J Med Genet 1998;35:218–221

Evidence of linkage of the inflammatory bowel disease susceptibility locus on chromosome 16 (IBD1) to ulcerative colitis

Muddassar M Mirza, John Lee, Dawn Teare, Jean-Pierre Hugot, Pierre Laurent-Puig, Jean-Frederic Colombel, Shirley V Hodgson, Gilles Thomas, Douglas F Easton, John E Lennard-Jones, Christopher G Mathew

Abstract

Crohn's disease (CD) and ulcerative colitis (UC) are inflammatory bowel diseases (IBD) of unknown aetiology which are characterised by chronic inflammation of the gastrointestinal tract. Epidemiological studies suggest the presence of a genetic component in the aetiology of both CD and UC. A susceptibility gene for Crohn's disease has recently been mapped to the pericentromeric region of chromosome 16 (IBD1), and this finding has been replicated in two subsequent studies. Although CD and UC are distinct clinical entities, the fact that both disorders occur in a significant proportion of families with multiple cases of IBD suggests that overlapping sets of susceptibility genes may be involved. We have addressed this question for IBD1 by typing eight microsatellite markers from the locus in 70 kindreds affected with either UC only or with both UC and CD and analysing the data for linkage by both non-parametric and parametric methods. Evidence for linkage was detected in families affected with only UC, with a mean proportion of 0.70 affected sib pairs sharing alleles identical by descent at D16S3136 (p=0.01), and a peak non-parametric linkage score of 2.02 at D16S3120 with the GENEHUNTER program (p=0.02). The estimated sib relative risk attributable to IBD1 in these families was 1.46. Surprisingly, no evidence of linkage was detected in the families affected with both UC and CD (p>0.2). The data suggest that IBD1 may also contribute to susceptibility to ulcerative colitis, and that it is likely to be located in the 12 cM interval between D16S419 and D16S409.

(*J Med Genet* 1998;35:218–221)

Keywords: ulcerative colitis; Crohn's disease; linkage; chromosome 16

Division of Medical & Molecular Genetics, UMDS, 8th Floor Guy's Tower, Guy's Hospital, London SE1 9RT, UK
M M Mirza
S V Hodgson
C G Mathew

St Mark's Hospital at Northwick Park, Watford Road, Harrow, Middlesex HA1 3UJ, UK
J Lee
J E Lennard-Jones

CRC Genetic Epidemiology Unit, Institute of Public Health, University of Cambridge, Cambridge CB2 2SR, UK
D Teare
D F Easton

Laboratoire de Génétique des Tumeurs, INSERM U434, Fondation Jean Dausset/CEPH, 27 Rue Juliette Dodu, 75010 Paris, France
J-P Hugot
P Laurent-Puig
G Thomas

Unit of Epidemiology, Centre Hospitalier Regional et Universitaire, 59037 Lille, France
J F Colombel

Correspondence to:
Dr Mathew.

Received 4 June 1997
Revised version accepted for publication 4 September 1997

Inflammatory bowel disease (IBD) is a chronic inflammation of the gastrointestinal tract which occurs as one of two disorders, Crohn's disease (CD) or ulcerative colitis (UC). In CD the inflammation may occur in any part of the gastrointestinal tract, whereas in UC it is confined to the colon and rectum. Several epidemiological studies have suggested that genetic susceptibility may make a significant contribution to the aetiology of IBD. Orholm *et al*[1] found that first degree relatives of both CD and UC patients had an increased risk of approximately ten-fold of developing the same disorder, whereas Meucci *et al*[2] found lower risks for relatives of CD and UC patients (λ_R=6.6 and 3.4, respectively). In the latter study, the risk to sibs was higher in CD (λ_S=15) than in UC (λ_S=6.0). In a British study, the λ_S for CD was 24.7.[3] Twin studies also indicate a genetic component, with concordance for CD in monozygotic twins from the Swedish Twin Registry being 58.3%, but only 6.3% for UC.[4] In a British study the MZ concordance rates were 20% and 16% for CD and UC, respectively.[5] The general conclusion from these studies is that genetic susceptibility contributes to the aetiology of both CD and UC, but that the effect is likely to be stronger in CD.

