows uneven sex-specific contribution from each source population at each generation.<sup>51</sup> We determined mitochondrial haplogroups using HaploGrep2 with Phylotree 17<sup>53</sup> and Y-haplogroups using yHaplo,<sup>54</sup> which uses phylogenetically informative SNPs to associate individuals with specific branches of the Y chromosome tree.<sup>55</sup> We assigned the origin of haplogroups based on previous studies associating haplogroups with either African, European, or indigenous American ancestry.<sup>24,52,56</sup>

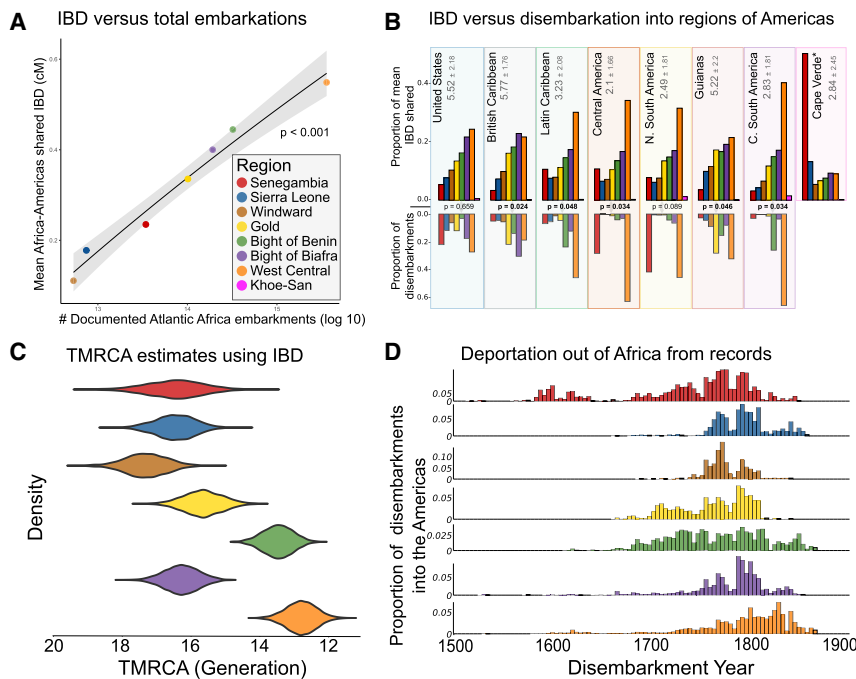
## Results

### Genetic Connections Compared to Historical Documents

The overall correlation between embarkation counts from Africa and mean proportion of identity by descent shared<sup>42</sup> ( $_{\text{ms}}\text{IBD}$ ) between regions of Atlantic Africa and all of the Americas (Figure 2A) is highly significant (Spearman's coefficient permutation;  $p < 0.001$ ). However, when we subset shipping records to specific disembarkation regions of the Americas,  $_{\text{ms}}\text{IBD}$  and enslaved disembarkation were not significantly correlated across all regions of the Americas (Figures 2B and S10). For example, the study cohorts representing the United States ( $p = 0.66$ ) and northern South America ( $p = 0.09$ ) have very little  $_{\text{ms}}\text{IBD}$  with Senegambia, compared to the proportion of captives from Senegambia disembarking in those two regions (Figure 2B). Many regions also demonstrate over-representation of  $_{\text{ms}}\text{IBD}$  with the Bights of Biafra and Benin compared to historical documents.

The mean number of IBD connections ( $_{\text{mc}}\text{IBD}_{\text{reg1-reg2}}$ ; the mean number of regions an individual shares a  $\geq 5$  cM IBD segment with) an individual in each region of the Americas has to the seven primary slave trading regions of Africa shows representative individuals from the United States match more slave trading regions ( $_{\text{mc}}\text{IBD}_{\text{US-Africa}} = 5.5 \pm 2.2$ ) than representative individuals from Central America and Mexico who tend to have IBD connections with fewer regions ( $_{\text{mc}}\text{IBD}_{\text{Central America - Africa}} = 2.1 \pm 1.7$ , Figure 2B; Table S5). Of note, there is a particularly strong discordance between disembarkations and  $_{\text{mc}}\text{IBD}$  in Central America and northern South America. There is also a stark difference between  $_{\text{mc}}\text{IBD}$  in the British Caribbean ( $_{\text{mc}}\text{IBD}_{\text{British Caribbean - Africa}} = 5.8 \pm 1.8$ ) versus Latin Caribbean ( $_{\text{mc}}\text{IBD}_{\text{Latin Caribbean - Africa}} = 3.2 \pm 2.1$ ), even though these regions are geographically proximate. These results correspond with comparisons between eight regions of western Europe and regions of the Americas, which indicated that Latin Americans tend to have higher  $_{\text{mc}}\text{IBD}_{\text{Latin America - Europe}}$  ( $x = 4.9 \pm 0.76$ ; primarily from Spanish Portuguese) than British occupied regions ( $x = 3.52 \pm 1.1$ ; Figures S11, S12; Table S6).

Our Bayesian MCMC approach to estimate time to most recent common ancestor living during the slave trade (TMRCA) from IBD segment lengths predicted that West Central Africa shares more recent common ancestors with individuals in the Americas, whereas Senegambia, Sierra Leone, the Windward Coast, and the Bight of Biafra



**Figure 2. IBD Sharing between the Americas and Atlantic Africa versus Documented Counts of Deportation**

(A) Comparison between the number of enslaved people embarking out of each region of Atlantic Africa to the mean shared IBD ( $m_{sIBD}$ ; cM) between each Atlantic Africa region and all individuals in the Americas.

(B)  $m_{sIBD}$  between regions of Africa and of the Americas (top) and the estimated proportions of enslaved disembarkation into each region of the Americas (bottom).  $p$  values indicate significant correlations between  $m_{sIBD}$  and disembarkation based on a permutation test using Spearman's coefficient of correlation. Numbers right of the region labels indicate the mean number ( $\pm$  STD) of  $m_{cIBD}$  connections ( $\geq 5$  cM) an individual representing that region of the Americas has to the seven primary slave trading regions.

