Genetic analyses of chromosome 12 loci in Crohn’s disease

S Lesage, H Zouali, J-F Colombel and EPIMAD, J Belaiche and GETAID, J-P Cézard and EPWGIBD, C Tysk, S Almer, M Gassull, V Binder, M Chamaillard, J Le Gall, G Thomas, J-P Hugot

Abstract

Background and aims—Inflammatory bowel disease (IBD) includes ulcerative colitis and Crohn’s disease, both of which are multifactorial diseases involving the interaction of genetic and environmental factors. A region on chromosome 12 centred around the marker locus D12S83 has previously been associated with IBD predisposition. The aim of the study was to investigate this genetic region in an independent panel of European families affected by Crohn’s disease.

Methods—A sample of 95 families with two or more affected relatives and 75 simplex nuclear families were genotyped for 19 microsatellite loci located on chromosome 12. A search for linkage and linkage disequilibrium was performed using non-parametric two point and multipoint analyses with the Analyze and Genehunter packages.

Results—No evidence of linkage or linkage disequilibrium was observed for any of the marker loci, including D12S83 (p=0.35 for the two point linkage test). Multipoint linkage analysis also failed to reveal positive linkage on chromosome 12. Power calculations allowed us to reject the hypothesis that the genetic region of chromosome 12 centred on D12S83 contains a susceptibility locus with a relative risk (R) equal to or greater than 2.0 in these families.

Conclusion—Failure to detect linkage or linkage disequilibrium in these families suggests that the chromosome 12 locus previously reported to be associated with genetic predisposition to IBD does not play a role in all European family samples. This observation is compatible with heterogeneity in the genetic basis of susceptibility to the disease and/or exposure to various environmental factors among Caucasian families.

Keywords: chromosome 12; inflammatory bowel disease; Crohn’s disease; linkage analyses; replication study

Inflammatory bowel disease (IBD) includes ulcerative colitis (UC) and Crohn’s disease (CD), both of which are multifactorial disorders whose aetiology appears to result from the interaction between environmental and genetic factors.1 However, the contributions of genetic background and environmental exposure in the development of the disease are not completely understood. Identification of the relevant environmental factors and susceptibility genes would be an important contribution to the understanding of the pathogenic mechanism involved in IBD.

To date, two susceptibility loci have been mapped on the genome with a non-ambiguous statistical threshold of significance. The first susceptibility locus, named IBD1, was localised in the pericentromeric region of chromosome 16.16 This finding has now been confirmed by five independent groups.17 18 19 20 21 IBD1 was reported to be involved in the susceptibility to CD but its contribution to UC is still being debated.22 23 24

A second locus has been mapped on the long arm of chromosome 12.25 26 27 This observation has now been reproduced in at least three independent published studies.28 29 30 This locus was reported to be involved in the genetic predisposition to both CD and UC. More recently, positive linkage disequilibrium has been reported between this locus and the microsatellite marker D12S83, thus refining the localisation of the predisposing gene.31 32

The aim of this work was to replicate the linkage and linkage disequilibrium on chromosome 12 using an independent set of 170 Caucasian CD families.

Materials and methods

Patients and families

One hundred and seventy families affected by CD were recruited through a large European consortium. However, most originated from France (n=148) and Belgium (n=17). Each contributor obtained approval of the relevant ethics committee to participate in the study and informed consent was obtained from each family member.

Diagnostic criteria for inclusion in the study have been published previously and take into account clinical, endoscopic, radiological, and pathological data.33 To limit the expected genetic heterogeneity, families containing individuals affected by UC were excluded.

Ninety five families had two or more affected members (multiplex families) and were used for linkage and linkage disequilibrium studies. In total, they included 329 healthy members

Abbreviations used in this paper: CD, Crohn’s disease; cM, centiMorgan; HRR, haplotype relative risk; IBD, inflammatory bowel disease; TDT, transmission disequilibrium test; UC, ulcerative colitis; PCR, polymerase chain reaction; NPL, non-parametric linkage.
and 231 affected relatives accounting for 128
affected sib pairs and 29 affected pairs of
second and third degree relatives. In 70% of
families, both parents of affected patients were
available; 18% had only one genotyped parent.
For most of the families with a missing
parent(s), the parental genotypes could be
inferred from the genotypes of the healthy sibs.
Thus the number of alleles identical by descent
for relative pairs could be inferred in almost all
cases.

Seventy five additional families with a single
affected member (simplex nuclear families) were
also used for association studies. The two
parents were available in all cases, making
intrafamilial association studies possible.

