Am. J. Hum. Genet. 63:1130–1138, 1998

An Autosomal Genomic Scan for Loci Linked to Type II Diabetes Mellitus and Body-Mass Index in Pima Indians

Robert L. Hanson,[1] Margaret G. Ehm,[2] David J. Pettitt,[1] Michal Prochazka,[1] D. Bruce Thompson,[1] David Timberlake,[3] Tatiana Foroud,[4] Sayuko Kobes,[1] Leslie Baier,[1] Daniel K. Burns,[2] Laura Almasy,[5] John Blangero,[5] W. Timothy Garvey,[6] Peter H. Bennett,[1] and William C. Knowler[1]

[1]Phoenix Epidemiology and Clinical Research Branch, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Phoenix: [2]Glaxo Wellcome, Inc., Research Triangle Park, NC; [3]Department of Statistical Genetics, Sequana Therapeutics, Inc., La Jolla; [4]Department of Medical Genetics, Indiana University School of Medicine, Indianapolis; [5]Department of Genetics, Southwest Foundation for Biomedical Research, San Antonio; and [6]Department of Medicine, Medical University of South Carolina and the Ralph H. Johnson Veteran's Affairs Medical Center, Charleston

Summary

Genetic factors influence the development of type II diabetes mellitus, but genetic loci for the most common forms of diabetes have not been identified. A genomic scan was conducted to identify loci linked to diabetes and body-mass index (BMI) in Pima Indians, a Native American population with a high prevalence of type II diabetes. Among 264 nuclear families containing 966 siblings, 516 autosomal markers with a median distance between adjacent markers of 6.4 cM were genotyped. Variance-components methods were used to test for linkage with an age-adjusted diabetes score and with BMI. In multipoint analyses, the strongest evidence for linkage with age-adjusted diabetes (LOD = 1.7) was on chromosome 11q, in the region that was also linked most strongly with BMI (LOD = 3.6). Bivariate linkage analyses strongly rejected both the null hypothesis of no linkage with either trait and the null hypothesis of no contribution of the locus to the covariation among the two traits. Sib-pair analyses suggest additional potential diabetes-susceptibility loci on chromosomes 1q and 7q.

Introduction

It is well recognized that type II diabetes mellitus has a substantial genetic component (Barnett et al. 1981; Knowler et al. 1981; Hanson et al. 1995*a*). Genes that predispose to some types of diabetes have been identified; these include several loci for type I diabetes (Davies et al. 1994) and for maturity-onset diabetes of the young (Froguel et al. 1992; Yamagata et al. 1996*a,* 1996*b;* Stoffers et al. 1997). However, the genes that cause the most common forms of diabetes remain unknown, and it is, therefore, likely that additional important diabetes-susceptibility loci remain to be identified. Moreover, the specific risk factors through which such genes influence the development of type II diabetes are also unknown. Obesity, as quantified by body-mass index (BMI) (kg/m^2), is a strong risk factor for type II diabetes (Knowler et al. 1981) and is also likely to have genetic determinants (Price et al. 1994). The present study represents a genomewide search for loci linked to diabetes and BMI in Pima Indians, a Native American population with a high prevalence of type II diabetes and obesity (Bennett et al. 1971; Knowler et al. 1978, 1991).

Received April 28, 1998; accepted for publication July 17, 1998; electronically published August 31, 1998.

Address for correspondence: Dr. Robert L. Hanson, Diabetes and Arthritis Epidemiology Section, National Institute of Diabetes, Digestive and Kidney Diseases, National Institutes of Health, 1550 East Indian School Road, Phoenix, AZ 85014. E-mail: rhanson@phx.niddk.nih.gov

0002-9297/98/6304-0027$02.00

Subjects and Methods

Subjects and Phenotypes

Since 1965, a longitudinal study of diabetes has been conducted among the residents of the Gila River Indian Community in central Arizona, most of whom are Pima or Tohono O'odham Indians (Bennett et al. 1971; Knowler et al. 1978). All individuals who are $\geqslant 5$ years old are invited to participate in a standardized health examination every 2 years. Genealogical information has been collected for all participants, and this allows construction of pedigrees for family and genetic studies. A 75-g orally administered glucose-tolerance test is interpreted according to World Health Organization criteria for the diagnosis of diabetes: a plasma glucose concentration $\geqslant 11.1$ mmol/liter, observed either in the 2-h postload venous plasma (World Health Organization 1985) or in the course of routine medical care (Knowler et al. 1978). Height and weight are measured, with the subject wearing light clothing and no shoes, for calculation of BMI (kg/m^2). In the present analysis, the max-

imum BMI observed in the longitudinal study was used as a measure of susceptibility to obesity; because the validity of BMI as a measure of obesity in young children is not well established, determination of the maximum was restricted to examinations at age >15 years.

A sample of 1,338 individuals who had participated in the longitudinal study was selected for genomic scans for loci linked to type II diabetes and obesity; criteria for inclusion were available DNA and membership in a nuclear family informative for diabetes or its metabolic correlates. The individuals constituted 332 nuclear families in 112 extended pedigrees. The present analyses involved the 264 nuclear families selected to be potentially informative for linkage studies of diabetes (i.e., $\geqslant 1$ sibling pair with $\geqslant 1$ affected sibling). There were 966 offspring in these nuclear families, 667 (69%) of whom had diabetes; among these families, there were 1,766 sib pairs informative for analyses of diabetes and 1,664 sib pairs informative for BMI. (There were more individuals in analyses of diabetes than in analyses of BMI, because a few persons of age <15 years were included.) The mean maximum BMI among offspring was 37.2 (SD 8.0) kg/m^2. The mean age at onset of diabetes among affected offspring was 34.0 (SD 10.6) years, and the mean age at last examination of nondiabetic offspring was 35.5 (SD 11.1) years. A subset of 225 of the individuals in the present analysis have been studied with respect to physiological components of obesity and diabetes, such as energy expenditure, adiposity, and insulin resistance and secretion; linkage analyses of these phenotypes have been reported elsewhere (Norman et al. 1997, 1998; Pratley et al. 1998).

