Am. J. Hum. Genet. 69:1080–1094, 2001

Markers for Mapping by Admixture Linkage Disequilibrium in African American and Hispanic Populations

Michael W. Smith,[1] James A. Lautenberger,[2] Hyoung Doo Shin,[1,*] Jean-Paul Chretien,[2,3] Sadeep Shrestha,[1,4] Dennis A. Gilbert,[5] and Stephen J. O'Brien[2]

[1]Intramural Research Support Program, Science Applications International Corporation–Frederick, and [2]Laboratory of Genomic Diversity, National Cancer Institute at Frederick, Frederick, MD; [3]Welch Center for Prevention, Epidemiology, and Clinical Research, and [4]Department of Epidemiology, Bloomberg School of Public Health, Johns Hopkins University, Baltimore; and [5]Celera Genomics, Foster City, CA

Population linkage disequilibrium occurs as a consequence of mutation, selection, genetic drift, and population substructure produced by admixture of genetically distinct ethnic populations. African American and Hispanic ethnic groups have a history of significant gene flow among parent groups, which can be of value in affecting genome scans for disease-gene discovery in the case-control and transmission/disequilibrium test designs. Disease-gene discovery using mapping by admixture linkage disequilibrium (MALD) requires a map of polymorphic markers that differentiate between the founding populations, along with differences in disease-gene allele frequencies. We describe markers appropriate for MALD mapping by assessing allele frequencies of 744 short tandem repeats (STRs) in African Americans, Hispanics, European Americans, and Asians, by choosing STR markers that have large differences in composite δ, log-likelihood ratios, and/or I*(2) for MALD. Additional markers can be added to this MALD map by utilization of the rapidly growing single-nucleotide–polymorphism databases and the literature, to achieve a 3–10-cM scanning scale. The map will be useful for studies of diseases, including prostate and breast cancer, diabetes, hypertension, and end-stage renal disease, that have large differences in incidence between the founding populations of either Hispanics or African Americans.

Introduction

The analysis of complex human diseases requires novel genetic strategies and approaches as we enter the known genomic sequence era. Approaches that involve the use of traditional family linkage analysis have yielded the locations of many genes, especially those that are highly penetrant and encode simple Mendelian disease phenotypes. More recently, use of sib-pair analysis, the transmission/disequilibrium test (TDT), and homozygosity mapping have made the identification of the genes involved in complex diseases more tractable (Risch and Merikangas 1996; Risch 2000). Whole-genome scans have identified genetic regions and genes involved in many diseases, including type I diabetes, asthma, prostate cancer, and others (e.g., Smith et al. 1996; Mein et al. 1998; Arngrimsson et al. 1999; The Tourette Syndrome Association International Consortium for Genetics 1999; Bellamy et al. 2000; Walder et al. 2000; Wiggs et al. 2000). Although these family-based approaches are powerful and make possible the identification of genes involved in many complex diseases, some diseases in which environmental and viral factors are important components may be best addressed by approaches that center around a case-control and TDT design.

The detection of polymorphic genes that influence quantitative traits, disease states, and other characters is the goal of population genetic association studies, but it depends upon the persistence of measurable linkage disequilibrium (i.e., haplotype allele association) between markers and undiscovered loci. In white populations, the extent and usefulness of linkage disequilibrium is generally limited to regions smaller than ~100 kb, because of recent population history (Bodmer 1986; Laan and Pääbo 1997; Huttley et al. 1999; Reich et al. 2001). The power of this approach depends upon how far linkage disequilibrium extends over a chromosomal interval which, in turn, determines the spacing and number of markers required for a genome scan.

One promising approach is mapping by admixture linkage disequilibrium (MALD), where the samples are collected from an admixed population in patient cohorts (Briscoe et al. 1994; Stephens et al. 1994; McKeigue 1997, 1998; Kaplan et al. 1998; Zheng and Elston 1999). These theoretical treatments and simulations

Received March 21, 2001; accepted for publication August 20, 2001; electronically published October 5, 2001.

Address for correspondence and reprints: Dr. Michael W. Smith, Science Applications International Corporation–Frederick, National Cancer Institute at Frederick, P. O. Box B, 7th Street Extension, Frederick, MD 21702-1201. E-mail: smithm@ncicrf.gov

* Present affiliation: Department of Epidemiology, SNP Genetics, Seoul, South Korea.

0002-9297/2001/6905-0017$02.00

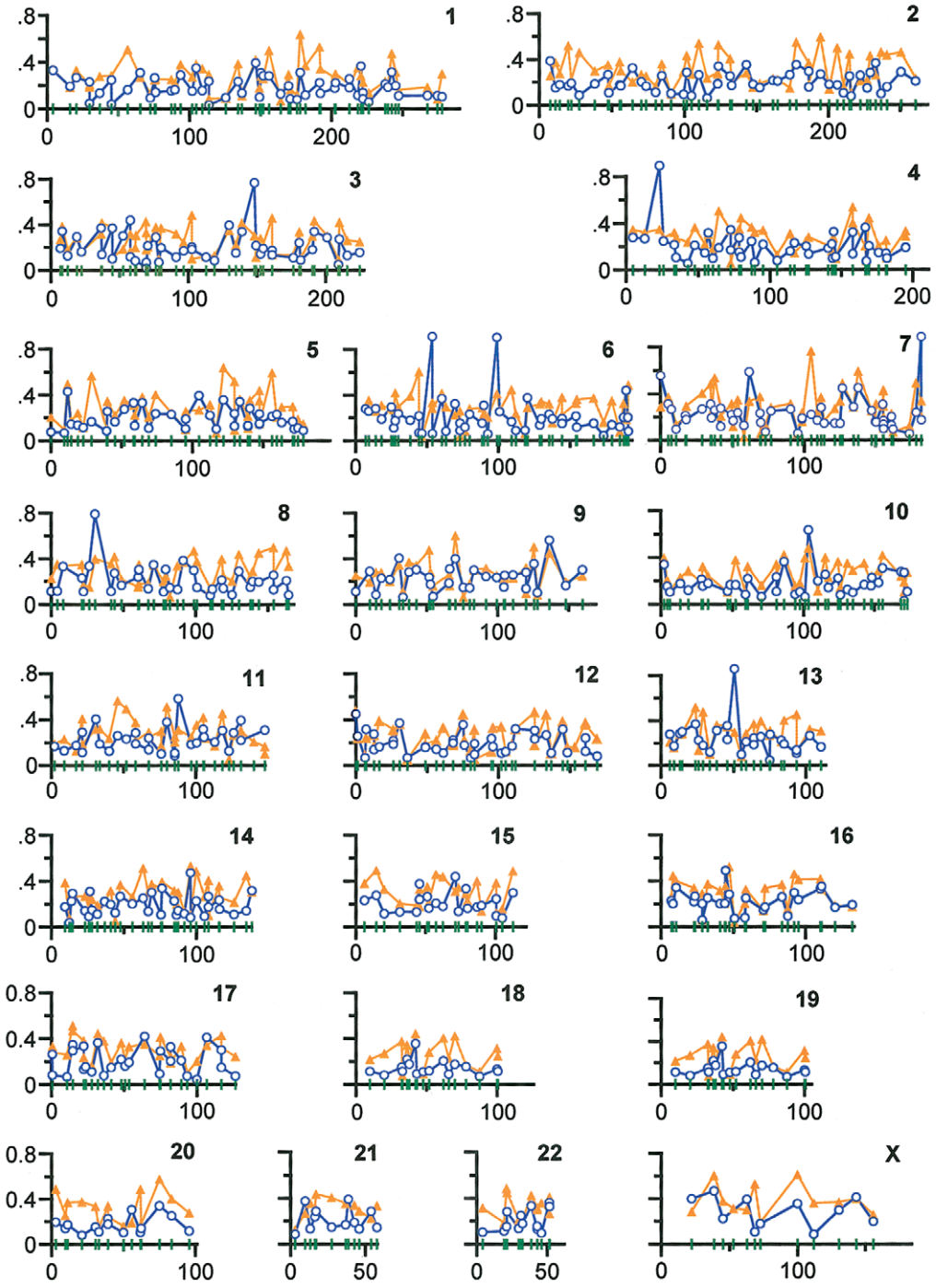

Figure 1 δ_c values for the loci examined across the human genome, in comparisons between European Americans and African Americans (*shaded triangles*) and between European Americans and Hispanics (*white circles*). δ_c values are shown on the Y-axis, and chromosome position (in centimorgans) is shown on the X-axis.

point out that recent admixture generates linkage disequilibrium that can extend for many centimorgans and can persist for as many as 20 generations. We have recently detected admixture linkage disequilibrium (ALD) across tens of centimorgans around the FY (Duffy) gene in African Americans (Lautenberger et al. 2000).

