

Ancestral Proportions and Admixture Dynamics in Geographically Defined African Americans Living in South Carolina

E.J. Parra,¹ R.A. Kittles,² G. Argyropoulos,³ C.L. Pfaff,¹ K. Hiester,¹ C. Bonilla,¹ N. Sylvester,² D. Parrish-Gause,² W.T. Garvey,³ L. Jin,⁴ P.M. McKeigue,⁵ M.I. Kamboh,⁶ R.E. Ferrell,⁶ W.S. Pollitzer,⁷ and M.D. Shriver^{1*}

¹Department of Anthropology, Pennsylvania State University, University Park, Pennsylvania 16802

²National Human Genome Center, Howard University, Washington, DC 20060

³Department of Medicine/Endocrinology, Medical University of South Carolina, Charleston, South Carolina 29403

⁴Human Genetics Center, University of Texas-Houston, Houston, Texas 77030

⁵Department of Epidemiology and Population Health, London School of Hygiene and Tropical Medicine, London WC1E 7HT, UK

⁶Department of Human Genetics, University of Pittsburgh, Pittsburgh, Pennsylvania 15261

⁷Department of Cell Biology and Anatomy, School of Medicine, University of North Carolina, Chapel Hill, North Carolina 27599

KEY WORDS South Carolina, African Americans, admixture, admixture linkage disequilibrium

ABSTRACT We analyzed admixture in samples of six different African-American populations from South Carolina: Gullah-speaking Sea Islanders in coastal South Carolina, residents of four counties in the “Low Country” (Berkeley, Charleston, Colleton, and Dorchester), and persons living in the city of Columbia, located in central South Carolina. We used a battery of highly informative autosomal, mtDNA, and Y-chromosome markers. Two of the autosomal markers (FY and AT3) are linked and lie 22 cM apart on chromosome 1. The results of this study indicate, in accordance with previous historical, cultural, and anthropological evidence, a very low level of European admixture in the Gullah Sea Islanders ($m = 3.5 \pm 0.9\%$). The proportion of European admixture is higher in the Low Country (m ranging between $9.9 \pm 1.8\%$ and $14.0 \pm 1.9\%$), and is highest in Columbia ($m = 17.7 \pm 3.1\%$). A sex-biased European gene flow and a small Native American contribution to the African-American gene pool are also evident in these data. We studied the pattern of

pairwise allelic associations between the FY locus and the nine other autosomal markers in our samples. In the combined sample from the Low Country ($N = 548$), a high level of linkage disequilibrium was observed between the linked markers, FY and AT3. Additionally, significant associations were also detected between FY and 4 of the 8 unlinked markers, suggesting the existence of significant genetic structure in this population. A continuous gene flow model of admixture could explain the observed pattern of genetic structure. A test conditioning on the overall admixture of each individual showed association of ancestry between the two linked markers (FY and AT3), but not between any of the unlinked markers, as theory predicts. Thus, even in the presence of genetic structure due to continuous gene flow or some other factor, it is possible to differentiate associations due to linkage from spurious associations due to genetic structure. *Am J Phys Anthropol* 114:18–29, 2001. © 2001 Wiley-Liss, Inc.

The determination of the extent of European admixture and the dynamics of the admixture process in African-American populations are of great anthropological, historical, and epidemiological interest. The increasing availability of markers showing high-frequency differentials between African and European populations (Shriver et al., 1997) has made possible more precise estimates of admixture proportions in African Americans living in different areas of the US (Parra et al., 1998). Previous studies indicated that the European genetic contribution to Southeastern African Americans (South Carolina and Georgia) has been lower than in other regions (Workman, 1968; Adams and Ward, 1973; Long, 1991; Parra et al., 1998). There are several historical reasons that may account for these differences. After

its founding in 1670, Charles Town (later to become Charleston) became one of the most important ports for the importation of enslaved persons from West and Central Africa. It has been estimated that 120,000 enslaved Africans were imported legally to Charleston, and that number could be much higher (around 200,000) when unrecorded and illegal trade

Grant sponsor: NIH/NIDDK; Grant number: DK53958; Grant sponsor: NIH/NHGRI; Grant number: HG02154.

*Correspondence to: Mark D. Shriver, Department of Anthropology, Pennsylvania State University, 409 Carpenter Bld., University Park, PA 16802. E-mail: Mds17@psu.edu

Received 11 May 2000; accepted 21 September 2000.

is taken into account (Pollitzer, 1999). This number represents a significant proportion of the total slave trade in the US, which is generally estimated to be in the range of 380,000–570,000 (Curtin, 1969; Johnson and Campbell, 1981). The demand for enslaved Africans in the southern British colonies, in order to work in the rice fields along the rivers and later in the cultivation of indigo and cotton, was very high through the 18th century and into the 19th century. As a consequence of this high influx of Africans into South Carolina and Georgia, the proportion of people of African ancestry in this area has been very high since the beginning of the colonies. In 1790, the percentage of persons of African origin with respect to the total South Carolina population was approximately 43%, but this percentage was much higher in several coastal areas, including the port of Charleston (ranging from 47–93%). On many plantations, the number of Africans vastly outnumbered Europeans, and the relative isolation of the people of African ancestry allowed the formation of a unique culture in which many African influences were preserved in terms of language, beliefs, folklore, and art. This Creole culture has survived to modern times in the Gullah-speaking Sea Islanders of South Carolina and Georgia (also known as Geechee in Georgia). A book discussing the genetic and cultural aspects of the Gullah people was recently published by one of the authors (Pollitzer, 1999).

We estimated the extent of European admixture in six different samples from South Carolina. The first sample corresponds to Gullah people from St. Helena and Beaufort. Four additional samples come from different counties of the Low Country (Berkeley, Charleston, Colleton, and Dorchester). Finally, the sixth sample comes from Columbia, the state capital, located in central South Carolina. In order to determine admixture proportions, we used a panel of specially selected autosomal markers showing high differences in frequency between African and European populations (APOA1, AT3, FY, ICAM1, LPL, D11S429, OCA2, RB1, Sb19.3, and GC). We also used a battery of informative mtDNA and Y-chromosome markers in order to infer sex-specific gene flow and to test for Native American contribution. Finally, we studied the pattern of pairwise associations between the 10 autosomal markers in our samples. This strategy has allowed us to test for admixture linkage disequilibrium between two linked markers (FY and AT3), and to test for the presence of associations between unlinked markers, which would be indicative of the existence of genetic structure in these populations.

SUBJECTS AND METHODS

Population samples

The selection of parental populations is a key aspect of any study of admixture. The samples used in our study as representative of parental populations are from areas known to have contributed substan-

tially to the populations of African and European ancestry currently living in the US.

African and European samples. The samples from Africa come from three geographical areas: Sierra Leone, Nigeria, and the Central African Republic.

Sierra leone. The sample from Sierra Leone includes several tribal groups. The best-represented groups are the Mende (N = 183) and the Temne (N = 79). There are also a limited number of individuals representing other groups (Creole, Fula, Limba, Loko, Mandingo, and Susu, among others). These samples were collected as part of an anthropological study to characterize the genetic diversity and ancestry of the Gullah Sea Islanders.

Nigeria. Two samples from Nigeria were analyzed in this study. The first sample was collected from a group of civil servants in Benin City, Nigeria (N = 46). The second sample comes from a traditional Yoruba community in the city of Ibadan, in southwestern Nigeria (N = 100).

Central african republic. This sample of Bantu speakers was collected as part of an anthropological survey of a village along the Oubangui River near the capital, Bangui (N = 49).

European samples. The samples of Europeans from Germany (N = 30), Ireland (N = 86), and England (N = 44) were collected at random as part of anthropological surveys (Parra et al., 1998).

