Genome-Wide Linkage and Admixture Mapping of Type 2 Diabetes in African American Families From the American Diabetes Association GENNID (Genetics of NIDDM) Study Cohort

Steven C. Elbein,1,2 Swapan K. Das,1,2 D. Michael Hallman,3 Craig L. Hanis,3 and Sandra J. Hasstedt4

OBJECTIVE—We used a single nucleotide polymorphism (SNP) map in a large cohort of 580 African American families to identify regions linked to type 2 diabetes, age of type 2 diabetes diagnosis, and BMI.

RESEARCH DESIGN AND METHODS—After removing outliers and problematic samples, we conducted linkage analysis using 5,914 SNPs in 1,344 individuals from 530 families. Linkage analysis was conducted using variance components for type 2 diabetes, age of type 2 diabetes diagnosis, and BMI and nonparametric linkage analyses. Ordered subset analyses were conducted ranking on age of type 2 diabetes diagnosis, BMI, waist circumference, waist-to-hip ratio, and amount of European admixture. Admixture mapping was conducted using 4,486 markers not in linkage disequilibrium.

RESULTS—The strongest signal for type 2 diabetes (logarithm of odds [LOD] 4.53) was a broad peak on chromosome 2, with weaker linkage to age of type 2 diabetes diagnosis (LOD 1.82). Type 2 diabetes and age of type 2 diabetes diagnosis were linked to chromosome 13p (3–22 cm; LOD 2.42 and 2.46, respectively). Age of type 2 diabetes diagnosis was linked to 18p (66 cm; LOD 2.96). We replicated previous reports on chromosome 7p (79 cm; LOD 2.93). Ordered subset analysis did not overlap with linkage of unselcted families. The best admixture score was on chromosome 12 (90 cm; P = 0.0003).

CONCLUSIONS—The linkage regions on chromosomes 7 (27–78 cm) and 18p overlap prior reports, whereas regions on 2p and 13p linkage are novel. Among potential candidate genes implicated are TCF7L1, VAMP5, VAMP8, CDK8, INSIG2, IPP1, PAX8, ILIR1, MAP4K4, and 13p linkage are novel. Among potential candidate genes implicated are TCF7L2, PAX1, VAMP5, VAMP8, CDK8, INSIG2, IPP1, PAX8, ILIR1, MAP4K4, and 13p linkage are novel. Among potential candidate genes implicated are TCF7L2, IL18R1, and 13p linkage are novel. Among potential candidate genes implicated are TCF7L2, IL18R1, and 13p linkage are novel. Among potential candidate genes implicated are TCF7L2, IL18R1, and 13p linkage are novel. Among potential candidate genes implicated are TCF7L2, IL18R1, and 13p linkage are novel. Among potential candidate genes implicated are TCF7L2, IL18R1, and 13p linkage are novel. Among potential candidate genes implicated are TCF7L2, IL18R1, and

Type 2 diabetes is marked by a clear genetic propensity, a high concordance in identical twins, tendencies for both diabetes and age of onset to be familial (1), and marked differences in prevalence among ethnic groups (2). Despite considerable evidence for a genetic predisposition, unraveling the genetic etiology has been daunting, with few confirmed genes identified from genome-wide linkage scans. Recent successes with genome-wide association scans (3) have greatly increased the number of confirmed genetic loci, but these successes have been limited primarily to Caucasian populations. With few exceptions (TCF7L2), these regions from genome-wide association scans overlap neither with published linkage scans nor with known candidate genes or pathways, and in many cases, the variants are of unknown functional significance. These confirmed diabetes genes are uniformly of small effect size, and together, they explain only a small portion of diabetes risk even in Caucasians (4). Only variants at TCF7L2 show evidence for an association in African American populations (5), but the association of TCF7L2 with diabetes and reduced insulin secretion has not been easily reproducible (6). Thus, the genetic predisposition for type 2 diabetes in minority populations remains largely unknown. Particularly for African-based populations, low levels of linkage disequilibrium increase the challenge of mapping genes by genome-wide association and argue for the importance of complementary methods.

