

Advances in Clinical and Medical Research

Genesis-ACMR-6(3)-109

Volume 6 | Issue 3

Open Access

ISSN: 2583-2778

Gullah Geechee, Progenitors for African North American Maternal Lineage

Jennifer Caldwell^{1*}, Imani Sykes¹, Robbie A. Beyl¹, Tomilowo Abijo², Kiara Charles¹, and Fatimah LC Jackson¹

¹Pennington Biomedical Research Center

²Fayetteville State University, Fayetteville North Carolina

*Corresponding author: Jennifer Caldwell, Pennington Biomedical Research Center

Citation: Caldwell J, Sykes I, Beyl RA, Abijo TT, Charles K, et al. Gullah Geechee, Progenitors for African North American Maternal Lineage. *Adv Clin Med Res.* 6(3):1-17.

Received: October 30, 2025 | **Published:** December 03, 2025

Copyright © 2025 Genesis Pub by Caldwell J, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC BY 4.0). This license permits unrestricted use, distribution, and reproduction in any medium, provided the original author(s) and source are properly credited.

Abstract

African American ancestral lineages are marked by resilience, shaped by centuries of forced migrations, the Transatlantic and domestic slave trades, and subsequent movements within the United States. These historical events have influenced both the cultural and genetic diversity of Legacy African Americans. This study examined mitochondrial DNA (mtDNA) haplogroups among 189 participants—79 Gullah Geechee and 110 African Americans—to explore the Gullah Geechee as potential progenitors of other Legacy African North American lineages. Haplogroup analysis revealed L2 and L3 as the most frequent among both groups, comprising 34.6% and 38.8% in African Americans and 28.0% and 54.7% in Gullah Geechee, respectively. Non-L haplogroups associated with European and Native American ancestry were present only among African Americans (<2%), indicating greater maternal admixture compared to the Gullah Geechee. A Chi-square test showed no statistically significant difference between groups ($\chi^2 = 115.62$, $df = 49$, $p = 2.6e-07$), suggesting shared maternal ancestry. These findings support the hypothesis that Gullah Geechee lineages represent a founding biocultural population for Legacy African Americans and highlight the importance of increasing representation of African-descended populations in genomic research to better understand ancestry and health disparities.

Keywords

Gullah Geechee; Progenitors; African American; Non-L haplogroups; Lineage portrays.

Introduction

African American ancestral lineage portrays a powerful story. Beginning in the lush lands of Africa, to forced migrations of the Trans-Atlantic Slave Trade, and mass exoduses of enslaved African Americans from the upper south to the lower south, and again during the Great African American migrations, African Americans' cultural story includes these historical migrations, followed by implanting deep indigenous roots in select parts of North America. Tracing the African American migration story is crucial to understanding the people on a macro and micro level. Macro-level being the shared experiences in daily lives and encounters that have shaped, sometimes forcefully, the cultural expression of these people. On a micro level, the understanding of the genetics of the various groups at different stages, from the different ethnic groups that crossed the Atlantic to the systematic admixtures that followed, is needed to understand how current genetic profiles interact within the current health climate.

Ancestry is a major component of analyzing diversity in health outcomes [1]. For many populations, ancestry is traceable via family documents, retained cultural practices, storytelling, and foodways. Advancement in science has allowed scientists, specifically geneticists, to observe the genome of individuals through genomic research. The understanding of the functions and responses of genes to the environment provides discovery of heredity, disorders, and medical prevention in relation to genomics (NIH, 2022). In Genomic research ancestry is commonly traced via mtDNA haplotypes and Y-DNA haplotypes; mtDNA traces maternal lineage and Y-DNA traces paternal lineage [2]. Haplotype analysis is a popular tool for determining ancestry; however, in the African Diaspora, more participants in genomic studies are needed to create a more robust picture of its true diversity [3,4,5,6].

For African North Americans the reconstruction of their ancestry is more complex. During the Trans-Atlantic Slave Trade (TAST), about 9 million Africans survived the forced migration to the Americas. Historical records suggest that most of the enslaved Africans came from West and West-Central Africa. Many died due to the harsh treatment and trauma of slavery, others were breeders who bore children for the plantation owners, and even more were sold off from their families at the whim of plantation owners. Centuries later, a 2004 study compared 1148 African Americans to African mtDNA databases. This research found that greater than 55% their participants have a West African ancestor, of their participants have a West African ancestor and less than 41% of the participants have lineage in West Central or Southwestern Africa [7].

In [4], published a comparison paper of West African ethnic groups' haplotypes to haplotypes present in their African American descendants. Their article compared African American haplotypes, and specifically Gullah-Geechee HVS-I haplotypes, with HVS-I mitochondrial DNA (mtDNA) haplotypes of Sub-Saharan Africans datasets and the occurrence of identical matches between groups. Their research showed <10% of African American mtDNA matches to a single African ethnic group or region. In addition, 40% of the African American mtDNA did not match the Sub-Saharan database, with common African haplotypes present. They recommended a much larger dataset be created to capture the variation in African and African American tDNA.

We postulate the events that lead to the dominant mitochondrial haplotypes in modern North African

Americans, the descendants of formerly enslaved Africans. This research continues the work done by [4] which established a correlation between West Africans and Gullah Geechee mitochondrial DNA haplotypes. Our study analyzes gene flow using the same databases analyzing Gullah Geechee mtDNA correlation to African American mtDNA while considering evolutionary events that presumably contributed to dominant maternal haplotypes, their mutations, selection events that contribute to admixture among African American groups currently spread across North America.

Haplogroups	Gullah Geechee	African Americans	Total
L0	0	3	3
L1	13	23	36
L2	21	34	55
L3	42	38	80
N=	79	110	174

Table 1: Overall HVS1 mtDNA Gullah Geechee and Africa Americans Frequencies.