(C) Posterior distributions of the estimated TMRCAs of IBD segments during the transatlantic slave trade between each Africa region and all of the Americas.

(D) The proportion of disembarkation of each Atlantic Africa region into the Americas over time, according to historical documents.

All Atlantic Africa regions correspond to the colors in key in (A).  $m_{sIBD}$  is the total, cumulative IBD shared between two populations (cM), divided by the total pairwise comparisons between two populations.

share earlier common ancestors (Figure 2C). This trend mirrors the majority of embarkation counts out of Africa from temporal historical documents (Figure 2D).

### Variation in African Ancestry across the Americas

Our estimations of African ancestry with 23andMe's ancestry inference algorithm indicated that mean sub-Saharan African ancestry in Americans with  $\geq 5\%$  African ancestry was highest in the British Caribbean ( $76\% \pm 20.3\%$ ), the United States ( $71.3\% \pm 22\%$ ), the Guianas ( $59.7\% \pm 32.6\%$ ), and Cape Verde ( $45.5\% \pm 12.6\%$ ; Figure 3). Mean sub-Saharan African ancestry was lowest in Latin America: Latin Caribbean ( $19.9\% \pm 12.9\%$ ), central South America ( $13.7\% \pm 10.5\%$ ), northern South America ( $11.9\% \pm 8.1\%$ ), and Central America plus Mexico ( $8.9\% \pm 8.7\%$ ). These proportions also correspond with the proportion of African haplogroups in each region (Tables S7 and S8).

Within 23andMe's Ancestry Composition hierarchy, there are four local ancestries corresponding to regions in Atlantic Africa: Nigerian (Nigeria), Senegambian (Gambia, Guinea, Guinea-Bissau, Senegal), Coastal West African (Sierra Leone, Ghana, Côte d'Ivoire, Liberia), and Congolese (Angola, Democratic Republic of the Congo). Ancestry inference in the Americas reveals that the majority of individuals within the United States (93%) as well as from the British and French Caribbean (82%) tend to have ancestry from all four of these Atlantic African populations, with Nigerian being the most common of the four. A high proportion of individuals from the Guianas (64%) and the Latin Caribbean (44%) also have ancestry from all four populations, with Coastal West African ancestry in the Guianas being most common relative to other Atlantic Af-

rican ancestries, and Senegambian being most common in the Latin Caribbean. The plurality (47%) of Mexican and Central American individuals have only Senegambian ancestry, whereas 32% of individuals from northern South America have both Senegambian and Nigerian ancestry. In central South America, the plurality (31%) of individuals have only Congolese ancestry (Figure 3; Table S7).

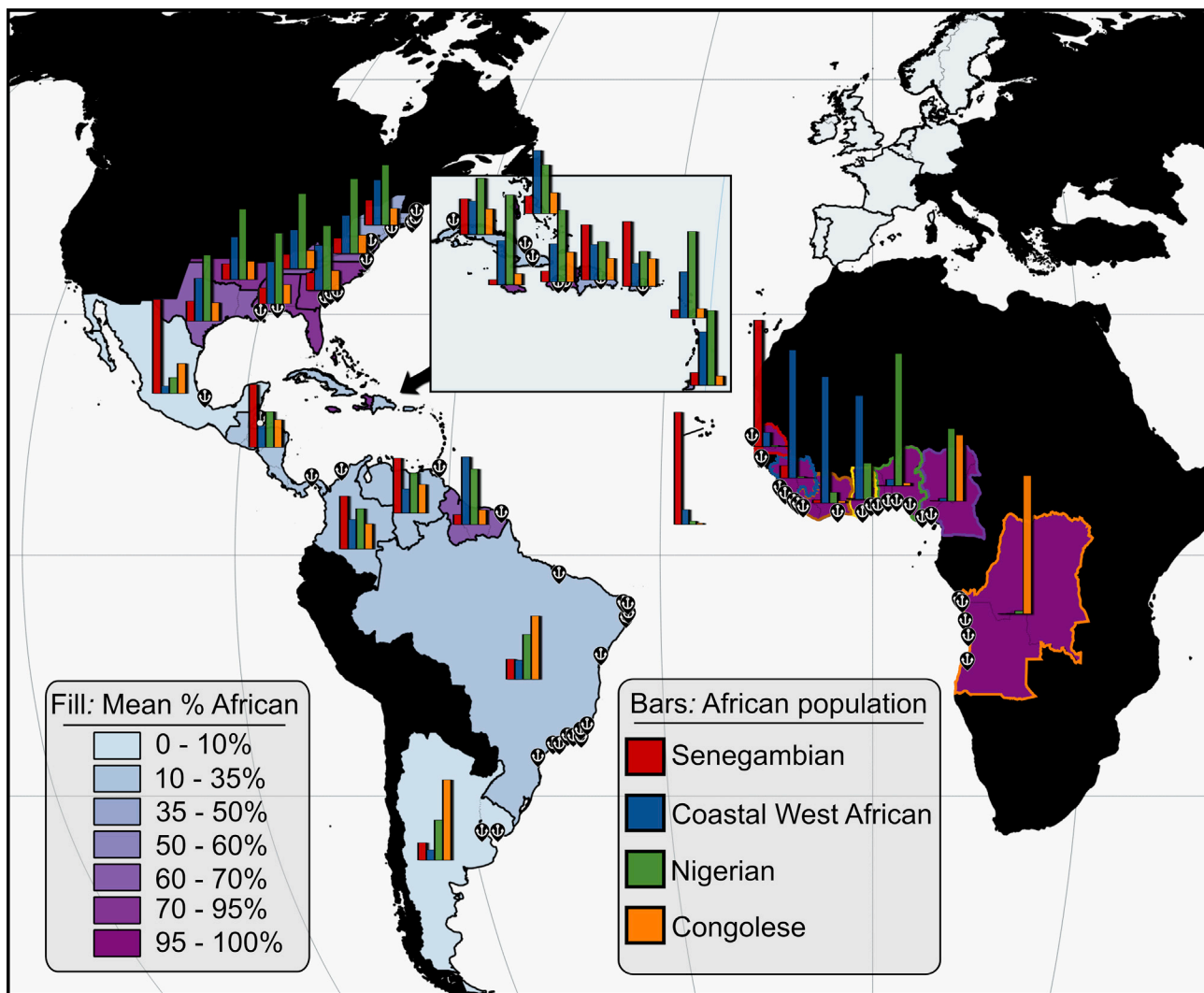
### Biases in Contributions of Males and Females to Current Gene Pools

