

Genome-Wide Linkage Analyses of Type 2 Diabetes in Mexican Americans

The San Antonio Family Diabetes/Gallbladder Study

Kelly J. Hunt,¹ Donna M. Lehman,¹ Rector Arya,¹ Sharon Fowler,¹ Robin J. Leach,² Harald H.H. Göring,³ Laura Almasy,³ John Blangero,³ Tom D. Dyer,³ Ravindranath Duggirala,³ and Michael P. Stern¹

The San Antonio Family Diabetes/Gallbladder Study was initiated to identify susceptibility genes for type 2 diabetes. Evidence was previously reported of linkage to diabetes on 10q with suggestive evidence on 3p and 9p in a genome-wide scan of 440 individuals from 27 pedigrees ascertained through a single diabetic proband. Subsequently, the study was expanded to include 906 individuals from 39 extended Mexican-American pedigrees, two additional examination cycles ~5.3 and 7.6 years after baseline, and genotypes for a new set of genome-wide markers. Therefore, we completed a second genome-wide linkage scan. Using information from a participant's most recent exam, the prevalence of diabetes in nonprobands was 21.8%. We performed genome-wide variance components-based genetic analysis on the discrete trait diabetes using a liability model and on the quantitative Martingale residual obtained from modeling age of diabetes diagnosis using Cox proportional hazard models. Controlling for age and age², our strongest evidence for linkage to the trait diabetes and the quantitative Martingale residual was on chromosome 3p at marker D3S2406 with multipoint empirical logarithm of odds scores of 1.87 and 3.76, respectively. In summary, we report evidence for linkage to diabetes on chromosome 3p in a region previously identified in at least three independent populations. *Diabetes* 54: 2655–2662, 2005

From the ¹Division of Clinical Epidemiology, Department of Medicine, University of Texas Health Science Center, San Antonio, Texas; the ²Department of Cellular and Structural Biology, University of Texas Health Science Center, San Antonio, Texas; and the ³Department of Genetics, Southwest Foundation for Biomedical Research, San Antonio, Texas.

Address correspondence and reprint requests to Kelly J. Hunt, PhD, Division of Clinical Epidemiology, Department of Medicine, University of Texas Health Science Center at San Antonio, 7703 Floyd Curl Dr., San Antonio, TX 78229-3900. E-mail: huntk@uthscsa.edu.

Received for publication 14 January 2005 and accepted in revised form 3 June 2005.

CIDR, Center for Inherited Disease Research; LOD, logarithm of odds; SAFDGS, San Antonio Family Diabetes/Gallbladder Study; WHO, World Health Organization.

© 2005 by the American Diabetes Association.

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.

Type 2 diabetes, a common multifactorial metabolic disease, is caused by both environmental and genetic factors. The incidence of type 2 diabetes continues to rise and increasingly affects individuals of all ages across all ethnic groups (1,2). Individuals from certain ethnic groups including Mexican Americans have an increased propensity toward developing type 2 diabetes. Furthermore, not only is the risk of diabetes two- to threefold higher in Mexican Americans than in non-Hispanic whites, but diabetes is more severe and more deadly in Mexican Americans than in non-Hispanic whites (3,4).

Genetic studies consistently indicate that diabetes is familial in nature. However, despite >20 genome-wide linkage analyses of type 2 diabetes, the genes influencing susceptibility to the common forms of type 2 diabetes remain largely unknown (5–21). Moreover, the increased risk and severity of type 2 diabetes in Mexican Americans may indicate an increased genetic susceptibility (22,23). Previously, in the San Antonio Family Diabetes Study, an extended pedigree study composed of Mexican-American families ascertained through a single proband with diabetes, Duggirala et al. (24) reported evidence of linkage to diabetes on 10q and suggestive evidence of linkage to diabetes on 3p and 9p. This earlier study was based on a single clinical examination of 440 individuals from 27 extended pedigrees with genotypic information from 379 microsatellite markers across the genome (24). Subsequently, the study has been expanded to include two additional examination cycles during which previous participants were invited to return, existing pedigrees were expanded, and new pedigrees were added. In addition, the Center for Inherited Disease Research (CIDR) has recently completed genotyping of 382 highly polymorphic markers distributed throughout the autosomes at ~10-cM intervals on over 96% of the 906 individuals who have participated in at least one clinical examination. Therefore, our objective was to perform a genome-wide linkage analysis of type 2 diabetes in the updated San Antonio Family Diabetes/Gallbladder Study (SAFDGS).

RESEARCH DESIGN AND METHODS

The SAFDGS design and population. The baseline SAFDGS population consisted of 579 individuals distributed across 31 extended Mexican-American families recruited and first examined between 1991 and 1994. Proband in the SAFDGS were low-income Mexican-American men and women identified as having diabetes in a prior population-based epidemiological study, the San Antonio Heart Study (25). In addition to having diabetes, the proband was required to have at least six first-degree relatives, excluding parents, who were at least 18 years old and living in the San Antonio area. First-, second-, and third-degree relatives of the proband were invited to participate.

Of the original 579 participants, 500 individuals returned for the first follow-up examination of the SAFDGS population that occurred between 1996 and 1999. In addition, 67 family members who were unable to participate in the baseline SAFDGS were recruited and participated in the first recall of the SAFDGS. Finally, between 1998 and 2001, a second follow-up examination occurred (known as the San Antonio Family Gallbladder Study), during which 265 new participants were recruited including both additional family members of the 31 original families as well as 8 additional families. In summary, there are 911 SAFDGS participants from 39 families who have participated in at least one of three clinic examinations. Of these 911 participants, 5 people have been excluded either because of chromosomal abnormalities or because of irresolvable pedigree uncertainties; hence, there are a total of 906 SAFDGS participants with phenotypic information available. The institutional review board of the University of Texas Health Science Center at San Antonio approved the study, and all subjects gave informed consent.

SAFDGS baseline and follow-up examinations. Each examination consisted of a standardized medical examination that included interviews, anthropometry, a fasting venipuncture, and an oral glucose tolerance test. Trained interviewers obtained information on medical history and medication use. Examinations occurred in the morning, after participants had fasted the 12 h before their examination. Measurement of BMI, waist circumference, fasting plasma glucose, and plasma glucose 2 h after a standardized oral glucose load have been previously described in detail (26).

