European Journal of Human Genetics (2009) 17, 1304 – 1308

www.nature.com/ejhg

ARTICLE

Specific association of a *CLEC16A/KIAA0350* polymorphism with *NOD2/CARD15*$^{-}$ Crohn's disease patients

Ana Márquez[1,6], Jezabel Varadé[1,6], Gema Robledo[2,6], Alfonso Martínez[1], Juan Luis Mendoza[3], Carlos Taxonera[3], Miguel Fernández-Arquero[1], Manuel Díaz-Rubio[3], María Gómez-García[4], Miguel Angel López-Nevot[5], Emilio G. de la Concha[1], Javier Martín[2,6] and Elena Urcelay*[,1,6]

[1]*Department of Immunology, Hospital Clínico San Carlos, Madrid, Spain;* [2]*Instituto de Parasitología y Biomedicina López Neyra (CSIC), Granada, Spain;* [3]*Department of Digestive Diseases, Hospital Clínico San Carlos, Madrid, Spain;* [4]*Department of Digestive Diseases, Hospital Virgen de las Nieves, Granada, Spain;* [5]*Department of Immunology, Hospital Virgen de las Nieves, Granada, Spain*

Independent genome-wide association studies highlighted the function of *CLEC16A/KIAA0350* polymorphisms modifying the risk to either multiple sclerosis (rs6498169) or type 1 diabetes (rs2903692). This C-type lectin gene maps to a linkage disequilibrium block at 16p13 and a functional role of this gene could be envisaged for other immune-related conditions, such as inflammatory bowel disease (IBD). The present study, aimed at investigating the association of those two polymorphisms with IBD, included 720 IBD patients and 550 ethnically matched healthy controls. The effect of rs2903692 previously described in diabetes was observed specifically for Crohn's disease (CD) patients lacking the main susceptibility factor described to date, that is, three polymorphisms within another pattern recognition gene, *NOD2/CARD15* (*NOD2*$^{-}$ *vs* *NOD2*$^{+}$ CD patients, G *vs* A: $P = 0.008$; OR (95% CI) = 1.54 (1.10–2.15); *NOD2*$^{-}$ CD patients *vs* controls: $P = 0.008$; OR (95% CI) = 1.37 (1.08–1.73)). Replication of these findings was performed in independent Spanish cohorts of 544 IBD patients and 340 controls and the combined data yielded significant differences (405 *NOD2*$^{-}$ *vs* 204 *NOD2*$^{+}$ CD patients, G *vs* A: $P = 0.0012$; $OR_{M\text{-}H}$ (95% CI) = 1.49 (1.17–1.90); *NOD2*$^{-}$ CD patients *vs* controls: $P = 0.0007$; $OR_{M\text{-}H}$ (95% CI) = 1.35 (1.13–1.60)). The pooled analysis of the ulcerative colitis patients *vs* controls also yielded a significant risk ($P = 0.0005$; OR (95% CI) = 1.52 (1.19–1.93)). These data would suggest that microbial recognition through different pathways seems to converge in the development of these polygenic bowel diseases.
European Journal of Human Genetics (2009) **17,** 1304–1308; doi:10.1038/ejhg.2009.50; published online 1 April 2009

Keywords: inflammatory bowel disease; *CLEC16A/KIAA0350*; *NOD2/CARD15*; polymorphisms and genetic susceptibility

Introduction

The environment, the patient's genetic profile, the intestinal microbiota and the immune system are key components involved in the generation of the chronic inflammatory reaction that characterizes inflammatory bowel diseases (IBDs). The balance between the commensal flora and host defensive responses seems to be critical in the pathogenesis of IBD. Evidence in support of this