South Carolina samples. The sample of Gullah-speakers (N = 83) comes from St. Helena and Beaufort, and was selected on the basis of the following criteria: 1, geographic location: both parents born in the Sea Islands (either an island or the immediate coast); 2, at least one parent speaks the Gullah language; and 3, individuals are unrelated.

The samples from the Low Country come from four different Counties: Charleston (N = 145), Berkeley (N = 182), Colleton (N = 121), and Dorchester (N = 93). The sample from Charleston includes individuals from five areas within Charleston County: the Sea Islands southwest of the city, east of the Cooper River, north of the city, St. Andrews and St. Paul west of the city, and the city itself. All samples came from African-American women living in these areas and participating in a lead determination and prenatal study.

The sample from Columbia (N = 90) consists of adult males residing in Columbia, South Carolina. All subjects were volunteer participants in a genetic study on prostate cancer and prostate cancer screening in the African-American population.

DNA analysis

Autosomal markers. Our strategy was to select the most informative single nucleotide polymor-

phisms (SNPs) and Alu insertions available, which can be genotyped easily in the laboratory, and to combine these markers to obtain a precise estimate of admixture. There are other genetic markers available for studying European admixture in African Americans, namely, immunoglobulin GM and the cDe complex of the RH system (Reed, 1969). However, these are complex markers for which haplotypes have to be characterized, which substantially increases the genotyping effort, without an equivalent gain in information content.

Ten autosomal markers (APOA1, AT3, FY, ICAM1, LPL, D11S429, OCA2, RB1, Sb19.3, and GC) were genotyped by standard PCR and electrophoretic separation of DNA fragments. Two of these markers, FY and AT3, are located on the long arm of chromosome 1, approximately 22 cM apart. Both markers are unlinked to any of the remaining loci. The primer sequences and PCR conditions for APOA1, AT3, FY, ICAM1, LPL, OCA2, RB1, Sb19.3, and GC were detailed in a previous paper (Parra et al., 1998). For typing D11S429, forward and reverse primer sequences were 5' CTG GCA TCT ATT AGA CCA GC 3' and 5' CGG TGC ACA ATG GAT GAA GGC 3', respectively. PCR reactions (94°C for 5 min, and then 30 cycles of 94°C for 30 s, 60°C for 30 s, and 72°C for 30 s, with a final extension at 72°C for 5 min) were performed in a 25- μ l volume containing 200 μ M dNTPs, 10 mM Tris-HCl (pH 8.9), 50 mM KCl, 1 U Taq polymerase, and 20 ng genomic DNA. The D11S429 polymorphism was detected by digesting the PCR product with 20 U *SacI* enzyme and running the resulting products in 3% agarose gels.

mtDNA and Y-chromosome specific markers. Six population-specific mtDNA haplogroups (L, H, A, B, C, and D) and the Alu insertion of the Y-chromosome (YAP) were typed using the primer sequences and PCR conditions described in Parra et al. (1998) (mtDNA haplogroups and YAP). DYS199, another very informative Y-chromosome specific marker, was characterized following the protocol of Lell et al. (1997).

Statistical analysis

The fit of genotype frequencies to Hardy-Weinberg proportions and the heterogeneity in allele frequencies in the four counties of the Low Country were analyzed by means of the HW and STRUC programs of the Genepop 2.0 computer package (Raymond and Rousset, 1995).

Because of the importance of an adequate estimate of the parental frequencies to obtain a correct admixture estimate, we would like to discuss our strategy for the estimation of the European and African parental frequencies. Ideally, we would like to have samples from all relevant areas that contributed to the migration of Europeans and Africans to the New World, and to weight the frequencies by the relative contribution of all relevant geographic and/or population areas (Reed, 1969). The limited

availability of samples, particularly from the populations of Africa, and the partial knowledge of contribution by areas, precludes our acting on this ideal. One way to overcome the difficulties inherent in having limited samples, and the uncertainty that they accurately reflect the relative contributions of different geographic areas, is to select genetic markers that show homogeneity within Africa and Europe. For this reason, we tried to use genetic markers that are informative for admixture (high-frequency difference between European and African populations) and that show similar frequencies in the samples that are available from Africa and Europe. In this way, we tried to avoid any bias in the parental African and European frequencies used to estimate admixture in African-American populations.

The admixture proportions of the samples were estimated using a weighted least squares method (WLS) (Elston, 1971; Long and Smouse, 1983; Long, 1991), implemented in the ADMIX program, kindly supplied by Dr. Jeffrey C. Long. This method incorporates the effect of the evolutionary and sampling variance in the admixture estimates and a χ^2 test of heterogeneity of admixture estimates from the different loci. Individual admixture proportions were calculated using a maximum likelihood method, as described by Chakraborty et al. (1986).

Haplotype frequencies and gametic disequilibrium coefficients for pairs of loci were estimated by using an expectation maximization algorithm described by Long et al. (1995). The gametic disequilibrium coefficient is the difference between the observed haplotype frequencies, inferred by means of the expectation maximization method, and the expected frequencies, which are a function of the observed allele frequencies for a pair of markers. Hypothesis testing was performed with the likelihood ratio statistic (G), which has a χ^2 distribution for large sample sizes. We used the 3LOCUS program, made available to us by Dr. Jeffrey C. Long.

To detect association between linked markers, we used a novel statistical method that distinguishes between the excess of association due to linkage and the potential association between unlinked markers due to population structure. In this method, it is necessary to estimate the overall admixture of each individual so that it is possible to control for confounding by the variation of admixture between individuals. More details on this Bayesian method can be found elsewhere (McKeigue, 1998; McKeigue et al., 2000).

RESULTS

Admixture proportions inferred from autosomal markers

The allele frequencies corresponding to the 10 autosomal markers analyzed in the six samples from South Carolina are depicted in Table 1. In the samples from the Low Country, we report frequencies for

TABLE 1. Allele frequencies of autosomal markers analyzed in present study¹

Population	N	APOA1*1	AT3*1	FY-Null*1	ICAM1*1	LPL*1	OCA2*1	RB1*1	Sb19.3*1	D11S429*1	GC*1F	GC*1S
Sierra Leone (Mende)	181	0.386	0.853	0.006	0.761	0.961	0.169	0.927	0.409	0.092	0.885	0.063
Sierra Leone (Temne)	98	0.356	0.816	0.000	0.695	0.974	0.116	0.944	0.409	0.074	0.907	0.050
Sierra Leone (average)		0.371	0.835	0.003	0.728	0.968	0.142	0.936	0.409	0.083	0.896	0.056
Nigeria-1	46	0.409	0.889	0.000	0.772	0.957	0.078	0.917	0.457	0.111	0.849	0.081
Nigeria-2	100	0.480	0.875	0.000	0.697	0.985	0.124	0.944	0.455	0.116	0.846	0.085
Nigeria (average)		0.454	0.883	0.000	0.734	0.971	0.101	0.933	0.456	0.114	0.847	0.083
Central African Republic	49	0.435	0.859	0.000	0.798	0.978	0.092	0.900	0.364	0.040	0.778	0.067
African average		0.420	0.859	0.001	0.753	0.972	0.112	0.923	0.410	0.079	0.841	0.069
England	44	0.934	0.291	1.000	1.000	0.528	0.695	0.294	0.949	0.576	0.203	0.622
Ireland	86	0.915	0.279	1.000	1.000	0.397	0.761	0.287	0.943	0.514	0.133	0.633
Germany	30	0.933	0.267	1.000	1.000	0.533	0.850	0.417	0.839	0.533	0.133	0.567
European average		0.927	0.279	1.000	1.000	0.486	0.769	0.333	0.910	0.541	0.156	0.607
Delta (p _{afr} - p _{eur}) ²		0.507	0.580	0.999	0.247	0.486	0.657	0.590	0.500	0.462	0.685	0.538
South Carolina												
Columbia	90	0.567	0.732	0.171	0.826	0.871	0.191	0.896	0.632	0.226	0.805	0.104
Berkeley County	182	0.468	0.777	0.088	0.753	0.912	0.176	0.827	0.513	0.150	0.806	0.107
Charleston County	145	0.490	0.768	0.075	0.683	0.921	0.152	0.827	0.510	0.135	0.781	0.130
Colleton County	121	0.504	0.742	0.107	0.790	0.908	0.180	0.789	0.548	0.113	0.772	0.134
Dorchester County	93	0.541	0.735	0.105	0.794	0.878	0.139	0.799	0.515	0.129	0.765	0.147
Low Country (total)		0.496	0.759	0.091	0.750	0.907	0.165	0.814	0.521	0.135	0.785	0.125
Gullah	83	0.475	0.867	0.036	0.759	0.939	0.171	0.922	0.410	0.080	0.822	0.072

¹ We followed the convention of defining the presence of Alu insertions and the absence of the polymorphic restriction sites as allele 1.

