A Genome-Wide Search for Type 2 Diabetes Susceptibility Genes in Utah Caucasians

Steven C. Elbein, Michael D. Hoffman, Kui Teng, Mark F. Leppert, and Sandra J. Hasstedt

Considerable evidence supports a major inherited component of type 2 diabetes. We initially conducted a genome-wide scan with 440 microsatellite markers at 10-cM intervals in 19 multigenerational families of Northern European ancestry with at least two diabetic siblings. Initial two-point analyses of these families directed marker typing of 23 additional families. Subsequently, all available marker data on the total of 42 families were analyzed using both parametric and nonparametric multipoint methods to test for linkage to type 2 diabetes. One locus on chromosome 1q21-1q23 met genome-wide criteria for significant linkage under a model of recessive inheritance with a common diabetes allele (logarithm of odds [LOD] = 4.295). Both pedigree-based nonparametric linkage (NPL) analysis and affected sib pair (MAPMAKER/SIBS) nonparametric methods also showed the highest genome-wide scores at this region, near markers CRP and APOA2, but failed to meet levels of genome-wide significance. The risk of type 2 diabetes to siblings of a diabetic person when compared with the population (λ_S) was estimated from MAPMAKER/SIBS to be 2.8 in these 42 families. Simulation studies using study data confirmed a genome-wide significance level of $P < 0.05$ (95% CI 0.005–0.0466). However, analysis of 20 similarly ascertained but smaller families failed to confirm this linkage. The LOD score with 50% heterogeneity for all 62 families considered together was only 2.25, with an estimated λ_S of 1.87. Our data suggest a novel diabetes susceptibility locus near APOA2 on chromosome 1 in a region with many transcribed genes.
***Diabetes* 48:1175–1182, 1999**

Considerable data support a major role for genetic susceptibility in the pathogenesis of type 2 diabetes. Several genes have been identified that may increase the risk of type 2 diabetes (1–4), but that do not appear to be major type 2 diabetes susceptibility loci (5–11). Although other genes do act as major loci in rare families (12–17), they are not important causes of type 2 diabetes with more typical onset (18–23).

Because candidate gene studies have not uncovered major type 2 diabetes susceptibility loci (19,21,24), many groups have undertaken genome-wide scans to identify unique loci for type 2 diabetes. Hanis et al. (9) reported a locus near the telomere of chromosome 2q (*NIDDM1*) in Hispanic type 2 diabetic sib pairs from Starr County, Texas. Mahtani et al. (25) found no significant linkage of diabetes to any region in a genome-wide scan of 26 Finnish type 2 diabetic families, but they reported a locus near *MODY3* on chromosome 12q (*NIDDM2*) among the subset of families with the lowest insulin response to glucose. In the San Antonio Hispanic population, Stern et al. (26) reported linkage of glucose to regions on chromosome 11 near the sulfonylurea receptor gene and on chromosome 6.

We used the genome-wide scan to test the hypothesis that a major susceptibility locus contributes to type 2 diabetes in multigenerational Utah families ascertained for Northern European ancestry and in which two or more siblings had onset of type 2 diabetes before age 65. This population is typical of other populations of Northern European ancestry (27) but is characterized by large sibships. We report the results of a 10-cM genome-wide search of 22 autosomes using both parametric and nonparametric multipoint analysis. Although we find some evidence for linkage proximal to the *NIDDM1* locus, our strongest evidence is for a major recessive type 2 diabetes locus on chromosome 1q21-q23.

RESEARCH DESIGN AND METHODS

Study population. The primary study population comprised 19 families with 468 members, 395 of whom were nonfounders (neither marry-in spouses nor parents of the proband) and 73 of whom were spouses or living parents (founders). Families were ascertained in Utah for Northern European ancestry and the presence of a type 2 diabetic sib pair with onset of diabetes before age 65. Available parents and siblings of the sib pair were studied, and offspring of the sib pair and their siblings were studied if they were over 18 years of age. Families in which both parents of the proband sib pair were known to be diabetic were not studied. All nondiabetic family members underwent a 75-g oral glucose tolerance test.

During the genotyping of these 19 families, an additional 23 families were similarly ascertained, except that in 7 families only a simple sib pair was available for study. Because these families were not available during the initial studies, we used a staged approach to genotyping additional families in which only selected markers were typed in all 42 familial type 2 diabetic kindreds (initial 19 families and 23 newly ascertained families). The extended family set thus comprised 42 families with 618 members, of whom 526 were nonfounders. Additional details on the study population have been reported elsewhere (8,28) and are available from the authors. No family was ascertained for maturity-onset diabetes of the young (MODY). The mean number of individuals/family was 14.3 nonfounders (range 2–24) with 4.0 affected individuals (mean 2–7).

An additional 20 families were sampled during the final genotyping on the initial 19 families and the extended (42-family) set, and thus were available only for replication studies. In addition to ascertainment at a later time, these families were smaller than the 42 families of the primary study (mean number of individuals tested was 13.5 for the primary set, 6.6 for the replication set), had fewer affected individuals (mean 4.1 in primary set vs. 3.2 in replication set), and had slightly

From the Division of Endocrinology (S.C.E., K.T.), University of Arkansas for Medical Sciences, and John L. McClellan Memorial Veterans Hospital, Little Rock, Arkansas; the Division of Endocrinology (M.D.H.), Department of Veterans' Affairs, and the University of Utah School of Medicine; and the Department of Human Genetics (M.F.L.), University of Utah School of Medicine, Salt Lake City, Utah.

Address correspondence and reprint requests to Steven C. Elbein, MD, Endocrinology 111J/LR, John L. McClellan Memorial Veterans Hospital, 4300 West 7th St., Little Rock, AR 72205. E-mail: elbeinstevenc@exchange.uams.edu.

Received for publication 2 October 1998 and accepted in revised form 3 February 1999.

Detailed marker information can be found in an on-line appendix at www.diabetes.org/diabetes/appendix.asp.

IGT, impaired glucose tolerance; LOD, logarithm of odds; MODY, maturity-onset diabetes of the young; NPL, nonparametric linkage; WHO, World Health Organization.

later onset (mean age of reported onset was 50.6 years in primary set, 51.5 years in replication set).

Before the linkage studies, the initial 19 families were screened for glucokinase mutations, and none were found (29). Additionally, no mitochondrial mutations were found (30). Linkage to both *MODY1* and *MODY3* regions was rejected before multipoint studies, but screening for *MODY3* (hepatocyte nuclear factor 1α) mutations in all 62 families was only completed after the analyses described here (see below). Families have not been screened for *MODY1* mutations.