*Correspondence: Dr E Urcelay, Department of Immunology, Hospital Universitario Clínico San Carlos, Martin Lagos s/n, 28040 Madrid, Spain. Tel: +34 91 330 3347; Fax: 34 91 330 3344;
E-mail: eurcelay.hcsc@salud.madrid.org
[6]These authors contributed equally to this work.
Received 16 April 2008; revised 23 February 2009; accepted 3 March 2009; published online 1 April 2009

hypothesis has been initially derived from several observations: the therapeutic benefit provided by antibiotic treatment, the enrichment of harmful bacteria (ie *Escherichia coli*) in patients compared with healthy subjects and the flora-dependent induction of colitis in murine strains. The simultaneous discovery of *NOD2/CARD15* as a susceptibility gene for Crohn's disease (CD) by two groups using positional cloning and candidate gene approaches[1,2] conferred specific support for the long-held theory that a genetically dysregulated host immune response to luminal bacteria results in CD. Muramyl dipeptide, a component of bacterial peptidoglycan, is recognized by the NOD2 receptor; nonetheless, the exact mechanism by which *NOD2* polymorphisms contribute to the increased propensity to develop CD remains incompletely understood. Recent data show that phagocytosis and subsequent bacterial killing and degradation can make ligands available to cytoplasmic Nod2.[3]

Genome-wide association scans for susceptibility loci have been successful in identifying genes modulating disease risk. These discoveries have revealed the importance of a 233 kb linkage disequilibrium block on chromosome 16p13 comprising the *CLEC16A/KIAA0350* gene in type 1 diabetes (T1D) and multiple sclerosis (MS) susceptibility.[4–7] The *CLEC16A* gene contains a C-type lectin domain and, although of unknown function yet, the encoded protein has been detected in immune cells.[8] C-type lectins that contain carbohydrate recognition domains have been shown to participate in the recognition of carbohydrates at cell surfaces, on circulating proteins as well as those present on pathogens and are consequently involved in diverse processes including plasma glycoprotein turnover, cell–cell adhesion and innate pathogen recognition.[9]

Receptors that recognize microbial molecular patterns are critical to innate immunity. In the past few years, mutations in Toll- and Nod-like receptors have been found associated with IBD.[10–12] Among those receptors, which evolved as a way to limit infection by invasive organisms, several C-type lectin-like molecules have been identified. The prevailing theory describes that microbes present a wide variety of surface antigens and metabolic products that may be recognized by the innate immune system, and as a general rule, no single innate immune receptor is ever likely to be the sole mediator of activation of the immune response.[13]

Given the potential implication of the *CLEC16A* gene in pathogen recognition, we aimed at investigating this gene as a predisposition factor for other immune-mediated polygenic diseases such as IBDs. A number of genes associated with CD, like *IL23R*, are also associated with other autoimmune disorders, indicative of common underlying mechanisms. Given the putative functional parallelism of the proteins encoded by the *CLEC16A* gene and the well-known susceptibility factor for CD, *NOD2/CARD15*, as molecules involved in pathogen recognition, we proceeded to check how *CLEC16A* polymorphisms may affect the *NOD2*-stratified subgroups. Two polymorphisms within the *CLEC16A* gene found previously associated either with T1D (rs2903692) or MS (rs6498169) were selected to perform a case–control study in our population. The better knowledge of the different genetic variants contributing to the overall risk will improve the definition of disease phenotypes.

Patients and methods

Patients and controls

The original study group consisted of 720 IBD patients consecutively recruited from one center (Hospital Clínico San Carlos, Madrid, Spain). A group of 550 healthy unrelated subjects from Madrid (mainly hospital employees and blood donors) were selected as controls. The replication cohorts (554 IBD patients and 340 controls) were recruited from Hospital Virgen de las Nieves (Granada, Spain). Diagnosis of IBD patients was based on standard clinical, radiologic, endoscopic and histologic criteria.[14] Cases and controls were all white Spanish (most of them described by Oliver *et al*[15]) and were included in this study after written informed consent. The ethics committee of Hospital Clínico (Madrid) approved the study.

CLEC16A/KIAA0350 and *NOD2/CARD15* polymorphisms

The *CLEC16A* (rs2903692 and rs6498169) and *NOD2/CARD15* (rs2066844, rs2066845 and rs22066847) polymorphisms were genotyped using TaqMan assays from Applied Biosystems Inc. (Foster City, CA, USA) following manufacturer's instructions and analyzed in an ABI 7900HT Fast Real-Time PCR System (Applied Biosystems Inc.).

