



## Evidence for linkage of regions on chromosomes 6 and 11 to plasma glucose concentrations in Mexican Americans.

M P Stern, R Duggirala, B D Mitchell, et al.

*Genome Res.* 1996 6: 724-734

Access the most recent version at doi:[10.1101/gr.6.8.724](https://doi.org/10.1101/gr.6.8.724)

---

**References** This article cites 39 articles, 10 of which can be accessed free at:  
<http://genome.cshlp.org/content/6/8/724.full.html#ref-list-1>

### License

**Email Alerting Service** Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or [click here](#).

---

An advertisement banner with a teal background. On the left, the text reads "CRISPR and RNAi Genetic Screening. Your new superpower." In the center, there is a white box with the words "LEARN MORE". On the right, there is a woman wearing a red and white superhero cape and mask, and the Cellecta logo, which consists of a cluster of green dots and the word "CELLECTA" below it.

---

To subscribe to *Genome Research* go to:  
<https://genome.cshlp.org/subscriptions>

---

Copyright © Cold Spring Harbor Laboratory Press

## RESEARCH

# Evidence for Linkage of Regions on Chromosomes 6 and 11 to Plasma Glucose Concentrations in Mexican Americans

Michael P. Stern,<sup>1,4</sup> Ravindranath Duggirala,<sup>1</sup> Braxton D. Mitchell,<sup>2</sup>  
Laurie J. Reinhart,<sup>1</sup> Sailaja Shivakumar,<sup>1</sup> Patricia A. Shipman,<sup>3</sup>  
Olga C. Uresandi,<sup>3</sup> Edgardo Benavides,<sup>3</sup> John Blangero,<sup>2</sup> and  
Peter O'Connell<sup>3</sup>

<sup>1</sup>Division of Clinical Epidemiology, Department of Medicine, and <sup>3</sup>Department of Pathology, University of Texas Health Science Center, San Antonio, Texas 78284; <sup>2</sup>Department of Genetics, Southwest Foundation for Biomedical Research, San Antonio, Texas 78245

The genetic factors involved in type II diabetes are still unknown. To address this problem, we are creating a 10 to 15 cM genetic map on 444 individuals from 32 Mexican American families ascertained on a type II diabetic proband. Using highly polymorphic microsatellite markers and a multipoint variance components method, we found evidence for linkage of plasma glucose concentration 2 hr after oral glucose administration to two regions on chromosome 11:  $\beta$ -hemoglobin (HBB) and markers DIIS899/DIIS324 near the sulfonylurea receptor (SUR) gene. lod scores at these two loci were 2.77 and 3.37, respectively. The SUR gene region accounted for 44.7% of the phenotypic variance. Evidence for linkage to fasting glucose concentration was also observed for two loci on chromosome 6, one of which is identical to a proposed susceptibility locus for type I diabetes (D6S290). When diabetics were excluded from the analyses, all lod scores became zero, suggesting that the observed linkages were with the trait diabetes rather than with normal variation in glucose levels. Results were similar whether all diabetics were included in the analyses or only those who were not under treatment with oral antidiabetic agents or insulin.

The Mexican American population, residing predominantly in the southwestern part of the United States, has both Native American and European Caucasian ancestry. This population of ~12 million people (U.S. Bureau of Census 1991) shares both cultural and genetic characteristics with the much larger population of Mexico (~80 million) and Central America. Estimates of Native American genetic admixture range from 20% to 40%, with higher levels of Caucasian admixture correlating with higher socioeconomic status (Chakraborty et al. 1986). The prevalence of type II (non-insulin-dependent) diabetes in Mexican Americans is approximately three times higher than in the general U.S. population (Stern and Haffner 1990). Because Native American

populations have diabetes prevalences as much as 10 times higher than the general U.S. population (Knowler et al. 1990; King and Rewers 1993), it is presumed that the excess type II diabetes prevalence in Mexican Americans derives from their Native American ancestry, and it has been shown that diabetes prevalence in several Mexican American populations is proportional to their average Native American genetic admixture (Gardner et al. 1984; Stern and Haffner 1990). Thus it appears that genes predisposing Mexican Americans to type II diabetes are derived from their Native American ancestors and that the risk of acquiring these genes is roughly proportional to the percent of the gene pool derived from this ancestral source. Which genes these might be, however, is completely unknown.

In 1993, we began a gene-mapping project in Mexican American families aimed at identifying chromosomal regions linked to diabetes and re-

<sup>4</sup>Corresponding author.  
E-MAIL [stern@uthscsa.edu](mailto:stern@uthscsa.edu); FAX (210) 567-6955.

lated traits in this ethnic group. Our plan was to perform a whole genome scan at ~10 to 15 cM intervals using 267 highly polymorphic microsatellite markers (Weber and May 1989; Economou et al. 1990). We have now completed the map for chromosomes 1–8, 11, 12, 17, 20, and 21 and have partial coverage on chromosomes 9, 10, 15, and 16. This represents coverage of nearly 75% of the entire genome.

Probands for this study were low-income Mexican American diabetics who had been identified as having type II diabetes in a prior epidemiologic survey (Gardner et al. 1984; Haffner et al. 1986; Stern et al. 1989; Stern and Haffner 1990). These individuals were approached in random order and invited to participate along with all of their first-, second-, and third-degree relatives. Thus, the families included in this study represent a random sample of low-income Mexican American families ascertained on a diabetic proband. A total of 579 individuals in 32 families were examined. Although no effort was made to enroll multiplex families, because of the high background prevalence of type II diabetes in this population (~15% in adults) (Gardner et al. 1984; Haffner et al. 1986; Stern and Haffner 1990), all but three families were multiplex. Fifteen families had diabetics in two generations and an additional five families had diabetics in three generations. Summary characteristics of the pedigrees are shown in Table 1 and clinical characteristics of the family members are shown in Table 2.

Plasma glucose concentration was measured following a 12-hr fast and again 2 hr after a standardized oral glucose load (Orangedex or Koladex, Custom Laboratories, Baltimore, MD), and diabetes was diagnosed according to the criteria of the World Health Organization (WHO) (World Health Organization Expert Committee

1985). Serum insulin concentrations were also measured at the fasting and 2-hr time points using a solid-phase radioimmunoassay (Diagnostic Products Corp., Los Angeles, CA) (Haffner et al. 1986). One hundred forty individuals met criteria for type II diabetes giving a crude prevalence of 24.2% (140/579) or 19.7% (108/547) if probands are excluded. Sixty-five diabetics were under treatment with oral antidiabetic agents or insulin.