We identified a bias toward European male and African female genetic contributions across regions of the Americas using the sex-specific model for admixture,<sup>22</sup> which assumes admixture events during the transatlantic slave trade. We observed an African female bias (African  $s_f/s_m > 1$ ) in every region of the Americas with highest levels in northern South America (17.17; Table 1; Tables S9 and S10). We also observed a European male sex bias (European  $s_m/s_f > 1$ ) in all explored regions except the Latin Caribbean and Central America. Additionally, haplogroup assignments in males indicated that Y chromosome haplogroups tend to be of European origin whereas mitochondrial haplogroups tend to be of African origin across all regions of the Americas (Table S8). Historical documents indicate the overall proportion of enslaved people on voyages were male ( $M_p > 0.60$ ) across all regions<sup>5</sup> (Table 1).

### Discussion

#### Overall Concordance with Historical Documents

We found that, overall, genetic evidence of Atlantic African ancestry across the Americas is consistent with historical



**Figure 3. Estimated Proportions of African Ancestry across Atlantic Africa, the Americas, and Europe**

Fill: Mean proportion of continent-level African Ancestry in each region. Bars: Mean local ancestry proportions of the four Atlantic African populations 23andMe identifies. Proportions are derived from local ancestry inference, which employs support vector machines (SVM) to assign one of 33 ancestries to each window along the genome and an autoregressive pair hidden Markov model to combine information across window-level SVM calls. Anchor symbols indicate major shipping port locations used during the slave trade.

documents of the transatlantic shipping of enslaved Africans.<sup>3,5,9</sup> The expectation that more embarkation out of Africa corresponds with more ancestral connections to the Americas is borne out by the comparison between  $msIBD_{\text{Africa-Americas}}$  and the number of captives deported from each African region (Figure 2A). For example, the highest proportion of  $msIBD$  is found between West Central Africa and the Americas, consistent with the highest proportion of enslaved embarkation from this slave-trading region (5.7 million people<sup>5</sup>). West Central Africa also experienced the largest proportion of deportations in the mid-1800s<sup>5</sup> (Figure 2D), consistent with our finding that people of African descent in the Americas share the most recent common ancestors with people from West Central Africa (Figure 2C). The mean number of IBD connections ( $mcIBD$ ) between regions of the Americas and Africa also support embarkation estimates from voyage records. The majority

of enslaved people arriving in Latin America (Latin Caribbean, Central America, N. South America, and C. South America) were generally deported from one or two slave trading regions, whereas enslaved people arriving in the United States and British Caribbean were taken from all regions of Atlantic Africa (Figure 2B).<sup>5,9,10</sup> Consequently, individuals from Latin America tend to have  $mcIBD$  with two or three Atlantic African regions, whereas individuals from the United States and the Caribbean tend to have genetic connections to five or six regions (Figure 2B).

#### Potential Inconsistencies with Historical Documents

While genetic connections between African populations and individuals from the Americas are strongly correlated with historical documents (Figure 2A), a granular breakdown of disembarkation in discrete regions of the Americas identifies discordance. For example,  $msIBD$  with the

**Table 1. Inferred Ancestry Sex Bias versus Documented Sex Bias**

America Region	African			European			Native American			$M_p$
	$s_f/s_m$	X	A	$s_m/s_f$	X	A	$s_f/s_m$	X	A	
Guianas	1.73	0.65	0.59	4.60	0.12	0.15	–	0.01	0.01	0.64
United States	1.47	0.76	0.71	3.03	0.21	0.25	–	0.02	0.01	0.69
British Caribbean	1.88	0.82	0.75	25.53	0.12	0.18	–	<0.01	<0.01	0.62
Latin Caribbean	13.31	0.24	0.19	0.90	0.53	0.52	3.78	0.12	0.06	0.66
C. South America	4.62	0.18	0.12	1.09	0.63	0.64	2.26	0.09	0.04	0.67
N. South America	17.17	0.14	0.11	2.27	0.43	0.49	3.38	0.30	0.17	0.63
Central America	15.60	0.11	0.08	0.89	0.32	0.31	28.04	0.44	0.27	0.68
Cape Verde	22.56	0.59	0.44	71.00	0.35	0.50	–	<0.01	<0.01	–

Female to male ( $s_f/s_m$ ) and male to female ( $s_m/s_f$ ) genetic contribution bias based on ancestry inference of the X chromosome and autosomes, assuming a single admixture event during the transatlantic slave trade (15 generations ago). For example,  $s_f/s_m = 2$  for African ancestry means ~2 African women for every African man. Mean Ancestry Composition based solely on the X chromosome (X) and solely on the autosomes (A) for each ancestry category.  $M_p$  is the proportion of enslaved males on voyages to each region of the Americas based on historical documents.  $s_f/s_m$  is not calculated for regions with less than 2% mean ancestry in either the X chromosome or autosomes. Assuming more recent admixture generations yielded near-identical results (Table S9).

