

## RESEARCH ARTICLE



WILEY

# Genetic landscape of Gullah African Americans

Kip D. Zimmerman<sup>1</sup> | Theodore G. Schurr<sup>2</sup> | Wei-Min Chen<sup>3,4</sup> | Uma Nayak<sup>3</sup> |  
 Josyf C. Mychaleckyj<sup>3,4</sup> | Queen Quet<sup>5</sup> | Lee H. Moultrie<sup>6</sup> | Jasmin Divers<sup>7</sup> |  
 Keith L. Keene<sup>8,9</sup> | Diane L. Kamen<sup>10</sup> | Gary S. Gilkeson<sup>10</sup> | Kelly J. Hunt<sup>11</sup> |  
 Ida J. Spruill<sup>12</sup> | Jyotika K. Fernandes<sup>10</sup> | Melinda C. Aldrich<sup>13,14,15,16</sup> |  
 David Reich<sup>17,18,19,20</sup> | W. Timothy Garvey<sup>21</sup> | Carl D. Langefeld<sup>1</sup> |  
 Michèle M. Sale<sup>3,4</sup> | Paula S. Ramos<sup>10,11</sup>

<sup>1</sup>Center for Precision Medicine, Wake Forest School of Medicine, Winston-Salem, North Carolina, USA

<sup>2</sup>Department of Anthropology, University of Pennsylvania, Philadelphia, Pennsylvania, USA

<sup>3</sup>Center for Public Health Genomics, University of Virginia, Charlottesville, Virginia, USA

<sup>4</sup>Department of Public Health Sciences, University of Virginia, Charlottesville, Virginia, USA

<sup>5</sup>Gullah/Geechee Nation, St. Helena Island, South Carolina, USA

<sup>6</sup>Lee H. Moultrie & Associates, North Charleston, South Carolina, USA

<sup>7</sup>Department of Health Services Research, New York University Winthrop Hospital, Mineola, New York, USA

<sup>8</sup>Department of Biology, East Carolina University, Greenville, North Carolina, USA

<sup>9</sup>Center for Health Disparities, East Carolina University Brody School of Medicine, Greenville, North Carolina, USA

<sup>10</sup>Department of Medicine, Medical University of South Carolina, Charleston, South Carolina, USA

<sup>11</sup>Department of Public Health Sciences, Medical University of South Carolina, Charleston, South Carolina, USA

<sup>12</sup>College of Nursing, Medical University of South Carolina, Charleston, South Carolina, USA

<sup>13</sup>Department of Thoracic Surgery, Vanderbilt University Medical Center, Nashville, Tennessee, USA

<sup>14</sup>Department of Medicine, Vanderbilt University Medical Center, Nashville, Tennessee, USA

<sup>15</sup>Department of Biomedical Informatics, Vanderbilt University Medical Center, Nashville, Tennessee, USA

<sup>16</sup>Vanderbilt Genetics Institute, Vanderbilt University Medical Center, Nashville, Tennessee, USA

<sup>17</sup>Department of Genetics, Harvard Medical School, Boston, Massachusetts, USA

<sup>18</sup>Howard Hughes Medical Institute, Harvard Medical School, Boston, Massachusetts, USA

<sup>19</sup>Broad Institute of MIT and Harvard, Cambridge, Massachusetts, USA

<sup>20</sup>Department of Human Evolutionary Biology, Harvard University, Cambridge, Massachusetts, USA

<sup>21</sup>Department of Nutrition Science, University of Alabama at Birmingham, Birmingham, Alabama, USA

## Correspondence

Paula S. Ramos, Division of Rheumatology & Immunology, Department of Medicine, Medical University of South Carolina, 96 Jonathan Lucas Street, MSC637, Charleston, SC 29425, USA.  
 Email: ramosp@musc.edu

## Abstract

**Objectives:** Gullah African Americans are descendants of formerly enslaved Africans living in the Sea Islands along the coast of the southeastern U.S., from North Carolina to Florida. Their relatively high numbers and geographic isolation were conducive to the development and preservation of a unique culture that retains deep African features. Although historical evidence supports a West-Central African ancestry for the

<sup>†</sup>This work is dedicated to Ida Spruill and Michèle Sale, deceased.

**Funding information**

National Center for Research Resources, Grant/Award Numbers: M01 RR001070, P20 RR-017696, UL1 RR029882; National Institute of Arthritis and Musculoskeletal and Skin Diseases, Grant/Award Numbers: K01 AR067280, P30 AR072582, P60 AR049459; National Institute of Diabetes and Digestive and Kidney Diseases, Grant/Award Number: R01 DK084350; W. M. Keck Foundation

Gullah, linguistic and cultural evidence of a connection to Sierra Leone has led to the suggestion of this country/region as their ancestral home. This study sought to elucidate the genetic structure and ancestry of the Gullah.

**Materials and Methods:** We leveraged whole-genome genotype data from Gullah, African Americans from Jackson, Mississippi, African populations from Sierra Leone, and population reference panels from Africa and Europe to infer population structure, ancestry proportions, and global estimates of admixture.

**Results:** Relative to non-Gullah African Americans from the Southeast US, the Gullah exhibited higher mean African ancestry, lower European admixture, a similarly small Native American contribution, and increased male-biased European admixture. A slightly tighter bottleneck in the Gullah 13 generations ago suggests a largely shared demographic history with non-Gullah African Americans. Despite a slightly higher relatedness to populations from Sierra Leone, our data demonstrate that the Gullah are genetically related to many West African populations.

**Discussion:** This study confirms that subtle differences in African American population structure exist at finer regional levels. Such observations can help to inform medical genetics research in African Americans, and guide the interpretation of genetic data used by African Americans seeking to explore ancestral identities.

**KEYWORDS**

admixture, ancestry, demography, slavery, West Africa

**1 | INTRODUCTION**

Knowledge about the genetic background of a population, including regional differences in ancestry, structure, and variation, is not only critical for medical and population genetic studies (Moreno-Estrada et al., 2014), but can also illuminate questions of social and cultural relevance. Multiple studies show wide variability in the levels of African ancestry in African American individuals in the United States (U.S.) at both state and regional levels (Baharian et al., 2016; Bryc, Durand, Macpherson, Reich, & Mountain, 2015; Dai et al., 2020; Han et al., 2017; Mathias et al., 2016; Micheletti et al., 2020; Patin et al., 2017). This variation in African ancestry underscores the importance of properly accounting for ancestry in historical and biomedical studies of diasporic populations (Reiner et al., 2005; Shriner, Tekola-Ayele, Adeyemo, & Rotimi, 2014).

While regional patterns in ancestry proportions in African Americans in the U.S. are broadly understood, fine-scale characterization of the ancestral diversity of discrete groups is lacking. In addition, as a result of the trans-Atlantic slave trade, African Americans were robbed of their African heritage and left with limited information about ancestors and homelands (Rotimi, 2003). The longing for identity and belonging leads many African Americans to actively draw together and evaluate various sources of genealogical information (historical, social and genetic) in order to weave together ancestry narratives (Nelson, 2008). Elucidating the African ancestries of African Americans can enrich the lives of African Americans by helping them to enrich a sense of identity, make connections to ancestral homelands, and ultimately foster reconciliation in the wake of emancipation (Nelson, 2016).

The Gullah are a culturally distinctive group of African Americans from the coastal Sea Islands of North Carolina, South Carolina, Georgia, and Florida. On many plantations of the coastal Sea Islands, Africans vastly outnumbered Europeans. Their relative isolation fostered the development of a unique culture in which many African influences were preserved, including language, folktales, religious beliefs, food preferences, music, dance, arts, and crafts (Jackson et al., 2005; Parra et al., 1998).

Marked by unique intonation and rhythm as well as syntax and lexicon, the Gullah language is hardly intelligible to the outsider, and remains the most characteristic feature of the sea islanders. The origin of this unique Creole language, like that of the Gullah people, is still debated. One hypothesis proposes that they descend from Krio ancestors originating in Sierra Leone (Hancock, 1969; Opala, 1987), a view supported by the fact that contemporary sea islanders can understand the Krio of Sierra Leone and vice versa. This striking linguistic resemblance, coupled with multiple cultural links (e.g., rice growing techniques, quilts, songs, stories), has led to a commonly held narrative that the Gullah are descendants of enslaved Africans from the African Rice Coast (Opala, 1987), the traditional rice-growing region which stretches south from Senegal to Sierra Leone and Liberia (Figure S1).

Other historical accounts support a diverse African ancestry of the Gullah (Brady, 1972; Nash, 2010; Pollitzer, 1999). The recorded legal slave trade into Charleston, South Carolina, documents approximately 39% of enslaved Africans as originating from West-Central Africa (present day Angola, Congo, and part of Gabon), 20% from

Senegambia (present day Senegal and Gambia), 17% from the Windward Coast (present day Ivory Coast and Liberia), 13% from the Gold Coast (present day Ghana), 6% from Sierra Leone (present day Sierra Leone and Guinea), and 5% from the Bights of Benin and Biafra (Togo, Benin, Nigeria, Cameroon, and part of Gabon) (Figure S1) (Pollitzer, 1999). In addition to cultural links (e.g., religious beliefs, arts and crafts), words and syntax support a larger role of West-Central Africa, the Gold Coast, and adjacent Nigeria in forming the Gullah language (Cassidy, 1980). While it has been proposed that the Congo-Angola area had an early cultural dominance with artifacts, lexicon and beliefs, the complexities of the Bantu grammar probably prevented its adoption in the Sea Islands. Senegambia, Sierra Leone, and the Windward Coast down through the Bight of Biafra contributed most in the latter half of the 18th century, when half of all slaves imported into Charleston arrived, adding more words, grammar, and even whole stories (Pollitzer, 1999). Thus, it is likely that, instead of Gullah deriving directly from Krio, both languages share a close common origin (Pollitzer, 1999).

In this context, it is currently unknown how significant of a genetic legacy that Sierra Leone ancestors might have left in present-day Gullah African Americans. Early genetic studies of autosomal, mtDNA, and Y-chromosome markers, indicated that the Gullah had high African ancestry (Parra et al., 1998; Parra et al., 2001) and a lower genetic distance to populations from Sierra Leone compared with African Americans from urban areas (McLean Jr. et al., 2005; McLean Jr. et al., 2003). Sierra Leone is also officially home to sixteen different ethnic groups, each with its own language and customs. The largest native ethnic groups include the Temne (35%) in northern Sierra Leone and areas around the capital, who arrived during the 11th and 12th centuries upon the fall of the Jalunkandu Empire in present day Republic of Guinea (Bangura, 2017; Rodney, 1970; Taylor, 2014; Wylie, 1977). The Mende (31%), who live mostly in the Southeast and the Kono District, appear to have early origins from people of the Empires of the Western Sudan (near present-day Mali), who migrated from the inland to the coast between the 2nd and 16th centuries to trade woven cloths for salt (Abraham, 2003; Fage, 2001), as well as more recent contributions from the Mane invasions of the 16th century (Dwyer, 2005; Ogot, 1992). The Limba (8%) are native to the savannah-woodland region in northern Sierra Leone, with many having moved to Freetown to escape capture by slave traders and transportation to North America (Alie, 1990; Finnegan, 1965). By contrast, the Fula (7%) are descendants of Fulani migrant from Guinea who settled in Sierra Leone during the 17th and 18th centuries (Fyfe, 1979; Knörr & Kohl, 2016). Likewise, the Kono (5%) and the Mandingo (2%) are descendants of Guinea migrants (Knörr & Kohl, 2016) (Figure S2).

