Published in final edited form as:

Diabetes Metab Res Rev. 2009 November ; 25(8): 740–747. doi:10.1002/dmrr.1031.

Genome-wide linkage scans for type 2 diabetes mellitus in four ethnically diverse populations; significant evidence for linkage on chromosome 4q in African Americans: the Family Investigation of Nephropathy and Diabetes (FIND) Research Group

Alka Malhotra¹, Robert P. Igo Jr.², Farook Thameem³, W.H. Linda Kao⁴, Hanna E. Abboud³, Sharon G. Adler⁵, Nedat H. Arar³, Donald W. Bowden⁶, Ravindranath Duggirala⁷, Barry I. Freedman⁶, Katrina A.B. Goddard², Eli Ipp⁵, Sudha K. Iyengar², Paul L. Kimmel⁸, William C. Knowler¹, Orly Kohn⁹, David Leehey¹⁰, Lucy A. Meoni⁴, Robert G. Nelson¹, Susanne B. Nicholas¹¹, Rulan S. Parekh⁴, Stephen S. Rich¹², Yii-Der I. Chen¹¹, Mohammed F. Saad¹³, Marina Scavini^{14,15}, Jeffrey R. Schelling², John R. Sedor², Vallabh O. Shah¹⁴, Kent D. Taylor¹¹, Denyse Thornley-Brown¹⁶, Philip G. Zager¹⁴, Amanda Horvath², and Robert L. Hanson¹ on behalf of the Family Investigation of Nephropathy and Diabetes Research Group¹⁷

¹National Institute of Diabetes and Digestive and Kidney Diseases, Phoenix, Arizona ²Case Western Reserve University, Cleveland Ohio ³University of Texas Health Science Center at San Antonio, San Antonio, Texas ⁴Johns Hopkins University, Baltimore, Maryland ⁵Harbor-University of California Los Angeles Medical Center, Torrance, California ⁶Wake Forest University, Winston-Salem, North Carolina ⁷Southwest Foundation for Biomedical Research, San Antonio, Texas ⁸National Institute of Diabetes and Digestive and Kidney Diseases Program Office, Bethesda, Maryland ⁹University of Chicago, Chicago, Illinois ¹⁰Loyola University, Chicago, Illinois ¹¹University of California Los Angeles, Los Angeles, California ¹²University of Virginia, Charlottesville, Virginia ¹³Caritas Carney Hospital, Dorchester, Massachusetts ¹⁴University of New Mexico, Albuquerque, New Mexico ¹⁵San Raffaele Scientific Institute, Milan (Italy) ¹⁶University of Alabama Birmingham, Birmingham, Alabama

Abstract

Background—Previous studies have shown that, in addition to environmental influences, type 2 diabetes mellitus (T2DM) has a strong genetic component. The goal of the current study is to identify regions of linkage for T2DM in ethnically diverse populations.

Methods—Phenotypic and genotypic data were obtained from African American (AA; total number of individuals (N)=1004), American Indian (AI; N=883), European American (EA; N=537), and Mexican American (MA; N=1634) individuals from the Family Investigation of Nephropathy and Diabetes. Nonparametric linkage analysis, using an average of 4,404 SNPs, was performed in relative pairs affected with T2DM in each ethnic group. In addition, family-based tests were performed to detect association with T2DM.

Corresponding Author: Sudha K. Iyengar Department of Epidemiology and Biostatistics Case Western Reserve University 10900 Euclid Avenue Cleveland, OH 44106 Phone: (216) 368-4388 ski@case.edu .

¹⁷A complete list of the FIND Research Group is given in the Appendix

CONFLICT OF INTEREST None

Results—Statistically significant evidence for linkage was observed on chromosomes 4q21.1 (LOD=3.13; genome-wide $p=0.04$) in AA. In addition, a total of eleven regions showed suggestive evidence for linkage (estimated at $\text{LOD}>1.71$), with the highest LOD scores on chromosomes 12q21.31 (LOD=2.02) and 22q12.3 (LOD=2.38) in AA, 2p11.1 (LOD=2.23) in AI, 6p12.3 (LOD=2.77) in EA, and 13q21.1 (LOD=2.24) in MA. While no region overlapped across all ethnic groups, at least five loci showing $\text{LOD}>1.71$ have been identified in previously published studies.

Conclusions—The results from this study provide evidence for the presence of genes affecting T2DM on chromosomes 4q, 12q, and 22q in AA, 6p in EA, 2p in AI, and 13q in MA. The strong evidence for linkage on chromosome 4q in AA provides important information given the paucity of diabetes genetic studies in this population.

Keywords

FIND; Type 2 Diabetes; linkage analysis; ethnicity

Introduction

Efforts to identify the genetic contributors to type 2 diabetes mellitus (T2DM) have been numerous, with over 30 genome-wide linkage studies identifying potential regions of linkage on almost every chromosome [1]. In addition, candidate gene association studies and, more recently, genome-wide association studies have identified single nucleotide polymorphisms (SNPs) with possible causal effects on T2DM [2,3]. However, there is a lack of clear replication in linkage studies which could be due to use of sample sizes that are too small to detect the effect of a gene that contributes a small proportion to the etiology of T2DM [4]. Furthermore, high rates of T2DM are observed in some ethnic groups including American Indians and Mexican Americans, and this suggests potential genetic differences among ethnic groups. Therefore, studying large samples from multiple ethnic backgrounds may allow better understanding of common genetic pathways that are critical to the underlying etiology of T2DM.

The Family Investigation of Nephropathy and Diabetes (FIND) was initiated to understand genetic factors underlying diabetic nephropathy, a common serious complication of diabetes [5]. FIND is a collaborative effort between eleven U.S. academic centers, where clinical and genetic data were collected from T2DM patients ascertained for diabetic nephropathy and their first-degree relatives, mostly siblings. Data from African American, American Indian, European American, and Mexican American families have been collected, providing a unique sample for genetic studies and permitting the exploration of ethnic differences. Using the available samples collected with diabetes and with or without nephropathy, the present study represents a nonparametric affecteds-only linkage analysis to identify regions linked to T2DM as well as a family-based association analysis for evidence of association in the presence of linkage.