A subset of 444 individuals from the largest pedigrees were selected for genotyping. This marker subset contained 103 sibships which, in turn, contained 349 individuals, 83 of whom were diabetic. These sibships yielded 545 sibpairs. Highly polymorphic microsatellite markers (heterozygosity > 0.7) were amplified by PCR (Saiki et al. 1985). The radiolabeled amplification products were separated on 7% denaturing polyacrylamide gels that were dried and exposed to x-ray film, and the genotypes were scored. Microsatellite markers were selected to cover the genome at ~10 to 15 cM intervals.

We used a multipoint variance components method to estimate the genetic variance attributable to the region around a given genetic marker. This method is based on specifying the expected genetic covariances between pairs of relatives as a function of their identity-by-descent (IBD) relationships at a given marker locus [assumed to be tightly linked to a quantitative trait locus (QTL)]. Although in principle this method can be applied to extended pedigrees, the analyses presented in this paper are limited to sibships to simplify the estimation of the IBD matrix. The variance components method has a number of advantages compared to the widely used sibpair method of Haseman and Elston (1972) (see Methods). We implemented this analytic technique using the program FISHER (Lange et al. 1988) and estimated the IBD relationships between sibpairs using the program SIBPAL in Statistical Analysis for Genetic Epidemiology (SAGE) (release 2.2, 1994. Louisiana State University Medical Center, New Orleans). The multipoint method requires that the distances between the markers be specified. To specify the genetic maps, we used the map distances reported in the Human Genome Data Base (GDB) (Fasman et al. 1994) supplemented by distances calculated from our genotype data with the program CRI-MAP VERSION 2.4 (1990; P. Green, K. Fallon, and S. Crooks, unpubl.) for markers not included in the GDB (see Methods).

**Table 1. Structure of Pedigrees**

Number of pedigrees	32
Number of examined individuals	579
Median number (range) of individuals per pedigree	12.5 (2–50)
Number of sibships	111 <sup>a</sup>
Median number (range) of individuals per sibship	3 (2–10)
Number of sibpairs	620 <sup>a</sup>

<sup>a</sup>103 sibships yielding 545 sibpairs were included in the marker subset (see text).

	Nondiabetics		Diabetics	
	men	women	men	women
Number	182	257	58	82
Age (years)	39.2 ± 17.1	38.4 ± 14.8	55.2 ± 13.4	57.2 ± 14.9
BMI (kg/m <sup>2</sup> )	29.1 ± 5.8	29.9 ± 6.9	31.0 ± 5.8	32.6 ± 6.8
Fasting glucose (mg/dl)	90.5 ± 11.6	87.9 ± 10.7	174.2 ± 62.9	186.2 ± 78.4
2-h glucose (mg/dl)	103.9 ± 34.3	110.7 ± 30.3	318.2 ± 96.5	324.8 ± 119.8

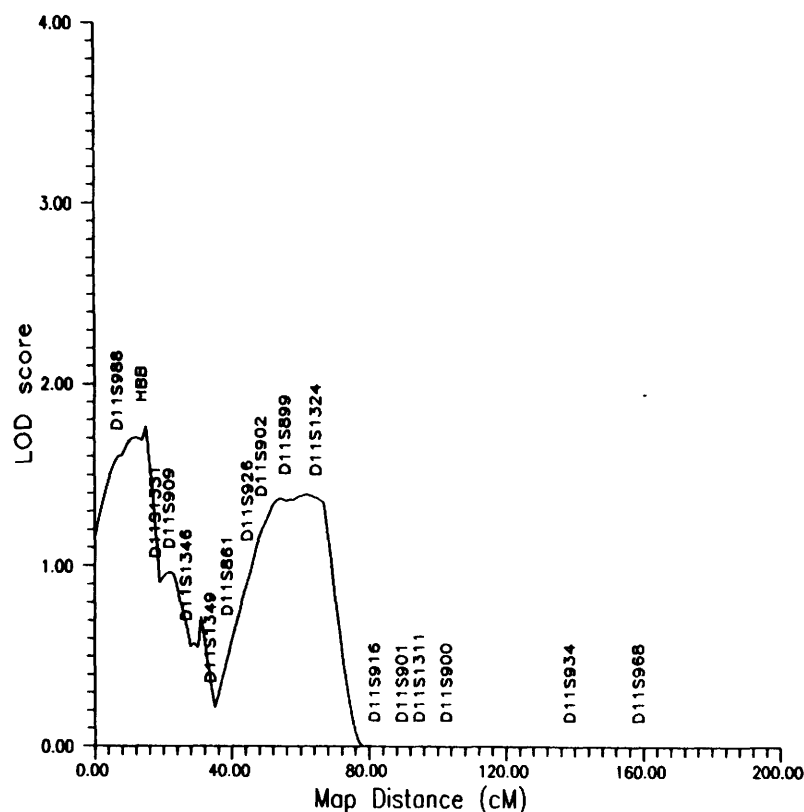
## RESULTS

For all traits we first compared a “sporadic” model, which assumes no familial resemblance beyond that expected by chance, with a “polygenic” model, which allows for additive genetic and shared environmental effects. For both fasting glucose and 2-hr glucose the sporadic models were strongly rejected ( $\chi^2 = 10.33$ ,  $df = 1$ ,  $P = 0.0013$  and  $\chi^2 = 5.61$ ,  $df = 1$ ,  $P = 0.0178$ , re-

spectively) indicating familial aggregation of plasma glucose concentrations. We next computed lod scores for linkage by comparing the multipoint linkage model with the polygenic model.

For plasma glucose concentration, lod scores  $>2.0$  were observed only on chromosomes 6 and 11. Figures 1 and 2 show the lod scores in favor of linkage for fasting and 2-hr glucose concentrations as a function of map distances on chromosome 11. Figure 1 shows two lod score peaks for fasting glucose on chromosome 11p. The first occurs at the  $\beta$ -hemoglobin locus (HBB, 11p15.4), 15.3 cM from the p-terminal end, and the second between D11S899 and D11S1324, 62.0 cM from the p-terminal end of chromosome 11. The lod scores for these peaks are 1.76 and 1.39, respectively. The same two peaks are seen for 2-hr glucose (Fig. 2), and here the lod scores are substantially higher at 2.77 and 3.37, respectively.

Figures 3 and 4 show the lod scores in favor of linkage for fasting and 2-hr glucose concentrations as a function of map distances on chromosome 6. Fasting glucose shows three peaks, two of which have lod scores  $>2.0$  (D6S300, lod score 2.17 and 80.5 cM from the p-terminal end of the chromosome and D6S290, lod score 2.15 and 167.4 from p-ter) (Fig. 3). Lower lod scores and a somewhat different profile are seen for 2-hr glucose (Fig. 4), but the peak at D6S290 is still present, albeit with a lower lod score of 1.50.



**Figure 1** lod scores as a function of map positions for fasting plasma glucose concentrations on chromosome 11. A total of 22 markers were typed. Only 17 are shown to avoid cluttering.

