European Journal of Human Genetics (2001) 9, 731 – 742

www.nature.com/ejhg

ARTICLE

Genetic refinement and physical mapping of a chromosome 16q candidate region for inflammatory bowel disease

Habib Zouali[1,16], Mathias Chamaillard[1,2,16], Suzane Lesage[1], Jean-Pierre Cézard[3], Jean-Frédéric Colombel[4], Jacques Belaiche and the GETAID[5], Sven Almer[6], Curt Tysk[7], Sean Montague[8], Miquel Gassull[9], Steen Christensen[10], Yigael Finkel[11], Corine Gower-Rousseau[4], Robert Modigliani[12], Jeanne Macry[13], Hana Selinger-Leneman[14], Gilles Thomas[1,2,15] and Jean-Pierre Hugot*[,1,2,3]

[1]Fondation Jean Dausset CEPH, Paris, France; [2]INSERM Unité 434, Paris, France; [3]European Paediatric Working Group on the Genetics of IBD, Department of Paediatric Gastroenterology, Hopital Robert Debré, Paris, France; [4]Registre EPIMAD, Hôpital Calmette, Lille, France; [5]Department of Gastroenterology, CHU de Liège, Belgium; [6]Division of Gastroenterology and Hepatology, IHM, Linköpings universitet, Linköping, Sweden; [7]Department of Gastroenterology, Örebro Medical Center hospital, Örebro, Sweden; [8]Department of Gastroenterology, Adelaide & Meath Hospital, Dublin, Ireland; [9]Department of Gastroenterology, Hospital Universitari Germans Trias i Pujol, Badalona, Spain; [10]Department of Gastroenterology Herlev Hospital, Herlev, Denmark; [11]Department of Gastroenterology, Karolinska Children's Hospital, Stockholm, Sweden; [12]Department of Gastroenterology, Hôpital Saint Louis, Paris, France; [13]INSERM Unité 458, Paris, France; [14]INSERM Unité 535, Le Kremlin-Bicêtre, France; [15]Department of Surgery, Hôpital Saint Antoine, Paris, France

Crohn's disease (CD) is a complex genetic disorder for which a susceptibility gene, IBD1, has been mapped within the pericentromeric region of chromosome 16. In order to refine the location of IBD1, 77 multiplex CD families were genotyped for 26 microsatellite markers evenly spaced by approximately 1 cM. Nonparametric linkage analyses exhibited a maximum NPL score of 3.49 (P=2.37 × 10^{-4}) in a region centred by markers D16S3136, D16S3117 and D16S770. Simulation studies showed that the probability for IBD1 to be located in a 5 cM region around these markers was 70%. A 2.5 Mb YAC and BAC contig map spanning this genetic region on chromosome band 16q12 was built. TDT analyses demonstrated suggestive association between the 207 bp allele of D16S3136 (P<0.05) and a new biallellic marker hb27g11f-end (P=0.01). These markers were located in the hb27g11 and hb87b10 BAC clones from the contig. Taken together, the present results provide a crucial preliminary step before an exhaustive linkage disequilibrium mapping of putatively transcribed regions to identify IBD1. *European Journal of Human Genetics* (2001) **9,** 731 – 742.

Keywords: Crohn's Disease; fine mapping; linkage disequilibrium; positional cloning; contig map

Introduction

Inflammatory Bowel Diseases (IBD), including Crohn's Disease (CD, MIM 266600) and Ulcerative Colitis (UC, MIM 191390) are common diseases occurring in young adults that are characterized by a chronic inflammation of the digestive tract.[1] Their combined prevalence has been estimated as high as 1 per 500 inhabitants in Western countries.[2] To date, the aetiology of IBD is unknown.[3] However, the increasing incidence of the disease in developed countries since the Second World War suggests the effect of environmental factors related to the modern Western way of life.[4] A lot of putative risk factors have

*Correspondence: Jean-Pierre Hugot, Fondation Jean Dausset/CEPH, 27, rue Juliette Dodu, 75010 Paris, France. Tel: +33 1 53 72 50 24; Fax: +33 1 53 72 50 58; E-mail: hugot@cephb.fr
[16]The two first authors contributed equally to this work.
Received 1 May 2001; revised 29 June 2001; accepted 3 July 2001

been investigated but to date, cigarette smoking is the only one confirmed to play a role in IBD.[5,6]

A genetic predisposition to IBD was initially suggested by ethnic and familial aggregations of the disease.[3] Twin studies further supported the hypothesis of a genetic component for IBD by demonstrating that the disease concordance rate was higher in monozygotic twins than in dizygotic twins.[7–9] Finally, IBD susceptibility genes were recently mapped by linkage studies to chromosomes 1, 3, 5, 6, 7, 12, 14, 16, 19 and X.[10–17]

The first susceptibility locus for IBD (called IBD1) was localised to the pericentromeric region of chromosome 16.[10] Several independent studies have replicated this result.[18–23] Recently, the analysis of 613 IBD families by an international genetic consortium confirmed these observations.[24] Although the IBD1 locus is involved in susceptibility to CD in homogeneous families, its role is still the subject of debate in UC and mixed families (ie families with members affected by either CD or UC).

In linkage studies the maximum value of the linkage statistic was uniformly located in a narrow genetic region centred on the microsatellite marker D16S411 (Table 1). In contrast, the confidence interval of this localization spans on a broad genetic region, estimated to be 20–40 cM. This region is *a priori* expected to contain several hundred genes. Additional work is thus required to refine the location of the IBD1 locus and to identify the gene. In order to pursue this positional cloning strategy, new linkage analyses were performed on a family set using closely spaced microsatellite markers located in the pericentromeric region of chromosome 16. A physical map of the region was also constructed. From this contig map,transmission disequilibrium tests allowed us to target the hb27g11 and hb87b10 BAC clones.

Subjects and methods

Subjects and ascertainment of phenotypic data

Because IBD1 may have only a lesser role in UC and mixed families, families in which only CD was segregating were selected for this study. Seventy-seven multiplex CD families were recruited through a European consortium using previously published standardised diagnostic criteria based on clinical, radiological, endoscopic and histological findings.[25] Most of them (n=59, 77%) were recruited in the North of France and Belgium. Each European investigator obtained the approval of the relevant ethics committees to participate in the study and written informed consent was obtained from each participating family member.

The 77 selected families contained at least two affected members. Altogether, the 77 families included a total of 261 healthy members and 179 CD patients accounting for 100 independent affected pairs. The affected family members were siblings (87 sibpairs), second-degree-relative (10 pairs), third-degree-relatives (two pairs) or half-sibs (one pair). In the subset of these 77 families, the proportions of patients with blood samples available for their two parents or only one parent were respectively 74% and 22%. For the remaining families, genotyping data from healthy sibs allowed the missing parental information to be inferred. Altogether, identity by descent information was available in nearly all studied families. For association studies, 31 additional simplex CD families were also included providing a total of 108 informative CD families.

Genotyping procedure of the microsatellite and biallelic markers

Twenty-six microsatellite markers (summarized in Table 2) located in the vicinity of D16S411 were selected from public genetic databases (Généthon: http://www.genethon.fr/, CHLC: http://www.chlc.org/, CEPH: http://www.cephb.fr/). Genomic DNAs were prepared from peripheral blood leukocytes[26] from the CD families and were genotyped using fluorescently labelled oligonucleotides, on an ABI 377 DNA Sequencer. Semi-automated fragment sizing was performed using the Genescan^{TM} 2.1 software (ABI) followed by allele identification using the Genotyper^{TM} 2.0 software (ABI). Each genotype was determined independently by two investigators. Conflicting data were either resolved or discarded.