Gullah Geechee as Progenitors for North Americans

Our theory uses the Trans-Atlantic Slave Trade as the bottleneck for filtering New World Africans and creating our modern African American gene pool. During the TAST millions of Africans were taken from numerous African Kingdoms and ethnic groups as prisoners of war. They were bought and bartered for war and luxurious materials like umbrellas, spices, and gunpowder. Many were brought from the inland parts of Africa to the West Coast to board slave ships headed to the Americas. Other enslaved Africans were taken north to Europe and millions to the Middle East. These migration patterns create a strong rationale to continue exploring the African Diaspora and reconstruct the histories of modern African people as they influence and diversify new lands [8].

Our aim in this article is to use the same datasets from [4]. and highlight the relationship between Gullah Geechee participants to their potential African American descendants; propose mechanisms that support the prevalence of cardiovascular disease seen in African North Americans. In a previous paper [9], we hypothesized Gullah Geechee as a progenitor population for North American Legacy Africans, specifically for those whose ancestors were forced into the Trans-Atlantic Slave Trade. The maternal lineage(mtDNA) is our focus for this research. From 1525 to 1856, approximately 132,000 enslaved women and girls were brought to continental North America directly from Africa or from the Caribbean. It is estimated that 66,000 of those women and girls survived reproductive age and bore children. Considering migrations, women who were breeders on plantations, and women who were separated from their families and forced to bear children at their new plantations, the maternal ancestors of today's 40 million Legacy African Americans are relatively small. We recognize that some Legacy African Americans stem from European, North American, and Native American maternal lineages Amerindian lineage. However, this research focuses on the African maternal line of Legacy African Americans as we attempt to reconstruct their ancestry.

Migrations influence Admixture of Legacy African Americans

First middle passage

Unlike most ethnic groups, reconstructing the heritage of Legacy African Americans is complex. When researching African North Americans, we understand that 9 million Africans made it to continental North America while many lost their lives at sea, were thrown overboard for insurance claims as they were only spoiled property, succumbed to plagues and infections in the Americas, or did not live long enough to bear children. Historical records indicate that many enslaved African were given new names either on the slave ships or after being sold into slavery. Their surname was often discarded or seen as remnants of a “savage” past, deemed inappropriate which had no place in the Americas and were removed to further disconnect Africans from their heritage. While the First Middle Passage set the stage for Legacy African American ancestry, the Second Middle Passage shuffled the ancestral lineages of African Americans once again [2,9].

Second middle passage

Contrary to many other countries with enslaved populations, the United States North African population grew exponentially. In some cities, the African American population outsized the European population 3:1 [2]. This feat was a part of the imperialist expansion and capitalistic nature of slavery. African men and women were often breeders for plantation owners. While some enslaved Africans bore children for enslavers, others' sole job on the plantation was to breed the next generation, often with one or more partners. Plantations became familial centers as much as economic factories. Enslaved Africans were sometimes genetic family members of enslavers but also could be bought or given to extended family members, continuing the trend and connection. Mating and reproduction often resulted from proximity; as a result, some Legacy African Americans' heritage parallels with plantation families where they worked.

The Domestic Slave trade was the largest intercontinental slave trade of United States history and began after the rise of King Cotton in 1793. The invention of the cotton gin and the westward expansion following the Louisiana Purchase led to the Second Middle Passage. During this expansion, droves of enslaved Africans were chained and migrated from the Upper South (Maryland, Virginia, Tennessee, and North Carolina) to the Lower South (Georgia, South Carolina, Alabama, Mississippi, Louisiana, Florida, Arkansas, and Texas) to attend the workload [10]. This forced migration disrupted the family structure, with over 1 million African Americans being relocated to the South. The family structure and thus their genetic composition was admixed once again, causing more evolutionary homogeneity over time [11,12]. Husbands and wives, children and parents, and siblings alike were dismembered from the unity they came to know. Similar to the standards of the TAST, the Second Middle Passage (SMP) called for a separation of unity, families, and communities even after they reached their new southern destination. The consequences of subsequent alienation, relocation, and separation were compounded through the Civil War and Reconstruction. One element of African American culture is seen in the uprooting of families, imprinted in their gene flow, time and time again.

Reconstruction and the African American Great Migration

After the Civil War, many African Americans went on a quest to reunite with their families. Some remained on the plantations where they had been enslaved, becoming sharecroppers, while others had dreams of

leaving the South for the “promised land” of advancement in northern states. Legacy African Americans were living in brutal and inhumane working conditions in the south, overt racism, “Black Codes” justice, and reinstatement of covert institutionalized slavery via the 13th Amendment and the “chain gang,” which left their lives vulnerable at the hand of their oppressors. The idea of the North became laws of the South, it was a way o, it was a way out of the debt system of sharecropping, and the promise of a new, autonomous life. While the Second Middle Passage was the largest forced exodus of people in the United States, the period of American reconstruction, along with the 1st and 2nd Great Migrations, significantly shaped the demography, redistribution, and social structure of Legacy African Americans and modern Americans indirectly, as the largest internal migration of the United 2003). Migration is not only an evolutionary process of human diversity but also a social influencer of biological processes like reproduction, morbidity, and mortality. Early population records show that 90% of Legacy African American (LAA) populations live almost exclusively in southern states with South Carolina, Louisiana, and Mississippi as the most heavily populated. While half of the Legacy African American population remained in the south from 1890 onward, approximately 6 million LAA’s moved Northeast, North, and West to more urban environments. As we reconstruct the ancestral lineage of Legacy African American (LAAs), it is important to note that migration often occurred in familial patterns, frequently following the railroad tracks. Few went by cars, some by foot, and even in the secret of the night as the threat of escapees being lynched was common practice [13].