Previous genome-wide linkage scans for type 2 diabetes in African American or African families have been limited. A small African American sample of 229 individuals (124 sibpairs) was included in the original microsatellite linkage scan of GENNID (Genetics of NIDDM) Study families (7) and showed linkage on chromosome 10 near microsatellite marker D10S1412. Sale et al. (8) examined 392 microsatellite markers in 675 individuals (638 affected sibpairs) and reported suggestive linkage on chromosome 6 near D6S1035 (163.5 cm); other regions on 7p showed an association with earliest onset diabetes. Rotimi et al. (9) examined 392 microsatellite markers in 675 individuals (638 affected sibpairs) and reported suggestive linkage on chromosome 6 near D6S1035 (163.5 cm); other regions on 7p showed an association with earliest onset diabetes. Rotimi et al. (9) examined 392 microsatellite markers in 675 individuals (638 affected sibpairs) and reported suggestive linkage on chromosome 6 near D6S1035 (163.5 cm); other regions on 7p showed an association with earliest onset diabetes. Rotimi et al. (9) examined 392 microsatellite markers in 675 individuals (638 affected sibpairs) and reported suggestive linkage on chromosome 6 near D6S1035 (163.5 cm); other regions on 7p showed an association with earliest onset diabetes. Rotimi et al. (9) examined 392 microsatellite markers in 675 individuals (638 affected sibpairs) and reported suggestive linkage on chromosome 6 near D6S1035 (163.5 cm); other regions on 7p showed an association with earliest onset diabetes. Rotimi et al. (9) examined 392 microsatellite markers in 675 individuals (638 affected sibpairs) and reported suggestive linkage on chromosome 6 near D6S1035 (163.5 cm); other regions on 7p showed an association with earliest onset diabetes. Rotimi et al. (9) examined 392 microsatellite markers in 675 individuals (638 affected sibpairs) and reported suggestive linkage on chromosome 6 near D6S1035 (163.5 cm); other regions on 7p showed an association with earliest onset diabetes. Rotimi et al. (9) examined 392 microsatellite markers in 675 individuals (638 affected sibpairs) and reported suggestive linkage on chromosome 6 near D6S1035 (163.5 cm); other regions on 7p showed an association with earliest onset diabetes. Rotimi et al. (9) examined 392 microsatellite markers in 675 individuals (638 affected sibpairs) and reported suggestive linkage on chromosome 6 near D6S1035 (163.5 cm); other regions on 7p showed an association with earliest onset diabetes. Rotimi et al. (9) examined 392 microsatellite markers in 675 individuals (638 affected sibpairs) and reported suggestive linkage on chromosome 6 near D6S1035 (163.5 cm); other regions on 7p showed an association with earliest onset diabetes. Rotimi et al. (9) examined 392 microsatellite markers in 675 individuals (638 affected sibpairs) and reported suggestive linkage on chromosome 6 near D6S1035 (163.5 cm); other regions on 7p showed an association with earliest onset diabetes. Rotimi et al. (9) examined 392 microsatellite markers in 675 individuals (638 affected sibpairs) and reported suggestive linkage on chromosome 6 near D6S1035 (163.5 cm); other regions on 7p showed an association with earliest onset diabetes. Rotimi et al. (9) examined 392 microsatellite markers in 675 individuals (638 affected sibpairs) and reported suggestive linkage on chromosome 6 near D6S1035 (163.5 cm); other regions on 7p showed an association with earliest onset diabetes. Rotimi et al. (9) examined 392 microsatellite markers in 675 individuals (638 affected sibpairs) and reported suggestive linkage on chromosome 6 near D6S1035 (163.5 cm); other regions on 7p showed an association with earliest onset diabetes. Rotimi et al. (9) examined 392 microsatellite markers in 675 individuals (638 affected sibpairs) and reported suggestive linkage on chromosome 6 near D6S1035 (163.5 cm); other regions on 7p showed an association with earliest onset diabetes.

From 1993 to 2003, the American Diabetes Association sponsored the collection of families and sibpairs with at
RESULTS

We first examined three traits, type 2 diabetes, age of type 2 diabetes diagnosis, and BMI, using variance components linkage in 5,914 markers for 1,344 members of 530 families. Given the close association of BMI with type 2 diabetes, the likely role in modifying age of type 2 diabetes diagnosis, and BMI, we conducted linkage analysis using variance component linkage analysis using 5,914 markers and NPL analyses using 5,870 markers. For variance component analysis, we used a likelihood analysis implemented in JPAP (14). Sample size was 1,344 for type 2 diabetes and BMI and 1,082 for age of type 2 diabetes diagnosis.

