Human Molecular Genetics, 2005, Vol. 14, No. 10 ***1245–1250***
doi:10.1093/hmg/ddi135
Advance Access published on March 24, 2005

Association between a complex insertion/deletion polymorphism in *NOD1* (*CARD4*) and susceptibility to inflammatory bowel disease

Dermot P.B. McGovern[1,2,*], Pirro Hysi[1], Tariq Ahmad[2], David A. van Heel[3], Miriam F. Moffatt[1], Alisoun Carey[4], William O.C. Cookson[1] and Derek P. Jewell[2]

[1]The Wellcome Trust Centre for Human Genetics, University of Oxford, Roosevelt Drive, Headington, Oxford OX3 7BN, UK, [2]Gastroenterology Unit, University of Oxford, Radcliffe Infirmary, Oxford, [3]Imperial College, Hammersmith Campus, London, UK and [4]Oxagen Ltd, Abingdon, Oxfordshire, UK

Received November 19, 2004; Revised March 11, 2005; Accepted March 21, 2005

The identification of the role of genetic variants within *NOD2* (*CARD15*) in Crohn's disease and ulcerative colitis susceptibility highlight the role of the innate immune system in inflammatory bowel disease (IBD) pathogenesis. *NOD1* (*CARD4*) is located on chromosome 7p14.3, in a region of known linkage to IBD and encodes an intracellular bacterial pathogen-associated molecular pattern receptor that is closely related to *NOD2*. We have identified strong association between haplotypes in the terminal exons of *NOD1* and IBD (multi-allelic $P = 0.0000003$) in a panel of 556 IBD trios. The deletion allele of a complex functional *NOD1* indel polymorphism (*ND_1 + 32656*1) was significantly associated with early-onset IBD ($P = 0.0003$) in unrelated cases and controls. *ND1 + 32656*1 was also associated with extra-intestinal manifestations of IBD ($P = 0.04$). These findings in two independent populations provide strong evidence for a role for *NOD1* variants in IBD susceptibility and reinforce the role of the innate immune system in IBD pathogenesis.

INTRODUCTION

The inflammatory bowel diseases (IBD), ulcerative colitis (UC) and Crohn's disease (CD) are related polygenic conditions sharing some but not all susceptibility alleles. They have a combined prevalence of 400/100 000 in western Europe (1) and are a significant cause of morbidity among young people. The identification of *NOD2*(*CARD15*) as the first susceptibility gene for CD (2,3) was a breakthrough in understanding IBD pathogenesis and for complex disease genetics in general.

The NOD2 protein is made up of a leucine rich repeat region (LRR) that recognizes bacterial muramyl dipeptide (4), a nucleotide binding domain and two caspase recruitment domains. Variants within or adjacent to the LRR muramyl dipeptide recognition domain are associated with altered NFκB activation and increased susceptibility to CD (2).

The CD association with *NOD2* has been widely replicated (5) and the association is particularly strong with CD of the small bowel (6,7). More recently, an association between *NOD2* variants and UC has also been described (8). The population attributable risks for *NOD2* variants are ~30 and 6% for CD (6) and UC (8), respectively. Genome-wide scans indicate that several other genetic loci for IBD exist (9–14).

A haplotype on 5q31 (*IBD5*) (15) has been widely replicated as an IBD locus and recently it has been suggested that variants in *OCTN1* and *OCTN2*, cation transporter genes, are the *IBD5* susceptibility genes (16,17). Variants in *DLG5*, encoding an epithelial scaffolding protein, on chromosome 10q23 have also been associated with susceptibility to IBD (16,17), although this finding has not yet been replicated. A further susceptibility locus for UC and CD has been identified on chromosome *7p14* in a British genome-wide scan for linkage to IBD (18). Further evidence for linkage to this region has been demonstrated in other genome-wide scans (9,10) and in a recent meta-analysis of genome scans for IBD (19).

