A 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\textsuperscript{1,2,3}, Lingyi Lu\textsuperscript{4}, Ida J. Spruill\textsuperscript{5}, Jyotika K. Fernandes\textsuperscript{5}, Kerry H. Lok\textsuperscript{6}, Jasmin Divers\textsuperscript{4}, Carl D. Langefeld\textsuperscript{4}, W. Timothy Garvey\textsuperscript{6,7}.

\textsuperscript{1}Center for Public Health Genomics, \textsuperscript{2}Department of Medicine, \textsuperscript{3}Department of Biochemistry and Molecular Genetics, University of Virginia, Charlottesville, VA; \textsuperscript{4}Department of Biostatistical Sciences, Wake Forest University School of Medicine, Winston-Salem, NC; \textsuperscript{5}Department of Medicine, Medical University of South Carolina, Charleston, SC; \textsuperscript{6}Department of Nutrition Sciences, University of Alabama at Birmingham, and the \textsuperscript{7}Birmingham VA Medical Center, Birmingham, AL.

Address correspondence to:
Michèle M. Sale, Ph.D.
Center for Public Health Genomics
University of Virginia
PO Box 800717
Email: msale@virginia.edu


Additional information for this article can be found in an online appendix at http://diabetes.diabetesjournals.org

This is an uncopyedited electronic version of an article accepted for publication in Diabetes. The American Diabetes Association, publisher of Diabetes, is not responsible for any errors or omissions in this version of the manuscript or any version derived from it by third parties. The definitive publisher-authenticated version will be available in a future issue of Diabetes in print and online at http://diabetes.diabetesjournals.org.

Copyright American Diabetes Association, Inc., 2008
ABSTRACT

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 (T2DM) and diabetic complications. Affected relative pairs with T2DM were recruited through Project SuGAR (the Sea Islands Genetic African American Registry).

Research Design and Methods: We conducted a genome-wide linkage scan, genotyping 5,974 SNPs in 471 affected subjects and 50 unaffected relatives from 197 pedigrees. Data were analyzed using MERLIN, and ordered subsets analyses (OSA) for age at T2DM diagnosis, waist circumference, waist-to-hip ratio (WHR), 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 T2DM (LOD 2.10) and the subset with later age at T2DM diagnosis (max. LOD 4.05). Two linkage peaks on chromosome 7 were detected at 44-45 cM for T2DM (LOD 1.18), and 78 cM for T2DM (LOD 1.64) and the subset with earlier age at T2DM diagnosis (max. 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), chromosome 3 at 121.0 cM (LOD 1.61).

Conclusions: This study revealed a novel T2DM locus in an AA population on 14q that appears to reduce age of disease onset, and confirmed two loci on chromosome 7.
There is little information available regarding genes contributing to type 2 diabetes (T2DM) in the indigenous or diasporic populations of sub-Saharan Africa. To date there have been only three reported linkage scans for T2DM in populations of African descent: two in African Americans (AA) (1; 2) and one in African families from Ghana and Nigeria (3). Although there have been several recent genome-wide association studies (GWAS) conducted in primarily European populations, none have been reported for AA, and relatively few diabetes genes have been found in AA populations using candidate gene approaches (4). Consequently we have few insights into genetic susceptibility factors in AA contributing to greater T2DM prevalence.

To better understand the genetics of type 2 diabetes in AA, 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 AA have high rates of T2DM, characterized by relatively high rates of diabetic complications, early age of onset, and a high relative risk to siblings, $\lambda_S$, of T2DM 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. While 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 United States, with Caucasian admixture below 3.5% (7-9), the lowest documented for any African American population. Analyses of mitochondrial and Y-chromosomal markers show the genetic distance between the Gullah and Sierra Leone tribes is measurably shorter than other African American populations (8; 10; 11).

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 Project SuGAR (the Sea Islands Genetic African American Registry). T2DM-affected sibling, half-sibling, or parent-child pairs were recruited, and assessed for medical, anthropometrical, and metabolic phenotypes in affected and non-affected family members, in order to conduct a whole genome linkage scan. This scan is the first to be conducted for T2DM in AA using the higher resolution SNP linkage panel.

