

ARTICLE

Novel isoforms of the *CARD8* (*TUCAN*) gene evade a nonsense mutation

Richard D Bagnall¹, Roland G Roberts¹, Muddassar M Mirza¹, Toshihiko Torigoe²,
Natalie J Prescott¹ and Christopher G Mathew^{*,1}

¹Department of Medical and Molecular Genetics, King's College London School of Medicine, Guy's Hospital, London, UK; ²Department of Pathology, Sapporo Medical University School of Medicine, Chuo-ku, Sapporo, Japan

CARD8 (*TUCAN*) is implicated in the regulation of apoptosis and inflammation, and is a positional and functional candidate gene for inflammatory bowel disease (IBD). Recent investigations have reported conflicting results of association between a *CARD8* nonsynonymous SNP, rs2043211, and IBD. SNP rs2043211 results in an A>T transversion in the *CARD8* template strand, which introduces a stop codon polymorphism (Cys10Stop), and genotyping of the Cys10Stop variant revealed that 9% of the control population was homozygous for the 'Stop' allele. The effect of the Stop allele on mRNA and protein expression of the two known isoforms of this gene was investigated. IBD patients homozygous for the Stop allele showed somewhat reduced expression of *CARD8* mRNA, but, contrary to expectation, expressed a 48 kDa protein isoform. A search of the EST database and reverse transcription-PCR analysis revealed a novel coding exon and three novel *CARD8* mRNA isoforms that are conserved in primates. The isoforms of *CARD8* differ in their N-termini, resulting in diverse predicted molecular weights (47, 48, 51, 54 and 60 kDa) and multiple outcomes for the variant including Cys10Stop, Cys34Stop, Phe52Ile and Phe102Ile; one isoform may arise through transcription and translation initiated downstream of rs2043211 to yield a novel protein isoform of ~47 kDa. The multiple isoforms and differing consequences for a predicted stop codon polymorphism underline the importance of detailed analysis of the effects of proposed functional variants on gene expression.

European Journal of Human Genetics (2008) 16, 619–625; doi:10.1038/sj.ejhg.5201996; published online 23 January 2008

Keywords: *CARD8*; *TUCAN*; isoforms; nonsense polymorphism; NMD

Introduction

The caspase-associated recruitment domain (CARD) is a protein–protein interaction motif found in a wide range of proteins primarily involved in the regulation of apoptosis and NF- κ B signalling.¹ CARD-containing proteins regulate apoptosis through direct interactions with caspase proteins, and modulate the expression of genes involved in

inflammation through regulation of the NF- κ B pathway. *CARD15*/*NOD2* activates NF- κ B in response to the peptidoglycan component of bacterial cell walls through its ligand muramyl dipeptide, and *CARD15* mutations that disrupt this response are associated with Crohn's disease (CD).^{2–4} Another CARD-containing protein, *CARD8*, also named *TUCAN* (tumour-upregulated CARD-containing antagonist of caspase-9) or *CARDINAL*, is implicated in the regulation of both apoptosis and inflammation, and is known to undergo differential splicing leading to at least two protein isoforms.^{5–8} A 48 kDa isoform (T48) expressed mainly in monocytes, placenta, lymph nodes and spleen⁸ interacts directly with caspase-1 and can induce apoptosis, whereas a larger 54 kDa isoform (T54) is overexpressed in

*Correspondence: Professor CG Mathew, Department of Medical and Molecular Genetics, King's College London School of Medicine, 8th Floor Guy's Tower, Guy's Hospital, London SE1 9RT, UK.
Tel: +44 207 188 3713; Fax: +44 207 188 2585;
E-mail: Christopher.Mathew@Genetics.kcl.ac.uk
Received 31 August 2007; revised 16 November 2007; accepted 5 December 2007; published online 23 January 2008

some cancers and suppresses caspase-mediated apoptosis.⁸ CARD8 also interacts with IKK γ /NEMO, leading to inhibition of NF- κ B activity.⁷ Its gene maps to a region that has shown linkage to inflammatory bowel disease (IBD) in several genome-wide linkage scans,^{9–12} confirmed in a meta-analysis of 10 genome-wide linkage studies.¹³

The CARD8/TUCAN protein is encoded by the 13 exons of the *CARD8* gene on chromosome 19q13. The two known isoforms of the protein have different transcription and translation initiation sites. The open reading frame (ORF) of T48 starts in exon 5 and encodes 432 amino acids, while the ORF of T54 starts in exon 4 and encodes 487 amino acids (see Figure 1). The N-termini of the resulting polypeptide sequences differ up to exon 7, after which both proteins have the same ORF and share the remaining 398 amino acids. The underlying cause of these differences is that the T48 transcript lacks the frame-shifting coding exon 6 of T54 (see Figure 1). The C-terminal CARD domain is present in both isoforms.