² Delta is defined as the absolute value of the difference between the average European and African frequencies.

the individual counties (Berkeley, Charleston, Colleton, and Dorchester), as well as the total frequencies. No significant differences were observed between the allele frequencies of the four counties for the autosomal markers, with the exception of ICAM1. The frequencies for the parental populations (European and African) indicated in the table were calculated as an unweighted average of three European samples (England, Ireland, and Germany) and three geographical areas in Africa (Sierra Leone (Mende and Temne), Nigeria (Benin City and Yoruba), and the Central African Republic (Oubangui River)), respectively. The δ values for the different markers, defined as the absolute value of the difference in frequency between the African and European populations, are presented in the last row of the table. These values are indicative of the information content of each marker for the estimation of the admixture proportions in the African-American populations.

Table 2 shows the European genetic contribution (m) to these six African-American samples from South Carolina, estimated using Long's method, with the upper and lower values of the 95% confidence interval. The admixture proportions observed in the four counties of the Low Country are also indicated. A cline in the percentage of European ancestry is evident from this analysis. The Gullah Sea Islanders show a very low European ancestry (3.5%), the lowest we have thus far observed in more than 16 African-American populations and Jamaica (Parra et al., 1998 and unpublished data). The average European contribution in the four Low Country counties is 11.8%. This value is similar to the m estimate observed in a subset of 95 randomly chosen individuals of this sample, and is lower than the m values of other African-American populations in the US (Chakraborty et al., 1992; Parra et al., 1998). The admixture estimates for these four counties are not significantly different, judging by the 95% con-

TABLE 2. European contribution to three African-American populations from South Carolina, with upper and lower values of 95% confidence interval

Population	M	95% C.I.
Gullah	3.5%	1.7–5.3%
Low Country		
Berkeley County	10.9%	8.5–13.3%
Charleston County	9.9%	6.4–13.4%
Colleton County	13.6%	10.3–16.9%
Dorchester County	14.0%	10.3–17.7%
Total (weighted)	11.8%	9.1–14.5%
Columbia	17.7%	11.6–23.8%

fidence intervals. Finally, the sample from Columbia shows the highest level of European ancestry ($m = 17.7\%$), which is similar to values previously observed in many northern African-American samples, using the same markers (Parra et al., 1998). The 95% confidence interval of the admixture estimates from Columbia and the Gullah Sea Islanders do not overlap, showing significant differences in admixture proportions.

The individual m values for the different markers show, in general, good concordance with the average m value (data not shown), with the exception of ICAM1, which tends to give lower m values than the remaining markers, in some cases even negative and significantly different when applying Long's χ^2 test. This could be due to a biased estimate of the African parental frequency (ICAM1 is monomorphic in Europeans), or some form of selection acting on this marker (Fernandez-Reyes et al., 1997). In any case, ICAM1 is the least informative of all the selected markers ($\delta = 0.24$), and makes only a small contribution to the estimate of admixture.

Individual admixture distributions

In addition to the admixture proportion of the samples, it is possible to calculate the admixture of each individual using a maximum likelihood

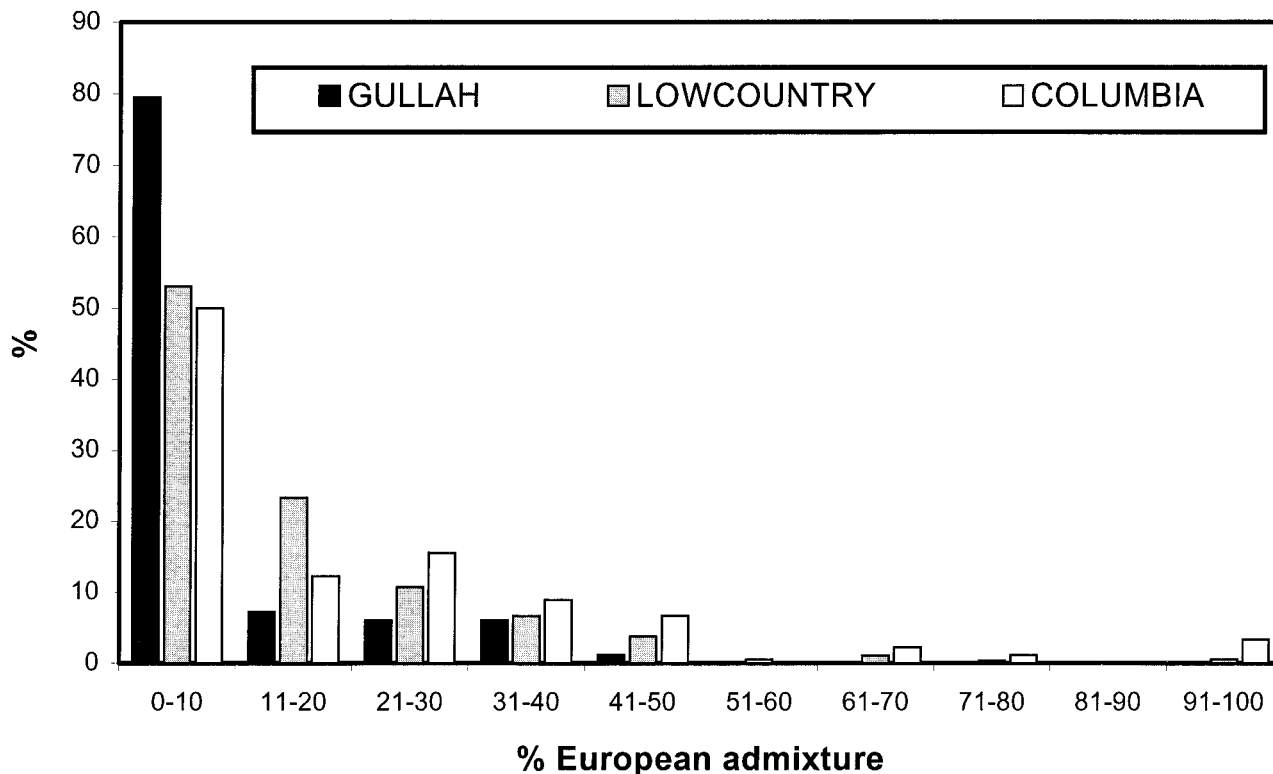


Fig. 1. Distribution of percentage of European alleles in African-American individuals from Beaufort and St. Helena (Gullah Sea Islanders), the Low Country, and Columbia. We created 10 groups to summarize individual admixture data. Given that there are 101 possible values of individual admixture (from 0–100% European contribution), the first group includes 11 individual admixture values (from 0–10% European contribution), and the remaining groups include 10 individual admixture values each (from 11–20%, 21–30%, and so on).