**RESEARCH DESIGN AND METHODS**

**Subjects.** 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 T2DM who belonged to families with multiple affected members (6). Inclusion criteria included self-described African American race, at least one T2DM-affected sibling pair, no more than one of the parents affected with T2DM, 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 non-affected
family members. The data were collected based on a multi-page questionnaire, detailed family history and medical history, standardized blood pressures, physical examination, body dimensions and estimation of percent body fat, and laboratory testing. Weights were determined using electronic calibrated scales (Detecto, Cleveland, OH) at 8-10AM after voiding and before breakfast. Heights were measured with a portable Harpenden stadiometer. Body mass index (kg/m²) 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/BUN, liver function tests, hemoglobin A₁C, fasting lipid panel (cholesterol, triglycerides, HDL), circulating islet cell antibodies (if diabetic), fasting glucose, and urine albumin/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 non-diabetic family members were evaluated with an oral glucose tolerance test (OGTT) 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 (IFG), impaired glucose tolerance (IGT), 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 non-diabetic relatives in the ascertained families to assist in generating accurate phase (and hence IBD) information. The mean pedigree size was 2.6 relatives, and pedigree sizes ranged from 2-7 individuals. One hundred twenty-one pedigrees contained more than 3 genotyped individuals. For the purposes of linkage analyses, phenotype categories were defined as affected (T2DM), unaffected (NGT), and unknown (used for ungenotyped relatives required to connect genotyped individuals).

**Genotyping.** DNA was extracted from 20-40 ml of venous blood using a standardized DNA isolation kit (Gentra Systems, Minneapolis, MN). The Project SuGAR registry includes 70 sib-pairs 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’s 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 (12). When the self-reported familial relationships were strongly inconsistent with the genotypic data for that pedigree, then 1) the pedigree was modified when the identity-by-descent statistics suggested a very clear alternative, or 2) a minimal set of genotypic data were 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-sib pairs, 55 half-sib pairs, 43 parent-offspring pairs, 6 grandparent-grandchild pairs, 65 avuncular pairs, 18 first-cousin pairs, 16 half-avuncular pairs, 2 half-first-cousin pairs. Each marker was examined for Mendelian inconsistencies using PedCheck (13), 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 (14). Where two SNPs displayed LD 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 MERLIN (multipoint engine for rapid likelihood inference) (15). 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 (16-18). The models without covariates test for excess allele sharing and are asymptotically equivalent to the MERLIN results.

**Multilocus tests of heterogeneity, and genome x 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 (16-18). 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 criterion at p-value = 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, upon 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 include the two loci and their statistical interaction into the model and compute 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) (19) were computed to investigate the influence of the mean age of T2DM 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, gender and their interaction as the trait of interest for the OSA.
The statistical significance of the change in
the 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 diabetes-affected
individuals are obese (median BMI 32.8
kg/m²), and have relatively poor glucose
control (median HbA1c 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 T2DM
in African Americans, which had a mean age
of onset of 51 (1), and comparable to the
mean age of the families described by Sale et
al. (2) of 41±12 years.

Primary linkage results. Genome-wide
linkage and multilocus conditional logistic
regression results are shown in Figure 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;
Figure 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 linkage at the other locus. A
comparison of the linkage and multilocus
conditional logistic regression results for
chromosomes 14 and 7 are shown in Figures
2 and 3 respectively. Conditional on the
model containing these two loci, no other
regions of the genome provided evidence of
linkage.

Genome x genome interaction analyses. Four
regions provided evidence for an interaction
between two chromosomal regions
(Supplementary Table 1). The interaction
two-dimensional response surface is shown in
Supplementary Figure 1. 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 to 0.02). None of the
regions identified in these analyses showed
single-locus evidence for linkage. These
analyses can be considered exploratory.

Ordered subsets analysis. The OSA found
differential evidence for linkage depending on
age at T2DM 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, $\Delta P$, 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 Figure 3. In contrast, subsetting on the 120 pedigrees (61%) with the latest age at T2DM diagnosis increased the chromosome 14 LOD score from 2.06 to 4.05 ($\Delta P = 0.0069$) at 123.1 cM (Figure 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 T2DM diagnosis ($\Delta P = 0.0074$ for the 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 ($\Delta P = 0.0049$; 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).