Genotypes

Five hundred three autosomal microsatellite markers were typed in the laboratory of J. Weber, at the Marshfield Medical Research Foundation (Schwengel et al. 1994; Dubovsky et al. 1995). An additional 13 markers were typed at Glaxo-Wellcome. The median distance between adjacent markers was 6.4 cM (range 0–25.6 cM); median heterozygosity was 68%. Genotypes for each marker were assessed by PCR and either fluorescent or radioactive-labeled specific primers. The reproducibility of the genotyping was evaluated in 76 duplicate samples, typed blindly for each marker. The median rate of agreement between duplicate samples was 97%, and no marker had an agreement rate <90%.

Inspection of the pattern of Mendelian errors over all markers was used to confirm that the genetic relationship among samples was compatible with the genealogical information. The distribution of marker alleles shared identical by state for each pair of siblings also was analyzed, to check for consistency (Ehm and Wagner 1998). After incompatible subjects were eliminated, additional Mendelian errors involving individual markers (presumably typing errors) were corrected by an algorithm that identifies family members whose genotypes could be responsible for the incompatibility and that iteratively (and arbitrarily) deletes these until all incompatibilities are resolved. Marker-allele frequencies were estimated on the basis of the genotypes of all remaining individuals.

The CRI-MAP program (Lander and Green 1987) was used to create maps with distances estimated on the basis of meiotic recombinants in the Pima data. In most cases, the order of the markers was taken as that on maps provided by the Marshfield Medical Research Foundation. For chromosomes 1, 4, 7, and 19, the data justified a slightly different marker order.

Linkage Analyses of Quantitative Traits

Linkage analysis of diabetes must account for the age-specific occurrence of the disease. This was accomplished with a cumulative-incidence method, which uses age and affection status to produce an "age-adjusted" diabetes score that can be analyzed as a quantitative trait (Hanson and Knowler 1998). The trait is defined as $Y - CI_x$, where $Y = 1$ if the individual is affected and $Y = 0$ if the individual is unaffected and where CI_x is the population cumulative incidence at age x, which is either the age at onset (for affected individuals) or the age at last examination (for unaffected individuals). The resulting continuous variable was analyzed by the transformation of Therneau et al. (1990), to produce a more symmetric distribution. The cumulative-incidence method is a powerful way to account for variable age at onset in sib-pair analysis, particularly when the gene of interest affects age at onset of the disease (Hanson and Knowler 1998). The natural logarithm of the maximum BMI at age >15 years also was analyzed, as a quantitative trait, and, prior to linkage analysis, was adjusted for age and sex, by linear regression.

For these quantitative traits, linkage analyses were conducted for sibships by means of variance-components methods (Amos 1994). In brief, the method involves fitting a linear "mixed" model, which, in the present study, involved estimating the trait mean (μ) and three components of variance. The variance was partitioned into (*a*) an additive monogenic component linked to the region of interest (σ^2_M), (*b*) a "polygenic" component that incorporates overall familial effects (σ^2_G), and (*c*) an "environmental" component that incorporates effects unique to the individual (σ^2_E). Under the assumption of no recombination between the trait and marker loci, the phenotypic variance-covariance matrix ($\mathbf{\Omega}$) for individuals in a pedigree is $\Omega = \Phi\sigma^2_G + \Pi\sigma^2_M + \mathbf{I}\sigma^2_E$, where $\mathbf{\Phi}$ is a matrix of the expected proportion of alleles shared identical by descent (IBD) (.5 for siblings), $\mathbf{\Pi}$ is a matrix

of the proportion of alleles shared IBD, as estimated on the basis of the genotypic data, and **I** is an identity matrix.

The parameters of these models were estimated, under the assumption that the distribution of the trait was multivariate normal, by maximizing the likelihood over all sibships, by use of the scoring algorithm (Lange et al. 1976). The null hypothesis of no linkage was assessed by comparing the full model to one in which σ^2_M was constrained to equal 0. Twice the difference in the natural logarithm of the likelihood between these two models has a distribution that is a $\frac{1}{2}:\frac{1}{2}$ mixture of a χ^2 variable and a point mass at 0, and the likelihood-ratio test (χ^2 with 1 df) can be used for hypothesis testing (Hopper and Matthews 1982). The LOD score (Z) for variance-component analyses was calculated by dividing the likelihood-ratio test for linkage by $2*\log_e(10)$.