## PLASMA GLUCOSE LINKED TO CHROMOSOMES 6 AND 11

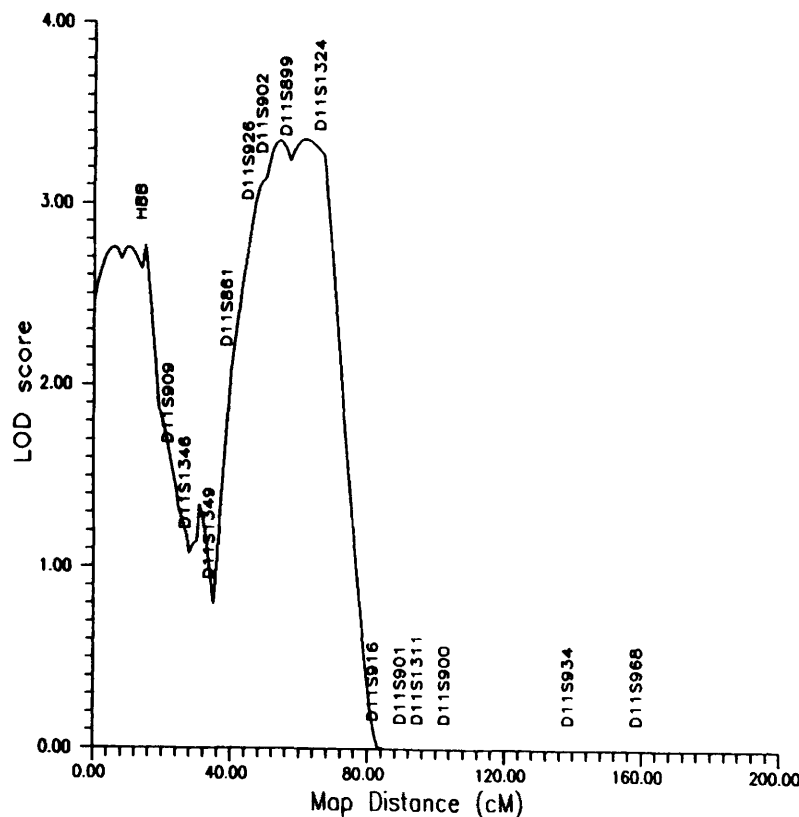
We also tested for evidence of linkage between chromosomes 6 and 11 and the following additional phenotypes: log of fasting and 2-hr glucose concentrations; fasting and 2-hr insulin concentrations; and body mass index [weight (kg)/height squared ( $m^2$ )]. The log of fasting and 2-hr glucose gave lod score profiles similar to those shown in Figures 1–4, albeit with lower lod scores. No evidence for linkage was observed on chromosomes 6 and 11 for the other phenotypes.

To assess whether the linkages shown in Figures 1–4 affect diabetes per se, and not merely quantitative variation in normal glucose levels, we repeated the analyses with the 83 diabetic subjects excluded. This reduced the number of sibpairs from 545 to 343. We also performed analyses in which we excluded the 42 diabetics who were under treatment with either oral agents or insulin. This reduced the number of sibpairs from 545 to 443. The results of these analyses are presented in Table 3. Exclusion of all diabetics completely obliterated all signals, suggesting that the observed linkages are with the trait diabetes

and not with quantitative variation in normal glucose levels. Excluding treated diabetics weakened the evidence for linkage with the HBB locus. The evidence for linkage to the region between D11S899 and D11S1324 persisted, strengthened in the case of fasting glucose and weakened in the case of 2-hr glucose. On chromosome 6, the evidence for linkage remained strongest for fasting glucose. The evidence for linkage actually became stronger when treated diabetics were excluded, although the peaks are shifted somewhat. The location of the fasting peak shifted from D6S300 at 80.5 cM to D6S257 at 57.2 cM, and the lod score increased from 2.17 to 2.78. The location of the second peak shifted from D6S290 at 167.4 cM to D6S1035 at 178.4 cM, and the lod score increased from 2.15 to 3.14.

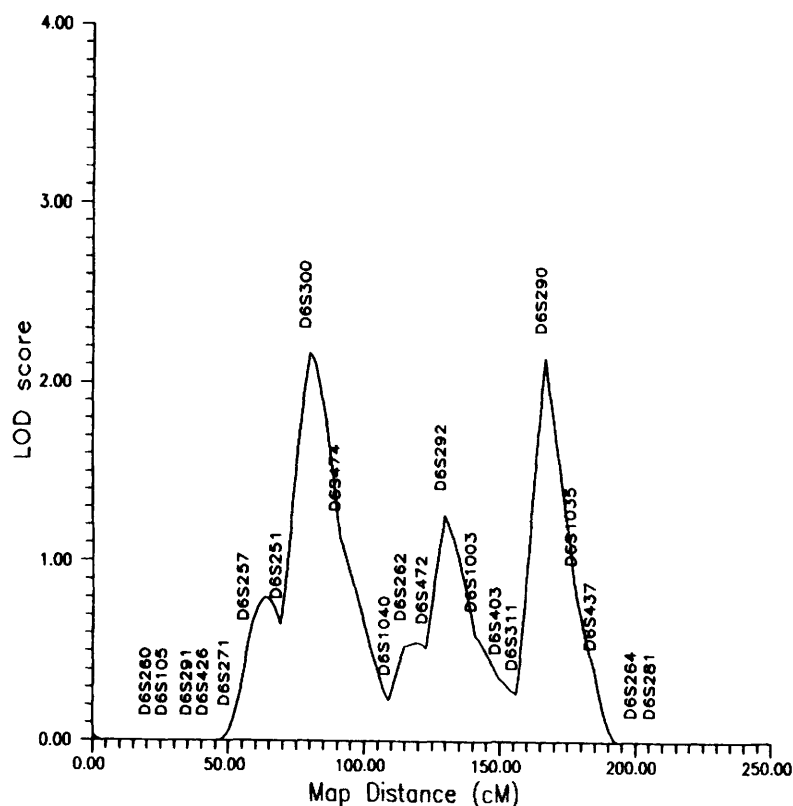
The number of lod score peaks observed for each trait suggested the possibility of oligogenic inheritance. To examine this possibility we performed a sequential oligogenic analysis by conditioning on the QTL exhibiting the maximal lod score and then performing two-locus linkage analyses in which the resultant profile lod function was evaluated (see Methods for details). These analyses were performed with all diabetics included. For 2-hr glucose, these analyses revealed that only the D11S899/D11S1324 peak at 62.0 cM on chromosome 11 (lod score, 3.37) was significant. All other peaks (including those on chromosome 6) were effectively eliminated. For example, the HBB lod score on chromosome 11 was reduced from its unconditional value of 2.77 to a conditional value of only 0.48. The percentage of phenotypic variance in 2-hr glucose levels that was attributable to the D11S899/D11S1324 peak was  $44.7 \pm 12.5\%$  (nominal  $P$  value of  $4.1 \times 10^{-5}$ ). The maximum likelihood estimate of the location of this QTL is at 61.0 cM with a 95% confidence interval from 43.5 to 70.5 cM. Two-locus analysis also failed to reveal any evidence for epistatic interactions among loci.

