Genome-Wide Linkage Scan in Gullah-Speaking African American Families With Type 2 Diabetes

The Sea Islands Genetic African American Registry (Project SuGAR)

Michèle M. Sale,[1,2,3] Lingyi Lu,[4] Ida J. Spruill,[5] Jyotika K. Fernandes,[5] Kerry H. Lok,[6] Jasmin Divers,[4] Carl D. Langefeld,[4] and W. Timothy Garvey[6,7]

OBJECTIVE—The Gullah-speaking African American population from the Sea Islands of South Carolina is characterized by a low degree of European admixture and high rates of type 2 diabetes and diabetic complications. Affected relative pairs with type 2 diabetes were recruited through the Sea Islands Genetic African American Registry (Project SuGAR).

RESEARCH DESIGN AND METHODS—We conducted a genome-wide linkage scan, genotyping 5,974 single nucleotide polymorphisms in 471 affected subjects and 50 unaffected relatives from 197 pedigrees. Data were analyzed using a multipoint engine for rapid likelihood inference and ordered subsets analyses (OSAs) for age at type 2 diabetes diagnosis, waist circumference, waist-to-hip ratio, and BMI. We searched for heterogeneity and interactions using a conditional logistic regression likelihood approach.

RESULTS—Linkage peaks on chromosome 14 at 123–124 cM were detected for type 2 diabetes (logarithm of odds [LOD] 2.10) and for the subset with later age at type 2 diabetes diagnosis (maximum LOD 4.05). Two linkage peaks on chromosome 7 were detected at 44–45 cM for type 2 diabetes (LOD 1.18) and at 78 cM for type 2 diabetes (LOD 1.64) and the subset with earlier age at type 2 diabetes diagnosis (maximum LOD 3.93). The chromosome 14 locus and a peak on 7p at 29.5 cM were identified as important in the multilocus model. Other regions that provided modest evidence for linkage included chromosome 1 at 167.5 cM (LOD 1.51) and chromosome 3 at 121.0 cM (LOD 1.61).

CONCLUSIONS—This study revealed a novel type 2 diabetes locus in an African American population on 14q that appears to reduce age of disease onset and confirmed two loci on chromosome 7. ***Diabetes*** **58:260–267, 2009**

From the [1]Center for Public Health Genomics, University of Virginia, Charlottesville, Virginia; the [2]Department of Medicine, University of Virginia, Charlottesville, Virginia; the [3]Department of Biochemistry and Molecular Genetics, University of Virginia, Charlottesville, Virginia; the [4]Department of Biostatistical Sciences, Wake Forest University School of Medicine, Winston-Salem, North Carolina; the [5]Department of Medicine, Medical University of South Carolina, Charleston, South Carolina; the [6]Department of Nutrition Sciences, University of Alabama, Birmingham, Alabama; and the [7]Birmingham VA Medical Center, Birmingham, Alabama.

Corresponding author: Michèle M. Sale, msale@virginia.edu.

Received 12 February 2008 and accepted 17 September 2008.

Published ahead of print at http://diabetes.diabetesjournals.org on 3 October 2008. DOI: 10.2337/db08-0198.

© 2009 by the American Diabetes Association. Readers may use this article as long as the work is properly cited, the use is educational and not for profit, and the work is not altered. See http://creativecommons.org/licenses/by-nc-nd/3.0/ for details.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

There is little information available regarding genes contributing to type 2 diabetes in the indigenous or diasporic populations of sub-Saharan Africa. To date, there have been only three reported linkage scans for type 2 diabetes in populations of African descent: two in African Americans (1,2) and one in African families from Ghana and Nigeria (3). Although there have been several recent genome-wide association studies (GWASs) conducted in primarily European populations, none has been reported for African Americans, and relatively few diabetes genes have been found in African American populations using candidate gene approaches (4). Consequently, we have few insights into genetic susceptibility factors in African Americans contributing to greater type 2 diabetes prevalence.

To better understand the genetics of type 2 diabetes in African Americans, we have studied Gullah-speaking African Americans living in coastal communities and on the sea islands of South Carolina. The ancestors of the Gullahs derived from the "grain coast" of West Africa and were forcibly imported because their rice-growing expertise was critical for the culture of this cash crop on low country plantations (5). Gullah-speaking African Americans have high rates of type 2 diabetes, characterized by relatively high rates of diabetic complications, early age of onset, and a high relative risk to siblings, λ_S, of type 2 diabetes at 3.3 (6). The diet is uniformly rich in animal fats, suggesting diabetes and obesity susceptibility alleles may more predictably produce a corresponding phenotype. Although there has been some emigration to northern American cities, there has been little immigration of African Americans born elsewhere into the Sea Islands. Studies of admixture indicate that the Gullah people are the most homogeneous population of African descent in the U.S., with Caucasian admixture below 3.5% (6–8), the lowest documented for any African American population. Analyses of mitochondrial and Y-chromosomal markers show that the genetic distance between the Gullah and Sierra Leonean tribes is measurably shorter than other African American populations (8–10).

Given the relatively low European admixture, diet high in animal fats, and increased prevalence and familial clustering of diabetes, studies of families from this population were anticipated to provide unique insights into predominantly "African"-derived diabetes loci. Thus, we initiated the Sea Islands Genetic African American Registry (Project SuGAR). Type 2 diabetes–affected sibling,

half-sibling, or parent-child pairs were recruited and assessed for medical, anthropometrical, and metabolic phenotypes in affected and nonaffected family members to conduct a whole-genome linkage scan. This scan is the first to be conducted for type 2 diabetes in African Americans using the higher resolution single nucleotide polymorphism (SNP) linkage panel.

RESEARCH DESIGN AND METHODS

This study was conducted under Institutional Review Board approval from the Medical University of South Carolina, the University of Alabama at Birmingham (UAB), and Wake Forest University School of Medicine and adhered to the tenets of the Declaration of Helsinki. Project SuGAR enlisted medical clinics, churches, and established organizations on the Sea Islands to aid in identifying patients with type 2 diabetes who belonged to families with multiple affected members (6). Inclusion criteria included self-described African American race, at least one type 2 diabetes–affected sibling pair, no more than one of the parents affected with type 2 diabetes, and at least one parent still living. Probands and their parents were all born and raised in the South Carolina low country.

