Am. J. Hum. Genet. 72:509–518, 2003

A Novel *NOD2/CARD15* Haplotype Conferring Risk for Crohn Disease in Ashkenazi Jews

Kazuhito Sugimura,[1] Kent D. Taylor,[1,2] Ying-chao Lin,[1] Tieu Hang,[1] Dai Wang,[1] Yong-Ming Tang,[1] Nathan Fischel-Ghodsian,[1,2] Stephan R. Targan,[1,2] Jerome I. Rotter,[1,2] and Huiying Yang[1,2]

[1]Division of Medical Genetics and Inflammatory Bowel Disease Research Center, Departments of Medicine and Pediatrics, Steven Spielberg Pediatric Research Center, Cedars-Sinai Medical Center, and [2]University of California at Los Angeles, Los Angeles

Crohn disease (CD) exhibits a 2–4-fold increased frequency in Jews as compared with other ethnic/racial groups. Three coding variants of the *NOD2/CARD15* have been reported as independent disease-predisposing mutations (DPMs), but these were found in only 30%–40% of patients with CD and could not account for all the linkage between CD and the *IBD1* locus. The aim of the present study was to explore whether additional DPMs at the *IBD1* locus exist in the high-risk Jewish group. Sixty-four Ashkenazi Jewish and 147 non-Jewish white families were studied. Six microsatellite markers spanning *IBD1* were genotyped for linkage analysis in subgroups stratified on *NOD2/CARD15* DPM status. SNPs in *NOD2/CARD15* (R702W, G908R, 1007*fs*, and S268P) were then genotyped in family and independent case-control samples. On the basis of initial results, sequencing was done on *NOD2/CARD15*-translated regions in 12 Jewish individuals. Subsequently, a new *NOD2/CARD15* variant was genotyped and analyzed. After excluding the influence of the three DPMs, significant linkage of *IBD1* to CD in Jews remained with two peaks at D16S403 (mean allele sharing [MAS] = 0.70] and D16S411 (MAS = 0.59). Further, we observed an increased frequency of a haplotype carrying only the 268S variant in Jewish patients (OR = 3.13, $P = .0023$) but not in non-Jews, suggesting the existence of a Jewish-specific additional disease-predisposing factor on this haplotype. Sequencing of this haplotype revealed a new variant (IVS8+158; JW1). The 268S-JW1 combination exhibited a further increased risk (OR = 5.75, $P = .0005$) and the highest population-attributable risk (15.1%) for CD among reported DPMs in Jews. In Ashkenazi Jews, unrecognized population-specific predisposing factor(s) exist on the 268S-JW1 haplotype at the *IBD1* locus. This factor may contribute to the higher risk for CD in Ashkenazi Jews as compared with non-Jews.

Introduction

Crohn disease (CD [MIM 266600]) is one of the two major forms of chronic inflammatory bowel disease (IBD) and can affect any part of the gastrointestinal tract. The cumulative risk of surgical treatment for patients with CD reaches 60% in 10 years after onset because of the fistulizing, perforating, stricturing, and/or obstructing complications of CD, thereby severely affecting patients' quality of life.

Although the pathogenesis of CD has not been fully elucidated, CD appears to be a multifactorial/oligogenic disease, with both environmental and genetic factors contributing to its etiology. The role of genetic factors in the etiology of CD has been supported by studies of ethnic differences, familial aggregation, and twin and spouse data (Yang and Rotter 2000). With regard to ethnic variation, a consistently increased incidence (2.0–4.0 times) and prevalence rate (2.2–9.4 times) of CD in the Jewish population have been documented, compared with other ethnic groups in the same geographic area (Rotter et al. 1992). These observations are considered to be evidence for a strong genetic predisposition to the etiology of CD and have led to the hypothesis that the higher risk of CD in the Jewish population is due, at least in part, to genetic factors.

Hugot et al. (1996) mapped the *IBD1* gene to the proximal region of the long arm of chromosome 16 (16q12) in the white population, utilizing genomewide scan linkage strategies. This finding has been replicated in many studies, including an international collaborative study reporting a remarkably high multipoint linkage score (MLS) for a complex disease (MLS = 5.7 at marker D16S411 in 16q12) (Ohmen et al. 1996; Parkes et al. 1996; Cho et al. 1997; Brant et al. 1998; Cavanaugh et al. 1998; Curran et al. 1998; Annese et al. 1999; Hampe et al. 1999; Akolkar et al. 2001; the IBD In-

Received September 19, 2002; accepted for publication November 26, 2002; electronically published February 7, 2003.

