Examination of Candidate Chromosomal Regions for Type 2 Diabetes Reveals a Susceptibility Locus on Human Chromosome 8p23.1

Marcus G. Pezzolesi,1 Moonsuk Nam,1,2,4 Terumasa Nagase,1,2 Tomasz Klupa,1,2 Jonathon S. Dunn,1 Wojciech M. Mlynarski,1,2 Stephen S. Rich,3 James H. Warram,1 and Andrzej S. Krolewski1,2

In a panel of large Caucasian pedigrees, we genotyped markers in eight chromosomal regions previously reported as supporting linkage with type 2 diabetes. We previously reported significant linkage on chromosome 20q (maximum logarithm of odds score [MLS] = 2.79) in this panel. In the present analysis, candidate regions on 1q, 2q, 3q, 5q, 9q, and 10q yielded little evidence for linkage; a region on 2p (MLS = 1.64, P = 0.01 at position 9.0 cM) gave suggestive evidence of linkage; and a region on 8p (MLS = 3.67, P = 2.8 \times 10^{-3}, at position 7.6 cM) gave significant evidence of linkage. Conditional analyses were performed for both 2p and 8p regions and the region reported on 20q. The MLS for 2p increased from 1.64 to 1.79 (empirical P = 0.142) when conditioned for heterogeneity on 20q. The case was similar for 8p, where the MLS increased from 3.67 to 4.51 (empirical P = 0.023) when conditioned on families without evidence of linkage at 20q. In conclusion, our data support a type 2 diabetes susceptibility locus on chromosome 8p that appears to be independent of other susceptibility loci. Although we were able to replicate linkage in our pedigrees on chromosome 2p, we did not find evidence of linkage for regions on 1q, 2q, 3q, 5q, 9q, or 10q. Diabetes 53:486–491, 2004

Type 2 diabetes is a disease characterized by hyperglycemia, insulin resistance, and impaired insulin secretion (1). Although the pathophysiology of this disease is not well understood, evidence of familial clustering indicates that genetic susceptibility plays a role in disease development (2,3). Identification of genetic susceptibility factors, however, has been difficult, and, to date, the putative genes contributing to the etiology of type 2 diabetes remain unknown. Efforts to identify the genetic factors underlying susceptibility to type 2 diabetes have used both candidate gene approaches and linkage studies (4). More than a dozen genomewide scans have been completed from which various chromosomal regions with suggestive or significant evidence of linkage with type 2 diabetes have emerged (5–23). However, despite intense efforts to clone the susceptibility genes, the identification of these genes has been unsuccessful, with the single exception of calpain-10 (CAPN10) or NIDDM1 (24).

Efforts to clone type 2 diabetes susceptibility loci have been plagued by inconsistent findings. Evidence for linkage with type 2 diabetes has been reported on most autosomes (25). However, there has been little overlap among linkage results, with the notable exceptions being the findings on chromosomes 1q (8,11,15,19,23) and 20q (26–30). Factors contributing to the failure to replicate linkage results include the presence of environmentally determined phenocopies of type 2 diabetes as well as the complex nature of the inheritance of type 2 diabetes. The latter includes genetic and allelic heterogeneity, epistasis (gene-to-gene interaction), and gene-to-environment interactions.

