ELSEVIER

FEMS Immunology and Medical Microbiology 41 (2004) 157–160

FEMS Immunology and Medical Microbiology

www.fems-microbiology.org

Polymorphism in NOD2, Crohn's disease, and susceptibility to pulmonary tuberculosis

Jemima C. Stockton [a], Joanna M.M. Howson [b], Agnes A. Awomoyi [c], Keith P. McAdam [c], Jenefer M. Blackwell [a], Melanie J. Newport [a,c,*]

[a] *Department of Medicine, University of Cambridge School of Clinical Medicine, Wellcome Trust/MRC Building, Addenbrooke's Hospital, Hills Road, Cambridge CB2 2XY, UK*

[b] *Diabetes and Inflammation Laboratory, Cambridge Institute for Medical Research, University of Cambridge, CB2 2XY, UK*

[c] *Medical Research Council Laboratories, P.O. Box 273, Banjul, The Gambia*

Received 25 July 2003; received in revised form 26 November 2003; accepted 27 February 2004

First published online 10 April 2004

Abstract

The nucleotide oligomerization binding domain 2 gene (*NOD2*) encodes an intracellular receptor for bacterial components, which is expressed in monocytes and is associated with Crohn's Disease (CD). This finding, along with epidemiological evidence, supports a role for infection in the pathogenesis of CD. Speculation that mycobacteria are involved in CD led us to investigate NOD2 in susceptibility to tuberculosis (TB), a global public health problem caused by *Mycobacterium tuberculosis*. CD-associated NOD2 variants were absent in a case-control study of 640 Gambians, where CD is rare. Novel NOD2 promoter polymorphisms were identified but showed no association with TB in this African population sample.

Keywords: Tuberculosis; Crohn's Disease; Genetic susceptibility; NOD2; Gambia

1. Introduction

Host genetic factors are important determinants of the outcome of infection with *Mycobacterium tuberculosis* (*M.tb*): most people remain well and contain infection while approximately 10% develop clinical disease [1]. It is hoped that the identification of the genes involved will lead to the development of new therapies and vaccines required for the eradication of tuberculosis (TB), given the poor efficacy of BCG in TB-endemic countries, and the emergence of drug-resistant *M.tb* strains [2,3].

Although controversial, a number of observations support a role for mycobacteria in the etiology of Crohn's Disease (CD) [4,5]. The granulomatous histological changes observed in CD strongly resemble those seen in TB, the major difference being the absence of overt acid fast bacilli in CD. There are many similarities between CD and bovine inflammatory bowel disease (Johne's disease), caused by *Mycobacterium avium* subspecies *paratuberculosis* (MAP). MAP nucleic acid has been identified in intestinal tissue from CD patients by a number of groups [6,7], and several studies suggest CD patients respond clinically to antimycobacterial chemotherapy [8].

CD, in common with TB, has a complex genetic inheritance with disease occurring in only a proportion of individuals exposed to the causative agent [9]. Nucleotide oligomerization binding domain 2 (NOD2) is an intracellular cytosolic receptor for bacterial components, which is expressed in monocytes. *NOD2* was recently identified as a CD susceptibility gene [10,11], adding to the evidence in support of an infective contribution to the etiology of CD, whether or not MAP is finally implicated. Genetic variants within or adjacent to the leucine-rich repeat (LRR) region of the *NOD2* gene

* Corresponding author. Tel.: +44-1223-331153; fax: +44-1223-331206.

E-mail address: melanie.newport@cimr.cam.ac.uk (M.J. Newport).