Bights of Benin and Biafra, as well as Nigerian ancestry, is overrepresented in the United States and Latin America considering the recorded number of enslaved people that embarked from ports along the coast of present-day Nigeria into these regions of the Americas (Figures 2B and 3). However, this discordance is explained by records from the intra-American slave trade, an intercolonial trade that occurred primarily between 1619 and 1807 whose impact has recently been understood through collections of port records, newspapers, and merchant accounts from disparate parts of the Americas.<sup>5,46</sup> Documented intra-American voyages indicate that the vast majority of enslaved people were transported from the British Caribbean to other parts of the Americas, presumably to maintain the slave economy as transatlantic slave trading was increasingly prohibited (Figure S13).<sup>46</sup> These voyages would spread African ancestry common in the British Caribbean to other regions of the Americas that were not in direct trade with specific regions of Africa. In Latin America, transatlantic slave trading voyages virtually ceased between 1660 and 1780, yet a large proportion of all documented captives continued to be transported throughout the early 1800s to Latin America primarily from the British Caribbean.<sup>10</sup> In the United States, nearly 10% of the total enslaved population arrived via intra-America trading with the British Caribbean and consequently  $msIBD$  between the United States and the British Caribbean (primarily the Bahamas, British Windward, and Leeward Islands) is high (Figure S14). Therefore, embarkation out of Africa and  $msIBD$  between Africa and all of the Americas is correlated (Figure 2A), whereas disembarkation in specific regions of the Americas and  $msIBD$  with Africa are not consistently correlated (Figure 2B) because of the asymmetric transportation of enslaved people within the Americas.

Another apparent discordance between historical documents and genetic data is seen in  $msIBD$  between most regions of the Americas and Senegambia. Senegambia is underrepresented given the overall proportion of enslaved

people who were deported from this region to the Americas (Figure 2B). This is likely because Senegambia was one of the first African regions from which large numbers of people were enslaved.<sup>5,9,57</sup> Historical documents and TMRCA estimates from IBD segments support this explanation as well, indicating that Senegambia's connection to the Americas is older (and less well documented) than that of other regions (Figures 2C and 2D). Having distant ancestors from Senegambia would increase the likelihood of people in the Americas sharing very short IBD segments with Senegambians, which, in turn, may be undetected by estimating IBD segments with reduced SNP densities.<sup>49</sup> This pattern of reduced  $msIBD$  and older TMRCA estimates is most evident in the disembarkation regions of northern South America and Central America where, according to records, nearly half of the enslaved people who disembarked at ports in these regions came from Senegambia, but less than 10% of  $msIBD$  between these regions and Africa is with Senegambia (Figure 2B). However, almost no captives from Senegambia disembarked directly in either northern South America or Central America after 1650,<sup>5,9,10</sup> consistent with our finding of an older TMRCA between Senegambia and these two regions, and the likelihood that earlier African disembarkation resulted in smaller IBD segments between Africans and Americans of the present day.

The underrepresentation of ancestral connections to Senegambia also occurs in regions of the Americas with more recent (1700–1865) disembarkation from Senegambia, such as the United States and Latin Caribbean. This pattern, in part, may result from the increasing rates of deportation of children from Senegambia over time (Figure S15). Enslavement of children and unsanitary conditions in the holds of ships led to high rates of malnourishment and illness, which in turn led to lower survival rates in later decades of the slave trade.<sup>58,59</sup> It is also possible that Senegambians experienced higher mortality due to dangerous plantation conditions. A large proportion of enslaved Senegambians were rice cultivators who



worked on rice plantations in the Americas, which were generally rampant with malaria.<sup>57,60,61</sup> Clearing of swamplands was also generally performed by men, which could explain higher male mortality.<sup>59</sup> While detailed records exist for very few plantations, reduced genetic representation from Senegambia in parts of the Americas is likely correlated with Senegambians being forced to work under life-threatening conditions.

### Ancestry Proportions Discordance

Continental African ancestry estimates are unexpected given the number of enslaved people disembarking in the Americas. Nearly 70% of the 10.5 million enslaved people who survived the Middle Passage disembarked in Latin America, yet regions of Latin America have the lowest overall proportion of African ancestry (Figure 3). On the contrary, individuals in this study from the United States and British Caribbean have the highest African ancestry proportion on average. Together, these results indicate that practices and treatment of enslaved people likely varied across the Americas. For instance, literature suggests higher overall mortality and smaller effective population sizes in enslaved populations brought to Latin America,<sup>62,63</sup> which would have reduced the contribution of African ancestry to the gene pool. In addition, enslaved indigenous people often admixed with enslaved Africans in Latin America,<sup>59</sup> which is evident in the Native American ancestry and haplogroups common in individuals from Latin America (Tables S7 and S8). Other investigations of genetic ancestry have identified comparable estimates of African ancestry across the United States;<sup>11–15</sup> however, other studies have found that African ancestry is higher in admixed populations in Latin America compared to our study.<sup>13,64</sup> This is likely due to the non-random representation of global populations in our study cohort, which consists primarily of US customers and does not target specific admixed populations in Latin America. Although our estimates are lower (about 5%–10% on average), the trend of lower African ancestry in Latin America is still supported by other research.<sup>64</sup>

### Sex-Bias Discordance

Despite more than 60% of enslaved people brought to each region of the Americas being men,<sup>5</sup> comparisons of ancestry estimates for the X chromosome and autosomes, as well as the comparison of mitochondrial (maternal) and Y (paternal) haplogroups, revealed a bias toward African female contributions to gene pools across all of the Americas.<sup>64,65</sup> However, this African female sex bias is more extreme in Latin America (between 4 and 17 African women for every African man contributing to the gene pool) than in British-colonized Americas (between 1.5 and 2 African women for every African man contributing to the gene pool; Table 1). An Americas-wide African female sex-bias can be attributed to known accounts of rape of enslaved African women by slave owners and other sexual exploitation.<sup>19,26,62,66</sup> Regional differences may be

due to higher mortality in enslaved males in Latin America as well as a common practice called *branqueamento*, or racial whitening, which involved women marrying lighter-skinned men with the intention of producing lighter-skinned children.<sup>1,25,27,62,67</sup> National *branqueamento* policies were implemented in multiple Latin American countries, funding and subsidizing European immigrant travels with the intention to dilute African Ancestry through reproduction with light-skinned Europeans.<sup>68</sup> Conversely, the smaller African female sex-bias seen in former British colonies could be due to the practice of coercing enslaved people to have children as a means of maintaining enslaved workforces nearing the abolition of the transatlantic slave trade.<sup>69</sup> In some areas, such as the United States, enslaved women were incentivized to reproduce with the promise of freedom following the birth of many children.<sup>20</sup> Furthermore, racist ideologies in the United States led to the segregation of people of African descent as opposed to promotion of European admixture.<sup>70</sup> Overall, the inhumane practices associated with institutionalized slavery, though differing across the Americas, all resulted in an African-female sex bias despite the preponderance of males among those enslaved.