Two definitions of diabetes were used. The first, referred to throughout this study as diabetes (fasting), required a fasting plasma glucose ≥ 7.0 mmol/l (126 mg/dl) (27), whereas the second, referred to as diabetes (World Health Organization [WHO]), required either a fasting plasma glucose ≥ 7.0 mmol/l (126 mg/dl) or a 2-h glucose after an oral glucose tolerance test ≥ 11.1 mmol/l (200 mg/dl) (28). For each definition, participants who did not meet these criteria but who self-reported physician-diagnosed diabetes and who reported current therapy with either oral antidiabetic agents or insulin were also considered to have diabetes in this investigation. For each definition, information from a participant's most recent examination was used to define his or her diabetes status with the caveat that individuals diagnosed with diabetes who when later examined no longer had diabetes were considered missing. Finally, age of diabetes diagnosis was modeled as a proxy for age of diabetes onset: for previously diagnosed diabetic participants, self-reported age of diagnosis was used as the age of onset; for diabetic participants initially diagnosed at a SAFDGS examination, the participant's reported age at that examination was used as the age of onset.

Genotyping and calculation of multipoint identities by descent. A complete genomic scan using 382 highly polymorphic markers distributed throughout the autosomes at ~ 10 -cM intervals has been completed by the CIDR on 872 of 906 participants. In 39 families, genotype data were available across five generations in 5 families, four generations in 19 families, three generations in 11 families, and two generations in 4 families. Genotype data were cleaned for both Mendelian and spurious double-recombinant errors using SimWalk2 (29). SimWalk2 provides a probability of error for each individual for each genotype based on the genotypes of other family members, the allele frequencies, and the marker map. For Mendelian error cleaning, genotypes were blanked at descending error probability thresholds, in increments of 0.1, until no more inconsistencies existed, resulting in the elimination of 1,247 of 364,934 genotypes, a blanking rate of 0.34%. For double recombinant cleaning, all genotypes with an error probability of ≥ 0.25 were blanked, resulting in the elimination of an additional 436 genotypes, a blanking rate of 0.12%. MultiMap/CRI-MAP (30,31) was used to construct sex-averaged marker maps using the cleaned genotype data. Allele frequencies were estimated by maximum likelihood methods implemented in SOLAR (sequential oligogenic linkage analysis routines) (32), and matrices of multipoint identity-by-descent sharing probabilities were estimated using Markov chain Monte Carlo methods implemented by LOKI (33).

In addition, in instances in which a statistically significant linkage signal was identified, additional microsatellite markers were typed in the 705 individuals for whom we have large quantities of DNA because of the lymphoblastoid cell lines available for these individuals. The locations of the highly polymorphic microsatellite markers used were determined from the

UCSC Human Genome Browser July 2003 assembly. PCR primer sequences for each marker were obtained from the genome database. Forward primers were labeled with fluorescent dyes. Multiplex PCR was performed using 240 ng of high-molecular weight genomic DNA isolated from lymphoblastoid cell lines. The resulting PCR products were run on an ABI 3100 Avant Genetic Analyzer, and alleles were called using the GeneMapper analysis software.

Statistical analyses. Using a variance decomposition approach implemented in SOLAR (32), we performed genetic analysis on the discrete trait diabetes, using a liability threshold model (6). This approach assumes that an individual belongs to a specific disease class if an underlying genetically determined risk or liability exceeds a certain threshold on a normally distributed liability curve. The liability is assumed to have an underlying multivariate normal distribution. The correlation in liability between individuals i and j is given by:

$$\rho_{ij} = \Pi_{mij}h_m^2 + 2\Phi_{ij}h^2 + I_{ij}e^2$$

where ρ_{ij} is the correlation in liability to disease between individuals i and j ; Π_{mij} is the proportion of alleles that individuals i and j share identical by descent at a marker, m , linked to a quantitative trait locus; h_m^2 is the heritability attributed to the quantitative trait locus near the marker locus; Φ_{ij} is the kinship coefficient for individuals i and j ; h^2 is the heritability attributed to residual additive polygenic effects; I_{ij} is an identity matrix (which equals 1 if i equals j and 0 if i does not equal j); and e^2 is equal to $1 - (h_m^2 + h^2)$. Maximum likelihood techniques were used to estimate variance components and covariate effects simultaneously.

In addition to performing genetic analysis on the discrete trait diabetes, we used SAS to model age of diabetes diagnosis as a proxy for age of diabetes onset with a Cox proportional hazards model (34). In the Cox proportional hazards models, self-reported age of diagnosis was used as the time of the event for previously diagnosed diabetic participants; for diabetic participants initially diagnosed at a SAFDGS examination, the participant's reported age at that examination was used as the time of the event; and finally, nondiabetic participants were censored at their most recent SAFDGS examination age. In addition, participants were left censored at their age 1 day before they entered the study or, in the case of an individual who entered with diabetes, 1 day before their age of reported diabetes onset. Moreover, the left censoring age (with a quadratic term) was adjusted for in the Cox proportional hazards model. Finally, using SOLAR, we performed standard multipoint variance components linkage analysis on the Martingale residual from the Cox proportional hazards model, a quantitative trait (34). In summary, Martingale residuals obtained from Cox proportional hazards models were used to model age of diabetes onset while accounting for age of entry into the study. Additional covariates of interest were incorporated into the SOLAR model.

Available covariates considered in the analyses included age (with a quadratic term) at a participant's most recent examination, sex, decade of birth (to account for secular trends in the prevalence of diabetes), and waist circumference at a participant's most recent examination. In each case, a P value of 0.05 was used as a nominal value for retention in the model. Finally, because it is well known that variance components-based linkage analyses can have elevated false-positive rates in analysis of traits that are not multivariate-normally distributed (35,36), for each multipoint linkage analysis completed on the discrete trait diabetes as well as the Martingale residuals for age of diabetes diagnoses, we have used a simulation-based approach implemented in SOLAR to adjust the nominal logarithm of odds (LOD) score (35). The LOD score distribution under the null hypothesis of no linkage was determined based on 10,000 replicates assuming a fully informative marker locus, and these empirical LOD scores were regressed on the theoretical LOD score distribution to obtain a correction constant, which was subsequently used to adjust the observed LOD score. Specifically, a large number of fully informative markers were simulated ($n = 10,000$) that were not linked to the trait of interest. For this marker, identity-by-descent information was calculated, and linkage analysis was conducted. The correction constant was obtained by comparing the distribution of the observed simulated LOD scores with the theoretical distribution and was used to adjust the observed LOD scores to empirical LOD scores. Throughout the study, empirical LOD scores are presented, because these are more reliable than conventional LOD scores. Chromosomal regions with an empirical LOD score >3 were considered to provide significant evidence of linkage; regions with an empirical LOD score >2 were considered to provide suggestive evidence for linkage, and regions with an empirical LOD score of 1.18 ($P < 0.01$) were considered "potentially interesting" as previously described by Vionnet et al. (37).