With the aid of contemporary media primarily from black-owned newspapers like the Chicago Defender, Pittsburg Courier, and California Eagle, communication about jobs, housing, and the luxuries of northern life spread throughout the south. Previous research notes the importance of the railroad system in Great Migration. Pullman porters often delivered newspapers from the north to southern rural areas [14]. Not only did people follow the closest railroad tracks to the urban cities in the north, but many migrants also used the trains legally and in secret to go north. Many people followed the system of chain migration, sending one family member ahead to become stable for the other family members to follow. [3], Showed the prevalence of the African originating HAP-L in major cities: Maywood, IL (n= 100, 16.61%), New York (n= 93, 18.23%), Pittsburgh (n= 84, 11.24%), Baltimore (n=96, 24.89), Charleston, S.C. (n=94, 12.92%), New Orleans (n=105, 10.75%). Of the major cities, Charleston S.C. has the lowest amount of European admixture followed by the northern city Philadelphia, Pennsylvania. The extent of heterogeneity, the limited weight of European admixture, and cultural dynamics of Legacy African Americans can be understood through the previously mentioned migration patterns, and by reconstructing reproductive events that altered gene flow. Not only was there a major flux of LAA’s who went north, once in urban cities the overall population grew rapidly.

MtDNA

Due to the frequent migrations of African North American descendants over the last 400 years, analyzing mitochondrial DNA (mtDNA) is one of the first steps to reconstructing their maternal ancestral story. The mtDNA haplotypes capture trends between ethnic groups and larger populations. As people and populations migrate and reproduce more frequently, more diversity is seen within ethnic groups while also allowing for haplotype sharing between ethnic groups. The location of mtDNA is within the mitochondria where it is present in hoop-like shape. A region of mtDNA known as the hypervariable

segment I (HVS-I) is specifically used to determine haplogroups. The HVS-I region is preserved between generations, is usually non-recombinant, and contains polymorphic sites that are used to infer migration patterns and molecular evolution of populations (McClean, 2003). Mitochondrial DNA, however, is limited in that it does not characterize nuclear DNA and the admixture dynamics of the entire human genome.

Our theory is that Legacy African Americans genetically, and more so maternally descend from the Sea Islands and coastal Low country that insulated the Gullah Geechee. Gullah Geechee is an amalgamation of African ethnic groups brought to Georgia, North and South Carolina, and parts of the upper Florida coast. They were expert rice cultivators, managed waterways, and cultivated indigo, and Sea-Island cotton. They forced to the United States from the Gold Coast, Senegambia, Bight of Benin, Bight of Bonny, Upper Guinea, and West-Central Africa.

(see figure 1). [4], found that over 40% came from West-Central Africa including nations like Angolans and Congolese. Research suggests that 120,000 legally documented and up to 200,000 illegally documented enslaved Africans came through the Port of Charleston (Curtin, 1969) (James and Campbell, 1981). Of the estimated 300,000 who came to the United States, roughly 40% came through the Port of Charleston, South Carolina. Today we can visualize the preservation of African genetics and biocultural diversity in the Gullah Geechee. Due to the extreme heat and malaria-infested swamps in the Low country, slave owners visited plantations less frequently. The protective mechanism of the Duffy allele kept the African descendants better fit for the environment. The insulation allowed for admixture to occur more exclusively between the diverse African ethnic groups; and it allowed cultural practices, rituals, and beliefs to be shared and remembered intergenerationally while forming a new culture in the process. In Charleston (and surrounding Sea Islands) the African population ranged from 43-93% of the total population. Often in many colonial port cities such as New Orleans, Charleston, and Chesapeake Bay, a caste system develops with various levels of European admixture which became indirectly proportionate to the social station [2,15]. In South Carolina, Gullah Geechee communities showed less overall European admixture (<10%) than compared with the Lowcountry area which is less than 30 miles inland. However, there was more interbreeding with Europeans as people left the sea island coast headed inland [2].

Method 1-2

Howard University approved this research as IRB-20-MED-78; however, the original review and approval were done by the University of South Carolina, as the African American Roots project. All participants were de-identified, and the dataset only included participant numbers and associated haplogroups. To complete an ancestral analysis of study populations, we will use secondary data from this dataset: Ely, B., Wilson, J. L., Jackson, F., & Jackson, B. A. (2006). African American mitochondrial DNAs often match mtDNAs found in multiple African ethnic groups. *BMC biology*, 4, 34. <https://doi.org/10.1186/1741-7007-4-34>. This data set comprises approximately n=79 African Americans who self-identified as Gullah-Geechee and a sample of African American DNAs (n=110) obtained from the Armed Forces DNA Identification Laboratory (AFDIL). A comparative analysis of mtDNA haplotypes was completed using R to complete Pearson's Chi-Squared and frequency tables. All additional statistical analyses were performed using SAS 9.1. To analyze the potential impact of Gullah Geechee on the Legacy African American population, population growth rates for colonial and present time periods were conducted.

Results

A total of 189 mtDNA samples were analyzed to assess the distribution of haplogroups between Gullah Geechee African Americans and African American individuals.

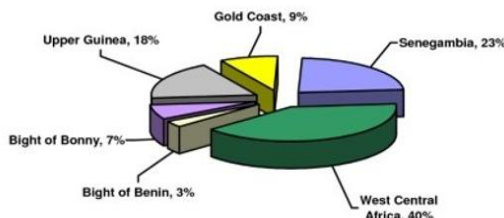


Figure 1: Proportions of enslaved Africans brought to historic Carolina coastal ports from 17th to 19th centuries CE, [16].

