

## INFLAMMATORY BOWEL DISEASE

# CARD4/NOD1 is not involved in inflammatory bowel disease

H Zouali, S Lesage, F Merlin, J-P Cézard and the EPWG-IBD group, J-F Colombel and the EPIMAD group, J Belaiche and the GETAID group, S Almer, C Tysk, C O'Morain, M Gassull, S Christensen, Y Finkel, R Modigliani, C Gower-Rousseau, J Macry, M Chamailard, G Thomas, J-P Hugot

*Gut* 2003;52:71-74

See end of article for authors' affiliations

Correspondence to: J-P Hugot, Fondation Jean Dausset, 27 rue Juliette Dodu, 75010 Paris, France; jean-pierre.hugot@cephb.fr

Accepted for publication 30 July 2002

**Background and aims:** Inflammatory bowel diseases (IBD), including Crohn's disease (CD) and ulcerative colitis (UC), are complex genetic disorders. CARD15/NOD2, a member of the Ced4 superfamily which includes Apaf-1 and CARD4/NOD1, has recently been associated with genetic predisposition to CD but additional genetic factors remain to be identified. Because CARD4/NOD1 shares many structural and functional similarities with CARD15, we tested its putative role in IBD.

**Patients and methods:** The 11 exons of CARD4 were screened for the presence of variants in 63 unrelated IBD patients. The only non-private genetic variation encoding for a substitution in the peptidic chain was genotyped in 381 IBD families (235 CD, 58 UC, 81 mixed, and seven indeterminate colitis families) using a polymerase chain reaction-restriction fragment length polymorphism procedure. Genotyping data were analysed by the transmission disequilibrium test.

**Results:** Five of nine sequence variations identified in the coding sequence of the gene encoded for non-conservative changes (E266K, D372N, R705Q, T787M, and T787K). Four were present in only one family. The remaining variant (E266K), which exhibited an allele frequency of 0.28, was not associated with CD, UC, or IBD. Furthermore, IBD patients carrying sequence variations in their CARD4 gene had a similar phenotype to those with a normal sequence.

**Conclusion:** Our results suggest that CARD4 does not play a major role in genetic susceptibility to IBD.

Crohn's disease (CD (MIM 266600)) and ulcerative colitis (UC (MIM 191390)) are complex genetic disorders caused by the interplay of genetic and environmental factors.<sup>1</sup> A genetic component to inflammatory bowel disease (IBD) susceptibility was first suggested by the higher concordance rate between monozygotic compared with dizygotic twins.<sup>2</sup> Subsequently, several CD susceptibility loci were localised in the genome (for review see Hugot and colleagues<sup>3</sup>) and more recently CARD15/NOD2 was identified as a susceptibility gene for CD.<sup>3-5</sup>

CARD15 is a member of the Ced4 superfamily, which includes APAF-1<sup>6</sup> and CARD4/NOD1.<sup>7,8</sup> Indeed, CARD4 and CARD15 proteins have a very similar structure characterised by one or two N terminal caspase recruitment domain(s) (CARD), a nucleotide binding domain (NBD), and 10 COOH terminal leucine rich repeats (LRR).<sup>7,8</sup> It has been demonstrated that they both have a role in activation of the nuclear factor  $\kappa$ B pathway and in apoptosis<sup>7,9</sup> induced by exposure to bacterial lipopolysaccharides (LPS). CARD4 is expressed in epithelial cells and has been reported to be involved in the recognition of *Shigella* LPS within the cell cytoplasm.<sup>10</sup> In addition, CARD4 has been mapped to chromosome bands 7p14-p15 (UniGene Cluster Hs 19405), a region which was previously reported to contain an IBD susceptibility locus in British families.<sup>11</sup> Thus for positional and functional reasons, CARD4 appeared to be a good candidate for an IBD susceptibility locus. To test this hypothesis, we screened the CARD4 coding sequence for point mutations and performed transmission disequilibrium analyses in families with IBD segregating

(n=26), Denmark (n=16), Italy (n=10), and Ireland (n=4). Of these 381 families, 306 contained several affected members and 75 were simplex families. Of the 306 multiplex families, 180 contained only CD patients, 45 contained only UC patients, and 81 were mixed families with UC, CD, and indeterminate colitis (IC) patients. In the 75 simplex families, 13 of the probands had UC, 55 had CD, and seven had IC.