The biallelic marker (hb27g11-f) was also genotyped. Briefly, 100 ng of genomic DNA was amplified by PCR on Thermal Cycler (MJ Research, PTC-200) using 1 μM of forward and reverse primers (hb27g11f-forward: AGCAAATTCTAG-GAGTTATG, hb27g11f-reverse: TGATTTCGTAAGA-

Table 1 Model free linkage analyses localising the IBD1 locus in the pericentromeric region of chromosome 16q

Authors	*Origin*	*Most linked marker*	*P value of the linkage test*	*Number of families (affected relative pairs)*
Hugot *et al*[10]	Europe	D16S411	0.000015	78 (111)
Ohmen *et al*[19]	USA	D16S411	0.02	48 (75)
Parkes *et al*[18]	UK	D16S411	0.002	67 (81)
Cavanaugh *et al*[22]	Australia	D16S409	0.002	54 (74)
Cho *et al*[12]	USA	D16S409	0.008	(72)
Brant *et al*[20]	USA	D16S769	0.007	(147)
Curran *et al*[21]	Europe	D16S411	0.005	134 (167)
Annese *et al*[23]	Italy	D16S419	0.026	58
Hampe *et al*[13]	Europe	D16S411	0.026	268 (239)

Table 2 Sex averaged genetic map of the IBD1 region. The order of microsatellite markers and genetic distances in cM Kosambi were calculated on the genotyping data from the CD families using the Cri-Map program

Markers	*Genetic interval (Kosambi : cM)*	*Genetic distance (Kosambi : cM)*
D16S3120 (Généthon, 1995)	2.9	0.0
D16S298 (GDB, 1992)	0.5	2.9
D16S299 (GDB, 1992)	0.5	3.4
SPN (GDB, 1989)	0.4	3.9
D162383 (GDB, 1993)	0.6	4.3
D16S753 (CHLC, 1995)	0.9	4.9
D16S3044 (Généthon, 1995)	0.0	5.8
D16S3105 (Généthon, 1995)	0.3	5.8
D16S409 (Généthon, 1995)	0.7	6.1
D16S261 (GDB, 1990)	0.1	6.8
D16S540 (CHLC, 1996)	0.1	6.9
D16S3080 (Généthon, 1995)	0.0	7.0
D16S517 (Généthon, 1995)	1.0	7.0
D16S411 (Généthon, 1995)	2.4	8.0
D16S541 (CHLC, 1995)	0.0	10.4
D16S3035 (Généthon, 1995)	1.0	10.4
D16S3136 (Généthon, 1995)	8.1	11.4
D16S3117 (Généthon, 1995)	0.9	11.5
D16S770 (CHLC, 1995)	0.8	12.4
D16S416 (Généthon, 1995)	1.8	13.2
D16S2623 (CHLC, 1995)	1.5	15.0
D16S419 (Généthon, 1995)	3.9	16.5
D16S771 (CHLC, 1995)	1.4	20.4
D16S390 (GDB, 1992)	3.8	21.8
D16S408 (Généthon, 1995)	12.8	25.6
D16S508 (Généthon, 1995)		38.4

CAAGTG). A volume of 10 μl of the PCR products was then digested using 0.067 U/μl of Bsr1 enzyme (NEB). After digestion, the fragments were electrophoresed on a 2% agarose gel and the genotypes deduced from the migration profile.

Genotyping data analyses

Mendelian segregation of the alleles was checked using the Unknown program version 5.23. Marker ordering and inter-marker distances were computed using the CRI-MAP program.[27] Any apparent double recombinants were identified and the corresponding families were genotyped again. The Kosambi map function was retained for map construction and the map computed by the CRI-MAP analysis was subsequently used for the linkage analyses. Nonparametric linkage analyses (NPL-score) were performed using the Genehunter computer program package version 2.0.[28,29]

The transmission disequilibrium test (TDT) and haplotype relative risk (HRR) were calculated using only one affected offspring per pedigree.[30,31] In multiplex families, the proband was thus chosen at random. In order to limit sample biases, 100 random samples were generated and analysed independently. For each studied polymorphic marker, the median *P*-values of the statistic were recorded and this *P*-value was corrected for the number of tested alleles (ie alleles with an observed allele frequency higher than 0.01) using the Bonferroni correction.

Simulation studies

In order to calculate the confidence interval of the IBD1 localization by linkage analysis, simulation studies were carried out. Because no genetic model of inheritance was available, we used parameters compatible with the observed data (location and identity by descent vector at the maximum NPL-score) for simulation studies.

For each microsatellite marker, genotypes were simulated for the set of 77 pedigrees using the observed allele frequency of the microsatellite markers to infer the genotypes of the founders. Non founder genotypes were then simulated according to Mendelian expectations. The affection status of each individual was determined stochastically according to the parameters of the genotype given above. One thousand simulations of the family data were carried out in this manner. For each simulation, the NPL-score was calculated as described in the genotyping data analysis section and the maximum of the score and its location were recorded.

Genomic library screening and contig construction

Five Yeast Artificial Chromosomes (YAC), previously mapped to the pericentromeric region of chromosome 16 were used in order to anchor the contig of Bacterial Artificial Chromosomes (BAC) clones. YAC clone DNAs (825_A_8, 867_H_8, 868_C_11, 841_F_11 and 953_C_10) were prepared in agarose blocks, and further isolated and purified using the GENE-

Table 3 STS markers located on the contig map. The alias numbers allow the identification of the STS on the contig map (Figure 3). The sequence of the primers and the size of the PCR products are indicated for each STS

(a) STS markers mapping on the contig.: microsatellite markers (♣), EST and genes (§), anonymous STS (&).

Alias	*STS name*	*Primer sequence for (5′-3′)*	*Primer sequence rev (5′-3′)*	*Size (bp)*
m5	d16s3035 ♣	AAAGCCAACCTTGCTTCA	TCTTGGAAACAGGTAAGTGC	20–280
m6	d16s3136 ♣	ATTGCCCTCAAGAACAGC	GTGCTATGCCATCCCAG	199–211
m7	d16s541 ♣	CCACACCAGCGTTTTTCTAA	CACACTTTACACACACCTATACC	144–164
m8	d16s3117 ♣	AAGCCATATTAGGTCTGTCCAT	GCTTGGGTTAAATGCGTGT	274–280
m9	d16s416 ♣	AGCAGTTTGGGTAAACATTG	AAATATGCCTTCTGGAGGTG	217–223
m10	d16s770 ♣	GGAGGATCAGGGGAGTTTAT	CAAAGTAAATGAATGTCTACTGC	131–135
m11	d16s2623 ♣	CCAACTCTGTAGTTTCAAAGAGC	TCACAGCCTACTTGCTTGGT	231–243
m13	wi-20979§	TCTCAATGGACTAAAAGAGGCTTT	ATTCGAGCTGGCAGTTCTGT	290
m16	wi-5812 &	ATAAAAATAAAGATTGCTACCCTC	CGAATGATTATTACCTTCATTTAC	174
m18	wi-16305§	TATGCAGCAAATAATTCGTTGG	AAACTGATTTCCTCTGTTTGGC	128
m19	shgc-17274§	GACAGTTTAATGGGATATTCCAGC	CAGTGCCAAGTTTAGTTTGTGC	137
m21	sgc-31023§	TTCAAGGCCACAAAGATGG	TAGAAGAGCAGTGGGGTAGAAT	150
m24	sgc34205§	GATCCAAGGACCCACACTTT	GAGCTGCTGCAGCCATCTT	150
m26	d16s766 ♣	TACCAAGCCTCCACAATTGT	GTTAGGGTTAGCTAGAGAAACG	139–163
m35	wi-1256 &	CAGCCTGTTTTCCTCATCTG	TCAGCACGGTGTGTCCTATA	150
m39	CHLC.ATA2 &	TAAAATTAGAGACAGACACAAGG	GGCAGCCATTTAGCTTCACT	271
m67	wi-9288§	TTAAGAGTACAGCTAAGTGCGGG	ATACACCTGCCCTTCCTCG	108
m113	stSG-32107§	TGTTCCTCAGGGAGCACAG	TCCTGGGTATGATCATGAAGC	134
m118	sgc-32374§	CATTTCTTTCCACAGCAGCA	AAGACCACCACATAGAATACCC	104
m119	stSG-30035§	TCATGAATCACGAACCGTGT	GGCACAGGGGCTTATGTG	169
m134	h61744§	GCAGCATTCTTTCAAAATAGTG	CCATACAAGGGCATCATTC	196
m192	adcy7-ex1§	CCGAGACGTGAAGCCCAG	GTTCCTGTCTGCTTCGTCTC	499
m193	adcy7-chlc§	TAAAGCAAAGCACTGTGCTG	AAGGAAGTCCTTGACCCATG	237