Research design and methods

Study participants

The FIND participants were identified and examined at eleven centers, and they encompassed families of African American, American Indian, European American, and Mexican American descent [5]. Participants gave informed consent and approval was received from the Institutional Review Board at each center. Data were collected from 362 African American families (2-8 members), 214 American Indian families (2-20 members),

202 European American families (2-7 members), and 481 Mexican American families (2-13 members). Characteristics of study participants in the current study are given in Table 1.

Phenotypes and genotypes

Diabetes was defined as a history of diabetes and current or previous use of antidiabetic medication (oral agents and/or insulin). Individuals who had Hemoglobin A1c (HbA1c) $\geq 7.0\%$ or fasting serum glucose concentration ≥ 126 mg/dl were also classified as affected (HbA1c was measured in all participants and fasting plasma glucose was measured in a subset). Individuals with no history of diabetes and HbA1c $< 6.0\%$ were considered nondiabetic, while for those with HbA1c 6.0-6.9%, a fasting plasma glucose < 126 mg/dl was required to confirm the absence of diabetes in accordance with the 1997 American Diabetes Association criteria [6].

Single nucleotide polymorphisms (SNPs) were genotyped throughout the genome by the Center for Inherited Disease Research (CIDR) using the Illumina linkage IVb panel (<http://www.illumina.com/>). Proximity of these markers to each other showed high linkage disequilibrium between some of the markers. Redundant markers and those in high linkage disequilibrium with another marker (i.e., $D' > 0.5$), but less informative, were removed yielding a total of 4,702, 3,830, 4,562, and 4,546 analyzed SNPs in the African American, American Indian, European American, and Mexican American populations, respectively. DeCode maps were used to obtain centimorgan distances for all populations. Information content and marker density were similar for all ethnic groups; the average information content across ethnic groups was 0.78 and average marker spacing was 0.85 centimorgans.

Statistical analysis

Genome-wide linkage scans were performed for T2DM within each ethnic group. Given the availability of full sibships and, in some cases, extended family members, linkage can be assessed by allele sharing among affected relatives or by methods that compare allele sharing among relatives concordant and discordant for affection status (e.g. regression methods). Many of the families in FIND had small numbers of nondiabetic siblings, resulting in few discordant relative pairs. As a consequence, instability was observed in the linkage results when regression methods were employed; only results from the affected pairs analyses are presented.

Prior to linkage analysis, genotypes that produced Mendelian errors were detected using the program Merlin [7]; these genotypes were assigned missing values. Nonparametric linkage (NPL) analysis was performed using the NPL_{all} statistic, as implemented in Merlin [7]. In this method, identity-by-descent (IBD) probabilities are estimated for all affected pairs across all inheritance patterns. The IBDs are used in a score statistic, which is then converted to a LOD score by the method of Kong and Cox [8].

Genome-wide significance of the linkage signal was estimated for each population by 'gene dropping' using 200 replicates while maintaining family structure and patterns of missing data [9]. LOD scores were estimated by performing genome-wide NPL linkage for each replicate followed by identification of the highest LOD score across the genome for each replicate. To estimate genome-wide p-values (ie, the number of times a test statistic at least as extreme as that observed in the data is seen by chance under the null hypothesis of no linkage at any location in the genome), the proportion of replicates for which the maximum LOD across the genome was greater than a given observed LOD was computed. Confidence intervals for this proportion were calculated from the binomial distribution.

Results over all ethnic groups were combined from each individual genome scan using Fisher's method [10]. The test statistic for the combined evidence for linkage is defined as

$-2 \sum \ln(p_i)$, where p_i is the p-value of the i th independent population (i.e., African American, American Indian, Mexican American, or European American) at a given chromosomal location. In the FIND data, the combined p-value is determined from a statistic that is asymptotically distributed as a chi-square with 8 degrees of freedom.

While calculation of an empirical p-value allows for an assessment of genome-wide statistical significance, it is also of interest to examine results for regions with evidence that, while not sufficient for genome-wide significance, is nonetheless suggestive of linkage. By analysis of the permuted data, it was estimated that one would expect to observe a result with $\text{LOD} > 1.71$ once per genomic scan, on average, within each ethnic group. This empirical criterion is comparable to the definition of “suggestive” linkage that has been widely employed [11]. Thus, all regions with $\text{LOD} > 1.71$ ($\text{LOD} > 1.56$ for Fisher’s test) are presented as demonstrating suggestive evidence for linkage.

A within-family association test was performed for each individual SNP in the genome-wide linkage scan using the Quantitative Transmission Disequilibrium Test (QTDT) for discrete traits [12]. This approach tests whether transmission to an affected individual from his/her parents (founders) differs from that expected by chance. Association tests were performed separately for each ethnic group using all SNPs, i.e., including those removed due to high LD. Although the present density of SNPs is adequate for linkage, it is much too sparse to capture the genomic haplotypic variation reliably for a full genome-wide association study. The present association study was, therefore, performed to identify any associated SNPs in regions identified by the linkage genome scan or in regions identified by previous or ongoing genome-wide association studies.

Results

The characteristics of participants within each FIND ethnic group are shown in Table 1. Nonparametric linkage analysis was performed for T2DM (Figure 1 and supplemental figure 1). In African Americans, the highest LOD score was observed on chromosome 4q21.1 ($\text{LOD} = 3.13$; genome-wide p-value = 0.04, 95% confidence interval [CI] [0.01, 0.07]) (Table 2). Other regions exhibiting suggestive evidence for linkage ($\text{LOD} > 1.71$) include 12q21.31 ($\text{LOD} = 2.02$; genome-wide p-value = 0.42, 95% CI [0.35, 0.49]), and 22q12.3 ($\text{LOD} = 2.38$; genome-wide p-value = 0.23, 95% CI [0.18, 0.29]). All three regions also had $\text{LOD} > 1.56$ (suggestive evidence for linkage for the Fisher’s test results) when the results from individual ethnic groups were combined using Fisher’s test (4q21.1 $\text{LOD} = 1.71$; 12q21.31 $\text{LOD} = 2.01$; 22q12.3 $\text{LOD} = 2.32$) (Table 3).