The study protocol was approved by the University of Utah Institutional Review Board, and all study participants provided informed consent before the study.

Determination of affection status. To achieve a single criterion for affection, individuals were considered diabetic for the linkage analysis if they met one of the following criteria: previously diagnosed type 2 diabetes on medical therapy; fasting glucose >7.8 mmol/l; or 2-h postchallenge (75 g glucose) glucose >7.8 mmol/l for <45 years of age, >11.1 mmol/l for 45–64 years of age, or >13.3 mmol/l for >64 years of age. These cutoffs correspond to World Health Organization (WHO) criteria for impaired glucose tolerance (IGT) for <45 years of age or type 2 diabetes (45–65 years of age), but they take into account the age dependence of postchallenge glucose (31). These a priori criteria, while nonstandard, allowed for a single diagnostic scheme that incorporated the population prevalence of IGT in determining affection status.

Nondiabetic subjects >45 years of age and subjects with IGT were restudied after a 3- to 5-year interval if available, and diagnostic status was updated by regular questionnaires during the study period. Subjects were considered affected if a diagnosis was established by repeat screening or if treatment with antidiabetic medication was initiated before multipoint analysis. Subjects with IGT or diabetes who did not meet age-specific diagnostic criteria, subjects with type 1 diabetes, and patients with either anti–islet cell or anti-GAD antibodies (32) were considered to be of unknown affection status. From the 19 families used for the primary genome scan, 50 individuals were of unknown affection status (35 nonfounders), and 119 were considered affected (105 nonfounders). Of those classified as affected for the primary analyses, 14 met WHO criteria for IGT but not diabetes. From the 42-family set used in extended studies, 62 were considered of unknown affection status (44 nonfounders), and 183 were considered affected (166 nonfounders), including 2 subjects diagnosed subsequent to the study visit and 18 individuals who had IGT by WHO criteria. Additionally, 14 elderly individuals with 2-h glucose values <13.3 mmol/l but normal fasting glucose were considered nondiabetic for linkage studies. For these studies, we also allowed for uncertainty in diagnosis for conflicting test results or oral glucose tolerance test results near cutoff points (Table 1). All classification decisions were made before data analysis and not altered.

Individuals with clear diabetes on therapy were considered affected regardless of subsequent testing and were not retested. Individuals with IGT or diabetes not requiring therapy and individuals over age 45 were invited to be retested, but retesting was biased to families ascertained early in the study. Additionally, regular questionnaires were distributed to family members to update their health status. The most recent test data were used for classification of affection status and assignment of liability class. Individuals with conflicting data on multiple tests were considered of unknown affection status. Because of updated diagnoses, affection status was different from previously published analyses (7,8).

Marker typing. DNA was prepared from peripheral lymphocytes or lymphoblastoid cell lines by standard methods (7). Microsatellite markers were chosen from published maps (33–37) for heterozygosity >0.7 and ~10-cM intervals for each chromosome. A total of 440 markers were typed at a mean interval of <9.2 cM. The largest gap was 22 cM (chromosome 1p between markers D1S233 and D1S199). Specific markers are available in an on-line appendix at www.diabetes.org/diabetes/appendix.asp. Microsatellite markers were typed using radioactive methods, as described previously (7,8,28).

Linkage analysis. As described above, the primary genome scan was conducted on 19 families. Each chromosome was tested using two-point parametric and nonparametric (affected pedigree member) (38) analyses as completed. The additional 23 families were entered into the study as they became available, with marker typing prioritized to regions in which two-point logarithm of odds (LOD) scores exceeded 1.0 under any parametric model (dominant, dominant with very low penetrance, recessive [7,8,28]) or $P < 0.001$ on affected pedigree member analysis. We also tested regions of suggestive linkage from other laboratories in all 42 families before multipoint analyses. Allele frequencies were determined from 60–100 spouse or unrelated controls of the same ethnic background as the families. Intermarker distances for multipoint analyses were estimated from published maps (35–37) and confirmed in disease families with the ILINK program from the LINKAGE 5.1 package (39–41).

Our primary analysis approach was multipoint parametric linkage analysis of families, performed with GENEHUNTER (42) under dominant and recessive parametric models, which allowed for inclusion of a penetrance function based on age modified by degree of obesity (Table 1). Although this model is empirical, inclusion of obesity reduces the variance in liability classes among affected members of a single family and fits empirical observations that obese family members have increased predisposition to diabetes. Families were trimmed (see below) as for nonparametric linkage (NPL) analysis. Although initial nonparametric analysis used the NPL_{ALL} statistic, based on subsequent reports (43,44) we present data using the modified statistic proposed by Kong and Cox (43). We also analyzed the data using multipoint affected sib pair analysis (MAPMAKER/SIBS [45]) under the "possible triangle" statistic based on reports that this method may be the most powerful nonparametric statistic (44). Exclusion mapping was performed using MAPMAKER/SIBS at recurrence risks of λ_s = 1.4, 1.6, and 2.8. Because affected sib pair and NPL analyses do not include penetrance, we conducted nonparametric analyses both including and excluding affected individuals in liability classes 7 and 8.

The limitations of GENEHUNTER on family size (twice the number of nonfounders minus the number of founders cannot exceed 16 [42]) required larger families to be trimmed before parametric and NPL analyses. First, each family was trimmed to make it unilineal by removing affected spouses and their children. If the family still exceeded the allowable size, unaffected individuals were succes-

TABLE 1
Parametric models for type 2 diabetes genome search

			Dominant or recessive model	
	Class	Percent sporadics	Penetrance (DD or Dd)	Sporadic penetrance, standard (dd)
Age range (years)				
<30	1	0	0.02	0
30–45	2	1	0.15	2.00E-05
45–55	3	5	0.30	0.0017
55–65	4	10	0.45	0.0050
65–75	5	20	0.60	0.0160
>75	6	40	0.60	0.0440
Certainty				
90%	7	20	0.58	0.1128
80%	8	20	0.56	0.2100

Parameters for dominant and recessive models are shown. Disease allele frequency was 0.05 for dominant and 0.25 for the recessive models. Penetrance is based on a linear function, with the maximum conservatively determined from twin studies. Uncertainty was used for individuals with conflicting laboratory tests or for those who just exceeded thresholds for affection status, according to the method of Ott (39) and Terwilliger and Ott (65). For affected sib pair and NPL analyses, individuals were considered affected regardless of liability class. Percent sporadics represents the percentage of all diabetic patients who were estimated to be phenocopies. The age-determined liability class was adjusted upward by one class for each 5 kg/m^2 of BMI ≥30, up to three classes, to derive the liability class used in the linkage analysis.

sively removed beginning with those most distantly related to any affected individual until the size limit was attained. After trimming, 341 members of 42 families for whom genotype data were available were included.

