Genome-wide scan for linkage to type 1 diabetes in 2,496 multiplex families from the Type 1 Diabetes Genetics Consortium

Patrick Concannon\textsuperscript{1,2}, Wei-Min Chen\textsuperscript{2,3}, Cécile Julier\textsuperscript{4}, Grant Morahan\textsuperscript{5}, Beena Akolkar\textsuperscript{6}, Henry A. Erlich\textsuperscript{7}, Joan E. Hilner\textsuperscript{8}, Jørn Nerup\textsuperscript{9}, Concepcion Nierras\textsuperscript{10}, Flemming Pociot\textsuperscript{9}, John A. Todd\textsuperscript{11}, Stephen S. Rich\textsuperscript{2}, and the Type 1 Diabetes Genetics Consortium

\textsuperscript{1}Department of Biochemistry and Molecular Genetics, and \textsuperscript{2}Center for Public Health Genomics, University of Virginia, Charlottesville, VA, USA
\textsuperscript{3}Department of Public Health Sciences, Division of Biostatistics and Epidemiology, University of Virginia, Charlottesville, VA, USA
\textsuperscript{4}Inserm U730, Centre National de Génotypage, Evry, France
\textsuperscript{5}Centre for Diabetes Research, The Western Australian Institute for Medical Research, and Centre for Medical Research, University of Western Australia, Perth, Australia
\textsuperscript{6}Division of Diabetes, Endocrinology, and Metabolic Diseases, The National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK), National Institutes of Health, Bethesda, MD, USA
\textsuperscript{7}Roche Molecular Systems, Pleasanton, CA, USA
\textsuperscript{8}Public Health Sciences, Wake Forest University School of Medicine, Winston-Salem, NC, USA
\textsuperscript{9}Steno Diabetes Center, Gentofte, Denmark
\textsuperscript{10}Juvenile Diabetes Research Foundation, New York, NY, USA
\textsuperscript{11}Juvenile Diabetes Research Foundation/Wellcome Trust Diabetes and Inflammation Laboratory, Department of Medical Genetics, Cambridge Institute for Medical Research, University of Cambridge, Cambridge, UK

Correspondence:
Patrick Concannon, Department of Biochemistry and Molecular Genetics, University of Virginia, Box 800733, Charlottesville, VA 22908 USA, E-mail: patcon@virginia.edu

Submitted 6 November 2008 and accepted 5 January 2009.

Additional information for this article can be found in an online appendix at http://diabetes.diabetesjournals.org
ABSTRACT

Objective. Type 1 diabetes arises from the actions of multiple genetic and environmental risk factors. Considerable success at identifying common genetic variants that contribute to type 1 diabetes risk has come from genetic association (primarily case-control) studies. However, such studies have limited power to detect genes containing multiple rare variants that contribute significantly to disease risk.

Research Design and Methods. The Type 1 Diabetes Genetics Consortium (T1DGC) has assembled a collection of 2,496 multiplex type 1 diabetes families from 9 geographical regions containing 2,658 affected sib-pairs (ASPs). We describe the results of a genome-wide scan for linkage to type 1 diabetes in the T1DGC family collection.

Results. Significant evidence of linkage to type 1 diabetes was confirmed at the HLA region on chromosome 6p21.3 (LOD=213.2). There was further evidence of linkage to T1D on 6q which could not be accounted for by the major linkage signal at the HLA class II loci on chromosome 6p21. Suggestive evidence of linkage (LOD≥2.2) was observed near CTLA4 on chromosome 2q32.3 (LOD=3.28) and near INS (LOD=3.16) on chromosome 11p15.5. Some evidence for linkage was also detected at two regions on chromosome 19 (LODS=2.84 and 2.54).

Conclusions. Five non-HLA chromosome regions showed some evidence of linkage to type 1 diabetes. A number of previously proposed type 1 diabetes susceptibility loci, based on smaller ASP numbers, showed limited or no evidence of linkage to disease. Low frequency susceptibility variants or clusters of loci with common alleles could contribute to the linkage signals observed.
Type 1 diabetes develops when the insulin-secreting β cells in the pancreas are depleted by an autoimmune process of unknown origin. Diagnosis occurs when sufficient β cell mass is lost that blood glucose levels cannot be maintained in the physiologic range. While insulin treatment for type 1 diabetes is life-saving, development of effective preventive therapies could be enhanced by improved prediction and a better understanding of the underlying disease mechanism, particularly events occurring during the extended preclinical period. Such considerations have motivated studies directed toward identifying the genetic risk factors that contribute to the disorder that might provide novel insights into disease etiology and targets for preventive therapy.