The gene encoding NOD1 (CARD4) is located within the chromosome 7p14 IBD locus. NOD1 is similar in structure to NOD2. It contains LRR and NOD domains but differs from NOD2 in the presence of a single CARD. NOD1, like NOD2, activates NFκB and enhances apoptosis. NOD1

*To whom correspondence should be addressed. Tel: +44 1865 287651; Fax: +44 1865 287533; Email: dermot@well.ox.ac.uk

Table 1. NOD1 polymorphisms tested in family panels

Position ID[a]	Database ID	Frequency[b]
ND1 − 664	rs2736726	0.754
ND1 + 233	rs2075817	0.751
ND1 + 18915	rs2975632	0.801
ND1 + 21658	rs3020207	0.709
ND1 + 21984	rs2075818	0.765
ND1 + 25816	rs2235099	0.766
ND1 + 26129	rs3020208	0.924
ND1 + 27053	rs2075821	0.769
ND1 + 27606	rs2075822	0.779
ND1 + 32656[c]	–	0.782
ND1 + 45343	rs2907748	0.799
ND1 + 50150	rs5743368	0.886

[a]The position is numbered from the first nucleotide of exon 1. The sequence is obtained from Golden Path (http://genome.ucsc.edu).
[b]Frequency in unrelated individuals (parents). Allele **1* is defined as the more common.
[c]Complex insertion/deletion polymorphism (indel), partially identified as rs6958571.

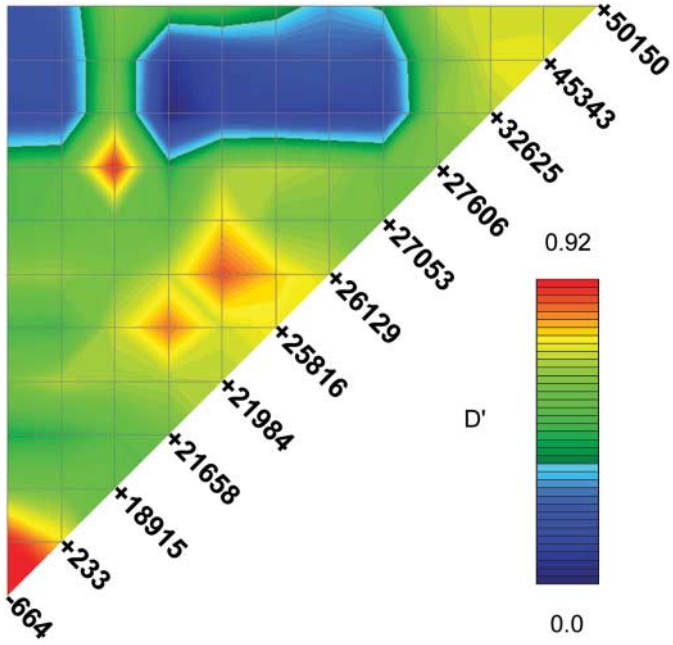

Figure 1. Linkage disequilibrium between markers. The marker positions are shown along the diagonal of the figure. Pairwise estimations of *D′* are shown from unrelated subjects in the family panel (the parents), on a scale of 1 (complete linkage disequilibrium: red) to 0 (blue). Marker positions are shown as a schematic rather than as actual distances apart.

detects a unique tripeptide motif (diaminopimelic acid) found in Gram-negative bacterial peptidoglycan (20). A previous study had typed a single *NOD1* polymorphism and failed to detect association with IBD (21). The aim of our study was therefore to assess, more comprehensively, genetic variation in *NOD1* for influences on IBD susceptibility.

RESULTS

We genotyped 12 previously identified *NOD1* polymorphisms (Table 1) (22) in a panel of 556 IBD trios containing 294 CD trios, 252 UC trios and 10 trios with a diagnosis of indeterminate colitis. We tested for association using the transmission disequilibrium test (TDT). We found that the common deletion allele of a complex polymorphism (*ND1 + 32656*1*) was significantly associated with IBD ($P = 0.02$) and with UC ($P = 0.01$). Associations were also observed between IBD and *ND1 + 233*1* ($P = 0.05$), *ND1 + 21984*1* ($P = 0.02$) and *ND1 + 27606*2* ($P = 0.05$). The *rs5743336* (*E266K*) *NOD1* polymorphism previously examined in IBD (21) had a frequency $<1\%$ in our subjects and was not tested.