In a recent study, we showed that the overall sibling genetic risk ratio (\(\lambda_S\)) for type 2 diabetes may be less than previously assumed (2,3) and is similar to that found in the general population. Moreover, the observed risk was comparable in siblings of index cases without a history of parental diabetes (>50% of all diabetic patients), suggesting a smaller overall genetic contribution to the diabetic etiology in these patients (3). In contrast, the sibling recurrence risk of diabetes in families having a diabetic parent was elevated (\(\lambda_S = 2.0\)). Furthermore, in a substantial proportion of these families, diabetes was transmitted vertically through three generations. A \(\lambda_S\) of 3.5 in this subset indicates a significantly higher risk to siblings. By studying families with a type 2 diabetes inheritance pattern consistent with a single, highly penetrant, dominant susceptibility locus, type 2 diabetes gene mapping may be achievable using classic linkage methods. Recently, we demonstrated the feasibility of this approach by collecting a panel of large families among whom type 2 diabetes was transmitted vertically. A putative locus for type 2 diabetes was identified on chromosome 20q13 using a panel of highly polymorphic microsatellite markers. Positive evidence for linkage was found for a 10-cM region between markers D20S119 and D20S428, with the strongest evi-
dence occurring at D20S196 (MLS = 2.79) (26,30). Using this approach of ascertaining highly informative (and more genetically homogeneous) families, it is anticipated that positional cloning of the putative susceptibility genes for type 2 diabetes will be achieved.

To further exploit this general approach and gain a better understanding of the inheritance of type 2 diabetes, chromosomal regions previously identified in families of European origin were investigated in our family collection. Criteria for selecting the chromosomal regions to be explored included suggestive evidence for linkage (MLS ≥1.5) in other studies (6,11–13,15,19,21,22). A subset of regions meeting these criteria were selected and subsequently examined in our panel of 43 large families with type 2 diabetes. A total of eight candidate chromosomal regions were considered in the present analysis: 1q, 2p, 2q, 3q, 5q, 8p, 9q, and 10q.

RESEARCH DESIGN AND METHODS

Families for the Joslin Study on the Genetics of Type 2 Diabetes were ascertained through index cases from the Joslin Clinic population; details of their recruitment have been previously described (26,30,31). Briefly, we sought to identify families with a pattern of occurrence of type 2 diabetes that was consistent with an autosomal-dominant mode of inheritance. The criteria used to identify eligible families consisted of 1) an index case and at least one sibling having been diagnosed with type 2 diabetes between ages 35 and 59 years, 2) the index case having been treated for at least 2 years with diet or oral agents, and 3) the occurrence of diabetes in at least three generations within each family.

All study participants completed medical and family history questionnaires and had fasting blood samples drawn for blood glucose determination, DNA extraction, and other biochemical measurements. All nondiabetic and diabetic individuals treated with oral agents or diet had blood samples drawn 2 h after a 75-g oral glucose challenge. Glucose levels were measured at the Joslin Clinic laboratory using a Synchron CX Delta system (Beckman Coulter, Fullerton, CA) or One-Touch home glucose meters (Lifescan, Newtown, PA). The latter measurements were re-expressed in terms of venous plasma glucose as measured in the Joslin Clinic laboratory. Serum insulin was measured by radioimmunoassay by Linco Research (St. Charles, MO).

In addition, measurements of height, weight, and blood pressure were taken. Diabetes and impaired glucose tolerance (IGT) status were diagnosed according to World Health Organization criteria (32).

A total of 534 individuals, including 247 with type 2 diabetes and 30 with IGT, from a collection of 43 families of European origin were included in the present study. This was the same panel of families studied by Khupu et al. (30).

To maintain the proportion of families with evidence of linkage from the unweighted analysis (while allowing any family to be considered “supporting” linkage), each simulation randomly assigned a weight of 1 to a subset of families. This new distribution of families supporting linkage was used to re-calculate the LOD score. The procedure was repeated 10,000 times, allowing the empirical P value to be estimated from the proportion of computer-generated replicates that exceeded the observed conditional MLS score.