PROCEDURES FOR GENOTYPING

Genomic DNA was prepared from peripheral
blood leucocytes using standard methods.
Nineteen polymorphic microsatellite loci were
selected from the Applied Biosystems (ABI,
Courtaboeuf, France) fluorescent labelled
human linkage mapping set version 1.0 (panels
17, 18, 19) or from the Genethon database14
and amplified by polymerase chain reaction
(PCR) (30 cycles) on an ABI 877 catalyst
(PCR) (30 cycles) on an ABI 877 catalyst
robotic workstation in 5 μl reactions containing
20 ng of genomic DNA, 3 mM MgCl2, 0.25
mM dNTPs (Pharmacia, Orsay, France), 1.5
pmol of each primer, and 0.4 units of
Amplitaq-Gold DNA polymerase (Perkin-
Elmer, Courtaboeuf, France) in 1×PCR Buffer
II (Perkin Elmer).

Pooled amplification products were then
subjected to electrophoresis in 5% polyacryla-
mid gels for two hours at 3000 V using an ABI
877 DNA Sequencer. Control DNA (indivi-
dual 1347-02 from the CEPH reference
families) was systematically used as an internal
standard. Semi-automated fragment sizing was
performed using Genescan 2.1 software (ABI)
followed by allele identification using Genoty-
per 2.0 software (ABI). Each genotype was
determined independently by two investigators
to confirm the accuracy of genotype. Conflic-
ting data were either resolved or discarded.

DATA ANALYSIS

Mendelian segregation of alleles was checked
using the Unknown program version 5.23.
Marker ordering and inter-marker distances
were computed using the CRI-MAP pro-
gram.13 Any apparent double recombinants
were identified and the genotypes rechecked.
Non-parametric multipoint score (NPL)
values, information content map, and exclu-
sion maps were produced by the Sibs compo-
nent of the Genehunter package version 2.016
using the map generated with the families
studied. The NPL statistics measure identity
by descent allele sharing between affected rela-
tives. Information content mapping measures
the fraction of the total inheritance information
extracted from all markers simultaneously.
Exclusion maps were constructed assuming no
dominance variance and using a locus specific
relative risks to siblings (λs) value of 2.0, as
previously reported for the chromosome 12
susceptibility locus. A locus was considered to
be excluded for linkage if the logarithm of the
odds score (LOD score) was less than −2.0 for
this specified value of λs.

Non-parametric two point analyses were
performed using the Analyze package version
3.0.18 It includes a two point affected sib pairs
analysis that tests for an excess of allele sharing
between affected siblings. A search for associ-
ation between the marker alleles and CD was
performed using the transmission/ disequilib-
ium test (TDT) and the haplotype relative risk
analysis (HRR) based on likelihood
ratio statistics. Intrafamilial association
tests compare the distribution of alleles trans-
mitt and not transmitted to the affected
patients from their parents. In the TDT, only
one allele is transmitted per patient and data
from heterozygous parents are included in
the analysis. In the HRR statistic, homozygous
parents are also taken into account.

Results

Members of the 170 CD families were
genotyped for 19 highly polymorphic markers,
evenly spread on chromosome 12 and includ-
ing the D12S83 locus that exhibited positive
linkage and linkage disequilibrium in previous
studies.10 12 The genetic map generated using
these genotypes spanned 157 centiMorgan
(cM), a genetic length slightly shorter than the
previously estimated length of chromosome 12
(174 cM, see Dib and colleagues14). The posi-
tions of the markers and the inter-marker
genetic distances were in agreement with the
published maps. The markers were 8.7 cM
apart on average (range 5.0–16.6 cM). As the
data from our families were more extensive
than previously published, we used the map
distances estimated from our data in subse-
quent statistical analyses (fig 1).

Non-parametric two point linkage analyses
performed on the 95 multiplex families did not
provide evidence of linkage for any of the poly-
morphic markers tested (table 1). The ob-
erved allele sharing between affected siblings