The Creole (2%) are descendants of freed African slaves from America who settled in Sierra Leone after 1787, and as such have multiple African origins (Fyfe, 1979). Following the American Revolutionary War (1775–1783), the British government freed Africans who served in the British armed forces and resettled them in Granville Town, the predecessor of Freetown and the present capital of Sierra Leone. Maroons, runaway enslaved Africans from the West Indies who formed independent settlements on different islands, were also

resettled in Freetown, as were over 50,000 “recaptives” brought there by the British navy (Cole, 2013). The subsequent generations born in Sierra Leone were called Krio, Kriole, or Creole.

Individuals from other Sierra Leone ethnic groups joined the Creole communities, thereby promoting a fusion of African and Western cultures (Dixon-Fyle & Cole, 2006). The Krio language unites the different ethnic groups for trade and interactions with each other (Oyètádé & Fashole-Luke, 2008). The present national boundary was only fixed in 1896, prior to which people moved freely through the coastal country, making their own settlements, and fixing their own boundaries between themselves and their neighbors (Fyfe, 1979).

This study sought to elucidate the population structure of the Gullah and their relationship to contemporary Sierra Leone ethnic groups and other West African populations using genome-wide genotype data. African Americans from the Jackson Heart Study (JHS) (Musunuru et al., 2010; Taylor Jr. et al., 2005) were also included to provide a comparison with a less geographically isolated, but regionally close, Southeastern U.S. African American sample. The results of this analysis support the complex African ancestry and reduced European admixture of the Gullah compared to other U.S. African American populations. These results are consistent with historical data (Pollitzer, 1999), which indicate that the Gullah are a mixture of numerous people from different genetic, ethnic, and linguistic currents who formed their own culture and language. Identifying the diverse people who played a role in shaping the Gullah has implications for all African Americans, and for the legacy of the African diaspora everywhere (Pollitzer, 1999).

## 2 | SUBJECTS AND METHODS

### 2.1 | Community engagement

This study was conducted in cooperation with and approval from the Sea Island Families Project (SIFP) Citizen Advisory Committee (Spruill et al., 2013). Interdisciplinary research teams from the Medical University of South Carolina (MUSC) developed community-engaged research projects between the academic researchers and Gullah African Americans residing in rural South Carolina, leading to the formation of SIFP Citizen Advisory Committee (Spruill et al., 2013). This partnership has been ongoing for over 20 years. Participants from three SIFP projects were included in our study: the Sea Island Genetic African American Registry (Project SuGAR) (Divers et al., 2010); the Center of Biomedical Research Excellence (COBRE) for Oral Health pilot project, “An Epidemiological Study of Periodontal Disease and Diabetes: Cytokine Genes and Inflammation Factors” (Fernandes et al., 2009); and the Systemic Lupus Erythematosus in Gullah Health (SLEIGH) Study (Kamen et al., 2008). The SIFP Citizen Advisory Committee meets quarterly for sharing of research results and providing guidance and recommendations to new research. The SIFP Citizen Advisory Committee agreed to this study aimed at understanding the population genetics and ancestry of the Gullah.

## 2.2 | Sample collection and SNP data generation

Self-identified Gullah African American subjects in our study from the aforementioned SuGAR (Divers et al., 2010), COBRE for Oral Health (Fernandes et al., 2009), and SLEIGH (Kamen et al., 2008) projects were recruited under ongoing protocols approved by the MUSC Institutional Review Board and adhered to the tenets of the Declaration of Helsinki. All self-identified Gullah African American participants and their parents were born and raised in the Sea Islands region of South Carolina, or South Carolina low country (along the coastal border and 30 miles inland). The vast majority of Gullah participants was recruited through scheduled community health fairs, recruitment events at local churches, medical clinics, and established organizations on the Sea Islands (Divers et al., 2010; Sale et al., 2009). The major sites of recruitment spanned the Charleston and Beaufort counties, including Saint James-Santee, Johns Island, Edisto Island, Saint Helena Island, and Daufuskie Island.

All subjects received a general medical examination, donated blood samples for DNA analysis, and provided basic demographic and ethnic information. DNA was extracted from blood using a standardized DNA isolation kit (Genra Systems, Minneapolis, MN). Sample collection and processing have been previously described for the African American subjects from the Jackson Heart Study (JHS) (Musunuru et al., 2010; Taylor Jr. et al., 2005), Native American Mixtec (Raghavan et al., 2015), and Sierra Leonean (Jackson et al., 2005) subjects. We note that, while Gullah African Americans represent a predominantly rural sample, JHS participants represent the urban, metropolitan area of Jackson, Mississippi. Of note, while the Creole from Sierra Leone were largely recruited in Freetown, and are thus a more urban group, the other ethnic groups were recruited in smaller towns and communities where ethnic communities had lived for many years and the ethnic affiliations were distinct and reliable, and are thus representative of more rural areas. In summary, demographic representativeness was attempted during sampling.

DNAs from the Gullah and Sierra Leone African subjects were genotyped using the Affymetrix Genome-Wide Human SNP Array 6.0 at the Children's Hospital of Philadelphia's (CHOP) Center for Applied Genomics. The Bayesian robust linear model with Mahalanobis (BRLMM-P) algorithm was used to generate SNP calls, and additional quality control (QC) was performed to exclude SNPs with low genotype call rates (<95%), low minor allele frequency (MAF < 0.05), and genotypes inconsistent with Hardy Weinberg Equilibrium (HWE) ( $P < 10^{-10}$ ). Samples were excluded for low call rate (<95%) and for gender inconsistencies between recorded and genotype-inferred sex. Duplicates, first-, second-, and third-degree relatives were also excluded, with unrelatedness being defined as pair-wise kinship coefficients smaller than 0.0442 estimated by Kinship-based Inference for GWAS (KING) software v2.1.2 (Manichaikul et al., 2010).

Prior to QC procedures, 1558 Gullah African Americans, 1775 JHS African Americans, 400 Sierra Leone Africans, and 8 Mixtec individuals were available for our analysis. The Mexican Mixtec individuals have >99% Native American ancestry estimated by ADMIXTURE (Alexander,

Novembre, & Lange, 2009). After QC procedures, 883 unrelated Gullah African Americans, 1322 unrelated JHS African Americans, 381 unrelated Sierra Leone Africans, and 7 Mixtec subjects were retained for analyses (Table S1). In addition, 125 unrelated Stanford-Human Genome Diversity Project (HGDP) (Li et al., 2008) and 386 unrelated HapMap III (release 3) (Altshuler et al., 2010) subjects were used for analysis. The geographic and linguistic distribution of the African populations used in this study are shown in Table S2 and Figure 1a.

## 2.3 | Data merging and SNP trimming

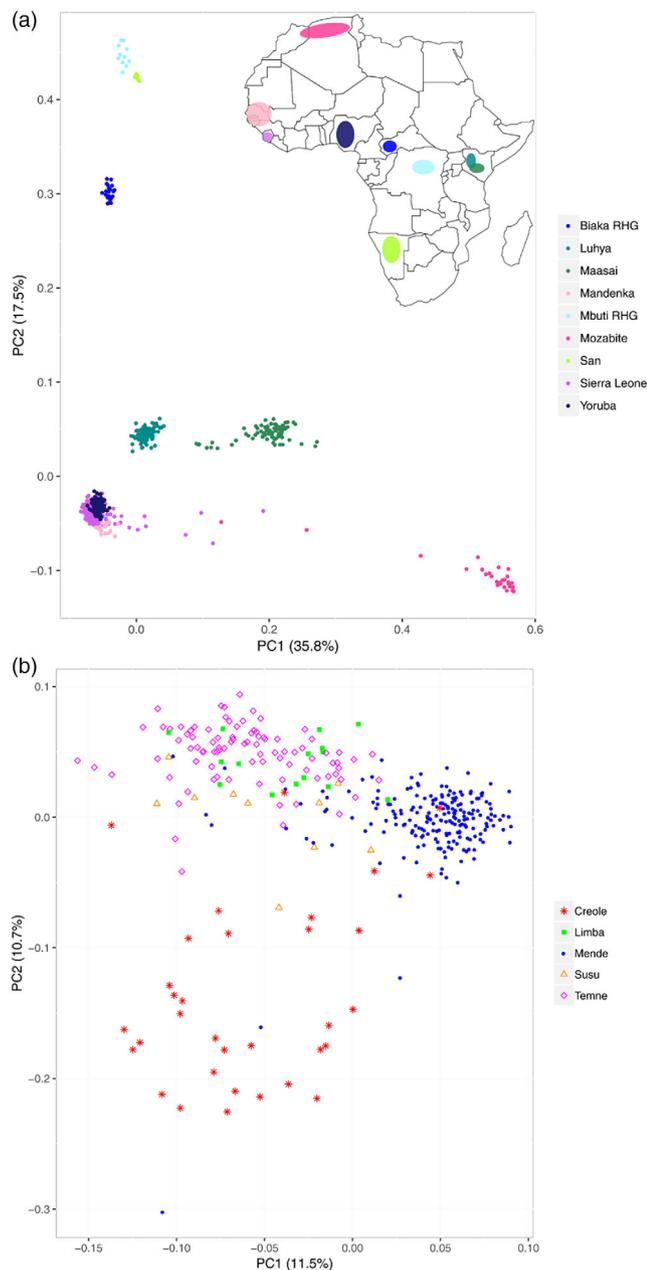
PLINK v1.9 (Chang et al., 2015) was used to combine our Affymetrix 6.0 data, HapMap III release 3 data, and HGDP data. After merging samples, 136,878 common variant SNPs were retained after removing SNPs (1) with strand problems, (2) occurring in the HLA region, (3) having a MAF < 0.01, (4) showing HWE <  $1 \times 10^{-10}$ , and (5) having a missing rate > 0.05. For methods that required a set of linkage disequilibrium (LD)-pruned SNPs (see below), we further removed SNPs with an  $r^2 > 0.1$ , leaving 64,303 uncorrelated SNPs for analysis. For more specific analyses, identical merging and filtering methods were used to combine the Affymetrix 6.0 data with the HapMap III data. This step resulted in a combined set of 27,374 pruned and filtered SNPs for the Gullah and 29,277 for the non-Gullah (JHS) African Americans.

## 2.4 | Principal component analysis for inference of population structure

Principal component analysis (PCA) as implemented in EIGENSOFT v6.0.1 (Patterson, Price, & Reich, 2006) was computed with the Sierra Leone African samples combined with all HGDP and HapMap III African samples, for African subpopulation structure inference. For African American population structure inference, HGDP and HapMap III European samples were added to the African samples as references. PCA analyses were performed using the set of LD-pruned 64,303 SNPs described above. To test for a significant shift in principal components between the Gullah and JHS, a PerMANOVA was computed for the sets of principal components between PC1 and PC6 (6 PCs explain >90% of the total variance). In addition, Wilcoxon signed-rank tests were computed on each individual principal component to test for differences between the Gullah and JHS.