An A>T transversion polymorphism (rs2043211 on the template strand) introduces a Cys>Stop at codon 10 of T48 (Cys10Stop), and recent investigations have provided conflicting evidence for an association with IBD. Association of the common 'Cys' allele was reported in Oxfordshire CD cases (359 controls, 35.6%; 366 CD, 29.1%, $P=0.0083$; 373 UC, 31.4%, $P=0.86$; 739 IBD, 30.2%, $P=0.12$).¹⁴ However, a recent British IBD case-control study found a highly significant association of the rare 'Stop' allele with IBD (1381 controls, 29.8%; 1304 CD, 34.0%, $P=0.0011$; 799 UC, 34.0%, $P=0.0046$; 2103 IBD, 34.0%, $P=0.0003$).¹⁵ Moreover, a combined study of German and Norwegian IBD patients found no significant association between the risk allele at rs2043211 and risk for CD or UC¹⁶ and no evidence for association of either allele of this SNP was found in a recent whole genome association scan of 1744 CD cases (MAF=0.323) and 2933 controls (MAF=0.318).¹⁷

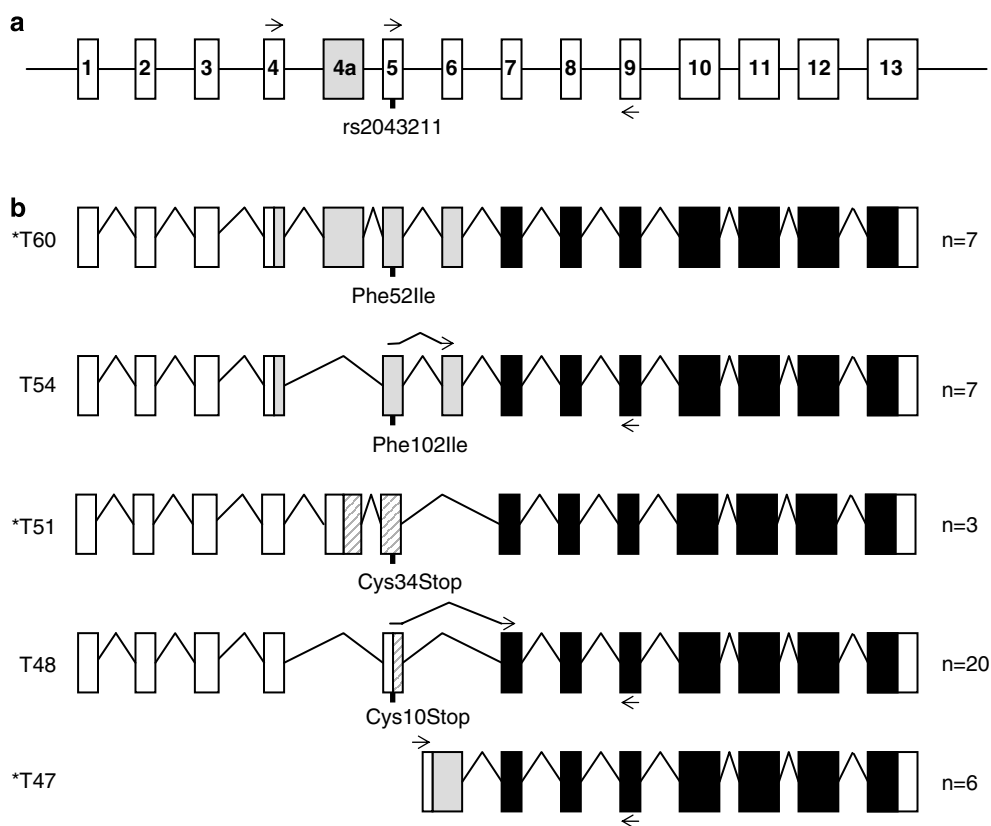


Figure 1 Structure of *CARD8* (*TUCAN*) gene and mRNA isoforms. (a) Genomic exon structure (not to scale). The novel alternatively spliced exon 4a is shown in grey. The position of SNP rs2043211 is indicated underneath. Horizontal arrows indicate RT-PCR primers used for amplifying *TUCAN* isoforms. The left primer in exon 5 (*TUCAN-EX5-F*) distinguishes isoforms with or without exon 6 only whereas the left primer in exon 4 (*TUCAN-EX4-F*) also distinguishes isoforms with or without the novel exon 4a. (b) *CARD8* mRNA isoforms. The left column shows isoform name with number of fully informative dbEST entries on the right. Asterisks indicate novel isoforms described for the first time in this work. 5' and 3' untranslated regions are shown as white boxes. Isoform-specific N-terminal coding regions (grey and hatched boxes indicate two different reading frames) and common C-terminal coding regions (black boxes) are shown. Horizontal arrows indicate isoform-specific RT-PCR forward primers. The outcome of the SNP rs2043211 with the nature of amino-acid substitution is indicated below each isoform.