Linkage disequilibrium between the markers was assessed in unrelated individuals (Fig. 1). LD was incomplete between all the markers, and in particular LD between the *ND1 + 32656*1* polymorphism and neighbouring markers was weak or absent. These findings are consistent with those previously observed in asthmatic families (22).

We further investigated the locus by examining two-, three- and four-marker haplotypes in a sliding window across the locus. Strong associations were observed between two-marker *ND1 + 32656/ND1 + 45343* haplotypes and CD (Multi-allelic TDT $P = 0.007$), UC ($P = 0.00007$) and IBD ($P = 0.0000003$) (Fig. 2). Examination of individual haplotypes indicated the presence of a strong protective effect of the *ND1 + 32656*2/ND1 + 45343*1* haplotype (frequency 7%) to IBD (Table 2). Extension of the haplotype to include other markers resulted in increased haplotype diversity, but progressively decreased the evidence for association. This result is consistent with as yet undiscovered effects in the interval between *ND1 + 32656* and *ND1 + 45343*.

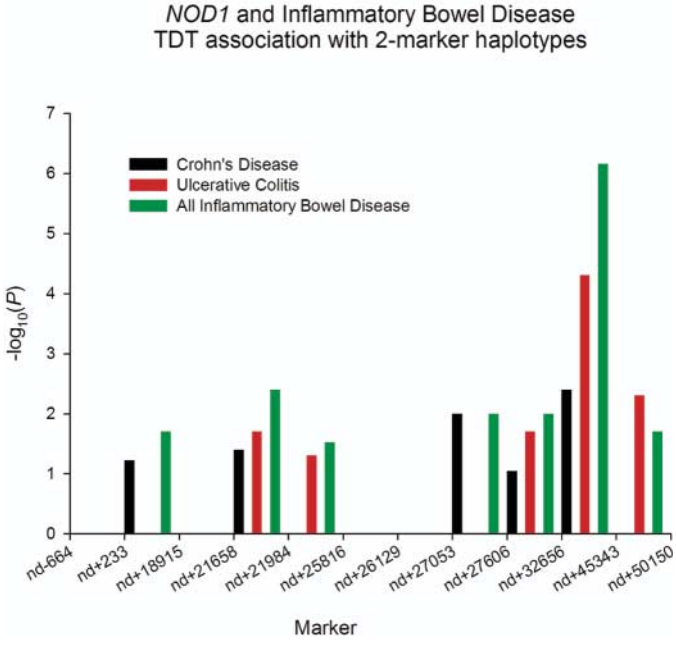

Figure 2. Family-based association study of *NOD1* two-marker haplotypes in a sliding window. *P*-values were calculated with the extended multi-allelic transmission disequilibrium test (ETDT).

Table 2. Transmission of *ND1 + 27606/ND1 + 32656* haplotypes in family panels

Haplotype	Frequency[a]	CD			CD age diagnosis <25			UC			IBD		
		T	NT	*P*-value	T	NT	*P*-value	T	NT	*P*-value	T	NT	*P*-value
1 (*1*1)	0.725	79	67	ns	40	44	ns	65	58	ns	146	127	ns
2 (*2*2)	0.156	73	64	ns	47	35	ns	61	45	ns	136	111	ns
3 (*2*1)	0.069	16	39	0.002	12	26	0.02	10	43	6×10^{-6}	26	83	5×10^{-8}
4 (*1*2)	0.050	24	22	ns	19	13	ns	22	12	0.09	47	34	ns

[a]Frequency in unrelated founders.

Table 3. Genotype frequencies of *ND1 + 32656* in healthy controls and patients with IBD

Status	*ND1 + 32656* Genotype			*P*-value[a]
	*1*1	*1*2	*2*2	
Controls	161	135	39	
IBD	358	259	47	0.017
UC	153	128	25	ns
CD	205	131	22	0.003
IBD onset <25	143	71	14	0.0003
CD onset <25	106	44	8	0.00004

[a]Mantel–Haenszel test for linear association compared with controls.

Weaker associations with other two-marker haplotypes were observed in a cluster between *ND1 + 21658* and *ND1 + 26129* (Fig. 1). These results suggested independent effects within *NOD1* on susceptibility to IBD.

