Age and sex based genetic locus heterogeneity in type 1 diabetes

Andrew D Paterson, Arturas Petronis

Abstract

Background—Two genome scans for susceptibility loci for type 1 diabetes using large collections of families have recently been reported. Apart from strong linkage in both studies of the HLA region on chromosome 6p, clear consistent evidence for linkage was not observed at any other loci. One possible explanation for this is a high degree of locus heterogeneity in type 1 diabetes, and we hypothesised that the sex of affected offspring, age of diagnosis, and parental origin of shared alleles may be the bases of heterogeneity at some loci.

Methods—Using data from a genome wide linkage study of 356 affected sib pairs with type 1 diabetes, we performed linkage analyses using parental origin of shared alleles in subgroups based on (1) sex of affected sibs and (2) age of diagnosis.

Results—Among the results obtained, we observed that evidence for linkage to IDDM4 on chromosome 11q13 occurred predominantly from opposite sex, rather than same sex sib pairs. At a locus on chromosome 4q, evidence for linkage was observed in sibs where one was diagnosed above the age of 10 years and the other diagnosed below 10 years of age.

Conclusions—We show that heterogeneity tests based on age of diagnosis, sex of affected subject, and parental origin of shared alleles may be helpful in reducing locus heterogeneity in type 1 diabetes. If repeated in other samples, these findings may assist in the mapping of susceptibility loci for type 1 diabetes. Similar analyses can be recommended in other complex diseases.


Keywords: type 1 diabetes; age of diagnosis; sex; parental origin of alleles

Type 1 diabetes is a complex disease caused by a combination of genetic and environmental factors. Over the past five years, genome wide linkage studies in type 1 diabetes have been performed with the aim of identifying novel susceptibility loci. These genome scans have confirmed the strong previous evidence that the HLA region (IDDM1) is involved in type 1 diabetes and highlighted a number of additional regions showing evidence for linkage. However, the consistency of these additional findings is low, suggesting that they may either be loci of small effect, population specific effects, or the result of type 1 error. Low penetrance and a high degree of genetic locus heterogeneity (when different genes confer susceptibility in different families) cannot be excluded in type 1 diabetes.

Several approaches have been used to reduce locus heterogeneity. Post hoc locus heterogeneity analysis can be useful to separate families that are linked to a specific locus from unlinked families, but this approach can only identify clearly linked (or unlinked) families when each family is sufficiently large. This strategy has also been used for small families; however, it is not possible to determine with any certainty which families are truly linked or unlinked to a particular locus. Another approach to reducing heterogeneity involves analysing subgroups based on pathogenic features, such as the presence of specific antibodies, or other biological measures. The power derived from dividing the sample into subgroups clearly depends on the choice of appropriate bases for splitting the data.

A third approach to reducing locus heterogeneity, searching for epistasis, has already been performed in one of the genome scans for type 1 diabetes as part of the primary analyses. Affected sib pair families were subdivided based on (1) sharing of either two or (2) one or no HLA haplotypes, whether the sibs were (3) DR3/DR4, (4) DR4/not DR4, or (5) DR3/not DR4, or (6) sharing of high or (7) low risk alleles at the insulin gene VNTR. Despite studying seven overlapping subgroups, this approach produced only a small number of loci which showed stronger evidence for linkage in certain subgroups. Furthermore, the undertaking of multiple statistical tests on many subgroups of the data leads to concerns about appropriate correction for the multiple analyses. Such extensive sample divisions have been criticised because of the failure to correct for the multiple statistical comparisons performed. Despite these concerns, heterogeneity analyses are important since they can assist in the identification and fine mapping of loci by providing more significant evidence for linkage, along with narrower confidence intervals regarding the position of the locus. Below we briefly justify the reasons for heterogeneity analyses in type 1 diabetes and provide a number of examples where such analyses appear helpful.