The finding that 9% of the control population was homozygous for the Stop allele and would therefore be predicted to lack a functional T48 protein was unexpected. This, together with the evidence for the T54 isoform, led us to a more detailed investigation of the expression of *CARD8*, which revealed a series of isoforms that impact on the functional consequences of the rs2043211 SNP.

Materials and methods

Reverse transcription PCR

Total RNA was isolated from lymphoblastoid cell lines of CD patients of known rs2043211 genotype, monocyte cell line THP1 and fresh peripheral blood lymphocytes from normal individuals, using RNA Isolator (Genosys, Cambridge, UK). Total RNA from ascending colon, duodenum and adult breast tissue was purchased from Stratagene (La Jolla, CA, USA). All RNA samples were quantified using an Agilent 2100 Bioanalyzer. Five hundred nanograms of RNA was heat denatured at 70°C for 5 min with 0.5 µg oligo-dT primer (Invitrogen, Paisley, UK) in a final volume of 20 µl and then chilled on ice. Reverse transcription was performed at 42°C for 1 h using 50 U M-MLV reverse transcriptase (Ambion, Huntingdon, UK), 3 µl 10 × first strand buffer, 10 U RNase inhibitor (Promega, Southampton, UK) and 0.5 mM each dNTP in a final volume of 30 µl. Five microlitre aliquots of cDNA were PCR amplified using 5 U AmpliTaq Gold[®] DNA polymerase (Applied Biosystems, Warrington, UK), 2 µl 10 × reaction buffer, 1.5 mM MgCl₂, 0.5 mM each dNTP and 40 ng each oligonucleotide primer in a final volume of 20 µl. Primers for reverse transcription (RT)-PCR amplification are shown in Supplementary Table 1. Reactions were heat denatured at 95°C for 10 min followed by 40 cycles of 95°C for 30 s, 65°C for 30 s, 72°C for 1 min.

DNA sequencing

Gel-purified RT-PCR products were sequenced using the BigDye v3.1 dye terminator kit (Applied Biosystems) according to the manufacturer's instructions, and analysed on an ABI 3730XL automated DNA sequencer.

Western blots

Frozen or fresh cultured cell pellets were lysed on ice in 20 mM Na₃PO₄ (pH 7.4), 1% SDS and Complete[™] protease inhibitor cocktail (Roche Diagnostics, Basel, Switzerland) and clarified at 13 000 r.p.m. for 10 min. Protein concentrations were determined using a BCA Protein Assay Kit (Perbio Science, Cramlington, UK). Fifty micrograms of protein was boiled in 80 mM Tris-HCl (pH 8.0), 4% glycerol, 0.2% SDS, 25% 2-mercaptoethanol and 0.2 mg/ml bromophenol blue and resolved through a 12.5% SDS-PAGE gel for 1.5 h at 150 V, then transferred to Hybond-ECL (Amersham Biosciences, Buckinghamshire, UK). The membrane was incubated at room temperature in PBS

containing 5% non-fat milk for 1 h, followed by PBS containing 5% non-fat milk and rabbit anti-TUCAN (*CARD8*) antibody at 4°C overnight. The TUCAN/*CARD8* antibody was raised against residues 99–115 of T54 (encoded by part of exon 7).⁸ The membrane was washed in PBS and incubated for 1 h at 4°C in PBS containing 5% non-fat milk and anti-mouse IgG antibody (Dako, Stockholm, Sweden), washed in PBS and incubated with the ECL detection kit (Amersham Biosciences). Chemiluminescence was visualised using Hyperfilm ECL (Amersham Biosciences).

Results

The rs2043211 A>T transversion in exon 5 of *CARD8* results in a predicted Cys>Stop at codon 10 in T48 and a Phe>Ile amino-acid substitution at codon 52 in T54. These two outcomes are likely to have different effects on the stability and structure of the T48 and T54 mRNA isoforms, as premature termination codons can target mRNAs for degradation via nonsense-mediated decay¹⁸ and can cause skipping of the exon containing the nonsense mutation.¹⁹ We therefore used the primer pair TUCAN-EX5-F in exon 5 and TUCAN-EX9-R in exon 9 (Figure 1), which direct amplification of a 531 and 571 bp product from T48 and T54 cDNA, respectively (the size difference reflects the exclusion or inclusion of exon 6), to investigate the expression of both TUCAN mRNA isoforms in different tissues, and in lymphoblastoid cell lines of CD patients who were homozygous for the presence or absence of the Stop allele or heterozygous. T48 and T54 mRNAs were expressed in all tissues examined (Figure 2) and the expression of T48 and T54 in CD patients homozygous for the Stop allele (TT; 939 and 090) could be seen in patients homozygous for the Cys allele (AA; 388 and 435).