(b) STS markers mapping on the contig.: BAC-end related sequences. STS are identified by the corresponding BAC clone references of the CEPH-BAC libraries

Alias	*BAC-end id.*	*Primer sequence for (5′-3′)*	*Primer sequence rev (5′-3′)*	*Size (bp)*
m42	hb938h2-f	CAGCTCTCATGTAAAACTAAC	GTCCCCTCACCGAAATGG	203
m43	hb938h2-r	TCTCTGGGCTGGGCATCTG	GCCTGTGTTGTCCTCCTG	488
m44	hb233b9-f	GCTTCCTTATGCCTACCTG	AACAGTACAACAGGGTGAG	484
m45	hb233b9-r	CAGAAATAAACGCCAAGGTG	TGTACGCTTCAGAGGACCCC	311
m46	hb27g11-f	TGATTTCGTAAGACAAGTG	AGCAAATTCTAGGAGTTATG	185
m47	hb27g11-r	GCTGAGTGTGTGGATAATG	TGCCTTGGAGCCCCCTTG	465
m48	bb795b2-f	CTCTAGGGAACTCTTTGGG	TCTCTGGCTTCTATGTCTG	195
m49	bb795b2-r	TCTTCCCATGTTCTTATCC	GCAACAGAGCAAGACCC	100
m50	bb234c12-f	AGCATGGATTTATAGTTCGG	CATGGATAACGGGCTGAGG	189
m51	bb234c12-r	GACCGCCGCGTGGAGAGTTG	CCCTTGGGTGAGTGGTGG	130
m52	hb710c6-f	TGGTAATTGCTAAACAGGTC	CAGACCAGAATCCAACCC	562
m53	hb710c6-r	CAAAGCAATGACAAGTTCTCC	TTGAAGACTGGGTGGTGGG	297
m54	bb42b11-f	TAACCTGGCTGTGGGCTTC	AGCCTCGTATAGTCATTCTC	553
m55	bb42b11-r	AGTGACGCCTTTCAGACAG	CACCTACAAGCCAATGGAC	426
m56	bb251a5-f	ACATGCAGGATTAACAACTAC	GTCTTGATATAATACCTGCTG	169
m57	bb251a5-r	CTTACTGAGCCTCCTATGTG	AGAACCAAGGCAATGCTGAC	265
m58	bb761b4-f	CATCAGGACCCCAAACCAG	AGCGTCAGAAACACCTCAG	176
m59	bb761b4-r	TGGTAAAGTTTCAATGGTTGC	ACCTTTCACTGACCCTATGG	179
m60	hb663b6-f	GCTGCCATAGTCATAGAGG	CTTGATGAGGAGGCTGTTG	322
m62	bb598f8-f	CTGTATAAACGAGTTGCTGG	CAGAGTAGTAGGTTGATGAG	235
m63	bb598f8-r	CTTCCAGCTTCTAGTAGACC	TTGGGGTGGCTAGTGACTG	518
m64	hb314b3-f	CATTGTTGGTAAGAGAGTGG	ACTTTGTTAGATCGTGCTC	163
m65	hb314b3-r	TCATCATTTCTTCAAGGTGG	CAGTGAGGAGTCCAAGAG	269
m68	hb87b10-f	TGTGGATTGCTTCTGGGAG	AGCCAATTGGAACCCTCTC	431
m70	hb319a11-r	AGGGAGATTACTGGTTTATG	GATTTACCTGAGACTTCTTG	256
m71	hb319a11-f	TGGCTAGAGCAGGGGGAAC	GATGTTATCCCAAGGAAAGC	293
m74	bb674f3-f	CCAGAACACTCAACTCAAAC	GATAGATTCTCCCTTGCACC	216
m76	hb87f5-f	GTAAAGAGTGAGTGAAGGAG	TACCTGACTGTTGATGCTTG	309
m78	hb276e3-f	AGAGCACAGGCATATTTGAG	TACAGTTGGTTTTCACGAGG	412
m80	hb783e11-f	CAAAATAGCCTGCGTGTAGC	CAAGCCCTATGGTTCTCTG	247
m82	hb133d1-f	TCACTCCCACCACCCTTTC	AGAAGTTTAGTGTGGCGTGG	369
m84	bb471f9-r	CCTGTGGAGTTTATAGTTATC	TCAGGCAGCCAATTCTTAAC	426

Continued

(Continued)