Association of each marker was calculated using the regression models described above and comparing the model with the SNP to one without the SNP. To address this issue, SNPs were clustered using r thresholds of 0.25, 0.50, and 0.70, as implemented in MERLIN, assuming no recombination within clusters. Nonparametric linkage (NPL) analysis implemented in MERLIN (13) for type 2 diabetes. We report variance component linkage analysis using 5,914 markers and NPL analyses using 5,870 markers. For variance component analysis, we used a likelihood analysis implemented in JPAP (14). Sample size was 1,344 for type 2 diabetes and BMI and 1,082 for age of type 2 diabetes diagnosis.

Association of each marker was calculated using the regression models described above and comparing the model with the SNP to one without the SNP. To address this issue, SNPs were clustered using r thresholds of 0.25, 0.50, and 0.70, as implemented in MERLIN, assuming no recombination within clusters. Nonparametric linkage (NPL) analysis implemented in MERLIN (13) for type 2 diabetes. We report variance component linkage analysis using 5,914 markers and NPL analyses using 5,870 markers. For variance component analysis, we used a likelihood analysis implemented in JPAP (14). Sample size was 1,344 for type 2 diabetes and BMI and 1,082 for age of type 2 diabetes diagnosis.

Association of each marker was calculated using the regression models described above and comparing the model with the SNP to one without the SNP. To address this issue, SNPs were clustered using r thresholds of 0.25, 0.50, and 0.70, as implemented in MERLIN, assuming no recombination within clusters. Nonparametric linkage (NPL) analysis implemented in MERLIN (13) for type 2 diabetes. We report variance component linkage analysis using 5,914 markers and NPL analyses using 5,870 markers. For variance component analysis, we used a likelihood analysis implemented in JPAP (14). Sample size was 1,344 for type 2 diabetes and BMI and 1,082 for age of type 2 diabetes diagnosis. Detailed methods are provided in supplementary information. Below we modeled time-dependent diabetes risk to account for age of type 2 diabetes diagnosis in affected pedigree members using a modification of the age at diagnosis regressive model. The families were collected in three phases.

For variance component analysis, we used a likelihood analysis implemented in JPAP (14). Sample size was 1,344 for type 2 diabetes and BMI and 1,082 for age of type 2 diabetes diagnosis.

Association of each marker was calculated using the regression models described above and comparing the model with the SNP to one without the SNP. To address this issue, SNPs were clustered using r thresholds of 0.25, 0.50, and 0.70, as implemented in MERLIN, assuming no recombination within clusters. Nonparametric linkage (NPL) analysis implemented in MERLIN (13) for type 2 diabetes. We report variance component linkage analysis using 5,914 markers and NPL analyses using 5,870 markers. For variance component analysis, we used a likelihood analysis implemented in JPAP (14). Sample size was 1,344 for type 2 diabetes and BMI and 1,082 for age of type 2 diabetes diagnosis.

Association of each marker was calculated using the regression models described above and comparing the model with the SNP to one without the SNP. To address this issue, SNPs were clustered using r thresholds of 0.25, 0.50, and 0.70, as implemented in MERLIN, assuming no recombination within clusters. Nonparametric linkage (NPL) analysis implemented in MERLIN (13) for type 2 diabetes. We report variance component linkage analysis using 5,914 markers and NPL analyses using 5,870 markers. For variance component analysis, we used a likelihood analysis implemented in JPAP (14). Sample size was 1,344 for type 2 diabetes and BMI and 1,082 for age of type 2 diabetes diagnosis.

Association of each marker was calculated using the regression models described above and comparing the model with the SNP to one without the SNP. To address this issue, SNPs were clustered using r thresholds of 0.25, 0.50, and 0.70, as implemented in MERLIN, assuming no recombination within clusters. Nonparametric linkage (NPL) analysis implemented in MERLIN (13) for type 2 diabetes. We report variance component linkage analysis using 5,914 markers and NPL analyses using 5,870 markers. For variance component analysis, we used a likelihood analysis implemented in JPAP (14). Sample size was 1,344 for type 2 diabetes and BMI and 1,082 for age of type 2 diabetes diagnosis.