Project SuGAR assessed medical, anthropometrical, and metabolic information on all consenting affected and nonaffected family members. The data were collected based on a multipage questionnaire, detailed family history and medical history, standardized blood pressures, physical examination, body dimensions, estimation of percent body fat, and laboratory testing. Weights were determined using electronic calibrated scales (Detecto, Cleveland, OH) at 8:00–10:00 A.M. after voiding and before breakfast. Heights were measured with a portable Harpenden statiometer. BMI (kg/m^2) was calculated. Standard arm, waist, hip, and thigh circumferences were recorded using a tension-controlled tape measure (Novel Products, Rockton, IL). Laboratory testing included complete blood count, electrolytes, creatinine/blood urea nitrogen, liver function tests, A1C, fasting lipid panel (cholesterol, triglycerides, and HDL), circulating islet cell antibodies (if diabetic), fasting glucose, and urine albumin-to-creatinine ratio. Diabetes was confirmed in cases using fasting glucose measures and/or need for diabetes medications coupled with review of medical records. All participating nondiabetic family members were evaluated with an oral glucose tolerance test or by fasting glucose. The criteria established by the National Diabetes Data Group as modified by the Expert Committee of the American Diabetes Association were used to define subjects as diabetic, impaired fasting glucose, impaired glucose tolerance, and normal glucose tolerance (NGT). The current genome scan involved a total of 521 individuals, including 471 affected subjects and 50 unaffected relatives who were recruited from 197 families. We included all phenotyped nondiabetic relatives in the ascertained families to assist in generating accurate phase (and hence identity by descent [IBD]) information. The mean pedigree size was 2.6 relatives, and pedigree sizes ranged from 2 to 7 individuals. One hundred twenty-one pedigrees contained 2 genotyped individuals, 48 pedigrees contained 3 genotyped individuals, and 28 pedigrees contained more than 3 genotyped individuals. For the purposes of linkage analyses, phenotype categories were defined as affected (type 2 diabetes), unaffected (NGT), and unknown (used for ungenotyped relatives required to connect genotyped individuals).

Genotyping. DNA was extracted from 20–40 ml venous blood using a standardized DNA isolation kit (Gentra Systems, Minneapolis, MN). The Project SuGAR registry includes 70 sibpairs plus available parents totaling 162 participants who were part of the Genetics of Non-Insulin Dependent Diabetes (GENNID) study. For the GENNID subjects, blood was sent to the central laboratory for lymphocyte transformation, and DNA extraction was performed by Coriell Cell Repositories.

A genome-wide scan was completed by the Center for Inherited Disease Research (CIDR) using Illumina HumanLinkage Panel IVb. A total of 5,974 SNPs were successfully genotyped, with a mean spacing of 0.65 cM (518 kb). The missing data rate was 0.26% (17,434 missing genotypes/6,626,408 total genotypes), and after correction or removal of likely misspecified relationships as determined using the genetic data (see below), the Mendelian consistency rate was 99.99% (535 events/6,292,704 study genotypes). The blind duplicate reproducibility rate was 99.998% (7 events/321,713 paired genotypes). Thirteen SNPs were removed from analyses because they violated Hardy-Weinberg assumptions ($P < 0.0001$).

Primary linkage analyses. Each pedigree was examined for consistency of familial relationships using the Pedigree Relationship Statistical Test (11). When the self-reported familial relationships were strongly inconsistent with the genotypic data for that pedigree, then *1*) the pedigree was modified when the IBD statistics suggested a very clear alternative, or *2*) a minimal set of genotypic data was converted to missing. A total of 46 pedigrees (23.7%) exhibited probable misspecified familial relationships and were modified as above, with 43.3% of these changes from a full sibling to half-sibling. After modifying all family relationships that appeared to be inconsistent with the genome scan data, there were the following affected relationship pairs: 152 full sibpairs, 55 half sibpairs, 43 parent-offspring pairs, 6 grandparent-grandchild pairs, 65 avuncular pairs, 18 first-cousin pairs, 16 half-avuncular pairs, and 2 half first-cousin pairs. Each marker was examined for Mendelian inconsistencies using PedCheck (12), and sporadic problem genotypes were converted to missing. Allele frequency estimates were computed using the maximum likelihood methods implemented in the software Recode (D. Weeks, personal communication). Map distances were based on the Rutger's genetic map (13). Where two SNPs displayed linkage disequilibrium values of $r^2 > 0.3$, we removed one SNP of the pair; 230 SNPs were removed for this reason.

The data were analyzed using the nonparametric linkage $(NPL)_{pairs}$ statistic and multipoint engine for rapid likelihood inference (MERLIN) (14). All results presented in the tables and figures represent multipoint analyses. We computed NPL regression analyses using the NPL_{pairs} statistics outputted from MERLIN, which we modified (15–17). The models without covariates test for excess allele sharing and are asymptotically equivalent to the MERLIN results.

Multilocus tests of heterogeneity and genome × genome interaction analyses. The NPL regression approach uses a conditional logistic regression likelihood with the family-specific NPL statistic at the locus of interest as the independent variable (15–17). The primary advantage of this regression-based approach is that it allows for the simultaneous evaluation of multiple loci and their interactions. That is, because NPL regression is a regression analysis, it allows for multiple loci to be in the model and tests for linkage at one locus adjusted for evidence for linkage at the other loci in the model. In this sense, it accounts for genetic heterogeneity. The multilocus model building was completed using stepwise conditional logistic regression allowing all autosomal loci in the genome at 0.5 cM spacing to be candidates to enter the model. Model building proceeded using standard stepwise regression methods with entry and exit criteria at $P = 0.05$. In stepwise methods, a locus enters the model if the locus provides evidence for linkage while adjusting for the evidence for linkage at all other loci in the model. Once a locus enters the model, all loci are tested for linkage, conditional on the other loci in the model. If on inclusion of a new locus, a previously significant locus is no longer significant, the latter is removed.

To test for an interaction, or epistasis, between two loci (genome by genome interaction analyses), we included the two loci and their statistical interaction into the model and computed the significance of the coefficient for the interaction term using a 1-degree of freedom test. As an exploratory tool, we computed all such two-locus interactions at every 2.5 cM across the entire genome. The shift to every 2.5 cM is due to the number of pairs of loci. Simulations show that little is lost in linkage analyses with this increased grain. Although a large number of comparisons were made, P values $<10^{-5}$ were considered indicative of epistasis between loci in these exploratory tests. A Bonferroni correction was applied for the number of comparisons made; however, this exploratory analysis should be viewed with caution given the large number of tests computed.