Diagnostic criteria were based on clinical, radiological, endoscopic, and histological findings, as previously described.<sup>12</sup> A standardised questionnaire was completed for all patients, including date of birth, sex, family history, age at onset, and details of disease location (at onset and at its maximal severity), granuloma formation, stenosis, transmural involvement, extradigestive symptoms, and therapeutic management. Stenosis and transmural involvement were defined by the occurrence in the digestive tract of at least one stricture or by the presence of a fistula or abscess, respectively, as shown by radiological, endoscopic, or pathological examinations during the evolution of the disease. The study was approved by the relevant ethics committees and informed consent was obtained from all participants.

### Genetic screening of CARD4 gene

A set of 63 unrelated index patients (including 54 CD and nine UC patients) was used for mutation screening. For each of the 63 index patients, the 11 exons and intron-exon boundaries of

## MATERIALS AND METHODS

### Patients and families

A large European consortium recruited 381 IBD families in France (n=246), Sweden (n=45), Belgium (n=34), Spain

**Abbreviations:** IBD, inflammatory bowel disease; CD, Crohn's disease; UC, ulcerative colitis; IC, indeterminate colitis; CARD4, caspase recruitment domain 4; TDT, transmission disequilibrium test. PCR, polymerase chain reaction; NBD, nucleotide binding domain; LRR, leucine rich repeats; LPS, lipopolysaccharide.

**Table 1** Primers used for mutation screening of CARD4 coding sequence and sizes of the corresponding polymerase chain reaction products

Exon	Forward primer	Backward primer	Size (bp)
1	TCATTGATCTCTCAGGGGC	GGTCCACACAATGCCATGC	378
2	CTTATGGTAGGAGGATGCCC	ACTGGTGCCTGCGGTCTTG	271
3	TGATGCTGTTTCTGCCTCTC	AATTGACCCTGCCTCTAG	374
3	TGAGACCATCTTCATCCTGG	CTTCCCACTGAGCAGGTTG	379
3	GTGCCTGACAGCTCCTGC	CAGGGTCATCGTGCAGTCG	372
3	CTCTCTGCTGGATCATCTTC	GTGGAAAACTCATAGGACTG	376
3	GGCTTCCTGCGGGGCTTTG	GGAAAAACAGGTGTGCCAC	392
3	CAGCCCTGAGGAGAAAAGCG	CCACCTGTTGCTCCCTTG	403
4	CATAATGAGTGCCTGCCCTG	GCCATCCGTGCCGATC	235
5	ACATCAGGAGCCAGAAAGTC	CAGAAGGGGGTGATCAAGAG	183
6	AGGAGGGTGACCCACATTG	CCACACACACAGCAGGTTG	199
7	CTTGGGGAGCTAACATTCAC	GAGCACGAATCACCCCTCC	208
8	GCTAAGATTGAATGTAGATGTG	CAAACAACAAATGAAATGACTC	175
9	CGCTCCTGTGAACTCTAAAAC	AAAGTAGCACAGTCTGAAATTG	197
10	TCAAATTTGATCCAGTATTATGG	GTGGTGGTGAGTAAACAGTC	193
11	ACCTGTGTACTTTTGTCTTCC	TTTGCTGCTGAGGCTCCAG	158

The large exon 3 was screened by amplification of six overlapping DNA fragments.

CARD4 (Genbank accession number NM\_006092) were amplified from their DNA by polymerase chain reaction (PCR) using 15 couples of primers (table 1) and sequenced on an ABI 377 automated sequencer using a Dye Terminator Cycle Sequencing Ready reaction kit (Perkin-Elmer Applied Biosystems, Foster City, California, USA). Sequence data were then aligned using the Sequence Navigator analysis software version 1.0.1 (Perkin-Elmer Applied Biosystems) and compared with the previously reported CARD4 sequence.

#### Genotyping methods

Families were then genotyped for the E266K polymorphism. Genomic DNA was amplified by PCR on a thermal cycler (PTC-200; MJ Research, Waltham, Massachusetts, USA) in a 20 µl mixture containing 100 ng of genomic DNA, 1× Taq-Gold polymerase buffer II (Perkin Elmer), 1.5 mM MgCl<sub>2</sub>, 0.2 mM dNTPs, 1 U Taq-Gold polymerase (Perkin Elmer), and 1 µM of forward and reverse primers (line 4, table 1). After denaturation (12 minutes at 95°C), the reaction consisted of 30 PCR cycles (30 seconds at 94°C, 30 seconds at 58°C, and 30 seconds at 72°C) followed by a final extension (seven minutes at 72°C). The E266K mutation was detected by taking advantage of the destruction of an *Ava*I restriction enzyme site by the single base pair substitution and 10 µl of the PCR products were digested with *Ava*I (NEB, Beverly, Massachusetts, USA) according to the manufacturer's recommendations. Genotypes were deduced from the migration profile on a 2% agarose gel:

wild-type DNA was visible as a double band (209 bp and 170 bp) while the mutated DNA was visible as a single 379 bp band.