Alias	*BAC-end id.*	*Primer sequence for (5′-3′)*	*Primer sequence rev (5′-3′)*	*Size (bp)*
m86	hb144e3-r	CTCCTTACCCCTGGTTCC	TCTGTTGTGAAAATAAGTCTG	211
m93	hb964d11-f	CCTGAAAATGGTGGAGATGC	AATATTACAGATGTTTGCAGG	372
m97	bb344f8-f	ACGCTTTGCTTCTGCTCTC	TGCAACACTTTACACGTCTC	463
m103	bb877a9-f	CAATGAGGGTGATCTGGTC	ATTGAGTTTTCCCCAGTACAG	535
m105	hb373c6-f	GTTCAGTTGGAGATAGTTGG	CCAGGCACTACAGTCACAG	186
m106	hb373c6-r	CAGAGGCATAGCAAGGAAAAC	CGTTCGGCCTGTTTGCTATG	117
m115	bb1060b2-f	CTGTGGTATAAAAGTTGGGG	TGGGTGCGTGTCCTCTCTC	115
m116	bb1060b2-r	CAGGAGAACTCTAAGGATAC	CCTTCCTGAGCCCCCTG	289
m117	bb230e1-r	TGAGGCAAAACAACTGAGAC	GCAGACATTGATTTTACACAG	109
m120	hb338f7-f	AACAGTCTGAAACAAACAACC	TTCCCACATACTGCTGCTTC	125
m122	bb370f9-f	AGAAAAGCCTCGGTTGAAAG	TGATTCCACAGCTCTTAACC	227
m123	bb370f9-r	TTAGCTGTCTCGCATCTGGG	ACATGGTCCTCAACGTGGTC	113
m124	bb998e2-f	TCAGCCACTACCTCCTTTAC	GATACAAGCAAGGATGAATAC	107
m127	hb395a11-r	TAGATGGAAAGTGATTGGTGT	AAGGGGTCTCCATCCAAGTC	180
m131	bb195h4-f	GATCAGAGCTACTTCCACTAAC	AGTGCTCATCTATCTTTTTCTC	117
m132	bb195h4-r	ATGTGTGTAGGGAAGCGGTG	ATTGTGATTTTTGGATATGGGG	133
m166	bb1118e3-f	TGAACACAAAAGCACACAAAC	AAGCACACATGGAACATTTAC	102
m167	bb1118e3-r	GCCAGGAGTTCAAGACCAG	CACACAGGGATGAAATGGG	264
m168	hb186b8-f	AAGATCGTGCCCTGCGTG	CGAGATTAAGCCAAGCCAAG	107
m169	hb186b8-r	CACGCTGACTTCCTACAAC	CTGGTCCTTTTGCTACTTTTG	71
m170	bb554g11-f	TCTGGCATTTAGGTGTTTGG	ATAAAAGGATGGGAAAGGTGG	356
m171	bb554g11-r	TTAACATTACAATCTTTGAAGG	TTAGACCGAATTTTCACATAC	175
m172	bb920h4-f	TACGTCCAGGCATTGTTATC	TTCCTCCATACTTTCTTCTTC	90
m173	bb920h4-r	ACTTCCAGCATCTCACATCC	CAGCCTCCCCTTTCTCCC	131
m174	hb930c6-f	TAGGGCTTATCTTTCTTTCTCA	CTTCAGCTCCTCAAACTCTC	131
m175	hb930c6-r	AATACATGCCACTTCCCTTCC	AGCACAGGGGTAAGTCAAAG	63
m178	hb20h7-f	TCCATCTGCCTTTACATTATC	GGTGCCACAATCTCCCATC	85
m182	bb797b9-f	GGGCATGATGAGGTTATTAAG	GAAGGCTGTGATACTAAACTG	108
m186	hb690f7-f	TCTCATGGCTATCCTAACTTC	TATTTCTCTATTTCCTTCTTCC	79
m190	bb541e-f	CTCTCATTCTGTCTTGGGTC	TATGTGGGGGGAGAAAAAAAG	205
m196	bb311c4-r	GTATTGCTTATGACATCCTTTG	CATGCAAATTAGAAAAAAGCTG	76
m198	hb364g8-r	GACCTCGTGCTCTTCTCTG	GCATGGAGGAAGAGTGTGG	116
m200	bb45b4-r	GGCCAATGAGGGAAAGCAG	CACAAGGTAGCAGTCACAG	143
m202	bb811a7-r	CCAGCAGCCAAGCCAGAG	GAGTCGGTGGAAGGGAAG	124
m204	bb555f4-r	TCCTGTCTTCTCCTCCATC	TGCACTGTGTTATTGTTGGG	129
m205	hb24e9-f	AAGCAGAAGTCATGGGAAGG	GAGAGTACACTAGAAAATGGG	139
m206	hb24e9-r	CTCAGATAAAGTGTGGATGCT	GAAACCACTGGCCTACCTC	77
m214	bb87h4-f	CATTAAGTTGGATTTTGTCAG	GTGGAAAAAGAAAGGAGTGAG	226
m215	bb87h4-r	GCTCTGCTGTCCAATGTTCC	CTTCATTAGCATCTTCATTTTAC	103
m216	bb916g11-f	AGAAACGCACTGATGAAAGG	ATTTTGCTTCTGCTTGTTCTG	135
m217	bb916g11-r	AACCATGACTCAACTTTCCC	ACCAGATGAGAGATGATGAG	82
m218	bb332b4-f	GTGGAGATGGCTGGAAGG	TGTGGGCAAGAACTGGGG	115
m219	bb332b4-r	GGCCGGAGTCACCAAAAG	TCAAAGCCTGCTCCACCC	62
m241	bb162a9-r	TTCCATCTGGTTTTCTTCTTTC	CTCTGGTATCAAGCATTTCGG	118
m242	bb724g12-f	ATCCTTTGTCCTCCATCCAG	ATTCTGAGGGACTGGCGAC	60
m244	hb406c3-f	ATGGATTTTATGCCCTTCTTG	TCCTGGAGGTTTTGTCTGTG	71
m245	hb406c3-r	TTACCACCTCACAACATATTC	TCTAAACTTGCCAGCCTACC	107

CLEAN® II kit (BIO 101, INC.). The clone inserts were then sized by pulsed-field gel electrophoresis experiments for confirmation of their published sizes.

A contig of BAC clones containing Human genomic DNA fragments was built using the Human BAC library developed at the Jean Dausset-CEPH foundation. This library was derived from the human cell line 1347-02 cloned into vector pBeloBAC11. Two cloning sites, defined by the endonuclease restriction enzyme *Bam*HI and *Hind*III, were used.

BAC clones mapping in the genetic region of interest were identified by an iterative PCR-based screening. At the starting point of the BAC contig construction, the BAC library was screened using the microsatellite markers cited in Table 3a. A search for STSs, genes and ESTs in published genetic maps,[32] provided 24 additional markers expected to map between D16S409 and D16S2623.

Positive BAC clones were then isolated and the ends directly sequenced using the fluorescent big dye terminator TM technology (Perkin Elmer, Applied Biosystems Division). These BAC end-sequences provided additional STSs, from which repeated elements were eliminated using the Repeat-Masker program (http://repeatmasker.genome.washington.edu/RM) and used to screen again the CEPH-BAC libraries. This chromosome walking procedure was repeated until the BAC contig spanned the genetic region from D16S541 to D16S263 without any gaps. The results of the PCR experi-

ments were entered in a local database (OMNIS Package version 7.0.1) and the alignment of all BAC clones was carried out by the Segmap version 3.35 software (UK HMGP resource centre).

BAC clone analyses

For each BAC clone mapping to the contig, a single colony of bacteria was cultured 16 h on LB medium containing 12.5 mg/l chloramphenicol (pH=7.3) at 37°C. DNA was then purified using the Plasmid Maxi kit (QIAGEN) after alkaline lysis and digested by the *Not*I enzyme. The obtained fragments were then run at 15°C for 16 h in a 0.7% agarose gel in 1X TAE buffer with linear ramping of 5 – 15 s at a field intensity of 6 V/cm on a Field Inverted Gel Electrophoresis (Biorad). The insert size was deduced by comparison of the migration profile of the digested DNA with the MidRange II PFG Marker (New England biolabs) molecular weight marker after staining in ethidium bromide.

The inserts were then tested for their marker content by a PCR-based approach using the set of identified genetic markers (see above). PCR was performed in a final volume of 20 μl containing 300 pg of BAC DNA, and 20 pmol of each primer.

In order to confirm the robustness of the PCR-based strategy, the marker content of *Eco*RI restricted DNA from 43 BAC clones was also deduced by Southern blots using radioactive random labelled probes chosen from BAC-end sequences not containing repeats or published STS.

Fluorescence *in situ* hybridisation (FISH)

The chromosome preparation and FISH analysis were performed using classical methods with slight modifications.[33] A 250 ng biotinylated BAC DNA was mixed with 10 μg Cot-1 DNA and 50 μg of sonicated salmon sperm DNA in 10 μl of hybridisation buffer containing formamide (50%), and dextran sulphate (10%) in 2 × SSC. The mixture was denatured at 70°C for 10 min and prehybridised for 1 h at 37°C. Hybridisation was carried out overnight at 37°C with normal human metaphase spreads on a microscope slide individually denatured for 2 min at 70°C in 2 × SSC/70% formamide. Hybridisation was detected using avidin-FITC and the chromosomes were counterstained with DAPI. The metaphase images with G-bands were obtained with Vysis Quips System software related to a CCD camera.

Results

Genetic analyses

Linkage analyses were performed using a total of 77 multiplex CD families genotyped for 26 microsatellite markers. No UC or mixed families were included in our study. The genetic map generated by this genotyping data set spanned 38.4 cM. The first 25 microsatellite markers were mapped to a 25 cM region (Table 2). Thus, the density of polymorphic markers was on average 1 per cM in the genetic region of interest. No major discrepancy between the genetic map drawn from this data set and the published maps was observed (data not shown). However, since the data from our families was more extensive than that previously published, we used the sex-averaged map estimated here in subsequent statistical analyses.

