nature genetics

Admixture mapping for hypertension loci with genome-scan markers

Xiaofeng Zhu[1], Amy Luke[1], Richard S Cooper[1], Tom Quertermous[2], Craig Hanis[3], Tom Mosley[4], C Charles Gu[5], Hua Tang[6], Dabeeru C Rao[5], Neil Risch[7,8] & Alan Weder[9]

Identification of genetic variants that contribute to risk of hypertension is challenging. As a complement to linkage and candidate gene association studies, we carried out admixture mapping using genome-scan microsatellite markers among the African American participants in the US National Heart, Lung, and Blood Institute's Family Blood Pressure Program. This population was assumed to have experienced recent admixture from ancestral groups originating in Africa and Europe. We used a set of unrelated individuals from Nigeria to represent the African ancestral population and used the European Americans in the Family Blood Pressure Program to provide estimates of allele frequencies for the European ancestors. We genotyped a common set of 269 microsatellite markers in the three groups at the same laboratory. The distribution of marker location–specific African ancestry, based on multipoint analysis, was shifted upward in hypertensive cases versus normotensive controls, consistent with linkage to genes conferring susceptibility. This shift was largely due to a small number of loci, including five adjacent markers on chromosome 6q and two on chromosome 21q. These results suggest that chromosome 6q24 and 21q21 may contain genes influencing risk of hypertension in African Americans.

Human hypertension is a complex biological trait that results from the joint influence of genetic and environmental factors. Important categories of environmental exposures have been identified, but the genetic architecture of this trait remains obscure. Genome-wide linkage analyses suggest that multiple chromosomal regions may have a role, although lack of consistency across studies undermines any generalizable conclusions[1–4]. Theoretical considerations have led several investigators to propose that the information generated by recent admixture of genetically distinct populations could be used to map disease-associated genes[5–8]. If, for example, a disease variant and a marker allele have substantially different frequencies across parental populations, strong linkage disequilibrium between the disease variant

and the marker may result and be preserved for several generations in the admixed group if the disease variant and the marker are sufficiently close. As proof of principle, the Duffy blood group allele was mapped based on these assumptions[5]. Because of its history of admixture, the African American population is useful in admixture mapping[9,10].

Abundant microsatellite markers that are informative for admixture mapping in African Americans[11,12] and a set of informative SNPs across the genome[13] have been identified. Statistical methods using admixed populations to map disease variants have also been proposed[7–17]. In this study, we carried out admixture mapping using an existing set of microsatellite markers designed for linkage
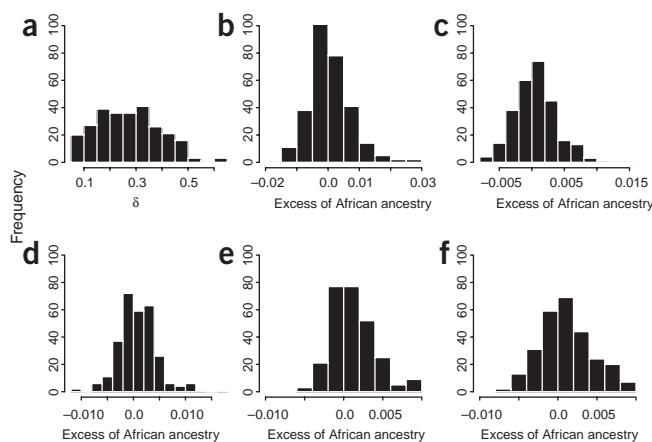

Figure 1 Distribution of δ values and excess African ancestry. (**a**) The distribution of microsatellite δ values between Nigerians and European Americans. (**b**) The excess of African ancestry in GenNet. (**c**) The excess of African ancestry in GENOA. (**d**) The excess of African ancestry in HyperGEN. (**e**) The excess of African ancestry in pooled cases. (**f**) The excess of African ancestry in pooled controls.

