ARTICLE

Association of the organic cation transporter OCTN genes with Crohn's disease in the Spanish population

Alfonso Martínez[1], Maria del Carmen Martín[1], Juan L Mendoza[2], Carlos Taxonera[2], Manuel Díaz-Rubio[2], Emilio G de la Concha[1] and Elena Urcelay*,[1]

[1]*Department of Clinical Immunology, Hospital Clinico San Carlos, Madrid, Spain;* [2]*Department of Gastroenterology, Hospital Clinico San Carlos, Madrid, Spain*

The *SLC22A4* and *SLC22A5* genes within the IBD5 risk locus encode the organic cation transporters OCTN1 and OCTN2. Two variants, 1672C>T in *SLC22A4* and −207G>C in *SLC22A5*, were shown to alter these genes' functions and were identified as genetic susceptibility factors for Crohn's disease (CD). We pursued to check both putative etiologic variants in an independent population through a case–control study with 309 Spanish CD patients and 408 ethnically matched healthy subjects. Both polymorphisms were found in partial linkage disequilibrium ($D' = 0.86$). The separate analysis of each OCTN variant evidenced no association. However, when the simultaneous presence of mutant variants in both genes was analyzed, an effect on CD susceptibility was observed ($P = 0.026$, odds ratio (OR) (95% confidence interval (CI)) = 1.59 (1.03–2.45)). The previously described predisposition conferred by the 5q31-risk haplotype increased in the absence of the etiologic 1672T and −207C alleles ($P = 0.0006$, OR (95% CI) = 10.14 (1.97–98.04)). Moreover, the risk contributed by these polymorphisms was higher in the IBD5 wild-type population ($P = 0.003$, OR (95% CI) = 2.65 (1.32–5.35)), arguing against the exclusive etiological role of the OCTN variants. The haplotype pattern inferred led to the consideration of these variants as susceptibility markers only in a defined genetic context. Our data support the interpretation of the 1672C>T *SLC22A4* and −207G>C *SLC22A5* polymorphisms as genetic markers of susceptibility/protection haplotypes.
European Journal of Human Genetics (2006) **14**, 222–226. doi:10.1038/sj.ejhg.5201529; published online 7 December 2005

Introduction

Crohn's disease (CD) is a complex genetic disorder with several susceptibility loci.[1] One of these loci mapped to human chromosome 5q31 (IBD5), where a common haplotype marked by 11 single-nucleotide polymorphisms (SNPs) in total linkage disequilibrium (LD) and spanning 250 kb in the cytokine gene cluster showed clear association with the disease.[2] This association was further supported by a murine model of colitis presenting an IBD-susceptibility locus in the syntenic region on mouse chromosome 11.[3] Initially, the strong LD within this 5q31 region precluded the identification of the CD etiological gene(s). More recently, resequencing the five genes included in that 250-kb locus pinpointed 10 SNPs with potential functional effects.[4] Among them, two SNPs located in the organic cation transporter genes, *SLC22A4* and *SLC22A5*, were considered pathogenic candidates to assess. The first one is a nonsynonymous substitution (1672C>T) causing an amino-acid change L503F in the *SLC22A4* gene and the other is a transversion (−207G>C) disrupting a heat shock element in the promoter region of the *SLC22A5* gene. They were found associated with CD in

*Correspondence: Dr E Urcelay, Department of Immunology, Hospital Clinico San Carlos, C/Martin Lagos sn, Madrid 28040, Spain.
Tel: +34 91 330 3347; Fax: +34 91 330 3344;
E-mail: eurcelay@hcsc.es*

the Canadian population that was originally tested for this extended locus and functional studies of these variants supported the relevance of both genes in CD susceptibility[4] and their influence in the phenotypic expression of CD in this population.[5] However, an ulterior study of these polymorphisms in Japan questioned their etiological role.[6]

We sought additional confirmation to these findings by examining both 1672C>T *SLC22A4* and −207G>C *SLC22A5* SNPs in another data set derived from a different population, white Spanish CD patients. Association of the 250-kb IBD5 locus with CD in our population has been already proven and we corroborated the total LD for two variants, IGR2060a_1G>C and IGR3081a_1T>G.[7]