Finally, the families were analysed for the three main CD associated mutations of CARD15/NOD2, as previously described.<sup>13</sup>

#### Linkage disequilibrium analysis

The transmission disequilibrium test (TDT) statistic<sup>14</sup> was computed using the TDT option of the Genehunter 2.0 package. Phenotype-genotype relationships were analysed by  $\chi^2$  tests (qualitative variables) or ANOVA analyses (quantitative variables).

#### RESULTS

Seven nucleotide changes were observed in exon 3 and two in exon 5 (table 2). A few variants located in introns were also found (data not shown). However, because they were not expected to alter splicing sites, they were assumed to have no functional relevance and were not studied further. Five of the nine variations identified in the coding sequences change the peptide chain of either the NBD (E266K, D372N) or LRR domains (R705Q, T787M, T787K) (table 2). Among these five variations, four amino acid changes (D372N, R705Q, T787M, and T787K) were observed in only one CD patient and were considered private mutations (table 2).

**Table 2** Sequence variations observed in the mutation screening procedure performed in 63 unrelated inflammatory bowel disease (IBD) cases

Location	Nucleotide change	Peptide change	Protein domain	Allele frequency	Heterozygous/homozygous patients	
					CD (n=54)	UC (n=9)
Exon 3	483 C>T	D161D	—	0.28	24/4	4/0
Exon 3	796 G>A	E266K*	NBD	0.28	24/4	4/0
Exon 3	1114 G>A	D372N*	NBD	0.008	1/0	0
Exon 3	1662 G>A	A554A	—	0.07	11/0	0
Exon 3	1722 G>A	A574A	—	0.28	22/4	3/1
Exon 3	2058 C>T	N686N	LRR	0.008	1/0	0
Exon 3	2114 G>A	R705Q	LRR	0.008	1/0	0
Exon 5	2360 C>T	T787M*	LRR	0.008	1/0	0
Exon 5	2360 C>A	T787K*	LRR	0.008	1/0	0

The A of the ATG of the initiator Met codon derived from the sequence published by Bertin and colleagues<sup>8</sup> and Inohara and colleagues<sup>7</sup> was denoted as "nucleotide+1".

\*Non-conservative variants.

CD, Crohn's disease; UC, ulcerative colitis; NBD, nucleotide binding domain; LRR, leucine rich repeats.

**Table 3** Transmission disequilibrium test (TDT) analyses performed in 381 inflammatory bowel disease (IBD) families. Analyses were performed for the ulcerative colitis (UC), Crohn's disease (CD), and IBD (UC, CD, and indeterminate colitis (IC)) phenotypes where appropriate

A allele*	Transmitted	Untransmitted	p Value
All IBD families (n=381)			
IBD phenotype	137	151	0.40
CD phenotype	102	106	0.78
UC phenotype	25	37	0.12
Pure CD families (n=235)			
CD phenotype	85	87	0.87
Pure UC families (n=58)			
UC phenotype	19	21	0.75

\*The A allele of the G796A polymorphism corresponds to the K allele of the E266K variation observed at the protein level.

Three genetic variations (E266K, D161D, and A574A) were frequent enough to perform association studies with reasonable power in our IBD families (table 2). These three polymorphisms were found to be in nearly complete linkage disequilibrium (data not shown). As the E266K variant was the only one to encode a changed protein, altering a glutamic acid residue (E) that is conserved in CARD15, we concentrated our study on this variant.

To investigate the possible role of the E266K variant in IBD predisposition, 381 IBD families were genotyped using a PCR-restriction fragment length polymorphism procedure. The E and K alleles were found to be in Hardy-Weinberg equilibrium. The frequency of the K variant was 0.23 in IBD patients and 0.25 in the panel of alleles not transmitted to the IBD patients used as controls (NS). In families with only CD or UC segregating, the frequency of the K allele was, respectively, 0.27 in CD and 0.22 in UC patients. These frequencies were identical in familial and sporadic cases allowing pooling of the multiplex and simplex families for the genotype-phenotype analyses. TDT analysis failed to demonstrate preferential transmission of the E or K allele for any of the three tested phenotypes (IBD, CD, or UC) (table 3).

