

Fine mapping of the chromosome 3p susceptibility locus in inflammatory bowel disease

J Hampe, N J Lynch, S Daniels, S Bridger, A J S Macpherson, P Stokkers, A Forbes, J E Lennard-Jones, C G Mathew, M E Curran, S Schreiber

Abstract

Background and aims—Genetic predisposition for inflammatory bowel disease (IBD) has been demonstrated by epidemiological and genetic linkage studies. Genetic linkage of IBD to chromosome 3 has been observed previously. A high density analysis of chromosome 3p was performed to confirm prior linkages and elucidate potential genetic associations.

Methods—Forty three microsatellite markers on chromosome 3 were genotyped in 353 affected sibling pairs of North European Caucasian extraction (average marker density 2 cM in the linkage interval). Marker order was defined by genetic and radiation hybrid techniques.

Results—The maximum single point logarithm of odds (LOD) score was observed for Crohn's disease at D3S3591. Peak multipoint LOD scores of 1.65 and 1.40 for the IBD phenotype were observed near D3S1304 (distal 3p) and near D3S1283 in the linkage region previously reported. Crohn's disease contributed predominantly to the linkage. The transmission disequilibrium test showed significant evidence of association ($p=0.009$) between allele 4 of D3S1076 and the IBD phenotype (51 transmitted *v* 28 non-transmitted). Two known polymorphisms in the CCR2 and CCR5 genes were analysed, neither of which showed significant association with IBD. Additional haplotype associations were observed in the vicinity of D3S1076.

Conclusions—This study provides confirmatory linkage evidence for an IBD susceptibility locus on chromosome 3p and suggests that CCR2 and CCR5 are unlikely to be major susceptibility loci for IBD. The association findings in this region warrant further investigation.

(Gut 2001;48:191-197)

Keywords: inflammatory bowel disease; fine mapping; chromosome 3

A genetic component in the aetiology of inflammatory bowel disease (IBD) has been clearly demonstrated by epidemiological and genetic linkage studies. Epidemiological investigations have consistently shown familial clustering¹ and an increased concordance of the IBD phenotype in monozygotic twins.^{2,3} Prevalence rates of IBD between ethnic and geographic isolates differ significantly. This is possibly due to a combination of different genetic and environmental factors operating in different popula-

tions. A genetic component underlying a portion of these differences is indicated by prevalence differences between ethnic groups living in the same geographic region, for example the 2-8-fold higher prevalence of IBD observed in Ashkenazi Jews versus non-Jews.⁴ Linkage and association analyses for several IBD predisposition regions have produced different results in independent family samples.⁵⁻⁷ The variance in linkage results may be attributed to several factors including: (1) differences in the genetic causation of IBD (that is, multiple disease genes); (2) use of different diagnostic criteria; (3) inability to recruit sufficient numbers of patients in some populations; (4) methodological limitations of linkage and association analyses for complex genetic disorders⁸; and (5) use of different genetic markers. To determine the significance of a given linkage result, replication of a genetic linkage observation using large independent sample sets is required. Several groups investigating the genetic components of IBD have recently completed thorough genome wide linkage analyses in large patient cohorts.^{5,6,9,10} These studies have defined three major IBD susceptibility loci located on chromosomes 6, 12, and 16. These findings are also supported by replication studies in a number of independent patient collections.^{5-7,11-17}

In addition to the well supported linkages to chromosomes 6, 12, and 16, a number of interesting secondary linkages have been defined.^{5,9,10} The linkage region on chromosome 3 was initially identified by Satsangi and colleagues⁵ who established linkage with supporting *p* value statistics of 0.0026 and 0.00021 in primary and follow on family sets. The minimum *p* value of 0.00021 corresponds to a logarithm of odds (LOD) score of 3.4 (as transformed according to the equation: $\chi^2 = \text{LOD} \times 2 \ln(10)$).¹⁸ This region encompasses the proximal segment of chromosome 3p and is defined by markers D3S1076 and D3S1573. Subsequent linkage studies have not supported the chromosome 3 linkage. A recent study in 161 IBD families of Canadian descent yielded LOD scores of 0.07-0.25.¹⁹ A second study in 58 families of Italian descent also failed to show evidence of linkage in this region.¹⁶ The genome wide study by Cho *et al* provided a peak multipoint LOD score of 1.0 in this region.⁹

Abbreviations used in this paper: IBD, inflammatory bowel disease; CD, Crohn's disease; UC, ulcerative colitis; HIV, human immunodeficiency virus; PCR, polymerase chain reaction; LOD, logarithm of odds; TDT, transmission disequilibrium test; IL5RA, interleukin 5 receptor alpha.

