Am. J. Hum. Genet. 66:1871–1881, 2000

Genomewide Search for Type 2 Diabetes Susceptibility Genes in Four American Populations

Margaret Gelder Ehm,[1] Maha C. Karnoub,[1] Hakan Sakul,[2,*] Kirby Gottschalk,[1] Donald C. Holt,[1] James L. Weber,[3] David Vaske,[3,†] David Briley,[1] Linda Briley,[1] Jan Kopf,[1] Patrick McMillen,[1] Quan Nguyen,[1] Melanie Reisman,[1] Eric H. Lai,[1] Geoff Joslyn,[2,‡] Nancy S. Shepherd,[1] Callum Bell,[2,§] Michael J. Wagner,[1] Daniel K. Burns,[1] and the American Diabetes Association GENNID Study Group[‖]

[1]Glaxo Wellcome, Inc., Research Triangle Park, NC; [2]Axys Pharmaceuticals, La Jolla, CA; and [3]Marshfield Medical Research Foundation, Marshfield, WI

Type 2 diabetes is a serious, genetically influenced disease for which no fully effective treatments are available. Identification of biochemical or regulatory pathways involved in the disease syndrome could lead to innovative therapeutic interventions. One way to identify such pathways is the genetic analysis of families with multiple affected members where disease predisposing genes are likely to be segregating. We undertook a genomewide screen (389–395 microsatellite markers) in samples of 835 white, 591 Mexican American, 229 black, and 128 Japanese American individuals collected as part of the American Diabetes Association's GENNID study. Multipoint nonparametric linkage analyses were performed with diabetes, and diabetes or impaired glucose homeostasis (IH). Linkage to diabetes or IH was detected near markers D5S1404 (map position 77 cM, LOD = 2.80), D12S853 (map position 82 cM, LOD = 2.81) and GATA172D05 (X-chromosome map position 130 cM, LOD = 2.99) in whites, near marker D3S2432 (map position 51 cM, LOD = 3.91) in Mexican Americans, and near marker D10S1412 (map position 14 cM, LOD = 2.39) in African Americans mainly collected in phase 1 of the study. Further analyses showed evidence for interactions between the chromosome 5 locus and region on chromosome 12 containing the MODY 3 gene (map position 132 cM) and between the X-chromosome locus and region near D12S853 (map position 82 cM) in whites. Although these results were not replicated in samples collected in phase 2 of the GENNID study, the region on chromosome 12 was replicated in samples from whites described by Bektas et al. (1999).

Introduction

In 1992, 4.5% of the population of the United States was diagnosed with type 2 diabetes (Raffel et al. 1996). Type 2 diabetes is associated with hyperglycemia and obesity, seldom leads to ketoacidosis, and generally occurs after age 40 years—in contrast to type 1 diabetes, which is associated with autoimmune destruction of the β-cells in the pancreas, with a juvenile age of onset (Expert Committee on the Diagnosis and Classification of Diabetes Mellitus 1998). Although the molecular basis of type 2 diabetes is unknown, it is thought to result from defects in both insulin action and insulin secretion. Segregation analysis and twin studies indicate that a genetic component exists, with a recurrence risk to first-degree relatives of ~3.5 (Rich 1990). In several relatively uncommon variants of the disease, causative mutations have been identified in genes such as the insulin and insulin-receptor genes (Turner et al. 1995), glucokinase (Froguel et al. 1993), and the HNF-1α (Yamagata et al. 1996*b*) and HNF-4α (Yamagata et al. 1996*a*) transcription factors. However, the genetic defects responsible for the vast majority of cases have not been identified.

Since type 2 diabetes is thought to be an oligogenic, multifactorial disease with a substantial environmental component, a large sample was required to ensure high power to locate susceptibility genes. The GENNID (Genetics of NIDDM) study is a multicenter project sponsored by the American Diabetes Association to collect a large family-based resource, containing multiple ethnic groups, to be used to identify the genetic components of type 2 diabetes (Raffel et al. 1996). We un-

Received December 20, 1999; accepted for publication April 3, 2000; electronically published May 2, 2000.

Address for correspondence and reprints: Dr. Margaret Gelder Ehm, Glaxo Wellcome, Inc., P.O. Box 13398, Research Triangle Park, NC 27709-3398; E-mail: mge37216@glaxowellcome.com

[*] Present affiliation: Parke-Davis Laboratory for Molecular Genetics, Alameda, CA.

[†] Present affiliation: Pioneer Hi-Bred International, Johnston, IA.

[‡] Present affiliation: Department of Human Genetics, University of California–Los Angeles, Los Angeles, CA.

[§] Present affiliation: National Center for Genome Resources, Sante Fe, NM.

[‖] See Acknowledgments for a list of GENNID Study Group members.

0002-9297/2000/6606-0016$02.00

Table 1

Listing of Individuals Used in Study

Group	Whites	Mexican Americans	Blacks	Japanese Americans	Whites 2	Mexican Americans 2
Families	77	53	65	18	85	64
Individuals	497	365	229	128	338	226
Diabetic individuals (sib pairs)	191 (117)	169 (173)	124 (62)	50 (45)	225 (130)	137 (86)
Diabetic or IH[a] individuals (sib pairs)	288 (198)	233 (239)	152 (109)	87 (82)	267 (254)	165 (137)

[a] Impaired glucose homeostasis.

dertook a genome scan of samples collected as part of the GENNID study to identify chromosomal regions linked to diabetes and possibly containing diabetes genes. This paper will describe the results of a genome scan on 1,783 individuals collected in four ethnic groups.

Subjects and Methods

Ascertainment of families in the GENNID study included two phases. Phase 1 enrolled families with at least two siblings affected with type 2 diabetes according to the National Diabetes Data Group criteria (fasting plasma glucose concentration >140 mg/dl on more than one occasion or 2-h and one other time point during an OGTT that was >200 mg/dl) (National Diabetes Data Group 1979). In addition, the families were required to have at least three available additional first-degree relatives and a nondiabetic, ethnically matched control individual. Phase 2 enrolled families either with at least two affected siblings and both parents or with one or no parents and at least two additional siblings.

All individuals collected as part of the GENNID study completed a family/medical history questionnaire (Raffel et al. 1996), gave fasting blood samples for establishment of cell lines, insulin, glucose, and lipid measurements, and underwent anthropometric and blood pressure measurements. Individuals not known to have diabetes had an oral glucose-tolerance test (Raffel et al. 1996).