The expression of individual mRNA isoforms was further investigated using the primer TUCAN-EX9-R in exon 9 with the forward primers TUCAN-54F or TUCAN-48F, which, by traversing exon boundaries, specifically amplify T54 and T48 cDNA, respectively (Figure 1). In keeping with our previous results, expression of T54 mRNA was similar in CD patients, irrespective of rs2043211 genotype status (Figure 2). Expression of the T48 isoform was somewhat lower relative to GAPDH in patients homozygous for the Stop allele (Figure 2). This suggests that this stop codon, contrary to expectations, may not induce significant nonsense-mediated decay in T48 transcripts, although the RT-PCR technique used here is at best semi-quantitative and, therefore, this observation requires further confirmation.

We also investigated the possibility that exon 5, which contains rs2043211, might be skipped in some T54 transcripts, as this could generate an in-frame mRNA and translation product of ~48 kDa. Amplification of T54 and T48 mRNA using a forward primer in exon 4 and a reverse

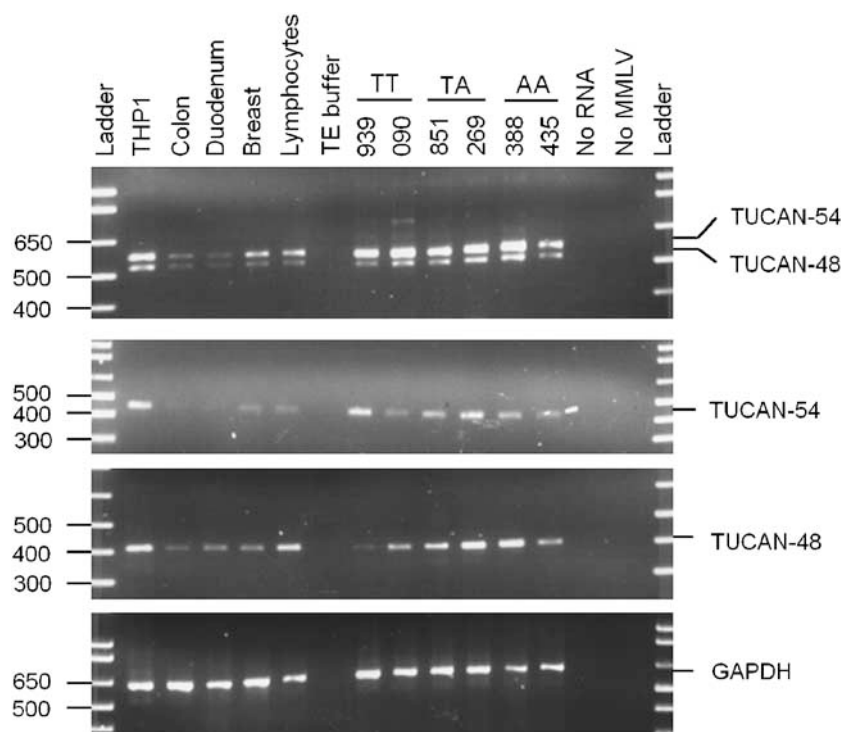


Figure 2 RT-PCR amplification of T54 and T48 isoforms of the *CARD8/TUCAN* gene. Panels show RT-PCR amplification of both T54 (upper band) and T48 (lower band) using primer pair TUCAN-EX5-F/TUCAN-EX9-R, T54 alone using primer pair TUCAN-54F/TUCAN-EX9-R, T48 alone using primer pair TUCAN-48F/TUCAN-EX9-R, and GAPDH as a loading control. The RNA template is from various tissues and from lymphoblastoid cell lines of CD patients homozygous (939 and 090; TT) or heterozygous (851 and 269; AT) for the 'Stop' allele or homozygous for the common 'Cys' allele (388 and 435; AA). Genotypes at rs2043211 are shown above and relevant size markers on the left. Lanes 'TE buffer', 'No RNA' and 'No MMLV' are negative control samples.

primer in exon 9 failed to yield any transcripts lacking exon 5, regardless of genotype, indicating that rs2043211 does not lead to exon skipping (results not shown).