A subgroup of 235 unrelated CD patients was randomly selected for more detailed analyses from families with only CD segregating. These patients were subdivided into three groups according to the number of K variant alleles they carried: 0, 1, or 2. No difference was observed between groups for the following variables: sex, age at onset, family history, disease location (at onset and at its maximal severity), granuloma formation, extradigestive symptoms, and therapeutic management. The only observed difference was an excess of the stricturing phenotype in patients carrying one or more K allele(s) ( $p=0.02$ ). However, this was not significant after applying the Bonferroni correction for multiple comparisons. Finally, we examined the relationship between CARD4 and CARD15 in CD patients to identify a putative gene interaction. The frequencies of the three CARD15 CD associated variants (R702W, G908R, and 1007fs) were similar in patients with the wild-type genotype or carrying the K allele of the E266K-CARD4 variation (data not shown).

## DISCUSSION

The recent identification of CARD15/NOD2 as a CD susceptibility gene makes its homologous gene CARD4/NOD1 a potential candidate gene for predisposition to IBD.<sup>3-5</sup> The function of CARD4 has recently been clarified. The gene encodes for a cytosolic protein involved in the pathway by which intracellular LPS activates nuclear factor  $\kappa$ B and JNK.<sup>10</sup>

Screening of all 11 exons and the flanking intronic sequences of CARD4 in 63 patients revealed nine nucleotide changes. Four are silent variations (D161D, A554A, A574A, and N686N) and have no expected functional effect. Five

mutations encoding amino acid changes (E266K, D372N, R705Q, T787M, and T787K) were identified, including three mutations occurring in the LRR domain of the protein, as was observed for the CARD15 mutations<sup>13</sup> in CD. Of these five variations, four occurred in only one patient (D372N R705Q, T787M, and T787K). In these patients, the medical records did not suggest a specific effect of these genetic variants. However, it is not possible to rule out an effect of these variations in the patient carriers without functional studies.

It is unlikely that additional mutations in the coding sequence of CARD4 with a large frequency in IBD patients would have been missed in our study and the E266K variation appears as the only frequent missense variant. However, the possible role of undetected sequence changes in the promoter or regulatory domains of the gene cannot be ruled out by this study and we only tested the hypothesis that, as for CARD15, mutations would be predominantly found in the coding sequence rather than in regulatory regions of the gene.

The G796A variation encodes a non-conservative change (E266K) in the NBD domain. The corresponding glutamic acid residue appears to be conserved in CARD15, suggesting a potential functional effect of the mutation. No transmission distortion of a particular allele of the E266K polymorphism was observed for any of the three tested phenotypes (IBD, CD, and UC). For the CD phenotype, this observation is unlikely to be due to lack of power considering the large number of families studied. In addition, when examining phenotypic subgroups, no phenotype-genotype relationship was evident in CD patients. For the UC phenotype, the power of this study was more limited but the data do not argue for a major role of CARD15 in UC. Finally, no CARD4/CARD15 interaction was identified in CD patients but the power of the statistical test was limited. Taken together, these results argue against a major role of CARD4 in IBD genetic susceptibility.

Considering the strong functional similarities between CARD4 and CARD15, which are both involved in the same pathway by which LPS activates nuclear factor  $\kappa$ B and apoptosis, the lack of involvement of this gene could be seen as surprising. In fact, the main known difference between CARD4 and CARD15 is their tissue specific expression. While CARD4 is widely expressed in epithelial cells, CARD15 is mainly expressed in monocytes.<sup>15</sup> Thus the present observations argue for a pivotal role of the monocyte lineage in the development of CD.

## ACKNOWLEDGEMENTS

We are grateful to the patients and their families for their participation in this study. Drs J Balanzo, B Bonaz, Y Bouhnik, G Cadiot, A Cortot, S Cucchiara, B Crusius, JJ Delchier, B Duclos, JL Dupas, JP Galmiche, JP Gendre, D Golfain, C Grännö, D Heresbach, A Lachaux, H Lutraite, C Lenaerts, E Lerebours, V Levy, R Löfberg, H Malchow, P Marteau, A Morali, F Pallone, S Pena, A Rotenberg, I

Rousseau, J Schmitz, F Shanahan, I Sobhani, H Svensson, A Van Gossum, M Van Winckel, M Veyrac, and gastroenterologists from Nord, Pas de Calais, Somme, and Seine Maritime, generously assisted in the recruitment of families for the study. We are grateful to L Cazes, M Legrand, C Vauray, A Martins, C de Toma, E Tubacher, and M Sahbatou. We thank Dr L Pascoe for critically reading the manuscript. This project received support from the following organisations: European Community, MENRT, INSERM, Direction Générale de la Santé, Association François Aupetit, IRMAD, and Swedish Society of Medicine.