To explore the effects of nonstandard diagnostic criteria and the effects of trimming families, we performed several post-hoc analyses of the chromosome 1 linkage. Analysis was repeated using the recessive model with affection defined by WHO criteria (pharmacological therapy or 2-h glucose >11.1 mmol/l); with original affection criteria but all unaffected individuals considered to be unknown ("affected-only"); and with four markers (D1S305, CRP, APOA2, and D1S196) but full pedigree information as implemented in VITESSE (46).

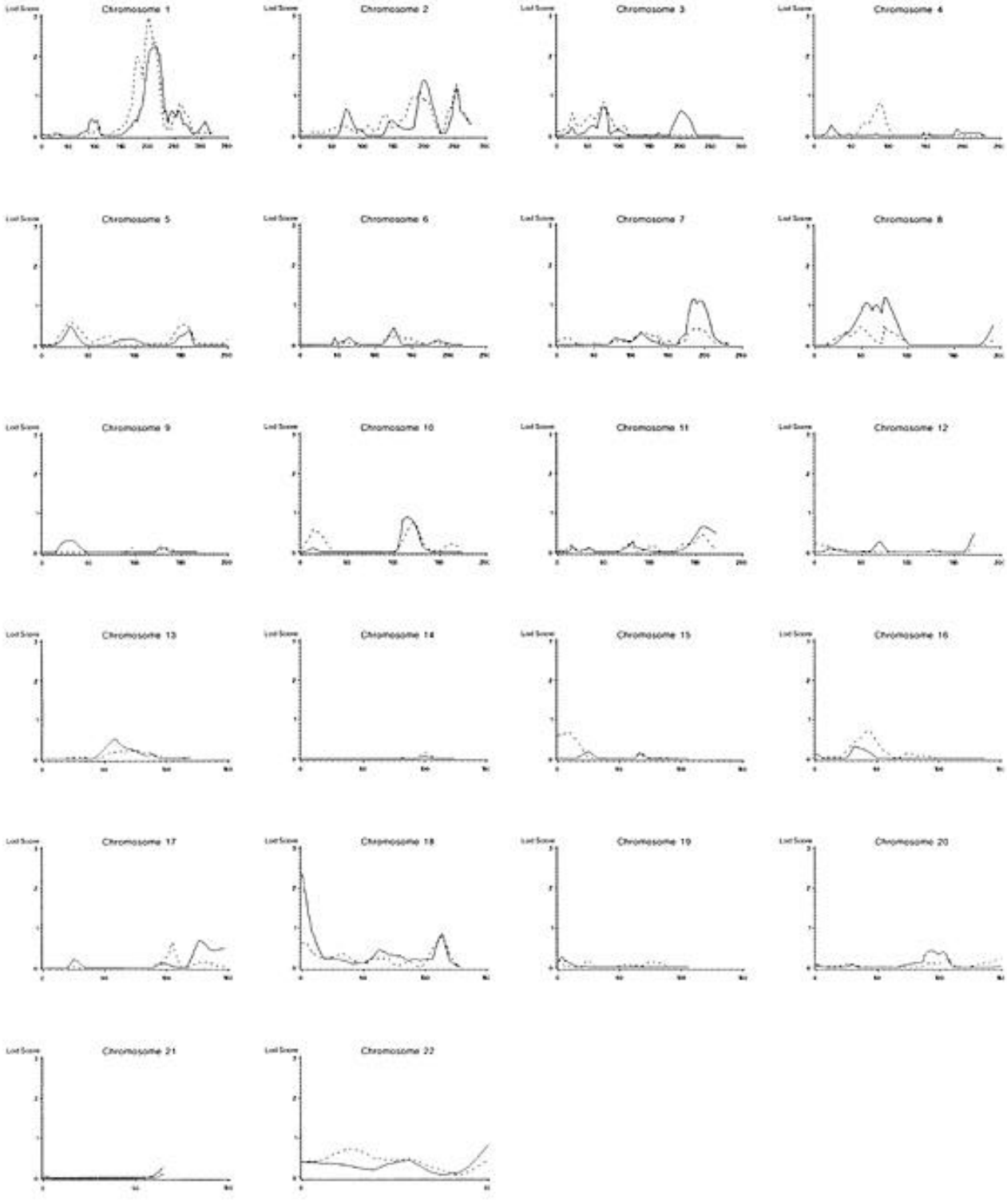

FIG. 1. The tracing of the modified NPL_{ALL} statistic expressed as the equivalent LOD score (——) and the MAPMAKER/SIBS maximized over λ (– – –) for each chromosome. Each analysis is shown with all individuals included. Distance is in cM from the most p-ter marker for each chromosome. See also Table 2 for highest scores on each model. Marker names and locations are available from the *Diabetes* Web site or from the authors.

Simulation analysis. Empirical significance of the peak LOD score under parametric multipoint analysis was determined by simulating 100 replicates of the genome-wide scan. We generated marker genotypes (47) for 100 replicates of all 22 autosomes (440 markers), assuming the marker distances and allele frequencies used in the analysis and the missing values found in the data, but without reference to disease phenotypes. Because of the computational time involved in conducting the full genome scan on all replicates with both dominant and recessive models, we used a modification of our actual analysis method. First, we used FASTLINK (41) to compute two-point LOD scores for linkage of each marker in each replicate to type 2 diabetes with recombination of 0.01, 0.05, 0.10, and 0.20 for both the dominant and recessive models. If any two-point LOD score on the chromosome exceeded 2.0, we conducted a multipoint analysis using GENEHUNTER (42) on the complete chromosome using the same model and recorded the maximum LOD score for that chromosome.

Replication studies. To determine whether results in the initial 42 families could be extended to a small number of similarly ascertained families, we typed 122 individuals from 20 newly ascertained families. Replication families were typed for the putative locus on chromosome 1q21-23 and the *NIDDM1* region on chromosome 2q, but they have not been typed at other regions.