Epidemiological evidence indicates that genetic factors determine the clinical phenotype in both CD and UC, with high concordance for age of debut, location and behavior of disease in multiple affected families.[8,9] The precise clinical phenotype of a patient depends most probably on the effect of a limited number of genes. Although the etiology of CD is unknown, a breakdown of tolerance to luminal bacteria in the digestive tract of genetically susceptible hosts has been proposed.[10] CD presents a dysregulated response of the intestinal mucosal immune system to innocuous luminal antigens in susceptible individuals. Defects in genes regulating barrier function might act synergistically leading to the characteristically inflammatory process of CD. The interplay of genetic and environmental (microbial) factors could culminate in sustained activation of the immune response probably facilitated by defects in the epithelial barrier. Moreover, being *SLC22A4* and *SLC22A5* genes widely expressed, this would not be the only possible cause of inflammation, but also their altered regulation in macrophages and T cells. The identification of the genes leading to disease and a better understanding of the underlying immunoregulatory abnormalities will be crucial to define different IBD profiles.

Methods
Patients and controls
The study group consisted of 309 unrelated adult white Spanish CD patients (53% women) with a median follow-up of 11.5 years (95% percentile values range from 3.4 to 26.9 years), recruited after informed consent from a single center. Diagnosis of CD was based on Lennard-Jones[11] criteria. Clinical history and personal interviews with patients allowed the assignment of disease phenotype following the Vienna Classification,[12] which groups patients in two subsets according to both location: L1 (terminal ileum), L2 (colonic), L3 (ileocolonic) and L4 (upper gastrointestinal) and behavior: B1 (inflammatory, nonstricturing non-fistulizing), B2 (stricturing) and B3 (fistulizing). Perianal disease was defined by the presence of perianal abscesses, fistulae and/or ulcers.

Patients and data are regularly followed up in the Inflammatory Bowel Disease Unit at Hospital Clínico San Carlos, Madrid. A group of 408 healthy white, unrelated subjects (60% women) from the Madrid region (mainly hospital employees and blood donors) were used as controls. The Hospital Ethical Committee approved this study.

Genotyping
5q31 *locus* Two variants, IGR2060a_1G>C and IGR3081a_1T>G, were independently analyzed in Urcelay *et al*.[7] As they were already found in complete LD, the second one is used as haplotypic maker in the present work.

SLC22A4 *and* SLC22A5 *polymorphisms* *SLC22A4* 1672C>T (in exon 9, L503F of OCTN1, rs1050152) was analyzed following the manufacturer's suggestions by TaqMan Assays-on-Demand (C___3170459_10_) and −207G>C in the *SLC22A5* promoter (OCTN2, rs2631367) by TaqMan Assays-by-Design, both from Applied Biosystems.

Statistical analysis
Case–control analysis was performed by applying the χ^2 statistics or Fisher exact test when necessary. The association between genotype and phenotypic characteristics of CD was estimated by the odds ratio (OR) with 95% confidence interval (CI). The statistical analysis used the Statistical Package for the Social Sciences (SPSS) version 12.0 for Windows (SPSS Inc., Chicago, IL, USA). The groups used in this case–control study (Table 1, 1672C>T: 300 patients/342 controls and −207G>C: 293 patients/402 controls) have >80% power to detect an association conferring a 1.5-fold increase in CD risk (at the 0.05 significance level) for alleles with the observed frequency of 43 and 48%, respectively, and calculated with one-side *P*-value (http://calculators.stat.ucla.edu).

D' was calculated as follows, f(ab)−f(a)f(b)/f(a) (1−f(b)), being f(a) the frequency of the minor allele and f(ab) the frequency of the haplotype containing both alleles.

Haplotypic frequencies were estimated using the Expectation–Maximisation (EM) algorithm[13] implemented in the Arlequin v2.000 software, with number of iterations set at 5000 and initial conditions at 50, with an epsilon value of 10^{-7}.