[1]Department of Preventive Medicine and Epidemiology, Loyola University Medical Center, 2160 S. First Ave., Maywood, Illinois 60153, USA. [2]Division of Cardiovascular Medicine, Stanford University School of Medicine, Stanford, California, USA. [3]Department of Genetics, University of Texas, Houston, Texas, USA. [4]Department of Medicine (Geriatrics), University of Mississippi Medical Center, Jackson, Mississippi, USA. [5]Division of Biostatistics, Washington University, St. Louis, Missouri, USA. [6]Division of Public Health Sciences, Fred Hutchinson Cancer Research Center, Seattle, Washington, USA. [7]Department of Genetics, Stanford University School of Medicine, Stanford, California 94305-5120, USA. [8]Division of Research, Kaiser Permanente, Oakland, California, USA. [9]Department of Medicine, University of Michigan Medical School, Ann Arbor, Michigan, USA. Correspondence should be addressed to N.R. (risch@lahmed.stanford.edu).

Table 1 Descriptive characteristics of the study subjects from each of three components of the FBPP

	Cases from individual networks			Pooled data	
	GenNet	GENOA	HyperGEN	Cases	Controls
Number (M/F)	193 (71/122)	553 (169/384)	594 (186/408)	737 (251/486)	573 (191/382)
Age	43.4 ± 11.2	55.9 ± 9.3	48.8 ± 10.9	53.1 ± 11.3	42.3 ± 13.4
BMI	33.5 ± 9.7	32.0 ± 7.0	33.0 ± 8.0	32.6 ± 7.6	28.7 ± 6.4
SBP	144.6 ± 19.4	136.4 ± 22.4	134.4 ± 23.1	147.5 ± 20.9	108.2 ± 7.6
DBP	88.7 ± 14.1	73.6 ± 11.7	77.1 ± 12.4	81.9 ± 11.6	64.3 ± 7.7
Number treated	107 (55%)	467 (84%)	520 (88%)	665 (90%)	0
African ancestry (%)	0.805	0.817	0.807	0.828	0.814

Data shown are mean ± s.d. F, female; M, male.

analysis in a large sample of African Americans who were part of a study on the genetics of hypertension.

The overall power of the marker density and δ distribution (**Fig. 1a**) that we used depends on the magnitude of genetic effects expected. Given the very large sample size in conjunction with reasonable marker density, this data set provided a useful opportunity to test the 'mapping by admixture linkage disequilibrium' approach on an important medical condition.

For the first, network-specific analyses, we randomly selected 193, 553 and 594 unrelated hypertensive African Americans from the available families in the GenNet, GENOA and HyperGEN substudies of the Family Blood Pressure Program (FBPP), respectively (**Table 1**). Blood pressures in GenNet were, on average, higher than in the other two networks, largely because a smaller proportion of hypertensive subjects from GenNet were taking antihypertensive medications (**Table 1**). Participants from GENOA were older than those from GenNet and HyperGEN; participants in all three networks had similar body mass indices (BMIs).

The estimated average African ancestry, defined as the mean of the African ancestry across individuals within networks, in the selected hypertensive individuals was 80.5%, 80.1% and 80.7% for GenNet, GENOA and HyperGEN, respectively, suggesting that there was little geographic effect. In the marker location–specific distributions of excess African ancestry in the selected hypertensive subjects for GenNet, GENOA and HyperGEN (**Fig. 1b–d**), several markers fell in the upper tail of the distribution. **Table 2** summarizes the markers with Z scores >2.5. Marker D4S3243 appeared in the results from both GenNet and HyperGEN, and D21S1437 appeared in both GenNet and HyperGEN, whereas GATA184A08 and D6S2436 on chromosome 6q occurred consistently in all three networks.