The epidemiological evidence for a genetic component in IBD has led to a variety of approaches aimed at the identification of susceptibility genes. The results of association studies of candidate genes such as the major histocompatibility complex in different populations have been conflicting, but recently the rare HLA alleles DRB1*103 and DRB1*12 were found to be associated with UC in the British population, with no association being found in CD.[6] We found no evidence for excess sharing of HLA haplotypes in 43 British IBD families with multiple affected cases, including 31 sib pairs with UC.[7] Recent work has focused on genome scans of IBD families with affected sib pairs or multiple cases in order to identify susceptibility loci which could serve as starting points for cloning the relevant genes. Hugot *et al*[8] mapped a locus for CD to chromosome 16, which they called IBD1, and Satsangi *et al*[9] found evidence for IBD loci on chromosomes 3, 7, and 12. The chromosome 16 finding has been replicated in two subsequent studies.[10 11]

An important issue regarding the genetics and pathogenesis of IBD is whether any or all susceptibility genes are shared by both CD and UC. There is evidence that UC is more common in the relatives of CD patients and vice versa, and approximately one third of extended pedigrees which we ascertained contained cases of both CD and UC.[12] The existence of the CD susceptibility locus on chromosome 16 (IBD1) allowed us to test the hypothesis of shared susceptibility genes in IBD for this locus. We have therefore typed markers from the IBD1 locus in kindreds with multiple cases of either UC only, or with both

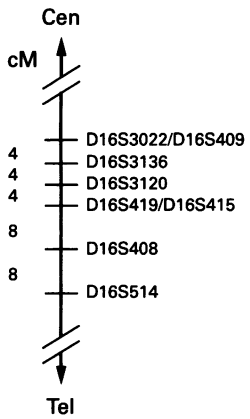

Figure 1 Linkage map of chromosome 16 markers.[9]

UC and CD, and analysed the data by both parametric and non-parametric linkage methods. We also investigated the unconfirmed additional CD susceptibility locus on chromosome 1p[8] in this family panel.

Methods

The ascertainment of IBD families, diagnostic criteria, and many of the pedigrees have been described previously in detail.[7 12] Briefly, a diagnosis of CD was made if granulomata were present with one or more accepted anatomical criteria,[13] and UC was diagnosed if there were classical symptoms of rectal bleeding with evidence of large intestinal inflammation in continuity with the rectum. The current study included 351 subjects from 70 extended IBD kindreds; 65 were white families of northwestern European origin (60 British and five French) and five families resident in the United Kingdom were of Asian origin. There were 32 kindreds with UC only (nine with two cases, 21 with three cases, and two with four cases) and 38 kindreds of "mixed" IBD phenotype, that is, with cases of UC and CD (12 with two cases, 19 with three cases, six with four cases, and one with five cases).

DNA was isolated and microsatellite markers typed in all 351 subjects as described previously.[7] Briefly, forward PCR primers were labelled with fluorescent dye primers, and loci amplified in simplex or multiplex 5 µl reactions in 96 well microtitre plates; separation and analysis of microsatellite alleles was carried out on an ABI 373 DNA analyser with Genescan 672 and Genotyper software (Applied Biosystems Inc). The markers typed were as follows: for chromosome 16, D16S3022, D16S409, D16S3136, D16S3120, D16S415, D16S419, D16S408, and D16S514; for chromosome 1, D1S435, D1S236, D1S206, D1S495, D1S239, D1S248, D1S221, and D1S502.[14] The allele frequencies of all markers were estimated from the founders in the 70 IBD families.

Statistical analysis of the genotype data was carried out using the GENEHUNTER program,[15] which is able to perform both a multipoint non-parametric analysis of linkage data in complex disorders (which does not require specification of genetic models) and a parametric analysis, over a fixed marker map. The non-parametric analysis is performed by comparing the estimated number of shared haplotypes between pairs of affected subjects with that which would be predicted by Mendelian segregation, which is generally more powerful than using only sib or relative pairs. The non-parametric linkage scores (NPL) presented here have been calculated using the "All" option, which gives increased weight to haplotype sharing in families with larger numbers of affected subjects. The normalised NPL scores and associated p values are indicated. In view of the extended nature of the IBD pedigrees in our sample, and the fact that segregation analysis supports single gene models,[16 17] the data were also analysed using four parametric models, allowing for genetic heterogeneity. This also permits estimation of the proportion of families which might be attributable to this locus. The basis of the models used has been described previously.[7] The models were: (1) common dominant, gene frequency (gf) 0.025, penetrance in carriers (pc) 0.2, penetrance in non-carriers (pnc) 0.0001; (2) common recessive, gf 0.2, pc 0.025, pnc 0.0001; (3) rare dominant, gf 0.006, pc 1.0, pnc 0.001; (4) rare recessive, gf 0.0004, pc 0.18, pnc 0.001. The chromosome 16 linkage was also investigated by calculation of the mean proportion of alleles shared identical by descent in affected sibs exactly as described in the original publication of this linkage,[8] so that the results of this and the previous studies could be compared directly. Finally, in the 30 UC families with at least one pair of affected sibs, the maximum likelihood estimates for pairs sharing 0, 1, and 2 alleles were used to estimate the sib relative risk attributable to IBD1,[18 19] assuming its coincidence with the marker D16S3120 (see multipoint analysis in table 1).