African Americans and Hispanics seem ideal for MALD-based association ascertainment. Studies have shown that African Americans represent an admixed

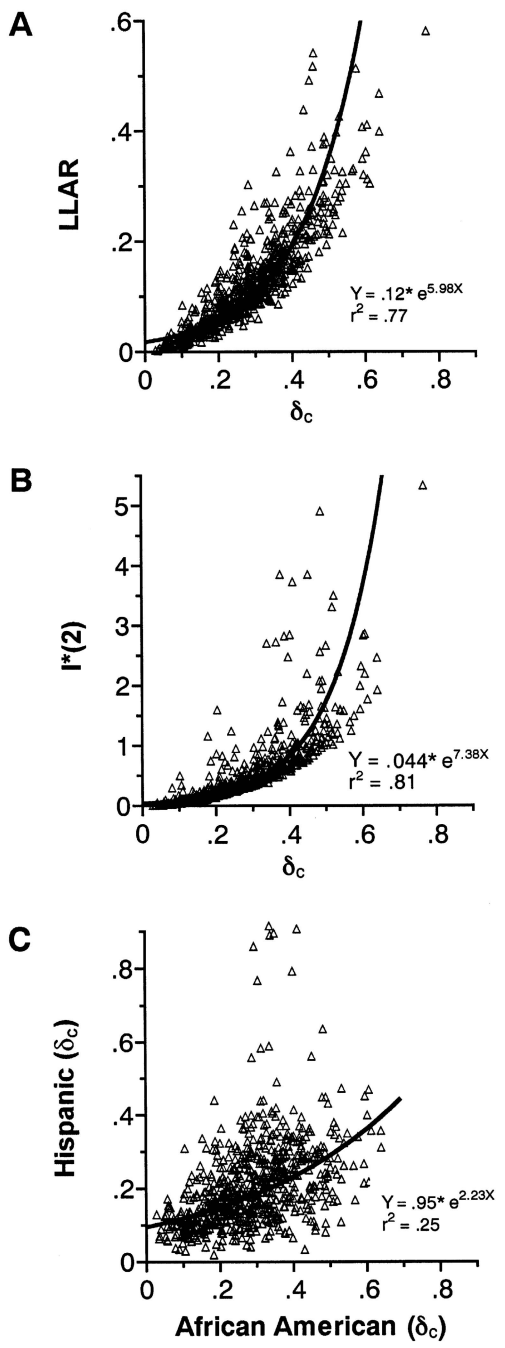

Figure 2 Relationship between differences seen at individual markers in δ_c and LLAR (*A*) and STR I*(2) (*B*) in African Americans, along with African American versus Hispanic δ_c values (*C*).

population with significant genetic contributions from both African and European ancestors (Chakraborty and Weiss 1988; Chakraborty et al. 1991). Recent estimates of the proportion of European genes in African American populations range from 6.8% for Sapelo Island in Georgia to 26% for Chicago (Long 1991; Chakraborty et al. 1992; Parra et al. 1998; Destro-Bisol et al. 1999). Hispanics—a complex U.S. ethnic group that includes Puerto Ricans, Cubans, Mexican Americans, and Spanish Americans—also constitute an admixed population of primarily European, 18%–31% Native American, and 3%–31% African origins (Hanis et al. 1991; Long et al. 1991), which is promising for MALD analysis.

Earlier studies of RFLPs suggested that establishing a collection of differentiating markers would be difficult to achieve with single-nucleotide polymorphisms (SNPs), where at most only 20% of 257 markers had large enough differences to be informative for MALD mapping (Dean et al. 1994), whereas subsequent work on short tandem repeat polymorphisms (STRs) suggested that about half had large differences (Bowcock et al. 1994). Current efforts of the SNP consortium (Altshuler et al. 2000) are likely to bring these biallelic markers to the forefront for MALD mapping in a case-control and TDT setting. However, the more-polymorphic STRs provide higher information content for TDT and case-control approaches, and, given the current state of genotyping technology, an STR-based MALD map provides a valuable gene-mapping resource.

In the present study, we sought to identify markers appropriate for MALD analysis, by genotyping of African Americans, Europeans, Hispanics, and Asians, using 421 STR loci and supplementing the data set with data from 323 markers from an asthma genome scan (Collaborative Study on the Genetics of Asthma 1997). These data were used to estimate allele frequencies and the usefulness of the loci for MALD mapping. Since MALD assessment provides remarkable potential for the discovery of novel genes involved in common diseases, the comprehensive set of markers with large differences between the founding populations for African Americans and Hispanics provides a foundation for future MALD gene localization studies.

Subjects and Methods

Patient DNAs were obtained from collections of human DNAs at the Laboratory of Genomic Diversity and included 45 African Americans, 45 Europeans, 45 Hispanics, and 40 Asians (Dean et al. 1994; Smith et al. 1997; O'Brien 2000; O'Brien et al. 2000). Early in the study, a different set of patients was used with fewer individuals (37 African Americans, 25 European Americans, 21 Hispanics, and 21 Asians), with the African American samples containing 18 parent/offspring pairs. DNAs from lymphoblastoid or fibroblast cell lines were extracted using methods we have published elsewhere (Dean et al. 1994). Some of the allele-frequency data have been reported elsewhere as part of an HIV-1/AIDS candidate gene analysis (Shin et al. 2000) or an asthma

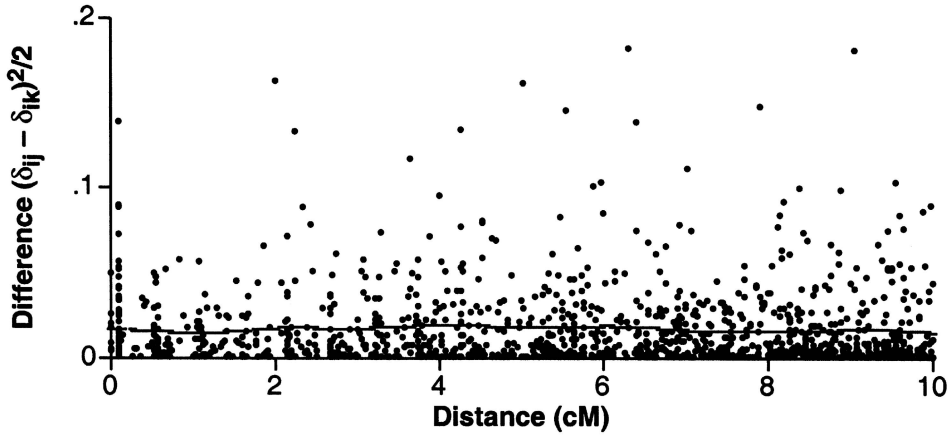

Figure 3 Variograms of marker δ_c in African Americans. For each point, the *X* and *Y* coordinates represent the map distance between markers *j* and *k* on chromosome *i* (map location of *j* > map location of *k*) and half the squared difference of δ_c between the markers, respectively (Diggle et al. 1994). All possible pairs of markers <50 cM apart were examined, and those at intervals of ≤10 cM are shown. One observation was off the scale, with a distance of 9 cM and a difference of .25. The line is the estimated kernel smoothing function.

genetics genome scan (Collaborative Study on the Genetics of Asthma 1997).

STR locus primers were obtained from a variety of sources, including (1) commercial STR panels that were in development (Applied Biosystems), (2) the Applied Biosystems X chromosome STR kit, (3) ongoing HIV-1/AIDS projects (O'Brien et al. 2000; Shin et al. 2000), (4) work around the FY gene (Lautenberger et al. 2000), and (5) experiments designed to fill gaps in the MALD map with additional STR loci. Amplification was performed with Perkin-Elmer 9600 thermal cyclers. Loci were amplified with Ampli*Taq* DNA polymerase under the following conditions: 2 min at 95°C; 10 cycles of 30 s at 94°C, 15 s at 55°C, and 15 s at 72°C; 20 cycles with a lowered (89°C) denaturation temperature, followed by a 72°C final extension for 10 min. In addition, a *Taq* gold (PE Biosystems) touchdown protocol was also used later in the project; this protocol consisted of 10 min at 95°C; 10 cycles of 30 s at 94°C, 30 s at 65°C, and 30 s at 72°C; 20 cycles of the same conditions but dropping the annealing temperature by 0.5°C, to 55°C; 15 cycles of annealing at 55°C; and a 72°C final extension for 10 min. Loci that yielded banding patterns characteristic of +A addition were tried again, using a 90-min final extension, no final extension, and/or by redesigning the unlabeled reverse primer to add a guanine or to finish with the sequence of GTTT (G/A/C) at the 5′ end (Brownstein et al. 1996; Magnuson et al. 1996). Primer sequences and allele size ranges for the primers we designed are available at the Laboratory of Genomic Diversity Web site. Fluorescently labeled PCR products (FAM, HEX, TET, and NED) were separated on Applied Biosystems 373 and 377 sequencers. Gels were analyzed with Genescan collection and analysis software, and genotypes were called using Genotyper software (Applied Biosystems). Alleles were binned using linear regression, visual examination, and Genotyper software. Data were analyzed using the Statistical Analysis System (SAS) (SAS Institute, Inc.). Estimates of composite δ (δ_c) and log-likelihood allelic ratio (LLAR) values (Shriver et al. 1997; Stephens et al. 1999) were computed by SAS. The δ_c value is defined as the sum of the absolute value of all *n* allelic frequency (f_i) differences divided by 2:

$$\delta_c = \frac{1}{2} \times \sum_{i=1}^{n} |f_{iA} - f_{iB}| \ ,$$

where f_{iA} and f_{iB} are the frequencies of the *i*th allele in the two groups, A and B, being compared at a locus. The LLAR statistic was calculated over all *n* alleles as

$$\text{LLAR} = \frac{1}{2}\sum_{i=1}^{n} f_{iA} \log \frac{f_{iA}}{f_{iB}} + \frac{1}{2}\sum_{i=1}^{n} f_{iB} \log \frac{f_{iB}}{f_{iA}} \ .$$