0928-8244/$22.00
doi:10.1016/j.femsim.2004.02.004

associate with CD. One variant in particular, a C insertion at nucleotide position 2936 (2936insC), results in a premature stop codon and a truncated protein lacking the terminal 33 amino acids of the LLR [12]. NOD2 shows structural homology to the cytosolic disease-resistant (R) proteins of plants [13], which protect plants from a range of invasive pathogens including bacteria, fungi and viruses. NOD2 has been shown to interact with several microbial components [14] including muramyl dipeptides, which are present in mycobacteria, to induce early innate immune responses that subsequently orchestrate the adaptive response. Thus, regardless of whether mycobacteria contribute to the etiology of CD or not, *NOD2* is a plausible TB-susceptibility candidate gene. We therefore tested for association between genetic variation in *NOD2* and pulmonary TB (PTB) in a Gambian population sample, where *M.tb* infection is endemic.

2. Methods

2.1. Study subjects

This study had the approval of the MRC/Gambia Government Ethics Committee. Three hundred and twenty men with pulmonary TB (mean age 36 years, range 19–58) were recruited from the TB clinic based at Serrekunda Health Centre, The Gambia. All cases had microscopically proven smear positive TB in the context of clinical and radiological findings consistent with pulmonary TB, in the absence of other chronic illness, HIV infection, or treatment with corticosteroids. Three hundred and twenty healthy male HIV negative blood donors (mean age 32 years, range 18–50) matched for ethnicity and age to the TB cases were simultaneously recruited at the Royal Victoria Hospital, Banjul. The rationale for selecting blood donors as controls is presented elsewhere [15]. Our study had 90% statistical power to detect a gene effect with an odds ratio of >1.5 at $P = 0.05$, assuming multiplicative model and minor allele frequency of 0.05 or more. Haplotypes frequencies were generated by estimation maximization and multiple imputation using the software COCAPHASE [16] (http://www.hgmp.mrc.ac.uk/~fdudbrid/software/unphased/).

2.2. Genotyping of three CD-associated single nucleotide polymorphisms

The three CD-associated single nucleotide polymorphisms (SNPs) (2936insC, R675W and G1881R) were genotyped using Invader technology (Third Wave Technologies™) as described elsewhere [17]. Genotypes were confirmed in a subset of 192 subjects by RFLP analysis as described elsewhere [18].

2.3. Sequencing of NOD2 exon 11 and promoter regions

The NOD2 promoter, covering a region 1475-bp upstream to 293-bp downstream of the first base of the NOD2 start codon, and exon 11, which contains the 2963insC SNP, were sequenced in 40 individuals using the ABI Big Dye Terminator v2 kit in conjunction with the ABI recommended protocol, the ABI 3700 capillary sequencer (Foster City, CA) and the Staden package for analysis (http://www.mrc-lmb.cam.ac.uk/pubseq/). PCR and sequencing primers designed from NCBI sequence data (http://www.ncbi.nlm.nih.gov/NM_022162, Table 1).

2.4. Genotyping of novel promoter regions SNPs

The promoter SNPs −274C>T and −709G>A were genotyped by RFLP. The primers AAGGACAATTTT-AGGAAACAGATCA (forward) and CTGTCTCT-GAGCAGGCATTG (reverse) amplify a 209-bp product across the −709 SNP. A one-base mismatch (in bold) was incorporated into the forward primer for SNP −709 in

Table 1
Primers used for sequencing the promoter and exon 11 regions of the NOD2 gene

Position (bp)[a]	Primer (direction)	Sequencing primers
Promoter sequencing		
−1475 to +293	ccctggcagcaagca (forward)	ccctggcagcaagca
	tcagctgtggaaacacca (reverse)	ccttgcacacaccaagccta
		tcagctgtggaaacacca
−1276 to −308	gtgtcaacagagtgctttac (forward)	gtgtcaacagagtgctttac
	ccttgcacacaccaagccta (reverse)	
−630 to +33	caagctggtgagcttcctga (forward)	gaagctggtgagcttcctga
	atcgtgagaggctgaaccac (reverse)	
Exon 11		
	ctgagcctttgttgatgagc (forward)	ctgagcctttgttgatgagc
	cctcaaaattctgccattcc (reverse)	

Further sequencing data available on request.
[a] Position relative to first base of start codon (+1).

order to create a restriction site for *Bsp*HI (New England Biolabs, UK), which cuts the common allele. The primers CAAGCTGGTGAGCTTCCTGA (forward) ATCGTGAGAGGCTGAAC (reverse) amplify a 663-bp product across the −274 SNP; *Bss*SI (New England Biolabs, UK), cuts the common allele. Statistics were generated by logistic regression using STATA software (Stata Corporation, Texas, USA).