Ist Medical Department, Christian-Albrechts-University, Kiel, Germany
J Hampe
N J Lynch
S Schreiber

Axys Pharmaceuticals Inc., La Jolla, California, USA
S Daniels
M E Curran

King's College School of Medicine, London, UK
S Bridger
A J S Macpherson

Academic Medical Centre, Amsterdam, the Netherlands
P Stokkers

Division of Medical and Molecular Genetics, GKT Medical School, Guy's Hospital, London, UK
C G Mathew

St Mark's Hospital, Harrow, UK
A Forbes
J E Lennard-Jones

Correspondence to: Professor S Schreiber, Ist Department of Medicine, Christian-Albrechts-Universität Kiel, Schittenhelmstr. 12, D-24105 Kiel, Germany. s.schreiber@mucosa.de

Accepted for publication 22 August 2000

Table 1 Family cohort used in the study; affected sibling pairs are ordered according to the size of the sibship and disease type (mixed refers to sibships with both ulcerative colitis (UC) and Crohn's disease (CD)). The number of families and affected sib pairs (in parentheses) are given in each category. In 55% of the 268 families, both parents were available for analysis.

Sibship size	CD	UC	Mixed (CD/UC)	Total
2	114 (114)	78 (78)	38 (38)	230 (230)
3	14 (42)	12 (36)	9 (27)	35 (105)
4	1 (6)	—	2 (12)	3 (18)
All	129 (162)	90 (114)	49 (77)	268 (353)

Table 2 Overview of the oligonucleotides used for typing of the *CCR2* and *CCR5* genes. The primer sequences were obtained from the literature^{21,22} and the database sequences

Polymorphism	Type	Oligonucleotides
CCR2 G to A change	FWD	TTGTGGGCAACATGATGG
	REV	GAGCCACAATGGGAGAGTA
CCR5 32 bp deletion	FWD	TACCTGGCTGTCGTCCATGC
	REV	TGACCATGACAAGCAGCG

In our previous genome wide linkage study, we reported a suggestive multipoint LOD score of 1.2 for this susceptibility region.¹⁰ The variable LOD scores observed for this interval are not unexpected for a complex disorder such as IBD^{8,20} and may be due to a reduced penetrance of this locus relative to others and methodological issues of non-parametric linkage analyses. Despite disparate genetic results for linkage of IBD to chromosome 3,^{16,19} the region is of significant interest. Two autoimmune disorders, multiple sclerosis and inflammatory arthritis, have been genetically linked to chromosome 3p. These independent observations suggest that a gene, or perhaps multiple genes, involved in regulating immune function and inflammatory response reside in this region of the genome. Additional interest in the 3p region stems from the observation that numerous candidate genes, including the chemokine receptor cluster and interleukin 5 receptor α , are located on chromosome 3p.²¹⁻²³

We have therefore performed a high resolution linkage and association study on chromosome 3p. Our goal was to better define the significance of the chromosome 3p locus and to directly test several of the candidate genes from the region of interest for involvement in IBD. Highly polymorphic microsatellite markers saturating chromosome 3p with an average spacing of 2 cM were genotyped in our large family cohort. Genetic linkage and a systematic association analysis using both single point and two point transmission disequilibrium test (TDT) statistics was then performed. In the vicinity of the single most significant association result, we selected two candidate genes and tested known variations that were previously associated with resistance to human immunodeficiency virus (HIV) infection.^{21,22} We demonstrate that these variants are not likely to be involved in susceptibility to IBD and suggest a more proximal localisation of the chromosome 3p susceptibility gene for IBD.

Materials and methods

FAMILY ASCERTAINMENT AND PHENOTYPES

The population studied here has been described previously.^{10,12,15,24,25} The cohort was recruited in Europe by several cooperative cen-

tres including: Charite University Hospital (Berlin, Germany), I Department of Medicine at the Christian-Albrechts-Universität (Kiel, Germany), King's College School of Medicine, Guy's Hospital, and St Mark's Hospital (London, UK), AMC (Amsterdam, the Netherlands). Forty six percent of families were of German origin, 6% were from the Netherlands, and 48% of families were recruited in the UK. All study participants gave informed written consent. The recruitment protocols were approved prior to study initiation by the respective institutional review boards. The diagnosis of IBD and classification into Crohn's disease (CD) and ulcerative colitis (UC) were determined by standard diagnostic criteria.^{26,27} Ascertainment criteria were determined prior to the initiation of patient collection. Medical records for all patients were reviewed by one or more of the principal investigators in the families originating from the UK and the Netherlands. For families of German extraction, patients were directly examined by one or more of the principal investigators, where possible. Alternatively, two written records containing a detailed disease history and results of all diagnostic procedures were obtained for each patient and reviewed by the principal investigators. A venous blood sample was obtained from the affected siblings and their parents if possible. An overview of the family cohort is given in table 1.