## 2.5 | Admixture inference

Unsupervised clustering as implemented in ADMIXTURE v1.3.0 (Alexander et al., 2009) was used to estimate global genetic ancestry of the European, African, and African-American populations. ADMIXTURE analysis was performed in European, African (including Sierra Leonean populations), and African American individuals, assuming 2 through 8 ancestral genetic clusters ( $k = 2$  through  $k = 8$ ) to determine the optimal number of ancestral reference groups. Five clusters



**FIGURE 1** Principal component analysis of all African samples. Principal component analysis (PCA) (EIGENSOFT) was applied to HGDP and HapMap III African and Sierra Leonean populations. (a) PCA showing the Mozabite cluster (along PC1), the West, Eastern and Southern, and Middle and South Western subpopulations clusters (along PC2). Insert shows approximate locations of sampled populations in Africa. (b) PCA of Sierra Leone ethnic groups with  $n > 10$  showing the Mende, Creole and Temne forming relatively different, but overlapping, clusters

(i.e.,  $k = 5$ ) gave the lowest cross-validation error (Figure S3). To help order the populations according to their genetic similarities, we used average linkage hierarchical cluster analysis based on the means of each of the five ancestral populations computed by ADMIXTURE v1.3.0 (Alexander et al., 2009) and inter-population similarity matrix of Euclidean distances. The resulting dendrogram is shown in Figure S4.

ALDER v1.03 (Loh et al., 2013) was used to infer admixture timing in Gullah and JHS African Americans on the basis of exponential decay of linkage disequilibrium. ALDER was run separately for the Gullah and JHS using the YRI and CEU as reference populations. ALDER was computed with a minimum number of four individuals from the test population needing successful genotype calls at a SNP in order for the SNP to be used. The bin size used was 0.0005 cM and the minimum genetic distances at which to start and stop curve fitting were specified to 0.1 cM and 0.5 cM, respectively.

## 2.6 | Inference of ancestry proportions

To generate estimates of ancestry on the autosomes and chromosome X for both the Gullah and JHS African-Americans, we ran qpAdm (Haak et al., 2015). This program leverages allele frequency correlations between the admixed and source populations with distant outgroups to eliminate potential biases due to genetic drift between the true source populations and the ones used as surrogates for them. The HapMap CEU, HapMap YRI, and Mixtec (kindly provided by Drs. Raghavan and Willerslev (Raghavan et al., 2015)) were used as source reference populations for European, African, and Native American ancestry, respectively. It is important to note that, unlike other admixed Native Americans, the Mexican Mixtec individuals (Raghavan et al., 2015) have >99% Native American ancestry as estimated by ADMIXTURE (Alexander et al., 2009). For the outgroups, the Luhya (LWK), Maasai (MKK), Han Chinese (CHB), Japanese (JPT), Gujarati Indians (GIH), and Toscani (TSI) populations from HapMap were used.

qpAdm (Haak et al., 2015) was run on 798 females from the JHS cohort and 680 females from the Gullah cohort. A total of 523,638 SNPs were used to estimate ancestry on the autosomes and 20,879 SNPs were used to estimate ancestry from the X-chromosome. To determine sex-biased admixture in the African American populations, we examined the African, European and Native American ancestry proportions between the X-chromosomes and the autosomes in both the Gullah and the JHS African Americans for equality. To aid with interpretation, the proportion of ancestry that came from males was estimated under a simple model where, in a population with equally many females and males, the mean X-chromosomal admixture fraction is a linear combination of female and male admixture parameters, with coefficients 2/3 and 1/3, respectively.

## 2.7 | Detection of genomic segments shared identical-by-descent (IBD)

We used GERMLINE v1.5.1 (Gusev et al., 2009) to infer IBD tracts of length 18 cM or longer shared between Gullah and JHS African American individuals. GERMLINE was run under all of the default parameters with the exception of the minimum match length ( $-\text{min}_m$ ) of 18 cM, treating each individual as two distinct and separate haplotypes ( $-\text{haploid}$ ), 32 bits ( $-\text{bits}$ ), and extending the match beyond the slice end to the first mismatching marker ( $-\text{w}_\text{extend}$ ). Long IBD segments

( $l \geq 18$  cM) are informative of recent relatedness. Segments longer than 5 cM identified by GERMLINE have a negligible number of false positives (Durand, Eriksson, & McLean, 2014). IBD estimation was performed using the set of LD-pruned 64,303 SNPs described above.

## 2.8 | Estimation of effective population sizes

IBDNe (Browning et al., 2018) was implemented to estimate ancestry-specific effective population size. The analysis excluded IBD segments with a “mincm” length shorter than 6, included 80 bootstrap samples, and set the maximum number of generations to estimate at 100. Ancestry-specific IBDNe (ASIBDNe) was implemented following the exact sequence of steps for running ASIBDNe outlined by the creators of the IBDNe software online (Browning et al., 2018). The analysis pipeline can be found at <http://faculty.washington.edu/sguy/asibdne/>. Briefly, data were phased and IBD segments were detected with Beagle (Browning & Browning, 2007) before applying RFMix (Maples, Gravel, Kenny, & Bustamante, 2013) to estimate ancestry by chromosomal segments. Once ancestry-specific IBD segments were estimated, IBDNe was applied to each set of ancestry-specific IBD segments.

## 2.9 | Genetic diversity and population differentiation

To assess the amount of genetic variation within the Gullah and Sierra Leone populations, heterozygosity (HET) and inbreeding coefficients ( $F$ ) were calculated using genotypic data on 273 healthy Gullahs and 381 Sierra Leoneans. After pruning SNPs in high LD ( $r^2 > 0.5$ ) with PLINK v1.9 (Chang et al., 2015), the remaining 395 K shared SNPs between these populations were used to calculate the HET and  $F$  statistics. For each individual, HET and  $F$  were estimated based on their expected and observed heterozygous calls, with  $F = (\text{HET}_{\text{exp}} - \text{HET}_{\text{obs}})/\text{HET}_{\text{exp}}$ . The mean HET and mean  $F$  were calculated for each population and a two-sample Wilcoxon test was used to test for a difference in HET and  $F$  between populations.

We computed the fixation index ( $F_{\text{ST}}$ ) (Weir & Cockerham, 1984) to inform about the genetic divergence between African Americans (Gullah and JHS) and their African and European ancestral populations. Using genotype data from 273 healthy Gullah, JHS African Americans, Sierra Leone, YRI and CEU samples, we computed the Weir and Cockerham's (1984)  $F_{\text{ST}}$  (Weir & Cockerham, 1984) as implemented in VCFtools v0.1.13 (Danecek et al., 2011). A total of 479 K autosomal SNPs were used for this analysis.

## 3 | RESULTS AND DISCUSSION

### 3.1 | Genetic structure of Sierra Leone Africans

Given the popularized narrative that Sierra Leone was the ancestral source of most Gullah African Americans, we first sought to

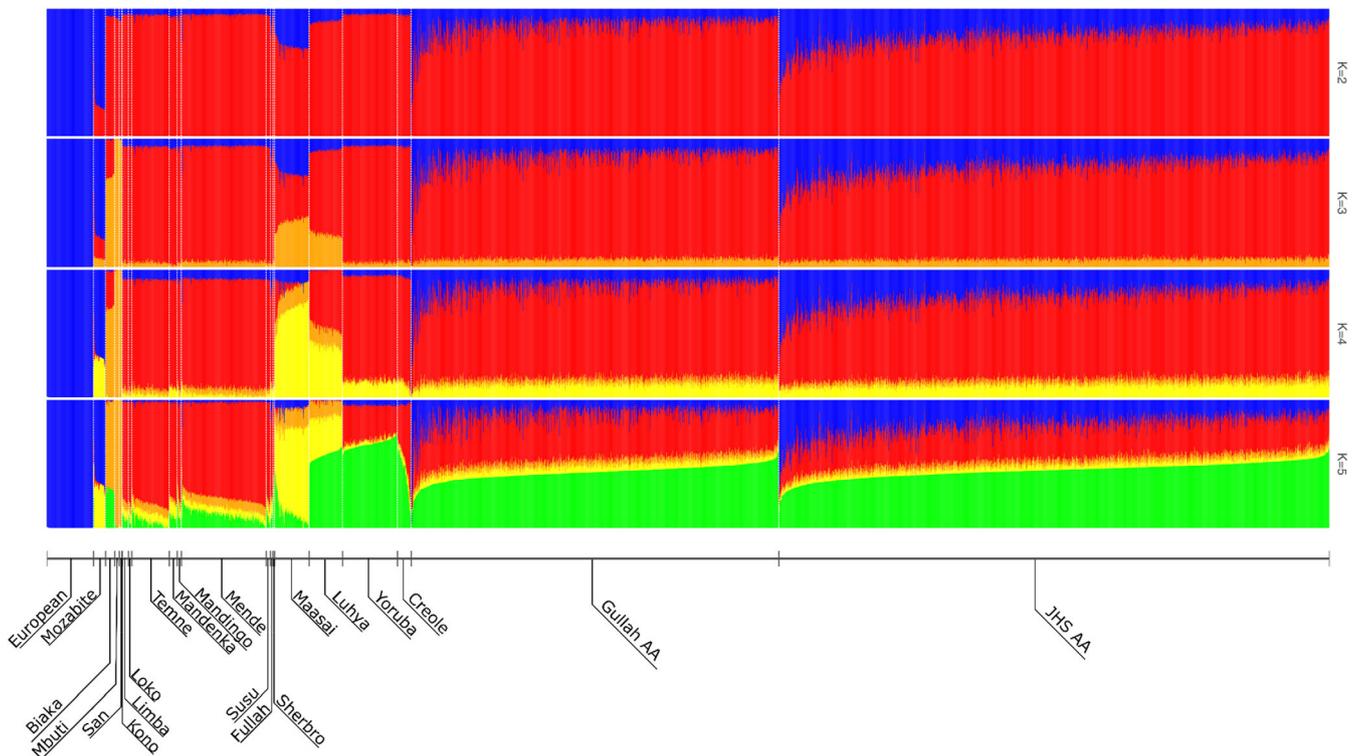
characterize the population structure of African ethnic groups from this region. We combined genotype data from Sierra Leone Africans with Africans from the HGDP and HapMap III studies (Table S1), and inferred patterns of population structure and individual ancestry by principal component analysis (PCA) (Figure 1) and ADMIXTURE (Alexander et al., 2009) (Figure 2). Consistent with previous reports (see the review by (Gomez, Hirbo, & Tishkoff, 2014), PCA distinguished geographic and linguistic African subpopulations, separating a combination of geographic groups and speakers of the four major language families (Afro-Asiatic, Nilo-Saharan, Niger-Kordofanian, and Khoisan) (Table S2).

As shown in Figure 1a, the first principal component (PC1) differentiated the Mozabites of North Africa from all other African populations. A few Mozabite and Sierra Leone individuals formed a geographical gradient, reflecting different levels of African and West-Eurasian-related admixture (Figure S5) (Henn et al., 2012). A slight partitioning of the East African groups also occurred, with the Maasai (Kenya) being separated from the Luhya (Kenya). PC2 further separated four main groups, with West Africans forming one cluster and the East African groups clustering together (Luhya and Maasai). The Biaka rainforest hunter-gatherers (Central African Republic) formed an individual cluster, whereas the Mbuti rainforest hunter-gatherers (Democratic Republic of the Congo) and San (Namibia) formed a more distant cluster. The relationship between population structure and geographic and linguistic factors is supported by the results from the global ancestry estimates (Figure 2 and Figure S4). ADMIXTURE (Alexander et al., 2009) (Figure 2) showed the highest West African-associated ancestry in the Mandenka of Senegal and Sierra Leone ethnic groups (red), the highest East African-associated ancestry in the Maasai and Luhya (yellow), and the highest Central and South African-associated ancestry in the Mbuti and San, respectively (*orange*).