Model-free linkage analyses confirmed that a susceptibility gene for IBD was located in the pericentromeric region of interest (maximum NPL score equal to 3.49, $P=2.37\times10^{-4}$). The maximum NPL-score was observed for a region containing the three microsatellite markers D16S3136, D16S3117 and D16S770 (Figure 1a). The first one (D16S3136) was located 3.4 cM from D16S411. Outside of this region, the NPL score decreased dramatically and the genetic interval exhibiting a NPL score higher than 2.5 was less than 6.6 cM. As expected for a high density set of polymorphic markers, the information content was higher than 90% in this genetic region (Figure 1b). At the maximum of the NPL-score, the identity by descent vector (z0, z1, z2) was equal to (0.19, 0.41, 0.4).

The genotyped family set was heterogeneous. It included 23 sibships that were used in our previous genome scan screening[10] and 54 newly recruited families. Out of these 77 families, six were included in the IBD International Genetics Consortium study.[24] Unfortunately, several of the families that we previously studied could not be genotyped since the corresponding DNA samples were no longer available. Because the family panel studied here was not independent of that used in our original study, the NPL-score was also computed for the 54 newly identified families only. The NPL-score curve showed a similar pattern to that observed for all the 77 families (Figure 1c). However, the maximum NPL-score was higher than for the total family set (NPL=4.67, $P=1.6\times10^{-6}$).

Simulation studies were carried out under the hypothesis of a disease gene mapping between D16S3117 and D16S770 and exhibiting properties compatible with the observed identity by descent vector. For 1000 replicates the maximum NPL-score ranged from 0 to 7.17 (mean: 3.78; 95% CI: 1.58 – 5.88). The location of the maximum NPL-scores varied over nearly the entire genetic region of interest from D16S3120 to D16S408. However, 28% of the positive replicates had the maximum NPL-score in the correct location and 70% of the maximum NPL-scores were located between D16S541 and D16S2623 (Figure 2).

Transmission Disequilibrium Test (TDT)[29] using only one affected member per family was performed for the 26 microsatellite markers. A weak association was observed for an allele of the markers D16S298 (P=0.02), D16S390 (P=0.05), D16S771 (P=0.03), D16S3080 (P=0.02) and D16S3136 (P=0.008). However, after applying the Bonferroni correction for the number of tested alleles, only a weak association between the 207 bp allele of the D16S3136 marker and the CD phenotype was remaining (P=0.05). The Haplotype Relative Risk (HRR) was also positive (P=0.006) for this allele.

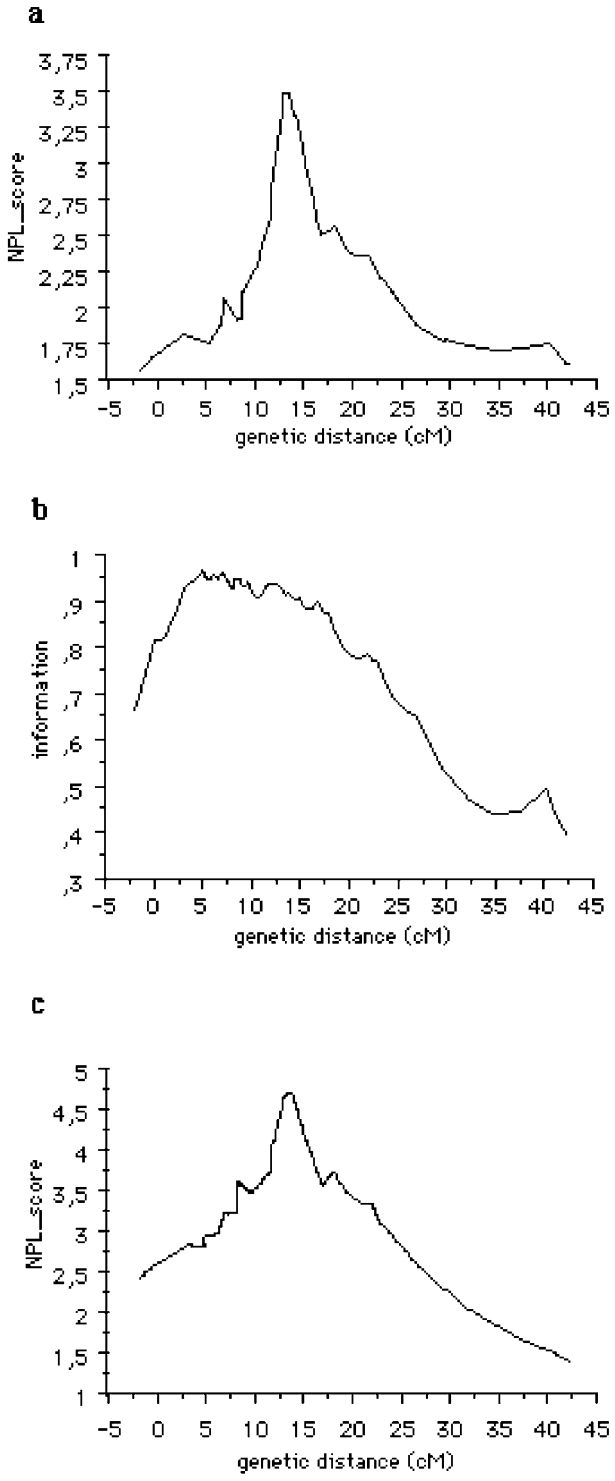

Figure 1 Multipoint linkage analyses. The genetic distances shown on the X-axis were calculated by the Crimap program using the genotyping data of 77 CD families for 26 microsatellite markers. The Multipoint Nonparametric Linkage (NPL) scores and the information content were calculated using the GENEHUNTER PACKAGE version 2.0. (**a**) NPL score computed for the sample of 77 CD families. (**b**) Corresponding information content. (**c**) NPL score obtained with a subgroup of 54 newly identified CD families and independent of the first linkage study.[10]

Since the initial association with D16S3136 was weak, we searched for a SNP in its vicinity. By direct sequencing of 12 unrelated individuals (10 CD patients and two unaffected individuals) we were able to locate a polymorphic sequence corresponding to the BAC end STS, hb27g11f. TDT analysis in CD families showed a significant association between the hb27g11-f marker and CD (P=0.01).

Altogether these results prompted us to further focus on the genetic region centred by D16S3117 and spanning the 4.6 cM from D16S541 to D16S2623. This region was considered small enough to justify construction of a high resolution physical map and linkage disequilibrium studies.

High-resolution physical map

We screened the CEPH-YAC and BAC libraries to identify clones in the region of interest. Initially, 5 YAC clones (825_A_8, 867_H_8, 868_C_11, 841_F_11 and 953_C_10) were identified using seven microsatellite markers (D16S3035 (m5), D16S3136 (m6), D16S541 (m7), D16S3117 (m8), D16S416 (m9), D16S770 (m10)), D16S2623 (m11). For six microsatellite markers (m5, m6, m7, m8, m9, m10) at least one clone was positive. The YAC clones were organized in a contig map of approximately 1.6 Mb (Figure 3). Their respective sizes are indicated in the figure.

Since BAC clones have been reported to be more stable and to contain smaller insert sizes than YAC clones, we also screened the CEPH-BAC libraries. Using the same seven microsatellite markers, we identified 43 BAC clones, which were then used as starting points for chromosomal walks. The contig was further extended by 32 iterative screenings of the BAC libraries using 21 newly identified STS derived from BAC end-sequences (Table 3b and Figure 3) and 11 additional STS markers expected to be mapped to this region (Table 3a).

This approach resulted in a BAC contig map that was 2.5 Mb wide, with a single gap between the m115 and m123 STSs (Figure 3). Considering the sizes of the YAC clones, this gap was estimated to be less than 100 kb. The contig consisted of 99 BAC clones with an average insert size of 130 kb and an average redundancy of 5.3 BACs per genetic locus in this pericentromeric region of chromosome 16.