RESULTS

As described above, because family ascertainment continued during initial marker typing, we used a staged approach in which all markers were typed in 19 multigenerational families. In addition, both two-point analysis from these studies and work from other laboratories (9,25,48,49) was used to direct typing of 23 additional families for chromosomes 1, 2, 7, 9, 11, 12, 19, and 20 before multipoint studies of all available data on 42 families. Two-point results from our laboratory have been partially reported elsewhere (6–8,28). Based on these and unpublished initial results, we typed additional markers on chromosomes 1, 7, 9, 11, and 19 and included 23 additional similarly ascertained families. Our primary multipoint linkage analysis was performed with all available marker data under dominant and recessive parametric methods using the GENEHUNTER program (42). However, to perform analyses comparable with other published studies, we also analyzed the data using affected sib pair (45) and NPL_{ALL} (42,43). Parameters for parametric analyses are shown in Table 1. Disease gene frequency was set to define a 5% population prevalence, which with a maximum of sporadic (noninherited) frequency of 40%, easily approximates the population prevalence of type 2 diabetes in Utah. Because parametric analyses excluded most regions other than chromosome 1, we show only results of the NPL_{ALL} (expressed as a LOD score) and affected sib pair analysis maximized over λ_s in Fig. 1. We show nonparametric statistics calculated when individuals with uncertain diagnoses (liability classes 7 and 8; see METHODS) were included, which was more powerful for most regions.

The maximum LOD for linkage was 4.295, which occurred under the recessive model on chromosome 1q21–23 at position 204 cM from p-ter marker, between markers CRP and APOA2 (Fig. 2). The one LOD unit CI (39) extended from position 202 to position 221 or approximately from marker CRP to marker D1S196. In contrast, linkage under the dominant model was excluded in this region (LOD < –2). Both NPL (LOD = 2.260, position 214) and affected sib=pair (LOD = 2.964, position 201) analyses peaked within the CI at position 213 but failed to meet genome-wide significance levels. Using MAPMAKER/SIBS (42), we estimated λ_s (the ratio of the risk to a sibling of a diabetic to the population risk) at 2.89, although the LOD score varied little between $\lambda_s = 2.4$ and $\lambda_s = 3.2$. We found no evidence for heterogeneity using the admixture test (39), and we still obtained a LOD score of 4.09 when we sub-

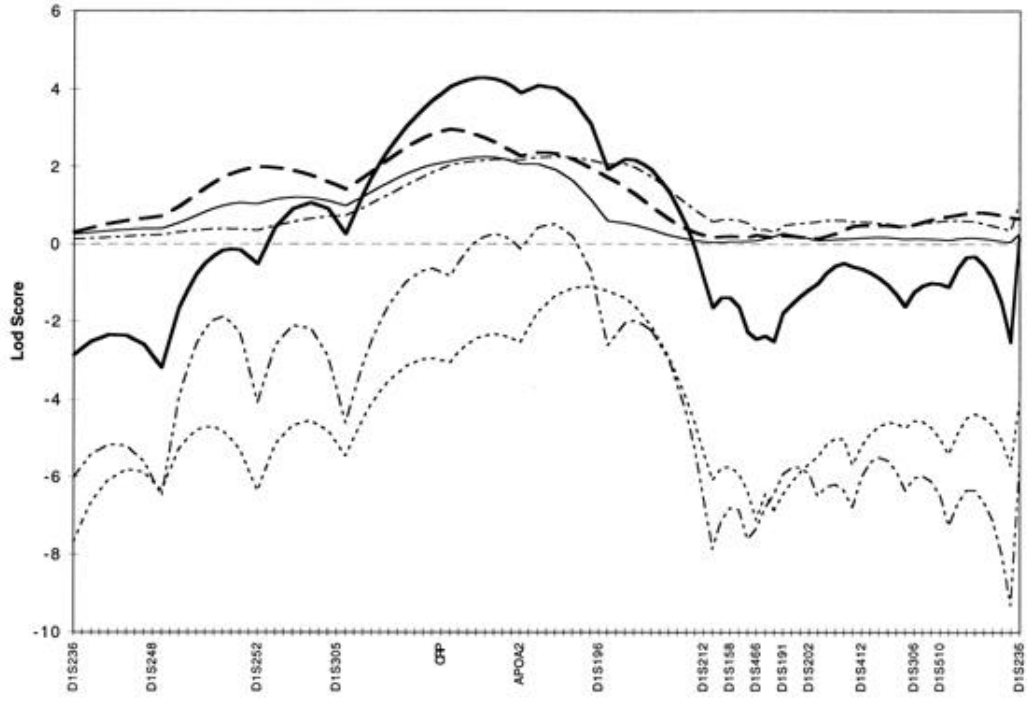

FIG. 2. The dominant, recessive, NPL_{ALL}, and MAPMAKER/SIBS tracings for chromosome 1 both for the 42 families and for the 42 families plus the additional replication families. Distances are taken from the literature but correspond closely to that calculated from disease families. Markers used in the analysis are shown on the *x*-axis. All analyses include all affected individuals, including liability classes 7 and 8. (——), 42 family recessive; (– –), 42 family MAPMAKER/SIBS; (——), 42 family NPL_{ALL}; (- - - -), 42 family dominant; (—·—), 62 family recessive with heterogeneity (upper line) and without heterogeneity (lower line).

stituted intermarker distances calculated from the ILINK (39) analysis of family linkage data for published map data.

No other region reached genome-wide significance levels under any method. All LOD scores ≥1 under both parametric and nonparametric analyses are shown in Table 2. We found no evidence for linkage at previously reported loci on chromosomes 2, 12, or 20 or at regions suggested in our previous two-point analyses (6–8,28). We excluded 82 and 84% of the genome scanned under our dominant and recessive models, respectively, and we excluded 53% of the genome but not the *NIDDM1* region at $\lambda_s = 2.8$ under MAPMAKER/SIBS. We were able to exclude <3% of the genome under MAPMAKER/SIBS at $\lambda_s = 1.6$ (data not shown).

We used simulation to determine the empirical genome-wide significance of our chromosome 1 linkage by using study map distances, allele frequencies, and pedigree data under both dominant and recessive models. One of 100 replicates of a genome-wide simulation exceeded our observed LOD score of 4.30 (4.61); an additional replicate showed a LOD score of 4.11; and 6/100 replicates produced LOD scores between 3 and 4. Based on these results, the 95% CI for the genome-wide significance of a LOD score of 4.295 is $0.0005 < P < 0.0466$, which meets the proposed standard of genome-wide significance of $P < 0.05$ (50).