Association of each marker was calculated using the regression models described above and comparing the model with the SNP to one without the SNP. To address this issue, SNPs were clustered using r thresholds of 0.25, 0.50, and 0.70, as implemented in MERLIN, assuming no recombination within clusters. Nonparametric linkage (NPL) analysis implemented in MERLIN (13) for type 2 diabetes. We report variance component linkage analysis using 5,914 markers and NPL analyses using 5,870 markers. For variance component analysis, we used a likelihood analysis implemented in JPAP (14). Sample size was 1,344 for type 2 diabetes and BMI and 1,082 for age of type 2 diabetes diagnosis.

Association of each marker was calculated using the regression models described above and comparing the model with the SNP to one without the SNP. To address this issue, SNPs were clustered using r thresholds of 0.25, 0.50, and 0.70, as implemented in MERLIN, assuming no recombination within clusters. Nonparametric linkage (NPL) analysis implemented in MERLIN (13) for type 2 diabetes. We report variance component linkage analysis using 5,914 markers and NPL analyses using 5,870 markers. For variance component analysis, we used a likelihood analysis implemented in JPAP (14). Sample size was 1,344 for type 2 diabetes and BMI and 1,082 for age of type 2 diabetes diagnosis.

Association of each marker was calculated using the regression models described above and comparing the model with the SNP to one without the SNP. To address this issue, SNPs were clustered using r thresholds of 0.25, 0.50, and 0.70, as implemented in MERLIN, assuming no recombination within clusters. Nonparametric linkage (NPL) analysis implemented in MERLIN (13) for type 2 diabetes. We report variance component linkage analysis using 5,914 markers and NPL analyses using 5,870 markers. For variance component analysis, we used a likelihood analysis implemented in JPAP (14). Sample size was 1,344 for type 2 diabetes and BMI and 1,082 for age of type 2 diabetes diagnosis.

Association of each marker was calculated using the regression models described above and comparing the model with the SNP to one without the SNP. To address this issue, SNPs were clustered using r thresholds of 0.25, 0.50, and 0.70, as implemented in MERLIN, assuming no recombination within clusters. Nonparametric linkage (NPL) analysis implemented in MERLIN (13) for type 2 diabetes. We report variance component linkage analysis using 5,914 markers and NPL analyses using 5,870 markers. For variance component analysis, we used a likelihood analysis implemented in JPAP (14). Sample size was 1,344 for type 2 diabetes and BMI and 1,082 for age of type 2 diabetes diagnosis.
abesity” loci). The study population was drawn from 10 sites (Table 1). Characteristics of the study population are shown in Table 2, and linkage results are shown in Fig. 1. The LOD scores for all three traits are summarized in Table 3. Type 2 diabetes was most strongly linked to chromosome 2 (124–132 Mb), with maxima of 4.53 at 127 cM and 2.98 at 108 cM, and this region showed suggestive linkage to age of type 2 diabetes diagnosis (LOD 1.82 at 84 cM and 1.66 at 115 cM). Both type 2 diabetes and age of type 2 diabetes diagnosis showed suggestive linkage also to chromosome 13p (3–22 cM; LOD 2.42 for type 2 diabetes and 2.46 for age of type 2 diabetes diagnosis). The strongest linkage for age of type 2 diabetes diagnosis was on chromosome 18p (LOD 2.96; 54–96 cM). Type 2 diabetes also showed suggestive linkage to chromosome 7p (79 cM, 2.93), a region implicated previously (18). Other suggestive type 2 diabetes peaks were on chromosome 4 (135 cM; LOD 2.26), 11 (123 cM, 2.36), and 16 (56 cM, 2.43). The chromosome 18 linkage for age of type 2 diabetes diagnosis corresponded with previous reports in Caucasians (19, 20). Best linkages for BMI were on chromosomes 1 (LOD 2.30, 69 cM), 18 (116 cM, at 2.45; telomeric to type 2 diabetes and age of type 2 diabetes diagnosis), and 21 (14 cM) but did not overlap with type 2 diabetes or age of type 2 diabetes diagnosis linkages except on chromosome 1 (70–90 cM and 123–130 cM; Table 3; Fig. 1).