A program written in Pascal was used to calculate the MALD-TDT (transmission/disequilibrium test) allele-collapsing statistic, I*(2) (Kaplan et al. 1998). Regression analysis of these comparison measures were first examined as linear models, and then curvilinear terms were added to better fit the residuals. Autocorrelation of δ_c values for the comparison of European Americans versus both African Americans and Hispan-

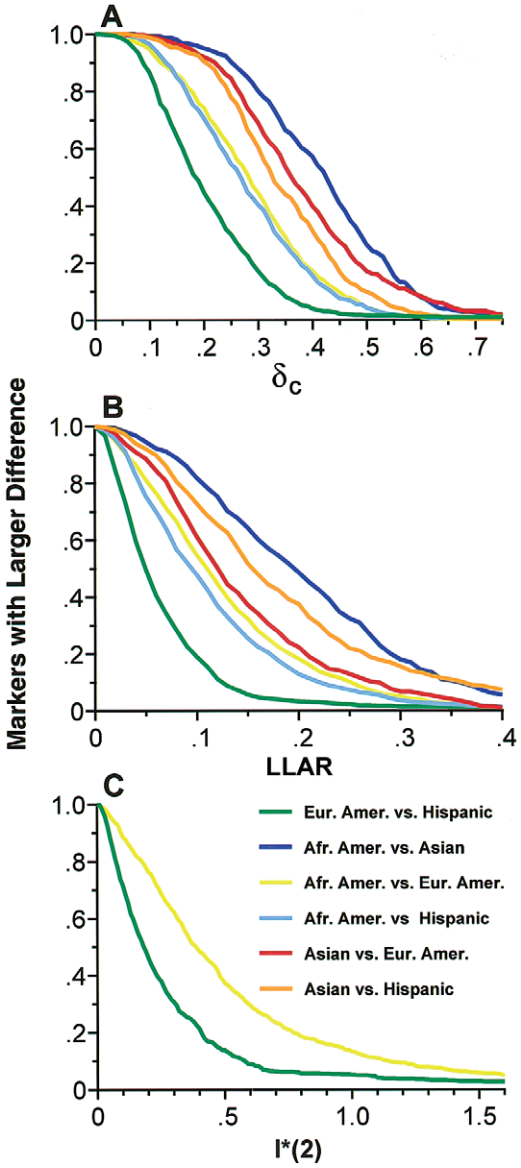

Figure 4 Cumulative frequency distributions of differences between African Americans, Asians, Hispanics, and European Americans are shown as δ_c (*A*), LLAR (*B*), and optimized STR allele-collapsing statistic I*(2) (*C*) (Kaplan et al. 1998).

ics was examined using longitudinal data analysis techniques (Diggle et al. 1994).

Results

Estimated allele frequencies from the 744 STR loci examined are available at the Laboratory of Genomic Diversity Web site. Those allele frequency estimates were used to determine differences between the four racial/ethnic groups. Comparisons of African Americans versus Asians, African Americans versus European Americans, African Americans versus Hispanics, Asians versus European Americans, Asians versus Hispanics, and Hispanics versus European Americans were calculated as (1) δ_c, one-half the sum of the absolute value of the allele frequency differences (Shriver et al. 1997; Stephens et al. 1999; Lautenberger et al. 2000) and (2) the LLAR estimate of the discrimination power of each locus derived from some of our previous work (Shriver et al. 1997). The comparisons of African Americans versus European Americans and of European Americans versus Hispanics were evaluated as the optimal I*(2) (Kaplan et al. 1998). Values of δ_c for the African American versus European American and the European American versus Hispanic comparisons are plotted by chromosome position in figure 1.

A comparison of the behavior of the three MALD statistics—δ_c, LLAR, and I*(2)—shows a high level of correlation. For example, in the comparison of 724 loci between African Americans and European Americans, the correlation coefficient of LLAR versus δ_c was .88, with $Y = 0.12 \times e^{5.98X}$ (fig. 2*a*). Similar results were obtained from the regression of I*(2) versus δ_c in the same ethnic group comparison ($r^2 = .81$; $Y = 0.044 \times e^{7.38X}$; fig. 2*b*). Some of the strengths and limitations of these different MALD statistics have been discussed elsewhere (Shriver et al. 1997; Kaplan et al. 1998; Stephens et al. 1999).

The distribution of δ_c was examined by chromosome and as a function of distance. No depression or elevation of all six δ_c comparisons was seen by chromosome in an analysis of variance (results not shown). An autocorrelation analysis of markers spaced at $\leq$50 cM showed no evidence of closely spaced markers having similar δ_c values in either admixed population in variograms. A representative comparison for African American versus European American differences in δ_c of marker pairs $\leq$10 cM apart is shown in figure 3. The lack of upward trend in the kernel smoothing line, which is flat in both populations out to 50 cM (not shown), indicates that the δ_c values of closely spaced marker pairs are no more similar than those of distantly spaced ones. The sample autocorrelation functions estimated with intrapair distances categorized into 1-cM-wide bins also displayed no evidence of positive autocorrelation in either population (analysis not shown).

The distribution of allelic differences conforms to our expectations, which are based upon the natural history of admixed Hispanics and African Americans (both including gene flow from Europeans) and nonadmixed Asian and European groups (fig. 4). Thus, the greatest difference is seen in the comparison between Asians and African Americans (who share little recent admixture), whereas the smallest differences occur between Hispanics and European Americans. For populations where MALD analysis would be feasible, appreciable diver-

Table 1

STR Markers Examined, Map Locations, δ_c, and MALD Map Status of Markers for European American versus African American and European American versus Hispanic Comparisons

Chromosome and Locus	Map Location (cM)	δ_c for European American versus	
		African American	Hispanic
Chromosome 1:			
D1S468[a]	4.2	.340	.330
D1S1612	16.2	.184	.193
D1S244[a]	20.6	.331	.269
D1S1597	29.9	.189	.048
D1S228	29.9	.209	.231
D1S3669	37.1	.281	.135
D1S199	45.3	.286	.248
D1S552	45.3	.283	.040
D1S1622[b]	56.7	.508	.164
D1S255[c]	65.5	.275	.310
D1S2130	72.6	.132	.095
D1S2134[c]	75.7	.249	.267
D1S197	76.3	.257	.143
D1S220	87.3	.254	.156
D1S1669	89.8	.222	.165
D1S209[a]	93.9	.381	.290
D1S1665	102.0	.277	.153
D1S216[a]	104.8	.384	.350
D1S1728[b]	109.0	.346	.158
D1S207	113.7	.293	.238
D1S551	113.7	.089	.034
D1S1588	125.5	.093	.097
D1S206[b]	134.2	.388	.236
D1S1631	136.9	.116	.136
D1S502[a]	146.5	.423	.395
D1S1675	149.2	.180	.102
D1S252[c]	150.3	.282	.310
D1S534[c]	151.9	.260	.285
D1S498[a]	155.9	.500	.283
D1S1653	164.1	.069	.129
D1S484	169.7	.296	.198
D1S1679	170.8	.166	.087
D1S1677	175.6	.123	.081
D1S2628[a]	177.9	.640	.311
D1S196[b]	181.5	.370	.119
D1S218[b]	191.5	.531	.225
D1S1589[b]	192.1	.344	.134
D1S518	202.2	.271	.178
D1S238[b]	202.7	.302	.224
D1S1660	212.4	.213	.182
D1S413[c]	212.4	.201	.258
D1S1678[b]	218.5	.355	.118
D1S249[c]	220.7	.286	.367
IL10-D	222.1	.150	.086
IL10-O	222.1	.229	.139
D1S1663	226.2	.139	.059
D1S229	237.7	.236	.214
D1S549	239.7	.206	.187
D1S213[a]	242.3	.476	.315
D1S1656[b]	245.1	.318	.199
D1S3462	247.2	.164	.114
D1S547	267.5	.186	.116
D1S1609	274.5	.086	.114
D1S423[b]	277.8	.302	.105
Chromosome 2:			
D2S319[c]	7.6	.257	.386
D2S1780[b]	11.2	.375	.151
D2S281	14.1	.278	.179
D2S162[b]	20.0	.519	.167
D2S423	22.1	.222	.191
D2S1400[b]	27.6	.459	.086
D2S1360	38.3	.296	.185
D2S165[a]	47.4	.351	.264
D2S405	48.0	.169	.108
D2S1788[b]	55.5	.377	.176
D2S2230	56.2	.298	.167
D2S1356[c]	64.3	.201	.323
D2S391	70.3	.266	.203
D2S1352	73.6	.252	.164
D2S406	80.2	.161	.111
D2S290[a]	84.4	.359	.255
D2S1394	90.8	.098	.099
D2S1777	99.4	.290	.093
D2S139[a]	101.6	.431	.283
D2S1790	104.8	.270	.080
D2S2181[a]	110.0	.538	.262
IL1RA-O	115.6	.234	.060
D2S160	123.0	.272	.182
IL1A	123.0	.282	.186
D2S121[a]	123.5	.526	.337
D2S347[a]	131.5	.405	.251
D2S1328	132.6	.200	.169
D2S114[c]	142.8	.279	.350
D2S442	147.4	.154	.179
D2S1399	152.0	.187	.150
D2S142	161.3	.228	.212
D2S1353	164.5	.220	.207
D2S1776[c]	173.0	.147	.262
D2S326[a]	177.5	.548	.350
D2S1391	186.2	.181	.155
D2S2273[a]	186.2	.366	.292
D2S117[a]	194.5	.593	.267
D2S1384	200.4	.138	.177
D2S157[b]	206.1	.498	.172
D2S2944	210.4	.163	.104
D2S164[a]	214.7	.442	.250
D2S434	215.8	.107	.076
D2S2197[c]	222.2	.208	.257
D2S1363	227.0	.217	.146
D2S401[a]	229.1	.429	.280
D2S396[c]	232.9	.262	.364
D2S427[b]	236.7	.460	.097
D2S206[b]	240.8	.436	.156
D2S338[a]	250.5	.458	.286
D2S125	260.6	.231	.206
Chromosome 3:			
D3S1270	7.0	.271	.194
D3S1297[a]	8.3	.387	.344
IL5RA	12.3	.184	.129
D3S1560[c]	19.0	.287	.295

