Saturation multipoint linkage mapping of chromosome 6q in type 1 diabetes

June L. Davies, Francesco Cucca, Juliet V. Goy, Zeinat A. A. Atta, Marilyn E. Merriman, Amanda Wilson, Anthony H. Barnett¹, Stephen C. Bain¹ and John A. Todd

Wellcome Trust Centre for Human Genetics, Nuffield Department of Surgery, University of Oxford, Windmill Road, Headington, Oxford OX3 7BN, UK and ¹Department of Medicine/Diabetes/Endocrinology, University of Birmingham, Birmingham Heartlands Hospital, Bordesley Green East, Birmingham B9 5SS, UK

Received March 13, 1996; Revised and Accepted April 29, 1996

Linkage analysis of type 1 diabetes sib pair families (n = 334) has suggested two separate regions of human chromosome 6q are linked to disease (designated IDDM5 and IDDM8). To test if these are false positive results, all available sib pair families (n = 429) were typed using a 92% informative map of chromosome 6q and multipoint analysis. The two regions still showed positive evidence of linkage, most notably the proterminal region, 6q27, corresponding to IDDM8 (MLS = 2.57, p = 0.0006; λs = 1.17). In addition, some evidence of transmission disequilibrium was seen with marker a046xa9 (IDDM5).

INTRODUCTION

Type 1 or insulin-dependent diabetes mellitus is a complex disorder in which both genetic and environmental factors contribute to the development of the disease. The degree of familial clustering of the disease (λs) can be estimated from the ratio of the risk for siblings of patients (6%) and the population prevalence (0.4%; p0). Using a candidate gene approach, susceptibility genes have been identified on chromosome 6p21 in the major histocompatibility complex (MHC) HLA region, HLA-DOB1 and -DRB1 (IDDM1) (2), and the insulin gene region (INS) (IDDM2) on chromosome 11p15 (3–5); IDDM2 has been identified as the variable number of tandem repeat (VNTR) sequence 596 bp 5’ to the INS ATG initiation codon (6). Assuming a multiplicative model of epistasis, IDDM1 and IDDM2 only account for 32% and 8% of familial clustering, respectively, in the UK population. Analysis using a multiplicative model is consistent with results obtained from the NOD mouse model of type 1 diabetes (7), the rapid decrease in risk from first to second to third degree relatives (8) and pair-wise analyses of IDDM loci (9).

The availability of affected sibpair families (10,11) combined with a dense map of polymorphic microsatellite marker loci (12) and advances in genotyping technology (13–15) has enabled genome-wide screening for regions linked to type 1 diabetes (16,17). Currently, evidence for five IDDM loci, unlinked to IDDM1 and IDDM2, have been published: IDDM3 on chromosome 15q26 (18), IDDM4 on 11q13 (16,17), IDDM5 on 6q25 (16), IDDM7 on 2q31 (19) and IDDM8 on 6q27 (20).

In our initial genome scan 11 chromosome 6q microsatellite markers (average spacing 8 cM) were analysed in 96 UK families yielding evidence of linkage at markers ESR (6q25; MLS = 1.82, p = 0.003; λs = 1.66) and D6S264 (6q27; MLS = 1.21, p = 0.013; λs = 1.28). Linkage to ESR was also obtained in USA families (n = 84; MLS = 1.18, p = 0.017; λs = 1.18) but not in an independent UK data set (n = 102) (16). There was significant heterogeneity at ESR between the three linkage data sets (p = 0.03), a result also seen at IDDM2 which is a true type 1 diabetes disease locus. The chromosome 6q25/ESR locus was designated IDDM5 (16) and warranted further analysis. In an independent study, 55 new families from the USA were analysed by Luo et al. (20) and additional support obtained for the D6S264 6q27-proterminal region (designated IDDM8).

In order to evaluate fully linkage of chromosome 6q to type 1 diabetes, we analysed all available affected sib pair families (n = 429, including 104 new UK families that have not previously been typed) using multipoint linkage analysis and a dense map of microsatellite markers.

RESULTS

Multipoint linkage mapping

The original genome scan results in 93 families (referred to here as data set UK1) for 11 chromosome 6q markers were reanalysed using the multipoint linkage program MAPMAKER/SIBS (21), giving maximal linkage to disease at ESR (MLS = 2.19, p = 0.001; λs = 1.76) (Fig. 1A). However, the average information content of the map based on 11 markers was only 73% (i.e. an average 73% of families were informative across the region), with the proterminal region poorly resolved (Fig. 1C). The 93 families were therefore genotyped with an additional 28 polymorphic microsatellite markers spanning a 90 cM region of chromosome 6q (average spacing 2.5 cM). A multipoint map was obtained with an increased average information content of 92% (Fig. 1D) and this provided greater evidence of linkage (a046xa9 MLS = 2.42, p = 0.0007; λs = 1.86) in the ESR region (map positions 15–45 cM) (Fig. 1B). The proterminal region (map positions 0–15 cM)
80–90 cM) also showed some positive evidence of linkage (D6Q27 MLS = 1.82, p = 0.003; λs = 1.23).