Table 1 Multipoint non-parametric linkage scores by marker position

Marker	*NPL score*	*p value*
(A) UC only families		
D16S3022	1.13	0.12
D16S409	1.33	0.09
D16S3136	1.93	0.02
D16S3120	2.02	0.02
D16S419	1.75	0.04
D16S415	1.63	0.05
D16S408	1.51	0.06
D16S514	1.15	0.12
(B) Mixed families		
D16S3022	−0.08	0.52
D16S409	−0.34	0.63
D16S3136	−0.11	0.54
D16S3120	−0.23	0.58
D16S419	0.08	0.45
D16S415	0.15	0.43
D16S408	0.27	0.38
D16S514	0.59	0.26
(C) All families		
D16S3022	0.69	0.24
D16S409	0.62	0.26
D16S3136	1.20	0.11
D16S3120	1.17	0.11
D16S419	1.23	0.10
D16S415	1.20	0.11
D16S408	1.21	0.11
D16S514	1.21	0.11

Results

We investigated the involvement of IBD1 in susceptibility to the broader IBD phenotype by typing eight highly informative microsatellite markers from this locus in 70 IBD kindreds, including 32 affected only with UC and 38 of "mixed" phenotype with both UC and CD. The markers chosen were four which had previously shown the most significant linkage to CD,[8] and four new markers[14] which were located close to or between the original markers. The linkage map of these markers, which covers a region of 28 cM, is shown in fig 1.

The NPL scores from the GENEHUNTER analysis and associated p values that were obtained for the eight markers in the "pure" UC families, in the "mixed" families, and in the complete set of families (pure UC and mixed) are shown in table 1. Evidence of linkage at several of the loci tested was obtained in the

Table 2 Multipoint parametric lod scores at marker position D16S3120

Model*	HLod	Alpha
(A) UC only families		
Common dominant	1.20	0.43
Rare dominant	0.88	0.27
Common recessive	0.20	0.23
Rare recessive	0.09	0.70
(B) "Mixed" families		
Common dominant	0.00	0.04
Rare dominant	0.01	0.02
Common recessive	0.00	0.00
Rare recessive	0.00	0.00
(C) All families		
Common dominant	0.63	0.19
Rare dominant	0.47	0.12
Common recessive	0.11	0.11
Rare recessive	0.01	0.08

*See Materials and methods for details.
Alpha=proportion of linked families.

UC families, with a maximum multipoint NPL score of 2.02 (p=0.02) at D16S3120. However, the scores in the mixed families did not provide any evidence of linkage (p>0.2 for all markers). The multipoint NPL scores for the combined set of 70 kindreds were positive for all markers tested (NPL max=1.23, p=0.10 at D16S419), but were not statistically significant. Inclusion of only the 60 British families had no effect on the analysis (NPL max=1.16, p=0.12). The parametric analysis (table 2) produced the highest lod scores at D16S3120 under a dominant model with heterogeneity. The lod score was 1.20 (p=0.0094) in the UC only families (α=0.43) and 0.63 (p=0.0443) in all families (α=0.19).

In order to facilitate a comparison between the data in the UC and mixed families from this study with the data from the CD families,[8] the mean proportion of alleles shared identical by descent for the eight markers is shown in table 3, together with the t statistics and associated p values. In the UC only families, the highest sharing was observed at D16S3136 (p=0.01). The variation in p values for adjacent markers results from the fact that only informative matings can be used for each marker. The maximum likelihood estimates for sib pairs sharing 0, 1, and 2 alleles were 0.17, 0.50, and 0.32, respectively. This translates to an estimated sib relative risk attributable to IBD1 in the UC families of 1.46. As in the GENEHUNTER analysis, no evidence for linkage to any of the chromosome 16 markers was detected in the mixed families.