GENOTYPING

Genomic DNA was prepared from whole venous blood using the Puregene system (Gentra Systems, Minneapolis, Minnesota, USA). Blood samples were stored at room temperature for up to one week or frozen and stored at -70°C for up to nine months prior to purification. Forty three polymorphic microsatellite markers covering chromosome 3 were genotyped using polymerase chain reaction (PCR) with fluorescent labelled primers, as described previously.²⁸ The primer sequences were derived from GDB or the literature.^{21,23} In brief, individual DNA samples were arrayed in 96 well microtitre plates and amplified by PCR with the respective primers. Product length of the PCR products was determined by electrophoresis on denaturing polyacrylamide gels using ABI 377 automated DNA sequencers. Data were collected using the PE Applied Biosystems (Foster City, California, USA) Genescan software. Allele analyses and individual allele calling were performed as described previously.^{28,29} Genotype errors as a result of non-mendelian segregation in pedigrees were detected and corrected as described by Hall and Nanthakumar.²⁸

Variants in *CCR2* were genotyped by PCR-RFLP using the primers in table 2. Amplification generated a 128 bp product, which was digested into 110 and 18 bp fragments with *Bsa* *BI* when isoleucine was substituted for valine at position 64 in the *CCR2* gene. The *CCR5* site was amplified with the primers in table 2 and the presence or absence of the 32 bp deletion was directly scored on 4% agarose gel.

Table 3 Single point logarithm of odds (LOD) scores and transmission disequilibrium test (TDT) results for the loci on chromosome 3. Single point LOD scores were obtained using Mapmaker/Sibs. TDT statistics were obtained from Genehunter 2.0. Nominal TDT results are only quoted for *p* values less than 0.01; for two locus TDT tests, values are noted in the row of the first marker (for example, the value of 0.004 quoted for D3S2338 refers to the haplotype D3S2338-D3S3726). The markers used also in the genome wide scan are indicated in italics. The most significant findings for single point and two point TDT analysis are indicated in bold type

Marker	Theta ^a	Two point MLS		All	TDT	
		UC	CD		1 locus	2 loci
IL5RA	0.025	0.00	0.96	0.09		
D3S3706	0.064	0.05	0.29	0.71		
D3S1304	0.043	0.00	0.20	0.28		
D3S3591	0.061	0.00	1.30	0.90		
<i>D3S1263</i>	0	0.00	0.30	0.00		
D3S3680	0.011	0.00	0.02	0.00		
D3S1259	0.006	0.00	0.07	0.00		
D3S3693	0.026	0.00	0.09	0.00		
<i>D3S3608</i>	0.031	0.00	0.00	0.44		
D3S1255	0.024	0.00	0.04	0.02		
D3S2338	0.019	0.00	0.00	0.00		0.004 (all)
<i>D3S3726</i>	0.063	0.27	0.11	0.10		
D3S2336	0.019	0.38	0.28	0.44		
D3S2337	0.024	0.00	0.05	0.10		0.004 (all)
<i>D3S1266</i>	0.011	0.00	0.04	0.04		0.004 (CD)
D3S1283	0.025	0.00	0.01	0.01		
D3S3547	0.009	0.00	0.28	0.32		0.002 (all)
D3S3727	0.008	0.01	0.23	0.39		
D3S1211	0.037	0.25	0.09	0.17		
D3S3518	0.015	0.28	0.10	0.26		
D3S1277	0.034	0.12	0.05	0.31		
D3S1298	0.019	0.36	0.22	0.23		
D3S3527	0.004	0.31	0.47	0.57		
<i>D3S3521</i>	0.006	0.70	0.30	0.79		
AFM206zf2	0.026	0.28	0.05	0.05		
D3S3624	0.012	0.87	0.44	0.95		
AFMb362wb9	0.004	0.00	0.00	0.00		
CCR2	0	0.00	0.23	0.00		0.004 (CD)
CCR5	0.016	0.00	0.19	0.00		0.003 (all)
D3S1076	0.012	0.42	0.05	0.16	0.009 (all)	
D3S1289	0.073	0.00	0.40	0.11		0.0009 (all)
<i>D3S1300</i>	0.089	0.00	0.57	0.30		0.006 (all)
<i>D3S1285</i>	0.065	0.02	0.01	0.09		
<i>D3S1566</i>	0.092	0.01	0.08	0.15		
<i>D3S3653</i>	0.084	0.32	0.27	0.64		
<i>D3S1271</i>	0.117	0.03	0.46	0.26		
<i>D3S1278</i>	0.088	0.03	0.78	0.16		
<i>D3S1267</i>	0.09	0.00	1.06	0.13		
<i>D3S1292</i>	0.092	0.20	0.27	0.35		
<i>D3S1569</i>	0.119	0.11	0.60	0.88		
<i>D3S1279</i>	0.142	0.12	0.14	0.23		
<i>D3S1565</i>	0.138	0.00	0.41	0.10		
<i>D3S1602</i>	0.067	0.23	0.00	0.00		
<i>D3S1580</i>	0.171	0.19	0.91	0.46		0.008 (all)
<i>D3S1311</i>	—	0.26	0.55	0.28		

^aMarker distances were calculated from the families analysed using the program Multimap.³⁰ Theta is the recombination fraction between two markers—that is, 0.025 corresponds to 25 recombination events in 1000 meioses. For small thetas the map distance in morgans is nearly identical to the theta—that is, 0.025=0.025 morgans=2.5 cM.