PCA was applied to data from Sierra Leone ethnic groups to further discriminate any potential clusters of genetic variation within the region (Figure 1b). Our results showed that the most populous ethnic groups (Mende and Temne) form relatively different, but overlapping clusters. The Limba clustered among the Temne, which was unexpected given their distinct, unrelated language, and the results of a previous analysis of mtDNA genetic diversity purporting that the Limba could be distinguished from the Mende, Temne, and Loko groups (Jackson et al., 2005). The Limba are indigenous to Sierra Leone, and their dialects are largely unrelated to the other languages in the region.

The Krio or Creole formed a relatively distinct cluster along PC2. This pattern of population structure was broadly consistent with individual ancestry estimates (Figure 2 and Figure S4), whereas the Temne and Mende showed similar ancestry proportions, and the Creole appeared more variable in their African ancestry than other groups (Figure 2 and Figure S4). The Creole were also slightly more similar to the Yoruba, while other Sierra Leone ethnic groups showed more genetic similarity to the Mandenka (Figure 3 and Figure S4).

Interestingly, the Creole were genetically intermediate between other Sierra Leone ethnic groups and the Yoruba, with ancestral diversity similar to that of the Gullah African Americans. This finding



**FIGURE 2** Ancestry estimates for European, African, and African American populations. ADMIXTURE analysis in Europeans, Africans (including Sierra Leone ethnic groups) and African Americans, assuming two through five ancestral genetic clusters ( $k = 2$  through  $k = 5$ ). The  $k = 5$  setting has the lowest cross-validation error of  $k = 2-8$ . Populations were ordered via hierarchical cluster analysis. The plot shows each individual as a thin vertical column colored in proportion to their estimated ancestry from one particular population. The initial distinction is between Europeans (blue) and Africans (other colors). Within Africans, red indicates a West African (aka African Rice Coast) ancestry (highest in the Mandenka of Senegal and Sierra Leone ethnic groups), orange a Central and South African ancestry (highest in the San of Namibia and Mbuti of Democratic Republic of the Congo), yellow an East African ancestry (highest in Maasai and Luhya of Kenya), and green a West-Central African ancestry (aka Bight of Benin) ancestry (highest in the Yoruba of Nigeria)

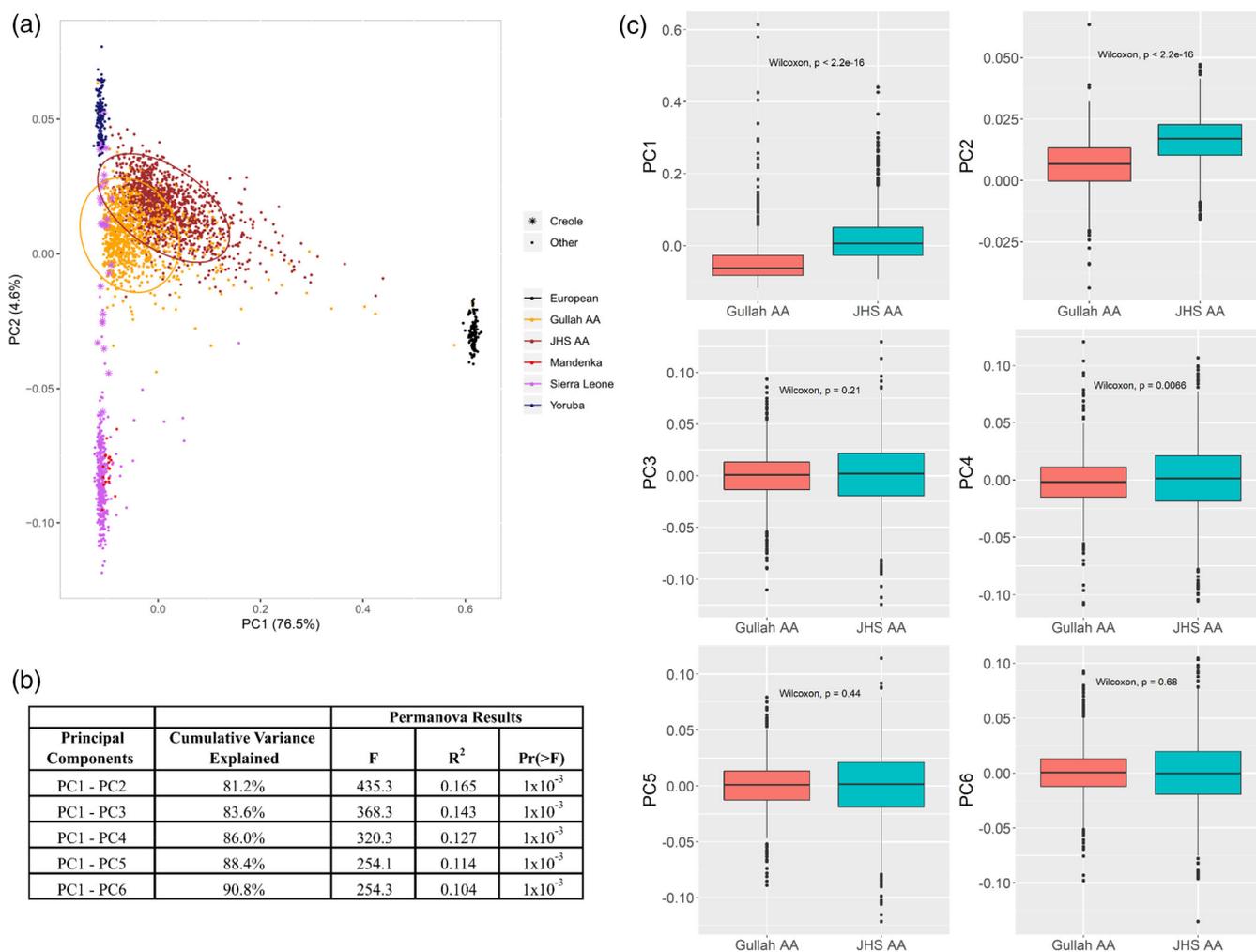
suggested that they were likely the descendants of individuals originating from various parts of Africa, including Sierra Leone and beyond, as well as African descended individuals with European admixture. This finding is also consistent with their demographic history, which suggests the Creole descended from freed enslaved Africans (Fyfe, 1979) who mixed with other ethnic groups (Dixon-Fyle & Cole, 2006). In summary, analysis of the African samples in this study showed that, despite the similarity among African Rice Coast populations (Mandenka and Sierra Leone ethnic groups), most of the recognized populations exhibit considerable genetic diversity.

### 3.2 | African ancestry estimates of Gullah African Americans

To characterize the global patterns of ancestry and population structure of Gullah African Americans, we used qpAdm (Haak et al., 2015; Patterson et al., 2012), ADMIXTURE (Alexander et al., 2009), and PCA, with data from Gullah living in South Carolina, and for comparison with a regionally close group, from JHS African Americans. This analysis confirmed previous genetic reports of autosomal, mtDNA, and Y-chromosome markers, i.e., that Gullah African Americans had

lower European admixture and higher African ancestry than other African American populations in the USA (McLean Jr. et al., 2003; McLean Jr. et al., 2005; Parra et al., 1998; Parra et al., 2001). The higher average proportion of African ancestry in the Gullah was evident from autosomal global ancestry inference from qpAdm (Haak et al., 2015; Patterson et al., 2012), with the average African contribution to the Gullah African Americans being 90.7% compared with 82.2% in JHS African Americans (Table 1). Among studies of African ancestry in different U.S. regions (Baharian et al., 2016; Bryc et al., 2015; Dai et al., 2020; Mathias et al., 2016; Micheletti et al., 2020; Patin et al., 2017), the high African ancestry proportion seen in the Gullah was nearly matched in a U.S. cohort of African Americans sampled within rural Southeast U.S. (89% in Florida and 88% in South Carolina) (Baharian et al., 2016). Thus, the Gullah show the highest average African ancestry proportion of any U.S. African American group studied to date.

In parallel with the higher average African ancestry, the European ancestry estimate in the Gullah was the lowest reported for African Americans in the U.S., while that for the JHS African Americans was similar to estimates in other groups (e.g., as low as 14% and 15%) (Baharian et al., 2016; Bryc et al., 2015; Mathias et al., 2016) (Table 1). Consistent with the ancestry estimates, genome-wide admixture



**FIGURE 3** Principal component analysis of Gullah and non-Gullah African Americans. (a) Principal component analysis (EIGENSOFT) using European (CEU samples from Utah) and African ancestral reference populations (YRI samples from Nigeria, Mandenka from Senegal, and Sierra Leone samples) illustrates the Gullah's closeness to Sierra Leone populations and the Mandenka, and also non-Gullah (JHS) African Americans' proximity to the Yoruba. (b) To demonstrate the difference between Gullah and JHS African Americans, a PERMANOVA analysis was computed for each set of principal components between PC1 and PC6 to test for differences between the two groups; the resulting analyses each demonstrated significant differences ( $1 \times 10^{-3}$ ), indicating a significant distance between the two groups. (c) Comparisons of each individual principal component between the Gullah and JHS. Wilcoxon p-values are displayed. PC1 and PC2 appear to be the primary drivers in the differences between the Gullah and JHS. AA: African American; JHS: Jackson Heart Study African Americans from Jackson, Mississippi

analyses showed that, despite the highly variable levels of European and West African ancestry in all African American groups, the Gullah had a lower average level of European admixture than JHS African Americans (Figure 2). The lower European contribution in the Gullah corroborates known differences in ancestry proportions among African Americans in different U.S. states (Baharian et al., 2016; Bryc et al., 2015; Dai et al., 2020; Mathias et al., 2016; Micheletti et al., 2020; Patin et al., 2017), and confirms that subtle differences in African American population structure can exist at finer regional levels.

The lower European admixture in the Gullah was further corroborated by estimates of admixture timing in Gullah and JHS African Americans. To obtain a relative date of the admixture in the African American groups, we used admixture linkage disequilibrium (LD) decay, as implemented in ALDER (Loh et al., 2013). The weighted

LD curve for the JHS African American group shows a steeper decay rate than that for the Gullah group as a function of genetic distance, suggesting earlier admixture in JHS African Americans (Figure S6). Accordingly, the admixture estimate for JHS African Americans dates to 4.5 generations ago, while for Gullah African Americans, the admixture estimate dates to 4.2 generations ago (Table 2). ALDER provides a relative date estimate for a single-pulse admixture event, producing a single average date in cases of multiple pulses of admixture. Hence, these admixture estimates do not preclude earlier admixture events. Regardless, these estimates are close to the estimation of admixture time for HapMap African Americans in Southwest USA (ASW) of 5.4 generations ago using the approach by Zhou and colleagues (Zhou et al., 2017). In summary, the low European ancestry in the Gullah is supported by a more recent admixture date.