CONCLUSIONS
The history of the Gullah-speaking African American population has resulted in relatively low European admixture that, when coupled with a diet rich in saturated fats, has produced high rates of T2DM. 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 T2DM 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 T2DM diagnosis (max. LOD 4.05). This locus does not appear to have been reported previously; any chromosome 14 linkages and significant genome-wide association study (GWAS) results for related phenotypes are more than 20 cM proximal to this region. The traits linked at this locus suggest 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 insulin-like growth factor 1 (IGF1) signaling (20; 21). One study of this gene in an Ashkenazi Jewish population did not find an association with T2DM (22).

The T2DM linkage peak identified on chromosome 7 at 77.5 cM (LOD 1.64) overlapped with a locus for earlier age at T2DM diagnosis (78.0 cM, max. LOD 3.93). Linkage with early age at T2DM diagnosis has previously been reported at 62 cM in a French population (23). Candidate genes under the LOD-1 intervals for the three chromosome 7 peaks in Table 2 include previously-identified T2DM genes glucokinase (GCK1 (24)), interleukin 6 (IL6 (25)), and growth factor receptor-bound protein 10 (GRB10) (26; 27), as well IGF binding proteins IGF2BP3, IGFBP1, IGFBP3. The IGF pathway is now suspected to play a role in diabetes because of observed associations with IGF2BP2 (28-30).

The modest linkage peak on chromosome 1 at 167.5 cM (LOD 1.51) is within the International Chromosome 1 Diabetes Genetics Consortium region (31) which includes an AA population from Arkansas (32), and is also close to the reported association with intergenic SNP rs2501354 (29). There were no other major loci that overlapped with prior T2DM linkage scans in populations of African descent (1-3), possibly due to the modest power of all AA 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 Leonese tribes (Mende, Temne, etc.) are quite short and measurably shorter than other African American populations (7-9), thus study specific loci may represent ancestral differences between the Gullah and the Ghanaian and Nigerian families of the Africa-America Diabetes Mellitus (AADM) 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 (33-35) – did not produce evidence for linkage in this population.

Although GWAS have proven effective in identifying novel T2DM genes in European populations, association with CDKAL1 SNP rs7756992 was not successfully replicated in a West African population (36), and "confirmed" diabetes genes calpain 10 (CAPN10), KCNJ11, PPARG, and HNF4A showed modest or no association in our prior studies of a different AA T2DM case-control population (33). A recent study in the same AA case-control population investigating T2DM loci identified from GWAS of European populations confirmed that the majority of these loci, with the exception of TCF7L2, do not have a major contribution to T2DM risk in AA (37). Currently there are no published reports of GWAS for T2DM in populations of African descent, although it is highly likely that future GWAS of African and African American populations will reveal novel T2DM susceptibility loci. In the absence of African American GWAS data for T2DM at present, the current linkage study adds to our knowledge of putative susceptibility-containing loci in this high-risk population. Due to the lack of overlap between linkage peaks and GWAS loci for an increasing number of disorders investigated using both approaches in well-powered studies, there is increasing speculation 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 since genotyping products are constructed to tag only common alleles and capture lower levels of variation in African-derived populations due to decreased LD. However, if the few known T2DM linkage loci in AA 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 AA populations and follow-up analyses in the Gullah-speaking AA population.

ACKNOWLEDGEMENTS

The authors 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 on 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 in Charleston SC, the St. James-Santee Family Health Center in McClellanville SC, the Sea Island Medical Center on St. Johns Island SC, the Family Health Center in Orangeburg SC, and the Beaufort-Jasper Comprehensive Health Services in Ridgeland SC and in particular the
Leroy E. Browne Medical Center branch on Saint Helena Island and the Elijah Washington Medical Center branch in 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, BS, the Project Sugar nurse coordinators Pam Wilson, LPN, Susan Cromwell, LPN, Frederica Hudges-Joyner RN, BSN, Karen Wilder-Smalls, LPN, Guinevere M. Maine, RN, Mattie Wideman, LPN, Gloria Smith, LPN, Deborah Daniels, LPN, Andrea Collins, LPN, Montrese Edwards, RN, BSN, and Janet Carter RD, and laboratory and analytic support from Barbara Wojciechowski, MA, George Argyropoulos, PhD, David McLean, PhD, Kerin McCormack MS, Nikki Rogers, PhD, 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.