#### Authors' affiliations

**H Zouali, S Lesage, F Merlin, M Chamillard, G Thomas**, Fondation Jean Dausset-CEPH, Paris, France

**J-P Cézard and the EPWG-IBD group**, European Paediatric Working Group of the Genetics of IBD, Department of Paediatric Gastroenterology, Hôpital Robert Debré, Paris, France

**J-F Colombel and the EPIMAD group, C Gower-Rousseau**, Registre EPIMAD, Service d'Épidémiologie et de Santé Publique, Hôpital Calmette, Lille, France

**J Belaiche and the GETAID group**, Groupe d'Études Thérapeutiques des Affections Inflammatoires Digestives, Department of Gastroenterology, CHU de Liège, Belgium

**S Almer**, Division of Gastroenterology and Hepatology, Institutionen för Molekylär och Klinisk Medicin, Linköpings Universitet, Linköping, Sweden

**C Tysk**, Department of Gastroenterology, Örebro Medical Centre Hospital, Örebro, Sweden

**C O'Morain**, Department of Gastroenterology, Adelaide and Meath Hospital, Dublin, Ireland

**M Gassull**, Department of Gastroenterology, Hospital Universitari Germans Trias i Pujol, Badalona, Spain

**S Christensen**, Department of Gastroenterology, Herlev Hospital, Herlev, Denmark

**Y Finkel**, Department of Gastroenterology, Karolinska Children's Hospital, Stockholm, Sweden

**R Modigliani**, Department of Gastroenterology, Hôpital Saint Louis, Paris, France

**J Macry**, INSERM U458, Hôpital Robert Debré, Paris, France

**J-P Hugot**, Fondation Jean Dausset-CEPH, Paris, France, and Department of Paediatric Gastroenterology, Hôpital Robert Debré, Paris, France

#### REFERENCES

- 1 **Hugot JP**, Zouali H, Lesage S, *et al*. Etiology of the inflammatory bowel diseases. *Int J Colorect Dis* 1999;**14**:2-9.
- 2 **Tysk C**, Lindberg E, Jarnerot 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.
- 3 **Hugot JP**, Chamillard M, Zouali H, *et al*. Association of NOD2 leucine-rich repeat variants with susceptibility to Crohn's disease. *Nature* 2001;**411**:599-603.
- 4 **Ogura Y**, Bonen DK, Inohara N, *et al*. A frameshift mutation in Nod2 associated with susceptibility to Crohn's disease. *Nature* 2001;**411**:603-6.
- 5 **Hampe J**, Cuthbert A, Croucher PJ, *et al*. Association between insertion mutation in NOD2 gene and Crohn's disease in German and British populations. *Lancet* 2001;**357**:1925-8.
- 6 **Zou H**, Henzel WJ, Liu X, *et al*. Apaf-1, a human protein homolog to C. elegans CED-4, participates in cytochrome c-dependent activation of caspase-3. *Cell* 1997;**90**:405-13.
- 7 **Inohara N**, Koseki T, del Peso L, *et al*. Nod1, an Apaf-1-like activator of caspase-9 and nuclear factor-kB. *J Biol Chem* 1999;**274**:14560-7.
- 8 **Bertin J**, Nir WJ, Fischer CM, *et al*. Human CARD4 protein is a novel CED-4/Apaf-1 cell death family member that activates NK-kB. *J Biol Chem* 1999;**274**:2955-8.
- 9 **Inohara N**, Ogura Y, Chen FF, *et al*. Human Nod1 confers responsiveness to bacterial lipopolysaccharides. *J Biol Chem* 2001;**276**:2551-4.
- 10 **Girardin SE**, Tournebize R, Mavris M, *et al*. CARD4/NOD1 mediates NF-kB and JNK activation by invasive *Shigella flexneri*. *EMBO Reports* 2001;**2**:736-42.
- 11 **Satsangi J**, Parkes M, Louis E, *et al*. Two stage genome-wide search in inflammatory bowel disease provides evidence for susceptibility loci on chromosome 3, 7 and 12. *Nature Genet* 1996;**14**:199-202.
- 12 **Lennard-Jones JE**. Classification of inflammatory bowel disease. *Scand J Gastroenterol* 1989;**170**(suppl):2-6.
- 13 **Lesage S**, Zouali H, Cézard JP, *et al*. CARD15/NOD2 Mutational analysis and genotype-phenotype correlation in 612 patients with inflammatory bowel disease. *Am J Hum Genet* 2002;**70**:845-57.
- 14 **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**:506-16.
- 15 **Ogura Y**, Inohara N, Benito A, *et al*. Nod2, a Nod1/Apaf-1 family member that is restricted to monocytes and activates NK-kB. *J Biol Chem* 2001;**276**:4812-18.