method, where the m value at which a particular multilocus genotype is most likely is taken as the individual admixture proportion. Since the allele frequencies and admixture proportions in the counties of the Low Country are similar, we combined them into a single sample. Figure 1 shows the individual admixture distributions observed in the Gullah Sea-Islander, Low Country, and Columbia samples. It is evident that there are important differences in the patterns observed in the three samples. In the Gullah, around 80% of the individuals have 10% or less European admixture. The proportion of individuals in this category is lower in the Low Country (around 53%) and even more so in Columbia (50%). On the contrary, the sample from Columbia has the highest proportion of individuals (around 7%) with estimates of m higher than 50%, while no individual of the Gullah sample falls in that category. The average estimates of individual admixture for the three samples are very close to the population estimates calculated by Long's method (Gullah, 6.05%; Low Country, 13.16%; Columbia, 18.73%).

mtDNA and Y-chromosome-specific markers

The analysis of mtDNA and Y-chromosome-specific markers can give additional information in terms of the dynamics of gene flow, and is especially

valuable for determining a potential sex bias in admixture contributions. We analyzed six mtDNA haplogroups, which are continent-specific: haplogroup L (African), haplogroup H (European), and haplogroups A, B, C, and D (Native Americans). Haplogroup L, which is common in Africa, consists of three major subgroups L1, L2, and L3 (Watson et al., 1997). The *Hpa*I site at np 3592 is observed in subgroups L1 and L2 but is lacking among the L3 group. Table 3 reveals that 45–55% of the African-American mtDNA types consist of L1/L2 lineages. The L3 lineage may represent a large proportion of the remaining mtDNA haplotypes in the African-American populations. The L3 lineage is pan-African, representing about 30–40% of African mtDNA sequences (R.A. Kittles, unpublished data). We also typed two Y-chromosome-specific markers: the YAP element is very useful for characterization of the male European contribution, and the DYS199 C-T transition is very informative for evaluating male Native American influence. The main results of this analysis are depicted in Table 3. The European female contribution seems to have been very low in the three populations. In the Gullah, we did not detect any H haplogroup among 83 individuals. In the Low Country, we observed only five H haplogroups in a sample of more than 540 individuals, and the female European contribution was esti-

TABLE 3. African-Americans allele frequencies of informative mtDNA haplogroups and Y-chromosome-specific markers in three South Carolina African-American populations and the European and Native American admixture proportions inferred from those frequencies

Population	mtDNA, maternal contribution									Y-specific paternal contribution				
	N	Hap-L1/L2	Hap-H	% Eur ¹	Hap-A	Hap-B	Hap-C	Hap-D	% Am ²	N	YAP+	% Eur ³	DYS199	% Am ⁴
Gullah	83	45.8%	0.0%	0.0%	0.0%	2.4%	0.0%	0.0%	2.4%	25	84.0%	5.6%	0.0%	0.0%
Low Country	541	55.2%	0.9%	2.1%	0.5%	0.7%	0.2%	0.2%	1.6%	0	n.a. ⁵	n.a. ⁵	n.a. ⁵	n.a. ⁵
Columbia	90	52.2%	1.1%	2.6%	1.1%	0.0%	0.0%	0.0%	1.1%	90	68.7%	24.1%	0.0%	0.0%

¹ European admixture estimated from Hap-H frequencies. ² Amerindian admixture estimated from Hap-A, -B, -C, and -D frequencies. ³ European admixture estimated from the Y alu insertion frequencies. ⁴ Amerindian admixture estimated from the DYS199 frequencies. ⁵ n.a., not available. Only females in the sample.

TABLE 4. G statistic evaluating extent of observed pairwise associations between FY and remaining markers examined in this study¹

Sample	G values, P								
	FY/AT3 ¹	FY/APO	FY/ICAM	FY/LPL	FY/L19.2	FY/OCA2	FY/RB2300	FY/SB19.3	FY/GC1F
Gullah	0.15, ns	0.18,ns	0.00,ns	3.19,ns	1.22,ns	2.91, ns	1.29, ns	1.52,ns	0.27,ns
Low Country	16.70***	4.83*	0.51,ns	1.21,ns	0.11,ns	14.15***	10.27***	0.87,ns	4.93*
Columbia	8.87**	0.86,ns	0.00,ns	1.05,ns	2.31,ns	2.69, ns	0.32, ns	0.19,ns	2.45,ns

¹ FY and AT3 (marked in bold type) are linked our chromosome 1. FY and the other autosomal markers are unlinked. ns, not significant. * $P < 0.05$. ** $P < 0.01$. *** $P < 0.001$.

mated at around 2%. Finally, in Columbia we detected one H haplogroup in a sample of 90 individuals, giving an approximate female European contribution of 2.6%. The male European contribution is higher than the female contribution. The estimates of European ancestry based on YAP frequencies are 5.6% in the Gullah, and 24.1% in Columbia. The sample from the Low Country is composed entirely of women, so in this case it was not possible to test for male-specific gene flow. With respect to the Native American influence, there is evidence of female Native American contribution in these southeast African-American populations. In the Gullah, we observed two individuals with haplogroup B, and in the Low Country all the Native American haplogroups (A, B, C, and D) were present, although at very low frequencies. In Columbia, we found one Native American A haplogroup. In the analysis of the DYS199 Y-chromosome specific locus, we did not observe any individual with the DYS199T variant, which is found at high frequencies in Native American populations.

Pairwise association between markers: testing for genetic structure

Two of the autosomal markers analyzed in this study, FY and AT3, are located on the long arm of chromosome 1, approximately 22 cM apart. Both markers are unlinked to the remaining loci. In order to test if the admixture process had created a detectable disequilibrium between these markers, and also to test if there is genetic structure in these African-American populations, we studied the pattern of pairwise association between FY, the most informative marker, and the nine remaining loci. Here we define genetic structure as the presence of a significant association between unlinked markers, which is not expected to occur in a random mating

population. Genetic structure can be caused by several factors, including population substructure, continuous gene flow, and assortative mating. The results are depicted in Table 4, in terms of the G statistic evaluating the extent of the observed association. In the Gullah sample no statistical association is observed between Duffy and any of the other genetic markers. In Columbia, the only significant association is observed between FY and AT3, which are linked on chromosome 1 ($G = 8.87$, $P < 0.01$). FY and AT3 also show a very significant association in the sample from the Low Country ($G = 16.70$, $P < 0.001$). Surprisingly, 4 of the 8 comparisons of Duffy and the unlinked markers are also significant in this much larger sample (FY/OCA, $G = 14.15$, $P < 0.001$; FY/RB1, $G = 10.27$, $P < 0.001$; FY/GC1F, $G = 4.93$, $P < 0.05$; FY/APOA1, $G = 4.83$, $P < 0.05$). These positive associations between unlinked markers suggest the presence of genetic structure. To further understand the observed patterns of disequilibrium, we plotted for all possible pairwise marker comparisons the D_t values vs. the D_0 values (Fig. 2). D_t is the observed disequilibrium value, and was estimated using the method of Long et al. (1995). D_0 is the initial disequilibrium created in the admixture event. Assuming no disequilibrium in the parental African and European populations, the initial admixture disequilibrium in African Americans would be described by the formula $D_0 = m(1 - m)\delta_A\delta_B$, where m is the admixture proportion, and δ_A and δ_B the frequency differences between the parental populations at loci A and B (Chakraborty and Weiss, 1988). A significant correlation is observed between D_0 and D_t for all possible marker comparisons for the combined sample from the four Low Country counties ($R^2 = 0.1926$, $P < 0.01$). Higher initial disequilibrium levels are associated with higher observed disequilibrium today.