In the next analysis, we pooled the African American subjects from all three networks and included both cases (hypertensive, $n = 737$) and controls (normotensive, $n = 573$; **Table 1**). We estimated the African ancestry at each marker location in the cases and controls separately (**Fig. 1e,f**). **Figure 2a** shows the plot of genome-wide Z scores along with ancestry information content. The distribution of marker location–specific African ancestry was shifted upward in cases versus controls, as indicated by the distributions of Z scores (**Fig. 2b,c**). The cases have 11 marker locations with Z scores >2.5 whereas the controls have 4 marker locations with Z scores >2.5 (**Table 3**). The highest Z scores (>4.0) in cases occurred at the locations of markers GATA184A08 and D21S1437. Neither of these locations had high Z scores in controls. Five adjacent marker locations on chromosome 6q and two on chromosome 21q had Z scores >2.5. The distributions of Z scores were shifted upward in cases compared with controls (using the nearest corresponding cut-off values in the

controls; **Table 4**). The five markers on chromosome 6 and the two markers on chromosome 21 seemed to be the main reason for the shift in distribution.

Because our analysis samples consisted of one randomly selected hypertensive individual from each of the eligible families, and families often contained more than one such individual, a different selection process might give rise to different results. To explore this issue, we randomly resampled cases in the same way and calculated Z scores for each of the marker locations, repeating this process 500 times. Markers GATA184A08 and D21S1437 were associated with a Z score >2.5 in 497 and 500 of 500 times, respectively, suggesting that the excess of African ancestry seen at these two markers is robust with regard to the way the hypertensive subjects were sampled.

In the final analysis, we carried out logistic regression on African ancestry at each marker location in the same sample of cases and controls described above, adjusting for covariates. Sex, age and BMI were all significantly associated with hypertension ($P < 0.05$;

Table 2 Marker locations associated with the largest excess of African ancestry in hypertensive subjects for each individual network

| Network and marker	Location (cM)	Excess African ancestry	Z score
GenNet			
D2S2952	2p24 (18)	0.018	2.6
D2S427	2q36 (237)	0.018	2.6
D4S3243	4q21 (88)	0.028	4.06
D4S2361	4q22.1 (93)	0.026	3.89
GATA184A08	6q24.1 (146)	0.021	3.08
D6S2436	6q25.1 (155)	0.021	3.08
D21S1437	21q21 (13)	0.017	2.55
GENOA			
D1S235	1q42.13 (255)	0.014	4.39
D6S1021	6q21 (112)	0.008	2.55
GATA184A08	6q24.1 (146)	0.011	4.23
D6S2436	6q25.1 (155)	0.010	3.01
D8S1132	8q22 (158)	0.009	2.71
HyperGEN			
D3S4529	3q12 (112)	0.010	2.84
D4S3243	4q21 (88)	0.014	3.76
D6S1009	6q24.1 (138)	0.011	2.98
GATA184A08	6q24.1 (146)	0.017	4.69
D6S2436	6q25.1 (155)	0.011	2.91
D6S305	6q25.3 (166)	0.011	3.11
D11S2002	11q13–14 (85)	0.010	2.82
D21S1437	21q21 (13)	0.011	2.88
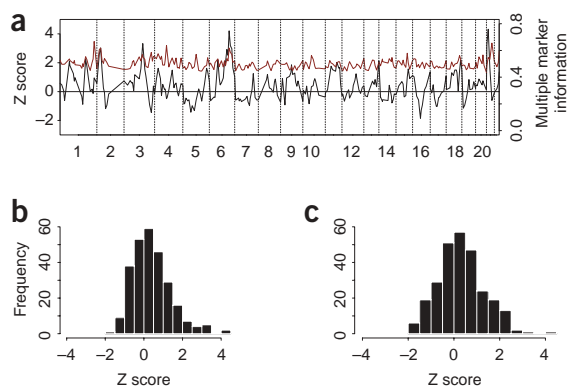
Figure 2 Genome-wide marker information content and distributions of Z scores. (**a**) Marker information content for admixture mapping (red line) and genome-wide Z score plot in pooled cases (black line). (**b**) The distribution of Z scores in pooled cases. (**c**) The distribution of Z scores in pooled controls.