RESULTS

Two-point and multipoint linkage analysis results obtained from analysis of the eight candidate chromosomal regions in all 43 families are summarized in Table 1. We were unable to replicate previously reported linkage of type 2 diabetes with markers on 1q, 2q, 3q, 5q, 9q, and 10q. However, some evidence for linkage was observed on 2p; the two-point linkage analysis yielded MLS = 1.95 at marker D2S166. This evidence was slightly reduced in multipoint analysis (MLS = 1.64, P = 0.01, at position 9.0 cM). The strongest evidence for linkage was found on chromosome 8p; the two-point linkage analysis yielded MLS = 1.72 (P = 0.05) for marker DSS1819 located at the telomeric boundary of the examined region. This evidence significantly increased in multipoint analysis (MLS = 3.62, P = 1.8 × 10^-4, at 83 cM).
position 9.0 cM). To examine this result further, we genotyped our families for eight additional microsatellite markers located in the p-telomeric region (0–10 cM). The linkage peak shifted to a more telomeric position (D8S1742, at position 7.6 cM) with the two-point MLS increasing to 3.08 ($P = 4.0 \times 10^{-5}$) and the multipoint MLS increasing to 3.67 ($P = 2.8 \times 10^{-5}$) (Fig. 1).

In all, 8 of the 43 families had an estimated MLS $\geq 0.5$ at marker D8S1742. One family with nine affected individuals in three generations, all sharing the same haplotype, contributed the greatest support for linkage. After removal of this family, the evidence for linkage in the seven remaining families was still significant (MLS = 2.23, $P = 8.4 \times 10^{-4}$), with the MLS occurring at the same chromosomal location.

To examine epistatic and heterogeneity effects among the chromosomal regions linked with type 2 diabetes, we performed conditional analyses using the results for 2p and 8p and the previously reported results for 20q (30) (Table 2). The $Z_{\text{h}}$ scores for each pedigree at the markers used for these analyses and a summary of families receiving weights of 1 and 0 under each model can be viewed in online appendix C. For each region, the multipoint MLS for the total family collection is provided along with the MLS for the family subsets. The MLS for the region on chromosome 2p decreased from 1.64 in all families to 0.72 when only families with negative $Z_{\text{h}}$ scores on 8p were examined; this finding supported the absence of an epistatic effect between the two chromosomal regions. Similarly, we found a decrease in MLS when linkage with 2p was analyzed in a subset of families with negative $Z_{\text{h}}$ scores on 8p, suggesting the absence of heterogeneity between these two regions. The results of the conditional analyses between 2p and 20q differed. Whereas the MLS for 2p under an epistatic model decreased significantly (almost to zero), it increased (empirical $P = 0.142$) above the MLS in all families when the analysis was carried out only in families with negative $Z_{\text{h}}$ scores on 20q (from 1.64 to 1.79).

A similar trend was observed when a conditional analysis was performed for chromosome 8p; however, the lack of an epistatic effect and the presence of a heterogeneity effect were much more pronounced (Table 2). Evidence of epistatic and heterogeneity effects between 8p and 2p were not significant. The MLS for 8p conditioned under an epistatic model with evidence at 20q decreased to 0.15 from 3.67. Significant evidence of heterogeneity (empirical $P < 0.023$) was observed between 8p and 20q; that is, evidence for linkage with the former region was particularly strong in families that did not show evidence for linkage with chromosome 20q. The MLS increased from 3.67 in all families to 4.51 in a subset of families not linked with 20q.

The results from the conditional analysis (heterogeneity model) for chromosomal regions on 8p and 20q are shown in Fig. 1. MLS statistics obtained in all 43 families for chromosome 8p are plotted over the 0–45 cM interval. The MLS is 3.67 at marker D8S1742, with a 10-cM support interval between markers DSS518 and DSS351. When families with negative $Z_{\text{h}}$ scores on chromosome 20 were removed from this analysis, the MLS for the remaining families increased to 4.51 ($P = 1.5 \times 10^{-5}$). The maximum MLS was observed at the same marker in the two analyses; however, the support interval was markedly reduced to $\approx 3.5$ cM (DSS518 to DSS277) when the families supporting apparent heterogeneity were removed.