Ordered subsets linkage analysis. A series of ordered subset analyses (OSAs) (18) were computed to investigate the influence of the mean age of type 2 diabetes diagnosis of affected family members, BMI, waist circumference, and waist-to-hip ratio (WHR) on linkage analyses. Analyses were conducted ranking the family level means for these parameters in ascending, and then in descending, order. For waist and WHR, we used the residuals computed from a linear model that predicts the trait as a function of age, sex, and their interaction as the trait of interest for the OSA. The statistical significance of the change in the logarithm of odds (LOD) score was evaluated by a permutation test under the null hypothesis that the ranking of the covariate is independent of the LOD score of the family on the target chromosome. Thus, the families were randomly permuted with respect to the covariate ranking, and an analysis proceeded as above for each permutation of these data. The resulting empirical distribution of the change in the LOD scores yielded a chromosome-wide P value (ΔP).

NPL regression and OSA methods are described in greater detail in the online appendix of Sale et al. (2).

RESULTS

Population characteristics. The clinical and phenotypic characteristics for the diabetes-affected individuals who were genotyped as part of the genome-wide scan are summarized in Table 1. The genotyped population was 76.8% female, probably reflecting participation bias. The

TABLE 1
Characteristics of diabetic African American subjects

Trait	*n*	Mean	Median	SD	Range
Age at study entry (years)	466	55.2	55.7	13.7	14.3–101
Age at diabetes diagnosis (years)	449	43.4	44.0	14.1	4–85
Duration of diabetes (years)	449	11.7	9.0	9.9	0–52
A1C (%)	401	9.0	8.8	2.2	4.1–20.6
BMI (kg/m^2)	436	33.6	32.8	7.2	17.3–53.5
Waist circumference (cm)	423	105.7	104.0	15.4	75–155
Waist-to-hip ratio	420	0.91	0.91	0.08	0.64–1.30

diabetes-affected individuals are obese (median BMI 32.8 kg/m^2) and have relatively poor glucose control (median A1C 8.8%, normal range 4.5–5.7). The median age at diagnosis (43 ± 15 years) was relatively young; 8 years earlier than the first published study of type 2 diabetes in African Americans, which had a mean age of onset of 51 (1) and comparable with the mean age of the families described by Sale et al. (2) of 41 ± 12 years.

Primary linkage results. Genome-wide linkage results are shown in Fig. 1, and all LOD scores >1 from linkage analyses are presented in Table 2. Six regions of the genome yielded LOD scores >1. Chromosome 14 at 123.6 cM had the strongest evidence for linkage with type 2 diabetes (LOD 2.10; Fig. 2). Other regions that provided modest evidence for linkage included chromosome 1 at 167.5 cM (LOD 1.51), chromosome 3 at 121.0 cM (LOD 1.61), and three peaks on chromosome 7 at 29.5 cM (LOD 1.15), 44.5 cM (LOD 1.18), and 78.0 cM (LOD 1.64).

Multilocus conditional logistic regression results. The results of the multilocus NPL regression model are also shown in Table 2. Two chromosomal regions (one on 14q and one on 7p) were retained in the model (using $P <$ 0.05 as our threshold) after adjusting for the evidence for

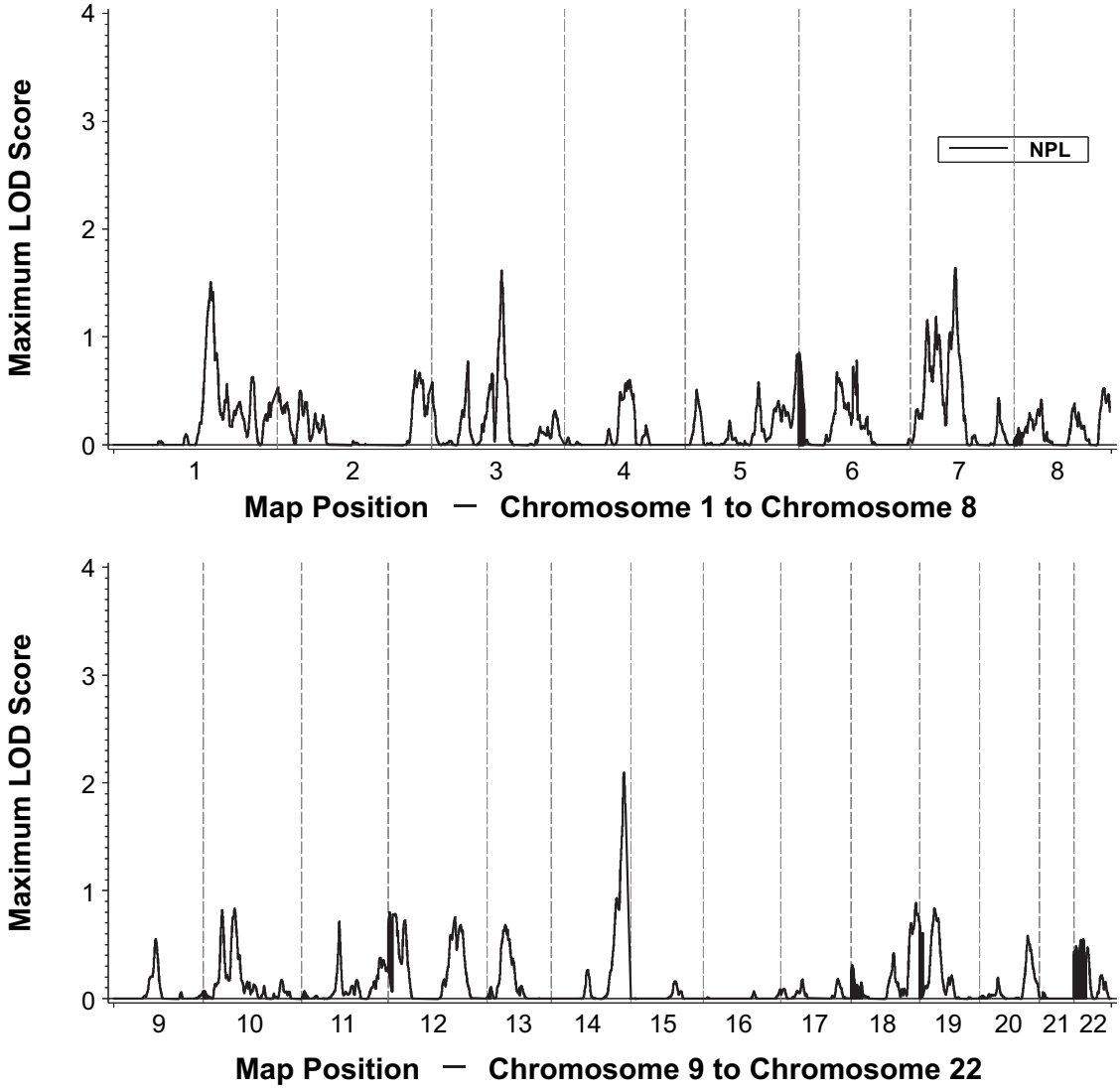

FIG. 1. Genome-wide linkage, NPL, results for type 2 diabetes.