Results
The two aforementioned SNPs within two nearby *SLC22A4* and *SLC22A5* genes were tested in the Spanish population. Their chromosomal location and those of the previously tested IBD5-haplotype polymorphisms (IGR3081a_1 and IGR2060a_1) are shown in Figure 1. The 1672C>T *SLC22A4* and −207G>C *SLC22A5* polymorphisms were

Table 1 Distribution of 1672C>T SLC22A4 (A), −207G>C SLC22A5 (B) and both polymorphisms (C) between Crohn's disease patients and healthy controls

(A)

1672C>T	CC	CT	TT	C	T
CD patients (n = 300)	86 (28.67%)	152 (50.67%)	62 (20.67%)	324 (54.00%)	276 (46.00%)
Controls (n = 342)	110 (32.16%)	167 (48.83%)	65 (19.01%)	387 (56.58%)	297 (43.42%)

(B)

−207G>C	GG	GC	CC	G	C
CD patients (n = 293)	86 (29.35%)	149 (50.85%)	58 (19.80%)	321 (54.78%)	265 (45.22%)
Controls (n = 402)	116 (28.90%)	186 (46.30%)	100 (24.90%)	418 (51.99%)	386 (48.01%)

(C)

	1672CC and −207GG	1672TX or −207CX	1672TX and −207CX
CD patients (n = 268)	46 (17.16%)	41 (15.30%)	181 (67.54%)
Controls (n = 303)	75 (24.75%)	31 (10.23%)	197 (65.02%)

OCTN mutant vs wild-type individuals: $P = 0.026$, OR (95% CI) = 1.59 (1.03–2.45).
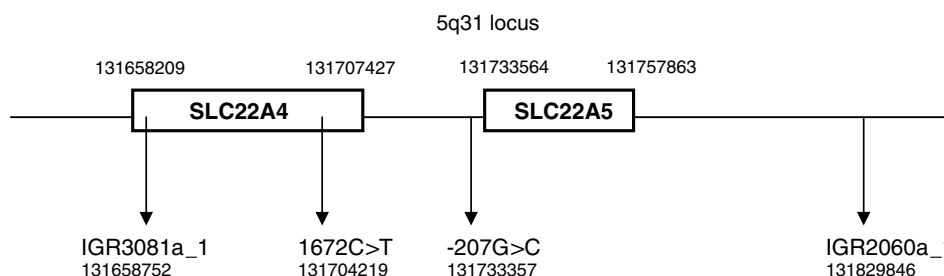
Figure 1 Schematic representation of the chromosomal region 5q31 with the location of the polymorphisms studied.

not in complete LD in our population ($D' = 0.86$). In our control cohort, the LD between both variants was strong but not complete ($P = 2.6 \times 10^{-32}$). Out of 228 control individuals with at least one mutated polymorphism, 197 subjects presented both, four presented the SLC22A4 gene polymorphism only and 27 individuals presented the SLC22A5 promoter variant only.

To check whether both were causative polymorphisms, we aimed at testing them individually. In Spain, neither the polymorphism at 1672C>T in the SLC22A4 gene nor the one at −207G>C in the SLC22A5 promoter showed any association with CD as a whole (Table 1A–B). The combined presence of 1672T SLC22A4 and −207C SLC22A5 mutations was then studied and a significant association was observed when wild-type individuals were compared with those presenting mutant variants (Table 1C). When patients were grouped following the Vienna classification in terms of location and behavior of the lesions, no significant difference was observed for any clinical phenotype, although stratification limited the power of the study (data not shown).

Table 2 IBD5-IGR3081a_1 association with CD in the SLC22A4 1672CC and −207GG SLC22A5 wild-type population

	IGR3081a_1 TT	IGR3081a_1 TG	IGR3081a_1 GG
CD patients (n = 46)	36 (78.26%)	6 (13.04%)	4 (8.69%)
Controls (n = 75)	73 (97.33%)	1 (1.33%)	1 (1.33%)

(GG+TG) vs TT: $P = 0.0006$, OR (95% CI) = 10.14 (1.97–98.04).