All available UK type 1 diabetes sib pair families (n = 299, including the original 93 and 102 UK family data sets (16) were typed for 39 markers. Figure 2A shows that there is still evidence of linkage in the IDDM5 region (15–45 cM region) but it now does not exceed MLS = 1 (p = 0.02; λs at ESR = 1.12). Evidence in the IDDM5 region (80–90 cM) was also weakened with the maximal evidence of linkage at D6Q27 (MLS = 1.52, p = 0.006; λs = 1.19).

Finally, all available USA families (n = 130, which included the 84 analysed previously only for ESR, D6S290 and D6S305) were typed with the 39 markers, providing support for both IDDM5 (D6S290 MLS = 1.9, p = 0.002; λs = 1.19) and IDDM8 (D6S281 MLS = 1.44, p = 0.008; λs = 1.21) (Fig. 2B). The evidence in support of linkage in all UK and USA families combined (n = 429) exceeded MLS = 1.5 (p = 0.006) in several regions of chromosome 6q (Fig. 2C); in the 15–45 cM IDDM5 region MLS = 1.77 (p = 0.004; λs = 1.1) at GATA2E12, less than 1 cM proximal to ESR, and MLS = 1.95 (p = 0.002; λs = 1.1) at D6S1007, approximately 21 cM distal to ESR, and in the IDDM8 region the peak MLS value was obtained at the most distal marker D6S281 (MLS = 2.57, p = 0.0006; λs = 1.17), which is over 40 cM distal to D6S1007.

Assuming a multiplicative model of epistasis, the IDDM8/D6S281 region contributes 5.3% of familial clustering in UK type 1 diabetes families (multipoint λs = 1.16), 7.2% of familial clustering in USA families (multipoint λs = 1.22) and 5.8% of familial clustering in UK and USA families combined (multipoint λs = 1.17). IDDM5 contributes 4.3% of familial clustering in UK families (multipoint λs = 1.12, based on data for marker ESR), 5.1% of familial clustering in USA families (multipoint λs = 1.15) and 4.3% of familial clustering in UK and USA families combined (multipoint λs = 1.12).

Transmission disequilibrium test

Fifteen markers showing greatest evidence of linkage of chromosome 6q to type 1 diabetes were used for transmission disequilibrium testing in all UK (n = 299) and USA (n = 130) families. Only the three most common alleles were examined for each marker, and these were a046xa9, D6S403, D6S250, GATA2E12, D6S400, D6S290, D6S245, D6S448, D6S473, D6S362, D6S275, D6S366, D6S363, IGF2R, D6S411, D6S305, a342vb5, D6S311, a046xa9, D6S476, GATA2E12, ESR, D6S440, D6S290, D6S441, D6S264, a304xb1, D6S297, a057vf9, D6S503, D6Q27, a304xb1, D6S446, D6S281. (C) and (D) Information content across a 90 cM region of chromosome 6q.
seen 81 times ($p = 0.0006$). Differences in allele transmissions between the UK and USA populations was seen with both $a046xa9$ (alleles 6 and 7) and $D6S281$ (allele 4) and may reflect differences in origin and degree of admixture between these populations.

**Table 1. Transmission of the three most common alleles of markers $a046xa9$ and $D6S281$ to affected offspring in UK and USA families**

<p>| | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$a046xa9$</td>
<td>$D6S281$</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>T</td>
<td>NT</td>
<td>$p$</td>
<td>T</td>
</tr>
<tr>
<td>Allele 6 - 133 mu</td>
<td></td>
<td></td>
<td></td>
<td>Allele 4 - 133 mu</td>
</tr>
<tr>
<td>UK1</td>
<td>88</td>
<td>82</td>
<td>53</td>
<td>51</td>
</tr>
<tr>
<td>ALL UK</td>
<td>236</td>
<td>214</td>
<td>167</td>
<td>127</td>
</tr>
<tr>
<td>USA</td>
<td>95</td>
<td>131</td>
<td>0.02</td>
<td>43</td>
</tr>
<tr>
<td>UK+USA</td>
<td>331</td>
<td>345</td>
<td>210</td>
<td>208</td>
</tr>
<tr>
<td>Allele 7 - 135 mu</td>
<td></td>
<td></td>
<td></td>
<td>Allele 6 - 137 mu</td>
</tr>
<tr>
<td>UK1</td>
<td>37</td>
<td>57</td>
<td>0.04</td>
<td>59</td>
</tr>
<tr>
<td>ALL UK</td>
<td>104</td>
<td>148</td>
<td>0.006</td>
<td>211</td>
</tr>
<tr>
<td>USA</td>
<td>62</td>
<td>68</td>
<td>1.04</td>
<td>140</td>
</tr>
<tr>
<td>UK+USA</td>
<td>166</td>
<td>216</td>
<td>0.01</td>
<td>351</td>
</tr>
<tr>
<td>Allele 8 - 137 mu</td>
<td></td>
<td></td>
<td></td>
<td>Allele 7 - 139 mu</td>
</tr>
<tr>
<td>UK1</td>
<td>32</td>
<td>20</td>
<td>32</td>
<td>32</td>
</tr>
<tr>
<td>ALL UK</td>
<td>82</td>
<td>48</td>
<td>0.003</td>
<td>101</td>
</tr>
<tr>
<td>USA</td>
<td>35</td>
<td>27</td>
<td>55</td>
<td>55</td>
</tr>
<tr>
<td>UK+USA</td>
<td>117</td>
<td>75</td>
<td>0.002</td>
<td>156</td>
</tr>
</tbody>
</table>