The samples analyzed for this study include white, Mexican American, and black families collected in the first two phases of the GENNID study (Raffel et al. 1996) and Japanese American families collected only in the first phase of the study. Our initial analysis included the phase 1 white, phase 1 Mexican American, phase 1 Japanese American, and combined phase 1 and 2 black samples. The phase 2 white and Mexican American individuals were used as replication samples. We did not analyze the phase 1 and 2 black samples separately because of insufficient numbers in each group.

We applied the American Diabetes Association criteria for the diagnosis of type 2 diabetes, impaired fasting plasma glucose, and impaired glucose tolerance (Expert Committee on the Diagnosis and Classification of Diabetes Mellitus 1998) to the data collected in the GENNID study, with the following modifications: individuals taking diabetic medications were identified as diabetic; confirmation of fasting or 2-h glucose levels on a subsequent day was not required; and individuals who were pregnant, were taking glucose-altering medications, were aged <25 years at their diagnosis or exam, or had a history of diabetes but were not taking diabetes medications or had normal fasting plasma glucose values were assigned a diagnosis of "unknown."

We limited our study to investigating the genome for linkage to two phenotypes: (a) diabetes, and (b) diabetes, impaired fasting glucose (IFG), or impaired glucose tolerance (IGT) (together termed "impaired glucose homeostasis" [IH]). Diabetes or IH was chosen as a phenotype, since IFG and IGT may be intermediate stages in the development of type 2 diabetes (Expert Committee on the Diagnosis and Classification of Diabetes Mellitus 1998). Table 1 summarizes the samples available for study. Table 2 gives a summary of the clinical characteristics of individuals sampled for phase 1. The characteristics for the phase 2 samples did not differ appreciably.

Whole-genome polymorphism scans were carried out at the Marshfield Medical Research Foundation by use of standard sets of highly informative short-tandem-repeat polymorphisms (Yuan et al. 1997). The genome scan consisted of 389–395 markers (screening set 6 for the phase 1 Mexican Americans, screening set 8 for all other groups). Additional markers were typed at Glaxo Wellcome, Inc., to fill gaps in the genome scan, to saturate regions of interest, and to ensure coverage of areas near candidate genes. The average spacing ranged from 8.2 to 9.0 cM, no gap was greater than 23.5 cM, and no more than 4 gaps were greater than 20 cM for any ethnic sample. The average marker heterozygosity for each sample was ~0.75.

Inspection of the Mendelian errors over all markers and the distribution of marker alleles shared identical by state for each pair of siblings was used to check for pedigree errors and monozygotic twins (Ehm and Wagner 1998). The pedigree errors were removed and all other Mendelian errors were eliminated by means of an

Table 2

Clinical Characteristics of Individuals Sampled

Ethnic Group and Characteristics	Mean	Standard Deviation	No.
Whites (n =497, 213 males):			
Age at exam	50.73	16.50	472
Age at onset	53.14	13.52	288
BMI	30.43	6.72	271
Fasting glucose (mg/dl)	150.61	64.58	283
Fasting insulin (μU/ml)	14.01	13.26	263
Mexican Americans (n = 365, 135 males):			
Age at exam	47.25	16.22	349
Age at onset	49.23	12.82	232
BMI	31.61	8.18	227
Fasting glucose (mg/dl)	175.82	84.11	226
Fasting insulin (μU/ml)	15.71	11.55	216
Blacks (n = 229, 66 males):			
Age at exam	51.90	14.56	218
Age at onset	51.02	12.43	152
BMI	33.04	8.69	142
Fasting glucose (mg/dl)	165.48	79.34	149
Fasting insulin (μU/ml)	17.39	21.42	138
Japanese Americans (n = 128, 67 males):			
Age at exam	56.17	15.95	125
Age at onset	57.91	13.67	87
BMI	25.84	4.01	84
Fasting glucose (mg/dl)	127.11	35.34	87
Fasting insulin (μU/ml)	8.74	6.91	86

NOTE.—The age of exam was available for all individuals sampled, while age of onset was only available for individuals who were diabetic or exhibited impaired glucose homeostasis. The summary statistics for body mass index (BMI), fasting glucose and fasting insulin included only individuals who were diabetic or had impaired glucose homeostasis.

algorithm which identifies individuals for whom setting their genotype for a particular marker to "missing" eliminates the errors (R. Idury, personal communication).

Allele frequencies were estimated for each ethnic sample by use of all individuals in the sample. Map distances used for linkage analyses of the white, Japanese American, and black samples were estimated by means of the phase 1 and 2 white samples. Map distances used for linkage analyses of the Mexican American samples were estimated with the phase 1 and 2 Mexican American samples. By use of the Marshfield map as a guide, along with information from other genetic and physical maps, a genetic map was created via CRIMAP (Lander and Green 1987). Additional markers were inserted, and the map with the most likely order and distance was chosen. Although the data represent few informative meioses, when the final map was compared to the Marshfield map there were no significant differences noted.

Multipoint nonparametric linkage analysis by means of affected relative pairs was performed by the Genehunter program (Kruglyak et al. 1996) in conjunction with Kong and Cox's modification, which calculates exact likelihoods and reports LOD scores on the basis of a one-parameter allele-sharing statistic, Z_{lr}, rather than using the perfect data approximation (Kong and Cox 1997). The asymptotically normal linkage statistic (NPL score) reported by GENEHUNTER was used in interaction analyses.

To investigate the possibility of interaction between the regions identified on individual chromosomes following Cox et al. (1999), we incorporated the linkage evidence for these regions in assessing the evidence for linkage at other positions. Instead of giving each family equal weight in Kong and Cox' approach (Kong and Cox 1997), families were weighted in three ways, according to the NPL score for each family at the position conditioned upon: (1) families with positive NPL scores were given a weight of 1, and those with negative NPL scores were given a weight of 0; (2) families with positive NPL scores were given a weight of 0, and those with negative NPL scores were given a weight of 1; and (3) families with positive NPL scores were given a weight of NPL/max(NPL), and those with negative NPL scores were given a weight of 0. P values associated with the increase in LOD scores were obtained by multiplying the difference in the weighted LOD scores by 2ln(10) which results in an asymptotically conservative χ^2 test with 1 df (Cox et al. 1999). The results represent exploratory analyses, and consequently the P values were not adjusted for multiple tests. Correlations between the family NPL scores from interesting regions were also calculated.