Table 3 Marker locations with the largest excess of African ancestry among hypertensive cases and the corresponding excess in controls in the pooled data

| | | Cases | | Controls | |
| | | Excess of | | Excess of | |
Marker	Location (cM)	African ancestry	Z score	African ancestry	Z score
D2S2952	2p24 (18)	0.007	2.56	0.002	0.55
D2S1400	2p23 (28)	0.009	3.07	–0.003	–0.79
D3S2460	3q13.3 (135)	0.008	2.94	0.003	0.71
D3S4523	3q21 (138)	0.009	3.26	0.0	–0.01
D6S1040	6q23 (129)	0.008	2.92	0.002	0.51
D6S1009	6q24.1 (138)	0.007	2.51	0.004	1.24
GATA184A08	6q24.1 (146)	0.012	4.14	0.003	0.83
D6S2436	6q25.1 (155)	0.009	3.12	0.0	0.13
D6S305	6q25.3 (166)	0.008	2.96	–0.003	–0.81
D21S1432	21q21 (3)	0.009	3.29	–0.004	–0.98
D21S1437	21q21 (13)	0.012	4.34	–0.002	–0.60

Supplementary Table 1 online). Marker locations on chromosome 2, 3, 6 and 21 were again significantly associated with hypertension in this analysis ($P < 0.05$). We also found that 22% of the marker locations gave significant evidence of association between African ancestry and hypertension at the 5% level, suggesting that type 1 error is substantially inflated in this analysis. This probably reflects the fact that the overall individual admixture rate was also associated with hypertension, although this associated was of borderline significance ($P = 0.054$).

Finally, we examined the δ values for the markers (**Tables 2** and **3**). The average δ value for the set of markers that were chosen on the basis of high Z scores was 0.4, compared with 0.27 for all the markers, indicating that collectively, they were informative about ancestry.

Our results identify two genomic regions (6q and 21q) that contain genes that probably influence the risk of hypertension in this population. We observed high Z scores for the adjacent marker locations at GATA184A08 and D6S2436 (9 cM apart) in hypertensives, but not in normotensives, in each of the three networks independently and in the pooled analysis. Supporting evidence in the pooled analysis was also provided by five adjacent marker locations, two on either side of GATA184A08, spanning a total of 37 cM. Therefore, these results are probably not an artifact of characteristics of the single locus GATA184A08 but instead suggest that a trait locus in the region of 6q24 contains variants that influence blood pressure. Similarly, we also found evidence for linkage with marker D21S1437 and its adjacent marker D21S1432 in the pooled analysis. We also found suggestive but

weaker evidence for linkage with two adjacent markers on chromosome 2p25.1 and 3q13.31–33.

In support of our conclusion that 6q24 is relevant to hypertension risk, several previous studies reported evidence of linkage with hypertension and related traits in this region. Evidence for linkage to blood pressure–related traits has been reported at the same location in Mexican Americans and European Americans[18,19]. More limited evidence exists for populations of African descent, although the volume of published research is much smaller. In an initial linkage genome scan[1], we found a lod score of 1.44 at 193 cM from p-ter on chromosome 6 for systolic blood pressure (SBP) among Nigerian families. In the second phase of this genome-scan study, after doubling the sample size, the lod score increased to 3.19 at 173 cM from p-ter, but this locus is 27 cM away from the peak location in this study (data not shown). Whether our results replicate the previous findings is therefore uncertain. The region 21q21.1 could be linked to hypertension, and further replication studies are required to confirm or refute the current evidence. Both regions contain a large number of potential gene candidates, too numerous to describe here.