Another rationale for analyzing mtDNA in Legacy African Americans is to find an association of evolutionary and genomic mechanisms that influence the prevalence of cardiovascular disease phenotypes in the broader African North American population. Cardiovascular Disease (CVD) is a major health disparity in ANA communities [17]. Years after the previously mentioned migrations, the Stroke Belt: [which is the southeast of the United States, consisting of the following states:] Georgia, North and South Carolina, Florida, Arkansas, Louisiana, Mississippi, and Alabama. This being still home to a disproportionate number of ANA's [18]. These descendants of former enslaved Africans tend to have an incidence of hypertension, stroke, and other cardiovascular co-morbidities. The Center for Disease and Control (CDC) noted that there is 2 to 4 times more like to experience mortality by stroke in comparison to other regions of the United States. Understanding that one of the Stroke Belt unresearched contributors being genetic influence. Although this dataset did not include cardiovascular health-related data, our aim is to also provide a theory on how CVD phenotypes potentially and disproportionately occur in African North Americans.

Haplogroups		African American	Gullah Geechee	Origins of mtDNA. Haplogroups: Regional or Ethnic groups with high frequencies YBP* Years Before Present
1.	A1	2	0	Asia (30,000-50,000 YBP): China, Japan
2.	A2	0	1	Asia (30,000-50,000 YBP): Ache, Waiwai, Zoro
3.	B	0	2	Asia (50,000 YBP): China, Southeast Asia, Siberia
4.	H	5	0	Southwest Asia (20,000-25,000 YBP): Near East and Southern Caucasus
5.	I	1	0	West Asia (21,000 YBP): El Molo and Rendille of Kenya, North Africa, Indo-European

6.	J	1	0	Near East or Caucasus (45,000 YBP): Ireland, Scotland, and Northeast Africans
7.	K; H	0	1	West Asia (26,000-4,300 YBP): Native American Amerindians, and Ashkenazi Jews
8.	L0a 1	1	0	Southern or Southeastern Africa (130 to 200 ka YBP): Mozambique, Guineans, and Yemen
9.	L0a 1a	1	0	Southern or Southeastern Africa (130 to 200 ka YBP): Tuareg, Libyans, and Afro-Asiatic
10.	L0a2	1	0	Southern or Southeastern Africa (130 to 200 ka YBP): Tanzania
11.	L1	1	0	Central Africa (107,000-174,300 YBP): Mbengaa Pygmies
12.	L1a1	0	1	Central Africa (107,000-174,300 YBP):
13.	L1a1-311	0	1	Central Africa (107,000-174,300 YBP):
14.	L1b	4	0	West Africa (10 kya YBP): Mozambique, Senegal, and Nigeria
15.	L1b-187	0	1	West Africa (10 kya YBP): Mozambique, Senegal, and Nigeria
16.	L1b1	5	3	West Africa (10 kya YBP): Mozambique, Senegal, and Nigeria
17.	L1b1-264	0	2	West Africa (10 kya YBP): Mozambique, Senegal, and Nigeria
18.	L1c	2	0	West and Central Africa (85 kya YBP): Mbenga Pygmy, Ba-Kola, Baka people of Gabon, and Angola
19.	L1c-294	0	1	West and Central Africa (85 kya YBP):
20.	L1c1	3	0	West and Central Africa (85 kya YBP): Mbenga Pygmy, Ba-Kola, Baka people of Gabon
21.	L1c1a	0	1	West and Central Africa (85 kya YBP): Mbenga Pygmy, Ba-Kola, Baka people of Gabon
22.	L1c2	6	0	West and Central Africa (85 kya YBP): Mbenga Pygmy, Ba-Kola, Baka people of Gabon
23.	L1c2-278	0	3	West and Central Africa (85 kya YBP): Mbenga Pygmy, Ba-Kola, Baka people of Gabon
24.	L1c3	2	0	West and Central Africa (85 kya YBP): Mbenga Pygmy, Ba-Kola, Baka people of Gabon

25. L2a	3	3	West Africa or Central Africa (80,000-111,000 YBP): Chad, Mozambique, Kenya, Uganda, and Tanzania
26. L2a1	16	2	West Africa or Central Africa (80,000-111,000 YBP): Chad, Mozambique, Kenya, Uganda, and Tanzania
27. L2a1a	1	1	West Africa or Central Africa (80,000-111,000 YBP): Chad, Mozambique, Kenya, Uganda, and Tanzania
28. L2a1a/b	0	1	West Africa or Central Africa (80,000-111,000 YBP): Chad, Mozambique, Kenya, Uganda, and Tanzania
29. L2b	3	3	West Africa or Central Africa (80,000-111,000 YBP)
30. L2b1	4	0	West Africa or Central Africa (80,000-111,000 YBP): Chad, Mozambique, Kenya, Uganda, and Tanzania
31. L2c	2	8	West Africa or Central Africa (80,000-111,000 YBP): Senegal, Cape Verde, and Guinea Bissau
32. L2c1	1	0	West Africa or Central Africa (80,000-111,000 YBP): Senegal, Cape Verde, and Guinea Bissau
33. L2c2	1	1	West Africa or Central Africa (80,000-111,000 YBP): Senegal, Cape Verde, and Guinea Bissau
34. L2d1	1	0	West Africa or Central Africa (80,000-111,000 YBP): Yemen, Mozambique, and Sudan
35. L2d1-399	0	1	West Africa or Central Africa (80,000-111,000 YBP): Yemen, Mozambique, and Sudan
36. L2d2	2	1	West Africa or Central Africa (80,000-111,000 YBP): Yemen, Mozambique, and Sudan
37. L3b	6	4	East Africa or Asia (80,000-60,000 YBP)
38. L3b1	5	0	East Africa or Asia (80,000-60,000 YBP)
39. L3b2	0	2	East Africa or Asia (80,000-60,000 YBP)
40. L3d	1	3	East to Central Africa (80,000-60,000 YBP): Akan People, Fulani, Chadians, Ethiopians, Berbers, and Yemenites
41. L3d1	1	0	East to Central Africa (80,000-60,000 YBP): Akan People, Fulani, Chadians, Ethiopians, Berbers, and Yemenites
42. L3d2	3	0	East to Central Africa (80,000-60,000 YBP): Akan People, Fulani, Chadians, Ethiopians, Berbers, and Yemenites