As expected, NPL analysis of type 2 diabetes gave overall smaller peaks and incomplete overlap with variance component scores. On chromosome 7 (27 cM; 1.73), type 2 diabetes linkage corresponded to a region identified in a preliminary report in African American Gullah-speak-

### Table 2

| Variable                          | Men (%) | Women (%) | Total |
|-----------------------------------|---------|-----------|-------|
| n                                 | 521 (34.8) | 975 (65.2) | 1,496 |
| Age (years)                       | 53.3 ± 13.1 | 53.8 ± 12.8 | 53.6 ± 12.9 |
| Diabetes                          | 406 (77.9) | 802 (82.3) | 1,208 (80.8) |
| Onset age of diabetes diagnosis (years) | 45.2 ± 11.7 | 44.8 ± 12.5 | 44.9 ± 12.3 |
| BMI (kg/m²)                       | 30.1 ± 6.2 | 34.4 ± 8.6 | 32.9 ± 8.1 |
| WHR                               | 0.96 ± 0.10 | 0.93 ± 0.13 | 0.94 ± 0.12 |
| Fasting glucose (mmol/l)          | 8.77 ± 4.53 | 8.59 ± 4.05 | 8.65 ± 4.23 |
| Fasting insulin (pmol/l)          | 136 ± 235 | 148 ± 224 | 144 ± 228 |

Data are n (%) or means ± SD.

![FIG. 1. Linkage plots for variance component and NPL methods. Fig. 1 shows all linkage analyses conducted in one figure. As noted in the text, numbers of individuals and numbers of markers varied across analyses. Variance component linkage for type 2 diabetes is black, for age of type 2 diabetes diagnosis is blue, and for BMI is green. LOD scores using MERLIN NPL scores are shown in red.]
ing families (18). Other regions on chromosome 13 (20 cM; LOD 1.63), 16 (51 cM; 1.84), and 22 (rs815550; 1.19) were near variance component peaks (Fig. 1; supplementary Table S2, available in the online appendix).

We performed ordered subset analysis on type 2 diabetes NPL scores by ranking families (both from low to high and high to low) by mean age of type 2 diabetes diagnosis, BMI, waist, WHR, diabetes duration, and admixture. Locations with scores that increased significantly are shown in Table 4, with the most prominent regions in a subset of at least 100 families on chromosomes 1 (30 cM; admixture), 3 (11 cM; age of type 2 diabetes diagnosis), 7 (131 cM; WHR), and 14 (57 cM; duration of diabetes). Ranking families by mean BMI produced no significant result. Only the region on chromosome 1 overlapped with a variance component linkage signal (Table 4).

Admixture mapping provides an alternative to family-based linkage to identify genes in recently admixed populations (21,22). Although our SNP-based linkage map was limited for admixture mapping by the relatively small proportion of ancestrally informative markers, the large number of markers nonetheless made mapping with the ADMIXMAP program feasible. We tested autosomal markers at 0.05-cM intervals to avoid confounding by linkage disequilibrium. Among the resulting 4,486 markers, the strongest association was for SNP rs1565728 (chromosome 12, 90 cM; \( P = 0.0003 \)). No other marker approached Bonferroni-corrected experiment-wide significance (supplemental data and supplementary Table S3, available in the online appendix), and no marker that achieved nominal uncorrected significance (\( P < 0.05 \)) mapped to a linkage peak. Individual admixture estimates varied from 0.010 to 0.614 but with mean estimates by study site ranging from 0.069 (South Carolina Gullah population) to 0.201 (Colorado) (supplemental data and supplementary Table S4, available in the online appendix). Heterogeneity of admixture proportions across sites was highly significant (\( \chi^2 = 154.76, P < 5 \times 10^{-11}; 10 \) df). Admixture