**TABLE 1** Estimates of African, European, and Native American ancestry in two African American cohorts (% with 95% confidence intervals)

	Ancestry		
	African	European	Native American
<b>JHS African Americans (n = 1322)</b>			
Autosomes (p-value = 0.00)	82.2% [82.1–82.3]	16.3% [16.2–16.4]	1.4% [1.3–1.5]
X-chromosome (p-value = 0.14)	86.5% [86.0–87.0]	11.8% [11.2–12.2]	1.7% [1.0–2.4]
Difference in mean X and autosomal ancestry	+4.3%	–4.5%	+0.3%
Estimated proportion of ancestry from males	42.15%	91.41%	17.86%
<b>Gullah African Americans (n = 883)</b>			
Autosomes (p-value = 0.00)	90.7% [90.6–90.8]	8.0% [7.9–8.1]	1.3% [1.2–1.4]
X-chromosome (p-value = 0.26)	93.4% [92.9–93.9]	4.3% [3.7–4.9]	2.2% [1.1–3.1]
Difference in mean X and autosomal ancestry	+2.7%	–3.7%	+0.9%
Estimated proportion of ancestry from males	45.53%	119.38%	–53.85%

Note: Mean estimates (95% confidence intervals) of African, European, and Native American ancestry are shown. qpAdm was used to estimate proportions of European, African, and Native American ancestry, and qpAdm rank P-values are listed for both cohorts for each of the autosomes and chromosome X analyses. The mean X-chromosomal ancestry minus the mean autosomal ancestry (%) is listed. To aid with interpretation, the proportion of ancestry that comes from males was estimated under a simple model in which, in a population with equally many females and males, the mean X-chromosomal admixture fraction is a linear combination of female and male admixture parameters, with coefficients 2/3 and 1/3, respectively. The proportion of European ancestry across all chromosomes is higher in the Jackson Heart Study (JHS) relative to the Gullah African Americans, but the proportion of European ancestry that comes from males is higher in the Gullah African Americans.

**TABLE 2** Inference of admixture timing using ALDER

Test population	Reference population		Z-score	Decay	Amplitude	Test status	Test p-value
	A	B					
JHS AA	YRI	CEU	22.44	4.54 ± 0.18	3.4 × 10 <sup>-04</sup>	Success	1.7 × 10 <sup>-111</sup>
Gullah AA	YRI	CEU	25.15	4.21 ± 0.15	2.2 × 10 <sup>-04</sup>	Success	1.3 × 10 <sup>-139</sup>

Note: Decay shows the estimate for admixture date (in generations) and its block jackknifing standard error. JHS: Jackson Heart Study African Americans from Jackson, Mississippi; AA: African American; CEU: European reference samples from Utah; YRI: African reference samples from Nigeria.

The higher mean level of African ancestry in the Gullah is likely the result of the historically higher proportion of African Americans living in the Sea Islands of South Carolina since the early 1700s. African descendants comprised the majority of the population of South Carolina until the Great Migration to northern industrial cities in 1910 (Rogers & Taylor, 1994). The demand for enslaved Africans to work in the rice fields, and later in the cultivation of indigo and cotton, was very high through the 18th century and into the 19th century. As a result, there was a large influx of Africans into South Carolina and Georgia since the early colonial period (Parra et al., 2001). In the first federal census of 1790, enslaved Africans comprised 18% of the nation's total population, but ranged from 47–93% in several coastal areas of South Carolina, including the port of Charleston, the center of American slavery (Parra et al., 2001). During that time, while South Carolina had 43% enslaved Africans, the percentage for the Beaufort and Charleston Districts was much higher (76%) and that for the parish including the Sea Islands even higher (93%) (Fredrickson, 1981). Moreover, on the Sea Islands, all of the plantation owners who could leave the plantations from late May to late June would do so to avoid the risk of contracting malaria. The absence of planters during this

period gave African descended individuals considerably more autonomy in developing their own culture there (Pollitzer, 1999).

Later, in 1860, when the enslaved African population in the US declined to 13%, that of South Carolina had risen to 57%. The approximately equal number of male and female slaves in these districts by 1810 suggests that the increase was due to reproduction of local populations rather than the importation of additional individuals from Africa. The increase in the number of Africans, their concentration in rural areas, the severity of slave codes, and the social alienation of Africans from Europeans, all contributed to the isolation of 18th century Gullah people. These conditions also provided an ideal context for creolization and the development of distinctive cultural attributes that continued into the 19th century and beyond (Pollitzer, 1999).

The higher mean level of African ancestry in the Gullah could be a direct effect of a potentially higher proportion of African Americans currently living in this region. Currently, in counties where the Gullah community members sampled for this study reside, African Americans encompass 20–50% of the population (United States Census Bureau, 2011). By contrast, the proportion of African Americans in the tri-county area sampled for the Jackson Heart Study ranges from

20 to 70% (United States Census Bureau, 2011). The slightly higher proportion of African Americans in the Jackson tri-county area suggests that the reported differences in population structure were not simply the result of current differential population proportions. Rather, the higher mean level of African ancestry in the Gullah reflects the historically higher proportion of African Americans living in the Sea Islands of South Carolina from the early 1700s through the mid-1900s.

Finally, since African Americans living in rural areas have a higher average African ancestry than those living in urban areas (Baharian et al., 2016), we further considered the effects of sampling on ancestry proportions. The Gullah are an intrinsically rural community, while JHS participants represent urban dwellers. Baharian et al. (2016) report that, for both African Americans sampled only in rural, or in both urban and rural regions, the average African ancestry proportions are higher in South Carolina than in Mississippi. In support of this study, the proportion of African ancestry in JHS (82%) is similar to that reported by Baharian et al. (2016) in their urban and rural samples from Mississippi (83%), while the proportion of African ancestry in Gullah (91%) is slightly higher than that reported in their rural samples from South Carolina (88%) (Baharian et al., 2016). The 88% African admixture estimate in rural samples from South Carolina (Baharian et al., 2016) falls outside of the 95% confidence interval around the mean African ancestry of the Gullah ([90.6–90.8%]), suggesting that the African ancestry in the Gullah is significantly different than that for rural non-Gullah African Americans in South Carolina. We thus infer that the slightly higher mean level of African ancestry in the Gullah might be an effect of their rural sampling, but also likely arose because of historical and sociocultural factors.

### 3.3 | Native American ancestry estimates in Gullah African Americans

Consistent with early mtDNA and Y-chromosome studies of Gullah African Americans (Parra et al., 2001), we found a small Native American contribution to the African American groups sampled for our study. We observed that the Gullah and JHS African Americans had slightly higher Native American ancestry than (~1.3–1.4%) had been reported in most African American groups in the U.S. (Baharian et al., 2016; Bryc et al., 2015; Dai et al., 2020; Mathias et al., 2016; Micheletti et al., 2020). This discrepancy may reflect the fact that previous studies have often used clustering methods like ADMIXTURE for estimating Native American ancestry proportions which are expected to give underestimates when the proxy population used for Native American ancestry (typically Mesoamerican) is highly genetically drifted from the true source population (Southeastern U.S. Native American). In contrast, the qpAdm ancestry estimation procedure explicitly accounts for genetic drift between the source population and the proxy population and produces an unbiased estimate. In the U.S., only African Americans living in the Southwest U.S. (ASW) from the 1000 Genomes Project had higher Native American ancestry (3.1%) (Martin et al., 2017).

This level of Native American ancestry is consistent with historical records about the Native American slave trade (Gallay, 2010; Pollitzer, 1999). In the early days of the American colonies, marriages were permitted between Europeans, Africans, and Native Americans (Pollitzer, 1999). Despite a 1671 law forbidding Native American slavery, Native Americans were publicly sold as slaves in Charleston, and their enslavement by colonists was common until the African slave trade accelerated in the 18th century. In fact, from 1670 to 1720, more Native Americans were shipped out of Charleston than Africans were imported (Gallay, 2010). After the Yamasee War (1715–17), Native American populations (Yamasee, Ochese, Waxhaw, Santee) declined in South Carolina, and most of the remaining Native American slaves were apparently absorbed into the African community (Pollitzer, 1999). The offspring of Africans and Native Americans were called *mustizoes* or *mustees*, in contrast to the *mulattoes* resulting from the union of Africans and Europeans.

From the 1730s through the 1780s, newspaper ads proclaimed that 2424 slaves had run away from their masters in South Carolina. As noted in 27% of the ads, a surprising 37% of runaways were described as being light, yellow, or mulatto in appearance, and 19% of them were said to be mustees (Pollitzer, 1999). The cultural influences of Native Americans are reflected in Gullah crafts, colono-ware, boat building techniques, and decoctions of healing herbs used to cope with illness (Pollitzer, 1999). Further support for Native American admixture in the South comes from the several socially distinct communities with European, African, and American Indian ancestry that have persisted to the present day (e.g., Brass Ankles and Turks in South Carolina). In summary, historical, ethnographic, and mtDNA and Y-chromosome data support a Native American contribution to the Gullah (Gallay, 2002; Gallay, 2010; Parra et al., 2001; Pollitzer, 1999), and this contribution is confirmed by our genome-wide data.

### 3.4 | Sex-biased admixture in Gullah African Americans

To aid with interpretation, the proportion of ancestry that comes from males was estimated under a simple model in which, in a population with equally many females and males, the mean X-chromosomal admixture fraction is a linear combination of female and male admixture parameters, with coefficients 2/3 and 1/3, respectively. We found evidence for patterns of sex-biased gene flow in the Gullah (Table 1), consistent with the reported higher male European and female African contributions in other U.S. African Americans (Baharian et al., 2016; Bryc et al., 2015; Dai et al., 2020; Mathias et al., 2016; Micheletti et al., 2020; Patin et al., 2017), as well as previous work with the Gullah (Parra et al., 2001). The evidence for sex-biased admixture is also in agreement with the smaller European paternal ancestry of African Americans in South Carolina relative to other North American and Caribbean groups reported in a large Y-chromosome study (Torres, Doura, Keita, & Kittles, 2012).

The decrease in European ancestry on the X-chromosome might imply a simultaneous European male bias and African female bias,

which is consistent with increased frequency of sexual interactions between European males and African females, including rape and/or coerced sexual interactions (Kennedy, 2003; Lind et al., 2007). However, as shown by Goldberg and Rosenberg (2015), the difference in X chromosomal and autosomal admixture might also be the result of male biases in both Europeans and Africans. This interpretation would be consistent with the aforementioned asymmetric mating practices involving African females and European males, as well as with the overrepresentation of males among African slaves brought to North America (about 70%) (Davis, 2001; Eltis & Richardson, 2010; Fredrickson, 1981; Painter, 2006) (e.g., see *Trans-Atlantic Slave Trade Database* website). Notably, these results show that Gullah and JHS African Americans have differing degrees of sex-biased ancestry contributions, with the Gullah exhibiting a greater male-biased European contribution (Table 1). These results are consistent with the negligible European female and higher European male contributions previously noted in the Gullah (Parra et al., 2001).

As recently reported (Micheletti et al., 2020), the extent of this sex bias towards European male and African female genetic contributions is known to vary across the Americas due to regional differences in slavery practices. Despite the lack of direct ethnographic records on mating patterns for the Gullah, the evidence that virtually all European X-chromosomes came from men suggests that few, if any, European descended women produced children in the Gullah community. Plantation owners favored males and children for commodity labor, whereas enslaved female were often situated in more domestic contexts to bear and raise children, while also being sold at greater frequency within the internal slave market in the U.S. (Davis, 1995; Joyner, 1985; Malone, 1992; Rosengarten, Chaplin, & Walker, 1986; Wood, 1974). Yet, given the complex mechanistic models of historical admixture (Goldberg & Rosenberg, 2015), the explanation for our genetic results will require further substantiation by other studies.