Age of diagnosis. There is a very strong rationale to look for age specific effects in genetic diseases. Attempts to reduce the degree of locus heterogeneity using the criteria of age of diagnosis have been helpful in the fine mapping and identification of a number of human disease genes including BRCA1, which is responsible for a proportion of early onset...
Table 1 Description of subgroups of families

<table>
<thead>
<tr>
<th>Basis of division and subgroups</th>
<th>Abbreviation</th>
<th>No of families</th>
</tr>
</thead>
<tbody>
<tr>
<td>All families (UK356)</td>
<td>All</td>
<td>356</td>
</tr>
<tr>
<td>Age both sibs with onset ≤ 10 years</td>
<td>E</td>
<td>130</td>
</tr>
<tr>
<td>Age both sibs with onset &gt; 10 years</td>
<td>L</td>
<td>95</td>
</tr>
<tr>
<td>One sib onset ≤ 10 years and one &gt;10 years</td>
<td>D</td>
<td>131</td>
</tr>
<tr>
<td>Sex male-male sib pairs</td>
<td>MM</td>
<td>103</td>
</tr>
<tr>
<td>Sex female-female sib pairs</td>
<td>FF</td>
<td>80</td>
</tr>
<tr>
<td>Mixed sex sib pairs</td>
<td>MF</td>
<td>170</td>
</tr>
</tbody>
</table>

The mean (SD) age of diagnosis in the E group is 5.4 (2.9) years; for the L group it is 16.7 (5.9) years.

breast cancer, and PS1 in early onset Alzheimer's disease. Genetic determination of age at diagnosis in type 1 diabetes is suggested by data from over 400 pairs of twins. In the NOD mouse model of type 1 diabetes, Idd4 is linked to early onset of diabetes. Using age as diagnosis as the basis for heterogeneity in the data from Mein et al, we have shown that this approach provides evidence for novel susceptibility loci for type 1 diabetes. Sib pairs both diagnosed >10 years of age show linkage to the region encompassing the Huntington’s disease locus (4p16.3), while linkage of sib pairs with onset both ≤10 years of age show linkage to the locus for Wolfart syndrome (Paterson and Petronis, unpublished observation).

Sex of affected sib pairs. Sex differences in the age specific incidence of type 1 diabetes have been described. Molecular studies have shown differences in the evidence for linkage of affected males and females to regions on the autosomes for hypertension, osteoarthritis, and alcoholism. Moreover, experimental animal models of a number of complex traits have identified loci that show evidence for linkage to affected offspring of only one gender, including hypertension, strokes, alcohol preference, obesity, nociception, insulin, arthritis, and glomerulonephritis. At IDDM1, significantly more males with type 1 diabetes have DR3/not DR4 than females. Using data provided by Dr Todd, we have previously reported that linkage to IDDM9 (chromosome 3q) was observed in female-female affected sib pairs, but not male sib pairs. Although the reasons for such differences are not clear, they may, in part, be because of genetic effects, rather than resulting from differential effects of hormones.

Parental origin of shared alleles. The differential penetrance of genes depending on their parental origin has been observed in a number of Mendelian disorders where it is termed genomic imprinting. Evidence for parental origin effects have been described at IDDM1 in type 1 diabetes, but these findings are controversial. At the insulin gene locus in both type 1 diabetes and polycystic ovarian syndrome, clear parental origin effects have been described. Molecular parental origin effects have been described in a number of other complex diseases, including coeliac disease, psoriasis, IgA deficiency, bipolar affective disorder, and asthma. A possible mechanism for differential parental effects of loci may arise from the observation that a number of genes display differences in susceptibility to mutation, depending on the parental origin of the gene. Similarly, in “genomic disorders”, parent of origin effects have been described. We have previously hypothesised that analysis of sharing of maternal and paternal alleles separately may be useful for complex traits and have provided such evidence for some loci for type 1 diabetes.

Based on the encouraging results of heterogeneity analyses presented above, we propose that further heterogeneity analyses may lead to additional insights into the genetics of type 1 diabetes and use data from a large genome scan for type 1 diabetes to illustrate this. Our rationale was to assess evidence for heterogeneity based on a number of factors which may be related to risk for type 1 diabetes.