### Table 3
Summary of linkage from variance component analysis

| Chromosome | Location | Trait                  | Maximum LOD | SNP         |
|------------|----------|------------------------|-------------|-------------|
| 1          | 32       | Type 2 diabetes        | 2.27        | Rs3766306   |
|            | 69       | BMI                    | 2.30        | Rs1707302   |
|            | 114      | AOD                    | 1.98        | Rs1215512   |
|            | 129      | BMI                    | 1.77        | Rs338491    |
| 2          | 84       | Type 2 diabetes        | 2.28        | Rs1919481   |
|            | 108      | Type 2 diabetes        | 2.98        | Rs981525    |
|            | 115      | AOD                    | 1.66        | Rs1567804   |
|            | 127      | Type 2 diabetes        | 4.53        | Rs925848    |
| 3          | 27       | BMI                    | 1.92        | Rs2251166   |
|            | 150      | BMI                    | 1.54        | Rs1479137   |
| 4          | 135      | Type 2 diabetes        | 2.26        | Rs426029    |
| 6          | 118      | Type 2 diabetes        | 1.82        | Rs2032558   |
| 7          | 79       | Type 2 diabetes        | 2.93        | Rs956523    |
|            | 135      | BMI                    | 1.84        | Rs1862083   |
| 11         | 123      | Type 2 diabetes        | 2.36        | Rs1073636   |
| 12         | 27       | Type 2 diabetes        | 2.14        | Rs732868    |
| 13         | 4        | Type 2 diabetes        | 2.33        | Rs765244    |
|            | 16       | Type 2 diabetes        | 2.42        | Rs306395    |
| 14         | 35       | AOD                    | 1.77        | Rs30222     |
| 15         | 56       | Type 2 diabetes        | 2.43        | Rs11901     |
| 16         | 116      | BMI                    | 2.96        | Rs920783    |
| 17         | 22       | AOD                    | 2.45        | Rs1572881/rs2491222 |
| 18         | 35       | BMI                    | 2.14        | Rs6517799   |
| 19         | 14       | Type 2 diabetes        | 1.77        | Rs1406966   |
| 20         | 4        | Type 2 diabetes        | 1.77        | Rs1215512   |
| 21         | 12       | Type 2 diabetes        | 2.14        | Rs737622/rs715550 |

### Table 4
OSA of type 2 diabetes

| Chromosome | Covariate | Peak  | Maximum OSA score | Maximum unconditional score | Difference | \( P \) value | Families and ranking |
|------------|-----------|-------|-------------------|----------------------------|------------|--------------|---------------------|
| 1          | Admix.    | 30.00 | 4.11              | 2.33                       | 1.79       | 0.0353       | −117/414            |
| 3          | AOD       | 10.96 | 3.88              | 2.17                       | 1.71       | 0.0279       | +175/404            |
| 4          | AOD       | 109.31| 3.93              | 1.80                       | 2.13       | 0.0419       | −20/404             |
| 7          | WHR       | 131.47| 4.29              | 2.72                       | 1.57       | 0.0326       | +211/371            |
| 11         | AOD       | 104.00| 4.10              | 2.58                       | 1.53       | 0.0334       | +18/404             |
| 14         | Duration  | 56.59 | 4.05              | 2.64                       | 1.41       | 0.0207       | +223/351            |
| 17         | Waist     | 56.10 | 3.95              | 1.43                       | 2.52       | 0.0161       | −58/373             |
| 20         | WHR       | 2.28  | 3.75              | 2.48                       | 1.27       | 0.0372       | −33/371             |

Data are Z scores; \( P \) values are based on 500 permutations. Best signals from ordered subset analyses are shown. Admix, Caucasian admixture; duration, duration of diabetes; waist, waist circumference. \*Families ranked from lowest to highest mean covariate value, −; families ranked from highest to lowest mean value, +.
proportion significantly affected both type 2 diabetes risk ($P = 0.01$) and age of type 2 diabetes diagnosis ($P = 0.003$) but had no effect on BMI ($P = 0.95$).

**DISCUSSION**

Considerable progress has been made in the identification of type 2 diabetes susceptibility genes for individuals of European ancestry. Nonetheless, only a small portion of the genetic risk has been explained, and of the 18 reported susceptibility genes (4), only TCF7L2 shows any evidence for a role in African-derived populations (5). Thus, most of the susceptibility genes in high-risk ethnic groups remain to be identified. Genome-wide association studies in populations of African descent may identify a subset of these genes, but a much larger number of SNPs will be required to overcome lower levels of linkage disequilibrium, with accompanying difficulties in separating real signals from noise in populations that are readily available. Family-based linkage studies, particularly SNP-based methods with higher information content, thus remain an important alternative. Based on European data, linkage peaks are unlikely to represent single common variants of large effect sizes. Instead, these peaks likely represent multiple rare variants of large effect, copy number polymorphisms, or multiple common variants that together provide a linkage signal. Only the latter model is amenable to genome-wide association studies, but even for such variants, the additional information of linkage studies may be essential to select significant associations from the large number of nominally associated SNPs.