In terms of both sample size and number of markers genotyped, this study is the largest admixture mapping exercise to date. The fact that we carried out identical admixture analyses on hypertensives and normotensives from the same sample of families but obtained differing results in the two groups lends support to our conclusions. We examined the correlation of admixture at specific genomic locations. This approach is more robust than tests of 'global' or average individual admixture, in which residual confounding from environmental factors is difficult to eliminate[20]. Thus, our investigation supports theoretical arguments that admixture mapping could be useful in studying complex traits such as hypertension. Our observation of excess African ancestry at 6q24 and 21q21 among hypertensive African Americans provides impetus for further investigation of these regions for loci influencing blood pressure regulation.

Table 4 The distribution of right-tail Z scores in controls and percentages exceeding those Z scores in cases for the pooled sample

| Controls | | Cases | Cases without seven markers* |
Extreme tail (%)	Z score	Extreme tail (%)	Extreme tail (%)
4.83	1.96	7.06	4.58
3.72	1.98	6.69	4.20
2.67	2.15	5.20	2.67
1.53	2.29	4.09	1.53
1.15	2.88	3.35	1.15

*The seven deleted markers included the five most significant markers on chromosome 6q and two on chromosome 21q.

METHODS

Subjects. The FBPP is a large multicenter genetic study of high blood pressure and related conditions in multiple 'racial' or 'ethnic' groups, including European Americans, African Americans, Mexican Americans, and Asians and Asian Americans. It includes four component networks: GenNet, GENOA, HyperGEN and SAPPHIRe. GenNet, GENOA and HyperGEN independently

collected samples from European American and African American families. GenNet sampled African American and European American nuclear families in Maywood, Illinois, and Tecumseh, Michigan, respectively, through identification of a young, middle-aged proband with elevated blood pressure. GENOA sampled African American sibships containing sibling pairs with essential hypertension from Jackson, Mississippi, and European American sibships containing hypertensive probands from Rochester, Minnesota. GENOA also sampled Mexican American sibships containing sibling pairs with hypertension from Starr County, Texas. HyperGEN recruited African American and European American hypertensive siblings and random unrelated subjects from five field centers (African Americans from Birmingham, Alabama and Forsyth County, North Carolina; European Americans from Salt Lake City, Utah; Minneapolis, Minnesota; and Framingham, Massachusetts). SAPPHIRe recruited two groups of Japanese American and Chinese and Chinese American sibling pairs, one group of sibling pairs concordant for hypertension and one group of sibling pairs discordant for hypertension, from the San Francisco Bay area, Hawaii and Taiwan. Our analyses focused only on the African American participants. To characterize ancestral African populations, we used data from Nigerians who took part in the International Collaborative Study on Hypertension in Blacks, as described elsewhere[1]. The study was approved by the Institutional Review Boards of all local participating institutions, and informed consent was obtained from all study subjects.

Genotyping. DNA was extracted from whole blood by standard methods by each of the four networks and was sent to the US National Heart, Lung, and Blood Institute's Mammalian Genotyping Service in Marshfield, Wisconsin, for genotyping. Screening Set 8 (372 highly polymorphic microsatellite markers with an average map distance of 10 cM) was used for all four networks. The samples from the International Collaborative Study on Hypertension in Blacks were also submitted to the US National Heart, Lung, and Blood Institute's Mammalian Genotyping Service, but Screening Set 9 was used for these analyses. For all analyses, we included only those markers that were genotyped uniformly in all study subjects (*i.e.*, overlapping markers between Screening Set 8 and Screening Set 9). This reduced the set used for the current analysis to 269 markers.