Chromosome 6p12.3 ($\text{LOD} = 2.77$; genome-wide p-value = 0.12, 95% CI [0.07, 0.17]) showed evidence for linkage in European Americans and was also identified in the combined analysis ($\text{LOD} = 2.88$). Five regions gave a LOD score > 1.71 in the American Indian population. The highest LOD score was observed on chromosome 2p11.1 ($\text{LOD} = 2.23$; genome-wide p-value = 0.28, 95% CI [0.22, 0.35]). In addition, suggestive evidence for linkage was observed on 7p15.1 ($\text{LOD} = 1.78$; genome-wide p-value = 0.61, 95% CI [0.54, 0.68]), 15q21.1 ($\text{LOD} = 1.79$; genome-wide p-value = 0.61, 95% CI [0.54, 0.67]), 16p12.2 ($\text{LOD} = 1.90$; genome-wide p-value = 0.46, 95% CI [0.39, 0.53]), and 21q21.1 ($\text{LOD} = 1.94$; genome-wide p-value = 0.43, 95% CI [0.36, 0.50]). The region on 16p12.2 was also identified in the combined analysis using Fisher’s method with $\text{LOD} = 2.13$.

In Mexican Americans, the strongest evidence for linkage was observed on chromosome 13q21.1 ($\text{LOD} = 2.24$; genome-wide p-value = 0.27, 95% CI [0.21, 0.34]). 7p22.1 ($\text{LOD} = 1.84$; genome-wide p-value = 0.51, 95% CI [0.44, 0.58]) was also identified in this population.

Four regions that showed LOD<1.71 in all individual population genome scans showed suggestive evidence for linkage in the combined analysis of all populations. These include 1p22.1 (LOD=2.32), 3q27.3 (LOD=1.65), 6q14.1 (LOD=1.81), 6q16.1 (LOD=1.87), and 12q22 (LOD=1.69).

For the association tests, SNPs showing $p < 0.0025$ (since LOD=1.71 is approximately equivalent to $p = 0.0025$, which was used to assess the linkage analysis results) were identified. No SNPs at this p -value cut-off overlapped with regions identified in linkage scans performed in the same study populations. However, five SNPs with $p < 0.01$ were in regions with LOD>1 (Supplemental table 1). Furthermore, 13 regions in which a SNP had $p < 0.01$ in the present analysis also contained a SNP with $p < 0.0001$ in previously published association studies (ie, within 1Mb of the SNP in the current study) (Supplemental table 2). For two of these variants (rs7313 on chromosome 7 and rs703990 on chromosome 10), the same SNP associated in the present study had nominal evidence for replication ($p < 0.05$) in the large DIAGRAM meta-analysis [13]. In MA, $p < 0.0025$ was observed for 11 SNPs with 5 SNPs showing $P < 0.001$ (rs412735, $p = 0.0009$; rs7899305, $p = 0.0009$; rs1886040, $p = 0.0006$; rs1389504, $p = 0.0007$; rs680798, $p = 0.0008$). The most significant p -value in EA was seen on chromosome 18 (rs1075470, $p = 0.0012$), in AA on chromosome 10 (rs749694, $p = 0.0011$), and in AI on chromosome 2 (rs1356056, $p = 0.0004$). Two, three, and seven SNPs were identified with $p < 0.0025$ in EA, AA, and AI, respectively (Supplemental table 3).

Discussion

The current genome-wide linkage study revealed LOD scores >1.71 for type 2 diabetes on chromosomes 1p, 2p, 4q, 6p, 7p, 12q, 13q, 15q, 16p, 21q, and 22q in at least one of the four ethnic groups; however none of these signals was strongly replicated among ethnic groups. Given the design of FIND, with families ascertained as informative for diabetic nephropathy, it is possible that linkage signals in an affecteds-only analysis could arise from loci that influence risk for diabetic nephropathy rather than diabetes *per se*. However, linkage analyses for diabetic nephropathy in these families did not show strong evidence in the regions identified as diabetes mellitus-linked [14]. Previous analyses in the initial 378 FIND families also did not show evidence for linkage of nephropathy in any of these regions [15].

Although individuals with type 1 diabetes were not specifically excluded in the FIND, type 1 diabetes is uncommon in members of these ethnic groups, except European Americans. Based on history of insulin use and Glutamic Acid Decarboxylase (GAD) autoantibody measurements, the prevalence of type 1 diabetes was estimated to be <4% in FIND participants [16]. The peak of the chromosome 6p linkage signal observed in European Americans is ~9 cM from the HLA locus, a major susceptibility locus for type 1 diabetes. It is possible that this finding reflects members with type 1 diabetes among European American FIND participants, but it was difficult to classify diabetes based on the available data.

The highest LOD score was observed on chromosome 4q21 in African Americans with a significant genome-wide p -value ($p < 0.05$). This result has important implications, given the small number of genetic studies of diabetes in this population. Chromosome 4q21 must therefore be further analyzed to localize a gene(s) that might affect diabetes in African Americans. Modest evidence for linkage in this region (ie, within a 1-LOD interval of the linkage peak in the current study) has been identified in two previous studies, one in families from Finland and Sweden showing a LOD score of 1.41 [17] and a second study in individuals of Chinese descent (NPL=1.94) [18]. The NK homeobox, family 6, member A (NKX6A) gene is present near this region which plays a role in insulin secretion in beta-cells

[19]. In addition, the 1-LOD interval contains the betacellulin (*BTC*) gene which has been studied with reference to diabetes [20,21]. Specifically, variants in the *BTC* gene were associated with diabetes in one African American study [20], but these results were not replicated in another independent study in African Americans [21].