3. Results

The three published SNPs shown to be associated with CD in a number of populations were absent in this Gambian population sample using both Invader and RFLP genotyping methods. We proceeded to sequence exon 11 in a subset of 40 PTB cases to confirm the absence of 2936insC; no other SNPs were revealed in this exon. The absence of known coding SNPs and the lack of new exon 11 variants in Gambians led us to identify novel promoter SNPs to investigate the role of variation in *NOD2* in susceptibility to TB. We identified a total of 8 SNPs in this region, 5 of which are present in the NCBI database (rs2066850, −926G>A; rs4785224, −709A>G; rs5743262, −363C>G; rs5743264, −133C>T and rs5743266, −59A>G). The other three SNPs comprise −162T>C, −197G>A and −274C>T. Four of these SNPs had minor allele frequencies of more than 4% (−709A>G, −133C>T, −274C>T and −59A>G). We successfully genotyped 2 of these SNPs (−709 and −274, Table 2) and found no significant difference in allele frequencies between PTB cases and controls. *Using TRANSFAC®* (http://www.gene-regulation.com/pub/databases.htmlslash#transfac), we were unable to identify any putative transcription factor binding sites involving these SNPs. Both SNPs were in Hardy–Weinberg equilibrium. None of the haplotypes generated were associated with susceptibility to PTB.

4. Discussion

NOD2 is an integral part of innate immune responses, acting as a detector for intracellular pathogen-derived molecules. It is expressed in monocytes, and activates the Nuclear Factor-κB (NF-κB) signaling pathway to induce inflammatory responses including the production of tumor necrosis factor. Since this pathway is important in mycobacterial immunity, we set out to determine whether the three functional *NOD2* variants known to be associated with CD were associated with TB. However, all three SNPs were absent in this Gambian population sample. We therefore sequenced the promoter region of the gene and identified a number of other SNPs within our population. We have failed to identify any association between two novel NOD2 promoter variants and PTB in a Gambian population sample.

There are a number of possible interpretations of this negative finding. Firstly, variation in *NOD2* may contribute to PTB susceptibility, but the effect is small and a much larger sample size is required to detect it. Secondly, this study has focused only on two SNPs in the promoter region and we therefore cannot exclude a causal variant elsewhere in the gene. The rationale for this approach was that the coding region variants associated with CD were not present in our populations, and sequencing of exon 11 and its intronic boundaries did not reveal any polymorphisms. Generally, there are more SNPs per kilobase of DNA in the promoter region of a gene compared to coding regions, and SNPs in this region may have more subtle effects on gene function that could underlie susceptibility to disease [19]. Meanwhile, the majority of the less frequent coding SNPs result in silent or conservative changes of little functional significance. Finally, it is possible that *NOD2* influences susceptibility to TB, but not PTB. We have investigated adults with pulmonary disease, which is the most important form of the disease both in terms of

Table 2
Allele and haplotype frequencies for the −709 and −274 SNPs in TB cases and controls

Allele frequencies[a]	TB cases		TB controls		*P*-value[c]
	Number[b]	%	Number	%	
-709^{+}	320	59.8	251	61.2	ns
-274^{+}	270	93.1	265	93.6	ns
Haplotype frequencies[d]					
$-709^{+}, -274^{+}$	263	54.4	265	54.8	ns
$-709^{+}, -274^{-}$	189	39.0	187	38.6	ns
$-709^{-}, -274^{+}$	28	5.7	29	6.0	ns
$-709^{-}, -274^{-}$	4	0.9	3	0.6	ns