A number of genome-wide scans for linkage to type 1 diabetes have been reported (1-7). All such studies consistently find evidence of linkage of type 1 diabetes to the HLA region on chromosome 6p21. An additional 15-20 regions with evidence of linkage to type 1 diabetes have been reported, but many of these regions have not been reproduced in independent studies. Several attempts have been made to address this problem through merging of data-sets from different smaller studies (5;7), but differences in marker sets and allele coding have limited the power of these efforts.

In order to address issues of sample size and consistency and to clarify the evidence for linkage to type 1 diabetes, the T1DGC has conducted the largest ASP linkage study in type 1 diabetes to date. This study includes many (N=1,189) of the type 1 diabetes families used in previous smaller linkage studies, but the majority of families (N=1,307) are newly recruited. More importantly, all samples, both retrospective and newly recruited, were independently genotyped expressly for this study at a single laboratory using a single platform and a common marker set. Analyses of these data provided evidence for linkage to type 1 diabetes in six regions on 4 different chromosomes.

**RESEARCH DESIGN AND METHODS**

**Subjects.** The families included in the analyses were ascertained from 9 different sources: the four T1DGC regional networks, Asia-Pacific (N=228), Europe (N=585), North America (N=370) and UK (N=124), the Diabetes UK Warren 1 Repository (N=429), the Human Biological Data Interchange repository (N=424), the Joslin Diabetes Center (N=112), the Steno Diabetes Center (N=146), and Sardinia (N=78). To be included in the study, a family had to have at least one affected sibling pair available for sampling; availability of one or both biological parents was preferred but not required. An individual was designated as affected if they had documented type 1 diabetes with onset at less than 35 years of age, had used insulin within 6 months of diagnosis, and had no concomitant disease or disorder associated with diabetes. A total of 2,496 families were genotyped. Most families, 2,287 (91.6%), had exactly two affected full siblings; 101 (4.0%) had three and 7 (0.3%) had four or more. Among 2,657 ASPs in total, both parents were available for 69.6% of the ASPs; 18.9% had only a single parent, and 11.6% had neither parent available. A few extended pedigrees also provided 13 affected half-sib pairs and 22 affected-affected avuncular pairs.

**Genotyping.** Genotyping was carried out at the Center for Inherited Disease Research (CIDR) using the Illumina Human Linkage-12 Genotyping Beadchip which contains 6,090 SNPs with an average genome-wide spacing of 0.58 cM.

**Statistical analysis.** Prior to statistical genetic analyses, genotype data were
evaluated for Mendelian errors using PedCheck (8). PREST (9) was used to estimate the likelihood of each specified relationship in pedigrees given the genotyping information. Based on these analyses, 43 families were removed from further analysis due to the presence of a duplicate family, a duplicate sample within the family, non-resolvable family errors or missing genotype information. Merlin (10) was used to identify and resolve inconsistencies within families. Finally, 199 affected individuals that were not genotyped were removed from the analysis. The Kong and Cox linear NPL model (11) implemented in Merlin was used for the NPL linkage analysis. Marker-marker linkage disequilibrium was modeled in the linkage analysis using the maximum-likelihood clustered marker approach implemented in Merlin with a cutoff of $r^2>0.5$.

RESULTS

Figure 1 shows the results of the genome-wide analysis for linkage to type 1 diabetes in the entire T1DGC family collection. The average information content across all chromosomes was 0.924 and ranged from 0.904 on chromosome 20 to 0.939 on chromosomes 6 and X. Table 1 lists the regions where evidence for linkage exceeded the threshold for suggestive linkage (LOD≥2.2) in an ASP linkage study. The HLA region on chromosome 6p21.3 was the only region to display significant evidence of linkage to type 1 diabetes (LOD=213.2). Four additional regions provided suggestive evidence of linkage. These included a region on chromosome 2q (LOD=3.28) where there have been multiple past reports of linkage to type 1 diabetes, and the region surrounding the insulin gene at 11p15.5 (LOD=3.16) that has been previously implicated in type 1 diabetes susceptibility by both linkage and association studies. The remaining two regions were on chromosome 19 (LODs of 2.84 and 2.54) and are separated from each other by 48.5 cM; evidence of linkage to chromosome 19 has been reported before but not to any level of confidence that justified specific attention. Other previously reported type 1 diabetes loci, which quite often were given IDDM designations, were not confirmed in the current study (Table 2).