TABLE 2
Linkage results with LOD > 1.0 and multilocus conditional logistic regression results

Chromosome	Position (cM)	Flanking markers	Primary linkage analysis			Multilocus conditional logistic regression analysis	
			LOD	LOD-1 interval	*P* value	LOD	LOD-1 interval
1	167.5	rs1319898/rs869714	1.51	155.0–181.5	0.00842		
3	121.0	rs1317244/rs12736	1.61	113.5–127.5	0.0064		
7	29.5	rs726395/rs1029718	1.15	7.0–96.5	0.0212	1.62*	23.0–55.0
7	44.5	rs1404282/rs1860759	1.18	7.5–60.5	0.0195		
7	78.0	rs1105305/rs517258	1.64	64.5–88.5	0.00598		
14	123.6	rs1132975/rs988131	2.10	117.1-tel	0.00189	2.52*	118.39-tel

*The evidence for linkage on 7p is adjusted for linkage on 14q, and similarly, the 14q locus is adjusted for linkage at 7p.

linkage at the other locus. Comparisons of the linkage and multilocus conditional logistic regression results for chromosomes 14 and 7 are shown in Figs. 2 and 3, respectively. Conditional on the model containing these two loci, no other regions of the genome provided evidence of linkage.

Genome × genome interaction analyses. Four regions provided evidence for an interaction between two chromosomal regions (supplementary Table 1, available in an online appendix at http://dx.doi.org/10.2337/db08-0198). The interaction two-dimensional response surface is shown in supplementary Fig. 1, available in the online appendix. The *P* values for these four instances of epistatic loci were considered robust relative to the number of comparisons per chromosome (corrected *P* value range 0.005–0.02). None of the regions identified in these analyses showed single-locus evidence for linkage. These analyses can be considered exploratory.

OSA. The OSA found differential evidence for linkage depending on age at type 2 diabetes diagnosis and BMI, but no increased evidence for linkage was detected subsetting on age-adjusted measures of waist or WHR. Regions displaying an increase in the LOD score equivalent to a chromosome-wide *P* value (ΔP) of <0.05 are shown in Table 3. Three of the four strongest results were seen with age at diagnosis. Subset analysis on the 105 pedigrees (54%) with the earliest age of diagnosis increased the chromosome 7p LOD score from 1.64 to 3.93 ($\Delta P = 0.0052$) at 78 cM, as shown in Fig. 3. In contrast, subsetting on the 120 pedigrees (61%) with the latest age at type 2 diabetes diagnosis increased the chromosome 14 LOD score from 2.06 to 4.05 ($\Delta P = 0.0069$) at 123.1 cM (Fig. 2). A third region on chromosome 18 at 91.0 cM also showed evidence for linkage in the subset of pedigrees with earliest age at type 2 diabetes diagnosis ($\Delta P = 0.0074$ for the

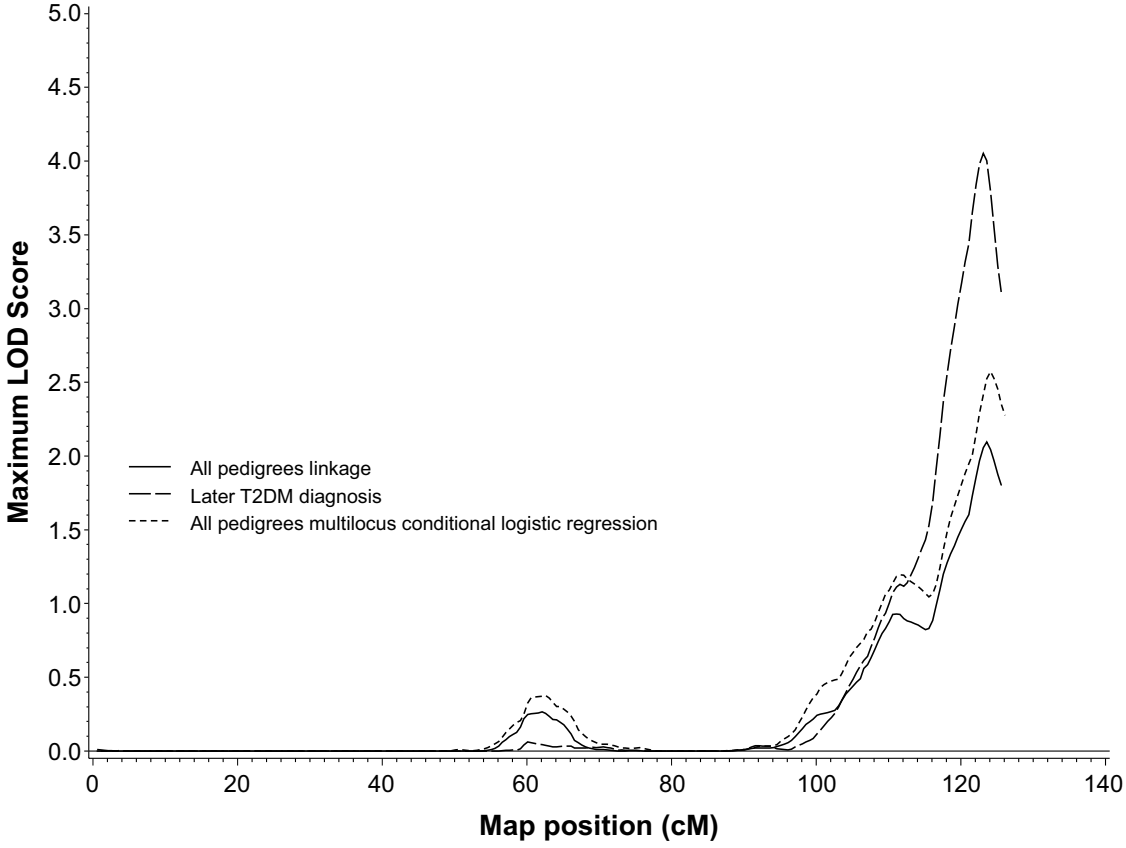

FIG. 2. Chromosome 14 results using the primary linkage approach (solid line), the multilocus conditional logistic regression model (dashed line), and the OSA analysis with later age at type 2 diabetes diagnosis (T2DM) (long dashed line).