### 3.5 | African ancestry of Gullah African Americans

We next tried to elucidate whether the genetic data supported a postulated Sierra Leone (Opala, 1987) or diverse African (Pollitzer, 1999) ancestry for the Gullah. Ancestry estimates (Figure 2) suggested that, relative to JHS African Americans, the Gullah had comparable Yoruba ancestry and higher ancestry from the African Rice Coast (from Senegal down to Liberia). As shown in Figure 3, a gradient in the clustering of the Gullah and non-Gullah African Americans indicated the Gullah's relative proximity to the Sierra Leone (especially Creole) and Mandenka samples, while non-Gullah African Americans' appeared closer to the Yoruba. A PerMANOVA revealed significant differences ( $p < 0.001$ ) in the principal components for each African American group, and analysis of the individual principal components demonstrated large shifts in PC1 and PC2 (Figure 3). Thus, although not forming individual clusters, the Gullah and non-Gullah samples were distributed along a gradient in this PCA, with the Gullah samples showing closer Sierra Leone relatedness than the JHS samples.

Close relatives are expected to share large identical-by-descent (IBD) segments, which can then be used to model recent ancestry and

elucidate population-level relatedness. Analysis of the mean number of shared IBD segments between pairs of Gullah and JHS African American individuals confirmed that, relative to JHS African Americans, Gullah individuals had a lower mean number of shared European segments and a higher number of shared African segments, including a slightly higher proportion of segments of Mandenka ancestry (Figure S7).

Furthermore, given the relationship between fixation index ( $F_{ST}$ ) estimates and admixture levels (Boca & Rosenberg, 2011), we computed the  $F_{ST}$  (Weir & Cockerham, 1984) to assess the genetic divergence between the admixed groups and their parental source populations. The  $F_{ST}$  estimates were smaller between the Gullah and the Yoruba and Sierra Leone populations than between the JHS African Americans and these same African populations (Table S3). The closeness to African populations parallels the higher African ancestry of the Gullah relative to the JHS African Americans. For both Gullah and non-Gullah African Americans,  $F_{ST}$  estimates for either Sierra Leone or Yoruba were similar, confirming the similar closeness of each African American group to both African populations. Collectively, these data support the closeness of the Gullah to putative ancestral African populations and the interpretation that the Gullah are not direct and exclusive descendants of populations from Sierra Leone. Instead, as postulated by Pollitzer (1999), the Gullah share a common ancestry with numerous populations from Sierra Leone and other regions of West Africa.

These results are consistent with the recently reported genetic ancestries of African Americans from the Southeast U.S. (Micheletti et al., 2020). In this large and representative cohort, Micheletti and colleagues found that Southeast African Americans had the highest African ancestry from Nigeria (26–30%), followed by Coastal West Africa (Sierra Leone and the Windward Coast, [~18%]), West Central Africa (~8%), and Senegambia (~7%). There are obvious discordances with the proportions of Africans that arrived in Charleston through the legal slave trade, who were mostly from West Central Africa (~39%), the Windward Coast and Sierra Leone (~23%), and Senegambia (~20%), and only ~5% from the Bights of Benin and Biafra (Pollitzer, 1999). The apparent overrepresentation of Nigerian ancestry can be explained by the trade of enslaved people from the British Caribbean (Micheletti et al., 2020), and the ethnic composition of Africans imported into the British West Indies indicating source areas in the Gold Coast and the Bights of Benin and Biafra (Pollitzer, 1999). On the other hand, the underrepresentation of Senegambian ancestry can be explained by accounts of early trading and high mortality from this region (Micheletti et al., 2020).

### 3.6 | Effects of geographic isolation on Gullah African Americans

Since the Gullah have remained a relatively isolated group over the past few centuries, we sought to determine whether this isolation has affected the genetic structure of their populations. Hallmarks of isolated populations include increased frequencies of recessive disorders, reduced genetic diversity, and higher identity-by-descent (IBD) as the result of founder events and population bottlenecks. There are no reports of the increased frequency of any recessive disorders in the

**TABLE 3** Genetic diversity in Gullah African American and Sierra Leone populations

Population	HETexp	HETobs	F
Gullah	0.332	0.333 <sup>†</sup>	-0.0018 <sup>†</sup>
Sierra Leone	0.332	0.334	-0.0045

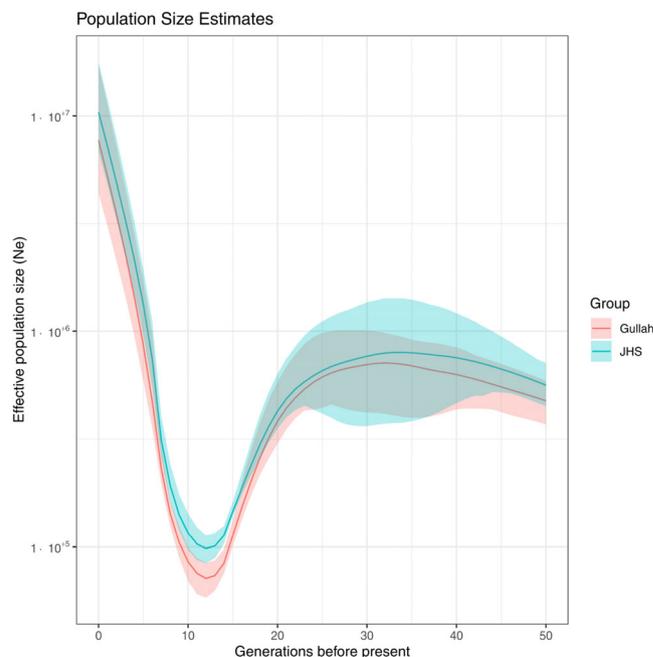
Note: Expected heterozygosity (HETexp), observed heterozygosity (HETobs), and inbreeding coefficient (F) for the Gullah African American and Sierra Leone populations.

<sup>†</sup> $p < 0.01$  compared with the Sierra Leone population (Wilcoxon test).

Gullah that would support the occurrence of founder events. We first compared the genetic diversity of the Gullah and Sierra Leone populations by measuring mean heterozygosity and inbreeding coefficients. We observed very similar, though slightly lower level of heterozygosity ( $p = 2.72 \times 10^{-3}$ ) (Table 3 and Table S4) and higher inbreeding coefficient ( $p = 8.78 \times 10^{-3}$ ) (Table 3 and Table S5, Figure S8), in Gullah compared with Sierra Leone individuals. The similarity in heterozygosity noted in Gullah and Sierra Leone individuals, despite the Gullah's admixture with European individuals, is not unexpected given their low levels of European admixture. This is also the case because the genetic divergence between two West African chromosomes is similar to that between a West African and a European chromosome.

A comparison of the different proportions of IBD segments shared in the Gullah and the JHS African Americans showed a lower mean number of shared European segments and a higher number of shared African segments in the Gullah (Figure S7). This increased number of long founding African haplotypes in the Gullah supports their increased proximity to West African populations, and is consistent with their relative geographic isolation (Pollitzer, 1999). When using IBDNe to explore IBD in African American samples over recent generations (Browning et al., 2018), we noted that the estimated effective population sizes, based only on the African ancestry-associated IBD segments, in African American groups were mostly similar to each other (Figure 4). This result suggested historical mixing within the larger African ancestry population that encompasses these groups, causing the two Southeast African American groups to have a shared demographic history.

This analysis further revealed a bottleneck event 13 generations ago for both groups (Figure 4), an estimate consistent with the turn of the 18th century. We infer that these bottlenecks mostly resulted from migration and death during the enslavement process. At the same time, these data also support somewhat different demographic histories for the two populations. Based on the African ancestry-associated IBD segments, Gullah have a slightly tighter bottleneck and lower estimated current effective size than JHS, a finding that is consistent with the higher number of long African segments of IBD among Gullah than JHS (Figure S7). Non-Gullah African Americans from the Southeast US might therefore have a more diverse geographical origin on average than the Gullah, leading to the larger current estimated effective population size. However, the more apparent bottleneck in the Gullah could also be a bias due to the reduced European admixture in the Gullah. Collectively, these results are consistent with the relative isolation of the Gullah.



**FIGURE 4** Estimation of ancestry-specific recent effective population size from segments of identity by descent (IBD) in Gullah and JHS African Americans. Plot displays the recent effective population size (Ne) in Gullah and JHS African Americans over the past 50 generations. The lines show the estimated effective population size based on IBD segments associated only with African ancestry, while the colored regions show 95% bootstrap confidence intervals. The graph displays a bottleneck event occurring nearly 13 generations ago, an estimate consistent with the turn of the 18th century. We used the ancestry-specific IBDNe pipeline to estimate the effective population sizes

## 4 | CONCLUSION

Our study helps clarify the debated ancestry of Gullah African Americans and makes clear that Sierra Leone is not the sole origin point. We confirmed their higher African, lower European, and small Native American ancestries, as well as a larger proportion of male-biased European admixture. We show that the Gullah have a diverse African ancestry, with increased proximity to West African populations than Southeastern non-Gullah African Americans. We found genomic evidence for a slightly tighter bottleneck in the Gullah consistent with a founder event(s) upon importation to the U.S. These findings are consistent with historical, cultural, and anthropological evidence indicating that their relative geographical isolation and strong community life allowed the Gullah to preserve many aspects of their African cultural heritage (Pollitzer, 1999). Although the subtle genetic differences relative to Southeastern non-Gullah African Americans support somewhat different demographic histories, these results also reveal largely shared common ancestries. As such, our data shows that the Gullah are not a genetically distinct group per se, but rather a culturally distinct group of African Americans with subtle variation in its genetic structure.

Broadly, this study reveals subtle differences in genetic structure and ancestry in African American populations. These differences can

have important implications for precision medicine, and further reveal the crucial need to include more ancestrally diverse individuals in medical genomic studies (Dai et al., 2020). Only a comprehensive understanding of the genetic architecture of these populations can ensure that they are not omitted from developments in new genetic technologies and clinical advancements, ultimately contributing to the closure of the health disparities gap as healthcare moves towards precision medicine. Finally, this study is important for Gullah and non-Gullah African Americans who were stripped of ancestral identities by the slave trade. Combined with socio-historical resources, this research can help to recover ancestral histories, and contribute to their new collective identities and ties to ancestral homelands, ultimately paving the road towards transforming lives and possibly reconciliation (Nelson, 2016).

## ACKNOWLEDGMENTS

We thank the Sea Islands Family Project Citizen Advisory Committee (SIFT-CAC) for their enthusiasm and support. We are grateful to the participants from Projects SuGAR, SLEIGH, COBRE and JHS. We thank Yiqi Huang and Kasia Bryc for their contributions to early design, analysis, drafting and revision of this work, as well as Kerry Lok, Xuanlin Hou, Fang Chen, and Yin Lin for technical and analytical assistance. We thank Fonte Felipe for historical insights. David Reich is an Investigator of the Howard Hughes Medical Institute. This work was supported by National Institutes of Health (NIH) research grants R01 DK084350 (MMS), K01 AR067280 (PSR), P20 RR-017696 and M01 RR001070 (South Carolina COBRE for Oral Health), P60 AR049459 (GSG and DLK), P30 AR072582 (GSG, DLK, PSR), UL1 RR029882 (GSG and DLK), and the W. M. Keck Foundation (WTG).

## CONFLICT OF INTEREST

The authors declare no conflicts of interest.