Marker order and distance separating each marker were defined by information derived from our pedigrees using the automated mapping program Multimap v 2.0.³⁰ The maps were confirmed by comparison with published genetic maps (CHLC and Genethon at <http://www.ncbi.nlm.nih.gov>). Placement of previously unlocalised markers was confirmed by radiation hybrid analysis using the Genebridge 3 radiation hybrid (Research Genetics, Huntsville, Alabama, USA) panel. RH vectors were analysed using the Stanford RH mapping server (at <http://www-shgc.stanford.edu/>) and by using the Radmap extension of the Multimap program.

STATISTICAL ANALYSIS

Genetic analyses were conducted using the two IBD subphenotypes CD and UC. A third category, “ALL”, contained CD/CD, UC/UC, and CD/UC (mixed) affected sibling pairs. The ALL category therefore represents IBD as a single phenotype for analysis. Allele frequen-

cies for each marker were calculated from the cohort genotype data using all individuals. Data were analysed with both single point and multipoint non-parametric allele sharing tests in affected sibling pairs using Mapmaker/Sibs with the “weighted pairs” option.³¹ For multipoint analysis, LOD scores were computed at 1 cM intervals along the chromosome. Mean information content across chromosome 3p was 93%.

Association statistics were calculated using the TDT (for single point TDT tests) and TDT2 (for two marker TDT tests) functions of the Genehunter 2.0 program.³²⁻³³ TDT provides a statistical measure of genetic association and linkage that is robust against population stratification, because it scores transmission events from parents to offspring. Data derived from less than 10 observed transmission events were excluded from the results to reduce the possibility of false positive findings. The program performs a classical TDT test using only heterozygote founders.³²

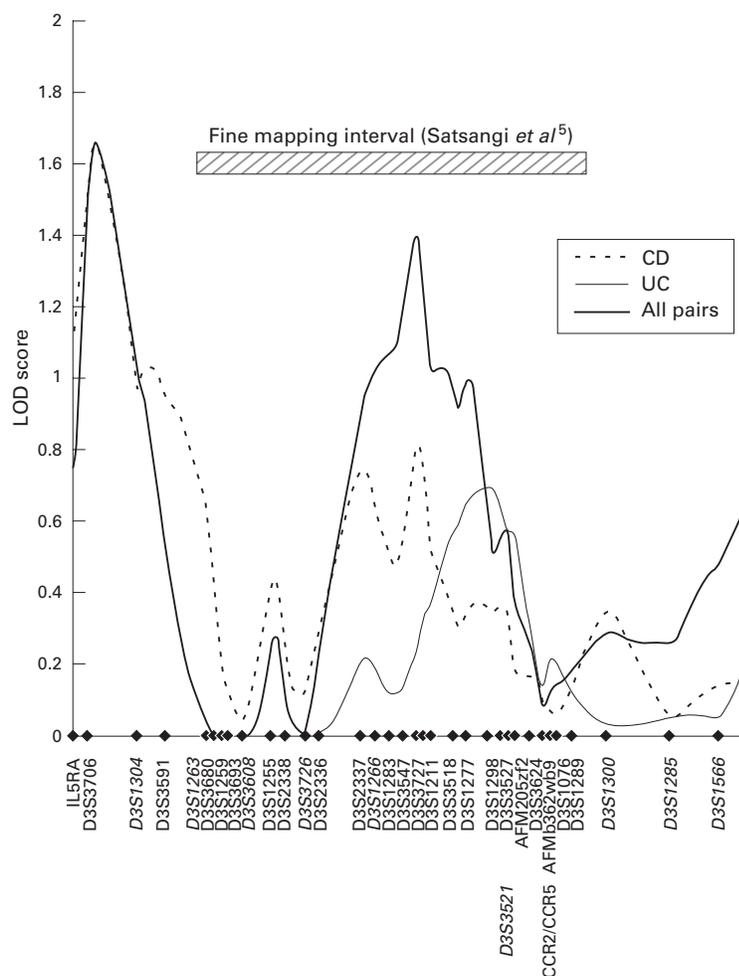


Figure 1 Multipoint MLS curves for the markers genotyped in the saturation region on chromosome 3p. Markers were genotyped in 353 affected sibling pairs and analysed for allele sharing using the "weighted pairs" option of Mapmaker/Sibs.²¹ Results for Crohn's disease (CD), ulcerative colitis (UC), and all pairs are shown. Genetic distances were estimated from the marker data using Multimap. The marker positions are indicated by filled diamonds. The markers also used in the genome wide analysis are indicated in italics.

This algorithm provides no correction for the testing of multiple alleles.

Results

A total of 43 microsatellite markers covering chromosome 3 were analysed for the study. Twenty of the markers—providing a general framework—were used in the genome wide linkage analysis¹⁰ and 23 additional saturation markers on chromosome 3p were added. Based on previously observed linkage of IBD to chromosome 3, linkage of other autoimmune disorders to chromosome 3p, and the presence of several interesting candidate genes in this region, a very high marker density was analysed in this area. The average marker density in the extended susceptibility region from D3S1304 to D3S1289 was 2 cM after estimation of the marker distances from the dataset using Multi-map.