To confirm these results we then tested *ND1 + 32656* in an independent panel of 664 subjects with IBD (358 CD and 306 UC) and 335 controls. In these subjects the *ND1 + 32656*1* allele was significantly associated with CD ($P = 0.003$) and IBD ($P = 0.017$) but not with UC (Table 3). However, although association to UC was not seen in these subjects, the trend was in the same direction as for CD, and the overall data from the family and case-control studies suggest that *ND1 + 32656*1* may be associated with general suscepibility to IBD rather than CD or UC alone. These findings are consistent with original linkage data (18,19).

We then compared the risk of CD conferred by *ND1 + 32656*1* and the known susceptibility loci *NOD2* and *IBD5* in the cases and controls by logistic regression. The OR for *ND1 + 32656*1* was 2.0 [95% confidence Interval (CI) 1.2–3.5], which was comparable with the effect of *IBD5* (OR = 2.3; 95% CI = 1.7–3.1) but less than that of *NOD2* (OR = 6.45; 95% CI = 4.1–10.2). There was no evidence of genetic interaction between the three loci.

Stratification of cases by age showed that early diagnosis (<25 years) of disease was strongly associated with *ND1 + 32656*1* for CD ($P = 0.00004$) and IBD ($P = 0.0003$) (Table 3). Wilcoxon life-table analyses demonstrated that the median age of diagnosis of CD was 32.3 years for the *ND1 + 32656*1*1* genotype, 36.2 for *1*2* and 49.6 for *2*2* ($P = 0.001$ for the difference) (Fig. 3). Similar results were seen if all the cases of IBD were examined together.

Examination of *NOD1, NOD2* and *IBD5* in a multiple regression showed that only *ND1 + 32656* was significantly associated with the age of diagnosis of CD ($P = 0.023$) (Table 4), although the trend was also for *NOD2* to predict early-onset of disease ($P = 0.06$). *IBD5* was not associated with early diagnosis of disease in these subjects ($P = 0.47$), in contrast to an earlier study (15). Univariate Spearman correlation coefficients were *NOD1,* $R = 0.174$, $P = 0.002$; *IBD5,* $R = 0.01$, $P = 0.923$ and *NOD2* $R = 0.057$, $P = 0.31$. No genetic interactions between the loci were found to affect the age of diagnosis of disease.

No significant independent association was seen when the IBD samples were stratified by the presence of stenotic or fistulating disease, disease location or smoking status. However, *ND1 + 32656*1* was associated with the presence of IBD extra-intestinal manifestations (EIMs: large and small joint arthritis, ankylosing spondylitis, occular inflammation, erythema nodosum, pyoderma gangrenosum and primary sclerosing cholangitis) (OR 1.35, 95% CI = 1.02–1.82, $P = 0.04$ when compared with individuals with IBD and without EIMs).

DISCUSSION

We have demonstrated association between genetic variants in *NOD1* and susceptibility to IBD in two independent cohorts. *ND1 + 32656*, which is located on the sliding two-locus haplotype significantly associated with IBD, is located at the beginning of intron IX and affects the binding of an unknown nuclear factor (22). NOD1 is expressed in both large and small bowel as a number of splice variants arising from exon IX (22). It has not yet been established whether *ND1 + 32656* alters the splicing of these products, but progressive skipping of exons X–XII results in proteins with a reduced number of LRR (22). This has similarities with the three common *NOD2* CD associated variants adjacent to or within the LRR coding region, in which the greatest risk for CD is conferred by the *fsinsC1007* frameshift mutation that truncates the protein's LRR.

It is of interest that the common *ND1 + 32656*1* allele was associated with susceptibility to IBD, whereas *ND1 + 32656*2* has been shown to confer susceptibility to asthma (22). These two diseases have quite distinctive patterns of inflammation with different patterns of cellular infiltration and a different cytokine milieu. This might suggest that changes in the structure or regulation of the NOD1 protein alter reactivity to particular antigens, or differentially regulate the nature of downstream inflammatory pathways.