Table 1 Two point linkage and linkage disequilibrium analyses

<table>
<thead>
<tr>
<th>Markers</th>
<th>Mean proportion of alleles shared</th>
<th>p Value</th>
<th>TDT test p value</th>
<th>HRR test p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>D12S352</td>
<td>0.55</td>
<td>0.13</td>
<td>0.50</td>
<td>0.50</td>
</tr>
<tr>
<td>D12S1685</td>
<td>0.50</td>
<td>0.02</td>
<td>0.50</td>
<td>0.50</td>
</tr>
<tr>
<td>D12S336</td>
<td>0.50</td>
<td>0.50</td>
<td>0.50</td>
<td>0.50</td>
</tr>
<tr>
<td>D12S77</td>
<td>0.50</td>
<td>0.35</td>
<td>0.35</td>
<td>0.35</td>
</tr>
<tr>
<td>D12S346</td>
<td>0.50</td>
<td>0.35</td>
<td>0.31</td>
<td>0.31</td>
</tr>
<tr>
<td>D12S1596</td>
<td>0.50</td>
<td>0.50</td>
<td>0.17</td>
<td>0.17</td>
</tr>
<tr>
<td>D12S345</td>
<td>0.50</td>
<td>0.50</td>
<td>0.50</td>
<td>0.50</td>
</tr>
<tr>
<td>D12S85</td>
<td>0.50</td>
<td>0.50</td>
<td>0.36</td>
<td>0.36</td>
</tr>
<tr>
<td>D12S368</td>
<td>0.51</td>
<td>0.44</td>
<td>0.50</td>
<td>0.50</td>
</tr>
<tr>
<td>D12S83</td>
<td>0.51</td>
<td>0.35</td>
<td>0.58</td>
<td>0.58</td>
</tr>
<tr>
<td>D12S1676</td>
<td>0.51</td>
<td>0.35</td>
<td>0.26</td>
<td>0.26</td>
</tr>
<tr>
<td>D12S292</td>
<td>0.50</td>
<td>0.47</td>
<td>0.38</td>
<td>0.38</td>
</tr>
<tr>
<td>D12S326</td>
<td>0.50</td>
<td>0.50</td>
<td>0.50</td>
<td>0.50</td>
</tr>
<tr>
<td>D12S351</td>
<td>0.50</td>
<td>0.50</td>
<td>0.17</td>
<td>0.17</td>
</tr>
<tr>
<td>D12S346</td>
<td>0.50</td>
<td>0.50</td>
<td>0.14</td>
<td>0.14</td>
</tr>
<tr>
<td>D12S78</td>
<td>0.50</td>
<td>0.50</td>
<td>0.28</td>
<td>0.28</td>
</tr>
<tr>
<td>D12S292</td>
<td>0.52</td>
<td>0.26</td>
<td>0.04</td>
<td>0.04</td>
</tr>
<tr>
<td>D12S86</td>
<td>0.54</td>
<td>0.15</td>
<td>0.39</td>
<td>0.39</td>
</tr>
<tr>
<td>D12S324</td>
<td>0.50</td>
<td>0.50</td>
<td>0.50</td>
<td>0.50</td>
</tr>
</tbody>
</table>

Results were obtained using 95 multiplex families with Crohn’s disease (CD) for the sib pair
analysis and 170 CD multiplex and simplex families for the transmission disequilibrium (TDT)
and haplotype relative risk (HRR) tests. Analyses were performed with the Analyze package ver-
sion 3.0. For the 19 markers of chromosome 12, p values of the linkage (sib pair analysis) and
association (TDT and HRR analyses) tests are indicated. The marker D12S83 which previously
demonstrated a positive linkage and association10,11 is in bold.

www.gutjnl.com

Lesage, Zouali, Colombel, et al
Genetic analyses of chromosome 12 loci in Crohn’s disease

was maximal at D12S1685 on the short arm of chromosome 12 (observed value 0.59 compared with an expected value of 0.50). However, this locus was 70 cm from D12S83 and was thus far from the genetic region of interest. In the genetic region of interest, centred around the marker D12S83, the maximum value of identity by descent between affected siblings was no more than 0.51 (expected value 0.50; p=0.35) providing no evidence of linkage.

Non-parametric multipoint linkage analyses are shown in fig 2. The information statistic measures the proportion of the possible inheritance information extracted at each location on the map using the panel of polymorphic markers in the families studied (fig 2A). The information is about 80% on average, indicating that the markers used are highly informative. The multipoint linkage statistic did not indicate linkage at any point on the chromosome (maximum value at D12S86, NPL=1.52, p=0.07) (fig 2B). In the genetic region surrounding the D12S83 locus, the NPL values were negative indicating that sharing of alleles identical by descent was less than that expected in this region. Because negative results may arise by chance when statistical power is low, the power to detect linkage using these data was analysed. For a locus specific increased relative risk of λ=2, a value previously attributed to the chromosome 12 susceptibility locus, the entire genetic region of interest was excluded (fig 2C).