RESULTS

A total of 906 individuals, aged 17–97 years, were examined. Individuals from 39 families and 32 unrelated individuals (i.e., individuals without offspring but married into

TABLE 1
Characteristics of individuals with and without type 2 diabetes (fasting) in the SAFDGS

	Proband	Family members	
<i>n</i>	39	666	186
Diabetes status (fasting)	Yes	No	Yes
Male (%)	38.5	40.7	41.9
Diabetes (WHO) (%)	100	3.8	100
Age (years)	63.2	40.6	55.9
Birth year	1935	1958	1942
Age of diabetes onset (years)	50.7	—	47.5
Fasting glucose (mg/dl)	171	95	184
2-h glucose (mg/dl)	367	118	325
BMI (kg/m ²)	31.3	30.1	32.5
Waist circumference (mm)	1,059	974	1,067

Data are means or %.

the 39 families) were included. The unrelated individuals were included to help evaluate covariate effects. The median number of members per pedigree for the 39 families was 17 and ranged from 2 to 64. Of the 906 participants, 10 were excluded from the analyses for diabetes (fasting) because they were initially diagnosed but at a subsequent examination failed to meet the criteria, whereas an additional 5 were excluded because of missing information (corresponding numbers for diabetes [WHO] were 13 and 8, respectively). Diabetes (fasting) was identified in 108 of 503 nonproband women and 78 of 349 nonproband men, whereas diabetes (WHO) was identified in 125 of 499 nonproband women and 90 of 347 nonproband men. For the trait diabetes (fasting) excluding parent-offspring pairs (which are uninformative for linkage), there are 9,505 relative pairs, with varying degrees of relatedness. Of these 9,505 relative pairs, 438 are affected-affected pairs, 5,913 are unaffected-unaffected pairs, and 3,154 are affected-unaffected pairs. Table 1 includes characteristics of the study participants with and without diabetes (fasting). Overall, point estimates indicate participants with diabetes (fasting) were older and had a higher BMI and a larger waist circumference (Table 1).

The age- and age²-adjusted heritability ($h^2 \pm SE$) for

diabetes (fasting) was 0.52 ± 0.121 ($P < 0.0001$), whereas the heritability for the Martingale residual (fasting) was 0.18 ± 0.050 ($P < 0.0001$). The Martingale residual for diabetes (fasting) age of onset meets the prerequisites of the variance component method used (all residuals within 4 SD; kurtosis, -0.49 ; residual kurtosis, -0.48 ; and skewness, 0.53). The results of the multipoint genome-wide linkage scans of the trait diabetes (fasting), adjusted for age and age², and the Martingale residual (fasting) are shown in Fig. 1.

The results of the multipoint genome-wide linkage scan of the trait diabetes (fasting) adjusted for age and age² revealed several potentially interesting chromosomal regions: the first on chromosome 3 and the second on chromosome 15. Moreover, the results of the multipoint genome-wide linkage scan of the Martingale residual (fasting) revealed a single statistically significant empirical LOD of 3.76 ($P = 0.00002$) on chromosome 3p near marker D3S2406, in the same region implicated in the analysis of the discrete trait diabetes (Fig. 2). Additional areas identified as potentially interesting are found in Table 2. Moreover, birth decade and waist circumference were covariates significantly associated with the trait diabetes (fasting) as well as the Martingale residual (fasting) modeling age of diabetes onset. Empirical LODs corresponding to the region on chromosome 3p identified as potentially interesting were slightly attenuated after further adjustment for birth decade (trait diabetes, empirical LOD = 1.93; Martingale residuals, empirical LOD = 3.55) and somewhat attenuated after further adjustment for both birth decade and waist circumference (trait diabetes, empirical LOD = 1.29; Martingale residuals, empirical LOD = 2.68).

The age- and age²-adjusted heritability ($h^2 \pm SE$) for diabetes (WHO) was 0.56 ± 0.112 ($P < 0.0001$), whereas the heritability for the Martingale residual (WHO) was 0.23 ± 0.051 ($P < 0.0001$). The Martingale residual for diabetes (WHO) age of onset meets the prerequisites of the variance component method used (within 4 SD; kurtosis, -0.69 ; residual kurtosis, -0.68 ; and skewness, 0.38). The results of the multipoint genome-wide linkage scans of the

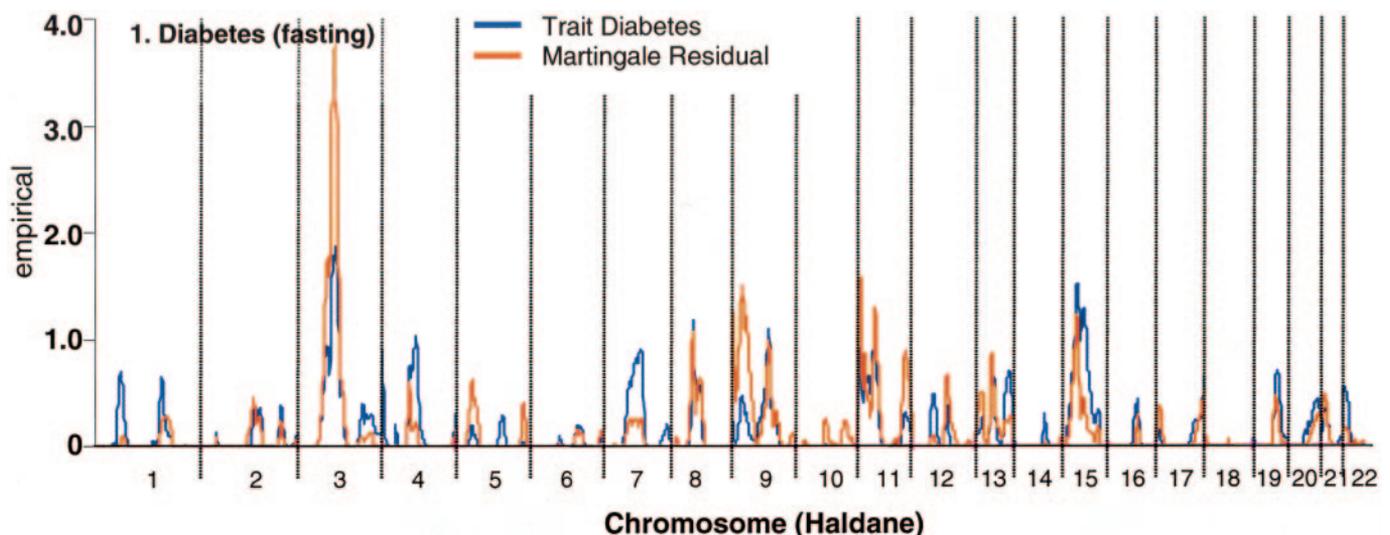


FIG. 1. Multipoint genome-wide linkage scan of the trait type 2 diabetes (fasting, blue, liability to disease) with adjustment for age and age² and the Martingale residual modeling age of diabetes onset (orange).