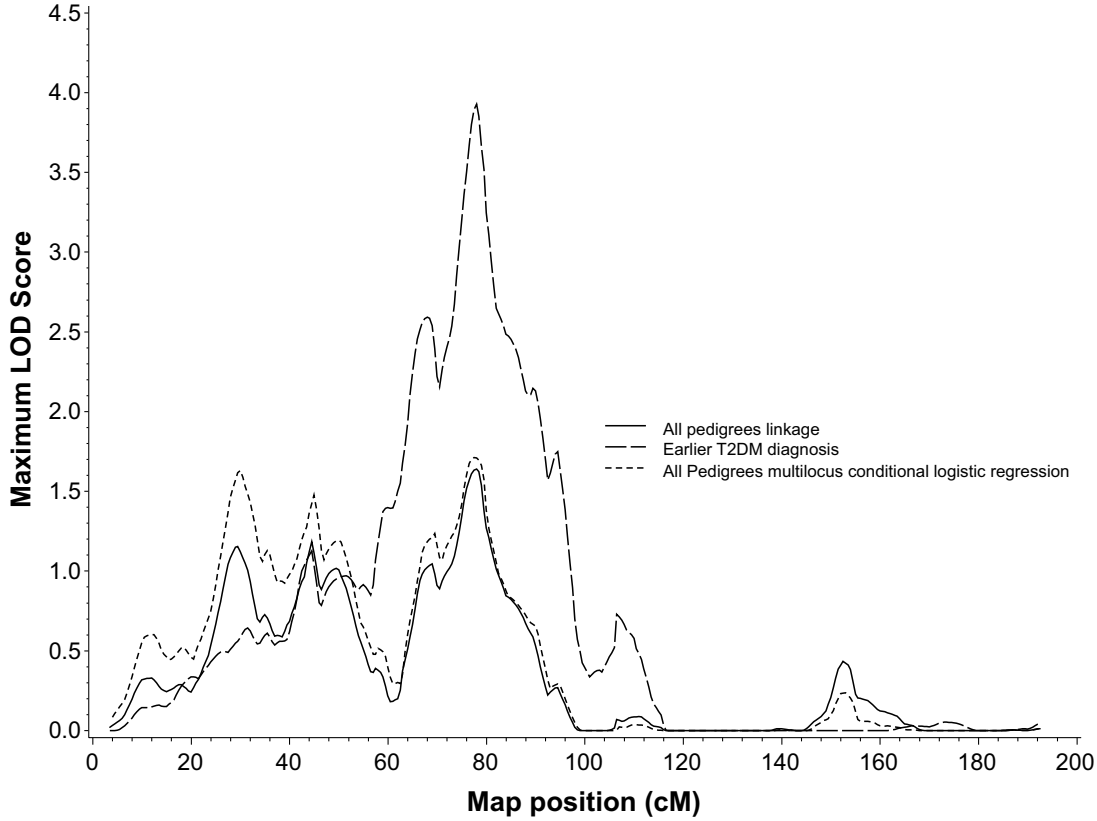

FIG. 3. Chromosome 7 results using the primary linkage approach (solid line), the multilocus conditional logistic regression model (dashed line), and the OSA analysis with earlier age at type 2 diabetes (T2DM) diagnosis (long dashed line).

change in LOD score from 0.09 to 3.81), although the number of pedigrees linked at this region was considerably fewer (16%, 32 pedigrees). Similarly, 50 pedigrees (26%) with the lowest mean BMI values showed increased evidence of linkage on chromosome 17 at 5.5 cM (ΔP = 0.0049; LOD score change 0.09 to 2.78). It is also interesting to note that the borderline increased evidence for linkage at 120 cM on chromosome 3 in the subset containing the 73% of pedigrees with earlier mean age at diagnosis ($\Delta P = 0.042$) overlaps with the chromosome 3 single locus result at 121 cM (Table 2).

DISCUSSION

The history of the Gullah-speaking African American population has resulted in relatively low European admixture

TABLE 3
Ordered subset analyses ($\Delta P < 0.05$) of age at diagnosis and BMI

Chromosome	Linked subset	Flanking markers	Position (cM)	Entire sample LOD	Maximum LOD	Optimal subset	Remaining families	Empirical *P* value for change	Proportion of pedigrees
3	Early age diagnosis	rs1512532/rs1398748	120.0	1.52	2.93	39.43 ± 8.19	56.00 ± 5.48	0.0419	0.73
4	High BMI	rs1456860/rs1450900	75.8	0.12	2.72	39.35 ± 3.70	30.56 ± 3.14	0.0167	0.36
7	Early age diagnosis	rs1105305/rs517258	78.0	1.64	3.93	36.54 ± 7.70	52.40 ± 6.00	0.0052	0.54
9	High BMI	rs994367/rs560764	53.0	0.02	2.30	39.47 ± 3.69	30.64 ± 3.18	0.0149	0.35
9	Early age diagnosis	rs2026406/rs927632	71.5	0.54	1.84	37.92 ± 7.93	53.81 ± 5.83	0.0321	0.62
9	Late age diagnosis	rs1819730/rs1407850	110.0	0	2.76	62.42 ± 4.40	42.16 ± 9.19	0.0486	0.09
12	High BMI	rs617022/rs1558776	15.0	0.62	2.84	37.08 ± 4.03	28.67 ± 2.39	0.0176	0.60
14	Late age diagnosis	rs1547350/rs6644	123.1	2.06	4.05	50.37 ± 6.40	33.74 ± 7.30	0.0069	0.61
16	Low BMI	rs870856/rs869048	131.1	0.07	1.95	26.07 ± 1.75	35.05 ± 4.65	0.0247	0.15
17	Low BMI	rs12939286/rs11062	5.5	0.09	2.78	27.43 ± 2.10	35.89 ± 4.38	0.0049	0.26
18	Early age diagnosis	rs1517162/rs565973	91.0	0.09	3.81	27.04 ± 6.48	47.22 ± 7.64	0.0074	0.16

Data are means ± SD.

that, when coupled with a diet rich in saturated fats, has produced high rates of type 2 diabetes. The first linkage scan performed in this population using a high-density SNP linkage panel has revealed a novel locus on 14q and two suggestive loci on chromosome 7 that appear to act independently and have stronger support in specific subsets.