When considering evidence for linkage to type 1 diabetes on chromosome 6 that might possibly be due to loci other than the classical HLA genes it was necessary to take into account the significant allele sharing at HLA, where LOD scores greater than 3.6 were detected over an interval of ~100 cM surrounding the locus. We searched for evidence of additional type 1 diabetes risk loci by calculating expected LOD (ELOD) scores based upon allele sharing at HLA along the length of chromosome 6 (Supplementary Figure 1). The observed LOD score exceeded the ELOD+3.6 in 3 regions. One of these regions, from 81.5 cM to 101.5 cM (maximum LOD-ELOD=10.8) overlapped with the location of the previously reported IDDM15 locus (6;12).

DISCUSSION

The T1DGC family collection includes the largest number of ASPs to be tested for linkage to type 1 diabetes in a genome-wide scan yet provides suggestive or significant evidence for only five loci linked to type 1 diabetes. In light of the numerous previous reports of loci linked to type 1 diabetes, we considered the relationship between the findings here and prior studies (Table 2).

The majority of studies, as in our current report, have observed significant evidence of linkage to type 1 diabetes at HLA and there is well-confirmed evidence of disease association in the same region. Several additional loci on chromosome 6, IDDM5, IDDM8 and IDDM15, have been proposed based on previous linkage studies (1;13;14). The current study provides little support for either IDDM5 or IDDM8 (Table 2). However,
we did obtain a regional maximum LOD of 5.43, at 190.0 cM on chromosome 6q27, where IDDM8 was previously localized (13), if linkage disequilibrium was not modeled in the analysis. This region is characterized by a relatively strong intermarker linkage disequilibrium which might explain this and previous reports of linkage in this region.

An additional region on chromosome 6q provided evidence for linkage to type 1 diabetes after adjusting for the effect of allele sharing at HLA. Previous studies have tested for linkage at IDDM15, which is located in this region, by calculating the expected LOD score at IDDM15 based upon the observed IBD sharing at HLA and the recombination fraction between HLA and D6S283 (7;12). The excess of the observed IBD sharing to that expected was then used as a measure of the independent contribution of IDDM15. Applying this approach to the current data set identified a broad (20 cM) region where the “adjusted” LOD exceeds 3.6. There remain uncertainties as to the precise location of IDDM15 since male and female recombination maps differ substantially in this region. Nevertheless, this region of increased LODs, after adjusting for the contribution of allele sharing at HLA, provides support for the existence of an additional linkage to type 1 diabetes in this region.

Chromosome 2q contains two confirmed type 1 diabetes risk loci near the IFIH1 and CTLA4 genes identified by association and three proposed loci, IDDM7, IDDM12 and IDDM13 for which suggestive evidence of linkage has been observed in some, but not all, studies (15-17). In the current report, there was no support for multiple distinct loci; suggestive evidence of linkage was observed in the 2q32.3 region, approximately midway between the previously proposed locations of IDDM7 and IDDM12. The INS locus at 11p15.5 is also well established as a risk locus for type 1 diabetes (18). Previous studies have provided significant evidence for both association and linkage to this site. The current study also finds suggestive evidence of linkage in this region.

Two regions on chromosome 19 displayed suggestive evidence of linkage to type 1 diabetes, which given our sample sizes and marker map, is the most interesting data yet for this chromosome. So far, no associations with genome-wide significance have been reported in these regions.

Recently, genome-wide association (GWA) studies have identified at least 18 genomic regions with significant evidence of association to type 1 diabetes but the majority of these loci, although common in the population, have relatively modest individual effects on risk (e.g., Odds Ratios (OR) of 1.1 - 1.3) (19-22). Linkage studies are not optimal for detecting such loci where the risk alleles are common and have small individual effect sizes. Thus, it is not surprising that none of these newly identified loci from GWA studies fall within regions of even suggestive linkage in the current study. Several of our linkage peaks do include loci previously detected by association but these tend to have large effects on risk such as HLA (OR=11.4) (23) or INS (OR=2.38) (18).