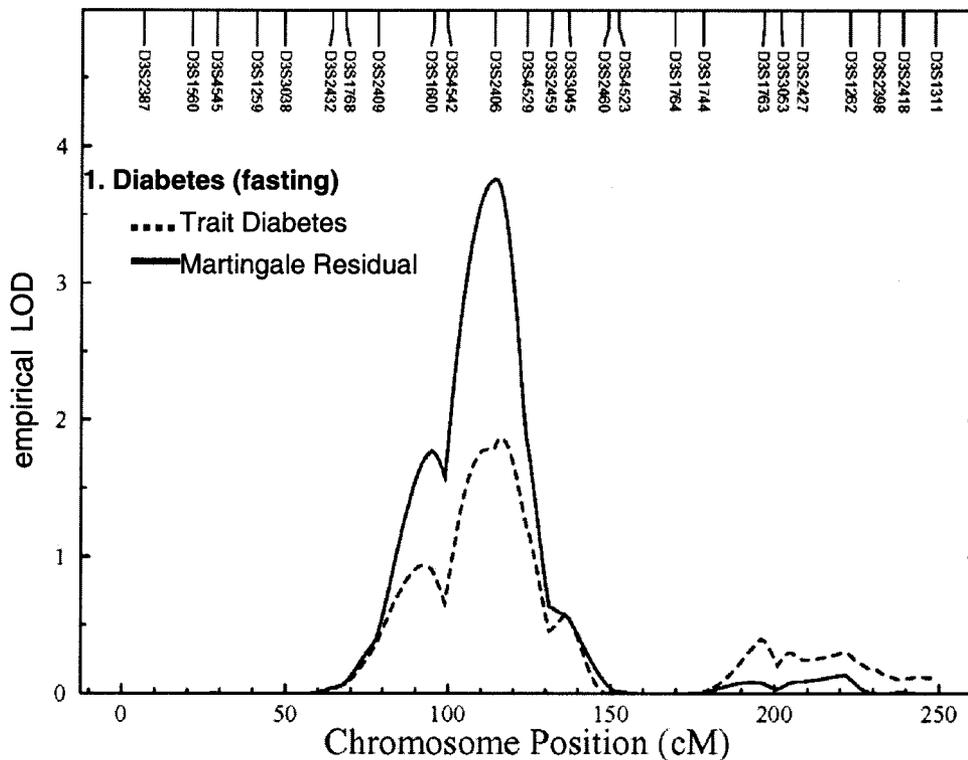


FIG. 2. Chromosome 3 multipoint linkage of the trait type 2 diabetes (fasting, dashed line, liability to disease) with adjustment for age and age² and the Martingale residual modeling age of diabetes onset (solid line).

trait as well as the Martingale residual for diabetes (WHO) were similar to those for diabetes (fasting) and are reported in Table 2. Moreover, empirical LODs corresponding to the region on chromosome 3p previously identified were slightly attenuated after further adjustment for birth decade (trait diabetes, empirical LOD = 2.07; Martingale residuals, empirical LOD = 3.03) and somewhat attenuated after further adjustment for both birth decade and

TABLE 2
Multipoint empirical LODs of ≥ 1.18 for the trait diabetes adjusted for age and age² and the Martingale residual modeling age of diabetes onset

Chromosome	cM (Haldane)	Nearest marker	Empirical LOD
Diabetes (fasting)			
3	116	D3S2406	1.87
15	47	D15S659	1.51
Martingale residual (fasting)			
3	114	D3S2406	3.76
9	28	GATA187D09	1.49
11	16	D11S1999	1.57
15	45	D15S659	1.22
Diabetes (WHO)			
3	117	D3S2406	1.90
4	114	D4S1647	1.43
11	61	D11S1993	2.25
15	43	D15S659	1.47
Martingale residual (WHO)			
3	115	D3S2406	3.29
9	40	D9S925	1.83
11	61	D11S1993	2.25
15	45	D15S659	1.42

Data are $P < 0.01$.

waist circumference (trait diabetes, empirical LOD = 1.40; Martingale residuals, empirical LOD = 2.15).

In addition to multipoint linkage analysis, two-point linkage analysis with each marker was completed for diabetes (fasting and WHO) adjusted only for age and age² as well as the (fasting and WHO) Martingale residual traits (Table 3). Furthermore, to follow up the linkage signal on 3p, we typed 10 additional microsatellite markers in the 705 individuals for whom we have ample DNA through lymphoblastoid cell lines. The markers encompassed a region of 3p from 49.2 to 81.9 Mb. Two-point empirical LODs for the 10 newly typed microsatellite markers on chromosome 3 in the region of interest are found in Table 3. Moreover, multipoint analysis including these additional markers resulted in a linkage profile that was similar to the original but slightly narrowed. The peak linkage signal for the (fasting) Martingale residual trait was at 112 cM (empirical LOD of 3.55) using the new map.

DISCUSSION

In summary, our strongest evidence for linkage to diabetes occurs on chromosome 3p near marker D3S2406 (103 cM, Marshfield map). This signal is present regardless of whether the diabetes definition includes or excludes information on 2-h glucose levels and is strengthened by accounting for age of diabetes onset.

For the diabetes (fasting) age of onset model, the 1.5-LOD support interval (~95% CI) around our peak spanned ~20 cM from marker D3S4542 to marker D3S4529 before typing the 10 additional microsatellite markers. However, with the 10 additional microsatellite markers, the 1.5-LOD support interval was narrowed and spans an ~10-cM region from marker D3S1296 to marker D3S3633 that harbors ~40 identified (known and predicted) genes.