Assessment of marker information. We assumed that Africa and Europe were the primary sources of ancestral populations for African Americans. To obtain allele frequency information for these ancestral populations, we randomly sampled a set of 236 unrelated individuals of Yoruban (Nigerian) ethnicity from the International Collaborative Study on Hypertension in Blacks and a set of 293 unrelated European American individuals from GenNet to represent African and European ancestors, respectively (allele frequencies for the European American subjects from GENOA and HyperGEN were similar to those from GenNet). To assess the marker information content for ancestry, we calculated the δ value of each marker, which is defined as

$$\delta = \sum_i |f_{i1} - f_{i2}|/2, \qquad (1)$$

where f_{i1} and f_{i2} represent the ith allele frequencies in the two ancestral populations, respectively. We calculated the allele frequencies of each marker using the gene counting method in the African American and European American samples separately and then used them to estimate the respective δ values.

The δ values provide information content for only a single marker at a time. To obtain a more refined estimate using multiple linked markers (multipoint), we ran STRUCTURE to obtain the vector $(p_0(x), p_1(x), p_2(x))$ of probabilities that an individual has zero, one or two alleles, respectively, of African origin at marker location x. These probabilities are derived using all adjacent markers and not just the marker at location x. We then calculated the variance $\sigma^2(x)$ of the distribution of $p(x)$ across the African American subjects. This variance will be low (near zero) when ancestry information is precise but will approach 0.5 with no information. Therefore, the formula $1 - 2\sigma^2(x)$ is used to represent the ancestry information content at location x. This formulation is similar to that used to define information content in linkage analysis[21].

The distributions of the δ values between European and African ancestral populations are presented in **Figure 1a** (mean 0.27, standard deviation 0.12). In

this set there were 44 markers (16.4%) with δ > 0.4, 111 (41.3%) markers with δ > 0.3 and 147 (54.6%) markers with δ > 0.25. The distributions of δ between the European American and the African American participants in GENOA and HyperGEN were similar (data not shown). The average marker spacing is 13 cM. This marker density provides information about admixture that has occurred on the order of ten generations ago, which is a reasonable estimate for the African American population. **Figure 2a** presents the ancestry information content across the genome. The average ancestry information content was 0.5 with standard error 0.04, suggesting that substantial information was obtained by simultaneous consideration of adjacent markers.

Analysis of individual networks: cases only. We obtained descriptive statistics using SAS (SAS Institute). For this analysis, we defined hypertensive status as current use of antihypertensive medications, SBP >140 mm Hg or diastolic blood pressure (DBP) >90 mm Hg. We then randomly selected hypertensives from the African American families, sampling one individual of age 20–60 y from each family. For each individual, we obtained a multipoint marker location–specific estimate of ancestry (African versus European) with the computer program STRUCTURE[22,23], which was run under the linkage model without haplotype phase information and using the Marshfield linkage map. We analyzed each network separately with the 236 Nigerians and 293 European Americans from GenNet. The average European admixture rate in this analysis was 17.2%, close to the previous estimates of 17.8% (ref. 22; using the same options in STRUCTURE) and 18.8% (ref. 24). The posterior mean of r was 0.085 with a 90% credible region of 0.074–0.095, corresponding to an average age of admixture of seven to ten generations ago. This result is consistent with the history of the African American population and previous estimates[22]. Each run of STRUCTURE consisted of 10,000 burn-in iterations followed by an additional 10,000 iterations, except for the simulations, for which we used 5,000 burn-in iterations followed by 10,000 additional iterations to save computation time. To determine whether we used sufficient iterations, for some analyses we also ran STRUCTURE with 50,000 burn-in iterations followed by an additional 100,000 iterations and obtained similar results. To determine the number of clusters K, we also examined the log-likelihoods for $K = 1, 2$ and 3, which were −1640092.4, −1613657.1 and −1614148.1, respectively, suggesting that the two-ancestral-population model was a reasonable assumption for African Americans, consistent with the literature[22,24].