Oligogenic analysis of fasting glucose levels was less conclusive. Conditioning on either of the two major peaks on chromosome 6 re-



**Figure 2** lod scores as a function of map positions for 2-hr plasma glucose concentrations on chromosome 11. A total of 22 markers were typed. Only 17 are shown to avoid cluttering.

STERN ET AL.



**Figure 3** lod scores as a function of map positions for fasting plasma glucose concentrations on chromosome 6. A total of 23 markers were typed. Only 21 are shown to avoid cluttering.

duced the other, although both retained nominal significance. All other fasting glucose peaks on chromosomes 6 and 11 were eliminated, however. The percentage of phenotypic variance in fasting glucose levels attributable to the putative locus at D6S300 was estimated to be  $30.1 \pm 12.3\%$  ( $P = 0.0075$ ), and the peak at D6S290 accounted for an additional  $29.6 \pm 11.5\%$  ( $P = 0.0079$ ) of the phenotypic variance. Allowing for epistasis appeared to significantly improve this two-locus model ( $P = 0.023$ ) with most of the genetic variance now being estimated in the epistatic component ( $56 \pm 26\%$ ). However, this inference regarding the possible existence of epistasis must be tempered by the relatively large error that is associated with this complex component. It is likely that sibships do not represent the most powerful design for detecting epistasis and that better estimates would be obtainable using extended pedigrees.

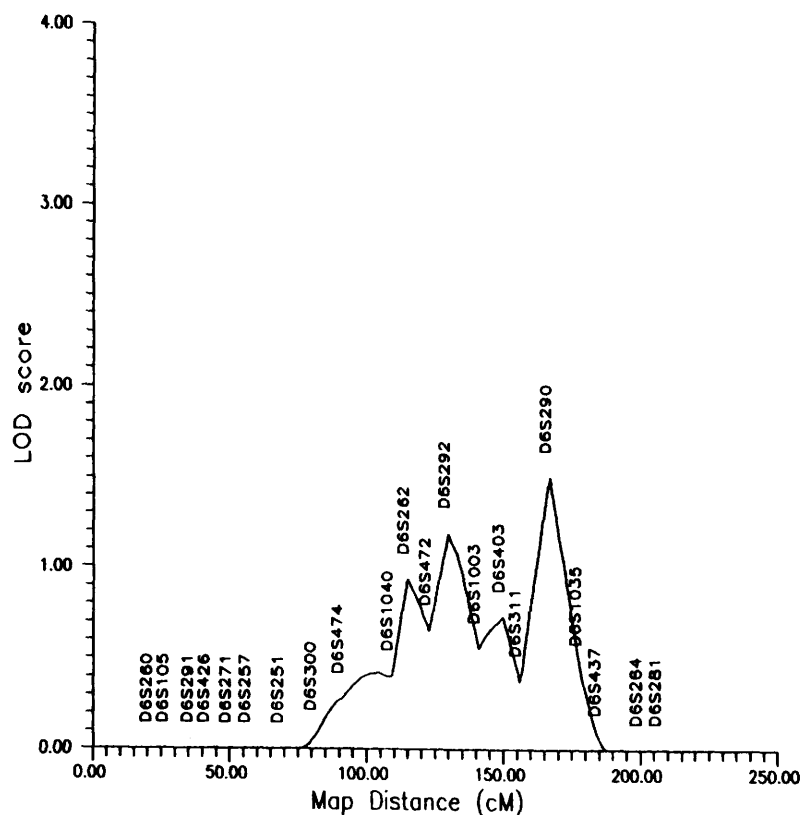
We also tested for linkage using a segregation model that we developed using the program REGTL in SAGE. This model posited a major

gene influencing age of diabetes onset (Stern et al. 1996). The best fitting model was a dominant one in which individuals carrying one or two copies of an early onset allele had an average age of diabetes onset of ~50 years and those without the allele had an average age of diabetes onset of ~80 years. This model was statistically indistinguishable from the general model, whereas a number of non-Mendelian models were rejected. The frequency of the putative early onset allele in this population was estimated to be 25%, and it was further estimated that this major gene accounted for ~70% of the variance in age of diabetes onset. The maximum likelihood parameter estimates of this model were used with the SAGE program LODLINK to test for linkage between the markers shown in Figures 1–4 and the putative diabetes age of onset gene. Only the HBB marker on chromosome 11 and its flanking markers showed positive lod scores. These lod scores are shown in Table 4. A maximum lod score of 2.1 at a recombination fraction of 0.05 was observed for the

marker HBB. Thus, the segregation model-based linkage analysis leaves open the possibility that there is an additional susceptibility locus near the HBB locus on chromosome 11.

## DISCUSSION

A large number of genetic studies are currently under way to identify susceptibility genes for type II diabetes and other complex traits. With so many linkage tests being carried out, the problem of false positives presents a serious and well-recognized challenge. Recently, Lander and Kruglyak (1995) addressed this problem by performing a series of simulations of whole genome scans with multiple tests for linkage. Based on these simulations, they proposed that the term “suggestive linkage” be used for statistical signals, that would be expected to occur by chance not more than once in the course of a whole genome scan. They proposed further that the term “significant linkage” be used for statistical signals that would be expected to occur by chance only



**Figure 4** lod scores as a function of map positions for 2-hr plasma glucose concentrations on chromosome 6. A total of 23 markers were typed. Only 21 are shown to avoid cluttering.

5% of the time in the course of a whole genome scan. Based on their simulation data, they estimate that for allele-sharing methods such as we have used, the lod scores corresponding to suggestive linkage would range from 1.9 to 2.4 and the lod scores corresponding to significant linkage would range from 3.3 to 3.8.

It is clear that the linkage signals that we have observed on chromosomes 6 and 11 depend on the presence of diabetics in the analyses, suggesting that these linkages are with the trait diabetes and not with quantitative variation in normal glucose levels. It is less clear whether it is preferable to include or exclude treated diabetics. Obviously, excluding the 42 of 83 diabetics who were under treatment reduces the statistical power, which may account for the weakening of some of our signals. On the other hand, all but two of these treated diabetics had glucose values above the normal range, that is, higher than all of the nondiabetic subjects. Thus, these high values contribute useful information in the linkage analysis, even though they are lower than they

would have been, without treatment, thereby distorting to some degree the natural glucose distribution.

Taking all of the above into consideration, we feel the best evidence for linkage is between 2-hr glucose concentration and the D11S899/D11S1324 region on chromosome 11 and between fasting glucose and the D6S290/D6S1035 region on chromosome 6.