Statistical analysis

Allele and genotype frequencies in patients and controls were compared by χ^2-test; *P*-values were considered significant at a level of <0.05. Odds ratio (OR) and *P*-values were calculated using a standard package (Epi Info version 6.02; CDC, Atlanta, GA, USA).

Meta-analysis to combine ORs across the sample sets used fixed effects of Mantel–Haenszel method.

Results

Both *CLEC16A* polymorphisms conformed to Hardy–Weinberg expectations in our collections. Table 1 shows the genotype and allele frequencies of the two tested *CLEC16A* polymorphisms in IBD patients. Similar genotypic distribution was found for either clinical phenotype, CD or ulcerative colitis (UC).

As indicated, a trend for association of rs2903692 with IBD patients was observed ($P=0.077$). When CD patients

Table 1 Genotype and allele frequencies of the *CLEC16A* polymorphisms in Spanish patients and controls (CD patients stratified by *NOD2/CARD15* status denoted as *NOD2*$^{-}$: wild type, and *NOD2*$^{+}$: heterozygous for R702W, G908R or L1007fs, homozygous or compound heterozygous)

	CLEC16A rs2903692					*CLEC16A rs6498169*				
Subjects	*GG*	*GA*	*AA*	*G*	*A*	*AA*	*AG*	*GG*	*A*	*G*
Controls ($n=550$)	187 (0.35)	247 (0.46)	98 (0.18)	621 (0.58)	443 (0.42)	258 (0.47)	227 (0.41)	63 (0.11)	743 (0.68)	353 (0.32)
IBD patients ($n=720$)	276 (0.38)	335 (0.47)	106 (0.15)	887 (0.62)	547 (0.38)[a]	340 (0.47)	298 (0.42)	79 (0.11)	978 (0.68)	456 (0.32)
UC patients ($n=370$)	142 (0.39)	170 (0.46)	55 (0.15)	454 (0.62)	280 (0.38)	181 (0.49)	152 (0.41)	36 (0.10)	514 (0.70)	224 (0.30)
CD patients ($n=350$)	134 (0.38)	165 (0.47)	51 (0.15)	433 (0.62)	267 (0.38)	159 (0.46)	146 (0.42)	43 (0.12)	464 (0.67)	232 (0.33)
NOD2^{-} CD ($n=220$)	92 (0.42)	105 (0.48)	23 (0.10)	289 (0.66)	151 (0.34)[b]	92 (0.42)	97 (0.44)	30 (0.14)	281 (0.64)	157 (0.36)
NOD2^{+} CD ($n=121$)	39 (0.32)	56 (0.46)	26 (0.22)	134 (0.55)	108 (0.45)	61 (0.51)	47 (0.39)	12 (0.10)	169 (0.70)	71 (0.30)

Abbreviations: IBD, inflammatory bowel disease; CD, Crohn's disease; UC, ulcerative colitis.
[a]G *vs* A, IBD *vs* controls: $P=0.077$; OR (95% CI) = 1.16 (0.98–1.36).
[b]G *vs* A, *NOD2*$^{-}$ CD *vs* *NOD2*$^{+}$ CD: $P=0.008$; OR (95% CI) = 1.54 (1.10–2.15).
G *vs* A, *NOD2*$^{-}$ CD *vs* controls: $P=0.008$; OR (95% CI) = 1.37 (1.08–1.73).

Table 2 Genotype and allele frequencies of the *CLEC16A* rs2903692 polymorphisms in patients and controls in other independent Spanish cohort (CD patients stratified by *NOD2/CARD15* status denoted as *NOD2*$^{-}$: wild type, and *NOD2*$^{+}$: heterozygous for G908R, R702W or L1007fs, homozygous or compound heterozygous)