43. L3e1	4	5	East Africa to West Africa (80,000-60,000 YBP)
44. L3e1a	3	0	East Africa to West Africa (80,000-60,000 YBP)
45. L3e1b	0	1	East Africa to West Africa (80,000-60,000 YBP)
46. L3e2	3	5	East Africa to West Africa (80,000-60,000 YBP)
47. L3e2b	6	9	East Africa to West Africa (80,000-60,000 YBP)
48. L3e3	0	<u>4</u>	East Africa to West Africa (80,000-60,000 YBP)
49. L3e4	0	1	East Africa to West Africa (80,000-60,000 YBP)
50. L3e4-223	0	1	East Africa to West Africa (80,000-60,000 YBP)
51. L3f	0	1	East Africa or Asia (80,000-60,000 YBP): Northeast Africa and Arabian Peninsula
52. L3f1	4	5	East Africa or Asia (80,000-60,000 YBP): Northeast Africa, Shael, and Arabian Peninsula
53. L3g1	1	0	East Africa or Asia (80,000-60,000 YBP)
54. L3h	1	0	East Africa (80,000-60,000 YBP)
55. U4	1	0	West Asia, North Asia, Northern Europe (21,000 – 14,000 YBP): Siberia
56. U6a1	1	0	Northwest Africa (31,000- 43,000 YBP): Canary Islands and Iberian Peninsula
57. W	1	0	Western Asia (23,9000 YBP): Europe, West/ South Asia, and Ancient Egyptians

Table 2: Complete AFDIL frequency chart of HVS1 mtDNA Gullah Geechee and African American Haplogroups.

H_0 = There is no difference in the frequency of the Gullah Geechee and African American haplotypes; and

H_1 = There is a difference between the frequency of the Gullah Geechee and African American haplotypes.

$N = 79$ (Gullah Geechee) + 110 (African Americans) = 189

Haplogroups = 57 and 49 analyzed after controlling and removing non-African haplogroups.

Haplogroup distribution

The most frequent haplogroups observed were L2 and L3, comprising 34.6% and 38.8% of AA individuals, and 28.0 and 54.7% of Gullah individuals (Table 1). Collectively, haplogroups L0–L3 accounted for 97% of African American mtDNA and 100% of Gullah mtDNA samples. Non-L haplogroups such as A2, B, H, I, J, K, U, and W were observed only in African American individuals, at low frequencies (<2% each), and were entirely absent in the Gullah group. These non-L haplogroups are typically associated with European and Native American lineages, indicating a higher degree of maternal admixture in the general African American population compared to the Gullah community.

Pearson's Chi-Squared X^2	df	p-value
115.62	49	2.63E-07

Chi Squared

The Chi Squared test found no statistically significant difference between the Gullah Geechee cohort and the Legacy African American cohort. While this is a good indication that Gullah Geechee is a potential continental progenitor for African Americans, we recognized that our sample size is small. Another consideration for the closeness seen in the mtDNA is that present-day Alabama was a part of Georgia territory until 1891. Alabama had similar slavery practices, has some similarities in environmental landscape, and could have possibly received many of enslaved Africans from Charleston, South Carolina and Savannah, Georgia which are in the current Gullah Geechee National Park Reserve.

Haplogroups	Gullah Geechee	African Americans
L1b1	3	5
L2a1	2	16
L2c	8	2
L3b	4	6
L3e2b	6	9
L3f1	5	4

Table 3: Most Frequent sub-Haplogroups in Gullah and AFDIL datasets.

Frequencies

After reviewing the frequency analysis, L3 (specifically L3e2b) was the most frequent haplogroup in the Gullah Geechee cohort, and L2 (specifically L2a1) was the most frequent haplogroup in the African American ADFIL cohort. These results are common in African Diasporic populations. L3 is an important marker for Bantu-speaking populations, originated 70,000 years ago during the Out-of-Africa expansion, and may have originated in Central Africa [7]. L3e2b is also common in Legacy African American, Afro-

Brazilian, and Caribbean ethnic groups. L2 haplogroup is the most common haplogroup of African descendants, making up about 1/3 of the total descendants. It originated in West Africa and Central Africa approximately 87-107,000 years ago. L2a1 is common in Wolof, Malinke, Hausa, Khoisan, Pygmies, Senegalese, Chad, and Akan ethnic groups. L2 is also common in Ethiopian ethnic groups and East Africans, including Sudan and Mozambique.

	1790	1820	1840	1860	2021
Black Population	108,895	265,301	335,314	412,320	1,412,875.09
Growth Rate * (Gullah)	1.29=130	-0.057= -5.76	-0.254= -25.44	-0.39= -39.36	250,000*
Growth Rate (Black)	11.9= 1197	4.32=432	3.21=321	2.42=242	

Table 4: Gullah Geechee and Black Population Growth Rate in South Carolina.

From: South Carolina Population 2021. South Carolina Population 2021 (Demographics, Maps, Graphs). (n.d.). <https://worldpopulationreview.com/states/south-carolina-population//>.