Project SuGAR could not have been accomplished without a grant from the W. M. Keck Foundation, Los Angeles CA. We acknowledge the support of the General Clinical Research Center at the Medical University of South Carolina (M01 RR-1070), and the Genetics Core Facility of the UAB Clinical Nutrition Research Unit (P30 DK-56336). The work was supported by research grant support from the National Institutes of Health including DK47461 (WTG), DK66358 (MMS), the South Carolina Center of Biomedical Excellence (COBRE) for Oral Health P20 RR17696 (Steve London, PI), and from the American Diabetes Association in the form of a GENNID Study Center grant (WTG) and Career Development Award (MMS). CIDR is funded through a federal contract from the NIH to Johns Hopkins University, contract number N01-HG-65403.
REFERENCES


17. 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
28. 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
Linkage scan for type 2 diabetes in African Americans


nucleotide polymorphisms from whole-genome association studies. Diabetes 57:2220-2225, 2008
### Table 1. Characteristics of diabetic African American subjects

<table>
<thead>
<tr>
<th>Trait</th>
<th>n</th>
<th>Mean</th>
<th>Median</th>
<th>SD</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age at study entry (years)</td>
<td>466</td>
<td>55.2</td>
<td>55.7</td>
<td>13.7</td>
<td>14.3-101</td>
</tr>
<tr>
<td>Age at diabetes diagnosis (years)</td>
<td>449</td>
<td>43.4</td>
<td>44.0</td>
<td>14.1</td>
<td>4-85</td>
</tr>
<tr>
<td>Duration of diabetes (years)</td>
<td>449</td>
<td>11.7</td>
<td>9.0</td>
<td>9.9</td>
<td>0-52</td>
</tr>
<tr>
<td>HbA1c (%)</td>
<td>401</td>
<td>9.0</td>
<td>8.8</td>
<td>2.2</td>
<td>4.1-20.6</td>
</tr>
<tr>
<td>BMI (kg/m²)</td>
<td>436</td>
<td>33.6</td>
<td>32.8</td>
<td>7.2</td>
<td>17.3-53.5</td>
</tr>
<tr>
<td>Waist circumference (cm)</td>
<td>423</td>
<td>105.7</td>
<td>104.0</td>
<td>15.4</td>
<td>75-155</td>
</tr>
<tr>
<td>Waist to hip ratio</td>
<td>420</td>
<td>0.91</td>
<td>0.91</td>
<td>0.08</td>
<td>0.64-1.30</td>
</tr>
</tbody>
</table>
Table 2. Linkage results with LOD > 1.0, and multilocus conditional logistic regression results

<table>
<thead>
<tr>
<th>Chromosome</th>
<th>Position of max. LOD (cM)</th>
<th>Flanking Markers</th>
<th>Primary linkage analysis</th>
<th>Multilocus conditional logistic regression analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>LOD</td>
<td>LOD-1</td>
</tr>
<tr>
<td>1</td>
<td>167.5</td>
<td>rs1319898/rs869714</td>
<td>1.51</td>
<td>155.0-181.5</td>
</tr>
<tr>
<td>3</td>
<td>121.0</td>
<td>rs1317244/rs12736</td>
<td>1.61</td>
<td>113.5-127.5</td>
</tr>
<tr>
<td>7</td>
<td>29.5</td>
<td>rs726395/rs1029718</td>
<td>1.15</td>
<td>7.0-96.5</td>
</tr>
<tr>
<td>7</td>
<td>44.5</td>
<td>rs1404282/rs1860759</td>
<td>1.18</td>
<td>7.5-60.5</td>
</tr>
<tr>
<td>7</td>
<td>78.0</td>
<td>rs1105305/rs517258</td>
<td>1.64</td>
<td>64.5-88.5</td>
</tr>
<tr>
<td>14</td>
<td>123.6</td>
<td>rs1132975/rs988131</td>
<td>2.10</td>
<td>117.1-tel</td>
</tr>
</tbody>
</table>

*The evidence for linkage on 7p is adjusted for linkage on 14q, and similarly, the 14q locus is adjusted for linkage at 7p.
Table 3. Ordered subset analyses (Δp<0.05) of age at diagnosis and BMI