In addition to purely statistical arguments, there are data from an animal model of type II diabetes that support our chromosome 11 linkage results. Two recent reports have characterized QTLs affecting diabetes-related traits in  $F_2$  cohorts bred from the spontaneously diabetic Goto-Kakizaki (GK) rat (Galli et al. 1996; Gauguier et al. 1996). As in our data, these studies detected separate loci affecting the fasting and the 2-hr plasma glucose concentration. Both of the GK rat studies indicated that a region on rat chromosome 1 was linked to post-load plasma glucose concentration. Interestingly, genes in this area show homology to the region on

human chromosome 11p (Levan et al. 1986; Hoovers et al. 1995), for which our results implicate a QTL for postload glucose levels. The GK rat results also suggested linkage between fasting glucose and regions on rat chromosomes 2, 10, and 17. Unfortunately, the rat-human homology relationships for the regions on human chromosome 6, for which our data implicate QTLs for fasting glucose concentration, are not presently known.

It is also of considerable interest that the region marked by D6S290 has been linked to type I diabetes (Davies et al. 1994; Todd 1995). This susceptibility site has been labeled IDDM5. Although the pathogenesis of type I and II diabetes are clearly different, the former having an immune basis and the latter not, they could, nevertheless, share certain susceptibility loci. Thus, although the primary susceptibility site for type I diabetes is in the HLA region on chromosome 6p (Davies et al. 1994; Todd 1995), additional loci, including some that might principally confer susceptibility to type II diabetes, could, nevertheless,

STERN ET AL.

**Table 3. Comparison of Multipoint lod Scores with All Diabetics or Treated Diabetics Excluded**

	Marker	Map position (cM)	lod Scores		
			all diabetics included	only untreated diabetics included	all diabetics excluded
<i>Chromosome 11</i>					
Fasting glucose	HBB	15.3	1.76	1.14	0
	D11S1324	66.5	1.39	2.15	0
2-h glucose	HBB	15.3	2.77	0.15	0
	D11S899/ D11S1324	62.0	3.37	1.90	0
<i>Chromosome 6</i>					
Fasting glucose	D6S257	57.2	0.53	2.78	0
	D6S300 <sup>a</sup>	80.5	2.17	0.86	0
	D6S290 <sup>a</sup>	167.4	2.15	2.74	0
	D6S1035	178.4	0.85	3.14	0
2-h glucose	D6S290 <sup>a</sup>	167.4	1.50	0.90	0
	D6S1035	178.4	0.43	1.20	0

<sup>a</sup>Original peaks.

contribute additional vulnerability to type I. Another susceptibility site for type I diabetes (IDDM2) has been mapped to the 5' regulatory region of the insulin gene on chromosome 11p15. However, our signal on 11p does not appear to be attributable to the insulin gene, as HBB is 12.2 cM centromeric to the insulin gene and the D11S899/D11S1324 peak is 57.9 cM centromeric to the insulin gene.

Elbein and coworkers (1996) tested for linkage between diabetes and 23 markers on chromosome 11 using a variety of model-based approaches and also the nonparametric, affected pedigree member (APM) method. They found no evidence for linkage to either the HBB gene or the

region where our main peak was observed. They did, however, find weak and somewhat inconsistent evidence of linkage to D11S916 and D11S901, at 82.8 and 91.1 cM, respectively. In the present study, the lod scores at these two loci were zero (see Figs. 1 and 2). Two important differences between the study by Elbein et al. and ours is that their study was carried out in a population of Northern European descent, and their pedigrees were ascertained on sibships containing at least two individuals with diabetes. Not only might populations of Northern European origin have different diabetes susceptibility genes than Mexican Americans, but ascertainment of multiplex families may identify pedigrees with

**Table 4. lod Scores Between Markers on Chromosomes 11 and a Putative Major Gene Influencing Diabetes Age of Onset as Specified by a Segregation Model**

Marker	Map position (cM)	Number of individuals genotyped	lod scores at recombination fractions of:				
			0.00	0.01	0.5	0.10	0.20
D11S1984	0	400	-0.70	-0.56	-0.16	0.16	0.37
D11S988	8.3	395	0.39	0.50	0.80	0.97	0.90
HBB	15.3	433	1.99	2.03	2.09	1.99	1.48
D11S1331	19.0	393	-0.89	-0.79	-0.48	-0.21	0.06

unique susceptibility loci not found, or found with lower frequency, in the general population.

Recently, Hanis et al. (1996) reported evidence of linkage between type II diabetes and a region at the q-terminal end of chromosome 2 in a population of Mexican Americans from Starr County, Texas, on the U.S.–Mexican border. We have been unable to find linkage between markers in this region, including the marker D2S125 where Hanis et al. observed their strongest signal, and either fasting or 2-hr glucose using the variance component technique. Also, tight linkage between this marker and the early age of diabetes onset gene (Stern et al. 1996) was rejected with lod scores of  $-4.42$  at a recombination fraction of 0 and  $-2.34$  at a recombination fraction of 0.1.

A number of plausible candidate genes are located in the regions where we observed linkage to plasma glucose concentrations. Two recent reports have linked familial persistent hyperinsulinemic hypoglycemia of infancy (PHHI), a disorder of unregulated insulin secretion associated with severe hypoglycemia, to chromosome 11p14–15.1 (Glaser et al. 1994; Thomas et al. 1995a). Thomas et al. (1995a) concluded that the PHHI locus lies in a region between the markers D11S926 and D11S899, that is, the region on chromosome 11p where our strongest 2-hr glucose signal was seen. Recently, the  $\beta$  cell high-affinity sulfonylurea receptor (SUR) gene, a regulator of insulin secretion, has been cloned, which maps to this same chromosomal region (Aguilar-Bryan et al. 1995; Thomas et al. 1995b). It appears that point mutations in SUR are responsible for abnormal insulin secretions in PHHI (Thomas et al. 1995b). Absence of linkage between type II diabetes and the SUR gene has been reported in a Japanese population (Iwasaki et al. 1996). Although this lack of concordance with our results could also reflect population differences, it should be noted that the negative report in Japanese was based on only 140 sibpairs and may, therefore, have lacked statistical power.

A potential candidate gene near D6S290 is the membrane glycoprotein PC-1 gene. It appears that this protein plays a role in the insulin resistance of type II diabetes (Maddux et al. 1995). The PC-1 gene has been localized to 6q22–q23 (Funakoshi et al. 1992).