## AUTHOR CONTRIBUTIONS

**Kip Zimmerman:** Data curation; formal analysis; investigation; validation; visualization; writing-original draft; writing-review & editing. **Theodore Schurr:** Investigation; writing-original draft; writing-review & editing. **Wei-Min Chen:** Conceptualization; data curation; investigation; methodology; writing-original draft; writing-review & editing. **Uma Nayak:** Data curation; investigation; writing-original draft; writing-review & editing. **Josyf Mychaleckyj:** Investigation; writing-original draft; writing-review & editing. **Queen Quet:** Investigation; writing-original draft; writing-review & editing. **Lee Moultrie:** Investigation; writing-original draft; writing-review & editing. **Jasmin Divers:** Data curation; investigation; writing-original draft; writing-review & editing. **Keith Keene:** Investigation; writing-original draft; writing-review & editing. **Diane Kamen:** Funding acquisition; resources; writing-original draft; writing-review & editing. **Gary Gilkeson:** Funding acquisition; resources; writing-original draft; writing-review & editing. **Kelly Hunt:** Investigation; resources; writing-original draft; writing-review & editing. **Ida Spruill:** Investigation; resources; writing-original draft; writing-review & editing. **Jyotika Fernandes:** Funding acquisition; resources; writing-original draft; writing-review & editing. **Melinda Aldrich:** Investigation; validation; writing-original draft; writing-review & editing. **David Reich:** Conceptualization; formal analysis; investigation;

methodology; writing-original draft; writing-review & editing. **W. Timothy Garvey:** Conceptualization; funding acquisition; investigation; resources; validation; writing-original draft; writing-review & editing. **Carl Langefeld:** Conceptualization; formal analysis; investigation; methodology; validation; writing-original draft; writing-review & editing. **Michele Sale:** Conceptualization; funding acquisition; investigation; project administration; resources; supervision; writing-original draft; writing-review & editing. **Paula Ramos:** Conceptualization; funding acquisition; investigation; project administration; supervision; validation; writing-original draft; writing-review & editing.

## DATA AVAILABILITY STATEMENT

All genotypic data used in this study have been deposited in public databases. Gullah African American and African (Sierra Leone) data can be accessed through dbGAP with the accession code phs000433.v1.p1 (The Sea Islands Genetic Network (SIGNET)).

## ORCID

Theodore G. Schurr  <https://orcid.org/0000-0001-9323-9237>

Melinda C. Aldrich  <https://orcid.org/0000-0003-3833-8448>

David Reich  <https://orcid.org/0000-0002-7037-5292>

Paula S. Ramos  <https://orcid.org/0000-0002-9433-6634>

## REFERENCES

- Abraham, A. (2003). *An introduction to the Precolonial history of the Mende of Sierra Leone*. Edwin Mellen Press.
- Alexander, D. H., Novembre, J., & Lange, K. (2009). Fast model-based estimation of ancestry in unrelated individuals. *Genome Research*, 19(9), 1655–1664.
- Alie, J. A. D. (1990). *A new history of Sierra Leone*. Macmillan Education.
- Altshuler, D. M., Gibbs, R. A., Peltonen, L., Dermitzakis, E., Schaffner, S. F., Yu, F., Bonnen, P. E., de Bakker, P. I., Deloukas, P., Gabriel, S. B., et al. (2010). Integrating common and rare genetic variation in diverse human populations. *Nature*, 467(7311), 52–58.
- 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., & Gravel, S. (2016). The great migration and African-American genomic diversity. *PLoS Genetics*, 12(5), e1006059.
- Bangura, J. J. (2017). *The Temne of Sierra Leone: African Agency in the Making of a British Colony*. Cambridge University Press.
- Boca, S. M., & Rosenberg, N. A. (2011). Mathematical properties of Fst between admixed populations and their parental source populations. *Theoretical Population Biology*, 80(3), 208–216.
- Brady, P. S. (1972). The slave trade and sectionalism in South Carolina. *The Journal of Southern History*, 38(4), 601–620.
- Browning, S. R., & Browning, B. L. (2007). Rapid and accurate haplotype phasing and missing-data inference for whole-genome association studies by use of localized haplotype clustering. *American Journal of Human Genetics*, 81(5), 1084–1097.
- Browning, S. R., Browning, B. L., Daviglus, M. L., Durazo-Arvizu, R. A., Schneiderman, N., Kaplan, R. C., & Laurie, C. C. (2018). Ancestry-specific recent effective population size in the Americas. *PLoS Genetics*, 14(5), e1007385.
- Bryc, K., Durand, E. Y., Macpherson, J. M., Reich, D., & Mountain, J. L. (2015). The genetic ancestry of African Americans, Latinos, and European Americans across the United States. *American Journal of Human Genetics*, 96(1), 37–53.
- Cassidy, F. G. (1980). The place of Gullah. *American Speech*, 55(1), 3–16.
- Chang, C. C., Chow, C. C., Tellier, L. C., Vattikuti, S., Purcell, S. M., & Lee, J. J. (2015). Second-generation PLINK: Rising to the challenge of larger and richer datasets. *GigaScience*, 4, 7.

- Cole, G. R. (2013). *The Krio of West Africa: Islam, culture, creolization, and colonialism in the nineteenth century*. Ohio University Press.
- Dai, C. L., Vazifeh, M. M., Yeang, C. H., Tachet, R., Wells, R. S., Vilar, M. G., Daly, M. J., Ratti, C., & Martin, A. R. (2020). Population histories of the United States revealed through fine-scale migration and haplotype analysis. *American Journal of Human Genetics*, 106(3), 371–388.
- 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. (2011). The variant call format and VCFtools. *Bioinformatics*, 27(15), 2156–2158.
- Davis, A. (1995). Reflections on the black woman's role in the community of slaves. In B. Guy-Sheftall (Ed.), *Words of fire: An anthology of African-American feminist thought* (pp. 200–218). The New Press.
- Davis, F. J. (2001). *Who is black? One nation's definition*. The Pennsylvania State University Press.
- Divers, J., Sale, M. M., Lu, L., Chen, W. M., Lok, K. H., Spruill, I. J., Fernandes, J. K., Langefeld, C. D., & Garvey, W. T. (2010). The genetic architecture of lipoprotein subclasses in Gullah-speaking African American families enriched for type 2 diabetes: The Sea Islands Genetic African American Registry (Project SuGAR). *Journal of Lipid Research*, 51(3), 586–597.
- Dixon-Fyle, M., & Cole, G. R. (2006). Introduction. In *New perspectives on the Sierra Leone Krio* (pp. 2–3). Peter Lang.
- Durand, E. Y., Eriksson, N., & McLean, C. Y. (2014). Reducing pervasive false-positive identical-by-descent segments detected by large-scale pedigree analysis. *Molecular Biology and Evolution*, 31(8), 2212–2222.
- Dwyer, D. (2005). The Mende problem. In A. G. B. Koen & J. Maniacky (Eds.), *Studies in comparative linguistics with special focus on Bantu and Mande* (pp. 29–42). Tervuren.
- Eltis, D., & Richardson, D. (2010). *Atlas of the transatlantic slave trade*. Yale University Press.
- Fage, J. D. (2001). *History of Africa*. Routledge.
- Fernandes, J. K., Wiegand, R. E., Salinas, C. F., Grossi, S. G., Sanders, J. J., Lopes-Virella, M. F., & Slate, E. H. (2009). Periodontal disease status in Gullah African Americans with type 2 diabetes living in South Carolina. *Journal of Periodontology*, 80(7), 1062–1068.
- Finnegan, R. H. (1965). *Survey of the Limba people of Northern Sierra Leone*. Her Majesty's Stationery Office.
- Fredrickson, G. M. (1981). *White supremacy: A comparative study of American and south African history*. Oxford University Press.
- Fyfe, C. (1979). *A short history of Sierra Leone*. Longman Group.
- Gallay, A. (2002). *The Indian slave trade: The rise of the English empire in the American South, 1670–1717*. Yale University Press.
- Gallay, A. (2010). *Indian slavery in the Americas. History now: America online*. The Gilder Lehrman Institute of American History Retrieved from <http://www.gilderlehrman.org/history-by-era/origins-slavery/essays/indian-slavery-americas>
- Goldberg, A., & Rosenberg, N. A. (2015). Beyond 2/3 and 1/3: The complex signatures of sex-biased admixture on the X chromosome. *Genetics*, 201(1), 263–279.
- Gomez, F., Hirbo, J., & Tishkoff, S. A. (2014). Genetic variation and adaptation in Africa: Implications for human evolution and disease. *Cold Spring Harbor Perspectives in Biology*, 6(7), a008524.
- Gusev, A., Lowe, J. K., Stoffel, M., Daly, M. J., Althuler, D., Breslow, J. L., Friedman, J. M., & Pe'er, I. (2009). Whole population, genome-wide mapping of hidden relatedness. *Genome Research*, 19(2), 318–326.
- Haak, W., Lazaridis, I., Patterson, N., Rohland, N., Mallick, S., Llamas, B., Brandt, G., Nordenfelt, S., Harney, E., Stewardson, K., Fu, Q., Mittnik, A., Bánffy, E., Economou, C., Francken, M., Friederich, S., Pena, R. G., Hallgren, F., Khartanovich, V., ... Reich, D. (2015). Massive migration from the steppe was a source for Indo-European languages in Europe. *Nature*, 522(7555), 207–211.
- Han, E., Carbonetto, P., Curtis, R. E., Wang, Y., Granka, J. M., Byrnes, J., Noto, K., Kermany, A. R., Myres, N. M., Barber, M. J., et al. (2017). Clustering of 770,000 genomes reveals post-colonial population structure of North America. *Nature Communications*, 8, 14238.
- Hancock, I. F. (1969). A provisional comparison of the English-derived Atlantic creoles. *African Language Review*, 8, 7–72.
- Henn, B. M., Botigue, L. R., Gravel, S., Wang, W., Brisbin, A., Byrnes, J. K., Fadhlaoui-Zid, K., Zalloua, P. A., Moreno-Estrada, A., Bertranpetit, J., et al. (2012). Genomic ancestry of North Africans supports back-to-Africa migrations. *PLoS Genetics*, 8(1), e1002397.
- Jackson, B. A., Wilson, J. L., Kirbah, S., Sidney, S. S., Rosenberger, J., Bassie, L., Alie, J. A., McLean, D. C., Garvey, W. T., & Ely, B. (2005). Mitochondrial DNA genetic diversity among four ethnic groups in Sierra Leone. *American Journal of Physical Anthropology*, 128(1), 156–163.
- Joyner, C. (1985). *Down by the riverside: A South Carolina slave community*. University of Illinois Press.
- Kamen, D. L., Barron, M., Parker, T. M., Shaftman, S. R., Bruner, G. R., Aberle, T., James, J. A., Scofield, R. H., Harley, J. B., & Gilkeson, G. S. (2008). Autoantibody prevalence and lupus characteristics in a unique African American population. *Arthritis and Rheumatism*, 58(5), 1237–1247.
- Kennedy, R. (2003). *Interracial intimacies: Sex, marriage, identity, and adoption*. Vintage Books.
- Knörr, J., & Kohl, C. (2016). *The upper Guinea coast in global perspective*. Berghahn Books.
- Li, J. Z., Absher, D. M., Tang, H., Southwick, A. M., Casto, A. M., Ramachandran, S., Cann, H. M., Barsh, G. S., Feldman, M., Cavalli-Sforza, L. L., & Myers, R. M. (2008). Worldwide human relationships inferred from genome-wide patterns of variation. *Science*, 319(5866), 1100–1104.
- 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., & Smith, M. W. (2007). Elevated male European and female African contributions to the genomes of African American individuals. *Human Genetics*, 120(5), 713–722.
- Loh, P. R., Lipson, M., Patterson, N., Moorjani, P., Pickrell, J. K., Reich, D., & Berger, B. (2013). Inferring admixture histories of human populations using linkage disequilibrium. *Genetics*, 193(4), 1233–1254.
- Malone, A. P. (1992). *Sweet chariot: Slave family and household structure in nineteenth-century Louisiana*. University of North Carolina Press.
- Manichaikul, A., Mychaleckyj, J. C., Rich, S. S., Daly, K., Sale, M., & Chen, W. M. (2010). Robust relationship inference in genome-wide association studies. *Bioinformatics*, 26(22), 2867–2873.
- Maples, B. K., Gravel, S., Kenny, E. E., & Bustamante, C. D. (2013). RFMix: A discriminative modeling approach for rapid and robust local-ancestry inference. *American Journal of Human Genetics*, 93(2), 278–288.
- Martin, A. R., Gignoux, C. R., Walters, R. K., Wojcik, G. L., Neale, B. M., Gravel, S., Daly, M. J., Bustamante, C. D., & Kenny, E. E. (2017). Human demographic history impacts genetic risk prediction across diverse populations. *American Journal of Human Genetics*, 100(4), 635–649.
- 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. (2016). A continuum of admixture in the Western Hemisphere revealed by the African Diaspora genome. *Nature Communications*, 7, 12522.
- McLean, D. C., Jr., Spruill, I., Argyropoulos, G., Page, G. P., Shriver, M. D., & Garvey, W. T. (2005). Mitochondrial DNA (mtDNA) haplotypes reveal maternal population genetic affinities of Sea Island Gullah-speaking African Americans. *American Journal of Physical Anthropology*, 127(4), 427–438.
- McLean, D. C., Jr., Spruill, I., Gevao, S., Morrison, E. Y., Bernard, O. S., Argyropoulos, G., & Garvey, W. T. (2003). Three novel mtDNA restriction site polymorphisms allow exploration of population affinities of African Americans. *Human Biology*, 75(2), 147–161.
- Micheletti, S. J., Bryc, K., Ancona Esselmann, S. G., Freyman, W. A., Moreno, M. E., Poznik, G. D., Shastri, A. J., and Me Research Team, Beleza, S., & Mountain, J. L. (2020). Genetic consequences of the transatlantic slave trade in the Americas. *American Journal of Human Genetics*, 107(2), 265–277.
- Moreno-Estrada, A., Gignoux, C. R., Fernandez-Lopez, J. C., Zakharia, F., Sikora, M., Contreras, A. V., Acuna-Alonzo, V., Sandoval, K., Eng, C.,