Table 3 compares clinical characteristics of the 55 diabetic individuals in the families linked with 8p with those of the 214 diabetic individuals in the remaining 37 unlinked families. The age at diagnosis was similar for affected diabetic members of the families linked to chromosome 8p and diabetic members of the unlinked fami-

![FIG. 1. Results from multipoint linkage analysis. Allele-sharing LOD scores are shown for chromosome 8p for all 43 families (——) and for families unlinked to chromosome 20q (—).](image-url)
lies. The affected individuals from the linked families appear to be leaner, and fewer were treated with insulin compared with the unlinked families. Affected members from the linked families also had higher fasting glucose and lower 2-h serum insulin levels. When we controlled for fasting plasma glucose values using multiple regression analysis, this difference (between diabetic family members linked and not linked with 8p) disappeared.

**DISCUSSION**

Selecting families through type 2 diabetes index cases and when they have a history of vertical transmission of diabetes through an affected parent has the possible benefit of reducing genetic and phenotypic heterogeneity. Reducing the variation while increasing the genetic effect may increase statistical power and facilitate the search for type 2 diabetes susceptibility genes (3). Previously, this study design has had success in identifying genes responsible for maturity-onset diabetes of the young (36). Applying a similar approach to the search for genes contributing to a form of type 2 diabetes that is more typical, yet with enrichment of genetic effects, should result in the identification of susceptibility genes that may lie in more important physiological pathways.

We have previously demonstrated that this approach can be profitable by identifying a putative locus for type 2 diabetes on chromosome 20q using a collection of large families with vertical transmission of disease (26,30). To exploit this approach further, we used our families to replicate evidence for linkage reported previously by other authors. Eight candidate chromosomal regions met our criteria and were chosen for examination in our panel of 43 large families with type 2 diabetes segregating as autosomal dominant disorder.

Our most significant finding was observed on chromosome 8p. We identified a major susceptibility locus while attempting to replicate initial findings from a genome scan reported by Wiltshire et al. in the U.K. (19). However, we identified a possible type 2 diabetes susceptibility region more telomeric to the original report, with a support interval of ~10 cM, between D8S158 and D8S351. The evidence for linkage reported in the U.K. genome scan, MLS = 2.55 at position 42.2 cM, is very broad and extends into the region observed in our families. In addition, Elbein et al. (11) reported modest evidence for linkage of type 2 diabetes to chromosome 8p. In multigenerational families of European ancestry, suggestive evidence for linkage was observed at marker D8S136 (LOD score = 1.35 at position 44 cM) and D8S87/D8S532 (LOD score = 1.36 at position 61–65 cM).

Presently, one potential candidate gene for type 2 diabetes has been identified in our support interval: protein

<table>
<thead>
<tr>
<th>Chromosomal region</th>
<th>Region and specific conditional marker</th>
<th>All families</th>
<th>Linked families at specific marker (epistasis)</th>
<th>Nonlinked families at specific marker (heterogeneity)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2p25 at D2S2166</td>
<td>8p23 at D8S1742</td>
<td>1.64</td>
<td>0.72</td>
<td>0.91</td>
</tr>
<tr>
<td></td>
<td>20q13 at D20S196</td>
<td>1.64</td>
<td>0.20</td>
<td>1.79*</td>
</tr>
<tr>
<td>8p23 at D8S1742</td>
<td>2p25 at D2S2166</td>
<td>3.67</td>
<td>3.39</td>
<td>0.47</td>
</tr>
<tr>
<td></td>
<td>20q13 at D20S196</td>
<td>3.67</td>
<td>0.15</td>
<td>4.51†</td>
</tr>
</tbody>
</table>

*Empirical \( P \leq 0.15; \) †empirical \( P \leq 0.05. \) Empirical \( P \) values were estimated based on the number of computer-generated replicates that exceeded the observed conditional MLS score.