Despite evidence for genome-wide significance of the linkage to chromosome 1, we were unable to replicate this finding in 20 newly ascertained nuclear families from the same population. The LOD score under the recessive model for these replication families was below –3.0 in the region of most probable linkage on chromosome 1q, and nonparametric analysis showed no evidence for linkage under NPL or affected sib pair methods. As expected from these results, the recessive LOD score for multipoint analysis of the combined 62-family data set was only 0.52 without heterogeneity and 2.25 with 50% heterogeneity. For the combined data, the maximum LOD scores were 0.508 under NPL and 1.681 under MAPMAKER/SIBS.

DISCUSSION

Late-onset type 2 diabetes probably results from the interplay of several genetic determinants and environmental factors, such as sedentary lifestyle and obesity. Although the number of genes involved and the interaction between these genes is uncertain, a small number of loci is consistent with available epidemiological data (51). A simple Mendelian model of typical late-onset type 2 diabetes is improbable. Consequently, most investigators have chosen nonparametric (model-free) approaches that examine allele sharing among affected (diabetic) siblings and that do not require specification of a model of inheritance as the primary approach to type 2 diabetes and similar complex genetic diseases (9,52). These methods do not include information from unaffected family members. However, the merits of this approach have been debated (53,54). Affected sib pair analyses discard much of the family information available from the larger families that characterize our study population. Parametric or model-based analysis methods utilize the additional information from unaffected family members that is excluded from allele-sharing methods. Furthermore, model-based analyses give more weight to the contributions of large families than do GENEHUNTER type analyses. Perhaps for these reasons, simulation studies suggest parametric, model-based analyses under simple dominant and recessive models with reduced penetrance retain high power compared with the correct model, even when the actual disease model is more complex (such as multiple loci) than the analytical model suggests (54). Furthermore, analysis under parametric models provides some information on mode of inheritance of the linked locus (55). Importantly, although data analysis under the wrong genetic model can reduce power, LOD scores are not falsely elevated by choosing incorrect model parameters (53). Thus, although our simple models of analysis are unlikely to reflect reality, and although this misclassification of model parameters may mask linkage to regions with genes of lesser effects, these models are unlikely to raise the LOD score.

We hypothesized that families with type 2 diabetes in a sib pair with onset before age 65 represent a unique subset of typical type 2 diabetes in which major genetic determinants play a more important role. We have tested this hypothesis by identifying and sampling a large number of multigenerational families, by typing both affected and unaffected individuals

TABLE 2
Highest LOD scores for genome-wide scan

LOD score	Method	Chromosome	Position	Markers
4.295	Recessive	1	204	CRP/APOA2
2.964	Sib pair	1	201	CRP/APOA2
2.260	NPL	1	214	CRP/APOA2
2.180	Sib pair*	2	253	D2S336
1.829	NPL*	2	253	D2S336
1.386	NPL	2	200	D2S117
1.063	Sib pair	2	193	D2S138/D2S117
1.348	NPL*	8	55	D8S136
1.365	NPL*	8	75	D8S87/FGFR1/D8S532
1.155	Recessive	11	158	D11S925/D11S912
2.361	NPL	18	0	D18S59

LOD scores >1.0 are shown for all methods. Recessive, LOD score calculated under recessive model in GENEHUNTER; sib pair, LOD score calculated from affected sib pair sharing under MAPMAKER/SIBS; NPL, *Z* score calculated from NPL_{ALL} program in GENEHUNTER, modified to provide improved power in the setting of missing parental data (see METHODS). The LOD score is calculated from the actual *Z* score, assuming an asymptotic distribution from the formula $Z^2/4.6$, and is taken from the ouput file of the modified GENEHUNTER program (43). Position is our estimated distance from the most p-ter marker for each chromosome. Markers are the closest typed markers in our set. The highest score is shown with liability classes 7 and 8 excluded.

for markers on all 22 autosomes, and by using a parametric approach to analyze the data under both dominant and recessive models that is very similar to recently proposed methods (54). We suggest a major type 2 diabetes locus on chromosome 1q21-1q23 that acts most like a high-frequency recessive gene (our model estimated a diabetes allele frequency of 0.25) with moderate penetrance (up to 60% at age 65). The evidence for this locus derives from a maximum LOD score of 4.295, which meets genome-wide significance under several criteria. Because we have tested for two models of dominance (dominant and recessive), a conservative correction would lower the actual value to 4.00 (56), which exceeds the recommended threshold of 3.3, corresponding to a single locus *P* value of 4.9×10^{-5}, or <1 false-positive result in 20 dense genome-wide scans (50). Our simulations using actual marker and pedigree data and both dominant and recessive models also support this conclusion. Arguably, we did not correct for the nonparametric analyses, but these were not our primary analytical methods. Nonetheless, these methods also show the highest genome-wide scores in this region, albeit without reaching genome-wide significance.

Based on our estimate of $\lambda_s = 2.89$ in the multipoint affected sib pair analysis of the initial 42 families—assuming a multiplicative model—and an estimated total λ_s of 3.5 (51), this locus could account for ~85% of type 2 diabetes among these families. By comparison, the *NIDDM1* locus was estimated to have a $\lambda_s = 1.36$ and to explain 30% of the familial clustering of diabetes in the Hispanic population. However, our estimate of λ_s for this locus might be as low as 1.41 in the initial 42 families, thus explaining only 27% of the familial clustering of type 2 diabetes. A more accurate estimate of the role of the chromosome 1 locus may come from our combined 62-family analysis, in which λ_S is only 1.87, thus explaining only 50% of type 2 diabetes under a multiplicative model and 35% under an additive model under the same assumptions.

Subsequent to the analyses described here, we identified 2 families of the 42 initial families with novel missense mutations of the hepatocyte nuclear factor 1α, which segregated with diabetes (20). Removal of these two families from the analysis indeed raises the LOD score for the initial 42 (now 40) families to 4.863, providing stronger evidence for a chromosome 1 locus. In additional post-hoc analyses, we excluded those with less certain diagnoses of diabetes, those whom we considered affected but who did not meet WHO criteria for diabetes, and all unaffected individuals. Each of these analyses lowered the LOD score to between 3.5 and 3.6 (data not shown). When all individuals removed to fit the GENEHUNTER restrictions were restored and only the four closest markers were analyzed using the recessive parametric analysis under VITESSE (46), the LOD score was 3.364. Although the reduced LOD score resulted in part from the analysis of only four markers, inclusion of some previously excluded unaffected individuals contributed to an overall lower LOD score. The differences in LOD scores among these post-hoc analyses reflects the significant role that unaffected individuals play in parametric analyses.