Previous studies of populations of African descent have been limited. As of this writing, no genome-wide association studies of African-based populations have been reported, and most European susceptibility loci do not appear to be important in African American populations (5). Among previous linkage studies, Sale et al. (8) used a microsatellite map to analyze 675 individuals in 247 families ascertained primarily for diabetic renal disease. The strongest signal from that study (chromosome 6q, 163 cM) is considerably telomeric to our strongest chromosome 6 peak (LOD 1.82 at 118 cM), whereas the second best signal from that study (chromosome 22, 32 cM, LOD 1.33) is in the same location as a minor peak in our current study (33 cM, LOD 1.68). Our strongest signals do not overlap with those found by Sale et al., however. A microsatellite scan in West African families (9) reported signals on chromosome 20 (LOD 1.8; 60–78 cM), which overlapped with a minor peak of similar size in our study at rs1406966 (LOD 1.77; 62 cM). A second minor peak on chromosome 4 from the West Africa study (LOD 1.37, 125 cM) also overlapped with a peak on chromosome 4 (135 cM; LOD 2.26) in our study. The African American families in the report of Ehm et al. (7) are entirely contained in the present multicenter study, and thus, any overlap in linkage signals would not be independent. More recently, Sale et al. (18) analyzed 471 individuals from the Gullah-speaking African American population of South Carolina for the same SNP markers as the present study. Chromosome 7 linkage to type 2 diabetes from that study overlaps with both NPL and variance component peaks near 27, 40, and 78 cM reported here (18) (M. Sale, personal communication). Because some of the GENNID samples included Gullah-speaking families from South Carolina, we repeated our chromosome 7 analysis excluding all samples from South Carolina. The maximum LOD score dropped from 2.93 to 2.03, consistent with a contribution to the total score from the South Carolina samples. Whereas Sale et al. (18) found chromosome 7 to be most strongly implicated in early-onset type 2 diabetes, ordered subset analyses in our sample with families ranked by age of type 2 diabetes diagnosis implicated no peak in this region.

In a recently published meta-analysis of 23 microsatellite linkage scans across ethnic groups that included both Chicago and Arkansas African American GENNID families (23), five regions of linkage overlapped with the current study: chromosomes 4 (135 cM), 6 (118 cM), 16 (56 cM), 20 (62 cM), and 22 (40 cM). Despite the partial overlap in samples, the only possible overlap in linkage signals between our study and the African American subgroup from the meta-analysis was on chromosome 16 (33–66 cM). Other African/African American peaks from the meta-analysis (chromosomes 4, 6, 10, and 14) were not observed in the present study.

Simulation studies and comparisons of microsatellite and SNP-based linkage scans have suggested that scans based on 10,000 SNPs provide superior information to the standard 10-cM microsatellite scans for diseases including prostate cancer (24), rheumatoid arthritis (25,26), and bipolar disorder (27). Each of the earlier studies used the 10,000-SNP Affymetrix platform, whereas we used the 5,914-SNP Illumina platform. We compared the information content between the Illumina SNP mapping panel and the earlier CIDR microsatellite panel (8) using Arkansas families that had been typed for both marker sets. The information content for all chromosomes was increased from 50 to 60% for the microsatellite scans to >80% for the SNP scan (supplemental data and supplementary Fig. S1). This increased information content would be expected to generate a larger number of significant or suggestive linkage peaks, narrower and better defined peaks, and perhaps could explain the lack of overlap with earlier, generally smaller linkage scans in African American families.

The typing of 5,914 SNPs at relatively even intervals allowed us to also map by admixture. In contrast to linkage, admixture mapping depends on the presence of extended haplotypes deriving from parental populations to map a disease such as type 2 diabetes that is more prevalent in the African than European populations. This approach has more power than linkage or genome scan approaches to identify genes that derive primarily or entirely from either the African or European populations when the markers also show large frequency differences between populations. Admixture mapping would not be a powerful method to identify susceptibility loci that are common and of similar frequency in both ethnic groups (22), and power would be reduced if marker frequencies are similar between parental populations. Admixture mapping has recently been used successfully in mapping hypertension susceptibility genes (28,29). The marker set in the current study was not specifically designed for admixture mapping, but the amount of data generated was sufficient to easily identify admixture. We found a strong negative correlation between European admixture and both type 2 diabetes risk and age of type 2 diabetes diagnosis. Furthermore, our data easily demonstrated significant heterogeneity in the amount of admixture across centers. Nonetheless, <10% of the markers showed a minor allele frequency difference between Yoruban and Caucasian HapMap samples exceeding 0.5. Thus, a specif-
ically designed admixture map might generate different and/or more significant signals.