### Discordance in Timing

Our Bayesian method of dating IBD segments estimates the number of generations to a shared common ancestor between individuals in the Americas and Africa and is generally consistent with counts of enslaved disembarkation over time, with some exceptions (Figures 2C and 2D). One clear discordance is early genetic connections between the Americas and the Windward Coast relative to historical documents, which indicate a more recent connection. This discordance may be due to ethnolinguistic representation of the Windward Coast in this study. Participants representing the Windward Coast indicated primary ethnolinguistic associations with the Grebo, Kpelle, and Mano peoples, none of whom were known to be significantly impacted by the transatlantic slave trade.<sup>4</sup> Therefore, the detected IBD connections with this region may be with more distant common ancestors from the Windward Coast. Other African regions are represented in this study by ethnolinguistic groups more directly impacted by the slave trade (approximately half of the significantly impacted tribes;<sup>4</sup> Table S1). In general, our African cohort appears to be a suitable proxy of historical populations impacted by the slave trade given population structure (Figures S1–S3). IBD shared within regions of Africa also highlights general stationarity of African populations (Figure S17) yet signifies some impact of intercontinental migration. For instance, some mixing is apparent between the Bight of Biafra, the Bight of Benin, and West Central Africa. High shared IBD between certain regions may be due to ancient movement, such as the Bantu Migration,<sup>71</sup> or the more recent transportation of enslaved people within Africa (8 million people were enslaved and retained in Africa).<sup>72</sup>

Other discordances highlight differences in mortality following the transatlantic voyage among enslaved people who originated from different regions of Africa. For instance, the earlier-than-expected TMRCA for the Bight of Biafra may be due to the high suicide rates documented in enslaved Igbo people.<sup>73,74</sup> Most documented suicides, such as the Igbo Landing mass suicide of 1803 in the United States, were recorded later in the slave trade. If suicide by means of revolt became more common in enslaved people from the Bight of Biafra as the slave trade continued, that TMRCA estimate would appear older than expected given recorded shipping of enslaved people from the Bight of Biafra.

### Regional Discordance

Our results indicate that African ancestry and  $m_{\text{IBD}}$  do not vary much within broad regions of the Americas despite enslaved people disembarking at different rates in different ports. This observation is expected, given that slavers transported enslaved people regionally, and extensive local migration followed the transatlantic slave trade.<sup>75</sup> For instance, the domestic slave trade in the United States between 1790 and 1860 would see the transportation of one million enslaved people across the southeastern United States.<sup>20</sup> Similarly, “The Trail of Tears” would force Native Americans and the enslaved people they owned west toward Oklahoma.<sup>76</sup> Also, the “Great Migrations” that occurred between 1916 and 1970 to escape segregation and poor economic conditions would bring former slaves and their descendants to all parts of the Northeast and Midwest United States<sup>18</sup> (Figure S16). Similar movements occurred elsewhere such as migrations from northeastern Brazil (Bahia and Pernambuco) to southern and central Brazil (Rio de Janeiro and Minas Gerais-Goiás).<sup>1,2</sup> These large migrations could easily obscure pre-existing local population structure in African-descendant communities as a result of admixture. Because the majority of individuals in this study have connections to multiple populations in Atlantic Africa, it is unlikely that members of distinct African communities remained isolated from one another after they reached the Americas. However, based on limited marriage records of enslaved people,<sup>21,77,78</sup> it is likely that communities from distinct Atlantic African regions initially remained isolated, but interbred later, which would lead to genetic evidence of post-transport mixing of African populations over generations.

Despite extensive genetic evidence of internal or local migrations, we found no evidence for long-distance migrations between broad regions in the Americas after disembarkation except between the United States and the British Caribbean (Figures S5, S6, and S15). Representative individuals in the United States and British Caribbean share little IBD with those in Latin America, and very little IBD sharing is found between regions within Latin America. This is striking when comparing the British Caribbean to the Latin Caribbean: although some countries are geographically proximate (e.g., Haiti and the Dominican

Republic), ancestry estimates and IBD connections within the Americas and to Atlantic Africa differ substantially.

### Discordance between IBD and Ancestry Composition

Local ancestry estimates across the Americas (except for central South America) indicate a lower mean proportion of Congolese (West Central African) ancestry than expected (Figure 3), considering West Central Africa shares more  $m_{\text{IBD}}$  with the Americas than does any other region in Atlantic Africa (Figure 2A). This pattern likely reflects the fact that 23andMe’s local ancestry analysis differentiates only four regions along the Atlantic African coast and can detect genetic connections at a deeper timescale than IBD analysis. Specifically, local ancestry inference identifies the reference population with haplotypes most similar to those of a test sample for small genomic regions, rather than identifying identical-by-descent segments, so the ancestry it detects is likely more distant in time than that of IBD-based methods.<sup>47</sup>

### Conclusions

We present a broad-scale investigation into the transatlantic slave trade, advancing the understanding of its consequences by comparing genetic connections with historical records. While the current genetic landscape of African ancestry in the Americas largely agrees with transatlantic voyage records, additional historical accounts shed light on discordances. For example, overrepresentation of Nigerian ancestry in parts of the Americas is explained by the intra-American trade of enslaved people from the British Caribbean, and underrepresentation of Senegambian ancestry across the Americas is supported by accounts of early trading and high mortality from this region. Patterns of variation across the Americas, such as lower African ancestry and a higher African female sex bias can be attributed to socioeconomic factors, non-African male admixture, and inhumane treatment of enslaved people. While transatlantic records have told us a large part of the story, insights from this study in combination with other historical accounts shed light on details of the genetic impact of the transatlantic slave trade on present-day populations in the Americas.