Results

Our analysis of the pedigree structures by means of the genotypes generated as part of the genome scan highlighted that, in each of the ethnic groups, there were individuals identified as males that were likely to be females (and vice versa), half siblings labeled as full siblings, and pedigree members that showed no relationship to their supposed pedigree. Given that not all of the parents were available for study, it was difficult to distinguish between parental errors and blood- or DNA-sample mixups. In summary, 24.4% of the families contained pedigree errors and 2.8% of the families contained errors in which an individual appeared to be unrelated to the rest of the members of the pedigree and were possibly blood-sample mixups. The percentages were consistent across all ethnic groups. In total, 212 individuals were removed from the pedigrees to eliminate these errors.

Results of multipoint nonparametric linkage analyses for diabetes or IH for all phase 1 ethnic samples are shown infigure 1. LOD scores are plotted versus map distance for all chromosomes. Five regions were significant at the $P < .001$ level (LOD > 2.07). Note that the

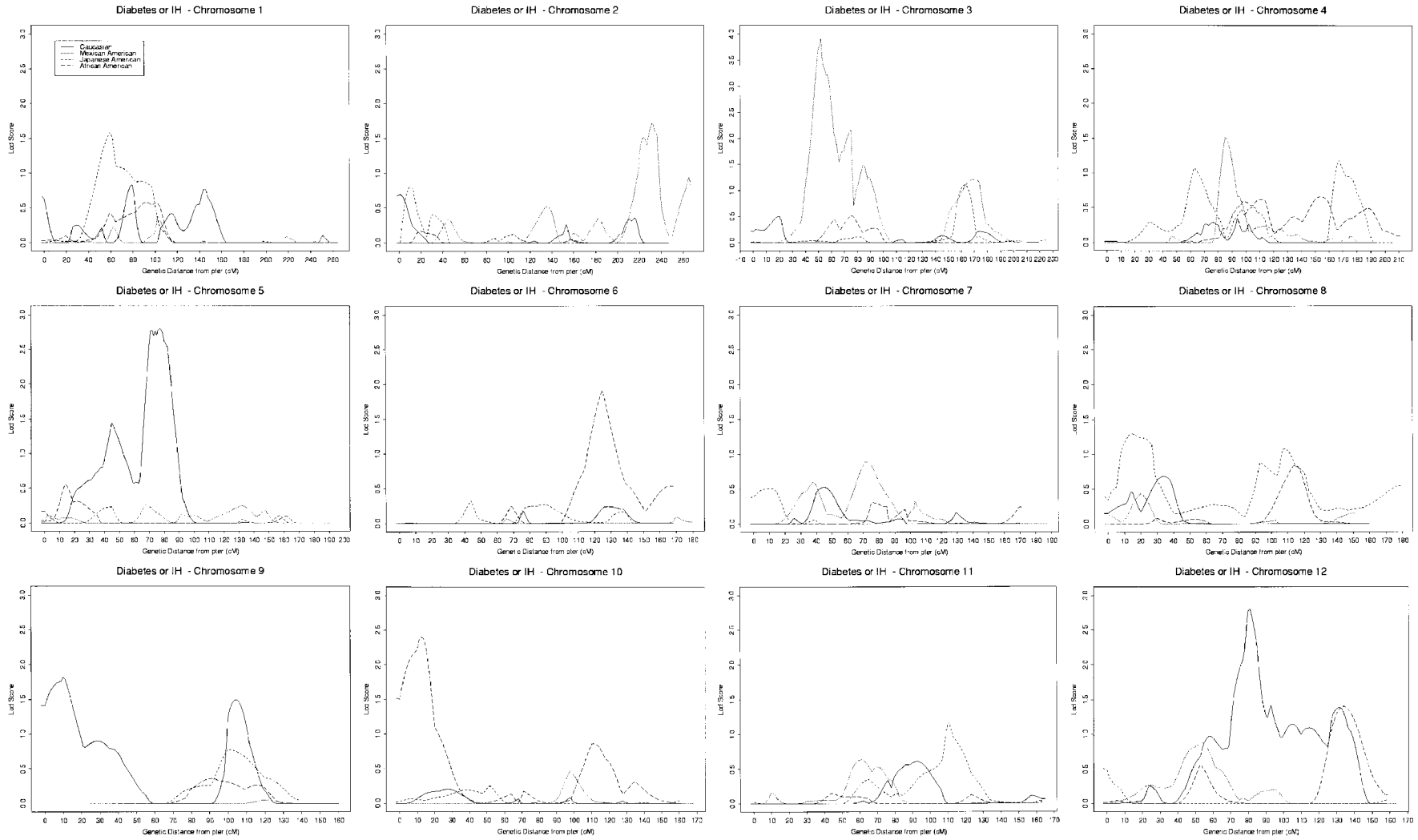
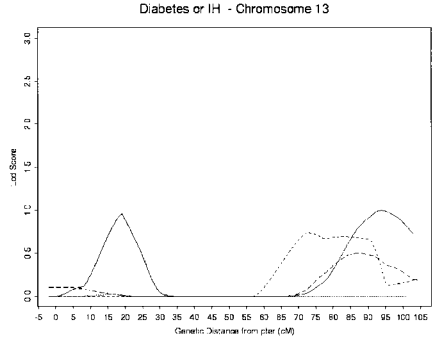
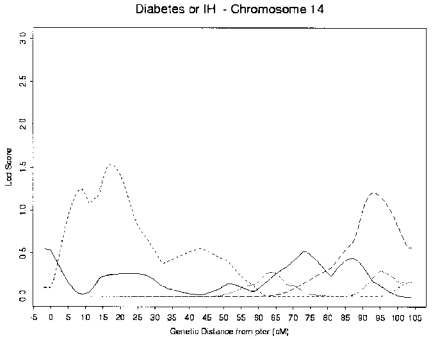
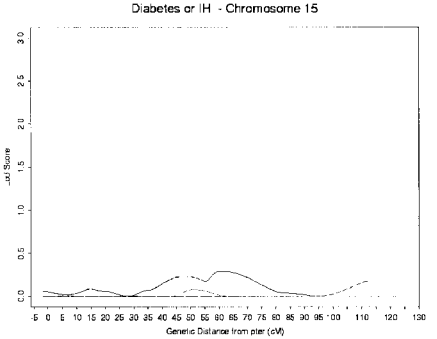
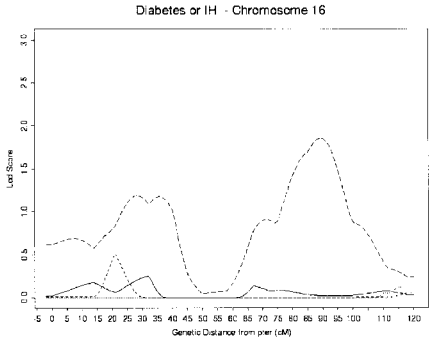