The large sample size studied here, 2,658 ASPs from 2,496 families, provided us with unprecedented power to detect linkage to type 1 diabetes across the genome. While we detected one locus with significant evidence of linkage, and five more regions with suggestive evidence, our study provides little support for the majority of previously proposed IDDM loci. Our power to detect evidence of linkage to type 1 diabetes comes at a cost; the low frequency of families with multiple affected siblings required a worldwide recruitment effort. Therefore, we cannot exclude the possibility that loci with strong effects in particular populations might be obscured in the current analysis. For example, LOD=3.21 was observed at 66.1 cM on
Type 1 diabetes linkage scan

chromosome 17 in the 370 T1DGC families recruited by the North American network while no linkage was detected in the other 8 populations (LOD scores ranging from -0.62 to 0.30).

In summary, the evidence for independent major gene effects in type 1 diabetes detectable by linkage (e.g., rare variants), that are shared across populations, appears to be limited to a few genes in the HLA complex, INS, CTLA4, IDDM15 and possibly two regions on chromosomes 19. The combination of these putative rare variant effects with common variants of smaller effect size, more readily detected by association-based approaches, accounts for the majority of risk for type 1 diabetes of which HLA class II alleles are by far the major contributor. Identification of candidate genes and variants in these and other regions will help to clarify the biochemical pathways and mechanisms that contribute to risk for this common autoimmune disease. The substantial collection of biospecimens and data assembled by the T1DGC, and available to diabetes researchers (www.t1dgc.org and www.t1dbase.org), should be an important resource for further characterizing those genomic regions identified by either linkage or association approaches as contributing to risk for type 1 diabetes.

ACKNOWLEDGEMENTS
This research utilizes resources provided by the Type 1 Diabetes Genetics Consortium, a collaborative clinical study sponsored by the National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK), National Institute of Allergy and Infectious Diseases (NIAID), National Human Genome Research Institute (NHGRI), National Institute of Child Health and Human Development (NICHD), and Juvenile Diabetes Research Foundation International (JDRF) and supported by U01 DK062418. Further support was provided by a grant from the NIDDK (DK46635) to PC and a joint JDRF and Wellcome Trust grant to the Diabetes and Inflammation Laboratory at Cambridge which also received support from the National Institute for Health Research Cambridge Biomedical Research Centre. Genotyping was performed by the Center for Inherited Disease Research (CIDR). CIDR is fully funded through a federal contract from the National Institutes of Health to the Johns Hopkins University, contract no. N01-HG-65403.

REFERENCES


19. Wellcome Trust Case Control Consortium: Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature* 447:661-678, 2007


### Table 1: Regions with suggestive or significant evidence of linkage to type 1 diabetes.

<table>
<thead>
<tr>
<th>Chromosome</th>
<th>Position (cM)</th>
<th>LOD</th>
<th>p-value</th>
<th>LOD-1 Support Interval</th>
<th>Flanking Markers&lt;sup&gt;a&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>194.25</td>
<td>3.28</td>
<td>5.0×10⁻⁵</td>
<td>191.3-197.8</td>
<td>rs1882395/rs1369842</td>
</tr>
<tr>
<td>6</td>
<td>52.0</td>
<td>213.2</td>
<td>8.0×10⁻²¹⁶</td>
<td>51.0-52.5</td>
<td>rs11908/rs412735</td>
</tr>
<tr>
<td>11</td>
<td>2.5</td>
<td>3.16</td>
<td>7.0×10⁻⁵</td>
<td>0 – 8.5</td>
<td>rs741737/rs1609812</td>
</tr>
<tr>
<td>19</td>
<td>9.5</td>
<td>2.84</td>
<td>0.00015</td>
<td>7.5-26</td>
<td>rs887270/rs1044250</td>
</tr>
<tr>
<td>19</td>
<td>58.0</td>
<td>2.54</td>
<td>0.0003</td>
<td>52-63</td>
<td>rs1019937/rs1878926</td>
</tr>
</tbody>
</table>

<sup>a</sup>Genotyped markers bounding the LOD-1 support interval.
# Table 2: Evidence for linkage in regions previously reported to harbor *IDDM* loci.