Results from single point analysis are shown in table 3. The peak single point LOD score was obtained at D3S3591 in the CD disease category. On multipoint analysis, a peak multipoint LOD score of 1.65 near marker D3S1304 was observed for the combined IBD

phenotype (ALL category). In the linkage region previously described by Satsangi *et al*, a peak multipoint LOD score of 1.4 was obtained, also for the IBD phenotype. The linkage at D3S1304 was almost exclusively derived from the CD phenotype. This was demonstrated by the single point LOD score of 1.3 obtained in the CD category at D3S3591 versus the 0.90 single point LOD obtained in the ALL category. For linkage in the previously implicated region on the proximal p arm of chromosome 3, a more equal contribution of the CD and UC phenotypes, with CD still being predominant, was observed. The multipoint linkage curve is shown in fig 1. Although there are several regions that are suggestive of genetic linkage, none of these regions meet strict criteria for defining a genetic locus³⁴ de novo.

The primary goals of this high density genotyping experiment were to better define the 3p susceptibility region and to attempt identification of disease related linkage disequilibrium. To facilitate the latter goal, a systematic single point TDT analysis using only heterozygous parents was performed for all markers. The nominal TDT results, with the most significant p value from all alleles and subphenotypes, is presented in table 3. The most significant nominal p value ($p=0.009$) was observed for allele 4 at D3S1076 with 51 transmitted alleles versus 28 non-transmitted alleles. D3S1076 is located near the chemokine receptor 2 (CCR2) and chemokine receptor 5 (CCR5) genes. Based on the TDT result, these genes were selected for direct analysis as potential IBD susceptibility genes. Two known variants, the I64V polymorphism in CCR2 and a 32 bp deletion in CCR5, were used for this analysis. Allele frequencies observed in our cohort were similar to data previously reported at 8.3% for CCR2 and 12% for CCR5.²¹ These markers were used for both linkage and association analysis. The linkage results are shown in table 3 and fig 1. TDT analysis of the CCR2 variant revealed a p value of 0.26 (46 transmissions versus 36 non-transmissions) in the ALL category and 0.13 (32 versus 21 transmissions) for the CD phenotype. Analysis of the CCR5 deletion yielded values of $p=0.15$ (64 transmissions versus 49 non-transmissions) for the ALL category and $p=0.32$ (37 versus 29 transmissions) in the CD category.

To obtain a more sensitive measure of association and increase the possibility of identifying less prevalent but important associated haplotypes, a systematic two locus association analysis was performed across the region. The association results with higher significance levels than 0.01 are given in table 3 together with the disease category from which they were obtained. Nominal p values as low as $p=0.0009$ (in the ALL category) were recorded. Interestingly, two clusters of association results were identified around D3S1076 and, to a lesser degree, around D3S2337. The cluster around D3S1076 contained nominal association values of $p=0.004$ for the CCR2-CCR5 haplotype in CD, $p=0.003$ for the CCR5-D3S1086,

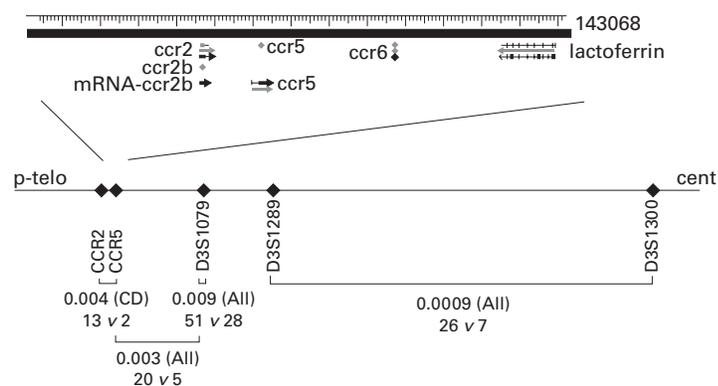


Figure 2 Detailed view of the putative association region around D3S1076. In the lower part of the figure, markers are arranged according to the genetic map. The brackets indicate single point and two point associations of markers. For each finding, the nominal p value with the phenotype category in which it was obtained and the number of transmitted versus non-transmitted alleles/haplotypes is indicated. The upper part of the figure offers a close up view of the CCR gene cluster and was derived from the NCBI website (www.ncbi.nlm.nih.gov, based on the Genbank record U95626).

and a p value of 0.0009 for the D3S1298-D3S1300 haplotype in the ALL category.

The interleukin 5 receptor alpha (IL5RA) gene is located on chromosome 3p in an area of suggestive linkage defined by our initial genome wide scan. This gene is involved in immunoregulation and is an interesting candidate for IBD. To investigate this candidate gene we genotyped an intragenic microsatellite²³ marker. No evidence of association (corrected p value >0.10) was detected. Hence IL5RA can also be excluded as a major risk determinant of IBD.

Discussion

We have reported here supporting evidence for the existence of an IBD susceptibility gene on chromosome 3p, described the results of genetic investigations for several important candidate genes, and provided a systematic high density association analysis of the chromosome 3p linkage region.