The 250-kb IBD5 risk haplotype in 5q31 marked by two polymorphisms in complete LD, IGR2060a_1 and IGR3081a_1, had been already confirmed in our population.[7] A striking result was obtained when, in the absence of the 1672T SLC22A4 and −207C SLC22A5 putative etiological alleles, the IGR3081a_1 polymorphism marked even stronger CD predisposition (Table 2) than before stratification (from Table 1 in Urcelay et al,[7] $P = 0.07$;

OR = 1.35). Moreover, when the TT wild-type IGR3081a_1 population was analyzed, the presence of the 1672T *SLC22A4* and −207C *SLC22A5* risk alleles increased CD susceptibility as well (Table 3), while no difference was observed for the IGR3081a_1 GG group (data not shown). Similar to what was described in the original report, in our population the susceptibility conferred by the 1672T *SLC22A4* and −207C *SLC22A5* mutations was paradoxically stronger after IGR3081a_1 stratification (Table 3) than before (Table 1C). The conditional frequencies (based on having the IBD5 nonrisk haplotype) in the control population are comparable in our study and in Peltekova's[4] (0.22 of controls having either risk variant vs 0.23). In Spain, when controls in the global cohort (Table 1C) were compared with those in the IGR3081a_1-wild-type population (Table 3), their distribution was very different ($P = 1.7 \times 10^{-71}$), showing strong LD in the 5q31 locus.

Given that Peltekova et al[4] reported that in the wild-type 1672CC *SLC22A4* and −207GG *SLC22A5* population the haplotype containing the IBD5 risk allele was equally frequent among affected individuals and controls (0.012 vs 0.013), while we found a significant difference (Table 2), we decided to study the haplotype distribution in our population. To infer the frequencies of the haplotypes identified by the three variants within the 5q31 locus, an EM algorithm was applied to both CD and control cohorts (Table 4). One of the two major haplotypes (HT1), lacking any susceptibility factor, confers protection against CD, as expected. Paradoxically, the one presenting three susceptibility factors (HT8) did not show CD association. Three minor HTs displayed strong association (HT4, HT5 and HT7).

Discussion

The chromosomal region 5q31 is of particular interest in CD because it harbors many genes involved in immune and inflammatory processes. The high- and low-affinity human carnitine transporters, OCTN2[14] and OCTN1,[15] respectively, are multifunctional polyspecific organic cation transporters acting in several tissues. Carnitine is an essential cofactor for β-oxidation of long chain fatty acids in mitochondria, resulting in suppressed ketone body production in the liver. These transporters are encoded by the *SLC22A5* and *SLC22A4* genes, which share 77% identity in their sequences. They act primarily in the elimination of cationic drugs and xenobiotics in kidney, intestine, liver and placenta and their dysregulation may have widespread implications. The association of two functional variants in these genes with CD has been recently reported.[4] Moreover, a different SNP that disrupts a RUNX1 binding site in an intronic region of the *SLC22A4* gene also increased the risk for another autoimmune inflammatory disease, rheumatoid arthritis, in a Japanese cohort.[16] A common physiological role of the *SLC22A4* gene in both inflammatory pathologies seems suggestive.

Ethnic variability in CD has been conclusively proven, for example, for the *NOD2/CARD15* susceptibility locus, successfully replicated in different populations throughout the world, but absent in Asian populations.[17] The study of the 1672C > T *SLC22A4* and −207G > C *SLC22A5* variants has been reported in Japan, where again these polymorphisms were absent.[6] Ethnic-specific allele frequencies at the OCTN locus have been reported.[18] Our results initially suggested association of both 1672T *SLC22A4* and −207C *SLC22A5* etiologic alleles with CD risk in the Spanish population (Table 1C). Further analysis of IBD5 locus susceptibility in the 1672CC *SLC22A4* and −207GG *SLC22A5* wild-type population provided with evidence for increased association of the 5q31 region with CD in these conditions (Table 2). These results could be explained by considering the simultaneous existence of two etiologic factors in the 250 kb locus, which would be not surprising

Table 3 Risk conferred by the 1672C > T *SLC22A4* and −207G > C *SLC22A5* polymorphisms in IBD5-IGR3081a_1 TT-wild-type population

	1672CC and −207GG	1672TX or −207CX	1672TX and −207CX
CD patients (n = 70)	36 (64.49%)	22 (26.03%)	12 (9.46%)
Controls (n = 99)	73 (73.74%)	22 (22.22%)	4 (4.04%)

OCTN mutant vs wild-type individuals: $P = 0.003$, OR (95% CI) = 2.65 (1.32–5.35).