Table 3 Allele sharing in affected sib pairs

Marker	Mean proportion of alleles shared ibd	t test	p value
UC only families			
D16S3022	0.69	2.09	0.02
D16S409	0.51	0.06	0.47
D16S3136	0.70	2.29	0.01
D16S3120	0.61	1.22	0.11
D16S419	0.66	1.86	0.03
D16S415	0.53	0.36	0.36
D16S408	0.49	−0.06	0.52
D16S514	0.47	−0.37	0.65
Mixed families			
D16S3022	0.51	0.14	0.44
D16S409	0.50	0.00	0.50
D16S3136	0.47	−0.33	0.63
D16S3120	0.42	−0.76	0.78
D16S419	0.43	−0.99	0.84
D16S415	0.46	−0.71	0.76
D16S408	0.40	−1.24	0.89
D16S514	0.45	−0.76	0.78

*ibd denotes alleles shared identical by descent.

The additional putative CD susceptibility locus on chromosome 1p was investigated by typing eight microsatellite markers from a 23 cM region (see Methods) in all 65 British kindreds and analysing the data with GENEHUNTER. The only positive NPL score was for the UC families at D1S206 (NPL=0.37, p=0.353); for the mixed families and for all families the NPL score was negative at all positions. Thus, no significant evidence for linkage was obtained with any of the markers tested in any group of families. There was also no evidence of linkage in the parametric analysis, allowing for heterogeneity.

Discussion

This study was designed to test the hypothesis that the susceptibility locus for Crohn's disease, which was identified on chromosome 16[8] and replicated by two other groups,[10 11] might also confer susceptibility to ulcerative colitis. The GENEHUNTER analysis of eight markers on chromosome 16, including those which showed the strongest evidence for linkage to CD[8] and additional markers from the region, produced a positive non-parametric linkage score in 32 kindreds which contained 89 subjects affected with UC, with an associated p value of 0.02 (table 1). The affected sib pair analysis of these data showed a significant excess of allele sharing for three markers in the region (table 3), with the highest proportion of sharing at D16S3136 (0.70, with p=0.01). These results are very similar to those obtained for the chromosome 16 markers in the two separate panels of Crohn's disease families reported by Hugot *et al*,[8] which are each of comparable size to the UC panel analysed in this study. The estimated sib relative risk of 1.46 in the UC families was also very similar to the estimate of 1.3 from the CD families.[8]

In view of the fact that only two regions of the genome were tested for linkage in this study (the IBD1 locus and the putative locus on chromosome 1p), the data provide good evidence that the IBD1 locus is also involved in susceptibility to UC.[20] It was therefore surprising that we did not find support for linkage to IBD1 in the 38 mixed families in which cases of both UC and CD were present (tables 1, 2, and 3). It is possible that families with the "mixed" phenotype have a higher proportion of sporadic cases in which genetic susceptibility is less important. An alternative, and we believe more likely, explanation is that since the degree of susceptibility conferred by IBD1 is modest (relative risk to sibs=1.3-1.46), its effect will not be shown in all data sets unless very large numbers of families are analysed.[20] This may also explain why our finding of evidence of linkage of the UC phenotype to IBD1 contrasts with two reports,[10 11] which, while replicating the linkage of CD to the IBD1 locus, did not detect linkage in UC. Ohmen *et al*[10] typed 23 sib pairs from 16 UC families, and combined these data with 19 sib pairs from 15 mixed families for their analysis, which showed no excess of haplotype sharing in affected sibs.

Their study differs from ours in that a much smaller number of families were studied; population specific differences may also have contributed, since there was no evidence of linkage of the Jewish CD families in their panel to IBD1. These distinctions do not apply to the Oxford study, which analysed 64 UC sib pairs from British families.[11] The Oxford and London families also appear to differ with respect to the contribution of HLA genes to susceptibility to UC.[6 7]

The multipoint analysis in the original report produced a peak between D16S409 and D16S419.[8] We analysed several additional markers in this interval (fig 1) and obtained a peak multipoint NPL score at D16S3120, which is 4 cM proximal to D16S419. This is in agreement with both replication studies,[10 11] which found the strongest evidence for linkage to CD at D16S411. Thus, IBD1 appears most likely to be located in the 12 cM interval between D16S409 and D16S419. This does not appear to contain any obvious candidate susceptibility genes at present. Further refinement by high resolution mapping of very large numbers of families and linkage disequilibrium analysis will be required to narrow it further and permit identification of candidates for IBD1.

Finally, the hypothesis of overlapping susceptibility genes for CD and UC can be tested further if the recent evidence of IBD loci on chromosomes 3, 7, and 12[9] is replicated in other studies.