The highest linkage peak for type 2 diabetes was seen on chromosome 14 at 123–124 cM (LOD 2.10), and this locus also showed increased evidence for linkage in a subset with later age at type 2 diabetes diagnosis (maximum LOD 4.05). This locus does not appear to have been reported previously; any chromosome 14 linkages and significant GWAS results for related phenotypes are more than 20 cM proximal to this region. The traits linked at this locus suggest that it may take some time to result in disease development. There are few obvious diabetes candidate genes under this peak, although this region does contain *AKT1*, a mediator of insulin and IGF-I signaling (19,20). One study of this gene in an Ashkenazi Jewish population did not find an association with type 2 diabetes (21).

The type 2 diabetes linkage peak identified on chromosome 7 at 77.5 cM (LOD 1.64) overlapped with a locus for earlier age at type 2 diabetes diagnosis (78.0 cM, maximum LOD 3.93). Linkage with early age at type 2 diabetes diagnosis has previously been reported at 62 cM in a French population (22). Candidate genes under the LOD-1 intervals for the three chromosome 7 peaks in Table 2 include previously identified type 2 diabetes genes glucokinase 1 (23), interleukin 6 (24), and growth factor receptor-bound protein 10 (25,26) and IGF binding proteins IGF2BP3, IGFBP1, and IGFBP3. The IGF pathway is now suspected to play a role in diabetes because of observed associations with IGF2BP2 (27–29).

The modest linkage peak on chromosome 1 at 167.5 cM (LOD 1.51) is within the International Chromosome 1 Diabetes Genetics Consortium region (30), which includes an African American population from Arkansas (31), and is also close to the reported association with intergenic SNP rs2501354 (28). There were no other major loci that overlapped with prior type 2 diabetes linkage scans in populations of African descent (1–3), possibly because of the modest power of all African American linkage studies to date, genetic heterogeneity, and/or differences in population history, including ancestral origins and population bottlenecks. Studies of mitochondrial and Y-chromosomal markers have determined that the genetic distance between the Gullah and Sierra Leonean tribes (Mende, Temne, etc.) is quite short and measurably shorter than other African American populations (8–10); thus, study-specific loci may represent ancestral differences between the Gullah and the Ghanaian and Nigerian families of the Africa America Diabetes Mellitus study (3). Interestingly, the region of chromosome 10 containing the transcription factor 7-like 2 (*TCF7L2*) gene—shown to be important in populations with African ancestry (32–34)—did not produce evidence for linkage in this population.

Although GWASs have proven effective in identifying novel type 2 diabetes genes in European populations, association with *CDKAL1* SNP rs7756992 was not successfully replicated in a West African population (35), and "confirmed" diabetes genes—calpain 10, K^+ inwardly rectifying channel, subfamily J, member 11 (*KCNJ11*), peroxisome proliferator–activated receptor-γ (*PPARG*), and hepatocyte nuclear factor 4α (*HNF4A*)—showed modest or no association in our prior studies of a different African American type 2 diabetic case-control population (32). A recent study in the same African American case-control population investigating type 2 diabetes loci identified from GWASs of European populations confirmed that the majority of these loci, with the exception of TCF7L2, do not have a major contribution to type 2 diabetes risk in African Americans (36). Currently, there are no published reports of GWASs for type 2 diabetes in populations of African descent, although it is highly likely that future GWASs of African and African American populations will reveal novel type 2 diabetes susceptibility loci. In the absence of African American GWAS data for type 2 diabetes at present, the current linkage study adds to our knowledge of putative susceptibility-containing loci in this high-risk population. Because of the lack of overlap between linkage peaks and GWAS loci for an increasing number of disorders investigated using both approaches in well-powered studies, speculation is increasing that linkage peaks may represent regions containing both allelic and genetic heterogeneity, i.e., multiple uncommon susceptibility variants in one or more genes. Thus, it is plausible that linkage analyses may identify novel loci containing multiple uncommon risk alleles of high penetrance that may not be captured under current GWAS SNP tagging approaches of common variants because genotyping products are constructed to tag only common alleles and capture lower levels of variation in African-derived populations due to decreased linkage disequilibrium. However, if the few known type 2 diabetes linkage loci in African Americans represent common alleles, they may be detected using a GWAS approach. In contrast to contemporary African populations, the relative homogeneity of ancestry and cultural factors such as diet in the Project SuGAR population is anticipated to result in increased expressivity of risk alleles, while still identifying susceptibility loci relevant to African-derived populations. Independent diabetes loci on chromosomes 14 and 7 warrant investigation in additional African American populations and follow-up analyses in the Gullah-speaking African American population.

ACKNOWLEDGMENTS

M.M.S. has received National Institutes of Health (NIH) Grant DK-66358 and a Career Development Award from the American Diabetes Association. W.T.G. has received NIH Grant DK-47461 and a GENNID Study Center grant from the American Diabetes Association. Project SuGAR has received a grant from the W.M. Keck Foundation (Los Angeles, CA). This work was supported by the General Clinical Research Center at the Medical University of South Carolina (Grant M01-RR-1070), by the Genetics Core Facility of the UAB Clinical Nutrition Research Unit (Grant P30-DK-56336), and by the South Carolina Center of Biomedical Research Excellence (COBRE) for Oral Health Grant P20-RR-17696. CIDR is funded through a federal contract from the NIH to Johns Hopkins University (contract N01-HG-65403).

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

Parts of this study were presented in poster form at the 68th Scientific Sessions of the American Diabetes Association, San Francisco, California, 6–10 June 2008.

We gratefully acknowledge the contribution of the Gullah families who participated in Project SuGAR, the Citizen Advisory Committee who provided guidance

throughout the life of the project, the Penn Community Center (Saint Helena Island, SC), and Federal Qualified Community Health Centers serving low country citizens that housed Project SuGAR staff and assisted in recruitment, including the Franklin C. Fetter Family Health Center (Charleston, SC), the St. James-Santee Family Health Center (McClellanville, SC), the Sea Island Medical Center (St. Johns Island, SC), the Family Health Center (Orangeburg, SC), the Beaufort-Jasper Comprehensive Health Services (Ridgeland, SC), and in particular the Leroy E. Browne Medical Center branch (Saint Helena Island, SC) and the Elijah Washington Medical Center branch (Sheldon, SC). We also thank multiple personnel in the Division of Endocrinology at the Medical University of South Carolina, including office and data manager Ann Smuniewski; data entry manager Cedric Rivers; the Project Sugar nurse coordinators Pam Wilson, Susan Cromwell, Frederica Hudges-Joyner, Karen Wilder-Smalls, Guinevere M. Maine, Mattie Wideman, Gloria Smith, Deborah Daniels, Andrea Collins, Montrese Edwards, and Janet Carter; and laboratory and analytic support from Barbara Wojciechowski, George Argyropoulos, David McLean, Kerin McCormack, Nikki Rogers, Miranda Marion, Joyce Hicks, Theresa Kearns, and Dell Curry. We are grateful for the assistance of other staff, friends, supporters, volunteers, social worker interns, and nursing and medical students. Finally, we are thankful for the operation of the Project SuGAR Van by Rosalyn Cato and Brother William Frazier.