	CLEC16A rs2903692				
Subjects	*GG*	*GA*	*AA*	*G*	*A*
Controls ($n=340$)	98 (0.29)	175 (0.52)	65 (0.19)	371 (0.55)	305 (0.45)
IBD patients ($n=544$)	200 (0.27)	271 (0.50)	69 (0.13)	671 (0.62)	409 (0.38)[a]
UC patients ($n=270$)	112 (0.42)	121 (0.46)	33 (0.12)	345 (0.65)	187 (0.35)
CD patients ($n=274$)	88 (0.32)	150 (0.55)	36 (0.13)	326 (0.59)	222 (0.41)
NOD2^{-} CD ($n=185$)	65 (0.35)	98 (0.53)	22 (0.12)	228 (0.62)	142 (0.38)[b]
NOD2^{+} CD ($n=83$)	19 (0.23)	50 (0.60)	14 (0.17)	88 (0.53)	78 (0.47)

Abbreviations: IBD, inflammatory bowel disease; CD, Crohn's disease; UC, ulcerative colitis.
[a]G *vs* A, IBD *vs* controls: $P=0.003$; OR (95% CI) = 1.35 (1.10–1.65).
[b]G *vs* A, *NOD2*$^{-}$ CD *vs* *NOD2*$^{+}$ CD: $P=0.06$; OR (95% CI) = 1.42 (0.97–2.1).
G *vs* A, *NOD2*$^{-}$ CD *vs* controls: $P=0.035$; OR (95% CI) = 1.32 (1.01–1.72).

were stratified by the most important CD genetic risk factor, the presence of three *NOD2/CARD15* polymorphisms, a significant difference was observed between *NOD2*$^{-}$ *vs* *NOD2*$^{+}$ CD patients, indicative of the difference between *NOD2*$^{-}$ patients and healthy controls (see Table 1). The effect of the minor allele of rs2903692 in this subset of patients moves in parallel to the originally described for T1D.[5]

No significant difference was detected for the other tested polymorphism, rs6498169, which is not surprising given that only a limited correlation exists between both markers ($r^2=0.26$), even though they map in the same linkage disequilibrium block ($D'=1$).

We aimed at replicating these findings in independent Spanish cohorts of 544 IBD patients and 340 healthy ethnically matched controls. As summarized in Table 2, the results obtained in the original cohort were validated. In the replication cohort, a significant difference was reached for the association with overall IBD ($P=0.003$; OR (95% CI) = 1.35 (1.10–1.65)); moreover, a highly significant effect on UC susceptibility was evidenced ($P=0.0005$; OR (95% CI) = 1.52 (1.19–1.93)). Again, the difference between *NOD2*$^{-}$ patients and controls was observed ($P=0.035$; OR (95% CI) = 1.32 (1.01–1.72)).

A Mantel–Haenszel analysis of the combined data from both cohorts yielded a significant difference between *NOD2*$^{-}$ patients and healthy controls ($P=0.0007$; $OR_{M\text{-}H}=$ 1.35 (1.13–1.60)) whereas no significant difference was detected for *NOD2*$^{+}$ patients *vs* controls ($P=0.35$; $OR_{M\text{-}H}=$ 0.90 (0.73–1.12)). The intracase subphenotypic analysis was therefore highly significant (*NOD2*$^{-}$ *vs* *NOD2*$^{+}$ patients: $P=0.0012$; $OR_{M\text{-}H}=1.49$ (1.17–1.90)). The pooled analysis of the UC patients *vs* controls also yielded a significant risk ($P=0.0005$; OR (95% CI) = 1.52 (1.19–1.93)).

Discussion

Genome-wide association studies performed during the past months are helping to unravel the pathogenesis of several inflammatory and immune-mediated conditions. Among them, two reports dealing with T1D patients evidenced association with a linkage disequilibrium block containing the *CLEC16A* gene in 16p13.[5,6] Different polymorphisms within this gene were tested and, interestingly, homozygosis for rs2903692 was found in a cell line that presented higher expression of *CLEC16A*,[5] arguing in