(*continued*)

(*continued*)

Table 1 (*Continued*)

Chromosome and Locus	Map Location (cM)	δ_c for European American versus	
		African American	Hispanic
D3S1304	22.3	.168	.163
D3S1259[a]	36.7	.321	.373
D3S2403[b]	37.2	.418	.140
D3S1293[a]	44.8	.381	.371
D3S3038	44.8	.112	.103
D3S1266[c]	52.6	.189	.304
D3S1211[a]	57.9	.320	.441
D3S2432	57.9	.170	.120
D3S1768	61.5	.202	.086
D3S1298[b]	62.1	.308	
D3S2354[b]	69.2	.430	.033
AFMb362wb9[b]	69.5	.333	.071
GAAT12D11	69.5	.182	.020
D3S2409	70.6	.210	.218
D3S3616[a]	76.5	.381	.293
D3S1766	78.6	.103	.073
D3S1300[b]	80.3	.373	.200
D3S1285[b]	91.2	.323	.118
D3S3544	96.7	.247	.174
D3S1284[b]	102.6	.486	.207
D3S2406	102.6	.103	.185
D3S3671	113.0	.124	.117
D3S2459	119.1	.128	.087
D3S1278[a]	129.7	.352	.397
D3S2460	134.6	.223	.153
D3S1267[a]	139.1	.417	.340
D3S3657[a]	148.2	.305	.767
D3S1238	149.3	.117	.219
D3S1764	152.6	.281	.144
D3S3546	154.5	.225	.193
D3S1744	161.0	.137	.136
D3S196[b]	161.0	.460	.176
D3S1763	176.5	.134	.111
D3S1282[b]	180.8	.340	.243
D3S3053	181.9	.105	.091
D3S3715[b]	190.4	.301	.183
D3S1232[a]	191.8	.435	.341
D3S1262[a]	201.1	.301	.289
D3S2398	209.4	.098	.051
D3S1294[a]	210.1	.423	.275
D3S2418	215.8	.264	.131
D3S1311	224.9	.252	.157
Chromosome 4:			
D4S412[a]	4.7	.348	.280
D4S2366[a]	12.9	.321	.270
D4S2949[a]	23.2	.350	.895
D4S403[c]	25.9	.276	.250
D4S419[b]	33.4	.323	.218
D4S2639	34.6	.274	.108
D4S2397	42.7	.241	.058
D4S2912[b]	47.6	.364	.212
D4S2632	54.6	.224	.146
D4S405[a]	57.0	.308	.318
D4S1627	60.2	.186	.101
D4S428[b]	64.2	.506	.190
D4S3248	72.5	.058	.171
D4S398[a]	72.5	.341	.345

(*continued*)

Table 1 (*Continued*)

Chromosome and Locus	Map Location (cM)	δ_c for European American versus	
		African American	Hispanic
D4S2367	78.4	.173	.153
D4S3018[a]	78.4	.352	.277
GC[b]	79.7	.447	.111
D4S3003[b]	87.1	.366	.247
D4S3243[b]	89.2	.323	.064
D4S1534[b]	95.1	.343	.220
D4S1647	104.9	.128	.080
D4S2623[b]	114.0	.312	.161
D4S2940	117.1	.145	.232
IL2	125.2	.250	.203
D4S2394	127.0	.281	.136
D4S1579	140.6	.210	.185
D4S1644	143.3	.162	.222
D4S1565	143.8	.268	.100
D4S424[c]	144.6	.275	.327
D4S1625	146.0	.098	.112
D4S1629[b]	158.0	.403	.137
D4S413[a]	158.0	.538	.324
D4S1566[c]	166.9	.204	.364
D4S2368	167.6	.081	.075
D4S1597[b]	169.4	.444	.206
D4S2431	176.2	.205	.147
D4S415	181.4	.255	.141
D4S2417	181.9	.104	.100
D4S1535[b]	195.1	.334	.193
D4S408	195.1	.289	.191
Chromosome 5:			
D5S2488	.0	.207	.077
D5S1492	9.4	.098	.068
D5S406[a]	11.9	.495	.432
D5S2505	14.3	.175	.146
D5S807	19.0	.242	.137
D5S817	22.9	.116	.116
D5S416[b]	28.8	.568	.168
D5S814	39.5	.105	.086
D5S419[a]	40.0	.348	.258
D5S1470	45.3	.222	.167
D5S426[a]	52.0	.355	.277
D5S418[a]	58.6	.311	.333
D5S1457	59.3	.224	.133
D5S407[a]	64.7	.384	.333
D5S2500	69.2	.207	.110
D5S647[b]	74.1	.405	.238
D5S1501	85.3	.244	.232
D5S1716	95.3	.099	.106
D5S428	95.4	.258	.190
D5S644[c]	104.8	.270	.396
D5S669[b]	112.5	.315	.222
D5S2501	117.0	.069	.103
D5S421[a]	122.0	.639	.357
D5S1505	129.8	.093	.129
D5S471[b]	129.8	.519	.238
D5S2059[a]	133.7	.365	.342
D5S816	139.3	.109	.134
IL9	139.3	.216	.126
D5S393[a]	140.7	.352	.286
D5S1480	147.5	.148	.150

(*continued*)

Table 1 (*Continued*)

Chromosome and Locus	Map Location (cM)	δ_c for European American versus	
		African American	Hispanic
D5S210[b]	147.5	.440	.145
D5S436[b]	147.5	.351	.236
D5S410[b]	156.5	.593	.214
D5S820	159.8	.207	.231
D5S1955	163.3	.295	.160
D5S2050[b]	171.1	.300	.169
D5S1471	172.1	.119	.104
D5S1456	174.8	.187	.112
D5S462	178.6	.146	.089
Chromosome 6:			
D6S1713[a]	7.0	.301	.281
SE30[a]	9.2	.354	.260
D6S309[c]	14.1	.259	.283
D6S470[b]	18.2	.311	.190
D6S443[c]	25.1	.284	.297
D6S1006	26.7	.272	.112
D6S259[b]	27.8	.422	.175
D6S260	29.9	.297	.239
D6S1588[b]	38.2	.392	.180
D6S1281	44.4	.077	.069
D6S276[b]	44.4	.606	.218
TNFB	46.4	.035	.063
D6S1019	53.8	.281	.061
D6S1610[a]	53.8	.337	.915
D6S426[a]	60.4	.417	.370
D6S1017	63.3	.298	.083
D6S459[c]	69.7	.225	.324
D6S1280	73.1	.261	.154
D6S427	73.1	.039	.089
D6S1960	76.6	.083	.119
D6S257	79.9	.282	.233
D6S1031	88.6	.289	.155
D6S286[c]	89.8	.233	.311
D6S1270	92.6	.206	.058
D6S1570[a]	99.0	.412	.907
D6S1043[c]	100.9	.267	.255
D6S434[b]	109.2	.449	.167
D6S1021	112.2	.236	.089
D6S474	118.6	.062	.090
D6S261[c]	120.3	.291	.377
D6S1040[b]	128.9	.323	.137
D6S262[b]	130.0	.341	.190
D6S976[b]	135.5	.323	.231
D6S1009	137.7	.157	.190
D6S1003[b]	144.5	.379	.192
D6S308	144.5	.225	.150
D6S441[b]	154.1	.327	.214
D6S2436[b]	154.6	.362	.117
D6S305[b]	166.4	.373	.152
D6S1277	173.3	.160	.041
D6S264[b]	179.1	.349	.139
D6S503	184.5	.074	.119
D6S1027[b]	187.2	.325	.204
D6S446[c]	189.0	.186	.441
D6S281	190.1	.221	.202
TBP[b]	190.5	.487	.081

(*continued*)

Table 1 (*Continued*)