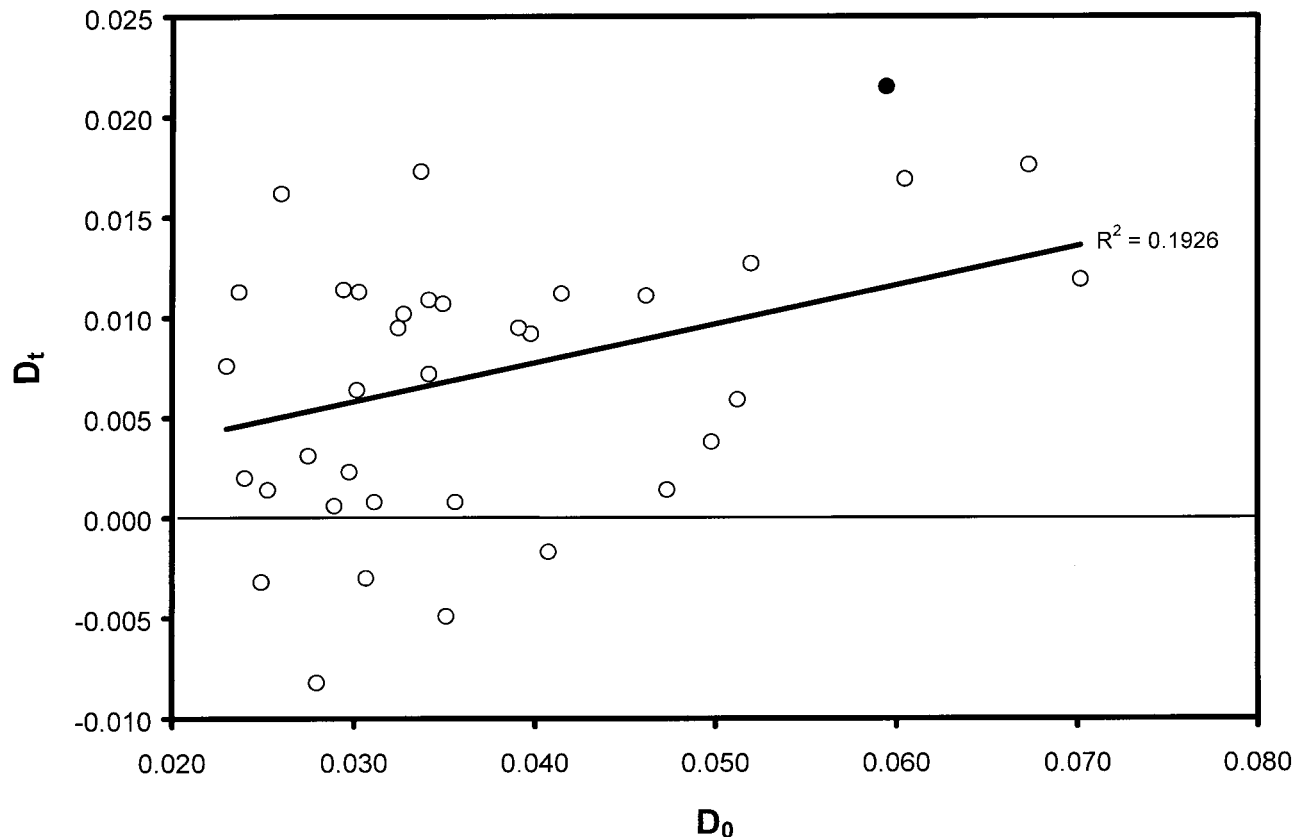


Fig. 2. Plot of relationship between initial disequilibrium (D_0) and current disequilibrium (D_t) between pairs of markers in the sample from the Low Country. Open circles, unlinked markers; solid circle, FY-AT3.

TABLE 5. Score tests for association of ancestry between locus FY and the other nine loci in the combined Low Country sample, but excluding those with missing genotype at the FY locus ($n = 509$)

Markers	Score	Percent information extracted	Z-test statistic	One-tailed P-value
FY/AT3¹	3.62	10%	1.89	0.03
FY/APOA1	1.46	5%	1.14	0.13
FY/ICAM1	-0.01	2%	-0.02	0.51
FY/LPL	-2.94	11%	-1.68	0.95
FY/D11S429	-1.25	5%	-1.09	0.86
FY/OCA	2.76	11%	1.53	0.06
FY/RB1	0.96	11%	0.51	0.31
FY/SB19.3	-0.82	4%	-0.76	0.78
FY/GC	0.84	11%	0.49	0.31

¹ FY and AT3 (in bold type) are linked on chromosome 1.

Table 5 shows the results of McKeigue's score test for association between locus FY and the other nine loci in the combined Low Country sample. As expected, no significant association is observed between FY and any of the unlinked loci, but the association between the two linked loci (FY and AT3) remains statistically significant at $P = 0.03$. The table also indicates the percentage of information extracted from the data. The analysis extracts only a proportion of the information (between 2–11%) that would be available if the ancestry (European or African) at each marker could be unambiguously assigned.

To increase the amount of information extracted, it would be necessary to type more markers in the vicinity of the available markers, more unlinked markers, and parent-offspring pairs in order to have information about phase (see McKeigue et al., 2000).

DISCUSSION

European genetic contribution to South Carolina African-Americans

We estimated the admixture proportions in six African-American samples from South Carolina: the Gullah-speaking Sea Islanders living in coastal South Carolina, four different Low Country counties, and the city of Columbia. In order to obtain good admixture estimates, the availability of informative markers and adequate samples of the parental populations are key requirements. We used a select panel of 10 autosomal markers showing high-frequency differences between African and European populations, in order to increase the precision of estimates of the genetic contribution of both parental populations to contemporary African Americans residing in these areas of South Carolina. The samples representative of the parental populations were selected on the basis of historical evidence indicating a substantial contribution to the contemporary US population of African and European ances-

try. With respect to the African continent, we based our estimates of ancestral African frequencies on data from five samples, corresponding to three different geographical regions in West Africa (Sierra Leone, Nigeria, and the Central African Republic). It has been estimated that a significant number of enslaved Africans entering the US came from these areas (Curtin, 1969). As shown in Table 1, the allele frequencies are similar for the 10 autosomal markers, both within regions and between regions in Africa, and this homogeneity reduces a possible bias in the estimates of ancestral frequencies caused by unequal contribution by areas. In order to estimate European ancestral frequencies, we used data from samples coming from England, Ireland, and Germany, which have been main sources of European migration to the US. The frequencies observed in the three European samples are not significantly different, with the exception of LPL and Sb19.3. It is important to mention that using European-American samples instead of European samples as representatives of the ancestral European population, the admixture estimates are remarkably similar to the values shown in Table 2, indicating the robustness of the results.

The results of the analysis show an interesting clinal admixture pattern. The Gullah Sea Islanders show a very low degree of European ancestry ($m = 3.5 \pm 0.9\%$), while the average of the four counties in the Low Country is $m = 11.8 \pm 1.4\%$ and in Columbia $m = 17.7 \pm 3.1\%$. No significant differences in allele frequencies or admixture proportions were observed in the neighboring counties around Charleston (Table 1), so we will use the term Low Country when referring to these pooled samples in the rest of the discussion. The standard errors associated with the admixture estimates are very low (ranging between 0.9–3.1%), emphasizing the importance of selecting informative markers for admixture studies.