The two known transcripts of *CARD8*, together with western blots showing its encoded protein migrating as a broad band or as a doublet,⁶ provide evidence of multiple protein isoforms. To assess the full isoform diversity of the N-terminus of *CARD8*, we analysed the structure of sequences in the human section of the expressed sequence tag (EST) database (dbEST, October 2006) identified by a BLASTN search using exons 4–7 of T54. Of the 69 cognate ESTs, 43 were fully informative for transcript structure in this region (Supplementary Table 2). Seven ESTs retain exon 6 and correspond to T54, whereas 20 ESTs skip exon 6 and correspond to T48. Interestingly, a further seven and three ESTs resemble T54 and T48 transcripts, respectively, but contain an additional 150bp exon between exons 4 and 5 (AC011466: 155973–156122), which is not described by either Yamamoto *et al*⁸ or Razmara *et al*.²⁰ The ORFs and splice sites are conserved in chimpanzee, orangutan and macaque, and this exon appears in the single orangutan (CR547228) and macaque (CJ459175) ESTs. As confirmation, RT-PCR amplification between exons 4 (TUCAN-EX4-F) and 9 (TUCAN-EX9-R) amplified the

expected 590 and 631bp products of T48 and T54, respectively, and further 740 and 781bp products (see Figure 3a), which we found by sequencing to correspond to T48 and T54 transcripts containing this additional exon. This would insert 50 amino acids (~6kDa) after residue 21 in the canonical T54 protein, yielding a novel protein, which we call T60 (EU118120), and would add a potential in-frame translation initiation codon 24 amino-acid residues upstream of the canonical T48 start codon to produce a protein of 51 kDa, which we call T51 (EU118121) (Figure 1). We have denoted this new exon as 'exon 4a' to maintain the published exon numbering.

Further examination of the human ESTs yielded the surprising result that six ESTs start within intron 5 (BF091373, BE168871, BE168868, BE168799, BI827664 and BG118619). These transcripts could arise from a promoter within this 1.9-kb intron, with the longest EST (BI827664) starting ~400bp upstream of exon 6. A search for human promoter sequences in intron 5 using the web-based 'core promoter' program (<http://rulai.cshl.edu/tools/genefinder/CPROMOTER/human.htm>) revealed a putative transcription start site 375bp upstream of exon 6. There is a potential translation initiation codon 20bp upstream of exon 6, which is in a good Kozak context in-frame with the

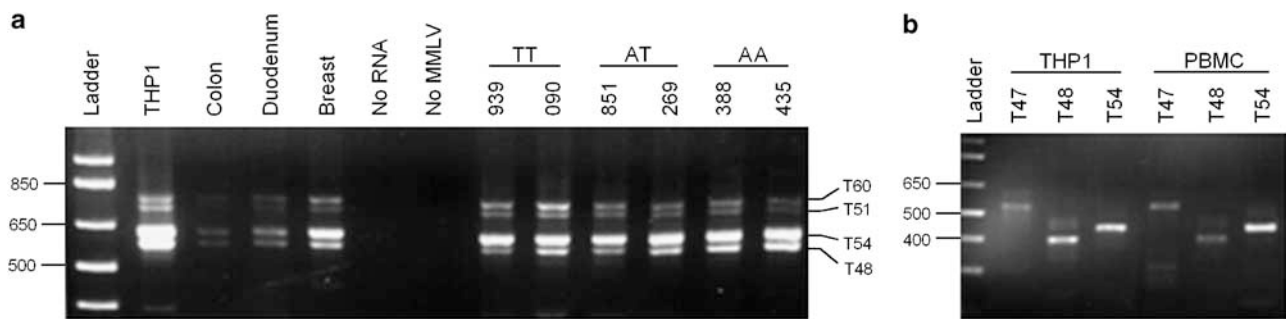


Figure 3 Comparative RT-PCR of all five *CARD8* transcript isoforms. (a) Amplification of exons 4–9 showing products of 781 bp (T60), 740 bp (T51), 631 bp (T54) and 590 bp (T48). Samples as in Figure 2. (b) Independent amplification of T47, T48 and T54 mRNA from THP1 and peripheral blood mononuclear cells using primer pairs TUCAN-47F/TUCAN-EX9-R, TUCAN-48F/TUCAN-EX9-R and TUCAN-54F/TUCAN-EX9-R, respectively.