	1819	1840	1861	2021
Black Population	38,400	225,571	435,080	1,314,469.02
Growth Rate	33.23 =3.323	4.8=482.7	2.02= 202.12	

Table 5: Alabama Black Population Growth Rate.

From: Alabama Population 2021. Alabama Population 2021 (Demographics, Maps, Graphs). (n.d.). <https://worldpopulationreview.com/states/alabama-population//>.

The population growth rate calculations are used to show the growth rate of Legacy African Americans from potential Gullah Geechee ancestors. The current Gullah Geechee population is between 250,000-500,000. According to African American Slave Population, the Gullah Geechee population grew during colonial slavery. We compared the past and present Gullah Geechee population growth changes and compared the current total African American populations in South Carolina and Alabama to the current Gullah Geechee population to infer a relationship between the Gullah Geechee to their potential Legacy African American descendants. Without considering migrations and descendants who no longer live in the south, there is a positive contribution of Gullah Geechee to the modern Legacy African American population located in the represented states. More work needs to be done to confirm these suggestions.

Maternal lineage connections across continents?

Many of the non-L haplogroups in the data originate from the L3 “Out of Africa” expansion which led to M and N haplogroups. Gullah Connection research, by Joseph Opala, observes that Gullah Geechee

migrated south to the Florida Everglades and Andros Island, Bahamas, to escape slavery. s to escape slavery. Other refugee groups, such as the Black Seminoles, the Muscogee Creek Freedman Indian Band, and Black Natives of Creek, Choctaw, and Cherokee descent also included runaway individuals as members of their tribes [18,20]. While further research to confirm their ancestral research is needed, this current study observes an Asian maternal influence in Haplogroups A, B and K from our GG dataset. Furthermore, Haplogroups A and B are seen in Native American Amerindian populations, leading more relevancy towards the migration pattern *mentioned in the* Gullah Connection. Haplogroup K includes ancestral markers originating in the Horn of Africa, Northern Africa, most of Europe, South and East Asia, Native American Amerindian, and Ashkenazi Jews which the last two have a contemporary presence in South Carolina. In the ADFIL African American sample, outlier haplogroups included: A, H, I, U and W. A few of these Haplogroups namely A, B, H and I represent Native American populations that inhabited the Southeastern corridor of the United States prior to the 1830 Indian Removal Act. The mtDNA shows evidence consistent with the anthropological and historical theories of the Gullah Connection and migration patterns previously mentioned [19].

Discussion

These results expand upon on the existing understanding that African Americans are the product of centuries of admixture among various African ethnic groups. There were no significant associations between ethnicity and all the haplotypes considered: ethnicity and hap sub (χ^2 (13) = 20.4294, p-value = 0.085); ethnicity and hap sub (χ^2 (3) = 20.4294, p-values = 6,0129) and ethnicity and hap_sub2 (χ^2 (4) = 7.547, p-values = 0.1097) (Table 2).

While this article did not dive into the modest gene flow represented from European and Native American Amerindian populations, it will highlight a few maternal outliers that reflect admixture outside of African ethnic groups. In line with the hypothesis, our results indicate that the Gullah Geechee (GG) show evidence of founding population for Legacy African Americans (LAA's). As GG began to thrive off the South Carolina coast, they served as an initiating stopping post for many enslaved Africans who entered the continental United States. Major slave ports such as New York and Chesapeake Bay had greater European influence and did not achieve the same economic success from agriculture as the sea islands. New Orleans, Louisiana also became a major slave port; however, this was not until after its founding in 1718. GG communities thrived during the early colonial periods. Gullah culture, food, Creole language and belief systems are the result of complementary aspects of the African ethnic groups brought over to the new for their specialization in rice cultivation. While 25% of Gullah language can be attributed to similarities in the Krio language from Sierra Leone, other aspects like counting can be traced to the Guinea of the Fula, songs can be traced to the Mende ethnic group; and the name Gullah, is also traced to the name "Gola" representing Angolans which in previous research shows a significant amount of maternal lineage in present day Gullahs Geechee [19,4,20]. In Charleston, South Carolina that enslaved Africans received orientation for the rules and customs of their new life. The mixture of different African ethnic groups learned to accommodate and mold to comfort Africans in the Americas while retaining practices of the world they once knew.

MtDNA and Mechanisms for Ancestral Diversity

Mitochondrial haplogroups emerge due to the accumulation of consistent mutations in maternal ancestry, often characterized by a single nucleotide polymorphism. These SNP's result in a unique subclade of the preceding haplogroup. Studying the female lineage and its unique set of conditions through the transatlantic slave trade and subsequent life under untold stress is necessary for the understanding of factors that could have led to accumulation and selection of mutations in the mitochondrial germline that is subsequently passed on to future generations.

Mitochondrial DNA (mtDNA) is passed from mother to offspring in every generation. MtDNA is presumably expressed as a single haplotype (homoplasmy). However, Heteroplasmy or chimeric heteroplasmy is known to occur in a number of animal species. Heteroplasmy, a state in cells in which their mitochondrial DNA is made up of more than one genome, is an important factor when characterizing the biochemical phenotypes that arise from the mitochondrial genotype. In humans, this may manifest as mtDNA containing components from multiple ethnic groups transferred by: "somatic mtDNA mutations, paternal leakage during fertilization, or be inherited from a heteroplasmic mother [21]." Moreover, research by Polovina et al suggests that heteroplasmy trends in familial groups over random occurrences. By analyzing smaller African descendant populations, families, and communities' scientists could reconstruct migrations and ancestral patterns of modern North American Africans.