favor of a functionality of this variant. Simultaneously, a genome-wide search for MS susceptibility factors pointed to the same gene, where another polymorphism, rs6498169, was shown to increase MS predisposition.[4] The investigation of the two aforementioned polymorphisms in IBD has provided us with evidence in support of the association of the putatively functional SNP in the *CLEC16A* gene with the subset of CD patients deprived of the main susceptibility factor described for Caucasian populations: the *NOD2/CARD15* polymorphisms (see Tables 1 and 2). In fact, NOD2 is an intracellular sensor of peptidoglycan, a function that putatively resembles that of the protein encoded by the *CLEC16A* gene, inasmuch as it could be involved in pathogen recognition according to its C-lectin domain. Concordantly, a functional variant in the promoter of another member of the C-type lectin receptor superfamily, DC-SIGN (dendritic-cell-specific ICAM3-grabbing nonintegrin) has also been shown to have a role in predisposition toward certain HLA phenotypes of either celiac disease[16] or UC patients,[17] indicative of the importance of pattern recognition processes in IBD risk. Therefore, either this *CLEC16A* polymorphism or some other in strong linkage disequilibrium with it would be responsible for the observed effect on NOD2$^-$ CD patients. The *CLEC16A* gene is flanked by the suppressor of cytokine signaling 1, *SOCS1* gene, which is expressed in inflamed tissues in an experimental colitis model and regulates the development of intestinal inflammation.[18,19] It has been reported that *SOCS1* impairs Toll-like receptor signaling, which as already mentioned has key functions in innate and adaptive immunity by detecting microbial pathogens.[20] Various microbial pathogens use the host's inhibitory SOCS proteins for manipulating cytokine receptor signaling, a helpful strategy to evade otherwise detrimental immune responses. Interestingly, located upstream of the *CLEC16A* gene is the human class II transactivator *MHC2TA* gene, and a haplotype conformed by two polymorphisms in the latter (rs3087456 and rs4774) was recently found associated with rheumatoid arthritis and MS.[21] Although no overall association with IBD was detected, stratification of CD patients by *NOD2/CARD15* polymorphisms could evidence an effect.

The obvious question that arises is why this locus was overlooked in the recent largest GWA study for CD.[22] The possible explanation being that as the observed effect involves only a subgroup of patients (*NOD2*$^-$ patients) the effect is blurred when the whole cohort of CD patients is investigated. Our own results in the original cohort support this explanation (overall IBD $P=0.077$) and, although rs2903692 was not analyzed in this GWA study, a good proxy for it (rs7203150, $r^2=0.74$) also showed a similar trend with diminished minor allele frequency in CD patients compared to controls (overall CD $P=0.14$). Interestingly, from the analysis of the replication cohort, it looks as if the association extends beyond the original *NOD2*$^-$ CD group to the entire IBD population, or in fact with even higher statistical significance to the UC subgroup, which coincidentally resembles the *NOD2*$^-$ phenotype. This variability could reflect the limited sample size of the study; nonetheless, it would hopefully prompt others to verify this association in other, larger cohorts of patients. Even though this gene has not been identified in GWA scans, coverage by the genotyping platforms used is incomplete, and the power of most of them was insufficient to detect modest effects, so a second tier of loci that require further assessment is still on the horizon.[23]

In conclusion, concordantly with the already reported function of this locus in T1D and MS, our study shows for the first time the association of a *CLEC16A* polymorphism with a distinct group of *NOD2*$^-$ CD patients. This new evidence argues in favor of the hypothesis that alterations in genetic factors involved in bacterial recognition through different pathways result in CD, which may allow more insight into processes critical to the pathogenesis of this IBD.

Acknowledgements

Alfonso Martínez holds a research contract of the Spanish Health Ministry (04/00175). Jezabel Varadé is a fellow and Elena Urcelay works for the Fundación para la Investigación Biomédica-Hospital Clínico San Carlos. This work was supported by grants from Fundación Mutua Madrileña, Fundación Ramón Areces, FIS PI07/0369, FIS PI07/0353, SAF2006-00398 and CTS-1180.