<table>
<thead>
<tr>
<th>Chr</th>
<th>Linked subset</th>
<th>Flanking Markers</th>
<th>Position (cM)</th>
<th>Entire Sample LOD</th>
<th>Maximized LOD</th>
<th>Optimal subset LOD</th>
<th>Remaining families</th>
<th>Empirical P value for change</th>
<th>Proportion of Pedigrees</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>Early Age Dx</td>
<td>rs1512532 / rs1398748</td>
<td>120.0</td>
<td>1.52</td>
<td>2.93</td>
<td>39.43±8.19</td>
<td>56.00±5.48</td>
<td>0.0419</td>
<td>0.73</td>
</tr>
<tr>
<td>4</td>
<td>High BMI</td>
<td>rs1456860 / rs1450900</td>
<td>75.8</td>
<td>0.12</td>
<td>2.72</td>
<td>39.35±3.70</td>
<td>30.56±3.14</td>
<td>0.0167</td>
<td>0.36</td>
</tr>
<tr>
<td>7</td>
<td>Early Age Dx</td>
<td>rs1105305 / rs517258</td>
<td>78.0</td>
<td>1.64</td>
<td>3.93</td>
<td>36.54±7.70</td>
<td>52.40±6.00</td>
<td>0.0052</td>
<td>0.54</td>
</tr>
<tr>
<td>9</td>
<td>High BMI</td>
<td>rs994367 / rs560764</td>
<td>53.0</td>
<td>0.02</td>
<td>2.30</td>
<td>39.47±3.69</td>
<td>30.64±3.18</td>
<td>0.0149</td>
<td>0.35</td>
</tr>
<tr>
<td>9</td>
<td>Early Age Dx</td>
<td>rs2026406 / rs927632</td>
<td>71.5</td>
<td>0.54</td>
<td>1.84</td>
<td>37.92±7.93</td>
<td>53.81±5.83</td>
<td>0.0321</td>
<td>0.62</td>
</tr>
<tr>
<td>9</td>
<td>Late Age Dx</td>
<td>rs1819730 / rs1407850</td>
<td>110.0</td>
<td>0.00</td>
<td>2.76</td>
<td>62.42±4.40</td>
<td>42.16±9.19</td>
<td>0.0486</td>
<td>0.09</td>
</tr>
<tr>
<td>12</td>
<td>High BMI</td>
<td>rs617022 / rs1558776</td>
<td>15.0</td>
<td>0.62</td>
<td>2.84</td>
<td>37.08±4.03</td>
<td>28.67±2.39</td>
<td>0.0176</td>
<td>0.60</td>
</tr>
<tr>
<td>14</td>
<td>Late Age Dx</td>
<td>rs1547350 / rs6644</td>
<td>123.1</td>
<td>2.06</td>
<td>4.05</td>
<td>50.37±6.40</td>
<td>33.74±7.30</td>
<td>0.0069</td>
<td>0.61</td>
</tr>
<tr>
<td>16</td>
<td>Low BMI</td>
<td>rs870856 / rs869048</td>
<td>131.1</td>
<td>0.07</td>
<td>1.95</td>
<td>26.07±1.75</td>
<td>35.05±4.65</td>
<td>0.0247</td>
<td>0.15</td>
</tr>
<tr>
<td>17</td>
<td>Low BMI</td>
<td>rs12939286 / rs11062</td>
<td>5.5</td>
<td>0.09</td>
<td>2.78</td>
<td>27.43±2.10</td>
<td>35.89±4.38</td>
<td>0.0049</td>
<td>0.26</td>
</tr>
<tr>
<td>18</td>
<td>Early Age Dx</td>
<td>rs1517162 / rs565973</td>
<td>91.0</td>
<td>0.09</td>
<td>3.81</td>
<td>27.04±6.48</td>
<td>47.22±7.64</td>
<td>0.0074</td>
<td>0.16</td>
</tr>
</tbody>
</table>
Figure 1. Genome-wide linkage results for type 2 diabetes.
Figure 2. Chromosome 14 results using the primary linkage approach (solid line), multilocus conditional logistic regression model (dashed line), and the OSA analysis with later age at T2DM diagnosis (long dashed line).
Figure 3. Chromosome 7 results using the primary linkage approach (solid line), multilocus conditional logistic regression model (dashed line), and the OSA analysis with earlier age at T2DM diagnosis (long dashed line).