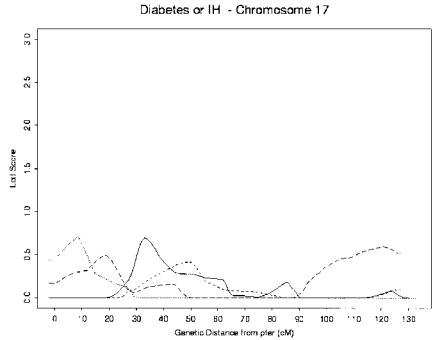
Figure 1 Multipoint linkage analysis results. Results for phase 1 samples for four ethnic groups obtained by analysis of the trait diabetes or impaired glucose homeostasis by multipoint nonparametric linkage analysis are shown for each chromosome. Subjects from four ethnic groups were genotyped for 389–395 markers. The average spacing was 8.2–9.0 cM.

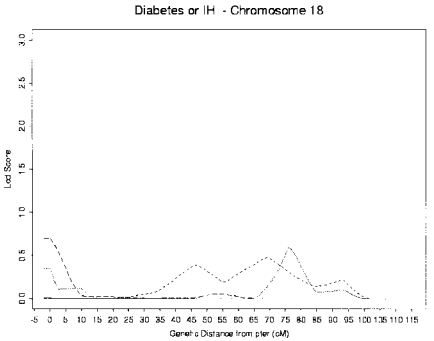

Diabetes or IH - Chromosome 13
Lod Score
Genetic Distance from pter (cM)

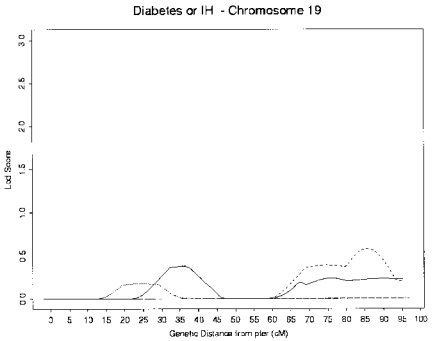

Diabetes or IH - Chromosome 14
Lod Score
Genetic Distance from pter (cM)

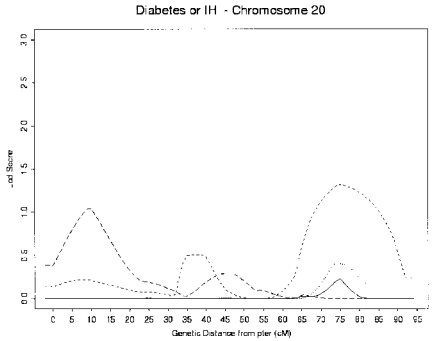

Diabetes or IH - Chromosome 15
Lod Score
Genetic Distance from pter (cM)

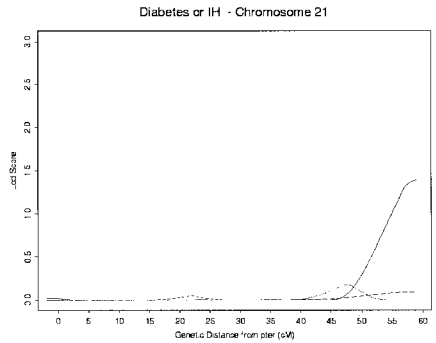

Diabetes or IH - Chromosome 16
Lod Score
Genetic Distance from pter (cM)

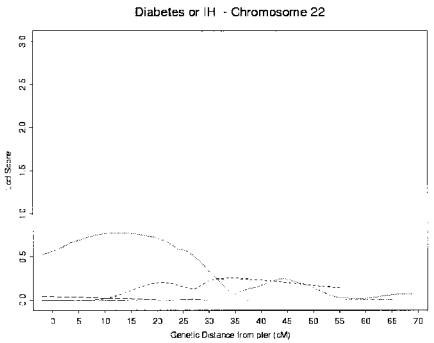

Diabetes or IH - Chromosome 17
Lod Score
Genetic Distance from pter (cM)

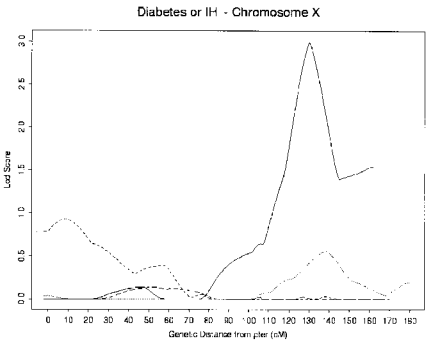

Diabetes or IH - Chromosome 18
Lod Score
Genetic Distance from pter (cM)


Diabetes or IH - Chromosome 19
Lod Score
Genetic Distance from pter (cM)


Diabetes or IH - Chromosome 20
Lod Score
Genetic Distance from pter (cM)


Diabetes or IH - Chromosome 21
Lod Score
Genetic Distance from pter (cM)


Diabetes or IH - Chromosome 22
Lod Score
Genetic Distance from pter (cM)


Diabetes or IH - Chromosome X
Lod Score
Genetic Distance from pter (cM)

results for the diabetes phenotype will be described in the text, but the data are not shown.

Markers near D3S2432 were linked in Mexican Americans to diabetes or IH (LOD = 3.91) and diabetes (LOD = 2.36). Neither result was replicated in the phase 2 Mexican Americans (diabetes or IH: LOD = 0.0028). Markers near D5S1404 were linked to diabetes or IH (LOD = 2.8) and to diabetes (LOD = 3.26) in whites. Neither the phase 2 white samples (diabetes or IH: LOD = 0.093) nor any other ethnic sample showed linkage to this region. Markers near D10S1412 were linked to diabetes or IH (LOD = 2.39) and to diabetes (LOD = 1.06) in African Americans. Markers near *D12S853* (82 cM) were linked to diabetes or IH in whites with a LOD score of 2.81 and to diabetes with a LOD of 1.91. Diabetes or IH was also linked to D12S342 (132 cM) with a LOD score of 1.38, although there was no evidence for linkage of diabetes (LOD = 0.24) to this region. This region is not significant at the $P = .001$ level, but these results are reported because a similar LOD score was obtained in the black sample, and it is the same region reported by Mahtani et al. (1996) to contain *NIDDM2* (MIM 601407). Markers near *GATA172D05* on the X chromosome were linked to diabetes or IH (LOD = 2.99) and to diabetes (LOD = 2.27) in whites. The phase 2 white samples did not show linkage to either region on chromosome 12 or the region on the X chromosome.