We used the marker location–specific estimates of ancestry resulting from STRUCTURE to calculate excess African ancestry and corresponding Z scores. Excess African ancestry is defined as the difference between the estimated marker location–specific African ancestry and its expectation based on the overall genome-wide average. We ranked the marker location–specific ancestries according to their Z scores. The Z score can be used to test for linkage between a trait and a marker, provided that all the markers have full information for ancestry[17]. Because markers do not generally have full information content for ancestry, using the Z score test will result in an inflated type 1 error rate. Therefore, to assess statistical significance, we generated a null distribution from normotensive subjects from the same sample from which the hypertensives were drawn.

Analysis of all networks pooled: cases and controls. We next pooled the African American samples from the three FBPP networks. In this analysis, we used a stricter definition of hypertensive status: current use of antihypertensive medications and SBP or DBP in the upper 50% of the distribution among medicated individuals or, for individuals not using antihypertensive medications, SBP or DBP in the upper 5% of the distribution among nonmedicated individuals. Controls were defined as untreated individuals with SBP < 120 and DBP < 80. We then randomly selected one control of age 20–60 y from each of the families from which controls were available. We randomly selected cases from among the siblings in each of the remaining families. Thus, all subjects are independent and come from different families. The average age of controls was 11 years younger than that of cases. We estimated the African ancestry at each marker location for each of the cases and controls as described above using multipoint analysis with the linkage model in STRUCTURE. We calculated excess African ancestry and corresponding Z scores for cases and controls separately and ranked them. We used the distribution of excess African ancestry and Z scores in controls as a reference

for comparison with cases to evaluate statistical significance. We recorded marker locations with Z scores >2.5.

Next, we carried out logistic regression analysis of the cases versus controls to test the association of African ancestry with hypertension for each of the marker locations, adjusting for the covariates sex, age and BMI. We also carried out logistic regression on the overall individual African admixture rate estimated from all markers together.

Calculation of Z scores from the output of STRUCTURE. Assume there are N individuals and L markers genotyped across the genome. Let M_i be the estimated genome-wide African ancestry for individual i given by STRUCTURE. Using the linkage model and specifying SITEBYSITE $= 1$, STRUCTURE outputs the joint posterior assignment probabilities of population origin for the two alleles at each marker location. Therefore, marker location–specific African ancestry can be easily calculated and designated as q_{il} for the lth location for the ith individual. Define

$$\Delta\Pi_l = \frac{1}{N}\sum_{i=1}^{N}(q_{il} - M_i) \qquad (2)$$

as the excess African ancestry for marker location l. We then define the Z score as

$$Z_l = \frac{\Delta\Pi_l}{\text{std}(\Delta\Pi_l)} \qquad (3)$$

for the lth marker. According to ref. 17, Z_l, where $l = 1, 2, \ldots, L$, has an approximate standard normal distribution under the null hypothesis of no linkage, provided that the markers have full information content for ancestry and are spaced more than 5 cM apart. We used

$$\frac{1}{N}\sum_{i=1}^{N}M_i \qquad (4)$$

as the estimator of African ancestry for a marker location unlinked to a trait locus. It usually leads to overestimation when a trait locus with different ancestral allele frequencies is present, because M_i is estimated using all the marker information across the genome (including markers that are linked to the trait locus). But the overestimate is likely to be modest. Furthermore, because the markers used here do not generally have full information content for ancestry, using the Z score test will result in an inflated type 1 error rate. Thus, we show only those markers with Z scores >2.5 without giving formal P values. Instead, to assess statistical significance, we generated a null distribution from normotensive subjects taken from the same sample from which the hypertensives were drawn. This method should be more robust than assuming the Z scores follow a standard normal distribution under the null hypothesis.

Note: Supplementary information is available on the Nature Genetics website.