Our results confirm previous studies indicating a very limited European influence in Gullah-speaking African Americans, as a result of their relative isolation through history. This isolation has resulted in the retention of numerous African characteristics in language, social organization, religion, magic, art, folklore, and music. Of particular interest is the linguistic evidence, first pointed out by Turner (1949), and extensively studied thereafter. Research in this area has indicated a very important contribution of several African languages to the Gullah Creole language, in terms of words, sounds, and grammar (reviewed by Pollitzer, 1999). In addition to the cultural evidence, anthropometric and serologic studies have also indicated a lower European contribution to the Gullah Sea Islanders than to other African-American populations in the US (Pollitzer, 1999). The 95% confidence intervals of the admixture estimates of our samples indicate that, even within South Carolina, the Gullah sample from Beaufort and Saint Helena shows a significantly lower percentage of European alleles than that ob-

served in Columbia and the neighboring counties of the Low Country (Table 2). The values obtained for the Low Country ($m = 11.8\%$), even if higher than in the Gullah, are lower than other values previously observed in other regions of the US, using the same autosomal markers (Parra et al., 1998). Given the known history of the region, it is very likely that some of the individuals of this sample are of recent Gullah ancestry. On the contrary, the m estimate of Columbia ($m = 17.7\%$) is in the range observed in many of the northern populations analyzed in our previous study (Detroit, 16.3%; Baltimore, 15.5%; Maywood, IL, 18.8%; and New York, 19.8%).

The analysis of individual admixture proportions using a maximum likelihood approach (Fig. 1) further reinforces the conclusions obtained from population-based admixture estimates. Individual admixture estimates are much less precise than population-based estimates (e.g., Long's method), and at least 40–50 markers, as informative as those analyzed here, would be necessary to have comparable precision in the estimate of the individual admixture. Thus, the distributions of individual admixture are built on the basis of estimates with high standard errors. Nevertheless, the comparison between distributions provides information regarding the differences in admixture patterns between the three South Carolina samples. The average individual admixture estimate in each sample agrees well with the values obtained using Long's method. In addition, the pattern of distribution of individual admixture further stresses the higher isolation of the Gullah Sea Islanders with respect to the other South Carolina populations. A much higher proportion of Gullah individuals (80%) shows a percentage of European ancestry equal to or lower than 10%, vs. 53% in the Low Country and 50% in Columbia (Fig. 1).

Additional information on the extent of admixture and the dynamics of the admixture process can be extracted by the analysis of continent-specific mtDNA haplogroups, which are maternally inherited, and informative Y-chromosome-specific markers, which are only transmitted from father to son. The combination of both kinds of markers allows the possibility of evaluating male- and female-specific gene flow. In accordance with previous data on other African-American populations through the US (Parra et al., 1998), a sex-biased European gene flow is also observed in the populations analyzed here. The European genetic contribution from males was higher than the female contribution (Table 3). In the Gullah Sea Islanders, the estimate of European admixture from YAP is 5.6%, higher than the estimate obtained from autosomal markers (3.5%), and we have not detected any European-specific H haplogroups. In the sample from the Low Country, which comprises only females (see Material and Methods), it was not possible to evaluate the male genetic contribution, but again the estimate based on

mtDNA (0.9%) is substantially lower than the autosomal estimate (11.8%). Finally, in Columbia the same pattern is also evident, with a much larger European contribution from males ($m = 24.1\%$) than from females ($m = 2.6\%$).

An interesting finding of this study has been the presence of Native American ancestry in these three African-American populations. In order to evaluate this influence, we typed the samples for the Native American haplogroups A, B, C, and D, and the Y-chromosome-specific DYS199 polymorphism. Most of the mtDNA of Native Americans belongs to the aforementioned haplogroups. Recently, a fifth founding haplogroup, termed haplogroup X, was confirmed in a number of studies (Bailliet et al., 1994; Forster et al., 1996; Scozzari et al., 1997; Smith et al., 1999). With the exception of some Canadian Ojibwa (i.e., those who were studied in this large geographically dispersed population), in which haplogroup X is relatively frequent (Scozzari et al., 1997), this haplogroup is rare among modern Native Americans (around 3%; Smith et al. 1999), and has not been found in relevant southeastern tribes, such as the Seminole, Cherokee, and Creek (Huoponen et al., 1997; Smith et al., 1999). Thus, an estimate of maternal Native American contribution in South Carolina African Americans can be calculated simply as the total percentage of Native American haplogroups observed in the samples. An important consideration is that one of the Native American-specific haplogroups, haplogroup B, is defined by a 9-bp deletion located in the intergenic region between the cytochrome oxidase II and lysine tRNA genes, and it is known that this deletion may have multiple geographic origins. This deletion, in addition to being common in Native Americans and other Asian populations, has occurred independently in individuals of African ancestry (Chen et al., 1995; Soodyal et al., 1996; Alves-Silva et al., 1999), and more rarely of European ancestry as well (Barrientos et al., 1995; Torroni et al., 1995; Alves-Silva et al., 1999). In order to distinguish the geographic source of the 9-bp deletion, we typed the individuals with the deletion for the sites *AluI* 10397 and *DdeI* 10394. Individuals with the 9-bp deletion of Native American origin (haplogroup B) typically lack both restriction sites (—), while individuals with the deletion of African or European origin lack the *AluI* 10397 site but have the *DdeI* 10394 site (—+) (Alves-Silva et al., 1999). A complementary perspective to the Native American female gene flow is given by the analysis of the nonpseudoautosomal region of the Y-chromosome, inherited only through the male line. A very informative polymorphism in this region for determining Native American male gene flow is the DYS199 C→T transition initially described by Underhill et al. (1996). The DYS199T allele is almost exclusively restricted to Native American populations, and is present in high frequencies in North, Central, and South America (Karafet et al., 1997; Lell et al., 1997; Rodriguez-Delfin et al., 1997;

Santos et al., 1999). In the Seminole tribe of Florida, the DYS199T frequency has been estimated at 48% (Lell et al., 1997).

We observed two B haplogroups in the Gullah sample, giving an approximate estimate of 2.4% Native American maternal contribution (Table 3). In the much larger sample from the Low Country ($n = 541$), we detected the presence of all Native American haplogroups (3 haplogroup A, 4 haplogroup B, 1 haplogroup C, and 1 haplogroup D), and the estimate of global Native American maternal contribution is 1.6%. In Columbia, the Native American influence is also evident (1 A haplogroup, $m = 1.1\%$). Interestingly, we have not observed any DYS199T genotype in any of the Gullah ($n = 25$) and Columbia ($n = 90$) males examined, suggesting that the male Native American contribution has not been significant. Unfortunately, these samples are relatively small, and the large sample from the Low Country is entirely comprised of women, so it is difficult to evaluate accurately sex bias in Native American gene flow to these African American populations.

Our results agree with historical facts indicating that admixture between Native Americans and people of African ancestry took place in South Carolina. Despite early laws against Native American slavery, many colonists disregarded the sanctions, and Native Americans were publicly sold as slaves in Charleston. After 1720, Native American slavery declined, and enslaved Indians were absorbed into the black community. The term *mustizoes* or *mustees* was coined to refer to the offspring of Indians and blacks (Pollitzer, 1999). Some early anthropological reports pointed out the high proportion of African American college students claiming some Native American ancestry (Herskovits, 1930; Meier, 1949). Judging from the mtDNA and Y-chromosome-specific data, the global Native American contribution to the southern African-American populations analyzed in this study has been quantitatively low, probably not higher than 1–2% (Table 4).