**DISCUSSION**

Two potential disease susceptibility regions have been identified on chromosome 6q, $IDDM5$ near the $ESR$ gene and $IDDM8$ at $D6S281$. Taking into consideration all the mapping data, it seems unlikely that the results for $IDDM5$ or $IDDM8$ are false-positives: $IDDM8$ shows positive evidence of linkage in all data sets and has a $\lambda_s$ value of 1.17 for UK and USA type 1 diabetes families combined, contributing 5.8% of familial clustering. The additional information obtained from saturating chromosome 6q with extra markers, analysing more families and the application of multipoint methods [over 60 000 new genotypes produced compared to 107 000 for the complete genome scan (16)] has also enabled us to exclude 90 cM of chromosome 6q for susceptibility loci with $\lambda_s \geq 2$ (data not shown; $\lambda_s$ value for HLA/$IDDM1 = 2.4$). However, no region could be excluded for a disease locus with $\lambda_s = 1.3$ (value equivalent to $IDDM2$ $\lambda_s$). It is evident that a map using all available microsatellite markers on chromosome 6q (average spacing of 2.5 cM and that genotype robustly) is required to extract the vast majority of information from families (>90%), where the collection of families is easily the most difficult and expensive step in the analysis of a complex disease.

Since linkage disequilibrium can generally be detected between markers or marker–disease loci up to 2 cM apart in a variety of populations with different histories and levels of admixture (22,23), any disequilibrium of marker alleles with disease seen in our UK and USA families, which have many different ethnic origins, would extend over a shorter distance and indicate close proximity of that marker to the aetiological mutation. However, of the transmission disequilibrium results, none would be statistically significant after correction for multiple marker loci (15 tested) and alleles (three most common tested), although the results for $a046xa9$ and $D6S281$ merit additional studies. $D6S281$ is one of the most distal loci on chromosome 6q and if no convincing evidence of linkage disequilibrium is found in additional families, then all neighbouring markers, up to a 10 cM region around $D6S281$, will need to be characterised and genotyped on a large number of families.

**MATERIALS AND METHODS**

DNA samples from UK and USA families are available from the repositories of the British Diabetic Association (BDA) and the Human Biological Database Interchange (HBDI), respectively.
To limit heterogeneity, the criteria for affected sib pair family selection was that at least one sibling was under 17 years of age at the onset of diabetes and the other under 29 years. The grandparents of all UK family members were Caucasian and born in the UK. Families with more than two affected sibs were not included. In the UK1 family data set, (originally n = 96), three families were excluded: family 142 subsequently became a triplex family, families 51 and 52 were excluded owing to DNA sample errors. Microsatellite markers were identified from the public databases and included dinucleotide and tetranucleotide repeats, except a040xa9, a342vb5, a085zd5, a290xb9, a057yf9 and a304xb1 provided by Généthon. Thirty-six new markers were initially identified but reproducible, robust genotyping results were obtained for only 28 markers.

Multipoint maximum lod score (MLS) values (24, 25) were calculated using the MAPMAKER/SIBS program (21), which also calculates map informativity and exclusion maps for effects with varying \( \lambda_a \) values and theoretical \( p \) values were assigned to the MLS scores (25). The data was analysed using the MAPMAKER/SIBS program assuming dominance variance and the graphs shown in Figures 1 and 2 were drawn in Microsoft Excel using output scores from MAPMAKER/SIBS. Additional parameters used were increment step 1 for the multipoint linkage maps 1(A), 1(B) and 2(A–C) (giving multipoint MLS values at marker locations), increment distance 0.5 for the information content maps 1(C) and 1(D) (giving multipoint values at 0.5 cm throughout the 90 cm region of chromosome 6q) and \( \lambda_a \) values of 1.1, 1.3, 1.5, 2, 2.4 and 3 for the exclusion maps (not shown).

The transmission disequilibrium test was carried out as described previously (26), and both affected sibs in each family were analysed.

ACKNOWLEDGEMENTS

We would like to thank Jackie Carr-Smith, Carolyn Smyth and Beth Rowe for collection of families, Sabine Fauré at Généthon and Helen Donis-Keller (D6Q27) for providing microsatellite markers; and the British Diabetic Association, the Juvenile Diabetes Foundation International, the Medical Research Council and the Wellcome Trust for financial support. JAT is a Wellcome Trust Principal Research Fellow and JLD is a British Diabetic Association—Warren Robertson. Autoimmunity, 7, 83–5.

REFERENCES