REFERENCES

1. Ehm MG, Karnoub MC, Sakul H, Gottschalk K, Holt DC, Weber JL, Vaske D, Briley D, Briley L, Kopf J, McMillen P, Nguyen Q, Reisman M, Lai EH, Joslyn G, Shepherd NS, Bell C, Wagner MJ, Burns DK: Genomewide search for type 2 diabetes susceptibility genes in four American populations. *Am J Hum Genet* 66:1871–1881, 2000
2. Sale MM, Freedman BI, Langefeld CD, Williams AH, Hicks PJ, Colicigno CJ, Beck SR, Brown WM, Rich SS, Bowden DW: A genome-wide scan for type 2 diabetes in African-American families reveals evidence for a locus on chromosome 6q. *Diabetes* 53:830–837, 2004
3. Rotimi CN, Chen G, Adeyemo AA, Furbert-Harris P, Parish-Gause D, Zhou J, Berg K, Adegoke O, Amoah A, Owusu S, Acheampong J, Agyenim-Boateng K, Eghan BA Jr, Oli J, Okafor G, Ofoegbu E, Osotimehin B, Abbiyesuku F, Johnson T, Rufus T, Fasanmade O, Kittles R, Daniel H, Chen Y, Dunston G, Collins FS: A genome-wide search for type 2 diabetes susceptibility genes in West Africans: the Africa America Diabetes Mellitus (AADM) Study. *Diabetes* 53:838–841, 2004
4. 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
5. Opala J: *The Gullah: Rice, Slavery, and the Sierra Leone-American Connection*. Freetown, Sierra Leone, United States Information Service, 1987
6. Garvey WT, McClean DC, Spruill I: The search for obesity genes in isolated populations: Gullah-speaking African Americans and the role of uncoupling protein 3 as a thrifty gene. In *Progress in Obesity Research*. Medeiros G, Halpern A, Bouchard C, Eds. Paris, France, John Libbey Eurotext, 2003, p. 373–380
7. 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: Ancestral proportions and admixture dynamics in geographically defined African Americans living in South Carolina. *Am J Phys Anthropol* 114:18–29, 2001
8. McLean DC Jr, Spruill I, Gevao S, Morrison EY, Bernard OS, Argyropoulos G, Garvey WT: Three novel mtDNA restriction site polymorphisms allow exploration of population affinities of African Americans. *Hum Biol* 75:147–161, 2003
9. Jackson BA, Wilson JL, Kirbah S, Sidney SS, Rosenberger J, Bassie L, Alie JA, McLean DC, Garvey WT, Ely B: Mitochondrial DNA genetic diversity among four ethnic groups in Sierra Leone. *Am J Phys Anthropol* 128:156–163, 2005
10. McLean DC Jr, Spruill I, Argyropoulos G, Page GP, Shriver MD, Garvey WT: Mitochondrial DNA (mtDNA) haplotypes reveal maternal population genetic affinities of Sea Island Gullah-speaking African Americans. *Am J Phys Anthropol* 127:427–438, 2005
11. Sun L, Wilder K, McPeek MS: Enhanced pedigree error detection. *Hum Hered* 54:99–110, 2002
12. O'Connell JR, Weeks DE: PedCheck: a program for identification of genotype incompatibilities in linkage analysis. *Am J Hum Genet* 63:259–266, 1998
13. Kong X, Murphy K, Raj T, He C, White PS, Matise TC: A combined linkage-physical map of the human genome. *Am J Hum Genet* 75:1143–1148, 2004
14. Abecasis GR, Cherny SS, Cookson WO, Cardon LR: Merlin-rapid analysis of dense genetic maps using sparse gene flow trees. *Nat Genet* 30:97–101, 2002
15. Langefeld CD, Davis CC, Brown WM: Nonparametric linkage regression. I: combined Caucasian CSGA and German genome scans for asthma. *Genet Epidemiol* 21 (Suppl. 1):S136–S141, 2001
16. Langefeld CD: *The Application of Statistical Methods for Highly Stratified Data to Linkage and Association Analyses of Complex Genetic Traits*. Ann Arbor, MI, University of Michigan, 1999
17. Langefeld CD, Boehnke M: Multiple trait locus nonparametric linkage regression (Abstract). *Am J Hum Genet* 65 (Suppl.):A25, 1999
18. Hauser ER, Watanabe RM, Duren WL, Bass MP, Langefeld CD, Boehnke M: Ordered subset analysis in genetic linkage mapping of complex traits. *Genet Epidemiol* 27:53–63, 2004
19. Cong LN, Chen H, Li Y, Zhou L, McGibbon MA, Taylor SI, Quon MJ: Physiological role of Akt in insulin-stimulated translocation of GLUT4 in transfected rat adipose cells. *Mol Endocrinol* 11:1881–1890, 1997
20. Galetic I, Andjelkovic M, Meier R, Brodbeck D, Park J, Hemmings BA: Mechanism of protein kinase B activation by insulin/insulin-like growth factor-1 revealed by specific inhibitors of phosphoinositide 3-kinase: significance for diabetes and cancer. *Pharmacol Ther* 82:409–425, 1999
21. Matsubara A, Wasson JC, Donelan SS, Welling CM, Glaser B, Permutt MA: Isolation and characterization of the human AKT1 gene, identification of 13 single nucleotide polymorphisms (SNPs), and their lack of association with type II diabetes. *Diabetologia* 44:910–913, 2001
22. Froguel P, Vaxillaire M, Sun F, Velho G, Zouali H, Butel MO, Lesage S, Vionnet N, Clement K, Fougerousse F, Tanizawa, Y, Weissenbach J, Beckmann JS, Lathrop GM, Passa P, Permutt MA, Cohen D: Close linkage of glucokinase locus on chromosome 7p to early-onset non-insulin-dependent diabetes mellitus. *Nature* 356:162–164, 1992
23. Hattersley AT, Turner RC: Mutations of the glucokinase gene and type 2 diabetes. *Q J Med* 86:227–232, 1993
24. Vozarova B, Fernandez-Real JM, Knowler WC, Gallart L, Hanson RL, Gruber JD, Ricart W, Vendrell J, Richart C, Tataranni PA, Wolford JK: The interleukin-6 (-174) G/C promoter polymorphism is associated with type-2 diabetes mellitus in Native Americans and Caucasians. *Hum Genet* 112: 409–413, 2003
25. Di Paola R, Ciociola E, Boonyasrisawat W, Nolan D, Duffy J, Miscio G, Cisternino C, Fini G, Tassi V, Doria A, Trischitta V: Association of hGrb10 genetic variations with type 2 diabetes in Caucasian subjects. *Diabetes Care* 29:1181–1183, 2006
26. Rampersaud E, Damcott CM, Fu M, Shen H, McArdle P, Shi X, Shelton J, Yin J, Chang YP, Ott SH, Zhang L, Zhao Y, Mitchell BD, O'Connell J, Shuldiner AR: Identification of novel candidate genes for type 2 diabetes from a genome-wide association scan in the Old Order Amish: evidence for replication from diabetes-related quantitative traits and from independent populations. *Diabetes* 56:3053–3062, 2007
27. The Wellcome Trust Case-Control Consortium: Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature* 447:661–678, 2007
28. Scott LJ, Mohlke KL, Bonnycastle LL, Willer CJ, Li Y, Duren WL, Erdos MR, Stringham HM, Chines PS, Jackson AU, Prokunina-Olsson L, Ding CJ, Swift AJ, Narisu N, Hu T, Pruim R, Xiao R, Li XY, Conneely KN, Riebow NL, Sprau AG, Tong M, White PP, Hetrick KN, Barnhart MW, Bark CW, Goldstein JL, Watkins L, Xiang F, Saramies J, Buchanan TA, Watanabe RM, Valle TT, Kinnunen L, Abecasis GR, Pugh EW, Doheny KF, Bergman RN, Tuomilehto J, Collins FS, Boehnke M: A genome-wide association study of type 2 diabetes in Finns detects multiple susceptibility variants. *Science* 316:1341–1345, 2007
29. Saxena R, Voight BF, Lyssenko V, Burtt NP, de Bakker PI, Chen H, Roix JJ, Kathiresan S, Hirschhorn JN, Daly MJ, Hughes TE, Groop L, Altshuler D, Almgren P, Florez JC, Meyer J, Ardlie K, Bengtsson Bostrom K, Isomaa B, Lettre G, Lindblad U, Lyon HN, Melander O, Newton-Cheh C, Nilsson P,