Pairwise association between autosomal markers: inferences about the admixture process

The analysis of the pattern of pairwise association between the 10 autosomal markers analyzed here can provide interesting insights on the admixture process and the genetic structure of South Carolina African-American populations. The association observed between FY and AT3 is indicative of the linkage disequilibrium created when the two parental populations admixed. Both theoretical (Stephens et al., 1994; Briscoe et al., 1994; McKeigue, 1997) and experimental studies (Parra et al., 1998; Lautenberger et al., 2000) have shown that this linkage disequilibrium resulting from admixture can span wide genomic intervals (more than 10 cM). Associations between unlinked markers can be informative about the extent of genetic structure and the pattern of gene flow. In a nonstructured population, we

would expect to find no associations between unlinked markers. In a population with genetic structure, the associations may be significant even between unlinked markers. Among the possible factors causing genetic structure, assortative mating and continuous gene flow could be relevant in South Carolina African Americans. We recently simulated two different models of admixture and its consequences in terms of associations between linked and unlinked markers (Pfaff and Shriver, 1999; Pfaff et al., in preparation). In admixed populations in which there is an initial process of admixture and further isolation from the parental populations (hybrid isolation model), after 15 generations a significant association is found between linked markers located 10 cM apart, and only low levels of association between unlinked markers (~5%). On the contrary, in admixed populations in which there is a constant input of alleles from one parental population to the admixed population every generation (continuous gene flow model), after 15 generations significant associations are observed between very close markers, but also in a large percentage of unlinked marker comparisons (~60%). Consequently, the analysis of the pattern of association, especially when using informative markers, can provide interesting insights into the admixture process.

In the Gullah sample, no significant association was detected either between FY and AT3 or any of the pair-wise comparisons between FY and the remaining nine markers. This is to be expected, given both the small sample size and the very low European contribution to the Gullah population, which is reflected in just six FY-null + alleles in a sample of 166 alleles. Thus, the power for detecting associations in this sample is very limited. In Columbia, the only significant association is observed between the linked markers FY and AT3 ($G = 8.87$, $P < 0.01$). The sample from the Low Country is much more informative in terms of sample size. In this case, there is a very significant association between FY and AT3 ($G = 16.70$, $P < 0.001$), a higher value than any other comparison of FY with the remaining markers (Table 4). Interestingly, in the Low Country, 4 of the 8 comparisons of FY with unlinked markers turned out to be significant as well (FY/APOA1, FY/GC, FY/OCA2, and FY/RB), indicating the presence of genetic structure. Although we have not observed the same pattern in previous studies with much smaller samples (including the sample from Columbia), recently in a study of more than 900 African Americans living in Jackson, Mississippi, 7 of 8 comparisons of FY with unlinked markers were also significant (Parra et al., 1999). In all these significant comparisons, both in Jackson and the Low Country, there is an excess of haplotypes having the two European-specific alleles, as expected from the admixture process. These results strongly resemble our simulations of the continuous gene flow model, in which even a small

input of European alleles per generation (around 2%) results in a high percentage of significant associations (data not shown). Additional evidence in support for this model comes from the plot of the current disequilibrium (D_t) vs. the initial disequilibrium (D_0) (Fig. 2). Under a hybrid isolation model, we would expect the disequilibrium between unlinked markers to be reduced drastically after some generations. The decay of disequilibrium follows the formula $D_t = (1 - r)^t D_0$, where D_t is the linkage disequilibrium t generations since admixture, D_0 is the initial disequilibrium, and r is the recombination fraction (Chakraborty and Weiss, 1988). This means that, for example, after 10 generations (a reasonable estimate of the minimum number of generations people of African ancestry have been living in the US), the disequilibrium at unlinked loci would be reduced to 0.1% of the initial level. Under the same number of generations, in loci located at 10 cM and 1 cM apart, the disequilibrium due to true linkage would be 34.9% and 90.4%, respectively, of the initial level. Consequently, we would expect to observe, for unlinked markers, a plot between D_0 and D_t in which the current disequilibrium would be near 0, irrespective of the initial disequilibrium. This is not the case in the sample from the Low Country, in which most D_t values are positive, relatively high, and significantly correlated with D_0 . This, again, is best explained under a model in which some disequilibrium is introduced each generation through continuous gene flow, although other mechanisms, such as assortative mating, could potentially produce this effect.

It is obvious that the genetic structure created as a result of continuous gene flow can pose problems for association studies in African-American populations. It is therefore necessary to control for such structure using appropriate statistical methods. One of the solutions is to use family-based association tests, such as the transmission-disequilibrium test (TDT) (Ewens and Spielman, 1995; McKeigue, 1997). Recently, another method for detecting linkage in admixed populations was proposed in the literature (McKeigue, 1998; McKeigue et al., 2000). We applied this method to our data and, as expected, controlling for the overall admixture of each individual eliminates associations between unlinked markers, but not the association between linked markers, as can be observed in the case of FY and AT3 (Table 5). Therefore, the excess association between linked markers can be distinguished from the background association between unlinked markers caused by genetic structure, and this kind of approach will be very useful for mapping genes involved in complex diseases showing prevalence differences in major population groups (e.g., hypertension, obesity, type 2 diabetes).

CONCLUSIONS

The current availability of very informative autosomal, mtDNA, and Y-chromosome-specific markers has made possible a much more detailed analysis of admixed populations than in the past. This analysis can give us interesting genetic information that can be compared with other historical, cultural, and anthropological evidence in order to get a better picture of admixture processes and dynamics. The results of our study in South Carolina confirm previous findings, based on multiple lines of evidence, indicating that the relative isolation of the Gullah Sea Islanders has resulted in a lower European admixture and a higher retention of African features than in other African-American populations. An interesting clinal admixture pattern is observed in three samples from South Carolina. Evidence of a sex-biased European gene flow and the presence of a low amount of Native American admixture are also obvious from the mtDNA and the Y-chromosome-specific analysis. Finally, the pattern of admixture seems to conform to a continuous gene flow model, which creates genetic structure and is reflected in significant associations between unlinked genetic markers, although other factors, such as assortative mating, could also explain the observed patterns. These spurious associations can pose a problem in case-control studies trying to elucidate the genetic basis of complex diseases, but by using appropriate statistical tests, it is possible to differentiate between the association due to linkage and the association due to genetic structure introduced by the continuous gene flow process.

ACKNOWLEDGMENTS

We thank Dr. Edward Williams (Director, Division of Analytical Chemistry, Bureau of Laboratories, South Carolina Department of Health and Environmental Control) for samples from the Low Country. We also thank Sally Weinrich, Ph.D., R.N., for providing samples from Columbia, SC. Finally, we express our gratitude to all the participants for their collaboration in this study. This work was supported in part by grants from NIH/NIDDK (DK53958) and NIH/NHGRI (HG02154) to M.D.S.