T54 and T60 ORF; this is conserved in chimpanzee, orangutan, rhesus monkey and marmoset. This would provide seven N-terminal amino acids before joining the exon 6 sequence to generate a putative protein isoform of 47.5 kDa, which we call T47 (EU118122) (Figure 1, Supplementary Table 2). T47 would differ from T48 only at its N-terminus, where the first 25 amino acids of T48 would be replaced by 20 new ones in T47 (Supplementary Figure 1). A conserved T47 cDNA has been isolated from the orangutan (CR858904) and has been predicted from the chimpanzee *CARD8* genomic sequence (XP_001170446). We sought to confirm the existence of the T47 transcript in humans by RT-PCR using a forward primer in intron 5 (TUCAN-47F), 104 bp upstream of exon 6, paired with a reverse primer in exon 9 (TUCAN-EX9-R); this amplified a 547 bp product (Figure 3b) and sequencing confirmed the expected sequence (data not shown). Thus, the prediction from the EST population alone is that T48 should be the most prevalent isoform, with less prevalent isoforms of T60, T54, T51 and T47 (Figure 1, Supplementary Table 2).

The various isoforms of *CARD8* differ only in their N-termini, and thus the rs2043211 SNP would cause a different outcome in each isoform; Cys10Stop in T48, Cys34Stop in T51, Phe52Ile in T54, Phe102Ile in T60 and, as transcription of T47 starts downstream of rs2043211, the sequence of the T47 isoform would be unaffected (Figure 1, Supplementary Figure 1).

In keeping with our analysis of the EST database, western analysis has previously shown that the predominant *CARD8* isoform expressed in normal cells is ~48 kDa.⁷ The Cys>Stop polymorphism is predicted to generate a severely truncated protein comprising the first nine amino acids of T48 followed by a translation stop codon; thus, individuals homozygous for the Stop allele would be expected to lack the T48 isoform. We investigated expression of the protein in lymphoblastoid cell lines of CD patients of the three possible Cys10Stop genotypes (AA, AT and TT or Cys/Cys, 'Cys/Stop' and 'Stop/Stop' respectively), using a polyclonal antibody raised against a peptide

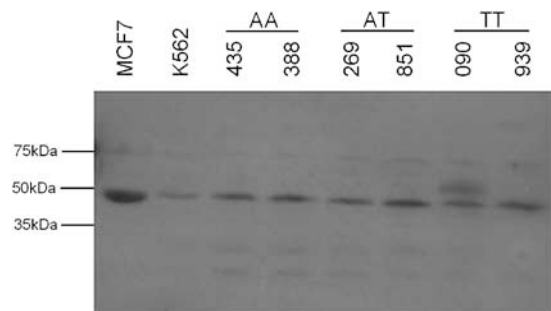


Figure 4 Effect of rs2043211 genotype on *CARD8* protein isoforms. Western analysis of *CARD8* in CD patients of various genotypes for the rs2043211/Cys10Stop variant. Genotypes at rs2043211 are shown above patient samples with relevant size markers on the left.

sequence common to all isoforms, encoded by part of exon 7. The MCF7 and K562 cell lines, which are AA (Cys/Cys) and AT (Cys/Stop) for rs2043211, respectively, were used as positive controls for the presence of T48, as both have previously been shown to express T48, with the former showing particularly high level expression.⁶ The cell lines MCF7 and K562, and CD patients who were also AA (Cys/Cys) or AT (Cys/Stop) at rs2043211, revealed an immunoreactive band of ~48 kDa; however, contrary to our initial expectations, so did patients who were homozygous for the Stop allele (Figure 4). In the light of the mRNA isoforms described above, this band may instead represent T47, which is only five amino acids shorter than the canonical T48 isoform and is unlikely to be resolved from it on the western blot. Interestingly, the lymphoblastoid cell line of one patient homozygous for the Stop allele showed an additional band of ~54 kDa. A formal possibility remains that skipping of downstream in-frame exons 10 and 12 from T54 transcripts could generate a protein of ~48 kDa. This was investigated by sequencing cDNA amplification products from exons 9 to 13, but no skipping of these downstream exons was observed.