Suggestive evidence for linkage was also observed in African Americans on chromosome 12q21. Nearby regions (ie, within 5cM of the 1-LOD interval of the linkage peak in the current study) have been identified in at least five published studies, four from Caucasian populations and one African American population from the Genetics of NIDDM study [22-26]. Many studies have observed linkage with T2DM in this broad region on chromosome 12q21-24, which contains strong candidates including hepatocyte nuclear factor-1-alpha (*HNF1- α*) implicated in maturity onset of diabetes of the young (MODY) [27].

Chromosome 22q also showed a LOD >1.71, previously identified for T2DM in two African American studies (LOD=1.3 and 1.68, respectively) [28,29] and Canadian Oji-Cree families (LOD ~1.22) [30]. This region was also observed for diabetes-related traits in Hypertension Genetic Epidemiology Study (LOD=2.0 for a quantitative diabetes status trait) [31] and Pima Indians (LOD=1.8 for fasting plasma glucose levels) [32]. Genes affecting food intake and weight loss are also present in this region including the galanin receptor 3 (*GAL3*) [33] and leukemia inhibitory factor (*LIF*) [34], and could affect susceptibility to T2DM.

A LOD score of 2.23 was detected on chromosome 2p11.1 in American Indians, in a region identified in a Finnish population [35] and the Framingham study for fasting and mean glucose levels [36]. Similarly, only one region showed LOD>1.71 in European Americans on chromosome 6p12.3. This region was also identified in the combined analysis using Fisher's method. While this exact region has not been identified in previous studies, suggesting a possibly novel T2DM locus, a region close to it (6p21) has been identified in three previously published studies [18,37,38].

In Mexican Americans, the highest LOD score was observed in a region on chromosome 13q21.1 (LOD=2.24), which has not been observed in previous studies, suggesting a possible novel T2DM locus. Furthermore, while not showing suggestive evidence for linkage in the current study, chromosome 1q22 (LOD=1.60) was identified in this population which is a well replicated region identified in over five genome scans [39-43]. This region has been extensively studied by the chromosome 1q consortium and candidate genes involved in glucose metabolism including the insulin receptor-related receptor gene (*INSRR*) [44] and B-cell leukemia transcription factor 1 (*PBX1*) [45] are located here.

In the present study we estimated empirical p-values for each population independently. Since there are four populations included, further correction to achieve study-wide significance would result in a more stringent cut-off for significance; for example, the empirical p-value for the linkage signal on chromosome 4q21.1 in African Americans would increase from p=0.04 to p=0.15. However, given the interest in identification of genetic determinants in each population and the fact that the "prior" probability of linkage is unchanged by the number of populations included in a particular analysis, it is our opinion that the increased type II error incurred by such a "study-wide" threshold is not warranted.

In the association tests, five SNPs were identified in regions showing nominal evidence for linkage. While none of the linkage peaks overlapped with previously published association studies, 12 out of 128 SNPs associated with T2DM (p<0.01) in the present study are found in regions that have been previously associated with type 2 diabetes, at p<0.0001, in genome-wide association studies [13,46-49]. This suggests some evidence for region-wise replication, although only 2 SNPs had evidence for SNP-wise replication in the large

DIAGRAM meta-analysis at modest levels of significance (Supplemental table 2). Given the multiple testing burden, it is likely that some of these associations are false positives and additional replication studies are needed. It is noteworthy that one of the regions identified in the association tests includes insulin-like growth factor 2 mRNA-binding protein 2 (*IGF2BP2*, chr3q28), a potential candidate gene which also contains variants reproducibly associated with T2DM in genome-wide association studies [50].

In contrast to linkage studies, association studies have more power to identify individual genetic variants of interest, particularly when the risk conferred by a functional susceptibility genetic element is modest, provided a marker in strong linkage disequilibrium with such a functional element is typed. Strong evidence for linkage was not identified in any of the regions in which common variants reproducibly associated with type 2 diabetes are located (Supplemental table 4); the extent of association for these variants is too modest to reliably detect with linkage with the present sample size. The present set of markers, while suitable for linkage analyses, is not sufficiently dense to reliably capture alleles in strong linkage disequilibrium with common genetic variants. Therefore, the present association study will need to be followed-up using a denser marker set to capture this information.

In the present study, linkage analysis was used to identify potential loci conferring susceptibility to diabetes mellitus. The strongest linkage signals were not shared among the different ethnic groups and this suggests potential genetic heterogeneity among ethnic groups, at least for loci detected by linkage analysis. However, several regions which were not strongly linked in any individual group did show evidence for linkage among all groups combined; these regions may reflect genetic determinants that are shared across ethnic groups.

The use of linkage provides a complementary analytic method to association studies, since linkage can identify regions in which multiple susceptibility variants are present, particularly if disease variants are rare. In addition, the use of family data avoids biases which might be introduced in case-control studies, including population stratification. Furthermore, even though most LOD scores observed in this study represent nominal evidence for linkage, some of the linkage signals, particularly on chromosomes 1q and 12q, replicate those seen in other studies and, thus, provide further support for the presence of diabetes-susceptibility variant(s) in these regions.

This report lists regions linked with diabetes mellitus in families recruited in the FIND, a cohort enriched for members with severe diabetic nephropathy. A strength of this analysis is that it included relatively large numbers of African American, Mexican American, and American Indian families, minority groups typically under-represented in genetic analyses. Although family members may not uniformly have had type 2 diabetes mellitus, at least 96% had clinical characteristics compatible with this disorder. More detailed investigation of these regions is required to identify potential diabetes susceptibility variants.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This study was supported by Research Grant U01DK57292-05 from the National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK) and, in part, by the Intramural Research Program of the NIDDK. This work was supported by the National Center for Research Resources for the General Clinical Research Center (GCRC) grants: Case Western Reserve University M01-RR-000080, Wake Forest University M01-RR-07122,

Harbor-UCLA Medical Center M01-RR-00425, College of Medicine—University of California Irvine M01-RR-00827-29, University of New Mexico HSC M01-RR-00997, and Frederic C. Bartter M01-RR-01346. The Indian Health Service (I.H.S.) provided use of facilities. The opinions expressed in this paper are those of the authors and do not necessarily reflect the views of the I.H.S. or other funding sources. Genotyping was performed by the Center for Inherited Disease Research (CIDR). CIDR is fully funded through a federal contract from the National Institutes of Health to the Johns Hopkins University, contract no. N01-HG-65403. The results of this analysis were obtained by using the S.A.G.E. package of genetic epidemiology software, which is supported by a U.S. Public Health Service resource grant (RR03655) from the National Center for Research Resources. We kindly thank all FIND participants.