Several aspects of our study suggest that our linkage on chromosome 1q might represent a false positive. First, false-positive LOD scores >4.3 were rarely observed in our simulations. Second, we were unable to replicate our linkage in 20 additional families that differed only in later ascertainment and a tendency to be smaller and older. Third, our LOD score was rather dependent on which unaffected individuals were included. On the other hand, significant linkage of type 2 diabetes to a locus in this same region was reported in Pima Indians using a discordant sib pair approach (57) (C. Bogardus, R. Hanson, personal communication). Furthermore, familial combined hyperlipidemia, which is an insulin-resistant state with a propensity to glucose intolerance, was recently mapped to the same region of chromosome 1 in Finland (58) and in the syntenic region of mice (59). Additional confirmation of our findings must await analysis of a large replication sample of several hundred families from a similar population analyzed with the same models.

Of other genome scans reported, both Hanis et al. (9) and Ghosh et al. (60) clearly excluded this region. In contrast, using the conservative NPL_{ALL} analysis, Mahtani et al. (25) reported *P* values of 0.053 at D1S305 and 0.073 at D1S484, both of which lie within the CI for our most significant linkage (25).

Although we did not find evidence for linkage at the proposed *NIDDM1* locus, we did find some evidence for a locus near marker D2S336 in 42 families when individuals with uncertain diagnoses were excluded (LOD = 2.180, MAPMAKER/SIBS; LOD = 1.830, modified NPL analysis). This locus is ~20 cM proximal to that found by Hanis et al. (9) and, even without correction for multiple analytical methods, would not meet genome-wide significance. Furthermore, these LOD scores also fall when our replication set is included (data not shown). Nonetheless, given map uncertainty in this region and suggestive evidence for a locus in this region in several studies (9,43,61), this region merits further study. Of the other regions with LOD scores exceeding 1, the region on chromosome 11q in which we previously reported evidence for linkage in two-point analyses (8) has been linked to BMI and diabetes in Pima Indians (57). No other region with evidence for linkage in our study showed evidence by all three analytical methods and evidence for linkage in another study population.

Chromosome 1q21-q23 has a large number of mapped genes, of which Apolipoprotein A2 (APOA2), located near the region of maximum LOD score, is the strongest candidate through potential control of free fatty acid levels (62,63). Other possible candidates include the hepatic form of pyruvate kinase (PKLR), a key regulatory enzyme in glycolysis, and LMX1, a regulator of insulin gene transcription (64). Work is in progress to examine the role of these candidate genes.

ACKNOWLEDGMENTS

These studies were supported by Grant DK39311 (S.C.E.) from the National Institutes of Health/NIDDK and by the Research Service of the Department of Veterans' Affairs. Pedigree ascertainment and testing were supported by the General Clinical Research Center (GCRC) of the University of Utah; Public Health Service Grant MO1-RR00064 from the National Center for Research Resources to the University of Utah; and in part by a Harold Rifkin Family Acquisition (GENNID) Grant from the American Diabetes Association. Primer synthesis was supported by PHS (GESTEC) Grant 5-P30-HG00199 to M.F.L. We also acknowledge the W.M. Keck Foundation for support to M.F.L.

We thank K. Wegner, T. Maxwell, and H. Tuckett for their valuable contributions to family ascertainment and DNA

preparation; Kelli Barrett, Denise Fuller, and Phaedra Yount for technical support; the nursing staff of the University of Utah GCRC; and the many volunteer subjects from Utah who made this study possible.