The contig provided a high-density STS map including 102 STSs (summarised in Table 3). The 99 BAC clones provided 79 new BAC end-related STSs (Table 3b). Seven EST clusters (Unigene cluster: Hs.269950 (m134), Hs.40505 (m21), stSG32107 (m113), Hs.261614 (m119), Hs.92888 (m18), Hs.146128 (m19 and m24), Hs.26295 (m118)), and two known genes (ADCY7 (m192, m67, m193) and KIAA0849 (m13)), were also mapped on the contig (Figure 3). On the other hand, one EST cluster (Hs.65885) and 15 known genes,

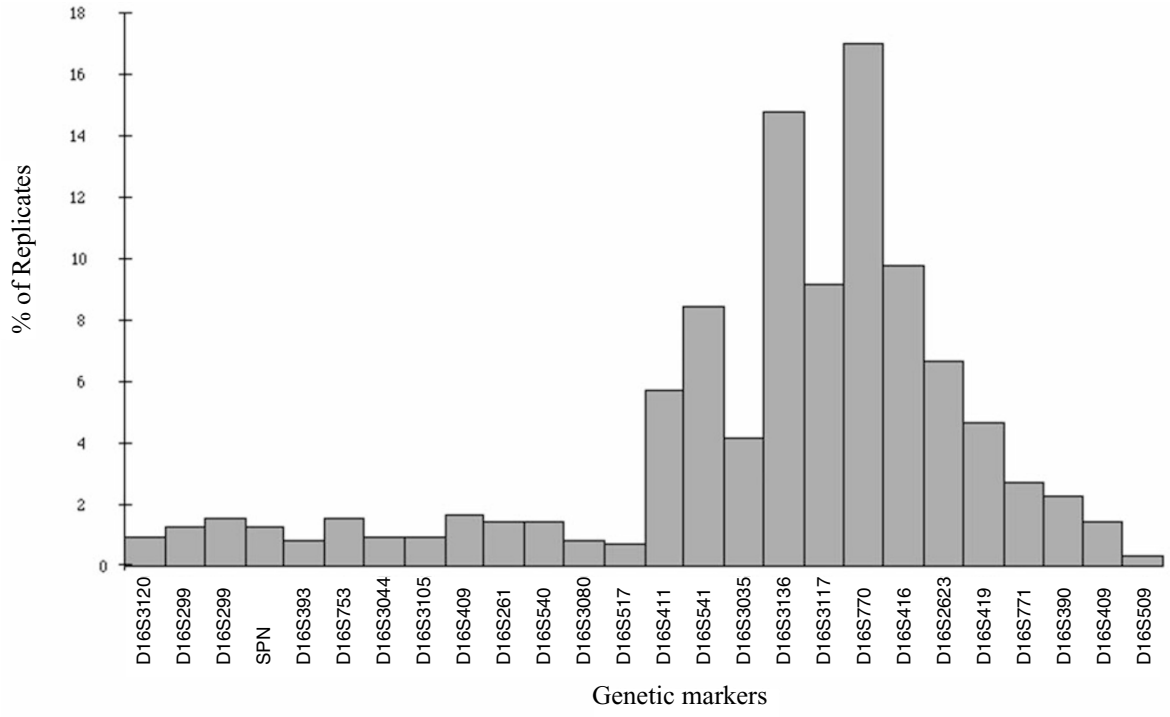

Figure 2 Results of the simulation studies. A disease gene with two alleles (frequency of the disease allele = 0.06), with an incomplete penetrance and mapping between D16S3117 and D16S770 was hypothesised. One thousand replicates of the set of 77 CD families were carried out. For each replicate, the location of the maximum of the NPL score was recorded. The percentages of replicates exhibiting a maximum NPL score at each marker position is reported. Markers are ordered according to the genetic map (Table 2).

(CBLN1 (Hs.662), CLTCL2 (Hs.178710), KIAA1005 (Hs.12328), GNAO1 (Genbank Accession No.: M60162), PHKB (Hs.78060), IL4R (Hs.75545), CES1 (Hs.76688) and CES2 (Hs.174170), MMP2 (Hs.111301), MMP15 (Hs.80343), MT1L (Hs.94360), NME3 (Hs.81687), SCYD1 (Hs.80420), SCYA22 (Hs.97203), and SCYA17 (Hs.66742), were excluded from the YAC/BAC contig map.

Validation of the physical map and FISH mapping

Southern-blot experiments using 59 STS-probes confirmed the efficiency of the PCR based screening approach for contig construction (data not shown). The only discrepancies observed after combining all PCR-STS data from the set of selected BAC clones were for clones hb233c8 and hb218c2. For these two clones, the suspicion of chimerism was confirmed by fluorescent *in situ* hybridisation experiments (data not shown). Four additional BAC clones (bb795b2, hb710c6, bb327a12 and bb1118e3) evenly spread across the contig were also analysed by FISH. They hybridised to the chromosome band 16q12 (Figure 4).

Discussion

The IBD1 gene, involved in the genetic predisposition to CD was previously mapped to the pericentromeric region on chromosome 16.[10] This first localisation has now been confirmed by numerous independent groups and by a large international collaborative study (Table 1). This study provides additional evidence that a CD susceptibility locus is located in this region. Indeed, model free linkage analyses, exhibited a very significant linkage (P=1.6 × 10^{-6}) in a newly recruited panel of 54 CD families.[34] Altogether, the results of the independent linkage studies, including this one, are highly concordant. This fact suggests that the effect of the IBD1 gene is detectable in most Caucasian populations and provides strong evidence that a susceptibility gene to IBD does reside in the pericentromeric region of chromosome 16. This conclusion is also supported by a large international consortium.[24]

Several candidate genes for IBD1 have been mapped to this region. CD19, CD43, IL4R and CD11, which encode for lymphocyte receptors, appeared as good candidates for IBD1, considering their involvement in the immune response. However, previous association studies of these genes and the disease were negative (and unpublished data).[35] Consequently, we decided to pursue the positional cloning strategy towards the identification of the IBD1 gene. Because the involvement of the IBD1 gene in UC and mixed families is still controversial, this study was limited to CD only families. In order to avoid selection biases, we studied jointly the 54