[a] + denotes wildtype: − denotes NOD2 variant at SNP −709 and −274. Both SNPS are in Hardy–Weinberg equilibrium.
[b] Discrepancies arise as a result of genotyping failure.
[c] Chi squared 2×2 and 4×2 contingency tables for allele and haplotype frequencies, respectively.
[d] $r^2 = 0.0144$. $D' = 0.5671$, a measure of linkage disequilibrium between the two SNPs, where a value of 1 indicates complete LD.

the numbers affected and disease transmission. Whilst the NF-κB pathway is important in this clinical phenotype, we have not investigated the role of *NOD2* in abdominal TB, which though much less common, is phenotypically more similar to CD.

The absence of CD-associated SNPs in Gambians the apparent resistance to CD in African populations. However, the known polymorphisms account for about 20% of the risk of CD, indicating that other factors, both genetic and environmental also contribute to disease susceptibility. Thus differences in diet or exposure to pathogens could also contribute to the lower incidence of CD in Africa. Interestingly, the three variants of NOD2 associated with susceptibility to CD were absent in Japanese and Chinese population samples, which also have a low incidence of CD [20,21]. From a population genetics perspective, the absence of CD-associated SNPs in *NOD2* in this and other African populations [22] suggests that the mutation occurred after the out-of-Africa expansion of *Homo sapiens*. Their presence in Caucasians but not Asians may have resulted from genetic drift. Alternatively, infection is a powerful selective force in human evolution and *NOD2* is a germline-encoded innate immune response gene, providing the first line of host defense. It is possible to speculate that *NOD2* variants, which afforded a selective advantage to acute bacterial pathogens in the past, now predispose to the development of chronic inflammation in a more hygienic environment. More detailed genetic analysis of the *NOD2* locus in different populations, including the role of these novel promoter SNPs in susceptibility to other infectious diseases in The Gambia, is required to address this possibility.

Acknowledgements

This work was supported by the Medical Research Council (UK). A.A. was supported by the Sir Halley Stewart Trust. We thank all study subjects for their participation. Musa Jawo and Gambia Sowe provided expert assistance in the field. Mr. Ebou Sarr, at the Blood Transfusion Centre, Royal Victoria Hospital, Banjul, provided assistance with collection of control samples. Sarah Nutland and Helen Rance assisted with DNA quantification, Chris Lowe with DNA sequencing, and Debbie Smyth with Invader genotyping. MJN holds a wellcome Trust advanced Clinical Fellowship.