References

1 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.
2 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.
3 Herskovits AA, Auerbuch V, Portnoy DA: Bacterial ligands generated in a phagosome are targets of the cytosolic innate immune system. *PLoS Pathog* 2007; **3**: e51.
4 Hafler DA, Compston A, Sawcer S *et al*: Risk alleles for multiple sclerosis identified by a genomewide study. *N Engl J Med* 2007; **357**: 851–862.
5 Hakonarson H, Grant SF, Bradfield JP *et al*: A genome-wide association study identifies *KIAA0350* as a type 1 diabetes gene. *Nature* 2007; **448**: 591–594.
6 Todd JA, Walker NM, Cooper JD *et al*: Robust associations of four new chromosome regions from genome-wide analyses of type 1 diabetes. *Nat Genet* 2007; **39**: 857–864.
7 Zoledziewska M, Costa G, Pitzalis M *et al*: Variation within the *CLEC16A* gene shows consistent disease association with both multiple sclerosis and type 1 diabetes in Sardinia. *Genes Immun* 2008; **10**: 15–17.
8 Robinson MJ, Sancho D, Slack EC, LeibundGut-Landmann S, Reis e Sousa C: Myeloid C-type lectins in innate immunity. *Nat Immunol* 2006; **7**: 1258–1265.
9 McGreal EP, Martinez-Pomares L, Gordon S: Divergent roles for C-type lectins expressed by cells of the innate immune system. *Mol Immunol* 2004; **41**: 1109–1121.
10 Franchimont D, Vermeire S, El Housni H *et al*: Deficient host–bacteria interactions in inflammatory bowel disease? The toll-like

receptor (TLR)-4 Asp299gly polymorphism is associated with Crohn's disease and ulcerative colitis. *Gut* 2004; **53**: 987–992.
11 Pierik M, Joossens S, Van Steen K *et al*: Toll-like receptor-1, -2, and -6 polymorphisms influence disease extension in inflammatory bowel diseases. *Inflamm Bowel Dis* 2006; **12**: 1–8.
12 Torok HP, Glas J, Tonenchi L, Mussack T, Folwaczny C: Polymorphisms of the lipopolysaccharide-signaling complex in inflammatory bowel disease: association of a mutation in the Toll-like receptor 4 gene with ulcerative colitis. *Clin Immunol* 2004; **112**: 85–91.
13 Underhill DM: Collaboration between the innate immune receptors dectin-1, TLRs, and Nods. *Immunol Rev* 2007; **219**: 75–87.
14 Lennard-Jones JE: Classification of inflammatory bowel disease. *Scand J Gastroenterol Suppl* 1989; **170**: 2–6; discussion 16–19.
15 Oliver J, Marquez A, Gomez-Garcia M *et al*: Association of the macrophage migration inhibitory factor gene polymorphisms with inflammatory bowel disease. *Gut* 2007; **56**: 150–151.
16 Nunez C, Rueda B, Martinez A *et al*: A functional variant in the CD209 promoter is associated with DQ2-negative celiac disease in the Spanish population. *World J Gastroenterol* 2006; **12**: 4397–4400.
17 Nunez C, Oliver J, Mendoza JL *et al*: CD209 in inflammatory bowel disease: a case–control study in the Spanish population. *BMC Med Genet* 2007; **8**: 75.
18 Chinen T, Kobayashi T, Ogata H *et al*: Suppressor of cytokine signaling-1 regulates inflammatory bowel disease in which both IFNgamma and IL-4 are involved. *Gastroenterology* 2006; **130**: 373–388.
19 Inagaki-Ohara K, Sasaki A, Matsuzaki G *et al*: Suppressor of cytokine signalling 1 in lymphocytes regulates the development of intestinal inflammation in mice. *Gut* 2006; **55**: 212–219.
20 Mansell A, Smith R, Doyle SL *et al*: Suppressor of cytokine signaling 1 negatively regulates Toll-like receptor signaling by mediating Mal degradation. *Nat Immunol* 2006; **7**: 148–155.
21 Martinez A, Sanchez-Lopez M, Varade J *et al*: Role of the *MHC2TA* gene in autoimmune diseases. *Ann Rheum Dis* 2007; **66**: 325–329.
22 Wellcome Trust Case Control Consortium: Genome-wide association study of 14 000 cases of seven common diseases and 3000 shared controls. *Nature* 2007; **447**: 661–678.
23 Mathew CG: New links to the pathogenesis of Crohn disease provided by genome-wide association scans. *Nat Rev Genet* 2008; **9**: 9–14.