REFERENCES

1. Almind K, Bjorbaek C, Vestergaard H, Hansen T, Echwald S, Pedersen O: Aminoacid polymorphisms of insulin receptor substrate-1 in non-insulin-dependent diabetes mellitus. *Lancet* 342:828–832, 1993
2. Groop LC, Kankuri M, Schalin Jantti C, Ekstrand A, Nikula Ijas P, Widen E, Kuismanen E, Eriksson J, Franssila Kallunki A, Saloranta C, Koskimies S: Association between polymorphism of the glycogen synthase gene and non-insulin-dependent diabetes mellitus. *N Engl J Med* 328:10–14, 1993
3. Hager J, Hansen L, Vaisse C, Vionnet N, Philippi A, Poller W, Velho G, Carcassi C, Contu L, Julier C: A missense mutation in the glucagon receptor gene is associated with non-insulin-dependent diabetes mellitus. *Nat Genet* 9:299–304, 1995
4. Inoue H, Ferrer J, Welling CM, Elbein SC, Hoffman M, Mayorga R, Warren-Perry M, Zhang Y, Millns H, Turner R, Province M, Bryan J, Permutt MA, Aguilar-Bryan L: Sequence variants in the sulfonylurea receptor (SUR) gene are associated with NIDDM in Caucasians. *Diabetes* 45:825–831, 1996
5. Stirling B, Cox NJ, Bell GI, Hanis CL, Spielman RS, Concannon P: Identification of microsatellite markers near the human ob gene and linkage studies in NIDDM-affected sib pairs. *Diabetes* 44:999–1001, 1995
6. Elbein SC, Hoffman M, Ridinger D, Otterud B, Leppert M: Description of a second microsatellite marker and linkage analysis of the muscle glycogen synthase locus in familial NIDDM. *Diabetes* 43:1061–1065, 1994
7. Elbein SC, Chiu KC, Hoffman MD, Mayorga RA, Bragg KL, Leppert MF: Linkage analysis of 19 candidate regions for insulin resistance in familial NIDDM. *Diabetes* 44:1259–1265, 1995
8. Elbein SC, Bragg KL, Hoffman MD, Mayorga RA, Leppert MF: Linkage studies of NIDDM with 23 chromosome 11 markers in a sample of whites of northern European descent. *Diabetes* 45:370–375, 1996
9. Hanis CL, Boerwinkle E, Chakraborty R, Ellsworth DL, Concannon P, Stirling B, Morrison VA, Wapelhorst B, Spielman RS, Gogolin-Ewens KJ, Shephard JM, Williams SR, Risch N, Hinds D, Iwasaki N, Ogata M, Ommoi Y, Petzold C, Rietzsch H, Schroeder H-E, Schulze J, Cox NJ, Menzel S, Boriraj VV, Chen X, Lim LR, Linder T, Mereu LE, Wang Y-Q, Xiang K, Yamagata K, Yang Y, Bell GI: 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
10. Lindner T, Gragnoli C, Schulze J, Rietzsch H, Petzold C, Schroeder H-E, Cox NJ, Bell GI: The 31-cM region of chromosome 11 including the obesity gene Tubby and ATP-sensitive postassium channel genes, SUR1 and KIR6.2, does not contain a major susceptibility locus for NIDDM in 127 non-Hispanic white affected sibships. *Diabetes* 46:1227–1229, 1997
11. Vionnet N, Hani EH, Lesage S, Philippi A, Hager J, Varret M, Stoffel M, Tanizawa Y, Chiu KC, Glaser B, Permutt MA, Passa P, Demenais F, Froguel P: Genetics of NIDDM in France: studies with 19 candidate genes in affected sib pairs. *Diabetes* 46:1062–1068, 1997
12. Steiner DF, Tager HS, Chan SJ, Nanjo K, Sanke T, Rubenstein AH: Lessons learned from molecular biology of insulin-gene mutations. *Diabetes Care* 13:600–609, 1990
13. Taylor SI, Kadowaki T, Kadowaki H, Accili D, Cama A, McKeon C: Mutations in insulin-receptor gene in insulin-resistant patients. *Diabetes Care* 13:257–279, 1990
14. Kadowaki T, Kadowaki H, Mori Y, Tobe K, Sakuta R, Suzuki Y, Tanabe Y, Sakura H, Awata T, Goto Y, Hayakawa T, Matsuoka K, Rawamori R, Kamada T, Horai S, Nonaka I, Hagura R, Akanuma Y, Yazaki Y: A subtype of diabetes mellitus associated with a mutation of mitochondrial DNA. *N Engl J Med* 330:962–968, 1994
15. Froguel P, Zouali H, Vionnet N, Velho G, Vaxillaire M, Sun F, Lesage S, Stoffel M, Takeda J, Passa P, Permutt MA, Beckmann JS, Bell GI, Cohen D: Familial hyperglycemia due to mutations in glucokinase: definition of a subtype of diabetes mellitus. *N Engl J Med* 328:697–702, 1993
16. 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* 384:455–458, 1996
17. Yamagata K, Furuta H, Oda N, Kaisaki PJ, Menzel S, Cox NJ, Fajans SS, Signorini S, Stoffel M, Bell GI: Mutations in the hepatocyte nuclear factor-4alpha gene in maturity-onset diabetes of the young (MODY1). *Nature* 384:458–460, 1996
18. Lesage S, Hani EH, Philippi A, Vaxillaire M, Hager J, Passa P, Demenais F, Froguel P, Vionnet N: Linkage analyses of the MODY3 locus on chromosome 12q with late-onset NIDDM. *Diabetes* 44:1243–1247, 1995
19. Elbein SC: The genetics of human noninsulin-dependent (type 2) diabetes mellitus. *J Nutr* 127:1891S–1896S, 1997
20. Elbein SC, Teng K, Yount P, Scroggin E: Linkage and molecular scanning analyses of MODY3/hepatocyte nuclear factor-1 alpha gene in typical familial type 2 diabetes: evidence for novel mutations in exons 8 and 10. *J Clin Endocrinol Metab* 83:2059–2065, 1998
21. Elbein SC: An update on the genetic basis of type 2 diabetes. *Curr Opin Endocrinol* 5:116–125, 1998
22. Glucksmann MA, Lehto M, Tayber O, Scotti S, Berkemeier L, Pulido JC, Wu Y, Nir WJ, Fang L, Markel P, Munnelly KD, Goranson J, Orho M, Young BM, Whitacre JL, McMenimen C, Wantman M, Tuomi T, Warram J, Forsblom CM, Carlsson M, Rosenzweig J, Kennedy G, Duyk GM, Thomas JD: Novel mutations and a mutational hotspot in the MODY3 gene. *Diabetes* 46:1081–1086, 1997
23. Kaisaki PJ, Menzel R, Linder T, Oda N, Rjasanowski I, Sahm J, Meincke G, Schulze J, Schmechel H, Petzold C, Ledermann HM, Sachse G, Boriraj VV, Kerner W, Turner RC, Yamagata K, Bell GI: Mutations in the hepatocyte nuclear factor-1α gene in MODY and early-onset NIDDM: evidence for a mutational hotspot in exon 4. *Diabetes* 46:528–535, 1997
24. Kahn CR, Vicent D, Doria A: Genetics of non-insulin-dependent (type-II) diabetes mellitus. *Annu Rev Med* 47:509–531, 1996
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, Munnelly K, Parkkonen M, Reeve-Daly MP, Weaver A, Brettin T, Duyk G, Lander ES, Groop LC: Mapping a gene for type 2 diabetes associated with an insulin secretion defect by a genome scan in Finnish families. *Nat Genet* 14:90–94, 1996
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. McLellan T, Jorde LB, Skolnick MH: Genetic distances between Utah Mormons and related populations. *Am J Hum Genet* 36:836–857, 1984
28. Elbein SC, Hoffman MD, Mayorga RA, Barrett KL, Leppert M, Hasstedt S: Do non-insulin-dependent diabetes mellitus (NIDDM) and insulin-dependent diabetes mellitus (IDDM) share genetic susceptibility loci? An analysis of putative IDDM susceptibility regions in familial NIDDM. *Metabolism* 46:48–52, 1997
29. Elbein SC, Hoffman M, Qin H, Chiu K, Tanizawa Y, Permutt MA: Molecular screening of the glucokinase gene in familial type 2 (non-insulin-dependent) diabetes mellitus. *Diabetologia* 37:182–187, 1994
30. Elbein SC, Hoffman MD: Role of mitochondrial DNA tRNA leucine and glucagon receptor missense mutations in Utah white diabetic patients. *Diabetes Care* 19:507–508, 1996
31. Harris MI, Hadden WC, Knowler WC, Bennett PH: International criteria for the diagnosis of diabetes and impaired glucose tolerance. *Diabetes Care* 8:562–567, 1985
32. Elbein SC, Wegner K, Miles C, Yu L, Eisenbarth G: The role of late onset autoimmune diabetes in white familial NIDDM pedigrees. *Diabetes Care* 20: 1248–1251, 1997
33. The Utah Marker Development Group: A collection of ordered tetranucleotide-repeat markers from the human genome. *Am J Hum Genet* 57: 619–628, 1995
34. Buetow KH, Weber JL, Ludwigsen S, Scherbier-Heddema T, Duyk GM, Sheffield VC, Wang Z, Murray JC: Integrated human genome-wide maps constructed using the CEPH reference panel. *Nat Genet* 6:391–393, 1994
35. Gyapay G, Morissette J, Vignal A, Dib C, Fizames C, Millasseau P, Marc S, Bernardi G, Lathrop M, Weissenbach J: The 1993–94 Genethon human genetic linkage map. *Nat Genet* 7:246–249, 1994
36. Matise 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
37. Cooperative Human Linkage Center: A comprehensive human linkage map with centimorgan density. *Science* 265:2049–2054, 1994
38. Weeks DE, Lange K: The affected-pedigree-member method of linkage analysis. *Am J Hum Genet* 42:315–326, 1988
39. Ott J: *Analysis of Human Genetic Linkage.* Revised ed. Baltimore, MD, Johns Hopkins University Press, 1991
40. Cottingham RW Jr, Idury RM, Schaffer AA: Faster sequential in genetic linkage computations. *Am J Hum Genet* 53:252–263, 1993
41. Schaffer AA, Gupta SK, Shriram K, Cottingham RW Jr: Avoiding recomputation in genetic linkage analysis. *Hum Hered* 44:225–237, 1994
42. Kruglyak L, Daly MJ, Reeve-Daly MP, Lander ES: Parametric and nonparametric linkage analysis: a unified multipoint approach. *Am J Hum Genet* 58:1347–1363, 1996
43. Kong A, Cox NJ: Allele-sharing models: LOD scores and accurate linkage tests. *Am J Hum Genet* 61:1179–1188, 1997
44. Davis S, Weeks DE: Comparison of nonparametric statistics for detection of