As suggested by Cox et al. (1999), we investigated the evidence for statistical interactions between regions of linkage by incorporating the linkage evidence with diabetes or IH for the chromosome 5 region identified in the phase 1 whites in assessing the evidence for linkage on all other chromosomes, using the phase 1 white samples. Similarly, we incorporated the evidence for linkage with diabetes or IH for the chromosome 12 (positions 82 cM and 132 cM) and X-chromosome regions when looking for interactions on all other chromosomes. Any significant interactions found were also assessed in the phase 2 whites. Similarly, we incorporated the linkage evidence with diabetes or IH for the chromosome 2 region hypothesized to contain *NIDDM1* (MIM 601283) (Hanis et al. 1996; Bell et al. 1997; Cox et al. 1999) and the chromosome 3 region in the phase 1 Mexican American samples in assessing the evidence for linkage on all other chromosomes using the phase 1 Mexican Americans. There was evidence for an epistatic interaction between position 130 cM on chromosome X and position 82 cM on chromosome 12, both linked to diabetes or IH (correlation = .335) in the phase 1 white samples. Position 130 cM on the X chromosome gave an original LOD score of 2.99, but, when families were weighted on the basis of evidence for linkage at 82 cM on chromosome 12, the LOD score increased to 4.43 ($P = .010$). There was also evidence for an epistatic interaction between position 132 cM on chromosome 12 and position 77 cM on chromosome 5, both linked to diabetes or IH (correlation = .353) in the phase 1 white samples. Position 132 cM on chromosome 12 gave a LOD score of 1.38 and a weighted LOD score of 2.38 ($P = .032$). An epistatic interaction indicates that the families that are linked to position 82 cM on chromosome 12 are many of the same families linked to position 130 cM on the X chromosome and the families linked to position 132 cM on chromosome 12 are similar to those linked to position 77 cM on chromosome 5. Neither interaction was replicated in the phase 2 white samples. No other regions showed evidence of interactions.

The results highlighted for whites and Mexican Americans in this study were not replicated in our additional samples of whites and Mexican-Americans. The results for the analysis of the phenotype diabetes or IH of the genome scan completed on the phase 2 samples (data not shown) of whites and Mexican Americans showed no multipoint LOD scores significant at the .001 level (LOD > 2.07) and only four regions significant at the .01 level (LOD > 1.175). Marker SE30 (0 cM) on chromosome 6 was linked to diabetes or IH in phase 2 whites (LOD = 2.04). Marker D14S599 (26 cM) was linked to diabetes or IH in phase 2 whites (LOD = 1.76). Marker D14S608 (17 cM) was linked to diabetes or IH in Japanese Americans (LOD = 1.53). Markers D18S843 (15 cM), D18S535 (50 cM), and D18S851 (63 cM) were linked to diabetes or IH in phase 2 whites with LOD scores of 1.40, 1.29, and 1.37, respectively.

The results described above and shown infigure 1 included data from additional markers typed in three regions which appeared to show evidence for linkage in a preliminary analysis of the genome scan in the phase 1 white samples (chromosome 5, position 84 cM; chromosome 9, position 10 cM; chromosome 12, position 88 cM) and in two regions highlighted in the phase 1 Mexican American samples (chromosome 2, position 259 cM; chromosome 3, position 58 cM). The initial genome scan and gap-filling markers were ~10 cM apart. We selected additional markers with the goal of obtaining one marker every 5 cM in each of the five regions. The LOD scores obtained from analyzing all of the markers were greater than the LOD scores obtained analyzing only the genome scan markers in four of the regions, with an average increase of 0.870 LOD units. In one of the regions, the peak moved 27 cM centromeric (from position 259 cM to position 232 cM on chromosome 2 in the Mexican American samples). In one of the regions the final LOD score was <1.0 (chromosome 9 in the white samples). In three of the regions, the 1–LOD unit support interval obtained from analyzing all the markers was smaller than that obtained analyzing only the scan markers: 11 versus 16 cM for

Table 3

Results of Published Linkages for Type 2 Diabetes

Ethnic Group	Sample	Peak Marker/Region	Trait	LOD
White	53 sibling pairs with a history of diabetic nephropathy (Bowden et al. 1997)	D12S86	Diabetes	1.45
		D20S197	Diabetes	1.5
French (white)	133 sibling pairs (Hani et al. 1997)	PCK1 (chromosome 20)	Diabetes	1.45
White	29 extended families (Elbein et al. 1999)	D20S197	Diabetes after 47	2.37
White	42 extended families (Elbein et al. 1999)	CRP-APOA2 (chromosome 1)	Diabetes (recessive)	4.3
White	26 families (Bektas et al. 1999)	D12S375-D12S1052	Diabetes	3.14
Finnish	26 families (Mahtani et al. 1996)	D12S1349	Diabetes (stratified on 30-min insulin)	3.3
Finnish	477 families (Ghosh et al. 1999)	20p	Diabetes	2.06
Mexican American	330 sibling pairs (Hanis et al. 1996; Bell et al. 1997; Cox et al. 1999)	D2S125	Diabetes	4.03
		CYP19 (chromosome 15)	Diabetes (weighted)	4.00
Mexican American	27 extended families (Duggirala et al. 1999)	D3S1566-GATA128C02	Diabetes onset age	2.51
			Diabetes	2.67
		D4S1615	Diabetes	1.94
		D9S288-D9S925	Diabetes onset age	2.06
			Diabetes	2.38
		D10S587	Diabetes onset age	3.75
			Diabetes	2.88
Pima Indian	264 nuclear families (Hanson et al. 1998)	D6S1009-D6S1003	Age-adjusted diabetes	1.39
		D9S299-D9S2026	Age-adjusted diabetes	1.22
		D11S4464-D11S912	Age-adjusted diabetes	1.66
		D1S127	Diabetes before age 25 years	4.1
		D7S1799	Diabetes	1.8

chromosome 3 in the Mexican American, 17 versus 77 cM for chromosome 5 and 12 versus 19 cM for chromosome 12 in the white samples).