**TABLE 3**

Clinical characteristics of diabetic members from families linked and unlinked to chromosome 8p

<table>
<thead>
<tr>
<th></th>
<th>Linked families</th>
<th>Unlinked Families</th>
<th>( P )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of families</td>
<td>8</td>
<td>35</td>
<td>—</td>
</tr>
<tr>
<td>Number of diabetic subjects (M/F)</td>
<td>28/27</td>
<td>109/105</td>
<td>—</td>
</tr>
<tr>
<td>Age at diagnosis (years)</td>
<td>50 ± 13</td>
<td>46 ± 15</td>
<td>0.06</td>
</tr>
<tr>
<td>Age at examination (years)</td>
<td>60 ± 12</td>
<td>59 ± 14</td>
<td>0.33</td>
</tr>
<tr>
<td>Percent IBW at examination</td>
<td>134 ± 23</td>
<td>140 ± 34</td>
<td>0.11</td>
</tr>
<tr>
<td>Lifetime maximum percent IBW</td>
<td>148 ± 28</td>
<td>156 ± 41</td>
<td>0.09</td>
</tr>
<tr>
<td>Systolic blood pressure</td>
<td>137 ± 17</td>
<td>136 ± 20</td>
<td>0.77</td>
</tr>
<tr>
<td>Diastolic blood pressure</td>
<td>79 ± 11</td>
<td>79 ± 10</td>
<td>0.98</td>
</tr>
<tr>
<td>Albumin/creatinine ratio</td>
<td>653 ± 1,902</td>
<td>475 ± 1,411</td>
<td>0.45</td>
</tr>
<tr>
<td>Treatment at examination (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Diet alone</td>
<td>20.0</td>
<td>17.3</td>
<td></td>
</tr>
<tr>
<td>Oral agents only</td>
<td>47.3</td>
<td>37.4</td>
<td>0.24</td>
</tr>
<tr>
<td>Insulin</td>
<td>32.7</td>
<td>45.3</td>
<td></td>
</tr>
<tr>
<td>Non-insulin-treated diabetic patients</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fasting blood glucose (mg/dl)</td>
<td>197 ± 80</td>
<td>156 ± 61</td>
<td>0.008</td>
</tr>
<tr>
<td>2-h blood glucose (mg/dl)</td>
<td>351 ± 90</td>
<td>270 ± 80</td>
<td>0.005</td>
</tr>
<tr>
<td>Fasting serum insulin (( \mu )U/ml)</td>
<td>12.4 ± 7.4</td>
<td>16.7 ± 12.9</td>
<td>0.12</td>
</tr>
<tr>
<td>2-h serum insulin (( \mu )U/ml)</td>
<td>22.4 ± 15.4</td>
<td>52.2 ± 65.7</td>
<td>0.01</td>
</tr>
</tbody>
</table>

Data are percent or means ± SD, unless otherwise indicated. IBW, ideal body weight.
phosphatase 1 regulatory (inhibitor) subunit 3 (PP1R3B) (37). Because serine/threonine protein phosphatase (PPase) inactivation promotes insulin exocytosis in the β-cell, alterations in this PPase inhibitor may influence glucose-sensitive insulin secretion (38). This gene is being examined for DNA sequence differences in our linked families. The other transcripts in our support interval on 8p remain incompletely characterized, so other positional candidate genes may exist and yet (at this point) remain to be examined.

The clinical features of diabetic patients from our 8p-linked families were similar to those of diabetic patients from families not supporting evidence for linkage to this region. This suggests that a defect(s) in the putative susceptibility gene on 8p results in a phenotype consistent with the common insulin-resistant form of type 2 diabetes. Families from our collection that have been linked with chromosome 20q share this same phenotype.

Modeling the complex pattern of inheritance of type 2 diabetes, as well as other genetically complex traits, is challenging. Genetic and allelic heterogeneity, epistasis, and environmental factors each interact to confer disease susceptibility. To improve the ability to map genes in such complex disorders, two-locus linkage analyses have been proposed (39,40). Using the approach described by Cox et al. (41), we investigated these mechanisms for loci where evidence of linkage was observed. Our most significant observation was that of heterogeneity among loci located on chromosomes 8p and 20q (empirical \( P < 0.023 \)). Interpreting the significance of this interaction, however, requires caution. Although our model approach attempted to minimize spurious observations, this possibility remains.