Chromosome and Locus	Map Location (cM)	δ_c for European American versus	
		African American	Hispanic
Chromosome 7:			
D7S2477[c]	.0	.288	.558
D7S531[a]	5.3	.373	.314
D7S517[a]	7.4	.308	.268
D7S2201	10.7	.137	.096
D7S2547	17.2	.196	.213
D7S513	17.7	.296	.180
D7S507[a]	28.7	.404	.270
D7S493[a]	34.7	.484	.313
D7S1802	35.3	.236	.194
D7S629[a]	37.5	.537	.262
D7S1808[c]	41.7	.210	.277
D7S2416	41.7	.285	.119
D7S526[b]	49.2	.314	.222
D7S817	50.3	.119	.169
D7S484[b]	53.5	.346	.238
D7S2846	57.8	.028	.129
D7S2469[a]	61.5	.335	.589
D7S519[b]	69.0	.376	.234
D7S1818	69.6	.060	.154
D7S1830	72.8	.129	.072
D7S2429[c]	76.7	.282	.253
D7S669[a]	90.4	.315	.268
D7S2212	95.4	.056	.060
D7S2485	98.4	.239	.182
D7S820	98.4	.161	.192
D7S657[b]	104.9	.767	.223
D7S821	109.1	.247	.172
D7S662[a]	111.8	.370	.282
D7S1799	113.9	.164	.147
D7S692	121.4	.273	.147
D7S2847[b]	125.2	.338	.146
D7S650[a]	126.8	.489	.449
D7S530[c]	134.6	.287	.289
D7S640[a]	137.8	.595	.452
D7S684[c]	147.2	.254	.256
D7S1824	149.9	.182	.161
D7S2195[b]	150.4	.312	.156
D7S661[a]	155.1	.324	.310
TCRB-6.1	155.6	.105	.099
TCRB-6.4	155.6	.296	.119
TCRB-6.7	155.6	.197	.138
TCRB-E[b]	155.6	.423	.170
TCRB-F[b]	155.6	.433	.119
D7S1805	161.2	.119	.130
D7S505	161.2	.231	.204
D7S1826	162.3	.071	.096
D7S3058	173.7	.122	.056
D7S550[b]	178.4	.493	.240
D7S2423[a]	182.0	.339	.890
D7S559	182.0	.207	.176
Chromosome 8:			
D8S504	.0	.228	.115
D8S262[b]	4.3	.353	.118
D8S277[a]	8.3	.337	.333
D8S550[b]	21.3	.351	.232

(*continued*)

Table 1 (*Continued*)

Chromosome and Locus	Map Location (cM)	δ_c for European American versus	
		African American	Hispanic
D8S1130	22.4	.166	.113
D8S1106[c]	26.4	.153	.336
D8S1827[a]	30.5	.399	.792
D8S258[b]	41.6	.356	.115
D8S136[a]	44.0	.419	.275
D8S1739	48.8	.219	.168
D8S1477	60.3	.155	.240
D8S283[a]	60.9	.350	.300
D8S505[c]	60.9	.181	.306
D8S1110	67.3	.185	.139
D8S285[a]	71.0	.346	.348
D8S1113	77.9	.202	.112
D8S260[c]	79.4	.249	.308
D8S1136	82.3	.066	.147
D8S1775[b]	87.5	.339	.132
D8S279[a]	91.5	.372	.387
D8S1697[a]	98.9	.471	.300
D8S1119[b]	101.0	.376	.150
GAAT1A4	110.2	.142	.080
D8S257	111.7	.153	.135
D8S1784[b]	118.2	.398	.215
D8S1132	119.2	.254	.143
D8S592	125.3	.205	.085
D8S514[a]	130.0	.439	.291
D8S508	137.9	.241	.153
D8S1128	139.5	.291	.202
D8S284[b]	143.8	.454	.198
D8S1100	154.0	.271	.131
D8S272[a]	154.0	.500	.257
D8S1741[b]	162.9	.472	.210
D8S373[b]	164.5	.336	.084
Chromosome 9:			
D9S1858	.0	.252	.113
D9S288[c]	9.8	.193	.290
D9S2169	14.2	.282	.085
D9S286	18.1	.279	.230
D9S269	24.1	.202	.218
D9S156[a]	30.6	.344	.405
IFNA	33.3	.115	.063
D9S1870[a]	37.6	.369	.283
D9S171[a]	42.7	.341	.305
D9S161[b]	51.8	.474	.238
D9S741	52.7	.284	.175
D9S319	54.5	.161	.068
D9S273	65.8	.258	.186
D9S301[c]	66.3	.163	.311
D9S175[a]	70.3	.602	.400
D9S1122	75.9	.061	.148
D9S922	80.3	.178	.144
D9S167[a]	83.4	.302	.300
D9S257[b]	91.9	.317	.243
D9S1781	99.4	.253	.235
D9S910	104.5	.223	.151
D9S176[c]	105.0	.239	.254
D9S938[c]	110.9	.229	.257
D9S1675[a]	120.0	.320	.280
D9S930	120.0	.094	.138
D9S154[a]	125.6	.502	.355
D9S934	128.0	.080	.102
D9S266[a]	136.5	.451	.562
D9S164	147.9	.196	.165
D9S1826[c]	159.6	.246	.303
Chromosome 10:			
D10S249[a]	2.1	.397	.341
D10S602	4.3	.209	.154
D10S1435	6.2	.115	.105
D10S1713	13.5	.232	.178
D10S189	19.0	.286	.121
D10S1412	28.3	.184	.215
D10S547[b]	29.2	.323	.157
D10S2325	32.8	.245	.180
D10S1423	46.5	.104	.112
D10S1662	48.4	.178	.167
D10S197[b]	52.1	.381	.167
D10S1426	59.0	.100	.083
D10S208[b]	60.6	.323	.220
D10S1220	70.2	.162	.066
D10S1225	80.8	.256	.110
D10S1652[b]	80.8	.342	.235
D10S1670[a]	86.2	.425	.365
D10S1432	93.9	.156	.085
D10S1699	97.3	.182	.107
D10S2327[b]	100.9	.398	.065
D10S1786[a]	103.4	.482	.636
D10S1739	110.0	.113	.200
D10S583[a]	115.3	.402	.255
D10S677	117.4	.229	.142
D10S192[b]	124.3	.364	.224
D10S1239	125.9	.146	.081
D10S1682[b]	130.9	.352	.126
D10S1237	134.7	.292	.101
D10S1230[b]	142.8	.352	.168
D10S587	147.6	.177	.158
D10S1213	148.2	.243	.229
D10S1223	152.9	.287	.185
D10S1703[a]	155.7	.421	.310
D10S1651[c]	168.8	.242	.279
D10S212	170.9	.105	.030
D10S555[c]	170.9	.189	.272
D10S169	173.1	.275	.108
Chromosome 11:			
D11S1984	2.1	.189	.168
D11S2362	8.9	.231	.129
D11S1999	17.2	.128	.176
D11S1981	21.5	.188	.116
D11S902[a]	21.5	.409	.290
D11S915[a]	30.9	.318	.405
D11S904[b]	33.6	.319	.189
D11S1776	40.1	.237	.126
D11S935[a]	45.9	.564	.261
D11S905[b]	52.0	.494	.232
D11S1313[a]	58.4	.376	.286
D11S1985	58.4	.219	.190
D11S4155	67.5	.291	.139

(*continued*)

Table 1 (*Continued*)

Chromosome and Locus	Map Location (cM)	δ_c for European American versus	
		African American	Hispanic
D11S987	67.5	.249	.241
D11S2371	76.1	.172	.103
D11S4207[b]	76.1	.301	
D11S937[a]	80.0	.511	.380
D11S1396	85.5	.254	.088
D11S2002	85.5	.205	.110
D11S4197[a]	87.9	.313	.583
D11S4134	96.9	.250	.185
D11S2000[b]	100.6	.357	.199
D11S1893[a]	105.2	.421	.317
D11S1986[c]	105.7	.238	.250
D11S1998	113.1	.171	.205
D11S925[a]	118.5	.453	.306
D11S4464	123.0	.054	.128
D11S934[c]	126.2	.268	.286
D11S1351[a]	131.3	.307	.396
D11S912	131.3	.297	.221
D11S968[c]	147.8	.105	.310
Chromosome 12:			
D12S352[a]	.0	.479	.449
D12S94[c]	1.2	.239	.257
D12S372	6.4	.092	.069
D12S1626[c]	7.1	.243	.316
D12S1673	12.6	.167	.138
D12S99[c]	12.6	.271	.275
CD4[b]	16.4	.390	.163
D12S358[b]	26.2	.313	.218
D12S391	26.2	.257	.185
D12S364[a]	30.6	.364	.370
D12S373	36.1	.052	.068
D12S1042	48.7	.280	.163
D12S1640	48.7	.189	.157
D12S1663	56.4	.220	.141
D12S85[b]	61.3	.332	.109
D12S1618	68.2	.199	.222
D12S398	68.2	.181	.197
D12S83[a]	75.2	.448	.356
D12S1294	76.1	.224	.177
D12S375	80.5	.133	.066
D12S1052	83.2	.056	.097
D12S92[b]	83.2	.315	.190
D12S1064	95.0	.123	.236
D12S95[b]	96.1	.331	.164
D12S1657[b]	102.0	.340	.107
D12S1300	105.0	.127	.120
PAH	109.5	.187	.169
D12S78[a]	111.9	.320	.318
D12S2070	125.3	.178	.232
D12S79[a]	125.3	.465	.305
D12S366[a]	133.3	.446	.265
D12S395	136.8	.111	.108
D12S342[a]	144.8	.387	.315
D12S2078	148.0	.161	.112
D12S1679	153.2	.162	
D12S1045[b]	160.7	.356	.123
D12S97[b]	160.7	.375	.240
D12S1638	168.8	.234	.082