Discussion

The GENNID study is unique, because samples from four distinct ethnic groups were ascertained, diagnosed, and genotyped and the resulting data was analyzed in a uniform and consistent manner. The individuals studied were sampled from outbred populations at eight centers throughout the United States. Consequently, the affected individuals are likely to be broadly representative of patients with type 2 diabetes, in contrast to studies in isolated populations, which may represent less common forms of disease. The phenotypes used for analysis include diabetes (using 1998 ADA criteria) and a broad diabetes phenotype, diabetes or impaired glucose homeostasis, which includes individuals with diabetes, impaired fasting glucose, or impaired glucose tolerance. Linkage analysis results obtained with these phenotypes (particularly diabetes or impaired glucose homeostasis) represent genes that confer susceptibility to common diabetic and prediabetic phenotypes. Although the GENNID study as a whole (phases 1 & 2) includes more families than many of the published reports, the number of families in each ethnic group is modest.

A number of authors have reported linkages to type 2 diabetes, each with varying levels of evidence (table 3). There are two reports of linkage on chromosome 3 in Mexican Americans. Hanis et al. (1996) and Bell et al. (1997) show nominal evidence (LOD >0.59) for a region around D3S2452 in their sample of Mexican Americans. This region is 20 cM centromeric to the marker (D3S2432) showing the peak LOD score in the GENNID Mexican Americans, but it overlaps substantially with the peak. Duggirala et al. (1999) show evidence for linkage of age of diabetes onset (LOD = 2.51) and diabetes (LOD = 2.67) to a region on 3p in Mexican Americans. However, their map appears to be significantly different than ours, and the region that they report is 55 cM centromeric (map position 108 cM on our map) to D3S2432 (Duggirala et al. 1999). Hanis et al. and Bell et al. results in Mexican Americans may replicate the result we report on chromosome 3, but the linkage reported by Duggirala et al. appears to be too far away.

Other authors have also provided replication for our results on chromosome 12. The first region on chromosome 12 (markers near D12S853: 82 cM) overlaps

with the region identified by Bektas et al. (1999) as linked to type 2 diabetes among early-onset white families with apparent autosomal dominant inheritance of diabetes (LOD = 3.14 at D12S1052, corresponding to map position 68 cM on our map). The second region, ~50 cM telomeric (D12S342) to the first, corresponds to the proposed location of *NIDDM2* (Mahtani et al. 1996) and contains the *MODY3* gene (MIM 600496) (Yamagata et al. 1996*b*), *HNF-1α*. Sequencing of the protein-coding exons and flanking intronic sequences of *HNF-1α* of unrelated Mexican Americans (60 diabetic patients and 14 controls) and whites (69 diabetic patients and 19 controls) from the GENNID sample revealed no known or novel mutations (E. Lai, unpublished data), making it unlikely that *HNF-1α* is responsible for these results. Note that the black samples also show linkage to diabetes or impaired glucose homeostasis at position 132 cM (D12S342) with a LOD score of 1.40. This is the only location in the entire genome scan for which two or more ethnic groups show linkage to the same region with LOD scores significant at the $P = .01$ level. Given the distance between the two regions of linkage on chromosome 12, the independent evidence for replication of these linkages, and the fact that the regions linked to D12S853 and D12S342 appear to be involved in epistatic interactions with genes on different chromosomes, it seems likely that there are at least two diabetes-susceptibility genes on chromosome 12.

Our results on chromosome 2 do not convincingly replicate the linkage identified by Hanis et al. (1996) and Bell et al. (1997) (putative disease gene named *NIDDM1*). We obtained a LOD score of 0.88 for D2S125 and a LOD score of 1.66 for D2S362 (30 cM centromeric) in Mexican Americans. Some of the families studied by Hanis et al. (1996) and Bell et al. (1997) are also included in the GENNID collection, so it is not surprising to find modest evidence for linkage near the *NIDDM1* gene at D2S125, and, therefore, this may not represent an independent replication. It seems unlikely that the higher LOD score we see at D2S362, 30 cM from D2S125, could also be due to the NIDDM1 gene.

The chromosome 3 linkage region and the one on chromosome 5 are relatively gene-rich and contain obvious candidate genes. *GNAI2* (MIM 139360), encoding Giα2, is the most interesting candidate for the chromosome 3 region. Giα2 knock-out mice are severely insulin resistant (Moxham and Malbon 1996), whereas mice overexpressing a constituitively active mutant (codon $205^{Gln \rightarrow Leu}$) in liver and skeletal muscle exhibited enhanced glucose tolerance compared with controls (Chen et al. 1997). The chromosome 5 linkage region contains three candidate genes of interest. *PIK3R1* (MIM 171833), which encodes the p85α regulatory subunit of PI3′-kinase, is of interest, since mice lacking this subunit exhibit increased insulin sensitivity and moderate hypoglycemia, presumably because of the compensatory overexpression of its alternately spliced p50α variant (Terauchi et al. 1999). Hansen et al. (1997) and Kawanishi et al. (1997) detected a codon $326^{Met \rightarrow Ile}$ variant in Danish and Japanese populations, respectively. The observed allele frequency did not differ between affected and unaffected individuals in either population. Hansen et al. did note, however, that individuals homozygous for the Ile allele exhibited reduced whole-body glucose disposal and intravenous glucose-disappearance constant. Hansen et al. (1997) also reported three silent polymorphisms in codons 261 (C→T), 663 (T→G) and 810 (A→G). Two additional genes in this region also merit comment. The *ISL1* (MIM 600366) gene encodes a tissue-specific transcription factor that regulates pancreas development and expression of the insulin gene. Tests for linkage of *ISL1* with diabetes were negative in Nigerian and African American populations (Tanizawa et al. 1994) and a white French population (Riggs et al. 1995). The follistatin gene, *FST* (MIM 136470), which regulates the relative proportion of endocrine and exocrine tissue in the pancreas, also lies within this region on chromosome 5. Urbanek et al. (1999) recently reported evidence of linkage between *FST* and *PCO1* (MIM 184700), polycystic ovary syndrome, which shares many features with type 2 diabetes, especially insulin resistance (Urbanek et al. 1999). To our knowledge no tests of linkage between *FST* and diabetes related phenotypes have been conducted to date. By contrast to the regions of chromosome 3 and chromosome 5 that are linked to diabetes, the two regions of linkage on chromosome 12 are relatively gene poor and, with the exception already noted of *HNF-1α* near the linkage peak at 132cM, do not contain obvious candidates. The linkage region on the X chromosome is a similarly gene-poor region.