Finally, because linkage disequilibrium was reported between the IBD2 locus and D12S83 by three independent groups, a search for associations was also performed (table 1). Because multiplex and simplex families are informative in association studies, 75 additional simplex families were included in this analysis. Using all of the 170 CD families, it was not possible to demonstrate an association, using either the TDT or HRR test, for any typed chromosome 12 loci.

As genetic heterogeneity has already been demonstrated in this disease, we also classified our families into different phenotypic subgroups and repeated the statistical analyses on each group. Neither linkage nor linkage disequilibrium was observed in any of the phenotypic subgroups, including: age at onset below or above 20 years; small intestine and/or colon involvement; or low or high severity during the disease course (data not shown).

Discussion

The study was performed to determine the contribution of a previously identified IBD susceptibility locus on chromosome 12 in an independent set of Caucasian families.

The initial study reporting linkage between IBD and a susceptibility locus on chromosome 12 used a total of 160 IBD families. In that study linkage was observed for IBD (p=0.0000003) as well as for UC (p=0.0025) and CD alone (p=0.003). Even if these results do not argue for a genetic heterogeneity between the two disorders, given the phenotypic differences between CD and UC, genetic heterogeneity in susceptibility cannot be ruled out. For this reason, we decided to limit our investigation to the CD phenotype. In the original study, linkage for CD was initially found in a group of 67 CD families containing 81 affected sib pairs. Because the number of families required to replicate a positive linkage may be significantly greater than the number initially used to find linkage, we studied a nearly twofold larger sample of families (95

Table 2 Results and power of the previously published linkage analyses on chromosome 12 for Crohn’s disease (CD) families

<table>
<thead>
<tr>
<th>Author</th>
<th>No of multiplex CD families</th>
<th>No of affected relative pairs</th>
<th>Results (p value of the test)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Satsangi et al</td>
<td>67</td>
<td>81</td>
<td>3×10^{-7}</td>
</tr>
<tr>
<td>Duerr et al</td>
<td>80</td>
<td>127</td>
<td>0.0002</td>
</tr>
<tr>
<td>Yang et al</td>
<td>48</td>
<td>75</td>
<td>0.0004</td>
</tr>
<tr>
<td>Curran et al</td>
<td>134</td>
<td>167</td>
<td>0.005</td>
</tr>
<tr>
<td>Cho et al</td>
<td>99</td>
<td>175</td>
<td>0.40</td>
</tr>
<tr>
<td>Brant et al</td>
<td>77</td>
<td>148</td>
<td>0.30</td>
</tr>
<tr>
<td>Roux et al</td>
<td>114</td>
<td>0.24</td>
<td></td>
</tr>
<tr>
<td>Present study</td>
<td>95</td>
<td>157</td>
<td>0.35</td>
</tr>
</tbody>
</table>

Figure 2 Multipoint linkage analyses of chromosome 12 in 95 multiplex families with Crohn’s disease (CD) using the Genehunter package. (A) Information content; (B) multipoint non-parametric linkage (NPL) statistics. The region implicated by Satsangi et al containing the inflammatory bowel disease (IBD) susceptibility locus is indicated by a shaded box. (C) Exclusion map calculated for a hypothetical locus with a relative risk λ=2. The exclusion threshold of LOD score (−2) is indicated.
families corresponding to 157 affected relative pairs) than that used in the initial study. Our sample was comparable in size to those in other linkage replication studies (table 2). Whereas the initial investigation was performed on affected sibling pairs only, the present study also examined second and third degree relative pairs, which is more informative. Thus these arguments suggest that the panel of CD families studied here was sufficient to detect linkage if it existed in this sample.

The density of markers used in our study was sufficient to extract, on average, 80% of the information gap in the region of interest. Thus it is unlikely that addition of new markers would reveal linkage on this region using the same sample. The linkage tests reported here are model free—that is, they do not require any assumption about the genetic model underlying the predisposition to CD. Thus the negative results we obtained are not likely to be due to incorrect parameterisation in the statistical calculation. Also, as identity by descent between affected relatives could be inferred unequivocally for most of the relative pairs, it is unlikely that allele frequency errors significantly affected the results.

We were unable to demonstrate linkage between CD and any locus mapped on chromosome 12. The linkage score was negative on most of the chromosome and reached its highest positive value of 1.52 (p=0.07) for a marker located on the short arm of chromosome 12, 70 cM away from the linkage peak reported by Satsangi and colleagues. Considering that it does not correspond to a region with a previously reported linkage, the value of the statistical test (p=0.02) did not reach the threshold of significant linkage accepted to demonstrate linkage in a genome wide screen (p<0.00002, see Kruglyak and Lander19).