Table 4 Arlequin-estimated frecuencies for haplotypes: *IBD5IGR3081a_1T > G// 1672C > T SLC22A4// −207G > C SLC22A5*

Haplotypes	CD Patients (%)	Controls (%)	
HT1-TCG	195 (38.00)	265 (48.05)	$P = 0.0009$; OR = 0.66
HT2-TCC	47 (9.23)	45 (8.23)	NS
HT3-TTG	0 (0.00)	6 (1.14)	NS
HT4-TTC	26 (5.10)	8 (1.46)	$P = 0.0008$; OR = 3.62
HT5-GCG	28 (5.46)	5 (0.93)	$P = 0.00002$; OR = 6.3
HT6-GCC	13 (2.57)	8 (1.48)	NS
HT7-GTG	13 (2.46)	3 (0.61)	$P = 0.007$; OR = 4.75
HT8-GTC	191 (37.19)	210 (38.10)	NS

given the amount of genes related to immunity clustered in that region. Alternatively, the polymorphisms could be regarded as mere genetic markers of a common causative gene, albeit not necessarily the *SLC22A4* and *SLC22A5* genes. Our data showed a CD predisposition effect of the 1672T *SLC22A4* and −207C *SLC22A5* polymorphisms independently of the 5q31 haplotype contribution (Table 3). Paradoxically, the risk due to the 1672T *SLC22A4* and −207C *SLC22A5* mutant alleles was higher within the IGR3081a_1 TT wild-type population (Table 3) than in the overall cohorts (Table 1C), similarly to what was seen in the Canadian population. This would be an argument against the unique etiologic role of these genes: being these variants the only causative ones, the conditioned analysis would only rest power to the comparison, and just the opposite effect is observed. However, one cannot formally discard the possibility that the OCTN polymorphisms were causative and some epistatic interaction would prevent the full expression of their action. This would explain the lack of susceptibility observed for the HT-8 with the three risk alleles (Table 4). The high frequency of the 1672T *SLC22A4* and −207C *SLC22A5* mutant alleles observed in the control population would suggest again the existence of an epistatic interaction leading to neutralization of the OCTN risk. Alternatively, considering those three polymorphisms (Table 4) as genetic markers of an undetermined susceptibility gene, the HT-8 would define a haplotype without effect in CD predisposition. A danger in these stratification type experiments is obviously multiple testing without the appropriate correction; however, the significance of our results would withstand an ample correction.

The frequencies of the haplotypes ascertained by the EM algorithm probably reflect the existence of different haplotypes: one protective and three conferring CD susceptibility (Table 4). These haplotypes act increasing CD risk only in a reduced percentage of the population, in keeping with our data. The estimated Spanish haplotype configuration encouraged us to suggest that the 1672T *SLC22A4* and −207C *SLC22A5* etiological variants would have an impact on CD susceptibility only within certain genetic context. They revealed three risk haplotypes affecting a minor subset of the Spanish CD population and one protective haplotype marked by the wild-type alleles: *IBD5-IGR3081a_1*T//* L503F *SLC22A4*C//* −207 *SLC22A5*G*. In fact, these data would lead to reinterpret the *IBD5 susceptibility haplotype* under the consideration of the influence of a protective haplotype carrying wild-type alleles.