Methods

Pedigree and genotyping data for 351 markers from the autosomes of 356 affected sib pair families ascertained in the UK by the British Diabetic Association Warren repository and made available by Dr Todd (http://diesel.cimr.cam.ac.uk/todd/) were used. These include 93 families from the original “UK96” genome scan, as well as 263 families. Age of onset data were provided by Charles Mein. Marker genotypes were scored uniquely in each family using a four allele system, numbering the alleles observed in each family 1-4.

Since we wanted to ensure that none of our results was dependent on possible misspecification of marker allele frequencies, we decided to base our analysis only on allele sharing that could be inferred to be identical by descent (IBD). Elsewhere, it has been pointed out that IBD analysis may tend to have reduced power compared to analysis which uses allele frequencies, because of the higher probability of being able to determine that 0 alleles are shared compared to 1 allele shared. However, in this paper we prefer the IBD sharing approach since it is robust. We used sib_ibd from ASPEX v 1.17 (Hinds and Risch, 1996) to perform this. This analysis provides separate counts of shared maternal and paternal alleles, as well as the combined maternal and paternal sharing. The number of alleles in the combined analysis in some cases is greater than the sum in each subgroup, because some alleles can be shared, but their parental origin is not known. We performed genome wide linkage analysis in each of the subgroups; more details about the characteristics of each subgroup are provided in table 1. The %IBD shared between groups at particular markers was compared using a 2 × 2 contingency table.

Because we generated six partially overlapping subgroups of the total data, correction for multiple tests cannot be performed using straightforward methods (for example, Bonferroni). In addition, since little is known about the molecular genetics of type 1 diabetes, our preliminary findings can be considered to be exploratory in nature, and can be used to generate specific hypotheses that can be tested in independent samples. Studies of complex traits
run the gauntlet of trying to find a balance between using a too stringent significance criterion and missing interesting results, or using a too liberal criterion and reporting many results which may turn out to be type 1 errors.\textsuperscript{50,51} In an attempt to control the type 1 error in our study, we decided to present only results where markers produced $\chi^2 > 10$ (p<0.0016) in any subgroup. Such a criterion is in agreement with “putative linkage” (p<0.0010).\textsuperscript{51}

### Results

No marked differences between the subgroups as regards allele sharing at microsatellite markers around the HLA locus on chromosome 6p21 were observed. However, it must be noted that specific HLA haplotypes and genotypes were not available for this study. Markers D16S3040 and D19S226 produced $\chi^2 > 10$ for all families, but no subgroup produced $\chi^2 > 10$ and these also will not be discussed further. The TH locus, close to IDDM2, also produced excess sharing of alleles in the whole sample (IBD=57%, $\chi^2=11.3$), and elsewhere we have previously shown that this derives predominantly from excess sharing of maternal alleles (IBD=61%, $\chi^2=12.5$) and not paternal (IBD=54%, $\chi^2=1.60$); however, no marked differences in the linkage of subgroups were noted (data not shown). The results from all other markers reaching our inclusion criterion are presented in tables 2 and 3.

Table 2 presents positive results from markers in the centromeric region of chromosome 10, encompassing IDDM10, and the results are not straightforward. At D10S588 there appears to be linkage in FF pairs; they show excess sharing of maternal alleles (IBD=76%, $\chi^2=8.8$), but this does not reach our critical level. In addition, mixed sex pairs show evidence for linkage throughout this region, with generally similar sharing of paternal and maternal alleles. Also, the linkage to markers in this region appears predominantly in E pairs rather than L or D pairs.