These studies indicate that despite the genetic complexity underlying type II diabetes in humans, specialized genetic studies can uncover clues to the locations of diabetes susceptibility

genes. In the present case, the Native American heritage common to the individuals under study may limit the genetic heterogeneity that broader-based studies in the United States may encounter. This study also suggests that in addition to highly polymorphic markers and maps, more sophisticated analytic tools to model IBD relationships, such as the multipoint variance components approach, may be helpful in the study of this complex disorder. Finally, these data suggest that the similarities between diabetes-related phenotypes in the GK rat and human type II diabetes may reflect a common underlying genetic basis and that continued study of both systems will prove synergistic in the elucidation of the genetic etiology of the disorder.

## METHODS

### Genotyping of Markers

DNA was extracted from white blood cells using proteinase K digestion/phenol extraction and alcohol precipitation in a semiautomated fashion on an ABI 341 RNA/DNA extractor. Genotyping was done with PCR of locus-specific microsatellite markers. Fifty nanograms of DNA was amplified in 20- $\mu$ l PCR assays containing 20 mM Tris-HCl (pH 8.3), 50 mM KCl, 1.0–2.5 mM MgCl<sub>2</sub>, 10 ng of each PCR primer (one labeled with [ $\gamma$ -<sup>32</sup>P]ATP using polynucleotide kinase), and 0.5 unit of *Taq* polymerase. Reactions were denatured at 94°C; annealing was carried out for 1 min at primer-specific temperatures; and extension was carried out for 1 min at 72°C. Amplification products were mixed with an equal volume of formamide/dye mixture, and 3- $\mu$ l aliquots were separated on a 7% denaturing polyacrylamide gel containing 34% urea and 32% formamide in 1 $\times$  Tris/borate/EDTA running buffer.

Gels were sequentially loaded with four sets of samples at 20 to 40 min intervals. To minimize sample handling errors, all steps of sample handling, PCR, and gel loading used eight channel tools and 96-well plates. Gels were fixed in 20% methanol/acetic acid, dried, and exposed to x-ray film. The genotypic data were entered into a data base and analyzed for discrepancies (i.e., violations of Mendelian inheritance), using the program INFER (PEDSYS) (Dyke 1993). Discrepancies were checked in the laboratory for mistyping, and markers for discrepant individuals were corrected.

### Multipoint Variance Components Linkage Analysis

A number of general multipoint variance components methods have been proposed that allow for locus-specific effects, residual genetic effects, covariate effects, and random environmental effects (Goldgar 1990; Schork 1993; Amos 1994). Simulation results (Goldgar 1990; Schork 1993; Blangero 1995a,b) have revealed that these methods are more powerful than the sibpair test of Haseman and

## STERN ET AL.

Elston (1972). In addition, the method is less prone to false positives, as it partially accounts for the well-known problem of nonindependence among sibpairs (Wilson and Elston 1993; Blangero 1995a,b). When we performed Hase-man–Elston sibpairs analyses on the markers implicated in this paper, the *P* values ranged from 0.01 to 0.24, but a consistent pattern did not emerge. Another major benefit of the variance component method is that it provides an unbiased estimate of the relative variance in a trait that is attributable to an underlying QTL (Blangero 1995b). Thus, it provides a useful metric with regard to the genetic effect size of a QTL.

Our approach is an extension of a method developed by Amos (1994), which is based on specifying the expected genetic covariances between relatives as a function of their IBD relationships at a marker locus (which is hypothesized to be linked to a locus influencing the quantitative trait). The phenotypic covariance matrix for a given sibship ( $\Omega$ ) is given by  $\Omega = \Pi\sigma_m^2 + 2\phi\sigma_g^2 + I\sigma_e^2$ , where  $\Pi$  is a matrix with elements ( $\pi_{mij}$ ) providing the proportion of genes that individuals *i* and *j* share IBD at a marker (*m*) linked to a quantitative trait locus;  $\sigma_m^2$  is the additive genetic variance attributable to the marker locus;  $\phi$  is the kinship matrix;  $\sigma_g^2$  is the variance attributable to residual additive genetic effects; *I* is an identity matrix; and  $\sigma_e^2$  is the variance attributable to random environmental effects. Although, in principle, this method can be used with extended pedigrees, estimating the IBD matrix for all relative pairs in an extended pedigree is highly computer intensive (Blangero 1995b). Therefore, we applied the method to sibships using the SIBPAL program in SAGE to obtain the IBD matrix for sibpairs.

Extensive power analyses using a nominal  $\alpha = 0.05$  cutoff criterion have been performed for this variance components method using computer simulations and the pedigree structures in the present data set. The results indicate excellent power to detect quantitative trait locus effects (e.g., >80% power to detect effects as small as those accounting for only 15% of the phenotypic variance) (J. Blangero, unpublished data).

Both variance components and covariate effects were estimated simultaneously by maximum likelihood techniques. A likelihood function based on multivariate normal density was numerically maximized to obtain the following nine parameter estimates: (1)  $\mu$  = the phenotypic mean; (2)  $\beta(\text{sex})$  = the regression coefficient for sex; (3)  $\beta(\text{ageM})$  = the regression coefficient for age in males; (4)  $\beta(\text{ageF})$  = the regression coefficient for age in females; (5)  $\beta(\text{age}^2\text{M})$  = the regression coefficient for age squared in males; (6)  $\beta(\text{age}^2\text{F})$  = the regression coefficient for age squared in females; (7)  $\sigma$  = the phenotypic standard deviation; (8)  $h^2$  = the genetic variance attributable to residual additive polygenic effects; and (9)  $h_m^2$  = the genetic variance attributable to the marker locus linked to a locus influencing the quantitative trait.

Maximum likelihood estimates that assume multivariate normality can be shown to yield consistent parameter estimates even when the distributional assumptions are violated. The estimated error covariance matrix, however, may be biased if normality does not hold. To obviate this difficulty, we calculated robust standard errors for these estimators by a modification of the covariance matrix of the parameters involving specification of the first two moments of the underlying distribution (Beatty et al.

1985). The hypothesis of no familial aggregation was tested by comparing the likelihood of a model in which both parameters  $h^2$  and  $h_m^2$  were constrained to values of 0 with a model in which the parameter  $h^2$  was estimated. The hypothesis of no linkage was tested by comparing the likelihood of a model in which only the parameter  $h_m^2$  was constrained to a value of 0 with a model in which this parameter was estimated. Twice the difference in the log likelihoods of these models yields a test statistic that is asymptotically distributed, approximating a  $\chi^2$  distribution [as a 1/2:1/2 mixture of a  $\chi^2$  and a point mass at 0 (Hopper and Mathews 1982; Self and Liang 1987)]. The degrees of freedom are equal to the difference in the number of parameters estimated in the two competing models. We have implemented the robust variance components linkage procedure using the computer program FISHER (Lange et al. 1988).