We thank the IBD families and their physicians for their cooperation in this study, and the following organisations for financial support: the Crohn's Disease in Childhood Research Association, the National Association of Colitis and Crohn's Disease, the Sir Halley Stewart Trust, the Cancer Research Campaign, the Association Aupetit, the Fonds de Recherche et de l'Enseignement Superieur, INSERM, and the Ligue National Contre le Cancer.

1 Orholm M, Munkholm P, Langholz E, Nielsen OH, Sorensen TIA, Binder V. Familial occurrence of inflammatory bowel disease. *N Engl J Med* 1991;**324**:84-8.
2 Meucci G, Vecchi M, Torgano G, *et al.* Familial aggregation of inflammatory bowel disease in northern Italy: a multicenter study. *Gastroenterology* 1992;**103**:514-19.
3 Satsangi J, Rosenberg WMC, Jewell DP. The prevalence of inflammatory bowel disease in relatives of patients with Crohn's disease. *Eur J Gastroenterol Hepatol* 1994;**6**:413-16.
4 Tysk C, Lindberg E, Jarnerot G, Floderus-Myrhed B. Ulcerative colitis and Crohn's disease in an unselected population of monozygotic and dizygotic twins. A study of heritability and the influence of smoking. *Gut* 1988;**29**:990-6.
5 Thompson NP, Driscoll R, Pounder RE, Wakefield AJ. Genetics versus environment in inflammatory bowel disease. *BMJ* 1996;**312**:95-6.
6 Satsangi J, Welsh KI, Bunce M, *et al.* Contribution of genes of the major histocompatibility complex to susceptibility and disease phenotype in inflammatory bowel disease. *Lancet* 1996:**347**:1212-17.
7 Naom I, Lee J, Ford D, *et al.* Analysis of the contribution of HLA genes to genetic predisposition in inflammatory bowel disease. *Am J Hum Genet* 1996;**59**:226-33.
8 Hugot JP, Laurent-Puig P, Gower-Rousseau C, *et al.* Mapping of a susceptibility locus for Crohn's disease on chromosome 16. *Nature* 1996;**379**:821-3.
9 Satsangi J, Parkes M, Louis E, *et al.* Two stage genome-wide search in inflammatory bowel disease provides evidence for susceptibility loci on chromosomes 3, 7 and 12. *Nat Genet* 1996;**14**:199-202.
10 Ohmen JD, Yang HY, Yamamoto KK, *et al.* Susceptibility locus for inflammatory bowel disease on chromosome 16 has a role in Crohn's disease, but not in ulcerative colitis. *Hum Mol Genet* 1996;**5**:1679-83.
11 Parkes M, Satsangi J, Lathrop M, Bell JI, Jewell DP. Susceptibility loci in inflammatory bowel disease. *Lancet* 1996; **348**:1588.
12 Lee JCW, Lennard-Jones JE. Inflammatory bowel disease in 67 families each with three or more affected first-degree relatives. *Gastroenterology* 1996;**111**:587-96.
13 Lennard-Jones JE. Classification of inflammatory bowel disease. *Scand J Gastroenterol* 1989;**24**(suppl 170):2-6.
14 Dib C, Faure S, Fizames C, *et al.* The Genethon human genetic linkage map. A comprehensive genetic map of the human genome based on 5,264 microsatellites. *Nature* 1996:**380**:152-4.
15 Kruglyak L, Daly MJ, Reeve-Daly MP, Lander ES. Parametric and nonparametric linkage analysis: a unified multipoint approach. *Am J Hum Genet* 1996;**58**:1347-63.
16 Orholm M, Iselius L, Sorensen TIA, Munkholm P, Langholz E, Binder V. Investigation of inheritance of chronic inflammatory bowel diseases by complex segregation analysis. *BMJ* 1993;**306**:20-4.
17 Kuster W, Pascoe L, Purrmann J, Funk S, Majewski F. The genetics of Crohn disease: complex segregation analysis of a family study with 265 patients with Crohn disease and 5,387 relatives. *Am J Med Genet* 1989;**32**:105-8.
18 Holmans P. Asymptotic properties of affected-sib-pair linkage analysis. *Am J Hum Genet* 1993;**52**:362-74.
19 Kruglyak L, Lander E. Complete multipoint sib-pair analysis of qualitative and quantitative traits. *Am J Hum Genet* 1995;**57**:439-54.
20 Lander E, Kruglyak L. Genetic dissection of complex traits: guidelines for interpreting and reporting linkage results. *Nat Genet* 1995;**11**:241-7.