At D3S1279 (IDDM9), it appears that the evidence for linkage derives predominantly from families with affected female sibs and that the allele sharing in these families is predominantly of maternal origin (table 3). At D4S512 (4p16.3) evidence for linkage appears in the late onset group, and the sharing at this marker is predominantly of maternal alleles (table 3). Finally, in the D group, there was marked excess non-sharing of alleles at D4S430 (IBD=34%, $\chi^2=18.1$, table 3). Both the E and L onset groups show no deviation from the null hypothesis (IBD=48% and 42%, respectively). Thus, at D4S430, the majority of the excess non-sharing at D4S430 derived from those pairs where one was diagnosed before and after the age of 10 years. At the FGF3 locus on chromosome 11q13 (IDDM4), excess sharing was observed predominantly arising from mixed sex pairs, while there was no evidence for linkage from same sex pairs (table 3).
## Locus heterogeneity in type 1 diabetes

Here we present a detailed heterogeneity analysis of a large genome scan for type 1 diabetes susceptibility loci, using sex of affected sibs and age of diagnosis as bases for our heterogeneity analyses. We also provide the degree of sharing of maternal and paternal alleles separately. Using a similar significance criteria to those suggested elsewhere, we found nine chromosomal regions which meet our significance criteria. Confirmation of the results from the exploratory results presented here are clearly required.

Although we observed no significant differences in allele sharing at markers close to HLA, this does not contradict previous findings since HLA genotypes were not available in these data. Age related differences in HLA haplotypes have been observed in type 1 diabetes, while parental origin effects at HLA have not been observed consistently. Definitive interpretation of differences in linkage of maternal and paternal meioses in pairwise linkage analysis requires a high density of markers to distinguish effects of sex specific genetic maps from a locus showing genomic imprinting. The differences in parental allele sharing at D3S1279 and D4S412 may be solely because of differences in the sex specific genetic distances between the marker and locus. For example, the markers flanking D3S1279 are greater than 10 cM either side of D3S1279 in both the male and female genetic maps. The sharing of maternal alleles in late onset sib pairs at D4S412 (chromosome 4p16.3) maps very close to the Huntington’s disease locus, for which clinical and molecular parental origin effects have been described. According to the Genetic Location Database, D4S412 maps 2 cM in both the male and female genetic maps from the Huntington’s disease gene (HD).

The excess non-sharing of alleles at D4S430 on chromosome 4q in sib pairs with discordant ages of onset is interesting since a previous analysis of part of the data studied here (UK96) reported that the extent of non-sharing at D4S430 (32%) was nearly as deviant from the null hypothesis as the excess sharing at HLA (73%). However, attempts to replicate the excess non-sharing in two separate family sets (UK102 and US84) did not produce excess non-sharing, and the finding in the UK96 was explained as probably the result of a type 1 error. Here we observe that the excess non-sharing derives predominantly from pairs with one sib diagnosed at age ≤10 years and one sib diagnosed older than 10 years of age (“discordant” for age of diagnosis). Excess non-sharing of alleles is expected when discordant sib pairs (one affected, one unaffected) are studied, and an example of this is the ADH3 locus in alcoholism. Deviation from the null hypothesis was greater in pairs discordant for alcoholism (IBD=46%) compared to concordant affected or concordant unaffected sib pairs (IBD=50% and 51%, respectively). The excess non-sharing in discordant sib pairs near the ADH3 locus has also been observed in another study of alcoholism. In relation to the

### Table 3: Pairwise analysis at D3S1279, D4S412, D4S430, and FGF3

<table>
<thead>
<tr>
<th>Marker</th>
<th>Subgroup</th>
<th>D3S1279</th>
<th>D4S412</th>
<th>D4S430</th>
<th>FGF3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>ALL</td>
<td>ALL</td>
<td>ALL</td>
<td>ALL</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mat v pat</td>
<td>Mat v pat</td>
<td>Mat v pat</td>
<td>Mat v pat</td>
</tr>
<tr>
<td></td>
<td></td>
<td>p</td>
<td>p</td>
<td>p</td>
<td>p</td>
</tr>
<tr>
<td></td>
<td></td>
<td>IBD1</td>
<td>IBD0</td>
<td>%IBD</td>
<td>IBD1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2 p</td>
<td>2 p</td>
<td>2 p</td>
<td>2 p</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.12</td>
<td>0.11</td>
<td>0.13</td>
<td>0.11</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.24</td>
<td>0.23</td>
<td>0.24</td>
<td>0.23</td>
</tr>
</tbody>
</table>

Note: Significance criteria to those suggested elsewhere, we found nine chromosomal regions which meet our significance criteria.
type 1 diabetes findings at D4S430, it is of note that there is evidence for linkage of type 2 diabetes in Mexican Americans to markers which map to within 10 cM of D4S430 in the sex averaged genetic map. In this study, data from unaffected sibs were not studied, but genotype data from unaffected subjects may assist in mapping the locus near D4S430.