Appendix

Members of the Family Investigation of Nephropathy and Diabetes Research Group

Genetic Analysis and Data Coordinating Center, Case Western Reserve University, Cleveland, Ohio: SK Iyengar*, RC Elston**, KAB Goddard**, JM Olson**, S Ialacci#, J Fondran, A Horvath, R Igo, Jr., G Jun, K Kramp, J. Molineros, SRE Quade.

Participating investigation centers:

Case Western Reserve University, Cleveland, OH: JR Sedor*, J Schelling**, A Pickens#, L Humbert, L Getz-Fradley.

Harbor-University of California Los Angeles Medical Center: S Adler*, E Ipp**, M Pahl**†, MF Seldin** §, S Snyder**, J Tayek**, E Hernandez#, J LaPage#, C Garcia, J Gonzalez, M Aguilar.

Johns Hopkins University, Baltimore, MD: M. Klag*, R. Parekh*, L Kao**, L Meoni**, T Whitehead, J Chester#.

National Institute of Diabetes and Digestive and Kidney Diseases, Phoenix, AZ: WC Knowler*, RL Hanson**, RG Nelson**, L Jones#, R Juan, R Lovelace, C Luethe, LM Phillips, J Sewemaenewa, I Sili, B Waseta, A Malhotra.

University of California, Los Angeles, CA: MF Saad*, SB Nicholas*, Y-D I Chen**, X Guo**, J Rotter**, K Taylor**, M Budgett, F Hariri#.

University of New Mexico, Albuquerque, NM: P Zager*, V Shah**, M Scavini**, A Bobelu#.

University of Texas Health Science Center at San Antonio, San Antonio, TX: H Abboud*, N Arar**, R Duggirala**, BS Kasinath**, F Thameem**, M Stern**.

Wake-Forest University, Winston-Salem, NC: BI Freedman*†, DW Bowden**, CD Langefeld**, SC Satko**, SS Rich**, S Warren#, S Viverette, G Brooks, R Young, M Spainhour.

Laboratory of Genomic Diversity, National Cancer Institute, Frederick MD: C Winkler*, MW Smith**, M Thompson, R Hanson#, B Kessing.

Minority Recruitment Centers:

Loyola University Chicago — DJ Leehey*, G Barone#

University of Alabama Birmingham — D Thornley-Brown*, C Jefferson#

University of Chicago — OF Kohn*, CS Brown#

National Institute of Diabetes and Digestive and Kidney Diseases program office: JP Briggs, PL Kimmel, R Rasooly

External Advisory Committee: D Warnock (chair), L Cardon, R Chakraborty, GM Dunston, T Hostetter, SJ O'Brien (ad hoc), J Rioux, R Spielman.

* Principal Investigator

** Co-investigator

Program Coordinator

§ University of California, Davis, CA

† University of California, Irvine, CA

‡ Study Chair

References

1. McCarthy MI. Growing evidence for diabetes susceptibility genes from genome scan data. *Curr Diab Rep.* 2003; 3:159–167. [PubMed: 12728642]
2. Frayling TM. Genome-wide association studies provide new insights into type 2 diabetes aetiology. *Nat Rev Genet.* 2007; 8:657–662. [PubMed: 17703236]
3. McCarthy MI, Zeggini E. Genome-wide association scans for Type 2 diabetes: new insights into biology and therapy. *Trends Pharmacol Sci.* 2007; 28:598–601. [PubMed: 17997168]
4. Stern MP. The search for type 2 diabetes susceptibility genes using whole-genome scans: an epidemiologist's perspective. *Diabetes Metab Res Rev.* 2002; 18:106–113. [PubMed: 11994901]
5. Knowler WC, Coresh J, Elston RC, Freedman BI, Iyengar SK, Kimmel PL, Olson JM, Plaetke R, Sedor JR, Seldin MF. The Family Investigation of Nephropathy and Diabetes (FIND): design and methods. *J Diabetes Complications.* 2005; 19:1–9. [PubMed: 15642484]
6. Report of the Expert Committee on the Diagnosis and Classification of Diabetes Mellitus. *Diabetes Care.* 1997; 20:1183–1197. [PubMed: 9203460]
7. Abecasis GR, Cherny SS, Cookson WO, Cardon LR. Merlin--rapid analysis of dense genetic maps using sparse gene flow trees. *Nat Genet.* 2002; 30:97–101. [PubMed: 11731797]
8. Kong A, Cox NJ. Allele-sharing models. LOD scores and accurate linkage tests. *Am J Hum Genet.* 1997; 61:1179–1188. [PubMed: 9345087]
9. Sawcer S, Jones HB, Judge D, Visser F, Compston A, Goodfellow PN, Clayton D. Empirical genomewide significance levels established by whole genome simulations. *Genet Epidemiol.* 1997; 14:223–229. [PubMed: 9181352]
10. Fisher, RA. *Statistical Methods for Research Workers.* 13th Edition. Oliver & Loyd; London: 1925.
11. Lander E, Kruglyak L. Genetic dissection of complex traits: guidelines for interpreting and reporting linkage results. *Nat Genet.* 1995; 11:241–247. [PubMed: 7581446]
12. Abecasis GR, Cookson WO, Cardon LR. Pedigree tests of transmission disequilibrium. *Eur J Hum Genet.* 2000; 8:545–551. [PubMed: 10909856]
13. Zeggini E, Scott LJ, Saxena R, Voight BF, Marchini JL, Hu T, de Bakker PI, Abecasis GR, Almgren P, Andersen G, Ardlie K, Bostrom KB, Bergman RN, Bonnycastle LL, Borch-Johnsen K, Burt NP, Chen H, Chines PS, Daly MJ, Deodhar P, Ding CJ, Doney AS, Duren WL, Elliott KS, Erdos MR, Frayling TM, Freathy RM, Gianniny L, Grallert H, Grarup N, Groves CJ, Guiducci C, Hansen T, Herder C, Hitman GA, Hughes TE, Isomaa B, Jackson AU, Jorgensen T, Kong A, Kubalanza K, Kuruvilla FG, Kuusisto J, Langenberg C, Lango H, Lauritzen T, Li Y, Lindgren CM, Lyssenko V, Marvelle AF, Meisinger C, Midthjell K, Mohlke KL, Morken MA, Morris AD, Narisu N, Nilsson P, Owen KR, Palmer CN, Payne F, Perry JR, Pettersen E, Platou C, Prokopenko I, Qi L, Qin L, Rayner NW, Rees M, Roix JJ, Sandbaek A, Shields B, Sjogren M,