The San Antonio Family Diabetes Study has previously

TABLE 3

Two-point empirical LODs of ≥ 1.18 , empirical LODs for markers in 3p region of interest, and empirical LODs for 10 newly typed microsatellite markers on chromosome 3 in the region of interest

Marker	cM (Haldane)	Fasting		WHO	
		Diabetes*	MR	Diabetes*	MR
Chromosome 3p					
D3S2409	78	0.27	0.28	0.07	0.10
D3S3561†	78	0.17	0.43	0.33	0.58
D3S1578†	78	0.65	0.99	0.48	0.61
D3S3721†	82	0.33	0.15	0.43	0.19
D3S1592†	84	1.07	0.83	1.27	0.91
D3S3631†	89	0.27	0.29	0.10	0.09
D3S1600	93	0.11	0.39	0.01	0.24
D3S4542	97	0.32	0.37	0.12	0.15
D3S3697†	103	0.40	1.00	0.28	0.87
D3S1296†	106	0.29	0.92	0.10	0.39
D3S3551†	107	1.42	3.30	1.02	2.14
D3S2406	111	1.54	4.11	1.79	3.81
D3S3633†	117	0.38	0.83	0.95	0.98
D3S1542†	117	0.96	1.33	1.60	1.39
D3S4529	120	0.74	1.62	1.10	1.58
D3S2459	128	0.13	0.28	0.01	0.00
Other locus					
D1S1619	301	1.07	0.41	1.33	0.53
D4S2361	102	1.16	0.62	1.68	0.64
D4S1652	231	0.59	0.64	1.44	0.97
D5S1470	49	0.10	0.75	0.52	1.58
D8S1477	65	1.38	1.16	1.41	0.46
GATA187D09	24	0.57	1.48	0.00	0.55
D11S1999	18	0.62	1.31	0.80	1.37
D11S1344	63	0.33	0.57	2.03	2.52
D11S912	145	0.24	1.30	0.63	1.51
GATA91H06	62	0.66	0.73	1.05	1.25
D12S1064	105	0.39	1.41	0.33	1.87

Bold values are $P < 0.01$. MR, Martingale residual trait. *Adjusted for age and age². †Newly typed microsatellite markers.

reported evidence for linkage to type 2 diabetes and nominal evidence for linkage to fasting-specific insulin in the 3p region implicated in this expanded SAFDGS. Specifically, Duggirala et al. (24) in 1999 reported a multipoint LOD of 2.56 for diabetes near marker GATA128C02 (Marshfield map, 112 cM) in 440 San Antonio Family Diabetes study participants who were examined at baseline and in 2001 reported a multipoint LOD ≥ 1.00 for fasting-specific insulin in 310 nondiabetic San Antonio Family Diabetes study participants at the same locus (38). Moreover, the San Antonio Family Heart Study comprised of a similar study population reports a LOD for serum insulin in 391 nondiabetic individuals of 3.07 on chromosome 3p in the region flanked by D3S1600 (86.0 cM, Marshfield map) and D3S1285 (91.0 cM, Marshfield map) (39). Similar to the SAFDGS, the San Antonio Family Heart study is an extended pedigree study of low-income Mexican Americans in San Antonio, Texas. Differences between the two studies include ascertainment without regard to health status of the proband in the San Antonio Family Heart Study as well as a younger mean age at baseline in the San Antonio Family Heart Study.

Moreover, previously in the San Antonio Family Diabetes Study, Duggirala et al. (24) not only reported suggestive evidence of linkage to diabetes on 3p but reported significant evidence of linkage on 10q as well as suggestive

evidence of linkage on 9p. Contrary to the earlier findings concerning 10q that have been replicated in a number of other study populations examining diabetes and related phenotypes (16,37,40–42), in the expanded SAFDGS, we do not find evidence of linkage to diabetes on 10q; however, we do replicate the earlier findings on both 3p and 9p. Explanations for discrepancies between previous and current findings on 10q include 1) changes in how diabetes is defined (i.e., the new fasting glucose threshold, 126 vs. 140 mg/dl), 2) changes in the study population (i.e., new pedigrees have been added and old pedigrees have been expanded) that could introduce further heterogeneity, 3) collection of phenotypic information at three versus a single time point and the subsequent aging of the study population, 4) the new set of microsatellite markers, and 5) secular trends in the prevalence of diabetes (1). Therefore, to help identify which explanation was most plausible, we completed a genome-wide linkage analysis on the 579 individuals at the San Antonio Family Diabetes study baseline examination using the new definitions of diabetes and the new genotypic information from CIDR. Results from this analysis replicate the previous 10q findings (i.e., Martingale residual [fasting], LOD = 3.00; Martingale residual [WHO], LOD = 2.57). Hence, neither the new definition of the phenotype nor the new marker set accounts for the loss of the 10q signal. Its disappearance, therefore, must relate to the expanded population, aging of our population, or secular trends in the incidence of diabetes. We attempted to control for the latter by controlling for birth decade. Different genes are likely to influence diabetes susceptibility at different ages; hence, the aging of our population over the three visits may have influenced the observed linkage signals. Moreover, because the gene pool is relatively stable, the rapid increase in the prevalence of diabetes in recent years reflects environmental changes that may call into play additional diabetes susceptibility genes through gene-environment interaction.

Other studies from outside San Antonio, Texas, have also reported evidence for linkage to type 2 diabetes or type 2 diabetes-related traits on chromosome 3p in the region implicated in this study. Our 3p signal has previously been identified with type 2 diabetes in a Chinese population (maximum reported LOD of 2.27) (21), a family study focused on identifying genes associated with maturity-onset diabetes of the young (maximum reported heterogeneity LOD of 1.81) (8), and in an ordered subset analysis of a Japanese population in which families were rank ordered based on average maximal BMI of sib-pairs and maximal BMI was adjusted for (maximum reported LOD of 2.42) (17). Additionally, our 3p region has been identified as having slight evidence for linkage to type 2 diabetes in several studies (43,44), to fasting glucose in the Framingham Offspring study (16), and to acute insulin response in the Pima Indians (45). Finally, four additional studies have reported evidence of linkage to type 2 diabetes near our region, ~50 cM toward the p-terminal end of chromosome 3 (12,13,20,46).