A multipoint LOD score of 1.4 in the chromosome 3 interval previously identified by Satsangi and colleagues⁹ was observed in the analysis of the ALL phenotype category. The original chromosome 3 linkage findings were driven predominantly by the CD subphenotype. In our dataset, a similar relationship regarding the contribution of CD and UC to the proximal chromosome 3p linkage was observed. We noted however that our family cohort contained an approximately 2:1 ratio of CD to UC patients. A second linkage peak on the long arm of chromosome 3 has recently been reported by Cho *et al.* We observed a multipoint LOD score of 1.0 in this region driven by analysis of the CD phenotype. The importance of this interval will be determined by further study.

There was a firm initial linkage finding for the proximal p arm with a p value of 0.00021 (corresponding to a LOD score of 3.4).¹⁸ The genome wide analyses of Cho *et al* produced a LOD score of 1.0, slightly distal to the Satsangi *et al* linkage. Additional studies in Canadian and Italian populations have failed to support linkage¹⁶⁻¹⁹ to chromosome 3. Our data provide

additional support for an IBD susceptibility gene in this genomic region. In total, these observations suggest that the chromosome 3p locus apparently has a smaller relative risk, compared with for instance the widely replicated linkage region on chromosome 16, although this locus also only accounts for a portion of the relative risk. Alternatively, the influence of the chromosome 3p locus may be significantly affected by interaction with other susceptibility genes which therefore limits the power of linkage studies. The final differentiation of these possibilities has to await the identification of the causative molecular variants that confer IBD susceptibility in this and other regions.

The primary objective driving the high density genotyping described here was the identification of disease associated disequilibrium. The methodology best suited to the identification of complex disease genes and the populations that may lead to the identification of the relevant genes are not yet clear.³⁵⁻³⁶ At this point, the use of single point and multipoint TDT testing seems to be a practical and robust way to develop an understanding of disease associated disequilibrium for a particular disease. The single point associations test yielded nominal p values to 0.009 (p value corrected for the presence of multiple alleles as given by TDTLIKE,³⁷ p=0.07). Given the number of tests performed, this result should not be over interpreted. Based on the assumption that possible founder haplotypes will be relatively rare in a modern admixed population, the fact that no significant single marker (that is, corrected p value <0.01) was identified is not unexpected. Therefore, two locus TDT analysis was performed for all markers. This approach will potentially identify more rare, but significant, haplotypes which cannot be discerned on the basis of a single polymorphic marker. Interestingly, the association findings having nominal p values <0.01 cluster at two points along the map: in the vicinity of D3S1076, and to a lesser degree, near D3S2337. We have observed association "signals" that point to this genetic region but we are not defining a single associated haplotype. This may be caused by: (i) an incorrect marker order which may occur for very closely spaced markers or (ii) the presence of multiple associated haplotypes in the population. The question of multiple testing biasing this result is clearly at issue but these findings may be used to refine the location of the chromosome 3p IBD locus and will facilitate direct candidate gene investigation and additional genetic experiments.

The chemokine receptors 2 and 5 are located in close proximity to the single marker association at D3S1076. Based on the suggestive association evidence and their functional importance in immunoregulation, these genes were chosen for direct investigation. In each of these genes, functional variants associated with resistance to HIV infection²² or disease progression (CCR2)²¹ have been identified. We investigated the putative role of CCR2 and CCR5 in IBD using these variants as polymor-

phic markers. Using TDT analysis in the 253 families with IBD, no significant influence of these mutations on the IBD phenotype was detected. Formal exclusion of CCR2 and CCR5 is not possible using the TDT method. The region around D3S1076 contains a number of suggestive association "signals" proximal to the chemokine receptor genes. Figure 2 describes the region in which the CCR2 and CCR5 genes are localised. Both reside on a fully sequenced BAC clone (Genbank record U95626) and a number of highly interesting candidate genes are located in their immediate vicinity. These include the lactotransferrin gene (fig 2), which is suspected to have a role in neutrophil function and bacterial defence, several genes of the ubiquitin complex (USP4, UQCRC1) which may play a critical role in antigen processing,³⁸ the cathelicidin antimicrobial peptide³⁹ and the TRAF interacting protein, which is important for tumour necrosis factor α signal transduction,⁴⁰ the mitogen activated protein kinase activated protein kinase 3,⁴¹ and the interferon α receptor 2.⁴² This region is extremely gene rich and harbours more than 200 transcripts within 3 cM of the chromosome (GeneMap99; <http://www.ncbi.nlm.nih.gov/genemap99/>). Given the linkage and association data implicating this region, it is possible that any one of these genes could have a role in IBD. Systematic investigation will be required to define the gene, or genes, from this region which are involved in modifying the IBD phenotype.