The predominant findings at markers on chromosome 10 (IDDM10) are that there is greater sharing in opposite sex than same sex pairs. However, this situation does not occur throughout the IDDM10 region; at D10S183 (table 2) MM and FF pairs both share 59% IBD, which is not significantly different from MF pairs which share 62%. According to the primary analysis, the second strongest linkage (after HLA) was to a region on chromosome 10, at IDDM10, with an MLS of 4.7. However, the shape of the multipoint linkage plot appears rather skewed (http://www.well.ox.ac.uk/~chaz/PICTURES/c10tot.GIF), although information content is relatively constant across the whole region (http://krusty.cimr.cam.ac.uk/~chaz/PICTURES/c10li.GIF). The shape of the multipoint plot differs from that which one would expect for a true peak reflecting a single underlying susceptibility locus. Possible explanations for this include the presence of more than one susceptibility locus in this region, and there are precedents for this from the NOD mouse model of type 1 diabetes. A mouse model of systemic lupus erythematosus, as well as type 1 diabetes in humans near IDDM1.

Our earlier analysis of the whole data set on chromosome 10 using parental sex specific multipoint analysis provided some evidence that there may be two distinct loci on chromosome 10 for type 1 diabetes, one at D10S191 and a second at D10S183,

which may be the result of type 2 error; however, it is difficult to speculate about mechanisms which may produce excess sharing in opposite rather than same sex pairs.

Over the last 10 years, a number of developments have made affected sib pair (ASP) studies the mainstay of linkage studies in complex diseases. For the main part, the reasons behind such a shift are compelling; larger numbers of families are available, results depend less on diagnosis of critical people, a lower probability of bilineal transmission at susceptibility loci, and the fact that genetic parameters are not required. Despite this, families with just affected sib pairs may represent a diluted sample of “genetic” families compared to larger multiplex families, with some affected sib pairs possibly resulting from environmental factors.

Although the results presented here are interesting, the data in this study are based on a relatively heterogeneous population. Two further approaches may additionally assist in reducing locus heterogeneity. First, the use of categorical “affected” or “unaffected” phenotypes may not be particularly powerful for complex diseases and, instead, biological markers of subgroups of disease may be more useful for genetic studies, since it is likely that the genes underlying such traits are fewer and the effects of at least some loci are larger. Secondly, the use of genetic isolates has not been exploited to a great extent in studies of type 1 diabetes. One exception has been the study of a large Bedouin Arab family that features a high degree of inbreeding. Such families may provide strong evidence for linkage to only a single locus, but the critical haplotype is usually large. Based on our results, we concluded that heterogeneity analysis may be useful to dissect the genetic architecture of complex human diseases.
Locus heterogeneity in type 1 diabetes


Age and sex based genetic locus heterogeneity in type 1 diabetes

Andrew D Paterson and Arturas Petronis

doi: 10.1136/jmg.37.3.186

Updated information and services can be found at:
http://jmg.bmj.com/content/37/3/186

These include:

References
This article cites 64 articles, 14 of which you can access for free at:
http://jmg.bmj.com/content/37/3/186#BIBL

Email alerting service
Receive free email alerts when new articles cite this article. Sign up in the box at the top right corner of the online article.

Topic Collections
Articles on similar topics can be found in the following collections

- Diabetes (105)
- Immunology (including allergy) (603)
- Metabolic disorders (329)
- Molecular genetics (1254)

Notes

To request permissions go to:
http://group.bmj.com/group/rights-licensing/permissions

To order reprints go to:
http://journals.bmj.com/cgi/reprintform

To subscribe to BMJ go to:
http://group.bmj.com/subscribe/