Since the analyses of single traits suggested that BMI and diabetes were linked to the same region, the extent to which the presumed gene in this region influences both traits, a phenomenon defined as "pleiotropy," was assessed. Bivariate linkage analyses therefore were conducted, by covariance-components models (Lange and Boehnke 1983). The parameters of the bivariate model include the mean and variance components of the first trait (μ_x, σ^2_{Gx}, σ^2_{Mx}, and σ^2_{Ex}), the mean and variance components of the second trait (μ_y, σ^2_{Gy}, σ^2_{My}, and σ^2_{Ey}), and the polygenic, monogenic, and environmental components of covariance between both traits (σ_{Gxy}, σ_{Mxy}, and σ_{Exy}). The null hypothesis of no linkage with either phenotype was assessed by comparing the full model to one in which σ^2_{Mx}, σ^2_{My}, and σ_{Mxy} are constrained to equal 0; the resulting χ^2 statistic was assessed on 3 df. The σ_{Mxy} parameter describes the covariance between trait x in one relative and trait y in another relative, as a function of their similarity at the chromosomal location of interest; that is, it reflects the pleiotropic effects that the linked locus has on both phenotypes. One therefore can test for this pleiotropy, by testing the null hypothesis $\sigma_{Mxy} = 0$. This bivariate-linkage method can substantially increase the power to detect genes that act pleiotropically to influence two traits (Almasy et al. 1997). As parameterized here, any potential direct effect of one trait on the other will be included in the "pleiotropic" effect.

Additional Linkage Analyses of Diabetes

To maximize the ability to detect diabetes-susceptibility loci, linkage analyses of diabetes were performed by several additional methods. Affected-sib-pair analyses were conducted by testing the null hypothesis that the mean proportion of alleles shared IBD among affected sib pairs is .5, against the alternative that it is >.5 (Elston 1984). To account for age in these analyses, individuals were considered to be affected only if the age at onset of diabetes was prior to an arbitrary threshold. For the present analyses, age thresholds of <45 years (for 551 sib pairs) and <25 years (for 55 sib pairs) were used. In addition, an analysis comparing sib pairs concordant for diabetes versus discordant sib pairs also was conducted, by means of the Haseman-Elston test (Haseman and Elston 1972; Elston 1984). In this analysis, individuals were considered to be affected if onset was at age <45 years and were considered to be unaffected if they were known to be nondiabetic at age >45 years (even if they subsequently developed diabetes). To account for the lack of independence introduced by the use of multiple sib pairs from the same family, the P value was assessed by use of a modified number of df (Wilson and Elston 1993).

Multipoint Analyses

The method of Fulker et al. (1995) was used to obtain approximate multipoint estimates of IBD. At each chromosomal location, this method estimates the proportion of alleles shared IBD for each sib pair, as a weighted average of the IBD estimates at each individual marker. In the present analyses, estimates of IBD for individual markers were generated by the SIBPAL program (SAGE 1994); missing data were imputed on the basis of flanking markers, prior to full multipoint estimation (Fulker and Cardon 1994). Haldane's (1919) mapping function was used to convert map distances into recombination fractions. This method produces an approximate estimate of the multipoint IBD distribution. Although an exact estimate can be obtained from the Lander-Green algorithm (Lander and Green 1987), as implemented in the MAPMAKER/SIBS program (Kruglyak and Lander 1995), the computational burden for larger sibships becomes excessive, so that 13 individuals have to be deleted from the larger sibships, to accommodate the limitations of the MAPMAKER/SIBS program. Analyses using the IBD estimates derived from these families produced results similar to those derived from the approximate method. However, only the results from the latter are presented, since these allow use of all individuals in the larger sibships. In many situations, multipoint analyses using the proportion of alleles shared IBD capture almost all of the linkage information that is available from use of the full IBD distribution (Fulker and Cherny 1996).

Results

Results of single-marker variance-components analyses giving $Z > 1.18$ ($P < .01$) with either age-adjusted diabetes or BMI are shown in table 1. Four markers, on chromosomes 6, 11, 13, and 14, showed evidence for linkage with age-adjusted diabetes, at $Z > 1.18$. Six

Table 1

Markers with Linkage of $Z > 1.18$ ($P < .01$) to Age-Adjusted Diabetes or BMI

Chromosome and Marker	Distance[a] (cM)	σ^2_M (Z[b]) for Diabetes[c]	BMI[d]
6:			
D6S1009	127.8	.22 (1.84)	.00 (.00)
11:			
D11S2000	108.7	.01 (.00)	.17 (1.19)
D11S1998	127.5	.05 (.07)	.15 (1.19)
D11S4464	136.6	.22 (1.87)	.27 (2.63)
D11S912	143.9	.18 (.70)	.27 (2.08)
13:			
D13S779	93.3	.20 (1.38)	.00 (.00)
14:			
D14S617	89.9	.21 (1.65)	.04 (.06)
16:			
D16S769	45.9	.02 (.02)	.16 (1.24)
D16S753	53.3	.08 (.21)	.16 (1.34)

[a] From the p-terminal end of the chromosome, according to a genetic map derived from the data from the data of the present study.

[b] Calculated on the basis of the likelihood-ratio test

[c] Data adjusted for age and sex, by a cumulative-incidence method.

[d] Data are adjusted for age and sex, by linear regression.

markers, four of which are on chromosome 11q and two of which are on chromosome 16, showed linkage with BMI, at $Z > 1.18$. The same marker, *D11S4464,* showed the strongest evidence for linkage with each trait.

The highest Z value obtained in multipoint variance-components linkage analyses of age-adjusted diabetes and BMI is shown, for each of the 22 autosomal chromosomes, in figure 1. Three regions showed evidence for linkage with age-adjusted diabetes, at $Z > 1.18$ ($P < .01$; table 2). These included regions on chromosome 11q ($Z = 1.7$, $P = .003$), chromosome 6q ($Z = 1.4$, $P = .006$), and chromosome 9q ($Z = 1.2$, $P = .009$). The chromosome 11q region that was linked to diabetes also was strongly linked to BMI ($Z = 3.6$, $P = 2.6 \times 10^{-5}$). No other chromosomal region showed $Z > 1.18$ for BMI, in multipoint analyses.