<table>
<thead>
<tr>
<th>Locus</th>
<th>Chrom</th>
<th>Defined by marker(s)</th>
<th>Marker tested</th>
<th>LOD</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>IDDM1</td>
<td>6p21.3</td>
<td><em>HLA-DQβ1</em></td>
<td>rs1019937</td>
<td>213.2</td>
<td>$8.0 \times 10^{-216}$</td>
</tr>
<tr>
<td>IDDM2</td>
<td>11p15.5</td>
<td><em>INS-VNTR</em></td>
<td>rs1004446</td>
<td>3.09</td>
<td>0.00008</td>
</tr>
<tr>
<td>IDDM3</td>
<td>15q26</td>
<td><em>D15S107</em></td>
<td>rs906310</td>
<td>0.00</td>
<td>0.6</td>
</tr>
<tr>
<td>IDDM4</td>
<td>11q13</td>
<td><em>D11S1917</em></td>
<td>rs923346</td>
<td>0.87</td>
<td>0.02</td>
</tr>
<tr>
<td>IDDM5</td>
<td>6q25</td>
<td><em>ESR</em></td>
<td>rs11902</td>
<td>0.00</td>
<td>0.5</td>
</tr>
<tr>
<td>IDDM6</td>
<td>18q</td>
<td><em>D18S64</em></td>
<td>rs1943418</td>
<td>0.68</td>
<td>0.04</td>
</tr>
<tr>
<td>IDDM7</td>
<td>2q31</td>
<td><em>D2S152</em></td>
<td>rs696092</td>
<td>1.91</td>
<td>0.002</td>
</tr>
<tr>
<td>IDDM8</td>
<td>6q27</td>
<td><em>D6S264</em></td>
<td>rs6910121</td>
<td>0.45</td>
<td>0.07</td>
</tr>
<tr>
<td>IDDM9</td>
<td>3q21-25</td>
<td><em>D3S1303</em></td>
<td>rs1983380</td>
<td>0.40</td>
<td>0.09</td>
</tr>
<tr>
<td>IDDM10</td>
<td>10cen</td>
<td><em>D10S193</em></td>
<td>rs959629</td>
<td>1.07</td>
<td>0.013</td>
</tr>
<tr>
<td>IDDM11</td>
<td>14q24.3</td>
<td><em>D14S67</em></td>
<td>rs1012921</td>
<td>0.13</td>
<td>0.2</td>
</tr>
<tr>
<td>IDDM12</td>
<td>2q33</td>
<td><em>CTLA4</em></td>
<td>rs1947983</td>
<td>1.05</td>
<td>0.014</td>
</tr>
<tr>
<td>IDDM13</td>
<td>2q34</td>
<td><em>IGFBP2</em></td>
<td>rs1851328</td>
<td>0.63</td>
<td>0.04</td>
</tr>
<tr>
<td>IDDM15</td>
<td>6q</td>
<td><em>D6S283</em></td>
<td>rs2207958</td>
<td>5.99</td>
<td>$7.5 \times 10^{-8}$</td>
</tr>
<tr>
<td>IDDM16</td>
<td>14q32.3</td>
<td><em>D14S293</em></td>
<td>rs872945</td>
<td>0.05</td>
<td>0.3</td>
</tr>
<tr>
<td>IDDM17</td>
<td>10q25</td>
<td><em>D10S592</em></td>
<td>rs151603</td>
<td>0.16</td>
<td>0.2</td>
</tr>
<tr>
<td>IDDM18</td>
<td>5q33-34</td>
<td><em>IL12B</em></td>
<td>rs270664</td>
<td>0.01</td>
<td>0.4</td>
</tr>
<tr>
<td></td>
<td>1q42</td>
<td><em>D1S1617</em></td>
<td>rs342817</td>
<td>0.00</td>
<td>0.5</td>
</tr>
<tr>
<td></td>
<td>16q22-24</td>
<td><em>D16S3098</em></td>
<td>rs1037973</td>
<td>1.24</td>
<td>0.008</td>
</tr>
</tbody>
</table>

*Not corrected for allele sharing at *IDDM1*. Adjustment based upon the expected LOD at *D6S283* and the recombination fraction with HLA yields an adjusted LOD of 2.7.
**Figure 1: Genome-wide evidence for linkage to type 1 diabetes.** Evidence for linkage was evaluated using Merlin and LOD scores were calculated using the Kong and Cox linear NPL model (10;11). LOD scores are plotted on the vertical axis versus chromosomal position on the horizontal axis. LOD scores at region 15.5-110.5cM on chromosome 6 are larger than 4 and not shown. The dotted horizontal line indicates a LOD of 2.2. The dashed line indicates LOD of 3.6.