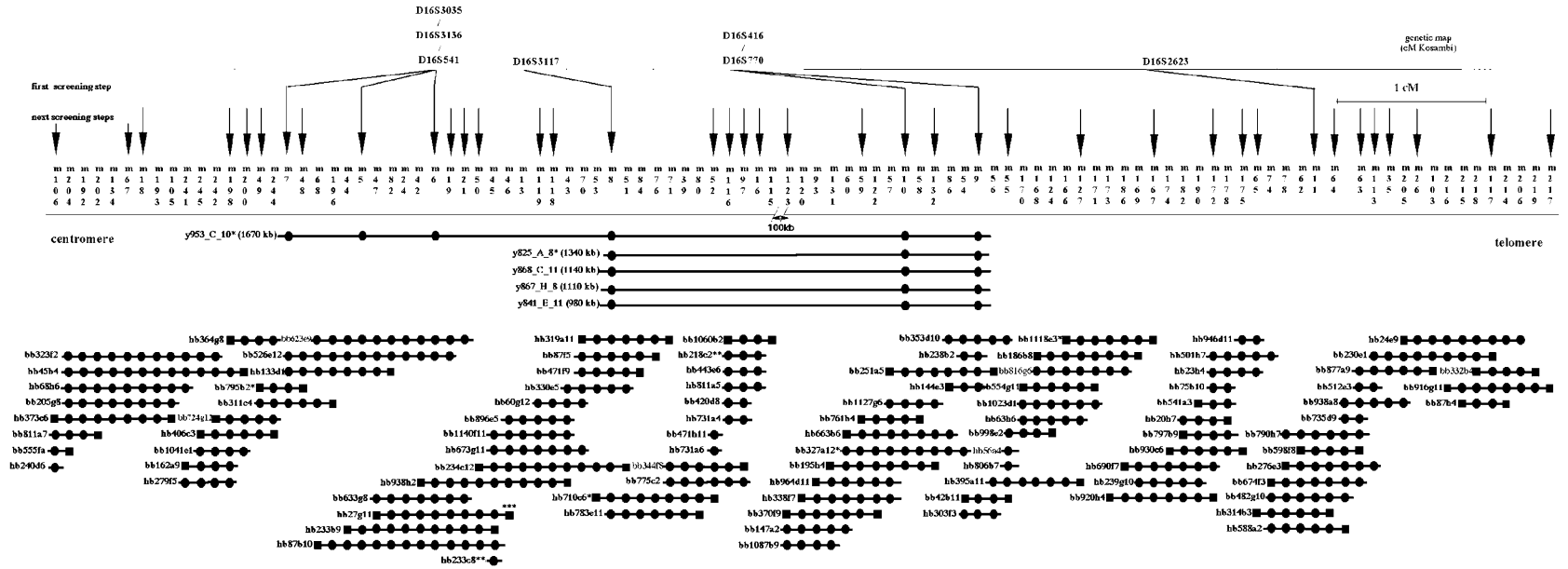

Figure 3 Physical map of the IBD1 region. At the top of the figure, a genetic map is represented. It includes seven microsatellite markers used for the first screening step of the BAC and YAC CEPH libraries. Markers (STSs, ESTs, and microsatellite markers) used for additional screenings the libraries are indicated by vertical arrows. The physical map is indicated by a solid line from centromere to telomere. Localised STSs are arbitrarily equidistant on the figure. They are indicated by anonymous numbers for which correspondence is available on Table 3. The five YAC and 99 BAC clones included in the contig construction are represented by horizontal solid lines and their identification number in the CEPH BAC libraries is reported. YAC clones are indicated by a prefix 'y' and BAC clones by a prefix, 'hb' or 'b'. The sizes of the YAC clones are mentioned between brackets. This contig map spans 2.5 Mb and is defined by 102 STSs. The STS content of a BAC clone was indicated by a circle on the horizontal BAC line under the corresponding STS. BAC-end related sequences are indicated by solid squares on the corresponding sequenced BACs. The BAC clones used in FISH experiments were indicated by a single asterisk for non chimeric clones, and a double asterisk for chimeric ones. The location of hb27g11f marker into hb27g11 BAC-end clone was labelled by a triple asterisk.

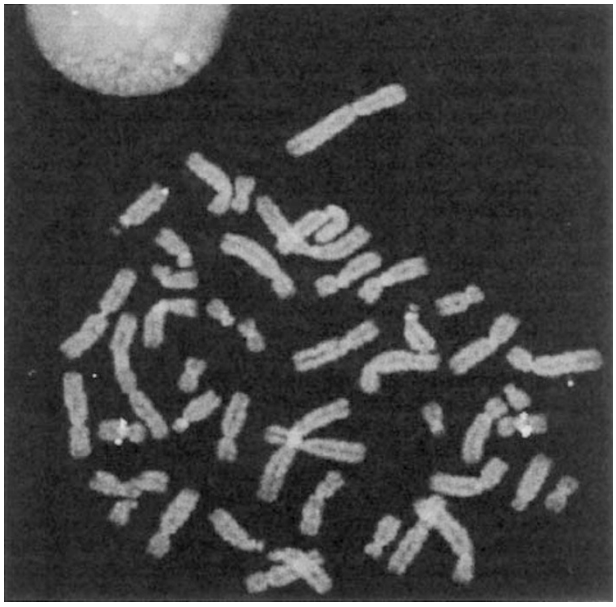

Figure 4 FISH analysis of the BAC clone bb795b2. The hybridisation signal revealed by avidin-FITC confirmed the cytogenetic location of the clone on chromosome band 16q12.

newly recruited families and the 23 available families recruited for the initial study.[10] The 1 cM spacing of the highly polymorphic markers used in this study provided a powerful tool for fine mapping of IBD1, as confirmed by the very high information content estimate. The maximum value of the nonparametric linkage statistic observed in these 77 CD families was located around the marker D16S3117. This marker is close (3.5 cM) to D16S411 which has often been reported as the most linked marker to IBD1 (Table 1). Considering that previous studies were carried out using markers spaced about 10 cM apart, this difference was not unexpected.

The plot of the NPL score exhibited a narrow peak suggesting that the confidence interval of the IBD1 location was small. Indeed, the 1- lod support interval covers a genetic distance of 6.6 cM. Simulation studies also argued for a small confidence interval showing that in 70% of cases, the maximum of the NPL-score was located in a 4.6 cM region around the gene. We judged that this interval was as small as could be expected using this approach and linkage studies were not pursued further.

Considering the uncertainty of the location that is typical of linkage studies when applied to complex genetic disorders, a larger linkage interval was expected. To explain this discrepancy, one can postulate that IBD1 may represent a strong risk factor for the disease. However, this hypothesis is not in accordance with the low calculated relative risk that can be attributed to the IBD1 locus (λs=1.3). A predominant effect of a small number of extended pedigrees linked to the disease could also be proposed, but detailed observation of the individual NPL-scores for the largest pedigrees failed to demonstrate a strong contribution of one or few families to the total NPL score (data not shown). The experimental observation thus remains to be explained.

In order to further characterise the genetic region, we constructed a BAC contig spanning from D16S541 to D16S2623. The reliability of the PCR based method for contig construction was confirmed by Southern blot experiments and the reliability of relative ordering of the BAC clones is supported by a non ambiguous integration in the contig of all the non chimeric clones. Only one small gap (less than 100 kb large) was unresolved in the present BAC contig. This lack of a corresponding clone in the CEPH BAC libraries may be the result of an unclonable sequence or a random under-representation of this sequence in the libraries. Further screening of other host-vector systems may help to resolve this point. The comparison between the genetic (4.6 cM) and physical (2.5 Mb) distances between the two more distant polymorphic markers suggests that this region has a high recombination rate (considering that the average recombination rate for the whole genome is estimated to be 1 cM for 1 Mb). FISH experiments using four independent BAC clones allowed us to map this contig to chromosome band 16q12.

The BAC contig was built using a PCR-based screening of the CEPH-BAC Human libraries. Thus, our data provide a practical evaluation of the quality of these libraries. In our experience, the average size of the clones is 130 kb with a redundancy higher than five genome equivalents. We observed a low rate of chimerical clones (2%) and a good coverage of the genetic region (only one gap 100 kb large in a 2.5 Mb contig). These results suggest that CEPH-BAC libraries are reliable tools for contig construction.

Cliff *et al*[36] have already constructed an initial probe-content BAC map of chromosome 16q providing $>85\%$ coverage of the long arm of this chromosome. Because these data were not available at the starting point of this study, we built the reported BAC contig for the purpose of IBD1 positional cloning. However, this new contig was built using a PCR based screening of independent libraries and thus provides new complementary data that may also contribute to reinforce the mapping projects of chromosome 16.[37,38] In particular, it may help to identify the IBD1 locus and the Blau Syndrome gene.[39]

In order to narrow the IBD1 localisation in this region, we looked for an association using the genotyped microsatellite markers. Because we looked for a linkage disequilibrium in the context of a known linkage, we used only unrelated CD patients (ie one trio per family) in the calculation. A weak positive association was observed for CD with the 207 bp allele of D16S3136 ($P<0.05$). This microsatellite marker is located 0.1 cM apart from D16S3117, corresponding to the maximum of the linkage tests. Considering the number of tested markers (n=26) a false positive result was probable. We thus developed a new SNP (hg27g11-f) in its immediate

vicinity and we were able to confirm a weak linkage disequilibrium using this marker. Thus, altogether, linkage and linkage disequilibrium studies pointed out the same genetic region centered by D16S3136 and D16S3117 and the two overlapping BAC clones hb87b10 and hb27g11.