Since on chromosome 11 there was evidence for linkage to both diabetes and BMI, a multipoint bivariate analysis was conducted to determine the extent to which a single locus may influence both traits (fig. 2). The strongest evidence for linkage was in the region between *D11S4464* and *D11S912*. The bivariate analysis strongly rejected the null hypothesis of no linkage with either phenotype ($Z = 5.0$, two-tailed $P = 1.8 \times 10^{-6}$). The null hypothesis of no pleiotropy also was strongly rejected ($Z = 3.7$, two-tailed $P = 3.4 \times 10^{-5}$), and the value of the pleiotropic covariance between the two phenotypes (σ_{Mxy}) implied that the genetic correlation was 1.0. These findings suggest that the same locus influences both phenotypes (*i.e.,* both a high BMI and an early age at onset of type II diabetes cosegregate with a genetic element in this region).

Additional linkage analyses of diabetes identified another potential diabetes-susceptibility locus, on chromosome 1 (fig. 3). The highest Z scores occurred near *D1S1677* (171 cM), in analyses comparing sib pairs concordant for diabetes versus discordant sib pairs ($Z = 2.5$, $P = .0004$), and near *D1S2127* (192 cM), in analysis of affected sib pairs with onset at age <25 years ($Z = 4.1$, $P = 7.4 \times 10^{-6}$). Analysis of sib pairs affected at age <45 years also suggested a potential diabetes-susceptibility locus, on chromosome 7, near *D7S1799* ($Z = 1.8$, $P = .002$; fig. 4). For chromosome 11q, the analysis of affected sib pairs with onset at age <45 years gave $Z = 0.4$ ($P = .094$), whereas the comparison of concordant and discordant sib pairs gave $Z = 1.9$ ($P = .0014$).

Discussion

The present genomic scan in Pima Indians provides (1) strong evidence that on chromosome 11q there is a locus influencing susceptibility to both obesity and type II diabetes and (2) some evidence that there are additional diabetes-susceptibility loci, on chromosomes 1q and 7q.

The study was conducted in a relatively large number of families from a population with a high prevalence of both diabetes and obesity. The longitudinal data allowed accurate assessment of both age at onset of diabetes and maximum BMI. The families came from a single Native American population in which there is probably less genetic heterogeneity in susceptibility to obesity and type II diabetes than there is in most other populations.

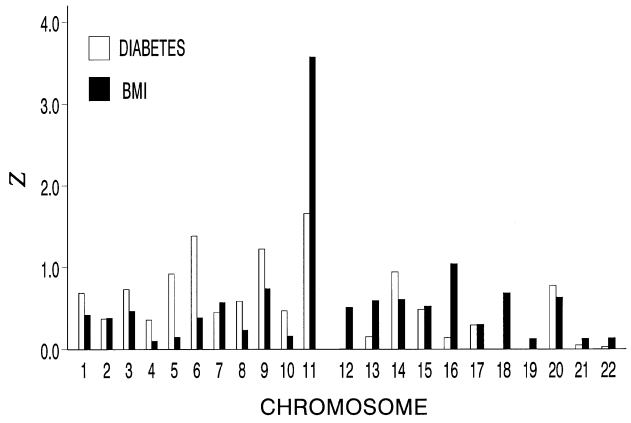

Figure 1 Maximum multipoint Z value, by chromosome, for diabetes and BMI.

The variance-components method used in the present analysis is a powerful tool for assessment of genetic linkage, for quantitative traits. It is often more powerful than sib-pair–based methods, such as the Haseman-Elston test (Amos et al. 1996; Pugh et al. 1997). Although it requires the assumption of multivariate normality, the method is generally robust to violations of this assumption (Beaty et al. 1985; Amos 1994; Amos et al. 1996). In fact, analysis of the present data that uses the Haseman-Elston method, which does not require the assumption of multivariate normality, identified the same regions as being linked with both BMI and diabetes (Hanson et al. 1997; Hanson and Pima Diabetes Genes Group 1997); for chromosome 11 the Z values were 2.0 and 2.4, respectively, for age-adjusted diabetes and BMI.

Criteria for assessment of statistical significance in genetic linkage studies of complex traits have been controversial. Some investigators have suggested that a relatively stringent threshold of $Z > 3.6$ is needed to obtain a genomewide P of <.05 (Lander and Kruglyak 1995), whereas others have maintained that the traditional criterion of $Z > 3.0$ is unlikely to be a false positive (Morton 1998). The issue becomes more complicated when, as in genetic studies in the Pima Indians, several correlated traits have been analyzed, because it is unclear whether—or how—one ought to adjust for multiple comparisons. With these caveats, the present analysis gives both evidence, on chromosome 11, for linkage to BMI ($Z = 3.6$) and strong evidence for linkage to the combined bivariate phenotype of diabetes and BMI ($Z = 5.0$). There is also strong evidence that the BMI locus pleiotropically affects diabetes ($Z = 3.7$). The present results, therefore, imply the existence, on chromosome 11, of a locus influencing susceptibility to obesity and type II diabetes. Statistical tests, however, do not in themselves establish causality, and, ultimately, to distinguish between etiologically important linkage and statistical artifact, the results must be either replicated in other populations or extended in the population that we have studied.