Orho-Melander M, Rastam L, Speliotes EK, Taskinen MR, Tuomi T, Guiducci C, Berglund A, Carlson J, Gianniny L, Hackett R, Hall L, Holmkvist J, Laurila E, Sjogren M, Sterner M, Surti A, Svensson M, Tewhey R, Blumenstiel B, Parkin M, Defelice M, Barry R, Brodeur W, Camarata J, Chia N, Fava M, Gibbons J, Handsaker B, Healy C, Nguyen K, Gates C, Sougnez C, Gage D, Nizzari M, Gabriel SB, Chirn GW, Ma Q, Parikh H, Richardson D, Ricke D, Purcell S: Genome-wide association analysis identifies loci for type 2 diabetes and triglyceride levels. *Science* 316:1331–1336, 2007

30. Das SK, Elbein SC: The search for type 2 diabetes susceptibility loci: the chromosome 1q story. *Curr Diab Rep* 7:154–164, 2007
31. Das SK, Chu WS, Hale TC, Wang X, Craig RL, Wang H, Shuldiner AR, Froguel P, Deloukas P, McCarthy MI, Zeggini E, Hasstedt SJ, Elbein SC: Polymorphisms in the glucokinase-associated, dual-specificity phosphatase 12 (DUSP12) gene under chromosome 1q21 linkage peak are associated with type 2 diabetes. *Diabetes* 55:2631–2639, 2006
32. Sale MM, Smith SG, Mychaleckyj JC, Keene KL, Langefeld CD, Leak TS, Hicks PJ, Bowden DW, Rich SS, Freedman BI: Variants of the transcription factor 7-like 2 (TCF7L2) gene are associated with type 2 diabetes in an African-American population enriched for nephropathy. *Diabetes* 56:2638–2642, 2007
33. Munoz J, Lok KH, Gower BA, Fernandez JR, Hunter GR, Lara-Castro C, De Luca M, Garvey WT: Polymorphism in the transcription factor 7-like 2 (TCF7L2) gene is associated with reduced insulin secretion in nondiabetic women. *Diabetes* 55:3630–3634, 2006
34. Helgason A, Palsson S, Thorleifsson G, Grant SF, Emilsson V, Gunnarsdottir S, Adeyemo A, Chen Y, Chen G, Reynisdottir I, Benediktsson R, Hinney A, Hansen T, Andersen G, Borch-Johnsen K, Jorgensen T, Schafer H, Faruque M, Doumatey A, Zhou J, Wilensky RL, Reilly MP, Rader DJ, Bagger Y, Christiansen C, Sigurdsson G, Hebebrand J, Pedersen O, Thorsteinsdottir U, Gulcher JR, Kong A, Rotimi C, Stefansson K: Refining the impact of TCF7L2 gene variants on type 2 diabetes and adaptive evolution. *Nat Genet* 39:218–225, 2007
35. Steinthorsdottir V, Thorleifsson G, Reynisdottir I, Benediktsson R, Jonsdottir T, Walters GB, Styrkarsdottir U, Gretarsdottir S, Emilsson V, Ghosh S, Baker A, Snorradottir S, Bjarnason H, Ng MC, Hansen T, Bagger Y, Wilensky RL, Reilly MP, Adeyemo A, Chen Y, Zhou J, Gudnason V, Chen G, Huang H, Lashley K, Doumatey A, So WY, Ma RC, Andersen G, Borch-Johnsen K, Jorgensen T, van Vliet-Ostaptchouk JV, Hofker MH, Wijmenga C, Christiansen C, Rader DJ, Rotimi C, Gurney M, Chan JC, Pedersen O, Sigurdsson G, Gulcher JR, Thorsteinsdottir U, Kong A, Stefansson K: A variant in CDKAL1 influences insulin response and risk of type 2 diabetes. *Nat Genet* 39:770–775, 2007
36. Lewis JP, Palmer ND, Hicks PJ, Sale MM, Langefeld CD, Freedman BI, Divers J, Bowden DW: Association analysis in African Americans of European-derived type 2 diabetes single nucleotide polymorphisms from whole-genome association studies. *Diabetes* 57:2220–2225, 2008