In summary, we have presented confirmatory linkage evidence for the existence of an IBD susceptibility gene on the proximal part of chromosome 3p. Neither CCR2, CCR5, nor IL5RA are likely to represent the IBD susceptibility gene in this region. Association results obtained in this dense microsatellite mapping experiment may facilitate identification of the risk gene located on chromosome 3p. Using a transcript based mapping approach, genes proximal to the CCR2 and CCR5 genes can be prioritised and investigated directly for involvement in IBD.

The authors thank the physicians, IBD patients, and their families for participating in this study. The cooperation of the German Crohn's and Colitis Foundation (DCCV e V), Professor Raedler/Hamburg, Professor Kruijs/Köln, Dr Theuer/Heilbronn, Dr Meckler/Gedern, Professor Lochs, Dr Wedel, T. Herrmann/Berlin, Dr Herchenbach/Recklinghausen, Professor Scheurlen/Würzburg, Dr Demharter/Augsburg, Dr Simon/Munich, Dr Purmann/Moers, Dr Jessen/Kiel, Dr Zehnter/Dortmund, Dr Lübke, Dr Weismüller/Koblenz, Dr Eiche/Denkendorf, Dr Schönfelder/Aachen, Professor Fleig/Halle all in Germany, Dr Wewalka/Linz, Dr Knofloch/Wels, both in Austria and Dr Hodgson, Dr Sanderson, Dr Pounder, Dr Forbes, Dr Forgacs/London, Dr Bird/Maidstone, Dr Hines/Haywards Heath, Dr Cairns, Dr Ireland/Brighton, Dr Barrison/St Albans and Dr Smith-Lang/Sidcup in the UK, are gratefully acknowledged. The authors acknowledge the great contribution of Dr JCW Lee in the collection of patients and the expert technical efforts of Jonalyn Matusalem, Larenia Pedriguez, Hye Jin Yang, Birte Köpke, Brigitte Mauracher, Tam Ho Kim, and Kirstin Schirmacher and the expert Macintosh support by Carl Manaster. This work was supported by Axys Pharmaceuticals Inc, the National Association for Colitis and Crohn's disease (UK), Crohn's in Childhood Research Association (UK), the Sir Halley Stewart Trust, the Deutsche Forschungsgemeinschaft (Schr 512/5-1, SFB 415), a Training and Mobility of Research (TMR) Network grant of the European Union (ERB-4061-PL-97-0389), by MFG, and a MedNet "Chronisch-entzündliche Darmerkrankungen" of the German Federal Department for Research and Education (BmBF).

- Orholm M, Munkholm P, Langholz E, et al. Familial occurrence of inflammatory bowel disease. *N Engl J Med* 1991;324:84-8.
- Tysk C, Lindberg E, Järnerot G, et al. 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.
- Thompson NP, Driscoll R, Pounder RE, et al. Genetics versus environment in inflammatory bowel disease: results of a British twin study. *BMJ* 1996;312:95-6.
- Roth MP, Petersen GM, McElree C, et al. Familial empiric risk estimates of inflammatory bowel disease in Ashkenazi Jews. *Gastroenterology* 1989;96:1016-20.
- 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.
- Hugot JP, Laurentpuig P, Gower-Rousseau C, et al. Mapping of a susceptibility locus for Crohn's disease on chromosome 16. *Nature* 1996;379:821-3.
- Cho J, Fu Y, Pickles M, et al. CAG repeat expansions in subsets of families with Crohn's disease. *Gastroenterology* 1997;112:A948.
- Hampe J, Wienker T, Nürnberg P, et al. Mapping genes for polygenic disorders: consideration for study design in the complex trait of inflammatory bowel disease. *Hum Hered* 2000;50:91-101.
- Cho JH, Nicolae DL, Gold LH, et al. Identification of novel susceptibility loci for inflammatory bowel disease on chromosomes 1p, 3q, and 4q: Evidence for epistasis between 1p and IBD1. *Proc Natl Acad Sci USA* 1998;95:7502-7.
- Hampe J, Schreiber S, Shaw SH, et al. A genome-wide analysis provides evidence for novel linkages in inflammatory bowel disease in a large European cohort. *Am J Hum Genet* 1999;64:808-16.
- 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.
- Curran ME, Lau KF, Hampe J, et al. Genetic analysis of inflammatory bowel disease in a large European cohort supports linkage to chromosomes 12 and 16. *Gastroenterology* 1998;115:1066-71.
- Cavanaugh JA, Callen DF, Wilson SR, et al. Analysis of Australian Crohn's disease pedigrees refines the localization for susceptibility to inflammatory bowel disease on chromosome 16. *Ann Hum Genet* 1998;62:291-8.
- Duerr RH, Barmada MM, Zhang L, et al. Linkage and association between inflammatory bowel disease and a locus on chromosome 12. *Am J Hum Genet* 1998;63:95-100.
- Hampe J, Shaw SH, Saiz R, et al. Linkage of inflammatory bowel disease to human chromosome 6p. *Am J Hum Genet* 1999;65:1647-55.
- Anness V, Latiano A, Bovio P, et al. Genetic analysis in Italian families with inflammatory bowel disease supports linkage to the IBD1 locus—a GISC study. *Eur J Hum Genet* 1999;7:567-73.
- Brant SR, Fu Y, Fields CT, et al. American families with Crohn's disease have strong evidence for linkage to chromosome 16 but not chromosome 12. *Gastroenterology* 1998;115:1056-61.
- Holmans P. Asymptomatic properties of affected-sib-pair linkage analysis. *Am J Hum Genet* 1993;52:362-74.
- Rioux JD, Daly MJ, Green T, et al. Absence of linkage between inflammatory bowel disease and selected loci on chromosomes 3, 7, 12, and 16. *Gastroenterology* 1998;115:1062-5.
- Kruglyak L, Lander ES. Limits on fine mapping of complex traits. *Am J Hum Genet* 1996;58:1092-3.
- Smith MW, Dean M, Carrington M, et al. Contrasting genetic influence of CCR2 and CCR5 variants on HIV-1 infection and disease progression. Hemophilia Growth and Development Study (HGDS), Multicenter AIDS Cohort Study (MACS), Multicenter Hemophilia Cohort Study (MHCS), San Francisco City Cohort (SFCC), ALIVE Study. *Science* 1997;277:959-65.
- Liu R, Paxton WA, Choe S, et al. Homozygous defect in HIV-1 coreceptor accounts for resistance of some multiply-exposed individuals to HIV-1 infection. *Cell* 1996;86:367-77.
- Epplen C, Frank G, Gomolka M, et al. Dinucleotide repeat polymorphisms in the IL2 and IL5RA genes. *Hum Mol Genet* 1994;3:679.
- Olavesen MG, Hampe J, Mirza MM, et al. Analysis of single nucleotide polymorphisms in the interleukin 4 receptor gene for association with inflammatory bowel disease. *Immunogenetics* 2000;51:1-7.
- Hampe J, Hermann B, Bridger S, et al. The interferon gamma gene as a positional and functional candidate gene for inflammatory bowel disease. *Int J Colorectal Dis* 1998;13:260-3.
- Podolsky DK. Inflammatory bowel disease. *N Engl J Med* 1991;325:928-37.
- Lennard-Jones JE. Classification of inflammatory bowel disease. *Scand J Gastroenterol Suppl* 1989;170:2-6.
- Hall J, Nanthakumar E. Automated fluorescent genotyping. In: Boyle AL, ed. *Current protocols in human genetics*. New York: John Wiley, 1997:2.8.1-2.8.19.
- Idury RM, Cardon LR. A simple method for automated allele binning in microsatellite markers. *Genome Res* 1997;7:1104-9.
- Matise TC, Perlin M, Chakravarti A. Automated construction of genetic linkage maps using an expert system