(*continued*)

Table 1 (*Continued*)

Chromosome and Locus	Map Location (cM)	δ_c for European American versus	
		African American	Hispanic
Chromosome 13:			
D13S175[c]	6.0	.209	.278
D13S787	8.9	.159	.172
D13S221[c]	12.9	.279	.285
D13S1254[c]	14.5	.283	.300
D13S260[a]	23.7	.517	.369
D13S1493	25.8	.137	.224
D13S219[b]	28.9	.477	.179
D13S894	33.5	.102	.119
D13S263[a]	38.3	.360	.311
D13S153[a]	45.6	.354	.355
D13S788[b]	45.6	.345	.231
D13S1309[c]	50.5	.295	.860
D13S800	55.3	.081	.099
D13S162	58.5	.284	.211
D13S170[b]	63.9	.371	.245
D13S317	63.9	.215	.186
D13S265[a]	68.7	.397	.256
D13S793	74.9	.271	.046
D13S154[a]	75.2	.308	.277
D13S779	82.9	.212	.224
D13S158[b]	84.9	.402	.190
D13S173	93.5	.119	.136
D13S796[b]	93.5	.455	.110
D13S1315[a]	102.7	.323	.267
D13S285[b]	110.6	.307	.165
Chromosome 14:			
D14S72[b]	9.4	.386	.175
D14S742	12.5	.078	.037
D14S283	13.9	.296	.226
D14S990[c]	14.6	.263	.292
D14S1041	23.2	.268	.141
D14S1280	25.9	.171	.091
D14S80[c]	26.6	.276	.311
D14S597	28.0	.240	.148
D14S297	31.8	.191	.111
D14S49	36.8	.222	.223
D14S1049[b]	40.9	.316	.197
D14S306	44.1	.050	.123
D14S288[a]	47.5	.369	.266
D14S587	55.8	.259	.202
D14S274[a]	63.3	.514	.251
D14S592[b]	66.8	.301	.135
D14S63[a]	69.2	.374	.301
D14S588[b]	75.6	.345	.106
D14S258[a]	76.3	.394	.339
D14S1036[b]	84.7	.318	.228
D14S53	86.3	.200	.104
D14S74[b]	87.4	.302	.143
D14S606	91.6	.138	.115
D14S610	95.9	.110	.082
D14S68[a]	95.9	.533	.473
D14S1044[b]	99.9	.487	.227
D14S617	105.5	.266	.093
D14S749	108.2	.101	.173
D14S81[a]	108.2	.408	.265
D14S51[b]	115.6	.359	.235

(*continued*)

Table 1 (*Continued*)

Chromosome and Locus	Map Location (cM)	δ_c for European American versus	
		African American	Hispanic
D14S611	115.9	.255	.155
D14S78	125.9	.217	.108
D14S260[b]	134.3	.448	.141
D14S1007[a]	138.2	.307	.317
Chromosome 15:			
D15S128[b]	6.1	.381	.232
D15S1002[a]	14.6	.494	.275
D15S165[b]	20.2	.327	.114
ACTC	31.5	.204	.130
D15S659	43.5	.187	.129
D15S126	45.6	.176	.239
D15S978[a]	45.6	.328	.376
D15S117[a]	51.2	.352	.263
D15S643	52.3	.216	.162
D15S1036[b]	57.4	.459	.206
D15S153[b]	62.4	.438	.178
D15S131[a]	71.3	.311	.438
D15S973[b]	73.5	.344	.134
D15S205[c]	78.9	.291	.333
D15S152	80.0	.229	.161
D15S127[b]	86.8	.399	.173
D15S652	90.0	.137	.185
D15S130[b]	100.6	.384	.244
D15S816	100.6	.120	.095
D15S657	104.9	.157	.076
D15S120[a]	112.6	.487	.297
Chromosome 16:			
D16S3024	7.1	.258	.227
D16S2622[b]	8.2	.445	.201
D16S423[a]	10.4	.401	.343
D16S748[b]	22.7	.339	.218
D16S3075[c]	23.3	.276	.266
D16S2619	28.3	.058	.058
D16S405	28.3	.289	.201
D16S3017[a]	32.1	.375	.252
D16S3046[b]	40.7	.316	.201
D16S403	43.9	.204	.202
D16S420[a]	44.5	.356	.491
D16S401[a]	46.9	.521	.282
D16S769	50.6	.036	.070
D16S753	57.8	.203	.078
D16S409[b]	58.5	.326	.249
D16S771	70.7	.141	.145
D16S3253[b]	71.8	.337	.171
D16S503[a]	83.6	.368	.259
D16S2624	87.6	.156	.093
D16S515[a]	92.1	.460	.297
D16S518[b]	95.1	.412	.236
D16S511[a]	110.4	.417	.330
D16S422[a]	111.1	.327	.352
GATA86C08	120.6	.185	.168
D16S3023	132.6	.174	.191
Chromosome 17:			
D17S1308	.6	.283	.081
D17S849[a]	.6	.334	.264
D17S1298	10.7	.258	.065
D17S796[a]	14.7	.470	.293

(*continued*)

Table 1 (*Continued*)

Chromosome and Locus	Map Location (cM)	δ_c for European American versus	
		African American	Hispanic
D17S938[a]	14.7	.513	.346
D17S1852[a]	22.2	.380	.334
D17S974	22.2	.236	.131
D17S1303	23.6	.159	.154
D17S969	27.8	.164	.110
D17S799[a]	32.0	.445	.364
D17S921[b]	36.1	.377	.077
D17S122	41.1	.190	.148
D17S959[b]	48.1	.362	.219
D17S1294	50.7	.219	.162
D17S798[b]	53.4	.329	.193
D17S791[a]	64.2	.354	.419
D17S809	74.5	.253	.092
D17S787[a]	75.0	.414	.291
D17S1290	82.0	.244	.204
D17S924[c]	82.0	.187	.328
D17S789[b]	89.3	.333	.208
D17S2059	93.3	.101	.073
D17S1301	100.0	.202	.042
D17S802[a]	106.8	.342	.411
D17S1822[a]	116.9	.426	.300
D17S784[b]	116.9	.322	.148
D17S928	126.5	.241	.073
Chromosome 18:			
D18S59	.0	.255	.123
D18S481[b]	6.9	.455	.208
D18S976[a]	12.8	.396	.270
D18S843	28.1	.114	.129
D18S464[a]	31.2	.390	.301
D18S877	54.4	.246	.110
D18S1135[b]	61.7	.487	.242
D18S57[a]	62.8	.313	.282
D18S535	64.5	.136	.123
D18S474[b]	71.3	.313	.211
D18S851	73.8	.193	.110
D18S69[b]	77.4	.303	.179
D18S858	80.4	.154	.122
D18S64[b]	84.8	.388	.084
D18S68[a]	96.5	.402	.329
GATA175B10[c]	96.5	.264	.270
D18S61[a]	105.0	.423	.287
ATA82B02	106.8	.251	.119
D18S1161[a]	114.3	.313	.338
D18S844	116.4	.149	.134
Chromosome 19:			
D19S591	9.8	.217	.112
D19S216	20.0	.270	.080
D19S413[b]	32.4	.374	.150
D19S586	32.9	.077	.113
D19S221[b]	36.2	.341	.218
ERBAL2	37.8	.211	.176
D19S226[a]	42.3	.443	.355
D19S714	43.1	.101	.091
D19S1037	47.7	.093	.110
D19S433	51.9	.280	.117
D19S220[b]	62.0	.403	.204
D19S198	65.8	.228	.088

(*continued*)

Table 1 (*Continued*)