Heteroplasmy offers protection against mitochondrial disease manifestation, due to mitochondrial genomic variants being unaffected by the disease mutation. Considering that people who live well past the average human life span show an above normal level of heteroplasmy, we argue that women, enslaved and transported across the Atlantic during the TAST, were more than likely to express a higher level of heteroplasmy. This is further supported by the fact that these women were subjected to extremely harsh conditions during transportation and arrival, and their subsequent nutritional conditions were well below the livable level conditions at subsequent of events provided a bottleneck condition for the selection of the traits that we propose the current mitochondrial haplogroup as one. Their ability to not only survive but also exhibit environmental fitness can be attributed to the presence of oplasmy. A g. A germline mitochondrial mutation, when paired with chattel slavery, provides the perfect conditions for mutated cells to dominate, as cells rapidly divide before and during development, leading to a provides the perfect conditions for mutated cells to dominate, as cells rapidly divide before and during development leading to a higher chance of a shift in mitochondrial genotype within a single generation. This mitochondrial bottleneck has given rise to the preponderance of the L1 and L2 haplogroups. The conditions that favor the selection of certain non-pathologic mutations over others under unique conditions are not well understood. In this case, the d, in this the case the selection of L1 and L2 from L0 in a heteroplasmic germline cell. In addition to ancestral properties of mitochondrial DNA, our next study focuses on the mechanisms of mitochondrial DNA that influence CVD phenotypes and pathologies [21].

Conclusion

We acknowledge that the observed consistencies in Gullah Geechee haplotype frequencies could result from several factors, including recruitment methods, endogamy, isolation from other ethnic groups, and a smaller sample size. Nonetheless, our findings indicate that Gullah Geechee mtDNA shows less European

admixture compared to other groups. Specifically, the Legacy African American participants retained approximately 89% of the L haplotype in their maternal lineage, while 94% of Gullah Geechee participants exhibited an L haplotype in their maternal lineage. The absence of the non- L haplogroups in Gullah cohort contrasting with their low presence in the LAA group suggests reduced maternal admixture from non-African sources. LAA's have more non-African maternal admixture, while Gullah Geechee have more homogenous maternal lineage. These findings align with the documented social and reproductive isolation of Gullah communities, which allowed for retained strong cultural and linguistic ties to West African traditions.

The diversity between GG and other LAA's reinforces the idea that African American genetic identity is not monolithic but rather reflects a spectrum of ancestral experiences shaped by geography, cultural preservation, and systemic social forces. Given the temporal context of forced migration and chattel slavery, the Gullah Geechee population displays the maternal mtDNA haplotypes found generally among West and Central African lineages. This supports the hourglass figure proposed by [9], which reveals a bottleneck and subsequent expansion of African genetic representation through maternal lines after Trans- Atlantic Slave Trade. Other studies show that a large majority of European admixture came from the paternal line, primarily as European slave owners often through coercive sexual exploitation of enslaved African women. Meanwhile, it was primarily the women, particularly those designated as breeders, that would be responsible for carrying the next generation of enslaved Africans. Making them instruments in creating the mitochondrial that is crucial for studying the preservation of African genetic heritage.

While the findings of this study indicated no link between ethnicity and the examined haplotypes, this could be attributed to the small sample size ($n = 189$) used in the study. This study can be regarded as merely a pilot investigation, and the data obtained may serve as a basis for designing a more extensive study. These insights emphasize the importance of increasing African and African- descended population representation in genomic studies for the better understand of health disparities, evolutionary biology, and ancestry. As the use of haplotype data continues to inform precision medicine and public health policy having an expanded dataset of underrepresented populations, like the Gullah Geechee, will be vital towards achieving both equity and scientific accuracy.

Data availability: The data generated and analyzed during the current study accompany this submission. Ethics approval, participant permissions, and all other relevant approvals were granted for this data sharing.

Ethics: This research was reviewed and approved by the Pennington Biomedical Research Center Institutional Review Board (Study #2023-051-PBRC). The study was determined to be Exempt under Category 4(ii), which permits secondary research using identifiable private information or biospecimens when data are recorded in a way that prevents direct or indirect identification of participants, and when no contact or re-identification occurs.

Acknowledgements

We gratefully acknowledge Dr. Bert Ely of the University of South Carolina for generously providing access to the mtDNA datasets foundational to this research and for establishing the comparative framework upon which this study builds. His earlier work on African and African American mitochondrial haplotypes made this analysis possible. We also thank the researchers and participants involved in the original African American Roots project, whose contributions continue to advance understanding of African-descended genetic diversity. Additional appreciation is extended to our colleagues and collaborators who offered guidance throughout this project.