Discussion

Our investigation of the EST database together with confirmatory RT-PCR analysis has revealed a novel coding exon and three novel mRNA isoforms of *CARD8*, which, like the two previously described isoforms, differ in their N-terminal residues but contain an identical C-terminal CARD domain. SNP rs2043211, for which there is conflicting evidence for an association with IBD, has different outcomes in the various isoforms, resulting in either a nonsense polymorphism or amino-acid substitution within the variable N-terminal regions. The variant outcomes affect not only the sequence of translated amino acids but also the stability of the mRNA isoforms. In keeping with previous observations, RT-PCR amplification in various tissues showed that the two known *CARD8* mRNA isoforms are widely expressed in adult human tissues, and there was a slightly reduced level of T48 transcripts that contained the Stop variant, indicating a degree of nonsense-mediated decay of these mRNAs. However, our finding of the apparently normal levels of an ~48 kDa protein in CD patients who are homozygous for the Stop allele was unexpected. One possible explanation for this would be skipping of in-frame exons; however, we found no evidence for such transcripts. Another explanation would be initiation of translation from an alternative site downstream of rs2043211, but no such candidate AUG codons are apparent. A possible solution to this was provided by a detailed examination of the human EST database, which revealed six ESTs of an isoform with an initiation codon in intron 5, which would encode a 47.5 kDa protein with a unique N-terminal sequence; as it is initiated downstream of exon 5 in which the rs2043211 is located, it should not be affected by the presence of this variant. In the T54 isoform, the polymorphism would result in an amino-acid substitution (Phe > Ile) rather than a stop codon; although we detected expression of T54 mRNA in a range of normal tissues, western blot analysis revealed that, with the possible exception of one CD patient (090) who was homozygous for the Stop allele, the T54 protein isoform does not appear to be translated in the cell lines we tested, and is not generally expressed in normal tissues. Additional faint immunoreactive bands seen on the western blot may result from other isoforms of the protein, such as the inclusion of exon 4a, which would generate an ~60 kDa protein when initiated from the T54 start codon in exon 4, or a 51 kDa protein when initiated from a start codon in exon 4a. In summary, the expected consequences of homozygosity for the Stop allele would therefore be loss of the T48 and T51 isoforms, expression of a mutated form of T54 and T60, but normal expression of T47 (Supplementary Figure 1).

Alternative splicing of mRNA can yield translation products that have different functions or localise to different cellular compartments. The two known and three novel protein isoforms of *CARD8* described here differ

from each other in their N-termini. Furthermore, as it is the N-terminal 320 amino acids of T48 that have been shown to be essential for inhibition of NF- κ B signalling,⁷ and as T48 and T54 have antagonistic effects on apoptosis,⁷ we might expect all five isoforms to differ in their effects on the regulation of the immune response. In addition, recent immuno-histochemical analysis of *CARD8* in non-small-cell lung cancer tissue revealed three patterns of localisation: exclusive cytoplasmic or nuclear localisation, or both cytoplasmic and nuclear localisation.²¹ The antibody in this study was raised against a sequence that is present in all five isoforms of *CARD8*, and it is therefore possible that the antibody may be detecting different isoforms in different cellular compartments.

The evolution of *CARD8* is obscure. It is a paralogue of the C-terminal ~400 amino acids of *NALP1* (ie including the CARD domain but not the pyrin, NACHT or LRR regions of *NALP1*); indeed exons 7–13 of all extant *CARD8* genes share an identical gene structure with the corresponding region of *NALP1* genes. Significant amino-acid homology to *NALP1* commences two-thirds of the way through the exon 7 region of *CARD8* and continues to its C-terminus. Most mammalian genome sequences examined (including the noneutherian mammals opossum and platypus) have a clearly defined *NALP1* gene. *CARD8*, on the other hand, while present in primates, carnivores, xenarthrans and afrotherians, is missing from the completed genome sequences of rodents, cetartiodactyls and marsupials. In birds and amphibians a *CARD8*-like gene is observed, but no *NALP1* orthologue is seen. This suggests that *NALP1* arose from a fusion event between ancestral *CARD8* and non-CARD-containing *NALP* genes, and that *CARD8* genes were later lost on several occasions. A corollary of this complex and dynamic history is that sequences upstream of *CARD8* exon 7 are poorly constrained and hard to identify. Dotplots reveal that although the introns and splice sites of *CARD8* exons 4, 4a, 5 and 6 are conserved in marmoset and dog (for example), the exonic sequences themselves are barely recognisable, only their size and reading frame being conserved (Supplementary Figure 1). Findings in elephant and armadillo are very similar. Combined with other data, the most parsimonious assumption is that the long T54/T60 isoforms are ancestral but that functional constraint on amino-acid sequence upstream of exon 7 is either weak or subject to positive selection. Either way, it is clear that many mammalian species (including mouse, rat and cow) can function perfectly well without *CARD8* genes.

In conclusion, this study demonstrates that individuals homozygous for a premature translation stop codon apparently affecting the 48 kDa major isoform of *CARD8* nevertheless retain an ~48 kDa immunoreactive protein. In fact, 9% of the general population are homozygous for what appears to be a loss-of-function polymorphism in a gene that is likely to have a significant role in the immune

response and apoptosis. This may reflect either some degree of redundancy in the functional pathways (notably, rodents do not have a *CARD8* gene) or partial rescue of *CARD8* function via alternative initiation of translation or splicing, leading to a functionally compromised but near full-length protein. This would be consistent with the presence of the ~48 kDa protein that we observed in apparent homozygotes for the Stop allele (TT). Our bioinformatic analysis and expression studies of the *CARD8* gene have provided evidence for the existence of three additional mRNA isoforms; T47, T51 and T60. The functional role of various isoforms of *CARD8* is not yet clear, and warrants further investigation, as does the possible role of the Cys>Stop polymorphism. Finally, this study shows that a detailed characterisation of the effect of disease-associated sequence variants on gene expression can reveal unexpected outcomes that may be highly relevant to their contribution to pathogenesis.