Our strongest signal was at rs1565728 on chromosome 12 (90 cM). When we examined different combinations of markers in this region, we obtained $P$ values from 0.00003 to 0.03, suggesting that no single marker was responsible for the signal. The only gene in the immediate vicinity of this marker is the transcription factor, E2F7, which is involved in cell cycle regulation but is expressed at low levels in most tissues (30). However, admixture and linkage mapping peaks did not overlap, and no admixture signal achieved genome-wide significance (supplementary Table S2). Chromosome 1 marker rs1007460, the second strongest admixture signal (supplementary Table S2), was close to a well-replicated linkage signal on chromosome 1q21-q24 (31). These findings will require follow-up with markers specifically selected for large differences in allele frequency between ancestral populations.

We tested for family-based association with type 2 diabetes, BMI, and age of type 2 diabetes diagnosis, although recognizing the limitations of the relatively low-density SNP linkage panel. No SNP attained Bonferroni-corrected significance (supplementary Table S5, available in the online appendix); but rs741923 on chromosome 19 ($P = 0.0001$) was suggestive for association with type 2 diabetes; and SNPs rs1459085 (chromosome 5; $P < 0.0002$), rs729958 (chromosome 9, $P = 0.00005$), and rs1941487 (chromosome 18, $P = 0.0002$) were suggestive for association with diabetes age of type 2 diabetes diagnosis. SNPs on chromosomes 2 (rs838715, $P = 0.0001$) and 8 (rs1483457, $P = 0.0001$) showed suggestive associations with BMI. SNPs rs1459085 (chromosome 5) and rs729958 (chromosome 9) were associated with both type 2 diabetes and age of type 2 diabetes diagnosis. Although association and linkage signals generally did not overlap, rs1051783 (type 2 diabetes, $P < 0.002$) was within the chromosome 2q linkage peak and near gene 

In summary, using a large multicenter collection of nuclear families, a SNP-based linkage panel, and analytical methods based on variance component, admixture mapping, and ordered subsets, we have identified multiple regions of interest. Phenotypes of type 2 diabetes and diabetes age of diagnosis frequently mapped to the same locations, lending additional confidence to these results. Based on current understandings of complex disease pathogenesis, the majority of these regions probably represent collections of causative variants, rather than single SNPs or genes that can explain the linkage signal. Dense maps suitable for narrowing linkage signals, further exploring admixture mapping, and testing family-based association using maps that tag all variants in known genes will be logical steps. With most genetic risk factors yet to be discovered, these studies will undoubtedly complement other genome-wide association efforts in identifying those genes that contribute to diabetes in African Americans and other populations.

ACKNOWLEDGMENTS

This work was supported by National Institutes of Health (NIH)/National Institute of Diabetes and Digestive and Kidney Diseases Grant DK-071100. CIDR is fully funded through NIH Contract HHSN268200782096C to The Johns Hopkins University.

No potential conflicts of interest relevant to this article were reported.

We thank the GENNID investigators who recruited and characterized the families, and we thank the American Diabetes Association for access to the samples used in this study. We thank Kristen Wheeler and Winston Chu for valuable assistance with DNA preparation for genotyping. Genotyping services were provided by the CIDR. Data from this study are available from the American Diabetes Association.

REFERENCES

1. Das SK, Elbein SC: The genetic basis of type 2 diabetes. Cellscience 2:100–131, 2006
2. Elbein SC: Evaluation of polymorphisms known to contribute to risk for diabetes in African and African-American populations. Curr Opin Clin Nutr Metab Care 10:415–419, 2007
3. Frayling TM: Genome-wide association studies provide new insights into type 2 diabetes aetiology. Nat Rev Genet 8:657–662, 2007
4. Lango H, Palmer CN, Morris AD, Zeggini E, Hattersley AT, McCarthy MI, Frayling TM, Weedon MN: Assessing the combined impact of 18 common genes that contribute to diabetes in African Americans and other populations.