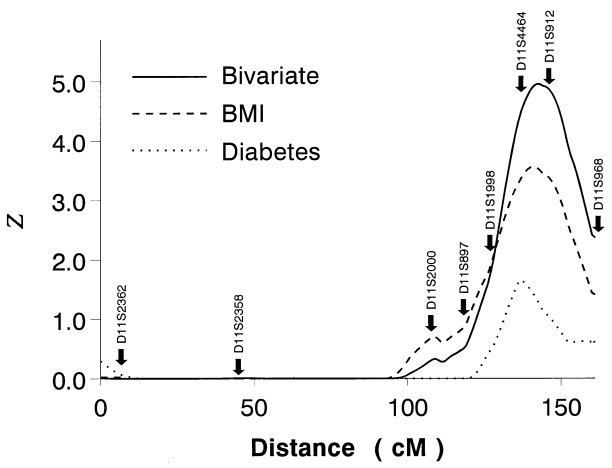

Figure 2 Multipoint results for chromosome 11 for BMI and for diabetes, adjusted for age and sex by a cumulative incidence method. The bivariate-analysis tests the null hypothesis of no linkage with either phenotype. The comparison of the full bivariate model with the model in which there is no linkage is in terms of χ^2 with 3 df; the Z value shown here has been calculated on the basis of the χ^2 (1 df) corresponding to the P value. Distances are from the p-terminal end of the chromosome, on the basis of a genetic map derived from data from the present study.

Current genetic maps (Murray et al. 1994; Dib et al. 1996) place the obesity-diabetes locus identified in the present study at 11q23-25. In a smaller, partially overlapping, sample of Pima Indians, this chromosome 11 region also was linked to 24-h energy expenditure (Norman et al. 1998), and a region ~30 cM centromeric was linked to percentage of body fat (Norman et al. 1997, 1998). It is tempting to speculate that the same locus is

Table 2

Chromosomal Regions with $Z > 1.18$ ($P < .01$) for Linkage with Age-Adjusted Diabetes or BMI, in Multipoint Analyses

Chromosome	Marker (Distance)/Marker (Distance)[a]	Distance[b] (cM)	σ^2_M (Z) for Diabetes[c]	σ^2_M (Z) for BMI[d]
6	D6S1009 (127.8 cM)/D6S1003 (139.1 cM)	128	.21 (1.39)	.00 (.00)
9	D9S299 (100.3 cM)/D9S2026 (107.4 cM)	105	.18 (1.22)	.00 (.00)
11	D11S4464 (136.6 cM)/D11S912 (143.9 cM)	139[e]	.21 (1.66)	.29 (3.57)

[a] Markers are those on either side (p terminal to q terminal) that are closest to the location of the peak mulitpoint Z value for the region; the locations are as determined on the basis of a genetic map derived from the data of the present study

[b] Location of peak multipoint Z value for region.

[c] Data are adjusted for age and sex, by a cumulative-incidence method.

[d] Data are adjusted for age and sex, by linear regression.

[e] The peak on chromosome 11 for diabetes occurred at 137 cM, whereas that for BMI occurred at 141 cM.

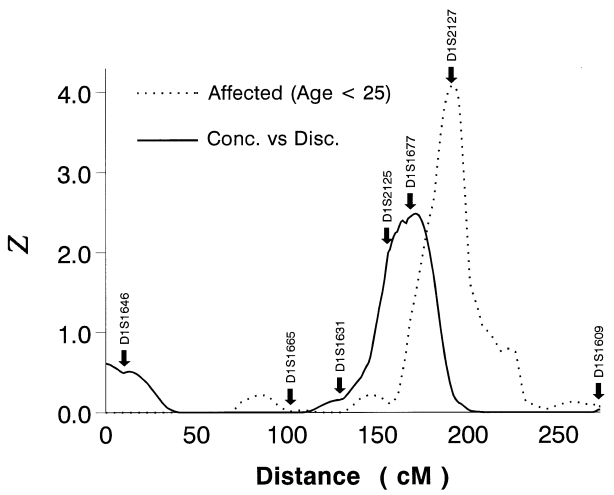

Figure 3 Multipoint results for linkage with diabetes, on chromosome 1. In analyses comparing concordant with discordant sib-pairs (Conclusion versus Discussion), onset was at age <45 years. Distances are from the p-terminal end of the chromosome, on the basis of a genetic map derived from data from the present study. *P* values have been converted to equivalent *Z* values under the assumption of a one-tailed test and a χ^2 distribution (Chotai 1984).

responsible for these results. However, to characterize further the contribution that the locus detected in the present study makes to the etiology of diabetes and obesity, the specific gene or genes responsible for the linkage results need to be identified. The metabolic studies also identified other regions, on chromosomes 1p, 18q, and 20q, as being linked to obesity-related traits, but none of these regions showed evidence for linkage with BMI in the present analyses.

The familial aggregation of diabetes in Pima Indians occurs in a manner that is partially separate from familial aggregation of obesity, and this suggests that there are additional genetic determinants of diabetes, which do not influence obesity (Hanson et al. 1995*b*). In the present analyses, two additional potential diabetes-susceptibility loci were identified, on chromosomes 1q and 7q, in regions that were not linked to obesity. The region on chromosome 1 showed very strong evidence for linkage with diabetes ($Z = 4.1$), in the 55 sib pairs who had onset of diabetes at age <25 years; the evidence was weaker, but still suggestive of linkage, in analyses comparing sib pairs concordant for diabetes versus discordant sib pairs. The analyses of affected sib pairs with onset of diabetes at age <45 years showed modest linkage evidence on chromosome 7 but little evidence on chromosome 1 ($Z = 0.5$). The evidence that chromosomes 1 and 7 are linked to diabetes in Pima Indians is thus weaker than the evidence for an obesity-diabetes locus on chromosome 11. The same chromosome 1 region, however, was linked to diabetes in a European American population in Utah (Elbein et al. 1998); the finding of linkage in two different populations strengthens the evidence for a diabetes-susceptibility locus in this region.