Acknowledgements

This work was supported by the National Association for Colitis and Crohn's Disease (UK), the Guy's and St Thomas' Charity and The Wellcome Trust.

References

- Hong GS, Jung YK: Caspase recruitment domain (CARD) as a bi-functional switch of caspase regulation and NF-kappaB signals. *J Biochem Mol Biol* 2002; **35**: 19–23.
- 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.
- Ogura Y, Bonen DK, Inohara N *et al*: A frameshift mutation in NOD2 associated with susceptibility to Crohn's disease. *Nature* 2001; **411**: 603–606.
- 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**: 1925–1928.
- Zhang H, Fu W: NDPP1 is a novel CARD domain containing protein which can inhibit apoptosis and suppress NF-kappaB activation. *Int J Oncol* 2002; **20**: 1035–1040.
- Pathan N, Marusawa H, Krajewska M *et al*: TUCAN, an anti-apoptotic caspase-associated recruitment domain family protein overexpressed in cancer. *J Biol Chem* 2001; **276**: 32220–32229.
- Bouchier-Hayes L, Conroy H, Egan H *et al*: CARDINAL, a novel caspase recruitment domain protein, is an inhibitor of multiple NF-kappa B activation pathways. *J Biol Chem* 2001; **276**: 44069–44077.
- Yamamoto M, Torigoe T, Kamiguchi K *et al*: A novel isoform of TUCAN is overexpressed in human cancer tissues and suppresses both caspase-8- and caspase-9-mediated apoptosis. *Cancer Res* 2005; **65**: 8706–8714.
- 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.
- Rioux JD, Silverberg MS, Daly MJ *et al*: Genomewide search in Canadian families with inflammatory bowel disease reveals two novel susceptibility loci. *Am J Hum Genet* 2000; **66**: 1863–1870.
- Duerr RH, Barnada MM, Zhang L, Pfutzer R, Weeks DE: High-density genome scan in Crohn disease shows confirmed linkage to chromosome 14q11–12. *Am J Hum Genet* 2000; **66**: 1857–1862.
- Van Heel DA, Dechairo BM, McGovern DPB *et al*: The IBD6 Crohn's disease locus demonstrates complex interactions with *CARD15* and *IBD5* disease-associated variants. *Hum Mol Genet* 2003; **12**: 2569–2575.
- Van Heel DA, Fisher SA, Kirby A *et al*: Inflammatory bowel disease susceptibility loci defined by genome scan meta-analysis of 1952 affected relative pairs. *Hum Mol Genet* 2004; **13**: 763–770.
- McGovern DP, Butler H, Ahmad T *et al*: TUCAN (*CARD8*) genetic variants and inflammatory bowel disease. *Gastroenterology* 2006; **131**: 1190–1196.
- Fisher SA, Mirza MM, Onnie CM *et al*: Combined evidence from three large British Association studies rejects TUCAN/*CARD8* as an IBD susceptibility gene. *Gastroenterology* 2007; **132**: 2078–2080.
- Franke A, Rosenstiel P, Balschun T *et al*: No association between the TUCAN (*CARD8*) Cys10Stop mutation and inflammatory bowel disease in a large retrospective German and a clinically well-characterized Norwegian sample. *Gastroenterology* 2007; **132**: 2080–2081.
- The Wellcome Trust Case Control Consortium: Genome-wide association study of 14 000 cases of seven common diseases and 3000 shared controls. *Nature* 2007; **447**: 661–683.
- Baker KE, Parker R: Nonsense-mediated mRNA decay: terminating erroneous gene expression. *Curr Opin Cell Biol* 2004; **16**: 293–299.
- Cartegni L, Chew SL, Krainer AR: Listening to silence and understanding nonsense: exonic mutations that affect splicing. *Nat Rev Genet* 2002; **3**: 285–298.
- Razmara M, Srinivasula SM, Wang L *et al*: CARD-8 protein, a new CARD family member that regulates caspase-1 activation and apoptosis. *J Biol Chem* 2002; **277**: 13952–13958.
- Checinska A, Oudejans JJ, Span SW *et al*: The expression of TUCAN, an inhibitor of apoptosis protein, in patients with advanced non-small cell lung cancer treated with chemotherapy. *Anticancer Res* 2006; **26**: 3819–3824.

Supplementary Information accompanies the paper on European Journal of Human Genetics website (<http://www.nature.com/ejhg>)