Chromosome and Locus	Map Location (cM)	δ_c for European American versus	
		African American	Hispanic
D19S412[b]	70.1	.421	.174
D19S246	78.1	.167	.154
D19S589	87.7	.111	.067
D19S887[b]	100.0	.314	.131
D19S254	100.6	.247	.113
Chromosome 20:			
D20S117[b]	2.8	.490	.197
D20S473	9.5	.257	.142
D20S116[b]	11.2	.371	.174
D20S115[b]	21.2	.376	.080
D20S189[b]	30.6	.336	.154
D20S604	32.9	.148	.111
D20S112[b]	39.3	.339	.229
D20S470	39.3	.240	.178
D20S477	50.1	.161	.104
D20S107[c]	55.7	.191	.307
D20S119[b]	61.8	.487	.105
D20S481	62.3	.180	.141
D20S196[a]	75.0	.576	.341
D20S120[a]	83.5	.403	.252
D20S171	95.7	.278	.117
Chromosome 21:			
D21S1432	3.0	.125	.087
D21S1414[c]	9.7	.270	.376
D21S1437[b]	13.1	.356	.134
D21S1918[b]	16.2	.319	.237
D21S214[a]	16.9	.436	.286
D21S1270[b]	27.4	.405	.148
D21S1440[b]	36.8	.354	.168
D21S167[b]	38.7	.391	
D21S156[b]	42.6	.338	.188
D21S266	45.9	.289	.132
D21S171[c]	53.9	.224	.286
D21S1446[b]	57.8	.334	.144
Chromosome 22:			
D22S420[b]	4.1	.315	.104
D22S1174	19.3	.186	.114
D22S264[b]	21.1	.485	.152
D22S315[a]	21.5	.453	.282
D22S1176	29.7	.211	.137
D22S280[c]	31.3	.233	.250
D22S685	32.4	.205	.178
D22S283[a]	38.6	.425	.333
IL-2RB[b]	42.8	.319	.130
IL2RBA[b]	42.8	.304	.156
D22S445[b]	45.8	.358	.096
D22S294[c]	51.4	.267	.361
D22S274[a]	51.5	.410	.327
Chromosome X:			
DXS987[c]	22.0	.286	.400
DXS1202[a]	38.4	.607	.469
DXS1214[b]	45.0	.378	.225
DXS1068[b]	52.6	.313	
DXS993[a]	62.5	.303	.395
PFC[b]	68.3	.529	.108
DXS1055	72.4	.192	.179
DXS990[a]	99.7	.613	.357
DXS1106[b]	111.8	.362	.087
DXS1001[a]	130.4	.371	.300
DXS1047[a]	143.2	.402	.414
DXS1227	155.9	.260	.200

NOTE.—Map positions were estimated from the Marshfield map, with some loci included by interpolation from radiation hybrid data. Primer sequences and additional data on all six δ_c and LLAR comparisons, along with I*(2) for African Americans and Hispanics, are available at the Laboratory of Genomic Diversity Web site.

[a] Markers which have δ_c values that meet the two criteria above for African Americans and Hispanics.

[b] African American MALD markers with $\delta_c \geq .30$ when compared to European Americans.

[c] Hispanic MALD markers with $\delta_c \geq .25$ when compared to European Americans.

gence is apparent. In the comparison of African Americans versus Europeans, 44% of STR loci show $\delta_c > .3$, and 74% of loci show $\delta_c > .2$. For the Hispanic-European comparison, 17% of loci have $\delta_c > .3$, and 45% have $\delta_c > .2$. These differences are critical, insofar as the size of δ and δ_c are the principal determinants of linkage-disequilibrium detection in admixed populations (Chakraborty and Weiss 1988; Chakraborty et al. 1991; Stephens et al. 1994, 1999). The operative δ_c for Hispanics and African Americans is almost certainly underestimated here, since our comparison utilized admixed populations and not the actual parent population—native Africans, in the case of African Americans. To illustrate this underestimation, consider the comparison of African Americans versus Asians (fig. 4*A*, *B*), which shows the greatest δ_c, since these populations do not share any recent gene flow. This comparison shows 80% of STR loci with $\delta_c > .3$ and 95% of the loci with $\delta_c > .2$. These values are a plausible surrogate estimator of similar mean distances between native African and European population structure. However, it is not expected that the same loci with high δ_c in the Asian versus African American comparison would be the same as those with high δ_c in other comparisons. This discordance is illustrated in figure 2*C*, where the correlation between STR δ_c values in comparisons of different ethnic groups is low ($r^2 = .25$), considering that both comparisons are with the same European American reference group.

Discussion

The development of allele frequency data for MALD mapping is critical to the advancement of the method-

ology for gene mapping studies. The theoretical basis of MALD mapping is now well established (Chakraborty and Weiss 1988; Chakraborty et al. 1991; Briscoe et al. 1994; Stephens et al. 1994; McKeigue 1997, 1998; Stephens et al. 1999; Zheng and Elston 1999). Empirical studies have also found MALD over large distances of as much as 30 cM around the FY gene in African Americans, and strong linkage disequilibrium was found with STRs in an 8-cM core around the FY gene (Parra et al. 1998; Hamblin and Di Rienzo 2000; Lautenberger et al. 2000; Wilson and Goldstein 2000). There is ample evidence that ongoing and differential levels of admixture across populations must be taken into account in any disease gene identification efforts (Parra et al. 2001; Pfaff et al. 2001). Others have attempted to identify markers appropriate for MALD (Dean et al. 1994; Collins et al. 2000), but the present study represents the largest to date. Taken together, these results suggest that the ~10-cM map of markers presented here makes a good foundation for MALD-based gene mapping in the African American and Hispanic populations.

The present study examines 744 markers, to identify those that are best able to differentiate between founding populations; such markers would be appropriate for MALD analysis in Hispanics or African Americans. Only weak correlations were found between δ_c, LLAR, or I*(2) in the European American versus African American and the European American versus Hispanic comparisons (fig. 4*C* and analyses not shown), so that the two groups of markers for MALD are nearly randomly overlapping. Those markers ($n = 315$) with a δ_c of $\geqslant .30$ have an average spacing of 11 cM in African Americans, and those with $\delta_c \geqslant .25$ ($n = 214$ markers) in Hispanics have an average spacing of 16 cM; these two groups share 153 markers in common (indicated in table 1). There is some concern that these STR-based markers will be supplanted by SNP; however, several factors work to the advantage of STRs. They are relatively easy to assay via direct PCR amplification and separation on commercial sequencers. In MALD-TDT applications, the diversity of alleles seen at STRs will make TDT trios more generally informative than biallelic SNP markers (Spielman et al. 1993; McKeigue 1997, 1998). Those multiallelic advantages of STRs could be counterbalanced by multiallelic haplotypes based on SNPs. However, STR technology is in hand and works quite well, whereas SNP genotyping technology is currently in a state of flux (Kristensen et al. 2001).

We have examined genomewide marker frequency data to explore the possibility of autocorrelation of marker δ_c values in African-Americans and Hispanics. This analysis was undertaken because the existence of positive autocorrelation could influence both historical inferences and the search for genetic regions that contribute to ethnic differences in phenotype distribution. Positive autocorrelation between closely spaced pairs of markers would have occurred if nearby markers tended to have similar δ_c values, yet neighbors are as similar as randomly selected loci in δ_c differences (fig. 3).

Biologically speaking, appropriate MALD markers depend on the disease model. In the case of African Americans, at least 30 diseases with a likely hereditary component have a higher prevalence in this minority group than in European Americans (Williams 1999). Thus, although searching for a European disease allele in African Americans has, theoretically, the most power, the empirical approach is to search for an African one. Markers most appropriate for this case have alleles with high frequencies in African Americans that are absent in European Americans.

Acknowledgments

We thank Drs. J. Coresh, M. Dean, G. Huttley, and G. Nelson, for their helpful discussions. We are grateful to G. Washburn for assistance in designing multiplex STR primer sets. We thank Dr. Stephen Rich and the Collaborative Study of Asthma Genetics for sharing their allele frequency data. Some computations used resources of the Advanced Biomedical Computing Center (Frederick, MD). The content of this publication does not necessarily reflect the views or policies of the Department of Health and Human Services, nor does mention of trade names, commercial products, or organizations imply endorsement by the U.S. government. This project has been funded in whole or in part with federal funds from the National Cancer Institute, National Institutes of Health, under contract NO1-CO-56000.

Electronic-Database Information

The URL for data in this article is as follows:

Laboratory of Genomic Diversity Web site, http://lgd.nci.nih.gov (for additional allele frequency data for each locus, a full set of difference statistics between the groups, and primer sequences)

References

Altshuler D, Pollara VJ, Cowles CR, Van Etten WJ, Baldwin J, Linton L, Lander ES (2000) An SNP map of the human genome generated by reduced representation shotgun sequencing. Nature 407:513–516

Arngrimsson R, Sigurard ttir S, Frigge ML, Bjarnadttir RI, Jonsson T, Stefansson H, Baldursdottir A, Einarsdottir AS, Palsson B, Snorradottir S, Lachmeijer AM, Nicolae D, Kong A, Bragason BT, Gulcher JR, Geirsson RT, Stefansson K (1999) A genome-wide scan reveals a maternal susceptibility locus for pre-eclampsia on chromosome 2p13. Hum Mol Genet 8:1799–1805

Bellamy R, Beyers N, McAdam KP, Ruwende C, Gie R, Samaai P, Bester D, Meyer M, Corrah T, Collin M, Camidge DR, Wilkinson D, Hoal-Van Helden E, Whittle HC, Amos W,

van Helden P, Hill AV (2000) Genetic susceptibility to tuberculosis in Africans: a genome-wide scan. Proc Natl Acad Sci USA 97:8005–8009

Bodmer WF (1986) Human genetics: the molecular challenge. Cold Spring Harbor Symposium. Quant Genet 51:1–13

Bowcock AM, Ruiz-Linares A, Tomfohrde J, Minch E, Kidd JR, Cavalli-Sforza LL (1994) High resolution of human evolutionary trees with polymorphic microsatellites. Nature 368:455–457