Several other chromosomal regions have been linked to some form of diabetes in other populations. These regions include genes for autosomal dominant diabetes with young age at onset (Bell et al. 1991; Froguel et al. 1992; Vaxillaire et al. 1995; Mahtani et al. 1996; Yamagata et al. 1996*a*, 1996*b*; Stoffers et al. 1997), the human leukocyte-antigen system on chromosome 6p (Tuomilehto-Wolf et al. 1993; Davies et al. 1994), and the *NIDDM1* locus on chromosome 2q (Hanis et al. 1996). The present study, however, shows no evidence for linkage with any of these loci (table 3). These negative findings suggest that these loci do not account for a large portion of the familial aggregation of diabetes in Pima Indians. These loci, however, still could have relatively small effects that are difficult to detect with linkage analysis, or they could provide the genetic "background" for the development of diabetes.

Genes that predispose to type II diabetes may be expected to be linked as well to its risk factors or metabolic constituents, such as obesity, low insulin sensitivity, and defective insulin secretion (Bogardus and Lillioja 1992; Ghosh and Schork 1996). In Pima Indians, there is suggestive evidence for linkage with insulin sensitivity and secretion in other chromosomal locations (Pratley et al. 1998), but there is little evidence for linkage with diabetes in these locations. However, the identification of

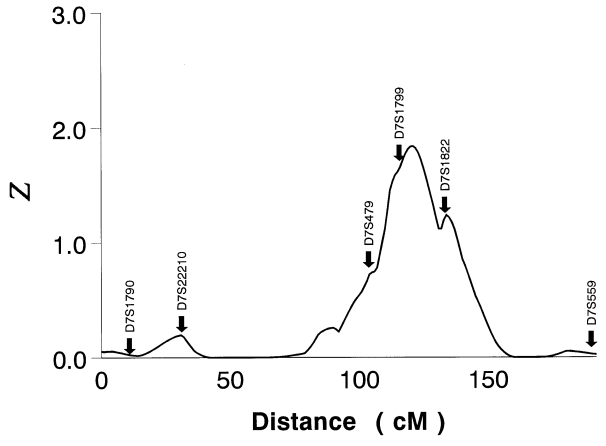

Figure 4 Multipoint results for linkage with diabetes, on chromosome 7. Results are for affected sib-pair analyses for sib-pairs with onset at age <45 years. Distances are from the p-terminal end of the chromosome, on the basis of a genetic map derived from data from the present study. *P* values have been converted to equivalent *Z* values under the assumption of a one-tailed test and a χ^2 distribution (Chotai 1984).

Table 3

Multipoint Sib-Pair Analyses for Selected Regions Linked or Associated with either Diabetes or Glucose Levels in Other Studies

Reference(s)	Chromosome	Distance[a] (cM)	σ^2_M (Z^b) for Diabetes[c]	σ^2_M (Z^b) for BMI[d]
Hanis et al. (1996)	2	247	.06 (.17)	.00 (.00)
Tuomilehto-Wolf et al. (1993), Davies et al. (1994)	6	50	.00 (.00)	.01 (.01)
Froguel et al. (1992)	7	62	.00 (.00)	.00 (.00)
Vaxillaire et al. (1995)	12	113	.00 (.00)	.03 (.07)
Stoffers et al. (1997)	13	34	.01 (.00)	.08 (.25)
Bell et al. (1991)	20	67	.08 (.20)	.05 (.09)

[a] Approximate locations for each candidate region, determined on the basis of nearby markers and available genetic maps.
[b] Calculated by the likelihood-ratio test.
[c] Data are adjusted for age and sex, by a cumulative incidence method.
[d] Data are adjusted for age and sex, by linear regression.

a locus that influences both obesity and diabetes suggests that the diabetogenic effect of this locus may be mediated, in part, through obesity. In fact, in the present study, analysis of the pleiotropic effect of the obesity locus on diabetes resulted in enhanced power to detect the diabetes-susceptibility locus. The present results, therefore, not only provide strong evidence for an obesity-diabetes locus on chromosome 11q but also demonstrate the importance of analyzing other risk factors, concomitantly with the disease itself, in linkage studies of complex diseases.

Acknowledgments

This work was supported partly under the terms of a Collaborative Research and Development Agreement between Glaxo-Wellcome, Plc, and the National Institute of Diabetes and Digestive and Kidney Diseases. Some of the results of this paper were obtained with the program package SAGE, which is supported by U.S. Public Health Service Resource Grant 1 P41 RR03655 from the National Center for Research Resources. W.T.G. is supported partly by National Institutes of Health (NIH) grant DK-47461, and L.A. is supported partly by NIH grant GM-18897. The authors thank the members of the Gila River Indian Community who have participated in these studies. Particular thanks are due to Vickie Andre, Sonja Antone, Linda Phillips, and the rest of the Diabetes and Arthritis Epidemiology Section, for conducting clinical examinations; to Warren Apel, Rachel Janssen, Vicky Ossowski, Nancy Riebow, Jeff Sutherland, Pam Thuillez, and Tina Xi (all of Phoenix), for technical assistance; to Philip Picciotti (Phoenix), Hakan Sakul (Sequana Therapeutics), Tai-He Xia (Glaxo-Wellcome), and Robert C. Williams (Arizona State University), for assistance with data analysis; to Don Holt (Glaxo-Wellcome), for preparation of the genetic maps; to Derek Traugher, Gopi Gopalam, and Madhu Sethi (all of Glaxo-Wellcome), for technical assistance; and to Ray Grimaila, Quan Nguyen, and David Yarnall (all of Glaxo-Wellcome), for selecting and genotyping markers. Additional advice was provided by Rodney Norman, Eric Ravussin, and Richard Pratley (all of Phoenix), Michael Wagner (Glaxo-Wellcome), and Lon Cardon (Sequana Therapeutics). Special thanks are also due to Clifton Bogardus (Phoenix), for advice and encouragement, and to Jeffrey C. Long (National Institute on Alcohol Abuse and Alcoholism), for advice.