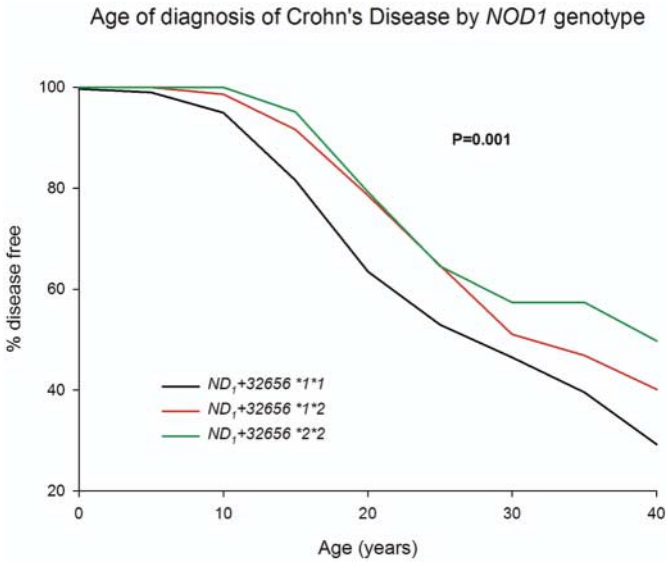

Figure 3. Life-table analyses of age of diagnosis of Crohn's Disease according to $ND_1 + 32656$ genotype.

NOD1 has been examined previously for a role in IBD susceptibility, with the conclusion that it does not influence disease susceptibility (21). However, these investigations were only concerned with rare coding polymorphism. The general recognition now is that complex disease susceptibility is expected often to be mediated through regulatory polymorphisms, as is the case in the present study.

NOD1 is expressed in large and small bowel (22) and plays a role in colonic epithelial defences against intracellular organisms, such as *Shigella flexeneri* (23) and enteroinvasive *E. coli* (24). The presence of bacterial flora is essential for IBD to develop in animal models (25), and antibiotics (26) and faecal diversion (27) are effective therapies for CD. The identification of associations between *NOD1, NOD2* and IBD suggest that altered recognition of intracellular bacterial pathogen-associated molecular pattern may be a key event in the pathogenesis of the disease. Further work is now required to establish the effect of these variants within the gastrointestinal tract and the role they play in IBD behaviour and response to treatment.

MATERIALS AND METHODS

Study populations

Five hundred and fifty-six IBD nuclear trios were recruited as previously described (5). An independent panel of 664 subjects with IBD [358 CD and 306 UC, median age of onset of disease 28.2 years (range 1.0–82.2 years), 44.4% male] and 335 controls was studied to confirm potentially positive results. All subjects and controls were Caucasian. The family collection was obtained from probands attending the Oxford IBD clinic and referred from gastroenterologists from around UK. Non-family cases were recruited patients attending the John Radcliffe Hospital, Oxford IBD clinic and were compared with healthy unrelated individuals recruited through the UK blood transfusion service and from healthy individuals attending 'well-person' screening clinics at their family doctors in Oxfordshire, UK.

Table 4. Effects of known susceptibility genes on age of diagnosis of CD: forward stepwise multiple regression analysis

Step	Variable[a]	R^2	R^2 change	*P*-value for change	Beta
1	NOD1[b]	0.017	0.017	0.023	0.130
2	NOD2[c]	0.028	0.011	0.060	−0.107

[a]Variables not entered in the equation are: sex (beta = 0.080, $P = 0.16$) and IBD5 (beta = 0.041, $P = 0.46$).
[b]$ND_1 + 32656$.
[c]*C2104T, G2722C* and *3020insC*.

All diseased cases were diagnosed as having IBD according to standard clinical (history of abdominal pain, weight loss, rectal bleeding, diarrhoea, abdominal mass, perianal disease with or without evidence of IBD EMIs), endoscopic (macroscopic evidence of rectal, colonic or terminal ileal mucosal inflammation), radiological (superficial or deep ulceration, presence of fistulae or strictures and distribution of disease) and histological findings (inflammation, inflammatory infiltrates, glandular architecture distortion, goblet cell depletion and the presence of granulomas). Patients were classified with indeterminate colitis in the presence of definitive IBD affecting the colon only but with histology ambiguous to the presence of UC or CD.

All individuals in the study gave written informed consent and ethical approval from the relevant hospital ethical committees had been obtained.