Our multipoint approach is an extension of the multipoint mapping strategy proposed by Fulker et al. (1995). The method uses information on all available markers separated by known map distances. Through standard regression procedures, the proportion of allele-shared IBD for a given chromosome location is estimated. Because the method requires knowledge of the distances between markers, we used the distances reported in the GDB (Fasman et al. 1994). Markers that were included in our data set, but are not mapped in the GDB, were placed on the map using the program CRI-MAP (P. Green, K. Falls, and S. Crooks, unpubl.), which employs Kosambi's mapping function. The variance components method was applied at 1-cM intervals. lod scores were obtained by converting the *ln* likelihood values into values of log to the base 10, and lod scores were plotted against map distances.

When the multipoint variance components analysis indicated multiple chromosomal regions with lod scores >2.0, we performed an oligogenic variance components analysis utilizing all suggestive markers. Joint analysis of these suggestive loci has two main benefits: (1) It helps to estimate false positives, as joint estimation can be shown to yield more accurate estimates of effect size (therefore, false linkages tend to be diminished); and (2) it increases the power to detect linkages, as conditional testing using profile lod scores has the effect of maximizing the relative signal-to-noise ratio of true linkages. We carried out conditional linkage tests as follows: Given a set of *n* markers whose marginal analysis is suggestive of linkage, we performed a multiple locus analysis in which a relative variance was estimated simultaneously for each underlying QTL. To obtain profile LOD scores, we then sequentially deleted the focal marker to yield a null model with *n* – 1 estimated QTL variance terms. The likelihood for this nested model was then compared to the general model, and the conditional (or profile) lod score was computed. Our experience with this approach shows that both types of results (the elimination of false positives and the increased signal of true linkages) can occur.

Our oligogenic models generally assume additivity. However, when two or more loci continued to suggest linkage with QTLs after conditional testing, we additionally allowed for epistasis (gene–gene interaction). Specifically, we tested for additive  $\times$  additive epistasis by defining a new variance component whose structuring coefficient matrix consisted of the Hadamard product of the two constituent markers' IBD probability matrices. The signifi-

cance of the epistatic component of variance was evaluated using a likelihood ratio test.

## ACKNOWLEDGMENTS

Some of the results reported in this paper were obtained using the program package S.A.G.E., which is supported by a U.S. Public Health Service Resource Grant (1 P41 RR03655) from the Division of Research Resources. This research was supported by grants from The National Institute of Diabetes, Digestive, and Kidney Diseases (R01 DK 42273, R01 DK 47482, and R29 DK 44297).

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

## REFERENCES

- Aguilar-Bryan, L., C.G. Nichols, S.W. Wechsler, J.P. Clement, A.E. Boyd, G. Gonzalez, H. Herrera-Sosa, K. Nguy, J. Bryan, and D.A. Nelson. 1995. Cloning of the  $\beta$  cell high-affinity sulfonylurea receptor: A regulator of insulin secretion. *Science* **268**: 423–426.
- Amos, C.I. 1994. Robust variance-components approach for assessing genetic linkage in pedigrees. *Am. J. Hum. Genet.* **54**: 535–543.
- Beaty, T.H., S.G. Self, K.G. Liang, M.A. Connolly, G.A. Chase, and P.O. Kwitrovich. 1985. Use of robust variance components models to analyze triglyceride data in families. *Ann. Hum. Genet.* **49**: 315–338.
- Blangero, J. 1995a. Genetic analysis of a common oligogenic trait with quantitative correlates: Summary of GAW9 results. *Genet. Epidemiol.* **12**: 689–706.
- . 1995b. Multivariate oligogenic linkage analysis of quantitative traits in general pedigrees. *Am. J. Hum. Genet.* **57**: A11.
- Chakraborty, R., R.E. Ferrell, M.P. Stern, S.M. Haffner, H.P. Hazuda, and M. Rosenthal. 1986. 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.
- Davies, J.L., Y. Kawaguchi, S.T. Bennett, J.B. Copeman, H.J. Cordell, L.E. Pritchard, P.W. Reed, S.C. Gough, S.C. Jenkins, S.M. Palmer, K.M. Balfour, B.R. Rowe, M. Farrall, A.H. Barnett, S.C. Bain, and J.A. Todd. 1994. A genome-wide search for human type 1 diabetes susceptibility genes. *Nature* **371**: 130–136.
- Dyke, B. *PEDSYS PGL Technical Report, no. 2*. 1993. Southwest Foundation for Biomedical Research. San Antonio, TX.
- Economou, E.P., A.W. Bergen, A.C. Warren, and S.E. Antonarakis. 1990. The polydeoxyadenylate tract of Alu repetitive elements is polymorphic in the human genome. *Proc. Natl. Acad. Sci.* **87**: 2951–2954.
- Elbein, S.C., K.L. Bragg, M.D. Hoffman, R.A. Mayorga, and M.F. Leppert. 1996. Linkage studies of NIDDM with 23 chromosome 11 markers in a sample of whites of northern European descent. *Diabetes* **45**: 370–375.
- Fasman, K.H., A.J. Cuticchia, and D.T. Kingsbury. The GDB™ Human Genome Data Base 1994. *Nucleic Acids Res.* **22**: 3462–3469.
- Fulker, D.W., S.S. Cherny, and L.R. Cardon. 1995. Multipoint interval mapping of quantitative trait loci using sib pairs. *Am. J. Hum. Genet.* **56**: 1224–1233.
- Funakoshi, I., H. Kato, K. Horie, T. Yano, Y. Hori, H. Kobayashi, T. Inoue, H. Suzuki, S. Fukui, M. Tsukahara, T. Kajii, and I. Yamashina. 1992. Molecular cloning of cDNAs for human fibroblast nucleotide pyrophosphatase. *Arch. Biochem. Biophys.* **295**: 180–187.
- Galli, J., L. Lou-sheng, A. Glasner, C.-G. Ostenson, H. Jiao, H. Fakhrai-Rad, H.J. Jacob, E.S. Lander, and H. Luthman. 1996. Genetic analysis of the non-insulin dependent diabetes mellitus of the GK rat. *Nature Genet.* **12**: 31–37.
- Gardner, L.I., M.P. Stern, S.M. Haffner, S.P. Gaskill, H.P. Hazuda, J.H. Relethford, and C.W. Eifler. 1984. Prevalence of diabetes in Mexican Americans: Relationship to percent of gene pool derived from Native American sources. *Diabetes* **33**: 86–92.
- Gauguier, D., P. Froguel, V. Parent, C. Bernand, M.-T. Bihoreau, B. Portha, M.R. James, L. Penicuaud, M. Lathrop, and A. Ktorza. 1996. Chromosomal mapping of genetic loci associated with non-insulin dependent diabetes in the GK rat. *Nature Genet.* **12**: 38–43.
- Glaser, B., K.C. Chiu, R. Anker, A. Nestorowicz, H. Landau, H. Ben-Bassat, Z. Shlomai, N. Kaiser, P.S. Thorton, C.A. Stanley, R.S. Spielman, K. Gogolin-Ewens, E. Cerasi, L. Baker, J. Rice, H. Donis-Keller, and M.A. Permutt. 1994. Familial hyperinsulinism maps to chromosome 11p14-15.1, 30 cM centromeric to the insulin gene. *Nature Genet.* **7**: 185–188.
- Goldgar, D.E. 1990. Multipoint analysis of human quantitative genetic variation. *Am. J. Hum. Genet.* **47**: 957–967.
- Haffner, S.M., M.P. Stern, H.P. Hazuda, J.A. Pugh, and J.K. Patterson. 1986. Hyperinsulinemia in a population at high risk for non-insulin dependent mellitus. *N.Engl. J. Med.* **315**: 220–224.
- Hanis, C.L., E. Boerwinkle, R. Chakraborty, D.L. Ellsworth, P. Concannon, B. Stirling, V.A. Morrison, B. Wapelhorst, R.S. Spielman, K.J. Gogolin-Ewens, et al. 1996. Genome-wide search for human non-insulin-dependent (type 2) diabetes genes reveals a major susceptibility locus on chromosome 2. *Nature Genet.* **13**: 161–167.
- Haseman, J.K. and R.C. Elston. 1972. The investigation