linkage in nuclear families: single-marker evaluation. *Am J Hum Genet* 61:1431–1444, 1997
45. Kruglyak L, Lander ES: Complete multipoint sib-pair analysis of qualitative and quantitative traits. *Am J Hum Genet* 57:439–454, 1995
46. O'Connell JR, Weeks DE: The VITESSE algorithm for rapid exact multilocus linkage analysis via genotype set-recoding and fuzzy inheritance. *Nat Genet* 11:402–408, 1995
47. Terwilliger JD, Ott J: A novel polylocus method for linkage analysis using the lod-score or affected sib-pair method. *Genet Epidemiol* 10:477–482, 1993
48. 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* 46:882–886, 1997
49. Ji L, Malecki M, Warram JH, Yang Y, Rich SS, Krolewski AS: New susceptibility locus for NIDDM is localized to human chromosome 20q. *Diabetes* 46:876–881, 1997
50. Lander E, Kruglyak L: Genetic dissection of complex traits: guidelines for interpreting and reporting linkage results. *Nat Genet* 11:241–247, 1995
51. Rich SS: Mapping genes in diabetes: genetic epidemiological perspective. *Diabetes* 39:1315–1319, 1990
52. Davies JL, Yoshihiko K, Bennett S, Copeman JB, Cordell HJ, Pritchard LE, Reed PW, Gough SCL, Jenkins C, Palmer SM, Balfour KM, Rowe BR, Farrall M, Barnett AH, Bain SC, Todd JA: A genome-wide search for human type 1 diabetes susceptibility genes. *Nature* 371:130–136, 1994
53. Greenberg DA, Hodge SE, Vieland VJ, Spence MA: Affecteds-only linkage methods are not a panacea. *Am J Hum Genet* 58:892–895, 1996
54. Greenberg DA, Abreu P, Hodge SE: The power to detect linkage in complex disease by means of simple LOD-score analyses. *Am J Hum Genet* 63: 870–879, 1998
55. Greenberg DA, Berger B: Using lod-score differences to determine mode of inheritance: a simple, robust method even in the presence of heterogeneity and reduced penetrance. *Am J Hum Genet* 55:834–840, 1994
56. Hodge SE, Abreu PC, Greenberg DA: Magnitude of type I error when single-locus linkage analysis is maximized over models: a simulation study. *Am J Hum Genet* 60:217–227, 1997
57. 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* 63:1124–1132, 1998
58. Pajukanta P, Nuotio I, Terwilliger JD, Porkka KV, Ylitalo K, Pihlajamaki J, Suomalainen AJ, Syvanen AC, Lehtimaki T, Viikari JS, Laakso M, Taskinen MR, Ehnholm C, Peltonen L: Linkage of familial combined hyperlipidaemia to chromosome 1q21-q23. *Nat Genet* 18:369–373, 1998
59. Castellani LW, Weinreb A, Bodnar J, Goto AM, Doolittle M, Mehrabian M, Demant P, Lusis AJ: Mapping a gene for combined hyperlipidaemia in a mutant mouse strain. *Nat Genet* 18:374–377, 1998
60. Ghosh S, Hauser ER, Magnuson VL, Ally DS, Valle T, Watanabe RM, Nylund SJ, Kohtamaki K, Hagopian WA, Bergman RN, Tuomilehto J, Collins FS, Boehnke M: Multipoint linkage analysis of NIDDM in 534 Finnish families in the FUSION study (Abstract). *Diabetes* 46 (Suppl. 1):76A, 1997
61. Hani EH, Hager J, Philippi A, Demenais F, Froguel P, Vionnet N: Mapping NIDDM susceptibility loci in French families: studies with markers in the region of NIDDM1 on chromosome 2q. *Diabetes* 46:1225–1226, 1997
62. Warden CH, Daluiski A, Bu X, Purcell-Huynh DA, De Meester C, Shieh B, Purpione DL, Gray RM, Reaven GM, Chen Y-DI, Rotter JI, Lusis AJ: Evidence for linkage of the apolipoprotein A-II locus to plasma apolipoprotein A-II and free fatty acid levels in mice and humans. *Proc Natl Acad Sci U S A* 90:10886–10890, 1993
63. McGarry JD: Disordered metabolism in diabetes: have we underemphasized the fat component? *J Cell Biochem* 55:29–38, 1994
64. German MS, Wang J, Fernald AA, Espinosa R 3rd, Le Beau MM, Bell GI: Localization of the genes encoding two transcription factors, LMX1 and CDX3, regulating insulin gene expression to human chromosomes 1 and 13. *Genomics* 24:403–404, 1994
65. Terwilliger JD, Ott J: *Handbook of Human Genetic Linkage*. Baltimore, MD, Johns Hopkins University Press, 1994