In contrast, no linkage was found to chromosome 12. The linkage score was negative on most of the chromosome and reached its highest positive value of 1.52 (p=0.07) for a marker located on the short arm of chromosome 12, 70 cM away from the linkage peak reported by Satsangi and colleagues.9 Considering that it does not correspond to a region with a previously reported linkage, the value of the statistical test (p=0.02) did not reach the threshold of significant linkage accepted to demonstrate linkage in a genome wide screen (p<0.00002, see Kruglyak and Lander19).

In contrast, no linkage was found to the genetic region previously implicated in susceptibility to IBD. To test if this negative result could be attributed to lack of power, an exclusion map was generated. The exclusion map allowed us to conclude that, in our families, no locus with \( \lambda \geq 2 \) could be located within the genetic region proposed by Satsangi and colleagues.9 However, a gene with a smaller effect cannot be ruled out.

An association between IBD and the microsatellite marker D12S83 has also been reported.12 This observation prompted us to replicate this result. Familial association studies have the advantage over case control studies with unrelated controls of avoiding biases due to population stratification. They are thus considered as more reliable for association studies. However, using 170 CD families, we failed to reproduce the reported association in our family panel.

Our findings indicate that in our family panel the IBD locus mapped on chromosome 12 had at most a small effect. Several other studies have also failed to find linkage on the long arm of chromosome 12 (table 2). Only three of seven replication studies detected the IBD locus on chromosome 12. In contrast, four of five replication studies detected the IBD1 locus on chromosome 16.

Several explanations may be proposed to explain the discrepancies between studies. Differences in diagnostic criteria may play a role. In most studies, including ours, the diagnostic criteria used were widely accepted and have been used by numerous investigators. Furthermore, several authors used the same diagnostic criteria and obtained discordant results.7 10 It has been suggested that the chromosome 12 locus might be more important for UC than CD.19 However, in most studies reporting a positive linkage, the contribution of the CD families was at least as important as that of the UC families.7 9 10 Furthermore, when different phenotypes were analysed separately in the present study, we were unable to detect linkage or linkage disequilibrium. It is thus unlikely that the present negative results are related to major clinical differences.

Differences between statistical analyses could also be considered. However, the statistics used in the above mentioned studies were based on model free methods and were often identical. Thus statistical methods are unlikely to explain the observed discrepancies.

Ethnic heterogeneity may also result in discordant results. All of the published studies were performed in Caucasian families of European origin. They were often the result of a multicentre effort of recruitment in the USA,10–19 Canada21 or Europe (Curran et al,7 Satsangi et al9, and the present study). However, ethnic differences cannot be ruled out. For example, it should be noted that the chromosome 12 locus was detected in two cohorts with a large proportion of families recruited in the UK and was less often identified in Ashkenazi families.10 It has also been suggested that the IBD susceptibility loci may be different in Ashkenazi Jews.3 But discordant results were also obtained in populations with comparable proportions of Ashkenazi Jews in North America.10–19–21

Linkage analyses on the chromosome 16 locus (IBD1) are often available in the family samples used for linkage studies on chromosome 12.4 7 20–22 With the exception of the Canadian set of families, all samples also provided evidence of linkage on chromosome 16 (including the present sample, data not shown). Thus linkage on chromosome 12 does not seem to be conditional on linkage on chromosome 16.

Finally, the discrepancy could be explained by the effect of differences in exposure to unknown environmental factors. Indeed, because IBD are multifactorial disorders, a susceptibility gene interacting with a specific environmental risk factor may be more easily discovered in an exposed population. In contrast, in a non-exposed population, the same gene may remain undetected because its contribution to the disease is small. Further investigations, stratifying the genetic analyses for environmental risk factors, are required to highlight this hypothesis.
This project received financial support from: the European Community (Contract Biomed 2 No BMH4-97-2098), the Ministère Français de l’Éducation Nationale et de la Recherche, INSERM, the Direction Générale de la santé (convention No 4T004C), the Association François Aupetit, and the Institut de Recherche des Maladies de l’Appareil Digestif. We are grateful to Drs Bruno Bonaz, Yoram Bouhnik, Jean-Pierre Cézard, Sal- ...

8th United European Gastroenterology Week
The UEGW abstract book (Gut 2000;47(suppl III)) has again been produced as a CD-ROM and can be found attached to the inside back cover of this issue.