Briscoe D, Stephens JC, O'Brien SJ (1994) Linkage disequilibrium in admixed populations: applications in gene mapping. J Hered 85:59–63

Brownstein MJ, Carpten JD, Smith JR (1996) Modulation of non-templated nucleotide addition by *Taq* DNA polymerase: primer modifications that facilitate genotyping. BioTechniques 20:1004-1006, 1008–1010

Chakraborty R, Kamboh MI, Ferrell RE (1991) 'Unique' alleles in admixed populations: a strategy for determining 'hereditary' population differences of disease frequencies. Ethn Dis 1:245–256

Chakraborty R, Kamboh MI, Nwankwo M, Ferrell RE (1992) Caucasian genes in American blacks: new data. Am J Hum Genet 50:145–155

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

Collaborative Study on the Genetics of Asthma (1997) A genome-wide search for asthma susceptibility loci in ethnically diverse populations. Nat Genet 15:389–392

Collins HE, Li H, Inda SE, Anderson J, Laiho K, Tuomilehto J, Seldin MF (2000) A simple and accurate method for determination of microsatellite total allele content differences between DNA pools. Hum Genet 106:218–226

Dean M, Stephens JC, Winkler C, Lomb DA, Ramsburg M, Boaze R, Stewart C, Charbonneau L, Goldman D, Albaugh BJ, Goedert JJ, Beasley RP, Hwang L-Y, Buchbinder S, Weedon M, Johnson PA, Eichelberger M, O'Brien SJ (1994) Polymorphic admixture typing in human ethnic populations. Am J Hum Genet 55:788–808

Destro-Bisol G, Maviglia R, Caglia A, Boschi I, Spedini G, Pascali V, Clark A, Tishkoff S (1999) Estimating European admixture in African Americans by using microsatellites and a microsatellite haplotype (CD4/Alu). Hum Genet 104:149–157

Diggle PJ, Liang K-Y, Zeger SL (1994) Analysis of longitudinal data. Oxford Statistical Science Series. Vol 13. Oxford University Press, Oxford

Hamblin MT, Di Rienzo A (2000) Detection of the signature of natural selection in humans: evidence from the Duffy blood group locus. Am J Hum Genet 66:1669–1679

Hanis CL, Hewett-Emmett D, Bertin TK, Schull WJ (1991) Origins of US Hispanics. Implications for diabetes. Diabetes Care 14:618–627

Huttley GA, Smith MW, Carrington M, O'Brien SJ (1999) A scan for linkage disequilibrium across the human genome. Genetics 152:1711–1722

Kaplan NL, Martin ER, Morris RW, Weir BS (1998) Marker selection for the transmission/disequilibrium test, in recently admixed populations. Am J Hum Genet 62:703–712

Kristensen VN, Kelefiotis D, Kristensen T, Borresen-Dale A-L (2001) High-throughput methods for detection of genetic variation. BioTechniques 30:318–332

Laan M, Pääbo S (1997) Demographic history and linkage disequilibrium in human populations. Nat Genet 17:435–438

Lautenberger JA, Stephens JC, O'Brien SJ, Smith MW (2000) Significant admixture linkage disequilibrium across 30 cM around the FY locus in African Americans. Am J Hum Genet 66:969–978

Long JC (1991) The genetic structure of admixed populations. Genetics 127:417–428

Long JC, Williams RC, McAuley JE, Medis R, Partel R, Tregellas WM, South SF, Rea AE, McCormick SB, Iwaniec U (1991) Genetic variation in Arizona Mexican Americans: estimation and interpretation of admixture proportions. Am J Phys Anthropol 84:141–157

Magnuson VL, Ally DS, Nylund SJ, Karanjawala ZE, Rayman JB, Knapp JI, Lowe AL, Ghosh S, Collins FS (1996) Substrate nucleotide-determined non-templated addition of adenine by *Taq* DNA polymerase: implications for PCR-based genotyping and cloning. BioTechniques 21:700–709

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

——— (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

Mein CA, Esposito L, Dunn MG, Johnson GC, Timms AE, Goy JV, Smith AN, Sebag-Montefiore L, Merriman ME, Wilson AJ, Pritchard LE, Cucca F, Barnett AH, Bain SC, Todd JA (1998) A search for type 1 diabetes susceptibility genes in families from the United Kingdom. Nat Genet 19: 297–300

O'Brien SJ (2000) Human genetic factors that impact HIV infection and progression. In: Phair JP, King E (eds) Medscape HIV/AIDS annual update 2000. Medscape, New York, pp 19–28

O'Brien SJ, Nelson GW, Winkler CA, Smith MW (2000) Polygenic and multifactorial disease gene association in man: lessons from AIDS. Annu Rev Genet 34:563–591

Parra EJ, Kittles RA, Argyropoulos G, Pfaff CL, Hiester K, Bonilla C, Sylvester N, Parrish-Gause D, Garvey WT, Jin L, McKeigue PM, Kamboh MI, Ferrell RE, Pollitzer WS, Shriver MD (2001) Ancestral proportions and admixture dynamics in geographically defined African Americans living in South Carolina. Am J Phys Anthropol 114:18–29

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

Pfaff CL, Parra EJ, Bonilla C, Hiester K, McKeigue PM, Kamboh MI, Hutchinson RG, Ferrell RE, Boerwinkle E, Shriver MD (2001) Population structure in admixed populations: effect of admixture dynamics on the pattern of linkage disequilibrium. Am J Hum Genet 68:198–207

Reich DE, Cargill M, Bolk S, Ireland J, Sabeti PC, Richter DJ, Lavery T, Kouyoumjian R, Farhadian SF, Ward R, Lander

ES (2001) Linkage disequilibrium in the human genome. Nature 411:199–204

Risch NJ (2000) Searching for genetic determinants in the new millennium. Nature 405:847–856

Risch N, Merikangas K (1996) The future of genetic studies of complex human diseases. Science 273:1516–1517

Shin HD, Winkler C, Stephens JC, Bream J, Young H, Goedert JJ, O'Brien TR, Vlahov D, Buchbinder S, Giorgi J, Rinaldo C, Donfield S, Willoughby A, O'Brien SJ, Smith MW (2000) Genetic restriction of HIV-1 infection and AIDS by promoter alleles of interleukin 10. Proc Natl Acad Sci USA 97:14467–14472

Shriver MD, Smith MW, Jin L, Marcini A, Akey JM, Deka R, Ferrell RE (1997) Ethnic-affiliation estimation by use of population-specific DNA markers. Am J Hum Genet 60:957–964

Smith JR, Freije D, Carpten JD, Gronberg H, Xu J, Isaacs SD, Brownstein MJ, Bova GS, Guo H, Bujnovszky P, Nusskern DR, Damber JE, Bergh A, Emanuelsson M, Kallioniemi OP, Walker-Daniels J, Bailey-Wilson JE, Beaty TH, Meyers DA, Walsh PC, Collins FS, Trent JM, Isaacs WB (1996) Major susceptibility locus for prostate cancer on chromosome 1 suggested by a genome-wide search. Science 274:1371–1374

Smith MW, Dean M, Carrington M, Winkler C, Huttley GA, Lomb DA, Goedert JJ, O'Brien TR, Jacobson LP, Kaslow R, Buchbinder S, Vittinghoff E, Vlahov D, Hoots K, Hilgartner MW, O'Brien SJ (1997) Contrasting genetic influence of CCR2 and CCR5 variants on HIV-1 infection and disease progression. Science 277:959–965

Spielman RS, McGinnis RE, Ewens WJ (1993) Transmission test for linkage disequilibrium: the insulin gene region and insulin-dependent diabetes mellitus (IDDM). Am J Hum Genet 52:506–516

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

Stephens JC, Smith MW, Shin HD, O'Brien SJ (1999) Tracking linkage disequilibrium in admixed populations with MALD using microsatellite loci. In: Goldstein DB, Schlötterer C (eds) Microsatellites: evolution and applications. Oxford University Press, Oxford, pp 211–224

Tourette Syndrome Association International Consortium for Genetics, The (1999) A complete genome screen in sib pairs affected by Gilles de la Tourette syndrome. Am J Hum Genet 65:1428–1436

Walder K, Hanson RL, Kobes S, Knowler WC, Ravussin E (2000) An autosomal genomic scan for loci linked to plasma leptin concentration in Pima Indians. Int J Obes Relat Metab Disord 24:559–565

Wiggs JL, Allingham RR, Hossain A, Kern J, Auguste J, DelBono EA, Broomer B, Graham FL, Hauser M, Pericak-Vance M, Haines JL (2000) Genome-wide scan for adult onset primary open angle glaucoma. Hum Mol Genet 9: 1109–1117

Williams DR (1999) Race, socioeconomic status, and health. The added effects of racism and discrimination. Ann NY Acad Sci 896:173–188

Wilson JF, Goldstein DB (2000) Consistent long-range linkage disequilibrium generated by admixture in a Bantu-Semitic hybrid population. Am J Hum Genet 67:926–935

Zheng C, Elston RC (1999) Multipoint linkage disequilibrium mapping with particular reference to the African-American population. Genet Epidemiol 17:79–101