- (MultiMap): a human genome linkage map. *Nat Genet* 1994;6:384–90.
- 31 Kruglyak L, Daly MJ, Lander ES. Complete multipoint sib-pair analysis of qualitative and quantitative traits. *Am J Hum Genet* 1995;57:439–54.
 - 32 Spielman RS, McGinnis RE, Ewens WJ. The transmission/disequilibrium test detects cosegregation and linkage. *Am J Hum Genet* 1994;54:559–60.
 - 33 Kruglyak L, Daly MJ, Reeve Daly MP, et al. Parametric and nonparametric linkage analysis: a unified multipoint approach. *Am J Hum Genet* 1996;58:1347–63.
 - 34 Lander E, Kruglyak L. Genetic dissection of complex traits: guidelines for interpreting and reporting linkage results. *Nat Genet* 1995;11:241–7.
 - 35 Slager SL, Huang J, Vieland VJ. Effect of allelic heterogeneity on the power of the transmission disequilibrium test. *Genet Epidemiol* 2000;18:143–56.
 - 36 Wright AF, Carothers AD, Pirastu M. Population choice in mapping genes for complex diseases. *Nat Genet* 1999;23:397–404.
 - 37 Terwilliger JD. A powerful likelihood method for the analysis of linkage disequilibrium between trait loci and one or more polymorphic marker loci. *Am J Hum Genet* 1995;56:777–87.
 - 38 Hochstrasser M. Ubiquitin-dependent protein degradation. *Annu Rev Genet* 1996;30:405–39.
 - 39 Bals R, Weiner DJ, Mosconi AD, et al. Augmentation of innate host defense by expression of a cathelicidin antimicrobial peptide. *Infect Immun* 1999;67:6084–9.
 - 40 Ye X, Mehlen P, Rabizadeh S, et al. TRAF family proteins interact with the common neurotrophin receptor and modulate apoptosis induction. *J Biol Chem* 1999;274:30202–8.
 - 41 Lewis AJ, Manning AM. New targets for anti-inflammatory drugs. *Curr Opin Chem Biol* 1999;3:489–94.
 - 42 Kotenko SV, Izotova LS, Mirochnitchenko OV, et al. The intracellular domain of interferon-alpha receptor 2c (IFN-alphaR2c) chain is responsible for Stat activation. *Proc Natl Acad Sci USA* 1999;96:5007–12.