- Steinthorsdottir V, Stringham HM, Swift AJ, Thorleifsson G, Thorsteinsdottir U, Timpson NJ, Tuomi T, Tuomilehto J, Walker M, Watanabe RM, Weedon MN, Willer CJ, Illig T, Hveem K, Hu FB, Laakso M, Stefansson K, Pedersen O, Wareham NJ, Barroso I, Hattersley AT, Collins FS, Groop L, McCarthy MI, Boehnke M, Altshuler D. Meta-analysis of genome-wide association data and large-scale replication identifies additional susceptibility loci for type 2 diabetes. *Nat Genet.* 2008; 40:638–645. [PubMed: 18372903]
14. Iyengar SK, Freedman BI. The FIND Consortium. A multi-ethnic genome-wide linkage study for diabetic nephropathy. The Family Investigation of Nephropathy and Diabetes (FIND) study. *J Am Soc Nephrol.* 2007; 18:608A. (Abstract).
 15. Iyengar SK, Abboud HE, Goddard KA, Saad MF, Adler SG, Arar NH, Bowden DW, Duggirala R, Elston RC, Hanson RL, Ipp E, Kao WH, Kimmel PL, Klag MJ, Knowler WC, Meoni LA, Nelson RG, Nicholas SB, Pahl MV, Parekh RS, Quade SR, Rich SS, Rotter JI, Scavini M, Schelling JR, Sedor JR, Sehgal AR, Shah VO, Smith MW, Taylor KD, Winkler CA, Zager PG, Freedman BI. Genome-wide scans for diabetic nephropathy and albuminuria in multiethnic populations: the family investigation of nephropathy and diabetes (FIND). *Diabetes.* 2007; 56:1577–1585. [PubMed: 17363742]
 16. Knowler WC, Ipp E. The FIND Research Group. Diabetes types and autoimmunity in the Family Investigation of Nephropathy and Diabetes. *J Am Soc Nephrol.* 2005; 16:149A. (Abstract).
 17. Parker A, Meyer J, Lewitzky S, Rennich JS, Chan G, Thomas JD, Orho-Melander M, Lehtovirta M, Forsblom C, Hyrkkö A, Carlsson M, Lindgren C, Groop LC. A gene conferring susceptibility to type 2 diabetes in conjunction with obesity is located on chromosome 18p11. *Diabetes.* 2001; 50:675–680. [PubMed: 11246890]
 18. Luo TH, Zhao Y, Li G, Yuan WT, Zhao JJ, Chen JL, Huang W, Luo M. A genome-wide search for type II diabetes susceptibility genes in Chinese Hans. *Diabetologia.* 2001; 44:501–506. [PubMed: 11357482]
 19. Schisler JC, Jensen PB, Taylor DG, Becker TC, Knop FK, Takekawa S, German M, Weir GC, Lu D, Mirmira RG, Newgard CB. The Nkx6.1 homeodomain transcription factor suppresses glucagon expression and regulates glucose-stimulated insulin secretion in islet beta cells. *Proc Natl Acad Sci U.S.A.* 2005; 102:7297–7302. [PubMed: 15883383]
 20. Silver K, Tolea M, Wang J, Pollin TI, Yao F, Mitchell BD. The exon Cys7Gly polymorphism within the betacellulin gene is associated with type 2 diabetes in African Americans. *Diabetes.* 2005; 54:1179–1184. [PubMed: 15793259]
 21. Elbein SC, Wang X, Karim MA, Chu WS, Silver KD. Analysis of coding variants in the betacellulin gene in type 2 diabetes and insulin secretion in African American subjects. *BMC Med Genet.* 2006; 7:62. [PubMed: 16869959]
 22. Bektas A, Suprenant ME, Wogan LT, Plengvidhya N, Rich SS, Warram JH, Krolewski AS, Doria A. Evidence of a novel type 2 diabetes locus 50 cM centromeric to NIDDM2 on chromosome 12q. *Diabetes.* 1999; 48:2246–2251. [PubMed: 10535461]
 23. Bowden DW, Sale M, Howard TD, Qadri A, Spray BJ, Rothschild CB, Akots G, Rich SS, Freedman BI. Linkage of genetic markers on human chromosomes 20 and 12 to NIDDM in Caucasian sib pairs with a history of diabetic nephropathy. *Diabetes.* 1997; 46:882–886. [PubMed: 9133559]
 24. 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.* 2000; 66:1871–1881. [PubMed: 10793009]
 25. Mahtani MM, Widen E, Lehto M, Thomas J, McCarthy M, Brayer J, Bryant B, Chan G, Daly M, Forsblom C, Kanninen T, Kirby A, Kruglyak L, Munnely K, Parkkonen M, Reeve-Daly MP, Weaver A, Brettin T, Duyk G, Lander ES, Groop LC. Mapping of a gene for type 2 diabetes associated with an insulin secretion defect by a genome scan in Finnish families. *Nat Genet.* 1996; 14:90–94. [PubMed: 8782826]
 26. Shaw JT, Lovelock PK, Kesting JB, Cardinal J, Duffy D, Wainwright B, Cameron DP. Novel susceptibility gene for late-onset NIDDM is localized to human chromosome 12q. *Diabetes.* 1998; 47:1793–1796. [PubMed: 9792550]