## STERN ET AL.

of linkage between a quantitative trait and a marker locus. *Behav. Genet.* **2**: 3–19.

Hoovers, J.M., E. Redeker, F. Speleman, J.W. Hoppener, S. Bholra, J. Blied, N. van Roy, N.J. Leschot, N.J. Westerwald, and M. Marmens. 1995. High Resolution chromosome localization of the calcitonin/CGRP/IAPP gene family members. *Genomics* **15**: 525–529.

Hopper, J.L. and J.D. Mathews. 1982. Extensions to multivariate normal models for pedigree analysis. *Ann. Hum. Genet.* **46**: 373–383.

Iwasaki, N., M. Kawamura, K. Yamagata, N.J. Cox, S. Karibe, H. Ohgawara, N. Inagaki, S. Seino, G.I. Bell, and Y. Omori. 1996. Identification of microsatellite markers near the human genes encoding the  $\beta$ -cell ATP-sensitive  $K^+$  channel and linkage studies with NIDDM in Japanese. *Diabetes* **45**: 267–269.

King, H. and M. Rewers. WHO Ad Hoc Diabetes Reporting Group. 1993. Global estimates for prevalence of diabetes mellitus and impaired glucose tolerance in adults. *Diabetes Care* **16**: 157–177.

Knowler, W.C., D.J. Pettitt, M.F. Saad, and P.H. Bennett. 1990. Diabetes mellitus in the Pima Indians: Incidence, risk factors, and pathogenesis. *Diabetes Metab. Rev.* **6**: 1–27.

Lander, E. and L. Kruglyak. 1995. Genetic dissection of complex traits: Guidelines for interpreting and reporting linkage results. *Nature Genet.* **11**: 241–247.

Lange, K., D. Weeks, and M. Boehnke. 1988. Programs for pedigree analysis: MENDEL, FISHER, and dGENE. *Genet. Epidemiol.* **5**: 471–472.

Levan, G., C. Szpirer, and M.C. Yoshida. 1986. Present status of chromosome localization of rat genes. *Rat News Lett.* **17**: 3–8.

Maddux, B.A., P. Sbraccia, S. Kumakura, S. Sasson, J. Youngren, A. Fisher, S. Spencer, A. Grupe, W. Henzel, T.A. Stewart, G.M. Reaven, and I.D. Goldfine. 1995. Membrane glycoprotein PC-1 and insulin resistance in non-insulin-dependent diabetes mellitus. *Nature* **373**: 448–451.

Saiki, R., S. Scharf, F. Taloona, K. Mullis, G. Horn, H. Erlich, and N. Arnheim. 1985. Enzymatic amplification of  $\beta$ -globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. *Science* **230**: 1150–1154.

Schork, N.J. 1993. Extended multipoint identity-by-descent analysis of human quantitative traits. Efficiency, power, and modeling considerations. *Am. J. Hum. Genet.* **53**: 1306–1309.

Self, S.G. and K.-Y. Liang. 1987. Asymptotic properties of maximum likelihood estimators and likelihood ratio tests under nonstandard conditions. *J. Am. Stat. Assoc.* **82**: 605–610.

Stern, M.P., J.K. Patterson, S.M. Haffner, H.P. Hazuda, and B.D. Mitchell. 1989. Lack of awareness and treatment of hyperlipidemia in type II diabetes in a community survey. *J. Am. Med. Assoc.* **262**: 360.

Stern, M.P. and S.M. Haffner. 1990. Type II diabetes and its complications in Mexican Americans. *Diabetes Metab. Rev.* **6**: 29–45.

Stern, M.P., B.D. Mitchell, J. Blangero, L. Reinhart, C.M. Kammerer, C.R. Harrison, P.A. Shipman, P. O'Connell, M.L. Frazier, and J.W. MacCluer. 1996. Evidence for a major gene for type II diabetes and linkage analyses with selected candidate genes in Mexican Americans. *Diabetes* **45**: 563–568.

Thomas, P.M., G.J. Cote, D.M. Hallman, and P.M. Mathews. 1995a. Homozygosity mapping to chromosome 11p of the gene for familial persistent hyperinsulinemic hypoglycemia of infancy. *Am. J. Hum. Genet.* **56**: 416–421.

Thomas, P.M., G.J. Cote, N. Wohllk, B. Haddad, P.M. Mathew, W. Rabl, L. Aguilar-Bryan, R.F. Gagel, and J. Bryan. 1995b. Mutations in the sulfonylurea receptor gene in familial persistent hyperinsulinemic hypoglycemia of infancy. *Science* **268**: 426–429.

Todd, J.A. 1995. Genetic analysis of type 1 diabetes using whole genome approaches. *Proc. Natl. Acad. Sci.* **92**: 8560–8565.

U.S. Bureau of the Census. 1991. *Current population reports, series P-20, no. 455. The Hispanic Population of the United States: March 1991.* U.S. Government Printing Office, Washington, D.C.

Weber, J. and P.E. May. 1989. Abundant class of human DNA polymorphisms which can be typed using the polymerase chain reaction. *Am. J. Hum. Genet.* **44**: 388–396.

Wilson, A.F. and R.E. Elston. 1993. Statistical validity of the Haseman-Elston sib-pair test in small samples. *Genet. Epidemiol.* **10**: 593–598.

World Health Organization Expert Committee. 1985. *Diabetes mellitus: Report of a WHO study group.* (Tech. Rep. Ser. no. 727) World Health Organization, Geneva, Switzerland.

Received April 19, 1996; accepted in revised form June 26, 1996.