SNP identification and genotyping

Polymorphisms in *NOD1* were identified and genotyping was performed as previously described (22). Alleles for the three common CD associated *NOD2* variants (Arg702Trp, Gly908Arg and the *fsinsC1007)* and for an *IBD5* CD risk haplotype tagging SNP were genotyped as previously described (5,15,28).

Statistics

Polymorphisms in the family panel were tested for Mendelian transmission and for Hardy–Weinberg equilibrium before inclusion in the analyses. Association to categorical traits within the families was examined by the TDT. Haplotypes were generated by the MERLIN computer program (29), and D' between markers estimated by the MERLIN utility HAPLOXT. The patterns of linkage disequilibrium between markers were visualized using GOLD (30). Two-, three- and four-marker haplotypes in the trios were generated across the locus in a sliding window by MERLIN (29), and coded as individual alleles before analysis by the multi-allelic TDT (ETDT) test (31).

In the case–control panel Hardy–Weinberg equilibrium was established before using the Mantel–Haenszel test for linear association to compare genotype frequencies and

life-table analyses with the Wilcoxon (Gehan) statistic to compare age of onset of disease between genotypes (SPSS for OSF, 6.1.4). The effects of *NOD1, NOD2* and *IBD5* genotypes on disease risk were estimated by logistic regression, and their effects on age of diagnosis were examined by multiple linear regression (SPSS for OSF, 6.1.4). Gene–gene interactions were tested as interaction terms within the regressions. The effects of phenotypic variants were examined by χ^2 estimations of contingency tables selected by phenotypic stratum.

Given the close correlations between various phenotypes and between the markers themselves, simple corrections for multiple comparisons were not applied to the initial test results and replication was used to determine true genetic effects. In order to maintain a conservative approach, two-tailed *P*-values were estimated for the replicate samples.

ACKNOWLEDGEMENTS

Support for this work was provided by the Medical Research Council (MRC) (D.Mc.G., D.v.H.), by an MRC LINK grant (Oxagen Ltd), The Digestive Diseases Foundation and Belmont Trust (D.Mc.G.), The International Organisation for the study of Inflammatory Bowel Disease (D.Mc.G.), The National Association for Colitis and Crohn's Disease (T.A.) and The Wellcome Trust (W.O.C.C. and M.F.M.). P.H. is the recipient of a Victor Blank scholarship.