References

[1] Bellamy, R. (2003) Susceptibility to mycobacterial infections. Genes Immun. 4, 4–11.
[2] Fine, P.E.M. (1993) Variation in protection by BCG: implications of and for heterologous immunity. Lancet 346, 1339–1345.
[3] Dye, C., Espinal, M.A., Watt, C.J., Mbiaga, C. and Williams, B.G. (2002) Worldwide incidence of multidrug-resistant tuberculosis. J. Infect. Dis. 185, 1197–1202.
[4] Hermon-Taylor, J. (2001) Protagonist: *Mycobacterium avium* subspecies paratuberculosis is a cause of Crohn's disease. Gut 49, 755–756.
[5] Quirke, P. (2001) Antagonist: *Mycobacterium avium* subspecies paratuberculosis is a cause of Crohn's disease. Gut 49, 757–760.
[6] Mishina, D., Katsel, P., Brown, S.T., Gilberts, E.C. and Greenstein, R.J. (1996) On the etiology of Crohn disease. Proc. Natl. Acad. Sci. USA 93, 9816–9820.
[7] Schwartz, D., Shafran, I., Romero, C., Piromalli, C., Biggerstaff, J. and Naser, N., et al. (2000) Use of short-term culture for identification of *Mycobacterium avium* subspecies paratuberculosis in tissue from Crohn's disease patients. Clin. Microbiol. Infect. 6, 303–307.
[8] Hulten, K., Almashhrawi, A., El-Zaatari, F.A. and Graham, D.Y. (2000) Antibacterial therapy for Crohn's disease: a review emphasizing therapy directed against mycobacteria. Dig. Dis. Sci. 45, 445–456.
[9] Hume, G. and Radford-Smith, G.L. (2002) The pathogenesis of Crohn's disease in the 21st century. Pathology 34, 561–567.
[10] Hugot, J.-P., Chamaillard, M., Zouali, H., Lesage, S., Cezard, J.-P. and Belaiche, J., et al. (2001) Association of NOD2 leucine-rich repeat variants with susceptibility to Crohn's disease. Nature 411, 599–603.
[11] Ogura, Y., Bonen, D.K., Inohara, N., Nicolae, D.L., Chen, F.F. and Ramos, R., et al. (2001) A frameshift mutation in NOD2 associated with susceptibility to Crohn's disease. Nature 411, 603–606.
[12] Hampe, J., Cuthbert, A., Croucher, P., Mirza, M.M., Mascheretti, S. and Fisher, S., et al. (2001) Association between insertion mutation in NOD2 gene and Crohn's disease in German and British populations. Lancet 357, 1925–1928.
[13] Dixon, M.S., Golstein, C., Thomas, C.M., van der Biezen, E.A. and Jones, J.D. (2000) Genetic complexity of pathogen perception by plants: the example of Rcr3, a tomato gene required specifically by Cf-2. Proc. Natl. Acad. Aci. USA 97, 8807–8814.
[14] Bonen, D.K., Ogura, Y., Nicolae, D.L., Inohara, N., Saab, L. and Tanabe, T., et al. (2003) Crohn's disease-associated NOD2 variants share a signalling defect in response to lipopolysaccharide and peptidoglycan. Gastroenterology 124, 140–146.
[15] Bellamy, R. and Hill, A.V.S. (1998) The NRAMP1 gene and susceptibility to tuberculosis. N. Engl. J. Med. 339, 199–201.
[16] Dudbridge, F. (2003) Pedigree disequilibrium tests for mulitlocus haplotypes. Genet. Epidemiol. 25, 115–121.
[17] Mein, C.A., Barratt, B.J., Dunn, M.G., Seigmund, T., Smith, A.N. and Esposito, L., et al. (2000) Evaluation of single nucleotide polymorphism (SNP) typing with Invader on PCR amplicons and its automation. Genome Res. 10, 330–345.
[18] Cavanaugh, J.A., Adams, K.E., Quak, E.J., Bryce, M.E., O'Callaghan, N.J. and Rodgers, H.J., et al. (2003) CARD15/NOD2 risk alleles in the development of Crohn's disease in the Australian population. Ann. Hum. Genet. 67, 35–41.
[19] Salisbury, B.A., Pungliya, M., Choi, J.Y., Jiang, R., Sun, X.J. and Stephens, J.C. (2003) SNP and haplotype variation in the human genome. Mutat. Res. 526, 53–61.
[20] Inoue, N., Tamura, K., Kinouchi, Y., Fukuda, Y., Takahashi, S. and Ogura, Y., et al. (2002) Lack of common NOD2 variants in Japanese patients with Crohn's disease. Gastroenterology 123, 86–91.
[21] Leong, R.W., Armuzzi, A., Ahmad, T., Wong, M.L., Tse, P. and Jewell, D.P., et al. (2003) NOD2/CARD15 gene polymorphisms and Crohn's disease in the Chinese population. Aliment. Pharmacol. Ther. 17, 1465–1470.
[22] Marsh, S. and McLeod, H.L. (2003) Crohn's disease: ethnic variation in CARD15 genotypes. Gut 52, 770.