We observed suggestive linkage to type 2 diabetes on chromosome 2p (marker D2S2166, MLS = 1.64, at position 9.0 cM). In this same region, MLS = 0.87 at 5.5 cM, modest evidence for linkage was reported by Ghosh et al. in a Finnish-U.S. study (13). Moreover, in the same study, the investigators examined evidence for linkage on chromosome 20q in sib-pairs conditional with evidence of linkage on chromosome 2p (epistatic effect). The MLS on 20q increased from 2.15 to 5.06. Repetition of this conditional analysis based on linkage at a locus closer to our locus (D2S2166) increased the MLS further. In contrast to these findings, we were unable to replicate evidence for epistasis upon repeating this two-locus analysis in our dataset.

Similar interactions among loci associated with type 2 diabetes have previously been reported. Cox et al. (41) identified an epistatic interaction between the NIDDM1 region of chromosome 2 and the CYP19 region of chromosome 15. Wilshire et al. (19) have also shown evidence of epistasis at loci on 1q24.2 and 10q23.3 and heterogeneity between loci on 1q24.2 and 5q32.

Evidence of a type 2 diabetes susceptibility locus residing on chromosome 1q has been reported by several groups. Elbein et al. (11) reported significant evidence for linkage on chromosome 1q between markers CRP and APOA2 in a study of 19 multigenerational families. Others (15,19,22,23) have since replicated this evidence in their respective populations. In addition, significant evidence for linkage to this region was reported from the Pima Indian genome scan (8). All six reports have identified linkage signals across a region spanning ~25 cM, from positions 163 cM to 189 cM on chromosome 1q. We observed little evidence for linkage to this region in our families (MLS = 0.05, at position 205 cM).

Several additional regions were examined and, as was the case with chromosome 1q, we were unable to replicate evidence for linkage in this dataset. These chromosomal regions included 2q (260–263 cM), 3q (201–214 cM), 5q (70–85 cM), 9q (51–84 cM), and 10q (75–156 cM) (6,12,15,19,21,22). The possibility cannot be dismissed that genes in these regions may contribute to disease susceptibility in our families; however, these contributions are likely to be modest. Given the different evolutionary histories of the multiple family collections, it is not expected that all type 2 diabetes susceptibility loci will be equivalently important in all populations. Because our family collection favored an autosomal dominant mode of inheritance, all gene effects may not be represented in these pedigrees. Our data indicate that genetic defects at two major loci, located on chromosomes 8p23.1 and 20q13.1, act independently to confer susceptibility to type 2 diabetes. These two major loci accounted for susceptibility to diabetes in almost 40% of the families in our collection. The major locus/loci responsible for diabetes in the remaining 60% of our families has yet to be identified. Through continued research efforts, including a genome-wide screen currently under way, we hope to be able to provide further insight toward understanding the genetic basis of disease in these families.

ACKNOWLEDGMENTS

This research was supported by National Institutes of Health Grants DK-47475 (A.S.K.) and DK-30836 (Genetics Core of the Diabetes Endocrinology Research Center at Joslin Diabetes Center).

We thank Betsy Palecek and Susie Tully for their efforts to recruit and examine new members of the Joslin Family Collection for the Genetics of Type 2 Diabetes. We would also like to thank Adam Smiles for maintaining the genetic and phenotypic databases for this study.

REFERENCES

8. Hanson RL, Ehm MG, Pettitt DJ, Prochazka M, Thompson DB, Timberlake
7. McCarthy MI: A genomewide scan for loci predisposing to type 2 diabetes
6. Hanson RL, Ehm MG, Pettitt DJ, Prochazka M, Thompson DB, Timberlake
8. Hanson RL, Ehm MG, Pettitt DJ, Prochazka M, Thompson DB, Timberlake
4. Bowden DW, Sales M, Howard TD, Qadri A, Spray BJ, Rothschild CB, Akots