### Data and Code Availability

There are restrictions to the availability of Angolan and 23andMe genotype data due to 23andMe consent and privacy guidelines. These data will not be made available.

### Consortia

The 23andMe Research Team: M. Agee, S. Aslibekyan, A. Auton, R. Bell, S. Clark, S. Das, S. Elson, K. Fletez-Brant, P. Fontanillas, P. Gandhi, K. Heilbron, B. Hicks, D. Hinds, K. Huber, E. Jewett, Y. Jiang, A. Kleinman, K. Lin, N. Litterman, J. McCreight, M. McIntyre, K. McManus, S. Mozafari, P. Nandakumar, L. Noblin, C. Northover, J. O’Connell,

A. Petrakovitz, S. Pitts, J. Shelton, S. Shringarpure, C. Tian, J. Tung, R. Tunney, V. Vacic, X. Wang, A. Zare.

## Supplemental Data

Supplemental Data can be found online at <https://doi.org/10.1016/j.ajhg.2020.06.012>.

## Acknowledgments

We thank the customers and others who consented to participate in research for enabling this study. We acknowledge other employees of 23andMe who contributed to the development of the infrastructure that made this research possible, and the suggestions of Scott Hadly, Altovise Ewing, Vence Bonham, Alvin Tillery, John Thornton, Linda Heywood, Beatriz Marcheco-Teruel, Richard Roberts, David Robinson, David Eltis, the [slavevoyages.org](https://www.slavevoyages.org) Scholars Group, and two anonymous reviewers. We also thank those who made additional samples available and provided guidance and discussion around those samples: Nathan Nunn (DRC), Bert Ely (Sierra Leone), and Brenna Henn (Khoe-San speakers). Finally, we would like to acknowledge the many people who were taken from Africa but never made it to the Americas, whose DNA is not directly part of this paper. Funding for this research was provided by 23andMe.

## Declaration of Interests

S.J.M., K.B., S.G.A.E., W.A.F., M.E.M., G.D.P., A.J.S., J.L.M., and members of the 23andMe Research Team are employees of and have stock, stock options, or both, in 23andMe. The remaining authors declare no competing interests.

Received: March 12, 2020

Accepted: June 15, 2020

Published: July 23, 2020

## Web Resources

23andMe's Ancestry Composition, <https://www.23andme.com/ancestry-composition-guide/>

1000 Genomes Project, <https://www.internationalgenome.org>

Ethical and Independent Review Services, <http://www.eandireview.com>

Human Genome Diversity Project, <https://www.hagsc.org/hgdp/>  
Intra-American Slave Trade Database, <https://slavevoyages.org/american/database/>

Transatlantic Slave Trade Database, <https://slavevoyages.org/voyage/database/>

## References

1. Martins, R.B. (1982). Growing in silence: the slave economy of nineteenth-century Minas Gerais, Brazil. *J. Econ. Hist.* 42, 222–223.
2. Bethell, L. (2009). *The Abolition of the Brazilian Slave Trade: Britain, Brazil and the Slave Trade Question* (Cambridge University Press).
3. Eltis, D. (2001). The volume and structure of the transatlantic slave trade: a reassessment. *William Mary Q.* 58, 17–46.
4. Heywood, L.M., and Thornton, J.K. (2007). *Central Africans, Atlantic creoles, and the foundation of the Americas, 1585–1660* (Cambridge University Press).
5. Eltis, D. (2007). A brief overview of the Trans-Atlantic Slave Trade. *Voyages: The Trans-Atlantic Slave Trade Database 1700–1810*. <http://www.slavevoyages.org/assessment/essay>.
6. Rotimi, C.N., Tekola-Ayele, F., Baker, J.L., and Shriner, D. (2016). The African diaspora: history, adaptation and health. *Curr. Opin. Genet. Dev.* 41, 77–84.
7. Acemoglu, D., Johnson, S., and Robinson, J. (2005). The rise of Europe: Atlantic trade, institutional change, and economic growth. *Am. Econ. Rev.* 95, 546–579.
8. Keller, A.G. (1908). *Colonization: a study of the founding of new societies* (Ginn).
9. Eltis, D., and Richardson, D. (2010). *Atlas of the Transatlantic Slave Trade* (Yale University Press).
10. Borucki, A., Eltis, D., and Wheat, D. (2015). Atlantic history and the slave trade to Spanish America. *Am. Hist. Rev.* 120, 433–461.
11. Gouveia, M.H., Borda, V., Leal, T.P., Moreira, R.G., Bergen, A.W., Kehdy, F.S.G., Alvim, I., Aquino, M.M., Araujo, G.S., Araujo, N.M., et al. (2020). Origins, admixture dynamics and homogenization of the African gene pool in the Americas. *Mol. Biol. Evol.* 37, 1647–1656.
12. Santos, H.C., Horimoto, A.V., Tarazona-Santos, E., Rodrigues-Soares, E., Barreto, M.L., Horta, B.L., Lima-Costa, M.F., Gouveia, M.H., Machado, M., Silva, T.M., et al.; Brazilian EPIGEN Project Consortium (2016). A minimum set of ancestry informative markers for determining admixture proportions in a mixed American population: the Brazilian set. *Eur. J. Hum. Genet.* 24, 725–731.
13. Fortes-Lima, C., Bybjerg-Grauholm, J., Marin-Padrón, L.C., Gomez-Cabezas, E.J., Bækvad-Hansen, M., Hansen, C.S., Le, P., Hougaard, D.M., Verdu, P., Mors, O., et al. (2018). Exploring Cuba's population structure and demographic history using genome-wide data. *Sci. Rep.* 8, 11422.
14. Wang, S., Ray, N., Rojas, W., Parra, M.V., Bedoya, G., Gallo, C., Poletti, G., Mazzotti, G., Hill, K., Hurtado, A.M., et al. (2008). Geographic patterns of genome admixture in Latin American Mestizos. *PLoS Genet.* 4, e1000037.
15. Ongaro, L., Scliar, M.O., Flores, R., Raveane, A., Marnetto, D., Sarno, S., Gneccchi-Ruscione, G.A., Alarcón-Riquelme, M.E., Patin, E., Wangkumhang, P., et al. (2019). The Genomic Impact of European Colonization of the Americas. *Curr. Biol.* 29, 3974–3986.e4.
16. Hall, G.M. (2005). *Slavery and African ethnicities in the Americas: restoring the links* (Univ of North Carolina Press).
17. Head, D. (2013). Slave Smuggling by Foreign Privateers: The Illegal Slave Trade and the Geopolitics of the Early Republic. *J. Early Repub.* 33, 433–462.
18. Baharian, S., Barakatt, M., Gignoux, C.R., Shringarpure, S., Erington, 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.
19. Baptist, E.E. (2001). “Cuffy,” “fancy maids,” and “one-eyed men”: rape, commodification, and the domestic slave trade in the United States. *Am. Hist. Rev.* 106, 1619–1650.
20. Deyle, S. (2005). *Carry me back: the domestic slave trade in American life* (Oxford University Press on Demand).