27. Yamagata K, Oda N, Kaisaki PJ, Menzel S, Furuta H, Vaxillaire M, Southam L, Cox RD, Lathrop GM, Boriraj VV, Chen X, Cox NJ, Oda Y, Yano H, Le Beau MM, Yamada S, Nishigori H, Takeda J, Fajans SS, Hattersley AT, Iwasaki N, Hansen T, Pedersen O, Polonsky KS, Bell GI. Mutations in the hepatocyte nuclear factor-1alpha gene in maturity-onset diabetes of the young (MODY3). *Nature*. 1996; 384:455–458. [PubMed: 8945470]
28. 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*. 2004; 53:830–837. [PubMed: 14988270]
29. Elbein SC, Das SK, Hallman DM, Hanis CL, Hasstedt SJ. 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. *Diabetes*. 2009; 58:268–274. [PubMed: 18840782]
30. Hegele RA, Sun F, Harris SB, Anderson C, Hanley AJ, Zinman B. Genome-wide scanning for type 2 diabetes susceptibility in Canadian Oji-Cree, using 190 microsatellite markers. *J Hum Genet*. 1999; 44:10–14. [PubMed: 9929969]
31. Avery CL, Freedman BI, Heiss G, Kraja A, Rice T, Arnett D, Miller MB, Pankow JS, Lewis CE, Myers RH, Hunt SC, Almasy L, North KE. Linkage analysis of diabetes status among hypertensive families: the Hypertension Genetic Epidemiology Network study. *Diabetes*. 2004; 53:3307–3312. [PubMed: 15561964]
32. Pratley RE, Thompson DB, Prochazka M, Baier L, Mott D, Ravussin E, Sakul H, Ehm MG, Burns DK, Foroud T, Garvey WT, Hanson RL, Knowler WC, Bennett PH, Bogardus C. An autosomal genomic scan for loci linked to prediabetic phenotypes in Pima Indians. *J Clin Invest*. 1998; 101:1757–1764. [PubMed: 9541507]
33. Mennicken F, Hoffert C, Pelletier M, Ahmad S, O'Donnell D. Restricted distribution of galanin receptor 3 (GalR3) mRNA in the adult rat central nervous system. *J Chem Neuroanat*. 2002; 24:257–268. [PubMed: 12406501]
34. Beretta E, Dhillon H, Kalra PS, Kalra SP. Central LIF gene therapy suppresses food intake, body weight, serum leptin and insulin for extended periods. *Peptides*. 2002; 23:975–984. [PubMed: 12084530]
35. Lindgren CM, Mahtani MM, Widen E, McCarthy MI, Daly MJ, Kirby A, Reeve MP, Kruglyak L, Parker A, Meyer J, Almgren P, Lehto M, Kanninen T, Tuomi T, Groop LC, Lander ES. Genomewide search for type 2 diabetes mellitus susceptibility loci in Finnish families: the Botnia study. *Am J Hum Genet*. 2002; 70:509–516. [PubMed: 11791216]
36. Meigs JB, Panhuysen CI, Myers RH, Wilson PW, Cupples LA. A genome-wide scan for loci linked to plasma levels of glucose and HbA(1c) in a community-based sample of Caucasian pedigrees: The Framingham Offspring Study. *Diabetes*. 2002; 51:833–840. [PubMed: 11872688]
37. Frayling TM, Lindgren CM, Chevre JC, Menzel S, Wishart M, Benmezroua Y, Brown A, Evans JC, Rao PS, Dina C, Lecoeur C, Kanninen T, Almgren P, Bulman MP, Wang Y, Mills J, Wright-Pascoe R, Mahtani MM, Prisco F, Costa A, Cognet I, Hansen T, Pedersen O, Ellard S, Tuomi T, Groop LC, Froguel P, Hattersley AT, Vaxillaire M. A genome-wide scan in families with maturity-onset diabetes of the young: evidence for further genetic heterogeneity. *Diabetes*. 2003; 52:872–881. [PubMed: 12606533]
38. Stern MP, Duggirala R, Mitchell BD, Reinhart LJ, Shivakumar S, Shipman PA, Uresandi OC, Benavides E, Blangero J, O'Connell P. Evidence for linkage of regions on chromosomes 6 and 11 to plasma glucose concentrations in Mexican Americans. *Genome Res*. 1996; 6:724–734. [PubMed: 8858347]
39. Elbein SC, Hoffman MD, Teng K, Leppert MF, Hasstedt SJ. A genome-wide search for type 2 diabetes susceptibility genes in Utah Caucasians. *Diabetes*. 1999; 48:1175–1182. [PubMed: 10331426]
40. Hanson RL, Ehm MG, Pettitt DJ, Prochazka M, Thompson DB, Timberlake D, Foroud T, Kobes S, Baier L, Burns DK, Almasy L, Blangero J, Garvey WT, Bennett PH, Knowler WC. An autosomal genomic scan for loci linked to type II diabetes mellitus and body-mass index in Pima Indians. *Am J Hum Genet*. 1998; 63:1130–1138. [PubMed: 9758619]
41. Hsueh WC, Jean PL, Mitchell BD, Pollin TI, Knowler WC, Ehm MG, Bell CJ, Sakul H, Wagner MJ, Burns DK, Shuldiner AR. Genome-wide and fine-mapping linkage studies of type 2 diabetes