References

Almasy L, Dyer TD, Blangero J (1997) Bivariate quantitative trait linkage analysis: pleiotropy versus coincident linkage. Genet Epidemiol 14:953–958

Amos CI (1994) Robust variance-components approach for assessing genetic linkage in pedigrees. Am J Hum Genet 54: 535–543

Amos CI, Zhu DK, Boerwinkle E (1996) Assessing genetic linkage and association with robust components of variance approaches. Ann Hum Genet 60:143–160

Barnett AH, Eff C, Leslie RDG, Pyke DA (1981) Diabetes in identical twins: a study of 200 pairs. Diabetologia 20:87–93

Beaty TH, Self SG, Liang KY, Connoly MA, Chase GA, Kwiterovich PO (1985) Use of robust variance components models to analyse triglyceride data in families. Ann Hum Genet 49:315–328

Bell GI, Xiang KS, Newman MV, Wu SH, Wright LG, Fajans SJ, Spielman RS, et al (1991) Gene for non-insulin-dependent diabetes mellitus (maturity-onset diabetes of the young subtype) is linked to DNA polymorphism on human chromosome 20q. Proc Natl Acad Sci USA 88:1484–1488

Bennett PH, Burch TA, Miller M (1971) Diabetes mellitus in American (Pima) Indians. Lancet 2:125–128

Bogardus C, Lillioja S (1992) Pima Indians as a model to study the genetics of NIDDM. J Cell Biochem 48:337–343

Chotai J (1984) On the lod score method in linkage analysis. Ann Hum Genet 48:359–378

Davies JL, Kawaguchi Y, Bennett ST, Copeman JB, Cordell HJ, Pritchard LE, Reed PW, et al (1994) A genome-wide search for human type 1 diabetes susceptibility genes. Nature 371: 130–136

Dib C, Faure S, Fizames C, Samson D, Drouot N, Vignal A, Millasseau P, et al (1996) A comprehensive map of the hu-

man genome based on 5264 microsatellites. Nature 380: 152–154

Dubovsky J, Sheffield VC, Duyk GM, Weber JL (1995) Sets of short tandem repeat polymorphisms for efficient linkage screening of the human genome. Hum Mol Genet 4:449–452

Ehm MG, Wagner M (1998) A test statistic to detect errors in sib-pair relationships. Am J Hum Genet 62:181–188

Elbein SC, Yount PA, Teng K, Hasstedt SJ (1998) Genome-wide search for type 2 diabetes susceptibility genes in caucasians: evidence for a recessive locus on chromosome 1. Diabetes 47 Suppl 1:A15

Elston RC (1984) Genetic analysis workshop II: sib-pair screening tests for linkage. Genet Epidemiol 1:175–178

Froguel P, Vaxillaire M, Sun F, Velho G, Zouall M, Butel MO, Lesage S, et al (1992) Close linkage of glucokinase locus on chromosome 7p to early-onset non-insulin-dependent diabetes mellitus. Nature 356:162–164

Fulker DW, Cardon LR (1994) A sib-pair approach to interval mapping of quantitative trait loci. Am J Hum Genet 54: 1092–1103

Fulker DW, Cherny SS (1996) An improved multipoint sib-pair analysis of quantitative traits. Behav Genet 26:527–532

Fulker DW, Cherny SS, Cardon LR (1995) Multipoint interval mapping of quantitative trait loci, using sib pairs. Am J Hum Genet 56:1224–1233

Ghosh S, Schork NJ (1996) Genetic analysis of NIDDM: the study of quantitative traits. Diabetes 45:1–14

Haldane JBS (1919) The combination of linkage values and the calculation of distances between the loci of linked factors. J Genet 8:299–309

Hanis CL, Boerwinkle E, Chakraborty R, Ellsworth DL, Concannon P, Stirling B, Morrison VA, et al (1996) 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

Hanson RL, Elston RC, Pettitt DJ, Bennett PH, Knowler WC (1995*a*) Segregation analysis of non-insulin-dependent diabetes mellitus in Pima Indians: evidence for a major-gene effect. Am J Hum Genet 57:160–170

Hanson RL, Knowler WC (1998) Analytic strategies to detect linkage to a common disorder with genetically determined age of onset: diabetes mellitus in Pima Indians. Genet Epidemiol 15:299–315

Hanson RL, Knowler WC, Pima Diabetes Genes Group (1997) Variance components and Haseman-Elston sib-pair linkage analyses in a genomic scan for markers linked to obesity. Am J Hum Genet Suppl 61:A278

Hanson RL, Pettitt DJ, Bennett PH, Narayan KMV, Fernandes R, de Courten M, Knowler WC (1995*b*) Familial relationships between obesity and NIDDM. Diabetes 44:418–422

Hanson R, Pima Diabetes Genes Group (1997) Genomic scan for markers linked to type II diabetes in Pima Indians. Diabetes 46 Suppl 1:51A

Haseman JK, Elston RC (1972) The investigation of linkage between a quantitative trait and a marker locus. Behav Genet 2:3–19