1. Cooper, R.S. *et al.* Genome scan among Nigerians linking blood pressure to chromosomes 2, 3 and 19. *Hypertension* **40**, 629–633 (2002).
2. Doris, P.A. Hypertension genetics, single nucleotide polymorphisms, and the common disease:common variant hypothesis. *Hypertension* **39**, 323–331 (2002).
3. Harrap, S.B. Where are all the blood pressure genes? *Lancet* **361**, 2149–2151 (2003).
4. Province, M.A. *et al.* A meta analysis of genome-wide linkage scans for hypertension: the National Heart, Lung and Blood Institute Family Blood Pressure Program. *Am. J. Hypertens.* **16**, 144–147 (2003).
5. Lautenberger, J.A., Stephens, J.C., O'Brien, S.J. & Smith, M.W. Significant admixture linkage disequilibrium across 30 cM around the FY locus in African Americans. *Am. J. Hum. Genet.* **66**, 969–978 (2000).
6. Pfaff, C.L. *et al.* Population structure in admixed populations: effect of admixture dynamics on the pattern of linkage disequilibrium. *Am. J. Hum. Genet.* **68**, 198–207 (2001).
7. Risch, N. Mapping genes for complex diseases using association studies in recently admixed populations. *Am J. Hum. Genet.* **51** (suppl),: 13 (1992).
8. Stephens, J.C., Briscoe, D., & O'Brien, S.J. Mapping by admixture linkage disequilibrium in human populations: limits and guidelines. *Am. J. Hum. Genet.* **55**, 809–824 (1994).
9. Fernandez, J.R. *et al.* Association of African genetic admixture with resting metabolic rate and obesity among women. *Obes. Res.* **11**, 904–911 (2003).
10. Shriver, M.D. *et al.* Skin pigmentation, biogeographical ancestry and admixture mapping. *Hum. Genet.* **112**, 387–399 (2003).
11. Collins-Schramm, H.E. *et al.* Ethnic-difference markers for use in mapping by admixture linkage disequilibrium. *Am. J. Hum. Genet.* **70**, 737–750 (2002).
12. Smith, M.W. *et al.* Markers for mapping by admixture linkage disequilibrium in African American and Hispanic populations. *Am. J. Hum. Genet.* **69**, 1080–1094 (2001).
13. Smith, M.W. *et al.* A high-density admixture map for disease gene discovery in African Americans. *Am. J. Hum. Genet.* **74**, 1001–1013 (2004).
14. Hoggart, C.J. *et al.* Design and analysis of admixture mapping studies. *Am. J. Hum. Genet.* **74**, 965–978 (2004).
15. McKeigue, P.M. 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 (1998).
16. Patterson, N. *et al.* Methods for high-density admixture mapping of disease genes. *Am. J. Hum. Genet.* **74**, 979–1000 (2004).
17. Zhu, X., Cooper, R.S. & Elston, R.C. Linkage analysis for complex disease using admixed populations. *Am. J. Hum. Genet.* **74**, 1136–1153 (2004).
18. Arya, R. *et al.* Factors of insulin resistance syndrome–related phenotypes are linked to genetics locations on chromosomes 6 and 7 in non-diabetic Mexican Americans. *Diabetes* **51**, 841–847 (2002).
19. Krushkal, J. *et al.* Genome-wide linkage analysis of systolic blood pressure using highly discordant siblings. *Circulation* **99**, 1407–1410 (1999).
20. Mountain, J.L. & Risch, N. Assessing genetic contributions to phenotypic differences among 'racial' and 'ethnic' groups. *Nat. Genet.* **36**, S48–S53 (2004).
21. Kruglyak, L. & Lander, E.S. Complete multipoint sib-pair analysis of qualitative and quantitative traits. *Am. J. Hum. Genet.* **57**, 439–454 (1995).
22. Falush, D., Stephens, M. & Pritchard, J.K. Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. *Genetics* **164**, 1567–1587 (2003).
23. Pritchard, J.K., Stephens, M., & Donnelly, P. Inference of population structure using multilocus genotype data. *Genetics* **155**, 945–959 (2000).
24. Parra, E.J. *et al.* Estimating African American admixture proportions by use of population-specific alleles. *Am. J. Hum. Genet.* **63**, 1839–1851 (1998).