Ideally, in an extended pedigree study of a discrete trait, complete information on who acquired the trait and their age of onset would be available. Therefore, when available, we used serial data to define diabetes status. This

enabled us to exclude individuals likely misclassified who were identified as having diabetes at one examination but who at a subsequent examination no longer met the criteria for having diabetes (~1% of the population for either definition). In addition, because we used information across three examinations that occurred over 10 years, we were able to identify a greater number of individuals as having diabetes (i.e., those who were initially nondiabetic but who developed diabetes at a later examination) as well as to determine more accurately their age of diabetes onset. Finally, by combining information across the three examinations, we maximized the number of individuals included in the analyses. One limitation of combining information collected over a 10-year period is that it tends to magnify the effect of secular trends. To help overcome this limitation, we adjusted for birth decade to help account for secular trends in the incidence of diabetes.

In comparison with modeling the discrete trait diabetes, which includes information only on whether or not an individual has diabetes (0.1), the Martingale residual can be viewed as a continuous measure that incorporates information based on a participant's age at entry into the study, whether they develop diabetes during the study, age of diabetes diagnosis if they develop diabetes, and age at the end of the study if they do not develop diabetes. For example, in our study, a participant that enters the study at age 21 and develops diabetes (fasting) at age 25 has a Martingale residual of 0.96, whereas a participant who enters the study at age 64 and is 72 when the study ends and still does not have diabetes has a Martingale residual of -0.71. Therefore, use of the Martingale residual allows us to model age of diabetes diagnosis because it assigns a value to individuals based on their age of diabetes diagnosis if they have diabetes or based on their obtained age without developing diabetes if they don't have diabetes. In summary, for each definition of diabetes (fasting and WHO), the LOD for the Martingale residual is quite a bit higher than the LOD for the trait; however, this is not unexpected because the Martingale residual is more precise, incorporating additional information.

We examined two definitions for diabetes because basing the definition on fasting only avoids the problems associated with 2-h glucose levels in the definition that increases the number of individuals with missing information on diabetes (due to missing 2-h values) and that has lower repeatability. However, it is reassuring that evidence for linkage to diabetes (fasting) is similar to evidence for linkage to diabetes (WHO). The slightly higher LODs for the Martingale residual for diabetes (fasting) relative to the Martingale residual for diabetes (WHO) (3.76 vs. 3.29 for the Martingale residual) may be explained by the lower percentage of people with missing data as well as the higher repeatability of the diabetes (fasting) definition. Moreover, although adjusting for birth decade to help account for secular trends in the incidence of diabetes only slightly attenuated the signals, additionally adjusting for waist circumference further reduced the LODs 24.5 and 29.0% for the diabetes age of onset model for fasting and WHO definitions, respectively. Interestingly, there is no evidence for linkage to waist circumference in the

SAFDGS in this region on chromosome 3 (K.J.H., unpublished data).

A number of genes in the region of linkage on chromosome 3p encode proteins that appear to have biological relevance to the pathophysiology of diabetes. Their function points to a role in either insulin secretion or insulin action. Succinyl-CoA synthetase (SUCLG2, MIM 603922) is expressed in the mitochondria and in the cytosol of pancreatic islet cells as well as numerous other tissues (47). Alterations in SUCLG2 could affect key functional steps in mitochondrial metabolism leading to insulin secretion (47–50). Forkhead box P1 (MIM 605515) is a novel member of a subfamily of winged helix/forkhead/HNF3 transcription factors (51). Mutations in two related members, FOXP3 and Foxo1 (mouse homolog of FOXO1), have been shown to be involved in the pathogenesis of diabetes (52,53). Glycogen-branching enzyme (MIM 607839) plays a role in glycogen synthesis and storage, a critical factor in energy homeostasis (54). Glycogen-branching enzyme is an attractive candidate because there is substantial evidence to suggest that impaired glycogen synthesis (non-oxidative glucose disposal) could be a primary metabolic defect in pre-diabetic individuals (55).

In summary, our genome-wide multipoint linkage scan for type 2 diabetes in a Mexican-American population found evidence for linkage on chromosome 3p. Moreover, this finding replicates prior linkage findings to type 2 diabetes and fasting-specific insulin in the San Antonio Family Diabetes Study, as well as replicates three independent studies that have reported evidence for linkage to type 2 diabetes in this region and a single independent study that has reported evidence for linkage to serum insulin in this region.

ACKNOWLEDGMENTS

This work was supported by National Institutes of Health Grants R01-DK-042273, R01-DK-047482, R01-DK-53889, and MH-59490. K.J.H. was partially supported by an American Diabetes Association Mentor-Based Award, an American Diabetes Association Research Award, and National Institute of Diabetes and Digestive and Kidney Diseases Grant K01-DK-064867. D.M.L. was supported by an American Diabetes Association Junior Faculty Award. CIDR is fully funded through a federal contract from the National Institutes of Health to The Johns Hopkins University (contract N01-HG-65403).

Genotyping services were provided by the CIDR. Yamin Ench, Korri Weldon, and Jeanette Hamlington provided excellent technical assistance. We thank the participants of the SAFDGS families for their support and cooperation.

REFERENCES

- Burke JP, Williams K, Gaskill SP, Hazuda HP, Haffner SM, Stern MP: Rapid rise in the incidence of type 2 diabetes from 1987 to 1996: results from the San Antonio Heart Study. *Arch Intern Med* 159:1450–1456, 1999
- Harris MI, Flegal KM, Cowie CC, Eberhardt MS, Goldstein DE, Little RR, Wiedmeyer HM, Byrd-Holt DD: Prevalence of diabetes, impaired fasting glucose, and impaired glucose tolerance in U.S. adults: the Third National Health and Nutrition Examination Survey, 1988–1994. *Diabetes Care* 21:518–524, 1998
- Hunt KJ, Williams K, Resendez RG, Hazuda HP, Haffner SM, Stern MP: All-cause and cardiovascular mortality among diabetic participants in the