Hopper JL, Matthews JD (1982) Extensions to multivariate normal models for pedigree analysis. Ann Hum Genet 46: 373–383

Knowler WC, Bennett PH, Hamman RF, Miller M (1978) Diabetes incidence and prevalence in Pima Indians: a 19-fold greater incidence than in Rochester, Minnesota. Am J Epidemiol 108:497–505

Knowler WC, Pettitt DJ, Saad MF, Charles MA, Nelson RG, Howard BV, Bogardus C, et al (1991) Obesity in the Pima Indians: its magnitude and relationship with diabetes. Am J Clin Nutr 53 Suppl:1543S-1551S

Knowler WC, Pettitt DJ, Savage PJ, Bennett PH (1981) Diabetes incidence in Pima Indians: contributions of obesity and parental diabetes. Am J Epidemiol 113:144–156

Kruglyak L, Lander ES (1995) Complete multipoint sib-pair analysis of qualitative and quantitative traits. Am J Hum Genet 57:439–454

Lander ES, Green P (1987) Construction of multilocus genetic linkage maps in humans. Proc Natl Acad Sci USA 84: 2363–2367

Lander E, Kruglyak L (1995) Genetic dissection of complex traits: guidelines for interpreting and reporting linkage results. Nat Genet 11:241–247

Lange K, Boehnke M (1983) Extensions to pedigree analysis. IV. Covariance components models for multivariate traits. Am J Med Genet 14:513–524

Lange K, Westlake J, Spence MA (1976) Extensions to pedigree analysis. III. Variance components by the scoring method. Ann Hum Genet 46:373–383

Mahtani MM, Widén E, Lehto M, Thomas J, McCarthy M, Brayer J, Bryant B, et al (1996) Mapping of a gene for type 2 diabetes associated with an insulin secretion defect by a genome scan in Finnish families. Nat Genet 14:90–94

Morton NE (1998) Significance levels in complex inheritance. Am J Hum Genet 62:690–697

Murray JC, Buetow KH, Weber JL, Ludwigsen S, Scherpbier-Heddema T, Manion F, Quillen J, et al (1994) A comprehensive human linkage map with centimorgan density. Science 265:2049–2054

Norman RA, Tataranni PA, Pratley R, Thompson DB, Hanson RL, Prochazka M, Baier L, et al (1998) Autosomal genomic scan for loci linked to obesity and energy metabolism in Pima Indians. Am J Hum Genet 62:659–668

Norman RA, Thompson DB, Foroud T, Garvey WT, Bennett PH, Bogardus C, Ravussin E, et al (1997) Genomewide search for genes influencing percent body fat in Pima Indians: suggestive linkage at chromosome 11q21-q22. Am J Hum Genet 60:166–173

Pratley RE, Thompson DB, Prochazka M, Baier L, Mott D, Ravussin E, Sakul H, et al (1998) An autosomal genomic scan for loci linked to prediabetic phenotypes in Pima Indians. J Clin Invest 101:1757–1764

Price RA, Charles MA, Pettitt DJ, Knowler WC (1994) Obesity in Pima Indians: genetic segregation analyses of body mass index complicated by temporal increases in obesity. Hum Biol 66:251–274

Pugh EW, Jaquish CE, Sorant AJM, Doetsch JP, Bailey-Wilson JE, Wilson AF (1997) Comparison of sib-pair and variance components methods for genomic screening. Genet Epidemiol 14:867–872

SAGE (1994) Statistical analysis for genetic epidemiology, release 2.2. Department of Epidemiology and Biostatistics, Case Western Reserve University, Cleveland

Schwengel DA, Jedlicka AE, Nathakumar EJ, Weber JL, Levitt

RC (1994) Comparison of fluorescence-based semi-automated genotyping of multiple microsatellite loci with autoradiographic techniques. Genomics 22:46–54

Stoffers DA, Ferrer J, Clarke WL, Habener JF (1997) Early-onset type II diabetes mellitus (MODY4) linked to IPF1. Nat Genet 17:138–139

Therneau TM, Grambsch PM, Fleming TR (1990) Martingale-based residuals for survival models. Biometrika 77: 147–160

Tuomilehto-Wolf E, Tuomilehto J, Hitman GA, Aulikki N, Stengård J, Pekkanen J, Kivinen P, et al (1993) Genetic susceptibility to non-insulin-dependent diabetes mellitus and glucose intolerance are located in the HLA region. Br Med J 307:155–159

Vaxillaire M, Boccio V, Philippi A, Vigouroux C, Terwilliger J, Passa P, Beckmann JS, et al (1995) A gene for maturity onset diabetes of the young (MODY) maps to chromosome 12q. Nat Genet 9:418–423

Wilson AF, Elston RC (1993) Statistical validity of the Haseman-Elston sib-pair test in small samples. Genet Epidemiol 10:593–598

World Health Organization Study Group on Diabetes Mellitus (1985) Diabetes mellitus. World Health Organization tech rep ser 727. World Health Organization, Geneva

Yamagata K, Furuta H, Oda N, Kaisaki PJ, Menzel S, Cox NJ, Fajans SS et al (1996*a*) Mutations in the hepatocyte nuclear factor-4-alpha gene in maturity onset diabetes of the young (MODY1). Nature 384:458–460

Yamagata K, Oda N, Kaisaki PJ, Menzel S, Furuta H, Vaxillaire M, Southam L, et al (1996*b*) Mutations in the hepatocyte nuclear factor-1-alpha gene in maturity onset diabetes of the young (MODY3). Nature 384:455–458