It remains to be proven whether mucosal inflammation can be caused by excessive stimulation through the OCTN transporters, acting as sensors of microorganisms and/or substances produced in the intestinal lumen. This function will be parallel to that of the *NOD2/CARD15* gene, which is involved in the recognition of bacteria.[19]

References

1 Brant SR, Shugart YY: Inflammatory bowel disease gene hunting by linkage analysis: rationale, methodology, and present status of the field. *Inflamm Bowel Dis* 2004; **10**: 300–311.
2 Rioux JD, Daly MJ, Silverberg MS *et al*: Genetic variation in the 5q31 cytokine gene cluster confers susceptibility to Crohn's disease. *Nat Genet* 2001; **29**: 223–228.
3 Mahler M, Leiter EH: Genetic and environmental context determines the course of colitis developing in IL-10-deficient mice. *Inflamm Bowel Dis* 2002; **8**: 347–355.
4 Peltekova VD, Wintle RF, Rubin LA *et al*: Functional variants of OCTN cation transporter genes are associated with Crohn's disease. *Nat Genet* 2004; **36**: 471–475. E-pub April 11 2004.
5 Newman B, Gu X, Wintle R *et al*: A risk haplotype in the solute carrier FAMILY 22A4/22A5 gene cluster influences phenotypic expression of Crohn's disease. *Gastroenterology* 2005; **128**: 260–269.
6 Yamazaki K, Takazoe M, Tanaka T *et al*: Association analysis of SLC22A4, SLC22A5 and DLG5 in Japanese patients with Crohn's disease. *J Hum Genet* 2004; **49**: 664–668.
7 Urcelay E, Mendoza JL, Martinez A *et al*: IBD5 polymorphisms in inflammatory bowel disease: association with response to infliximab. *World J Gastroenterol* 2005; **11**: 1187–1192.
8 Bayless TM, Tokayer AZ, Polito II JM, Quaskey SA, Mellits ED, Harris ML: Crohn's disease: concordance for site and clinical type in affected family members—potential hereditary influences. *Gastroenterology* 1996; **111**: 573–579.
9 Bonen DK, Cho JH: The genetics of inflammatory bowel disease. *Gastroenterology* 2003; **124**: 521–536.
10 Bouma G, Strober W: The immunological and genetic basis of inflammatory bowel disease. *Nat Rev Immunol* 2003; **3**: 521–533.
11 Lennard-Jones JE: Classification of inflammatory bowel disease. *Scand J Gastroenterol Suppl* 1989; **170**: 2–6; discussion 16–19.
12 Gasche C, Scholmerich J, Brynskov J *et al*: A simple classification of Crohn's disease: report of the Working Party for the World Congresses of Gastroenterology, Vienna 1998. *Inflamm Bowel Dis* 2000; **6**: 8–15.
13 Excoffier L, Slatkin M: Maximum-likelihood estimation of molecular haplotype frequencies in a diploid population. *Mol Biol Evol* 1995; **12**: 921–927.
14 Tamai I, Ohashi R, Nezu J *et al*: Molecular and functional identification of sodium ion-dependent, high affinity human carnitine transporter OCTN2. *J Biol Chem* 1998; **273**: 20378–20382.
15 Tamai I, Yabuuchi H, Nezu J *et al*: Cloning and characterization of a novel human pH-dependent organic cation transporter, OCTN1. *FEBS Lett* 1997; **419**: 107–111.
16 Tokuhiro S, Yamada R, Chang X *et al*: An intronic SNP in a RUNX1 binding site of SLC22A4, encoding an organic cation transporter, is associated with rheumatoid arthritis. *Nat Genet* 2003; **35**: 341–348.
17 Inoue N, Tamura K, Kinouchi Y *et al*: Lack of common NOD2 variants in Japanese patients with Crohn's disease. *Gastroenterology* 2002; **123**: 86–91.
18 Urban TJ, Giacomini KM, Risch N: Haplotype structure and ethnic-specific allele frequencies at the OCTN locus: implications for the genetics of Crohn's disease. *Inflamm Bowel Dis* 2005; **11**: 78–79.
19 Inohara N, Ogura Y, Fontalba A *et al*: Host recognition of bacterial muramyl dipeptide mediated through NOD2. Implications for Crohn's disease. *J Biol Chem* 2003; **278**: 5509–5512.