- San Antonio Heart Study: evidence against the "Hispanic paradox." *Diabetes Care* 25:1557-1563, 2002
4. Haffner SM, Mitchell BD, Pugh JA, Stern MP, Kozlowski MK, Hazuda HP, Patterson JK, Klein R: Proteinuria in Mexican Americans and non-Hispanic whites with NIDDM. *Diabetes Care* 12:530-536, 1989
 5. Busfield F, Duffy DL, Kesting JB, Walker SM, Lovelock PK, Good D, Tate H, Watego D, Marczak M, Hayman N, Shaw JT: A genomewide search for type 2 diabetes-susceptibility genes in indigenous Australians. *Am J Hum Genet* 70:349-357, 2002
 6. Duggirala R, Williams JT, Williams-Blangero S, Blangero J: A variance component approach to dichotomous trait linkage analysis using a threshold model. *Genet Epidemiol* 14:987-992, 1997
 7. Francke S, Manraj M, Lacquemant C, Lecoecur C, Lepretre F, Passa P, Hebe A, Corset L, Yan SL, Lahmidi S, Jankee S, Gunness TK, Ramjuttun US, Balgobin V, Dina C, Froguel P: A genome-wide scan for coronary heart disease suggests in Indo-Mauritians a susceptibility locus on chromosome 16p13 and replicates linkage with the metabolic syndrome on 3q27. *Hum Mol Genet* 10:2751-2765, 2001
 8. Frayling TM, Lindgren CM, Chevre JC, Menzel S, Wishart M, Benmezroua Y, Brown A, Evans JC, Rao PS, Dina C, Lecoecur C, Kanninen T, Almgren P, Bulman MP, Wang Y, Mills J, Wright-Pascoe R, Mahtani MM, Prisco F, Costa A, Cagnet 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* 52:872-881, 2003
 9. Hsueh WC, Mitchell BD, Shuldiner AR: Use of genome scans to identify susceptibility genes for type 2 diabetes. In *Genetics of Diabetes Mellitus*. Lowe WL, Ed. Dordrecht, the Netherlands, Kluwer Academic Publishers, 2001, p. 231-250
 10. Hsueh WC, St 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* 52:550-557, 2003
 11. Imperatore G, Hanson RL, Pettitt DJ, Kobes S, Bennett PH, Knowler WC: Sib-pair linkage analysis for susceptibility genes for microvascular complications among Pima Indians with type 2 diabetes: Pima Diabetes Genes Group. *Diabetes* 47:821-830, 1998
 12. Iwasaki N, Cox NJ, Wang YQ, Schwarz PE, Bell GI, Honda M, Imura M, Ogata M, Saito M, Kamatani N, Iwamoto Y: Mapping genes influencing type 2 diabetes risk and BMI in Japanese subjects. *Diabetes* 52:209-213, 2003
 13. 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* 70:509-516, 2002
 14. Martin LJ, Comuzzie AG, Dupont S, Vionnet N, Dina C, Gallina S, Houari M, Blangero J, Froguel P: A quantitative trait locus influencing type 2 diabetes susceptibility maps to a region on 5q in an extended French family. *Diabetes* 51:3568-3572, 2002
 15. McCarthy MI, Froguel P: Genetic approaches to the molecular understanding of type 2 diabetes. *Am J Physiol Endocrinol Metab* 283:E217-E225, 2002
 16. 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* 51:833-840, 2002
 17. Mori Y, Otabe S, Dina C, Yasuda K, Populaire C, Lecoecur C, Vatin V, Durand E, Hara K, Okada T, Tobe K, Boutin P, Kadowaki T, Froguel P: Genome-wide search for type 2 diabetes in Japanese affected sib-pairs confirms susceptibility genes on 3q, 15q, and 20q and identifies two new candidate Loci on 7p and 11p. *Diabetes* 51:1247-1255, 2002
 18. Parker A, Meyer J, Lewitzky S, Rennie JS, Chan G, Thomas JD, Orholm-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* 50:675-680, 2001
 19. Stern MP: The search for type 2 diabetes susceptibility genes using whole-genome scans: an epidemiologist's perspective. *Diabetes Metab Res Rev* 18:106-113, 2002
 20. Watanabe RM, Ghosh S, Langefeld CD, Valle TT, Hauser ER, Magnuson VL, Mohlke KL, Silander K, Ally DS, Chines P, Blaschak-Harvan J, Douglas JA, Duren WL, Epstein MP, Fingerlin TE, Kaleta HS, Lange EM, Li C, McEachin RC, Stringham HM, Trager E, White PP, Balow JJ, Birznieks G, Chang J, Eldridge W: The Finland-United States investigation of non-insulin-dependent diabetes mellitus genetics (FUSION) study: II. An autosomal genome scan for diabetes-related quantitative-trait loci. *Am J Hum Genet* 67:1186-1200, 2000
 21. Xiang K, Wang Y, Zheng T, Shen K, Jai W, Li J, Lin X, Wu S, Zhang G, Wang S, Lu H: Genome wide scan for type 2 diabetes susceptibility loci in Chinese. (Abstract). *Diabetes* 51 (Suppl. 2):A262, 2002
 22. Gardner LI Jr, Stern MP, Haffner SM, Gaskill SP, Hazuda HP, Relethford JH, Eifler CW: Prevalence of diabetes in Mexican Americans: relationship to percent of gene pool derived from Native American sources. *Diabetes* 33:86-92, 1984
 23. Chakraborty R, Ferrell RE, Stern MP, Haffner SM, Hazuda HP, Rosenthal M: Relationship of prevalence of non-insulin-dependent diabetes mellitus to Amerindian admixture in the Mexican Americans of San Antonio, Texas. *Genet Epidemiol* 3:435-454, 1986
 24. Duggirala R, Blangero J, Almasy L, Dyer TD, Williams KL, Leach RJ, O'Connell P, Stern MP: Linkage of type 2 diabetes mellitus and of age at onset to a genetic location on chromosome 10q in Mexican Americans. *Am J Hum Genet* 64:1127-1140, 1999
 25. Stern MP, Rosenthal M, Haffner SM, Hazuda HP, Franco LJ: Sex difference in the effects of sociocultural status on diabetes and cardiovascular risk factors in Mexican Americans: the San Antonio Heart Study. *Am J Epidemiol* 120:834-851, 1984
 26. 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* 6:724-734, 1996
 27. American Diabetes Association: Clinical practice recommendations 2002. *Diabetes Care* 25:S1-S148, 2002
 28. World Health Organization: *Report of a WHO Consultation: Definition, Diagnosis and Classification of Diabetes Mellitus and Its Complications. Part I: Diagnosis and Classification of Diabetes Mellitus*. Alwan A, King H, Eds. Geneva, World Health Org., Department of Noncommunicable Disease Surveillance, 1999, 1-59
 29. Sobel E, Lange K: Descent graphs in pedigree analysis: applications to haplotyping, location scores, and marker-sharing statistics. *Am J Hum Genet* 58:1323-1337, 1996
 30. Lander ES, Green P: Construction of multilocus genetic linkage maps in humans. *Proc Natl Acad Sci U S A* 84:2363-2367, 1987
 31. Matisse TC, Perlin M, Chakravarti A: Automated construction of genetic linkage maps using an expert system (MultiMap): a human genome linkage map. *Nat Genet* 6:384-390, 1994
 32. Almasy L, Blangero J: Multipoint quantitative-trait linkage analysis in general pedigrees. *Am J Hum Genet* 62:1198-1211, 1998
 33. Heath SC: Markov chain Monte Carlo segregation and linkage analysis for oligogenic models. *Am J Hum Genet* 61:748-760, 1997
 34. Yoo B, Pankratz VS, de Andrade M: Practical application of residuals from survival models in quantitative trait linkage analysis. *Genet Epidemiol* 21 (Suppl. 1):S811-S816, 2001
 35. Blangero J, Williams JT, Almasy L: Variance component methods for detecting complex trait loci. *Adv Genet* 42:151-181, 2001
 36. Allison DB, Neale MC, Zannolli R, Schork NJ, Amos CI, Blangero J: Testing the robustness of the likelihood-ratio test in a variance-component quantitative-trait loci-mapping procedure. *Am J Hum Genet* 65:531-544, 1999
 37. Vionnet N, Hani E, Dupont S, Gallina S, Francke S, Dotte S, De Matos F, Durand E, Lepretre F, Lecoecur 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* 67:1470-1480, 2000
 38. Duggirala R, Blangero J, Almasy L, Arya R, Dyer TD, Williams KL, Leach RJ, O'Connell P, Stern MP: A major locus for fasting insulin concentrations and insulin resistance on chromosome 6q with strong pleiotropic effects on obesity-related phenotypes in nondiabetic Mexican Americans. *Am J Hum Genet* 68:1149-1164, 2001
 39. Mitchell BD, Cole SA, Hsueh WC, Comuzzie AG, Blangero J, MacCluer JW, Hixson JE: Linkage of serum insulin concentrations to chromosome 3p in Mexican Americans. *Diabetes* 49:513-516, 2000
 40. Ghosh S, Watanabe RM, Valle TT, Hauser ER, Magnuson VL, Langefeld CD, Ally DS, Mohlke KL, Silander K, Kohtamaki K, Chines P, Balow JJ, Birznieks G, Chang J, Eldridge W, Erdos MR, Karanjawala ZE, Knapp JJ, Kudenko K, Martin C, Morales-Mena A, Musick A, Musick T, Pfahl C, Porter R, Rayman JB: The Finland-United States investigation of non-insulin-dependent diabetes mellitus genetics (FUSION) study: I. An autosomal genome scan for genes that predispose to type 2 diabetes. *Am J Hum Genet* 67:1174-1185, 2000
 41. Wiltshire S, Hattersley AT, Hitman GA, Walker M, Levy JC, Sampson M,