LITERATURE CITED

- Adams J, Ward RH. 1973. Admixture studies and detection of selection. *Science* 180:1137-1143.
- Alves-Silva J, Santos MS, Pena SDJ, Prado VF. 1999. Multiple geographic sources of region V 9-bp deletion haplotypes in Brazilians. *Hum Biol* 71:245-259.
- Bailliet G, Rothhammer F, Carnese FR, Bravi CM, Bianchi NO. 1994. Founder mitochondrial haplotypes in Native Americans populations. *Am J Hum Genet* 54:27-33.
- Barrientos A, Casademont J, Solans A, Moral P, Cardellach F, Urbano-Marquez A, Estivill X, Nunes V. 1995. The 9-bp deletion in the region V of the mitochondrial DNA: evidence of mutation recurrence. *Hum Genet* 96:225-228.
- Briscoe D, Stephens JC, O'Brien SJ. 1994. Linkage disequilibrium in admixed populations: applications in gene mapping. *J Hered* 85:59-63.
- Chakraborty R, Weiss KM. 1988. Admixture as a tool for finding linked genes and detecting that difference from allelic association between loci. *Proc Natl Acad Sci USA* 85:9119-9123.
- Chakraborty R, Ferrell RE, Stern MP, Haffner SM, Hazuda HP, Rosenthal M. 1986. Relationship of prevalence of non-insulin dependent diabetes mellitus to Native American admixture in the Mexican Americans of San Antonio, Texas. *Genet Epidemiol* 3:435-454.
- Chakraborty R, Kamboh MI, Nwankwo M, Ferrell RE. 1992. Caucasian genes in American blacks: new data. *Am J Hum Genet* 50:145-155.
- Chen Y-S, Torroni A, Excoffier L, Santachiara-Benerecetti AS, Wallace DC. 1995. Analysis of mtDNA variation in African-populations reveals the most ancient of all human continent-specific haplogroups. *Am J Hum Genet* 57:133-149.
- Curtin P. 1969. *The Atlantic slave trade*. Madison, WI: University of Wisconsin Press.
- Elston RC. 1971. The estimation of admixture in racial hybrids. *Ann Hum Genet* 35:9-17.
- Ewens WJ, Spielman RS. 1995. The transmission/disequilibrium test: history, subdivision and admixture. *Am J Hum Genet* 57:455-464.
- Fernandez-Reyes D, Craig AG, Kyes SA, Peshu N, Snow RW, Berendt AR, Marsh K, Newbold CI. 1997. A high frequency African coding polymorphism in the N-terminal domain of ICAM1-1 predisposing to cerebral malaria in Kenya. *Hum Mol Genet* 6:1357-1360.
- Forster P, Harding R, Torroni A, Bandelt H-J. 1996. Origin and evolution of Native American mtDNA variation: a reappraisal. *Am J Hum Genet* 59:935-945.
- Herskovits M. 1930. *The anthropometry of the American Negro*. Columbia Univ Contrib Anthropol 11.
- Huoponen K, Torroni A, Wickman PR, Sellitto D, Gurley DS, Scozzari R, Wallace DC. 1997. Mitochondrial DNA and Y chromosome-specific polymorphisms in the Seminole tribe of Florida. *Eur J Hum Genet* 5:25-34.
- Johnson DM, Campbell RR. 1981. *Black migration in American: a social demographic history*. Durham, NC: Duke University Press.
- Karafet T, Zegura SL, Vuturo-Brady J, Posukh O, Osipova L, Wiebe V, Romero F, Long JC, Harihara S, Jin F, Dashnyam B, Gerelsaikhan T, Omoto K, Hammer MF. 1997. Y chromosome markers and trans-Bering Strait dispersals. *Am J Phys Anthropol* 102:301-314.
- Lautenberger JA, Stephens JC, O'Brien SJ, Smith MW. 2000. Significant admixture disequilibrium across 30 cM around the FY locus in African Americans. *Am J Hum Genet* 66:969-978.
- Lell JT, Brown MD, Schurr TG, Sukernik RI, Starikovskaya YB, Torroni A, Moore LG, Troup GM, Wallace DC. 1997. Y chromosome polymorphisms in Low Country and Siberian populations: identification of Low Country Y chromosome haplotypes. *Hum Genet* 100:536-543.
- Long JC. 1991. The genetic structure of admixed populations. *Genetics* 127:417-428.
- Long JC, Smouse PE. 1983. Intertribal gene flow between the Ye'cuana and Yanomamö: genetic analysis of an admixed village. *Am J Phys Anthropol* 61:411-422.
- Long JC, Williams RC, Urbanek M. 1995. An E-M algorithm and testing strategy for multiple-locus haplotypes. *Am J Hum Genet* 56:799-810.
- McKeigue PM. 1997. Mapping genes underlying ethnic differences in disease risk by linkage disequilibrium in recently admixed populations. *Am J Hum Genet* 60:188-196.
- McKeigue PM. 1998. Mapping genes that underlie ethnic differences in disease risk: methods for detecting linkage in admixed populations by conditioning on parental admixture. *Am J Hum Genet* 63:241-251.
- McKeigue PM, Carpenter JR, Parra EJ, Shriver MD. 2000. Estimation of admixture and detection of linkage in admixed populations by a Bayesian approach: application to African-American populations. *Ann Hum Genet* 64:171-186.
- Meier A. 1949. A study of the racial ancestry of the Mississippi college Negro. *Am J Phys Anthropol* 7:227-240.

- Parra EJ, Marcini A, Akey J, Martinson J, Batzer MA, Cooper R, Forrester T, Allison DB, Deka R, Ferrell RE, Shriver MD. 1998. Estimating African American admixture proportions by use of population-specific alleles. *Am J Hum Genet* 63:1839–1851.
- Parra EJ, Hiester KG, Norton HL, Baker JT, Bonilla C, Gulden FO, Hutchinson RG, Ferrell RE, Boerwinkle E, Shriver M. 1999. African-American hypertensive subjects have less European ancestry than normotensive controls. *Am J Hum Genet [Suppl]* 65:33.
- Pfaff CL, Shriver MD. 1999. Power for mapping by admixture linkage disequilibrium: a case/control simulation study. *Am J Hum Genet [Suppl]* 65:440.
- Pollitzer WS. 1999. The Gullah people and their African heritage. Athens, GA: University of Georgia press.
- Raymond M, Rousset F. 1995. Genepop, version 2.0.
- Reed TE. 1969. Caucasian genes in American Negroes. *Science* 165:762–768.
- Rodriguez-Delfin L, Santos SEB, Zago MA. 1997. Diversity of the human Y chromosome of South American Amerindians: a comparison with blacks, whites, and Japanese from Brazil. *Ann Hum Genet* 61:439–448.
- Santos FR, Pandya A, Tyler-Smith C, Pena SD, Schanfield M, Leonard WR, Osipova L, Crawford MH, Mitchell RJ. 1999. The central Siberial origin for Native American Y chromosomes. *Am J Hum Genet* 64:619–628.
- Scozzari R, Cruciani F, Santolamazza P, Sellitto D, Cole DEC, Rubin LA, Labuda D, Marini E, Succa V, Vona G, Torroni A. 1997. mtDNA and Y chromosome-specific polymorphisms in modern Ojibwa: implications about the origin of their gene pool. *Am J Hum Genet* 60:241–244.
- Shriver MD, Smith M, Jin L, Marcini A, Akey JM, Deka R, Ferrell RE. 1997. Forensic ethnic affiliation estimation by use of population-specific allele DNA markers. *Am J Hum Genet* 60:957–964.
- Smith DG, Malhi RS, Eshleman J, Lorenz JG, Kaestle FA. 1999. Distribution of mtDNA haplogroup X among Native North Americans. *Am J Phys Anthropol* 110:271–284.
- Soodyall H, Vigilant L, Hill AV, Stoneking M, Jenkins T. 1996. mtDNA control-region sequence variation suggests multiple independent origins of an “Asian-specific” 9-bp deletion in sub-Saharan Africans. *Am J Hum Genet* 58:595–608.
- Stephens JC, Briscoe D, O’Brien SJ. 1994. Mapping by admixture linkage disequilibrium in human populations: limits and guidelines. *Am J Hum Genet* 55:809–824.
- Torroni A, Petrozzi M, Santolamazza P, Sellitto D, Cruciani F, Scozzari R. 1995. About the “Asian”-specific 9-bp deletion of mtDNA. *Am J Hum Genet* 57:507–508.
- Turner LD. 1949. Africanisms in the Gullah dialect. Chicago: University of Chicago Press.
- Underhill PA, Jin L, Zemans R, Oefner P, Cavalli-Sforza LL. 1996. A pre-Columbian Y chromosome-specific transition and its implications for human evolutionary history. *Proc Natl Acad Sci USA* 93:196–200.
- Watson E, Forster P, Richards M, Bandelt HJ. 1997. Mitochondrial footprints of human expansions in Africa. *Am J Hum Genet* 61:691–704.
- Workman PL. 1968. Gene flow and the search for natural selection in man. *Hum Biol* 40:260–279.