References

1. Campbell MC, Tishkoff SA. (2008) African Genetic Diversity: Implications for Human Demographic History, Modern Human Origins, and Complex Disease Mapping. *Ann Rev Genomics Human Gene.* 9(1):403-33.
2. Parra EJ, Kittles RA, Argyropoulos G, Pfaff CL, Hiester K, et al. (2001) Ancestral proportions and admixture dynamics in geographically defined African Americans living in South Carolina. *Am J Phys Anthropol.* 114(1):18-29.
3. Parra EJ, Marcini A, Akey J, Martinson J, Batzer MA, et al. (1998) Estimating African American Admixture Proportions by Use of Population-Specific Alleles. *Am J Human Gene.* 63(6):1839-51.
4. Ely B, Wilson JL, Jackson F, Jackson BA. (2006) African-American mitochondrial DNAs often match mtDNAs found in multiple African ethnic groups. *BMC Biology.* 4(1).
5. Bryc K, Durand EY, Macpherson JM, Reich D, Mountain JL. (2015) The Genetic Ancestry of African Americans, Latinos, and European Americans across the United States. *Am J Human Gene.* 96(1):37-53.
6. Landry LG, Ali N, Williams DR, Rehm HL, Bonham VL. (2018) Lack of diversity in genomic databases is a barrier to translating precision medicine research into practice. *Health Aff.* 37(5):780-85.
7. Salas A, Carracedo Á, Richards M, Macaulay V. (2005) Charting the Ancestry of African Americans. *Am J Human Gene.* 77(4):676-80.
8. Jackson FL. (1991) An evolutionary perspective on salt, hypertension, and human genetic variability. *Hypertension.* 17(1):129-32.
9. Caldwell J, Jackson FLC. (2021) Evolutionary Perspectives on African North American genetic diversity: origins and prospects for future investigations. *Evol Anthropol.* 30(4):242-252.
10. Gates HL. (2013). *The Second Middle Passage.* PBS.
11. Slatkin M. (1987) Gene flow and the geographic structure of natural populations. *Sci.* 236(4803):787-92.
12. Baharian S, Barakatt M, Gignoux CR, Shringarpure S, Errington J, et al. (2016) The Great Migration and African-American Genomic Diversity. *PLoS Genet.* 12(5):e1006059.
13. Tolnay SE. (2003) The African American "Great Migration" and beyond. *Ann Rev Sociol.* 29(1):209-32.
14. Wilkerson I. (2020) *The warmth of other suns: the epic story of America's great migration.* Penguin Books.
15. Jackson FL. (2008) Ancestral links of Chesapeake Bay region African Americans to specific Bight of Bonny (West Africa) Microethnic groups and increased frequency of aggressive breast cancer in both regions. *Am J Human Bio.* 20(2):165-73.
16. Jackson FL. (2004) Human genetic variation and health: new assessment approaches based on ethnogenetic layering. *British Med Bulletin.* 69(1)2;15-35.
17. Yancy CW. (2018) Cardiovascular Disease Disparities: The Gap Remains. *JAMA Cardiol.* 3(12):1183.
18. Karp DN, Wolff CS, Wiebe DJ, Branas CC, Carr BG, et al. (1942) Reassessing the Stroke Belt. *Stroke.* 47(7):1939-42.
19. Opala JA. (1987) *The Gullah: Rice, Slavery and the Sierra Leone-American Connection.* USIS.

20. Amos AM. (2014) Black Seminoles: The Gullah Connections. The Black Scholar.
21. Polovina ES, Parakatselaki ME, Ladoukakis ED. (2020) Paternal leakage of mitochondrial DNA and maternal inheritance of heteroplasmy in *Drosophila* hybrids. *Sci Rep*. 10(1):2599.
22. Stewart J, Chinnery P. (2015) The dynamics of mitochondrial DNA heteroplasmy: implications for human health and disease. *Nat Rev Genet*. 16:530-42.
23. African-Americans – Slave Population. South Carolina. (n.d.).
24. Alabama Population 2021. Alabama Population 2021 (Demographics, Maps, Graphs). (n.d.).
25. Behar DM, Hammer MF, Garrigan D, Villems R, Bonne-Tamir B, et al. (2004) MtDNA evidence for a genetic bottleneck in the early history of the Ashkenazi Jewish population. *Europ J human Gene*. 12(5):355-64.
26. Bray AW, Ballinger SW. (2017) Mitochondrial DNA mutations and cardiovascular disease. *Curr Opin Cardiol*. 32(3):267-74.
27. Jackson FL. (2008) Ethnogenetic layering (EL): an alternative to the traditional race model in human variation and health disparity studies. *Ann Human Bio*. 35(2):121-44.
28. Johnson DM, Campbell RR. (1981) Black migration in America: A social demographic history (Vol. 4). Durham, NC: Duke University Press.
29. McLean DC, Spruill I, Argyropoulos G, Page GP, Shriver MD, et al. (2005) Mitochondrial DNA (mtDNA) haplotypes reveal maternal population genetic affinities of Sea Island Gullah-speaking African Americans. *Am J Phys Anthropol*. 127(4):427-38.
30. McLean DC, Spruill I, Gevaio S, Morrison EY, Bernard OS, et al. (2003). Three Novel mtDNA Restriction Site Polymorphisms Allow Exploration of Population Affinities of African Americans. *Human Bio*. 75(2):147-61.
31. Parra EJ, KG Hiester, HL Norton, JT Baker C, Bonilla, et al. (1999) African-American hypertensive subjects have less European ancestry than normotensive controls. *Am J Human Gene*. 65(4):A14-A14.
32. Pollitzer WS. (2005) The Gullah people and their African heritage. University of Georgia.
33. Schmid P, Ravenell KR, Sheldon SL, Flegel WA. (2011) DARC alleles and Duffy phenotypes in African Americans. *Transfusion*. 52(6):1260-67.
34. South Carolina Population 2021. South Carolina Population 2021 (Demographics, Maps, Graphs). (n.d.).
35. Tishkoff SA, Reed FA, Friedlaender FR, Ehret C, Ranciaro A, et al. (2009) The genetic structure and history of Africans and African Americans. *Sci*. 324(5930):1035-44.
36. Venter M, van der Westhuizen FH, Elson JL. (2018). The aetiology of cardiovascular disease: a role for mitochondrial DNA? *Cardio J Africa*. 29(2):122-32.
37. Wallace DC. (2000) Mitochondrial defects in cardiomyopathy and neuromuscular disease. *Am Heart J*. 139(2).
38. Weiss) Bolnick DA, Smith DG. (2003) Unexpected patterns of mitochondrial DNA variation among Native Americans from the Southeastern United States. *Am J Phys Anthropol*. 122(4):336-54.