- O'Rahilly S, Frayling TM, Bell JI, Lathrop GM, Bennett A, Dhillion 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* 69:553–569, 2001
42. An P, Hong Y, Weisnagel SJ, Rice T, Rankinen T, Leon AS, Skinner JS, Wilmore JH, Chagnon YC, Bergman RN, Bouchard C, Rao DC: Genomic scan of glucose and insulin metabolism phenotypes: the HERITAGE Family Study. *Metabolism* 52:246–253, 2003
43. Elbein SC, Hoffman MD, Teng K, Leppert MF, Hasstedt SJ: A genome-wide search for type 2 diabetes susceptibility genes in Utah Caucasians. *Diabetes* 48:1175–1182, 1999
44. Hanis CL, Boerwinkle E, Chakraborty R, Ellsworth DL, Concannon P, Stirling B, Morrison VA, Wapelhorst B, Spielman RS, Gogolin-Ewens KJ, Shepard JM, Williams SR, Risch N, Hinds D, Iwasaki N, Ogata M, Omori Y, Petzold C, Rietzch H, Schroder HE, Schulze J, Cox NJ, Menzel S, Boriraj VV, Chen X: A genome-wide search for human non-insulin-dependent (type 2) diabetes genes reveals a major susceptibility locus on chromosome 2. *Nat Genet* 13:161–166, 1996
45. 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* 101:1757–1764, 1998
46. 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
47. Kowluru A: Adenine and guanine nucleotide-specific succinyl-CoA synthetases in the clonal beta-cell mitochondria: implications in the beta-cell high-energy phosphate metabolism in relation to physiological insulin secretion. *Diabetologia* 44:89–94, 2001
48. Lang J: Molecular mechanisms and regulation of insulin exocytosis as a paradigm of endocrine secretion. *Eur J Biochem* 259:3–17, 1999
49. Proks P, Eliasson L, Ammala C, Rorsman P, Ashcroft FM: Ca(2+)- and GTP-dependent exocytosis in mouse pancreatic beta-cells involves both common and distinct steps. *J Physiol* 496:255–264, 1996
50. Wollheim CB, Ullrich S, Meda P, Vallar L: Regulation of exocytosis in electrically permeabilized insulin-secreting cells: evidence for Ca2+ dependent and independent secretion. *Biosci Rep* 7:443–454, 1987
51. Wang B, Lin D, Li C, Tucker P: Multiple domains define the expression and regulatory properties of Foxp1 forkhead transcriptional repressors. *J Biol Chem* 278:24259–24268, 2003
52. Nakae J, Biggs WH III, Kitamura T, Cavenee WK, Wright CV, Arden KC, Accili D: Regulation of insulin action and pancreatic beta-cell function by mutated alleles of the gene encoding forkhead transcription factor Foxo1. *Nat Genet* 32:245–253, 2002
53. Wildin RS, Ramsdell F, Peake J, Faravelli F, Casanova JL, Buist N, Levy-Lahad E, Mazzella M, Goulet O, Perroni L, Bricarelli FD, Byrne G, McEuen M, Proll S, Appleby M, Brunkow ME: X-linked neonatal diabetes mellitus, enteropathy and endocrinopathy syndrome is the human equivalent of mouse scurfy. *Nat Genet* 27:18–20, 2001
54. Pederson BA, Csitkovits AG, Simon R, Schroeder JM, Wang W, Skurat AV, Roach PJ: Overexpression of glycogen synthase in mouse muscle results in less branched glycogen. *Biochem Biophys Res Commun* 305:826–830, 2003
55. DeFronzo RA: Pathogenesis of type 2 diabetes: metabolic and molecular implications for identifying diabetic genes. *Diabetes Rev* 5:177–269, 1997