REFERENCES

1. Rubin, G.P., Hungin, A.P., Kelly, P.J. and Ling, J. (2000) Inflammatory bowel disease: epidemiology and management in an English general practice population. *Aliment. Pharmacol. Ther.*, **14**, 1553–1559.
2. Hugot, J.P., Chamaillard, M., Zouali, H., Lesage, S., Cezard, J.P., Belaiche, J., Almer, S., Tysk, C., O'Morain, C.A., Gassull, M. *et al.* (2001) Association of NOD2 leucine-rich repeat variants with susceptibility to Crohn's disease. *Nature*, **411**, 599–603.
3. Ogura, Y., Bonen, D.K., Inohara, N., Nicolae, D.L., Chen, F.F., Ramos, R., Britton, H., Moran, T., Karaliuskas, R., Duerr, R.H. *et al.* (2001) A frameshift mutation in NOD2 associated with susceptibility to Crohn's disease. *Nature*, **411**, 603–606.
4. Inohara, N., Ogura, Y., Fontalba, A., Gutierrez, O., Pons, F., Crespo, J., Fukase, K., Inamura, S., Kusumoto, S., Hashimoto, M. *et al.* (2003) Host recognition of bacterial muramyl dipeptide mediated through NOD2. Implications for Crohn's disease. *J. Biol. Chem.*, **278**, 5509–5512.
5. van Heel, D.A., McGovern, D.P., Cardon, L.R., Dechairo, B.M., Lench, N.J., Carey, A.H. and Jewell, D.P. (2002) Fine mapping of the IBD1 locus did not identify Crohn disease-associated NOD2 variants: implications for complex disease genetics. *Am. J. Med. Genet.*, **111**, 253–259.
6. Ahmad, T., Armuzzi, A., Bunce, M., Mulcahy-Hawes, K., Marshall, S.E., Orchard, T.R., Crawshaw, J., Large, O., de Silva, A., Cook, J.T. *et al.* (2002) The molecular classification of the clinical manifestations of Crohn's disease. *Gastroenterology*, **122**, 854–866.
7. Lesage, S., Zouali, H., Cezard, J.P., Colombel, J.F., Belaiche, J., Almer, S., Tysk, C., O'Morain, C., Gassull, M., Binder, V. *et al.* (2002) CARD15/NOD2 mutational analysis and genotype–phenotype correlation in 612 patients with inflammatory bowel disease. *Am. J. Hum. Genet.*, **70**, 845–857.
8. McGovern, D.P., van Heel, D.A., Negoro, K., Ahmad, T. and Jewell, D.P. (2003) Further evidence of IBD5/CARD15 (NOD2) epistasis in the susceptibility to ulcerative colitis. *Am. J. Hum. Genet.*, **73**, 1465–1466.
9. Cho, J.H., Nicolae, D.L., Gold, L.H., Fields, C.T., LaBuda, M.C., Rohal, P.M., Pickles, M.R., Qin, L., Fu, Y., Mann, J.S. *et al.* (1998) 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*, **95**, 7502–7507.
10. Rioux, J.D., Silverberg, M.S., Daly, M.J., Steinhart, A.H., McLeod, R.S., Griffiths, A.M., Green, T., Brettin, T.S., Stone, V., Bull, S.B. *et al.* (2000) Genomewide search in Canadian families with inflammatory bowel disease reveals two novel susceptibility loci. *Am. J. Hum. Genet.*, **66**, 1863–1870.
11. Hampe, J., Schreiber, S., Shaw, S.H., Lau, K.F., Bridger, S., Macpherson, A.J., Cardon, L.R., Sakul, H., Harris, T.J., Buckler, A. *et al.* (1999) A genomewide analysis provides evidence for novel linkages in inflammatory bowel disease in a large European cohort. *Am. J. Hum. Genet.*, **64**, 808–816.
12. Ma, Y., Ohmen, J.D., Li, Z., Bentley, L.G., McElree, C., Pressman, S., Targan, S.R., Fischel-Ghodsian, N., Rotter, J.I. and Yang, H. (1999) A genome-wide search identifies potential new susceptibility loci for Crohn's disease. *Inflamm. Bowel. Dis.*, **5**, 271–278.
13. Duerr, R.H., Barmada, M.M., Zhang, L., Pfutzer, R. and Weeks, D.E. (2000) High-density genome scan in Crohn disease shows confirmed linkage to chromosome 14q11-12. *Am. J. Hum. Genet.*, **66**, 1857–1862.
14. van Heel, D.A., Dechairo, B.M., Dawson, G., McGovern, D.P., Negoro, K., Carey, A.H., Cardon, L.R., Mackay, I., Jewell, D.P. and Lench, N.J. (2003) The IBD6 Crohn's disease locus demonstrates complex interactions with CARD15 and IBD5 disease-associated variants. *Hum. Mol. Genet.*, **12**, 2569–2575.