- Romero-Hidalgo, S., Ortiz-Tello, P., Robles, V., Kenny, E. E., Nuno-Arana, I., Barquera-Lozano, R., Macin-Perez, G., Granados-Arriola, J., Huntsman, S., Galanter, J. M., ... Bustamante, C. D. (2014). Human genetics. The genetics of Mexico recapitulates native American substructure and affects biomedical traits. *Science*, 344(6189), 1280–1285.
- Musunuru, K., Lettre, G., Young, T., Farlow, D. N., Pirruccello, J. P., Ejebe, K. G., Keating, B. J., Yang, Q., Chen, M. H., Lapchyk, N., Crenshaw, A., Ziaugra, L., Rachupka, A., Benjamin, E. J., Cupples, L. A., Fornage, M., Fox, E. R., Heckbert, S. R., Hirschhorn, J. N., ... NHLBI Candidate Gene Association Resource. (2010). Candidate gene association resource (CARE): Design, methods, and proof of concept. *Circulation. Cardiovascular Genetics*, 3(3), 267–275.
- Nash, R. (2010). South Carolina indigo, European textiles, and the British Atlantic economy in the eighteenth century. *The Economic History Review*, 63(2), 362–392.
- Nelson, A. (2008). Bio science: Genetic genealogy testing and the pursuit of African ancestry. *Social Studies of Science*, 38(5), 759–783.
- Nelson, A. (2016). *The social life of DNA: Race, reparations, and reconciliation after the genome*. Beacon Press.
- Ogot, B. A. (1992). *General history of Africa, Vol. V: Africa from the sixteenth to the eighteenth century*. UNESCO: University of California Press.
- Opala, J. (1987). *The Gullah: Rice, slavery, and the Sierra Leone-American connection*. US Information Service: Fort Sumter National Monument.
- Oyètàdè, B. A., & Fashole-Luke, V. (2008). Sierra Leone: Krio and the quest for national integration. In A. Simpson (Ed.), *Language and national identity in Africa* (pp. 122–140). Oxford University Press.
- Painter, N. I. (2006). *Creating black Americans: African-American history and its meanings, 1619 to the present*. Oxford University Press.
- Parra, E. J., Kittles, R. A., Argyropoulos, G., Pfaff, C. L., Hiester, K., Bonilla, C., Sylvester, N., Parrish-Gause, D., Garvey, W. T., Jin, L., McKeigue, P. M., Kamboh, M. I., Ferrell, R. E., Pollitzer, W. S., & Shriver, M. D. (2001). Ancestral proportions and admixture dynamics in geographically defined African Americans living in South Carolina. *American Journal of Physical Anthropology*, 114(1), 18–29.
- Parra, E. J., Marcini, A., Akey, J., Martinson, J., Batzer, M. A., Cooper, R., Forrester, T., Allison, D. B., Deka, R., Ferrell, R. E., & Shriver, M. D. (1998). Estimating African American admixture proportions by use of population-specific alleles. *American Journal of Human Genetics*, 63(6), 1839–1851.
- Patin, E., Lopez, M., Grollemund, R., Verdu, P., Harmant, C., Quach, H., Laval, G., Perry, G. H., Barreiro, L. B., Froment, A., Heyer, E., Massougbdji, A., Fortes-Lima, C., Migot-Nabias, F., Bellis, G., Dugoujon, J. M., Pereira, J. B., Fernandes, V., Pereira, L., ... Quintana-Murci, L. (2017). Dispersals and genetic adaptation of Bantu-speaking populations in Africa and North America. *Science*, 356(6337), 543–546.
- Patterson, N., Moorjani, P., Luo, Y., Mallick, S., Rohland, N., Zhan, Y., Genschoreck, T., Webster, T., & Reich, D. (2012). Ancient admixture in human history. *Genetics*, 192(3), 1065–1093.
- Patterson, N., Price, A. L., & Reich, D. (2006). Population structure and eigenanalysis. *PLoS Genetics*, 2(12), e190.
- Pollitzer, W. S. (1999). *The Gullah people and their African heritage*. The University of Georgia Press.
- Raghavan, M., Steinrucken, M., Harris, K., Schiffels, S., Rasmussen, S., DeGiorgio, M., Albrechtsen, A., Valdiosera, C., Avila-Arcos, M. C., Malaspina, A. S., Eriksson, A., Moltke, I., Metspalu, M., Homburger, J. R., Wall, J., Cornejo, O. E., Moreno-Mayar, J. V., Korneliusen, T. S., Pierre, T., ... Willerslev, E. (2015). Population genetics. Genomic evidence for the Pleistocene and recent population history of Native Americans. *Science*, 349(6250), aab3884.
- Reiner, A. P., Ziv, E., Lind, D. L., Nievergelt, C. M., Schork, N. J., Cummings, S. R., Phong, A., Burchard, E. G., Harris, T. B., Psaty, B. M., & Kwok, P. Y. (2005). Population structure, admixture, and aging-related phenotypes in African American adults: The cardiovascular health study. *American Journal of Human Genetics*, 76(3), 463–477.
- Rodney, W. (1970). *A history of the Upper Guinea Coast, 1545–1800*. Monthly Review Press.
- Rogers, G. C. J., & Taylor C. J. (1994). *A South Carolina chronology, 1497–1992*. University of South Carolina Press.
- Rosengarten, T., Chaplin, T., & Walker, S. (1986). *Tombee: Portrait of a cotton planter*. William Morrow & Company.
- Rotimi, C. N. (2003). Genetic ancestry tracing and the African identity: A double-edged sword? *Developing World Bioethics*, 3(2), 151–158.
- Sale, M. M., Lu, L., Spruill, I. J., Fernandes, J. K., Lok, K. H., Divers, J., Langefeld, C. D., & Garvey, W. T. (2009). Genome-wide linkage scan in Gullah-speaking African American families with type 2 diabetes: The Sea Islands Genetic African American Registry (Project SuGAR). *Diabetes*, 58(1), 260–267.
- Shriner, D., Tekola-Ayele, F., Adeyemo, A., & Rotimi, C. N. (2014). Genome-wide genotype and sequence-based reconstruction of the 140,000 year history of modern human ancestry. *Scientific Reports*, 4, 6055.
- Spruill, I. J., Leite, R. S., Fernandes, J. K., Kamen, D. L., Ford, M. E., Jenkins, C., Hunt, K. J., & Andrews, J. O. (2013). Successes, challenges and lessons learned: Community-engaged research with South Carolina's Gullah population. *Gateways*, 6. <https://doi.org/10.5130/ijcre.v6i1.2805>
- Taylor, B. K. (2014). *Sierra Leone: The Land, Its People and History*. New Africa Press.
- Taylor, H. A., Jr., Wilson, J. G., Jones, D. W., Sarpong, D. F., Srinivasan, A., Garrison, R. J., Nelson, C., & Wyatt, S. B. (2005). Toward resolution of cardiovascular health disparities in African Americans: Design and methods of the Jackson Heart Study. *Ethnicity & Disease*, 15(4 Suppl. 6), S6-4-17.
- The Trans-Atlantic Slave Trade Database (2018). *SlaveVoyages*. Retrieved from <https://www.slavevoyages.org>.
- Torres, J. B., Doura, M. B., Keita, S. O., & Kittles, R. A. (2012). Y chromosome lineages in men of west African descent. *PLoS One*, 7(1), e29687.
- United States Census Bureau (2011). In Commerce UDo (Ed.), *The Black population: 2010 census briefs* Retrieved from <http://www.census.gov/prod/cen2010/briefs/c2010br-06.pdf>
- Weir, B. S., & Cockerham, C. C. (1984). Estimating F-statistics for the analysis of population structure. *Evolution*, 38(6), 12.
- Wood, P. H. (1974). *Black majority: Negroes in colonial South Carolina from 1670 through the Stono rebellion*. W. W. Norton & Company.
- Wylie, K. C. (1977). *The Political Kingdoms of the Temne: Temne Government in Sierra Leone, 1825–1910*. Africana Publishing Company.
- Zhou, Y., Yuan, K., Yu, Y., Ni, X., Xie, P., Xing, E. P., & Xu, S. (2017). Inference of multiple-wave population admixture by modeling decay of linkage disequilibrium with polynomial functions. *Heredity (Edinb)*, 118(5), 503–510.

## SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of this article.

**How to cite this article:** Zimmerman, K. D., Schurr, T. G., Chen, W.-M., Nayak, U., Mychaleckyj, J. C., Queen Quet, Moultrie, L. H., Divers, J., Keene, K. L., Kamen, D. L., Gilkeson, G. S., Hunt, K. J., Spruill, I. J., Fernandes, J. K., Aldrich, M. C., Reich, D., Garvey, W. T., Langefeld, C. D., Sale, M. M., & Ramos, P. S. (2021). Genetic landscape of Gullah African Americans. *American Journal of Physical Anthropology*, 175(4), 905–919. <https://doi.org/10.1002/ajpa.24333>