21. Edwards, L.F. (1996). "The Marriage Covenant is at the Foundation of all Our Rights": The Politics of Slave Marriages in North Carolina after Emancipation. *Law Hist. Rev.* 14, 81–124.
22. 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.
23. Lind, J.M., Hutcheson-Dilks, H.B., Williams, S.M., Moore, J.H., Essex, M., Ruiz-Pesini, E., Wallace, D.C., Tishkoff, S.A., O'Brien, S.J., and Smith, M.W. (2007). Elevated male European and female African contributions to the genomes of African American individuals. *Hum. Genet.* 120, 713–722.
24. 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.
25. Meade, T.A. (2016). *History of modern Latin America: 1800 to the present* (John Wiley & Sons).
26. Malarek, V. (2011). *The natashas: The horrific inside story of slavery, rape, and murder in the global sex trade* (Skyhorse Publishing Inc.).
27. Montalvo, F.F., and Codina, G.E. (2001). Skin color and Latinos in the United States. *Ethnicities* 1, 321–341.
28. Skidmore, T.E. (1993). *Black into white: race and nationality in Brazilian thought* (Duke University Press).
29. 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.
30. DeFrancesco, L. (2015). 23andMe protects consumers' data. *Nat. Biotechnol.* 33, 1220.
31. Eriksson, N., Macpherson, J.M., Tung, J.Y., Hon, L.S., Naughton, B., Saxonov, S., Avey, L., Wojcicki, A., Pe'er, I., and Mountain, J. (2010). Web-based, participant-driven studies yield novel genetic associations for common traits. *PLoS Genet.* 6, e1000993.
32. 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., Abecasis, G.R.; and 1000 Genomes Project Consortium (2015). A global reference for human genetic variation. *Nature* 526, 68–74.
33. Cann, H.M., de Toma, C., Cazes, L., Legrand, M.-F., Morel, V., Piouffre, L., Bodmer, J., Bodmer, W.F., Bonne-Tamir, B., Cambon-Thomsen, A., et al. (2002). A human genome diversity cell line panel. *Science* 296, 261–262.
34. van Dorp, L., Lowes, S., Weigel, J.L., Ansari-Pour, N., López, S., Mendoza-Revilla, J., Robinson, J.A., Henrich, J., Thomas, M.G., Nunn, N., and Hellenthal, G. (2019). Genetic legacy of state centralization in the Kuba Kingdom of the Democratic Republic of the Congo. *Proc. Natl. Acad. Sci. USA* 116, 593–598.
35. Jackson, B.A., Wilson, J.L., Kirbah, S., Sidney, S.S., Rosenberger, J., Bassie, L., Alie, J.A., McLean, D.C., Garvey, W.T., and Ely, B. (2005). Mitochondrial DNA genetic diversity among four ethnic groups in Sierra Leone. *Am. J. Phys. Anthropol.* 128, 156–163.
36. Henn, B.M., Gignoux, C.R., Jobin, M., Granka, J.M., Macpherson, J.M., Kidd, J.M., Rodríguez-Botigué, L., Ramachandran, S., Hon, L., Brisbin, A., et al. (2011). Hunter-gatherer genomic diversity suggests a southern African origin for modern humans. *Proc. Natl. Acad. Sci. USA* 108, 5154–5162.
37. Knight, A., Underhill, P.A., Mortensen, H.M., Zhivotovsky, L.A., Lin, A.A., Henn, B.M., Louis, D., Ruhlen, M., and Mountain, J.L. (2003). African Y chromosome and mtDNA divergence provides insight into the history of click languages. *Curr. Biol.* 13, 464–473.
38. Beleza, S., Campos, J., Lopes, J., Araújo, I.I., Hoppfer Almada, A., Correia e Silva, A., Parra, E.J., and Rocha, J. (2012). The admixture structure and genetic variation of the archipelago of Cape Verde and its implications for admixture mapping studies. *PLoS ONE* 7, e51103.
39. 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.
40. Murdock, G.P. (1959). *Africa its peoples and their culture history* (McGraw Hill Text).
41. Durand, E.Y., Do, C.B., Mountain, J.L., and Macpherson, J.M. (2014). Ancestry Composition: A Novel, Efficient Pipeline for Ancestry Deconvolution. *bioRxiv*. <https://doi.org/10.1101/010512>.
42. Browning, B.L., and Browning, S.R. (2013). Improving the accuracy and efficiency of identity-by-descent detection in population data. *Genetics* 194, 459–471.
43. McInnes, L., Healy, J., and Melville, J. (2018). Umap: Uniform manifold approximation and projection for dimension reduction. *ArXiv Preprint ArXiv:1802.03426*.
44. Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M.A., Bender, D., Maller, J., Sklar, P., de Bakker, P.I., Daly, M.J., and Sham, P.C. (2007). PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* 81, 559–575.
45. Alexander, D.H., Novembre, J., and Lange, K. (2009). Fast model-based estimation of ancestry in unrelated individuals. *Genome Res.* 19, 1655–1664.
46. O'Malley, G.E. (2014). *Final passages: the intercolonial slave trade of British America, 1619-1807* (UNC Press Books).
47. Do, C., Durand, E.Y.-M., and Macpherson, J.M. (2016). Scalable pipeline for local ancestry inference (Google Patents).
48. Loh, P.-R., Palamara, P.F., and Price, A.L. (2016). Fast and accurate long-range phasing in a UK Biobank cohort. *Nat. Genet.* 48, 811–816.
49. Powell, J.E., Visscher, P.M., and Goddard, M.E. (2010). Reconciling the analysis of IBD and IBS in complex trait studies. *Nat. Rev. Genet.* 11, 800–805.
50. Palamara, P.F., Lencz, T., Darvasi, A., and Pe'er, I. (2012). Length distributions of identity by descent reveal fine-scale demographic history. *Am. J. Hum. Genet.* 91, 809–822.
51. Goldberg, A., Verdu, P., and Rosenberg, N.A. (2014). Autosomal admixture levels are informative about sex bias in admixed populations. *Genetics* 198, 1209–1229.
52. Haber, M., Jones, A.L., Connell, B.A., Asan, Arciero, E., Yang, H., Thomas, M.G., Xue, Y., and Tyler-Smith, C. (2019). A Rare Deep-Rooting D0 African Y-Chromosomal Haplogroup and Its Implications for the Expansion of Modern Humans Out of Africa. *Genetics* 212, 1421–1428.
53. Weissensteiner, H., Pacher, D., Kloss-Brandstätter, A., Forer, L., Specht, G., Bandelt, H.-J., Kronenberg, F., Salas, A., and Schönherr, S. (2016). HaploGrep 2: mitochondrial haplogroup classification in the era of high-throughput sequencing. *Nucleic Acids Res.* 44 (W1), W58–63.