15. Rioux, J.D., Daly, M.J., Silverberg, M.S., Lindblad, K., Steinhart, H., Cohen, Z., Delmonte, T., Kocher, K., Miller, K., Guschwan, S. *et al.* (2001) Genetic variation in the 5q31 cytokine gene cluster confers susceptibility to Crohn disease. *Nat. Genet.*, **29**, 223–228.
16. Stoll, M., Corneliussen, B., Costello, C.M., Waetzig, G.H., Mellgard, B., Koch, W.A., Rosenstiel, P., Albrecht, M., Croucher, P.J., Seegert, D. *et al.* (2004) Genetic variation in DLG5 is associated with inflammatory bowel disease. *Nat. Genet.*, **36**, 476–480.
17. Peltekova, V.D., Wintle, R.F., Rubin, L.A., Amos, C.I., Huang, Q., Gu, X., Newman, B., van Oene, M., Cescon, D., Greenberg, G. *et al.* (2004) Functional variants of OCTN cation transporter genes are associated with Crohn disease. *Nat. Genet.*, **36**, 471–475.
18. Satsangi, J., Parkes, M., Louis, E., Hashimoto, L., Kato, N., Welsh, K., Terwilliger, J.D., Lathrop, G.M., Bell, J.I. and Jewell, D.P. (1996) Two stage genome-wide search in inflammatory bowel disease provides evidence for susceptibility loci on chromosomes 3, 7 and 12. *Nat. Genet.*, **14**, 199–202.
19. van Heel, D.A., Fisher, S.A., Kirby, A., Daly, M.J., Rioux, J.D. and Lewis, C.M. (2004) Inflammatory bowel disease susceptibility loci defined by genome scan meta-analysis of 1952 affected relative pairs. *Hum. Mol. Genet.*, **13**, 763–770.
20. Chamaillard, M., Hashimoto, M., Horie, Y., Masumoto, J., Qiu, S., Saab, L., Ogura, Y., Kawasaki, A., Fukase, K., Kusumoto, S. *et al.* (2003) An essential role for NOD1 in host recognition of bacterial peptidoglycan containing diaminopimelic acid. *Nat. Immunol.*, **4**, 702–707.
21. Zouali, H., Lesage, S., Merlin, F., Cezard, J.P., Colombel, J.F., Belaiche, J., Almer, S., Tysk, C., O'Morain, C., Gassull, M. *et al.* (2003) CARD4/NOD1 is not involved in inflammatory bowel disease. *Gut*, **52**, 71–74.
22. Hysi, P., Kabesch, M., Moffatt, M.F., Schedel, M., Carr, D., Zhang, Y., Boardman, B., von Mutius, E., Weiland, S., Leupold, W. *et al.* (2005) NOD1 variation, immunoglobulin E and asthma. *Hum. Mol. Genet.*, **14**, 935–941.
23. Girardin, S.E., Tournebize, R., Mavris, M., Page, A.L., Li, X., Stark, G.R., Bertin, J., DiStefano, P.S., Yaniv, M., Sansonetti, P.J. *et al.* (2001) CARD4/Nod1 mediates NF-kappaB and JNK activation by invasive Shigella flexneri. *EMBO Rep.*, **2**, 736–742.
24. Kim, J.G., Lee, S.J. and Kagnoff, M.F. (2004) Nod1 is an essential signal transducer in intestinal epithelial cells infected with bacteria that avoid recognition by toll-like receptors. *Infect. Immun.*, **72**, 1487–1495.
25. Kuhn, R., Lohler, J., Rennick, D., Rajewsky, K. and Muller, W. (1993) Interleukin-10-deficient mice develop chronic enterocolitis. *Cell*, **75**, 263–274.
26. Rutgeerts, P., Hiele, M., Geboes, K., Peeters, M., Penninckx, F., Aerts, R. and Kerremans, R. (1995) Controlled trial of metronidazole treatment for prevention of Crohn's recurrence after ileal resection. *Gastroenterology*, **108**, 1617–1621.

27. Rutgeerts, P., Goboes, K., Peeters, M., Hiele, M., Penninckx, F., Aerts, R., Kerremans, R. and Vantrappen, G. (1991) Effect of faecal stream diversion on recurrence of Crohn's disease in the neoterminal ileum. *Lancet*, **338**, 771–774.
28. Negoro, K., McGovern, D.P., Kinouchi, Y., Takahashi, S., Lench, N.J., Shimosegawa, T., Carey, A., Cardon, L.R., Jewell, D.P. and van Heel, D.A. (2003) Analysis of the IBD5 locus and potential gene–gene interactions in Crohn's disease. *Gut*, **52**, 541–546.
29. Abecasis, G.R., Cherny, S.S., Cookson, W.O. and Cardon, L.R. (2002) Merlin-rapid analysis of dense genetic maps using sparse gene flow trees. *Nat. Genet.*, **30**, 97–101.
30. Abecasis, G.R. and Cookson, W.O. (2000) GOLD-graphical overview of linkage disequilibrium. *Bioinformatics*, **16**, 182–183.
31. Sham, P.C. and Curtis, D. (1995) An extended transmission/disequilibrium test (TDT) for multi-allele marker loci. *Ann. Hum. Genet.*, **59**, 323–336.