As described in the results section, we chose to type additional markers in three regions in the phase 1 white samples and two regions in the phase 1 Mexican American samples. The regions were selected on the basis of the strength of the LOD score in the region and other published reports concerning the region. Given the results that we report for these five regions, typing additional markers in regions showing suggestive or significant evidence for linkage may increase the evidence for linkage and narrow the 1-LOD-unit support interval. In two of these regions (chromosomes 5 and 12 in the whites), we typed additional markers so that the average distance between markers was 2 cM. The evidence for linkage derived from these maps was very similar to that derived from markers spaced ~5 cM apart, suggesting that typing markers <5 cM apart may have limited utility for linkage mapping.

The results obtained for the phase 1 samples of the

GENNID study generally were not replicated in the other ethnic groups. This may be because the ethnic groups have very distinct population histories and, in each population, different genes are responsible for disease susceptibility. Results shown for the phase 1 samples were not replicated in the phase 2 samples. This may be due to the fact that the ascertainment criteria in the second phase resulted in collecting mostly sibling pairs, in contrast to the more extended families (containing sibling pairs) collected in the first phase (see Methods). Of 1,568 relative pairs with DNA in the phase 1 white sample, 884 pairs represent first-degree relative relationships, and 636 pairs represent second-degree relationships. For the phase 2 white sample, 531 of the 570 possible pairs represent first-degree relationships and 36 represent second-degree relationships. In the phase 1 Mexican American sample, 663 and 554 of the 1,252 possible pairs represent first-degree and second-degree relationships, respectively. In the phase 2 Mexican American sample, 325 and 5 of the 338 possible pairs represent first- and second-degree relationships, respectively. It is also possible that the lack of replication in the phase 2 samples is because of the number of genes likely to be involved in the etiology of diabetes and the difficulties of replicating results without very large sample sizes (Suarez et al. 1994). Nonetheless, in our analyses, the more extended families provided more information and flexibility. It has been our experience that linkage results are rarely replicated in other populations, leading us to the conclusion that perhaps very large family collections are more desirable than replication datasets that are limited largely to sibling pairs. In addition to large numbers of families, it is critical that samples suitable for fine mapping (in the same population) also be collected in conjunction with the initial family resource. This could be accomplished by collecting unaffected siblings, parents, or unrelated control individuals.

While none of the studies undertaken to date is likely to have the power to identify all diabetes genes (Suarez et al. 1994), this study identified linkages on chromosome 3 and 12 that were highlighted by others. Linkage on chromosome 3 (LOD score = 3.91) is significant at the most stringent level (Lander and Kruglyak 1995) and linkages on chromosome 5, 12, and X are suggestive. Further, these results have identified interactions, although unreplicated ones, between these linkages, which, according to Cox et al. (1999), provides additional evidence that loci from these regions contribute to disease susceptibility. Although the biological model underpinning the evidence for epistatic interactions between potential disease-susceptibility loci on chromosomes X and 12 and 5 and 12 is not readily apparent, these interactions may prove useful in identifying the disease-susceptibility genes. The strength of the evidence for linkage to broad diabetes phenotypes and the additional reports in the literature suggesting these regions are linked in other populations give us confidence that these regions harbor genes that are involved in type 2 diabetes.

This study represents a collaborative effort between industry, academia, and a nonprofit organization (the American Diabetes Association). Given the likely need for large and diverse family collections to identify and confirm linkages to type 2 diabetes and other genetically complex diseases, this type of collaboration could provide a model for further work. The genotype and phenotype data described in this paper are publicly available by request from the American Diabetes Association.

Acknowledgments

The authors gratefully acknowledge Gopi Gopalam, Kavita Gopalam, Ray Grimaila, Pam St. Jean, Derek Traughber, David P. Yarnall (Glaxo Wellcome), and Gerald Asp (Axys), for technical assistance; and Lee Babiss, Steven Jacobs, Fred Fiedorek (Glaxo Wellcome), Jeff Hall, and Tim Harris (Axys), for advice and support. The authors also thank the investigators and staff at the GENNID Family Acquisition Centers and the participants in that study of the genetics of diabetes. The patient collection was supported in part by the American Diabetes Association. Genetic material was collected by and families were characterized by the American Diabetes Association GENNID Study Group, which is comprised of Eric Boerwinkle, Ph.D., University of Texas Health Science Center; John Buse, M.D., Ph.D., University of North Carolina; Ralph DeFronzo, M.D., University of Texas Health Science Center; David Ehrmann, M.D., University of Chicago; Steven C. Elbein, M.D., University of Utah/University of Arkansas; Wilfred Fujimoto, M.D., and Steven E. Kahn, M.B., Ch.B., University of Washington; Craig L. Hanis, Ph.D., University of Texas Health Science Center; Richard A. Mulivor, Ph.D., and Jeanne C. Beck, Ph.D., Coriell Cell Repositories; Jill Norris, Ph.D., University of Colorado School of Medicine; M. Alan Permutt, M.D., Washington University School of Medicine; Leslie Raffel, M.D., Cedars-Sinai Medical Center; and David C. Robbins, M.D., Medlantic Research Institute.

Electronic-Database Information

Accession numbers and URLs for data in this article are as follows:

Center for Medical Genetics, Marshfield, Wisconsin, http://www.marshmed.org/genetics (for the Marshfield map)

Online Mendelian Inheritance in Man (OMIM), http://www.ncbi.nlm.nih.gov/Omim (for NIDDM1 [601283], NIDDM2 [601407], MODY3 [600496], GNAI2 [139360], PIC3R1 [171833], ISL1 [600366], FST [136470], and PCO1 [184700])

References

Bektas A, Sprenant M, Wogan LT, Plengvidhya N, Rich SS, Doria A (1999) Evidence of a novel type 2 diabetes locus 50 cM centromeric to *NIDDM2* on chromosome 12q. Diabetes 48:2246–2251

Bell G, Cox N, Lindner T, Concannon P, Spielman R, Boerwinkle E, Hanis C (1997) Genetics of NIDDM in the Mexican-Americans of Starr County, Texas: an update. Diabet Rev 5:277–283