- and glucose traits in the Old Order Amish: evidence for a new diabetes locus on chromosome 14q11 and confirmation of a locus on chromosome 1q21-q24. *Diabetes*. 2003; 52:550–557. [PubMed: 12540634]
42. Vionnet N, Hani EH, Dupont S, Gallina S, Francke S, Dotte S, De Matos F, Durand E, Lepretre F, Lecoeur C, Gallina P, Zekiri L, Dina C, Froguel P. Genomewide search for type 2 diabetes-susceptibility genes in French whites: evidence for a novel susceptibility locus for early-onset diabetes on chromosome 3q27-qter and independent replication of a type 2-diabetes locus on chromosome 1q21-q24. *Am J Hum Genet*. 2000; 67:1470–1480. [PubMed: 11067779]
 43. Wiltshire S, Hattersley AT, Hitman GA, Walker M, Levy JC, Sampson M, O'rahilly S, Frayling TM, Bell JI, Lathrop GM, Bennett A, Dhillon R, Fletcher C, Groves CJ, Jones E, Prestwich P, Simecek N, Rao PV, Wishart M, Bottazzo GF, Foxon R, Howell S, Smedley D, Cardon LR, Menzel S, McCarthy MI. A genomewide scan for loci predisposing to type 2 diabetes in a U.K. population (the Diabetes UK Warren 2 Repository): analysis of 573 pedigrees provides independent replication of a susceptibility locus on chromosome 1q. *Am J Hum Genet*. 2001; 69:553–569. [PubMed: 11484155]
 44. Whitmore TE, Maurer MF, Day HL, Jelmberg AC, Dasovich MM, Sundborg LM, Burkhead SK, Heipel MD, Madden KL, Kramer JM, Kuijper JL, Xu WF, Jaspers SR, Holly RD, Lok S. The assignment of the human insulin receptor-related receptor gene (INSRR) to chromosome 1q21-->q23 by the use of radiation hybrid mapping. *Cytogenet Cell Genet*. 1999; 87:93–94. [PubMed: 10640820]
 45. Thameem F, Wolford JK, Bogardus C, Prochazka M. Analysis of PBX1 as a candidate gene for type 2 diabetes mellitus in Pima Indians. *Biochim Biophys Acta*. 2001; 1518:215–220. [PubMed: 11267683]
 46. Hayes MG, Pluzhnikov A, Miyake K, Sun Y, Ng MC, Roe CA, Below JE, Nicolae RI, Konkashbaev A, Bell GI, Cox NJ, Hanis CL. Identification of Type 2 Diabetes Genes in Mexican Americans Through Genome-wide Association Studies. *Diabetes*. 2007; 56:3033–3044. [PubMed: 17846124]
 47. Rampersaud E, Damcott CM, Fu M, Shen H, McArdle P, Shi X, Shelton J, Yin J, Chang CY, 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*. 2007; 56:3053–3062. [PubMed: 17846126]
 48. Saxena R, Voight BF, Lyssenko V, Burt 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 BK, 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, 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*. 2007; 316:1331–1336. [PubMed: 17463246]
 49. Wellcome Trust Case Control Consortium. Genome-wide association study of 14 000 cases of seven common diseases and 3 000 shared controls. *Nature*. 2007; 447:661–678. [PubMed: 17554300]
 50. Rodriguez S, Gaunt TR, Day IN. Molecular genetics of human growth hormone, insulin-like growth factors and their pathways in common disease. *Hum Genet*. 2007; 122:1–21. [PubMed: 17534663]

Figure 1.
Genome-wide linkage results for type 2 diabetes mellitus

Table 1Characteristics of FIND participants ^a

	African American	American Indian	European American	Mexican American
N	1004	883	537	1634
Number of families	362	214	202	481
Family size	2-8	2-20	2-7	2-13
Age at recruitment (years)	59.1 ±11.2	51.4 ±12.8	61.6 ±11.4	56.9 ±11.9
Gender (Male/Female)	339/665	346/537	252/285	696/938
^b BMI (kg/m ²)	32.8 ±8.2	32.9 ±8.3	31.3 ±7.8	30.9 ±7.0
Age at diabetes diagnosis (years)	40.6 ±14.0	37.0 ±12.6	42.7 ±15.3	41.4 ±12.9
Diabetes duration (years)	18.5 ±10.6	14.4 ±12.3	18.8 ±10.6	15.5 ±10.6

For each trait the mean ± standard deviation is shown

^a All participants have diabetes^b Body Mass Index (kilograms/meters²)

Table 2

Linkage analysis results for T2DM (multipoint LOD score >1.71)

Ethnic group	Chromosome	Distance (cM)	Closest marker(s)	Cytogenetic location	Multipoint LOD
African American	4	87.15	rs1566485	4q21.1	3.13
	12	111.16	rs1492254; rs2063508	12q21.31	2.02
	22	39.87	rs138777	22q12.3	2.38
American Indian	2	103.14	rs7583314	2p11.1	2.23
	7	31.35	rs1375237	7p15.1	1.78
	15	53.26	rs472579	15q21.1	1.79
	16	32.09	rs1001937	16p12.2	1.90
	21	31.43	rs145472	21q21.1	1.94
European American	6	58.53	rs1885615	6p12.3	2.77
Mexican American	7	6.3	rs798485	7p22.1	1.84
	13	56.32	rs641385	13q21.1	2.24

Table 3

Linkage analysis results for T2DM of the combined data using Fisher's test (multipoint LOD score >1.71)

Chromosome	Distance (cM)	Closest marker(s)	Cytogenetic location	Multipoint LOD
1	97.18	rs947420	1p22.1	2.32
4	87.15	rs1566485	4q21.1	1.71
6	66.19	rs1563788	6p12.1	2.88
6	93.77	rs1884831	6q14.1	1.81
6	108.31	rs2040431	6q16.1	1.87
12	111.35	rs2063508	12q21.31	2.01
16	32.09	rs1001937	16p12.2	2.13
22	42.21	rs933224; rs2413411	22q13.1	2.32