More recently, going on with the positional cloning approach of IBD1, we were able to confirm that the BAC clone hb87b10 contains not only several additional biallelic markers in linkage disequilibrium with CD but also the IBD1 gene.[40]

Acknowledgements

We gratefully acknowledge the participation of the patients and their families. Families for the study were kindly recruited by Drs Joaquin Balanzo, Vibeke Binder, Bruno Bonaz, Yoram Bouhnik, Guillaume Cadiot, Antoine Cortot, Bart Crusius, Salvatore Cucchiara, Jean-Jacques Delchier, Bernard Duclos, Jean-Louis Dupas, Jean-Paul Galmiche, Jean-Pierre Gendre, Denis Golfain, Christer Gränö, Denis Heresbach, André Lachaux, Hervé Lautraite, Catherine Lenaerts, Eric Lerebours, Victor Levy, Robert Löfberg, Helmut Malchow, Philippe Marteau, Alain Morali, Francesco Pallone, Salvatore Pena, André Rotenberg, Isabelle Rousseau, Jacques Schmitz, Fergus Shanahan, Iradj Sobhani, Hans Svensson, André Van Gossum, Myriam Van Winckel and Michel Veyrac. For their excellent technical assistance we are grateful to: Jean Christophe Beaudoin, Christophe Billon, The Hung Bui, Lucien Cazes, Fabrice Chareyre, Catherine Giudicelli, Laetitia Gressin, Sonia Kemergi, Marie Legrand, Anna Martins, Catherine Massart, Mourad Sahbatou, Emmanuel Tubacher, Héra Der Sarkissian. This project received financial support from: the European Community (Contract Biomed 2 no BMH4-97-2098), the Ministère Français de l'Education Nationale et de la Recherche, the 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 and the Swedish Society of Medicine.

References

1 Podolsky DK: Medical progress: Inflammatory bowel disease. *N Engl J Med* 1991; **325**: 928–1016.
2 Calkins BM, Mendelhoff AI: The epidemiology of idiopathic inflammatory bowel disease. in; Kirsner JB, Shorter RG (eds) Inflammatory bowel disease. Williams & Wilkins, Baltimore, 1995, pp 31–68.
3 Hugot JP, Zouali H, Lesage S *et al*: Etiology of the inflammatory bowel diseases. *Int J Colorectal Dis* 1999; **14**: 2–9.
4 Kyle J, Ewen SW: Crohn's disease in the northeastern and northern Isles of Scotland: an epidemiological review. *Gastroenterology* 1992; **103**: 392–399.
5 Cosnes J, Carbonnel F, Beaugerie L *et al*: Effects of cigarette smoking on the long-term course of Crohn's disease. *Gastroenterology* 1996; **110**: 424–431.
6 Tobin MV, Logan RF, Langman MJ *et al*: Cigarette smoking and inflammatory bowel disease. *Gastroenterology* 1987; **93**: 316–321.
7 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**: 990–999.
8 Orholm M, Binder V, Sorensen IA *et al*: Inflammatory Bowel Disease in a Danish twin register. *Gut* 1996; **39**(Suppl 3): A187.
9 Thompson NP, Driscoll R, Pounder RE *et al*: Genetics versus environment in inflammatory bowel disease: results of a British twin study. *BMJ* 1996; **312**: 95–96.
10 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**: 821–823.
11 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.
12 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**: 7502–7507.
13 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**: 808–816.
14 Hampe J, Shaw SH, Saiz R *et al*: Linkage of inflammatory bowel disease to human chromosome 6p. *Am J Hum Genet* 1999; **65**: 1647–1655.
15 Duerr RH, Barmada MM, Zhang L *et al*: High-Density Genome Scan in Crohn Disease Shows Confirmed Linkage to Chromosome 14q11-12. *Am J Hum Genet* 2000; **66**: 1857–1862.
16 Rioux JD, Silverberg MS, Daly MJ *et al*: Genomewide Search in Canadian Families with Inflammatory Bowel Disease. *Am J Hum Genet* 2000; **66**: 1863–1870.
17 Vermeire S, Satsangi J, Peeters M *et al*: Evidence for inflammatory bowel disease of a susceptibility locus on the X chromosome. *Gastroenterology* 2001; **120**: 834–840.
18 Parkes M, Satsangi J, Lathrop GM *et al*: Susceptibility loci in inflammatory bowel disease. *Lancet* 1996; **348**: 1588.
19 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**: 1679–1683.
20 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**: 1056–1061.
21 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**: 1066–1071.
22 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**: 291–298.
23 Annese 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**: 567–573.
24 The IBD International Genetics Consortium: International collaboration provides convincing linkage replication in complex disease through analysis of a large pooled data set: Crohn disease and chromosome 16. *Am J Hum Genet* 2001; **68**: 1165–1171.
25 Lennard-Jones JE: Classification of inflammatory bowel disease. *Scand J Gastroenterol* 1989; **170**(Suppl): 2–6; discussion 16–19.
26 Miller SA, Dykes DD, Polesky HF: A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Res* 1988; **16**: 1215.
27 Lander E, Green P: Construction of multilocus genetic linkage maps in humans. *Proc Natl Acad Sci USA* 1987; **84**: 2363–2367.
28 Kruglyak L, Daly MJ, Reeve-Daly MP *et al*: Parametric and nonparametric linkage analysis: a unified multipoint approach. *Am J Hum Genet* 1996; **58**: 1347–1363.
29 Kruglyak L, Lander ES: Faster multipoint linkage analysis using Fourier transforms. *J Computat Biol* 1998; **5**: 1–7.
30 Spielman RS, McGinnis RE, Ewens WJ: Transmission test for linkage disequilibrium: the insulin gene region and insulin-dependent diabetes mellitus (IDDM). *Am J Hum Genet* 1993; **52**: 506–516.

31 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 – 787.
32 Deloukas P, Schuler GD, Gyapay G *et al*: A physical map of 30,000 human genes. *Science* 1998; **282**: 744 – 746.
33 Pinkel D, Straume T, Gray JW: Cytogenetic analysis using quantitative, high-sensitivity, fluorescence hybridization. *Proc Natl Acad Sci USA* 1986; **83**: 2934 – 2938.
34 Kruglyak L, Lander ES: High-resolution genetic mapping of complex traits. *Am J Hum Genet* 1995; **56**: 1212 – 1223.
35 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.
36 Cliff SH, Sutherland RD, Jewett PB *et al*: Construction of a BAC contig map of chromosome 16q by two-dimensional overgo hybridization. *Genome Res* 2000; **10**: 714 – 721.
37 Han CS, Sutherland RD, Jewett PB *et al*: Construction of a BAC contig map of chromosome 16q by two-dimensional overgo hybridization. *Genome Res* 2000; **10**: 714 – 721.
38 Cao Y, Kang HL, Xu X *et al*: A 12-Mb complete coverage BAC contig map in Human chromosome 16p13.1-p11.2. *Genome Res* 1999; **9**: 763 – 774.
39 Tromp G, Kuivaniemi H, Raphael S *et al*: Genetic linkage of familial granulomatous inflammatory arthritis, skin rash, and uveitis to chromosome 16. *Am J Hum Genet* 1996; **59**: 1097 – 1107.
40 Hugot JP, Chamaillard M, Zouali H *et al*: Association of NOD2 leucine-rich repeat variants with susceptibility to Crohn's Disease. *Nature* 2001; **411**: 599 – 603.