54. Poznik, G.D. (2016). Identifying Y-chromosome haplogroups in arbitrarily large samples of sequenced or genotyped men. *bioRxiv*. <https://doi.org/10.1101/088716>.
55. Consortium, Y.C.; and Y Chromosome Consortium (2002). A nomenclature system for the tree of human Y-chromosomal binary haplogroups. *Genome Res.* 12, 339–348.
56. Kivisild, T. (2015). Maternal ancestry and population history from whole mitochondrial genomes. *Investig. Genet.* 6, 3.
57. Baum, R.M. (1999). *Shrines of the slave trade: Diola religion and society in precolonial Senegambia* (Oxford University Press).
58. King, W. (2011). *Stolen childhood: Slave youth in nineteenth-century America* (Indiana University Press).
59. Schneider, D., and Schneider, C.J. (2014). *Slavery in America* (Infobase Publishing).
60. National Research Council (1996). *Lost crops of Africa: volume I: grains* (National Academies Press).
61. Carney, J. (2000). The African origins of Carolina rice culture. *Ecumene* 7, 125–149.
62. Bergad, L. (2007). *The comparative histories of slavery in Brazil, Cuba, and the United States* (Cambridge University Press).
63. Browning, S.R., Browning, B.L., Daviglus, M.L., Durazo-Arvizu, R.A., Schneiderman, N., Kaplan, R.C., and Laurie, C.C. (2018). Ancestry-specific recent effective population size in the Americas. *PLoS Genet.* 14, e1007385.
64. Marcheco-Teruel, B., Parra, E.J., Fuentes-Smith, E., Salas, A., Buttenschøn, H.N., Demontis, D., Torres-Español, M., Marín-Padrón, L.C., Gómez-Cabezas, E.J., Álvarez-Iglesias, V., et al. (2014). Cuba: exploring the history of admixture and the genetic basis of pigmentation using autosomal and uniparental markers. *PLoS Genet.* 10, e1004488.
65. Conomos, M.P., Laurie, C.A., Stilp, A.M., Gogarten, S.M., McHugh, C.P., Nelson, S.C., Sofer, T., Fernández-Rhodes, L., Justice, A.E., Graff, M., et al. (2016). Genetic Diversity and Association Studies in US Hispanic/Latino Populations: Applications in the Hispanic Community Health Study/Study of Latinos. *Am. J. Hum. Genet.* 98, 165–184.
66. Cash, F.G. (1853). *Slavery a System of Inherent Cruelty* (Leeds Anti-Slavery Society).
67. Klimentidis, Y.C., Miller, G.F., and Shriver, M.D. (2009). Genetic admixture, self-reported ethnicity, self-estimated admixture, and skin pigmentation among Hispanics and Native Americans. *Am. J. Phys. Anthropol.* 138, 375–383.
68. Agier, M. (1995). Racism, culture and black identity in Brazil. *Bull. Lat. Am. Res.* 14, 245–264.
69. Marable, M. (2015). *How capitalism underdeveloped Black America: Problems in race, political economy, and society* (Haymarket Books).
70. Laveist, T.A. (1993). Segregation, poverty, and empowerment: health consequences for African Americans. *Milbank Q.* 71, 41–64.
71. Vansina, J. (1995). New linguistic evidence and ‘the Bantu expansion.’ *J. Afr. Hist.* 36, 173–195.
72. Manning, P. (1981). The enslavement of Africans: a demographic model. *Canadian Journal of African Studies/La Revue Canadienne Des Études Africaines* 15, 499–526.
73. Rucker, W.C. (2008). *The river flows on: Black resistance, culture, and identity formation in early America* (LSU Press).
74. Snyder, T.L. (2010). Suicide, slavery, and memory in North America. *J. Am. Hist.* 97, 39–62.
75. Gates, H.L. (2011). *Life upon these shores: Looking at African American history, 1513-2008* (Alfred A Knopf Incorporated).
76. Henderson, G., and Olasiji, T.D. (1995). *Migrants, immigrants, and slaves: Racial and ethnic groups in America* (University Press of America).
77. Graham, R. (1976). Slave families on a rural estate in colonial Brazil. *J. Soc. Hist.* 9, 382–402.
78. Holt, K. (2005). Marriage choices in a plantation society: Bahia, Brazil. *Int. Rev. Soc. Hist.* 50, 25–41.