Bowden D, Sale M, Howard T, Qadri A, Spray B, Rothschild C, Akots G, et al (1997) 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

Chen J, Guo J, Moxham C, Wang H-Y, Malbon C (1997) Tissue-specific expression of Q205L Gia2 in vivo mimics insulin action. J Mol Med 75:283–289

Cox N, Frigge M, Nicolae D, Concannon P, Hanis C, Bell G, Kong A (1999) Loci on chromosomes 2 (NIDDM1) and 15 interact to increase susceptibility to diabetes in Mexican Americans. Nat Genet 21:213–215

Duggirala R, Blangero J, Almasy L, Dyer T, Williams K, Leach R, O'Connell P, et al (1999) Linkage of type 2 diabetes mellitus and age of diabetes onset to a genetic location on chromosome 10q in Mexican Americans. Am J Hum Genet 64:1127–1140

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

Elbein S, Hoffman M, Teng K, Leppert M, Hasstedt S (1999) A genome-wide search for type 2 diabetes susceptibility genes in Utah Caucasians. Diabetes 48:1175–1182

Froguel P, Zouali H, Vionnet N, Velho G, Vaxillaire M, Sun F, Lesage S, et al (1993) Familial hyperglycemia due to mutations in glucokinase. Definition of a subtype of diabetes mellitus. N Engl J Med 328:697–702

Ghosh S, Watanabe R, Hauser E, Valle T, Magnuson V, Erdos M, Langefield C, et al (1999) Type 2 diabetes: evidence for linkage on chromosome 20 in 716 Finnish affected sib pairs. Proc Natl Acad Sci USA 96:2198–2203

Hani EH, Clement K, Velho G, Vionnet N, Hager J, Philippi A, Dina C, et al (1997) Genetic studies of the sulfonylurea receptor gene locus in NIDDM and in morbid obesity among French Caucasians. Diabetes 46:688–694

Hanis C, Boerwinkle E, Chakraborty R, Ellsworth D, 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

Hansen T, Andersen CB, Echwald SM, Urhammer SA, Clausen JO, Vestergaard H, Owens D, et al (1997) Identification of a common amino acid polymorphism in the p85alpha regulatory subunit of phosphatidylinositol 3-kinase. Diabetes 46:494–501

Hanson R, Ehm M, Pettitt D, Prochazka M, Foroud T, Kobes S, Baier L, et al (1998) An autosomal genomic scan for loci linked to type 2 diabetes mellitus and body mass index in Pima Indians: an obesity-diabetes locus at 11Q23-25. Am J Hum Genet 63:1130–1138

Kawanishi M, Tamori Y, Masugi J, Mori H, Ito C, Hansen T, Andersen CB, et al (1997) Prevalence of a polymorphism of the phosphatidylinositol 3-kinase p85 alpha regulatory subunit (codon 326 Met→Ile) in Japanese NIDDM patients. Diabetes Care 20:1043

Kong A, Cox N (1997) Allele-sharing models: LOD scores and accurate linkage tests. Am J Hum Genet 61:1179–1188

Kruglyak L, Daly M, Reeve-Daly M, Lander E (1996) Parametric and nonparametric linkage analysis: a unified multipoint approach. Am J Hum Genet 58:1347–1363

Lander E, 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

Mahtani MM, Widen 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

Moxham C, Malbon C (1996) Insulin action impaired by deficiency of the G-protein subunit Gia2. Nature 379:840–844

National Diabetes Data Group (1979) Classification and diagnosis of diabetes mellitus and other categories of glucose intolerance. Diabetes 28:1039–1057

Raffel L, Robbins D, Norris J, Boerwinkle E, DeFronzo R, Elbein S, Fujimoto W, et al (1996) The GENNID Study: a resource for mapping the genes that cause NIDDM. Diabetes Care 19:864–872

Rich S (1990) Mapping genes in diabetes: genetic epidemiological perspective. Diabetes 39:1315–1319

Riggs A, Tanizawa Y, Aoki M, Wasson J, Ferrer J, Rabin D, Vaxillaire M, et al (1995) Characterization of the LIM/homeodomain gene islet-1 and single nucleotide screening in NIDDM. Diabetes 44:689–694

Suarez B, Hampe C, Eerdewegh PV (1994) Problems of replicating linkage claims in psychiatry. In: Gershon E, Cloninger C (eds) Genetic approaches to mental disorders. American Psychiatric Press, Inc., Washington, DC, pp 23–46

Tanizawa Y, Riggs AC, Dagogo-Jack S, Vaxillaire M, Froguel P, Liu L, Donis-Keller H, et al (1994) Isolation of the human LIM/homeodomain gene islet-1 and identification of a simple sequence repeat polymorphism. Diabetes 43:935–941

Terauchi Y, Tsuji Y, Satoh S, Minoura H, Murakami K, Okuno A, Inukai K, et al (1999) Increased insulin sensitivity and hypoglycaemia in mice lacking the p85 alpha subunit of phosphoinositide 3-kinase. Nat Genet 21:230–235

Expert Committee on the Diagnosis and Classification of Diabetes Mellitus (1998) Report of the Expert Committee on the Diagnosis and Classification of Diabetes Mellitus. Diabetes Care 20:1183–1197

Turner R, Hattersley A, Shaw J, Levy J (1995) Type II diabetes: clinical aspects of molecular biological studies. Diabetes 44: 1–10

Urbanek M, Legro R, Driscoll D, Assis R, Ehrman D, Norman R, Strauss JF 3d, et al (1999) Thirty-seven candidate genes for polycystic ovary syndrome: strongest evidence for linkage is with follisatin. Proc Natl Acad Sci USA 96:8573–8578

Yamagata K, Furuta H, Oda N, Kaisaki P, Menzel S, Cox N,

Fajans S, et al (1996*a*) Mutations in the hepatocyte nuclear factor-4α gene in maturity-onset diabetes of the young (MODY1). Nature 384:458–460

Yamagata K, Furuta H, Oda N, Kaisaki PJ, Menzel S, Cox NJ, Fajans SS, et al (1996*b*) Mutations in the hepatocyte nuclear factor-1α gene in maturity-onset diabetes of the young (MODY 3). Nature 384:455–458

Yuan B, Vaske D, Weber J, Beck J, Sheffield V (1997) Improved set of short tandem repeat polymorphisms for screening the human genome. Am J Hum Genet 60:459–460