Address for correspondence and reprints: Dr. Huiying Yang, Medical Genetics Birth Defect Center, SSB 378, Cedars-Sinai Medical Center, 8700 Beverly Boulevard, Los Angeles, CA 90048. E-mail: Huiying.Yang@cshs.org

0002-9297/2003/7203-0002$15.00

ternational Genetics Consortium 2001). The *IBD1* locus has also been demonstrated to be significantly linked with CD in Jewish families (Cho et al. 1997; Akolkar et al. 2001). Recently, the *IBD1* gene (*NOD2/CARD15*) was simultaneously identified by a positional-cloning strategy (Hugot et al. 2001) and a positional candidate gene strategy (Hampe et al. 2001; Ogura et al. 2001*a*). The NOD2/CARD15 molecule is expressed primarily in monocytes and activates NF-kB through interaction with its N-terminal caspase recruitment domains (CARDs) (Ogura et al. 2001*b*). Three coding polymorphisms or variants (R702W, G908R, and 1007*fs*, denoted elsewhere as SNP8, 12, and 13, respectively; Hugot et al. 2001) in *NOD2/CARD15* were shown to be independently associated with CD (Hugot et al. 2001), with the most significant evidence for 1007*fs*, which results in a frameshift in the 10th leucine-rich repeat and is followed by a premature stop codon. This truncation mutation in the NOD2/CARD15 molecule appears to result in an inability to activate NF-κB in response to bacterial lipopolysaccharides (Ogura et al. 2001*a*). All three of these disease-predisposing mutations (DPMs) are found on the same background haplotype, which can be identified by several other SNPs, including P268S (denoted elsewhere as SNP5; Hugot et al. 2001). However, an average of 60%–70% of patients with CD do not have any of the three *NOD2/CARD15* DPMs (Abreu et al. 2002; Ahmad et al. 2002; Cuthbert et al. 2002; Hampe et al. 2002*b*; Lesage et al. 2002; Vermeire et al. 2002). These three predisposing mutations do not account for all of the linkage between CD and the *IBD1* locus, since residual evidence of linkage in the region is observed after families with the variants are removed (Hugot et al. 2001; Hampe 2002*a*).

Because of the unique evolutionary history of the Ashkenazi Jewish population (Motulsky 1979, 1995), one might expect that the genes predisposing to CD in the Jewish population would be more homogeneous than those in the general white population (Diamond and Rotter 2002). A study of the Ashkenazi Jewish population with high incidence of CD should yield increased power and more specific haplotype information to identify genetic variants that are associated with CD, in a manner analogous to the specific *BRCA1* and *BRCA2* mutations occurring in the Jewish population (Struewing et al. 1997; Moslehi et al. 2000). Thus, to determine if additional DPMs exist at the *IBD1* locus in the high-risk Jewish group, we first tested if there was remaining evidence for linkage after excluding the influence of the three known causative mutations (702W, 908R, and 1007*fs*). We then examined the association between CD and the known causative mutations and a background variant. Significant association of CD with a background haplotype in the Jewish population then led us to sequence the translated region of *NOD2/CARD15* to identify additional DPMs in this gene. These studies resulted in identification of a new haplotype for *NOD2/CARD15* that is a major factor predisposing to CD in Ashkenazi Jews.

Subjects and Methods

Study Subjects

A total of 211 families with CD (64 Ashkenazi-Jewish and 147 non-Jewish white families) were studied, consisting of 373 patients with CD and 672 unaffected relatives. The probands of all families were ascertained from the IBD Center at Cedars-Sinai Medical Center or referred to us by gastroenterologists or the Crohn's Colitis Foundation of America nationwide. These families have at least one family member affected with CD and do not have any known individuals affected with ulcerative colitis. In these 211 families, 91 multiplex families with CD that included sib pairs with CD were available for linkage analysis (28 Ashkenazi-Jewish and 63 non-Jewish families). In addition, an independent case-control panel (112 Ashkenazi-Jewish and 166 non-Jewish patients with CD and 79 Ashkenazi-Jewish and 143 non-Jewish control individuals) was also included in this study. Controls were recruited from spouses, married-in relatives, or acquaintances who had no known IBD or other autoimmune diseases or family history of IBD. The present study, as a part of an ongoing IBD genetic study, was reviewed and approved by the Human Subject Institutional Review Board at Cedars-Sinai Medical Center.

Genotyping and Sequencing Analysis

To conduct the linkage study, we genotyped six microsatellite markers spanning the *IBD1* locus, covering 34 cM on chromosome 16 (D16S403, D16S753, D16S409, D16S411, D16S419, and D16S408) using an ABI 377 automated DNA analyzer and associated software (Applied Biosystems). For association analysis of *NOD2/CARD15,* we genotyped all families with CD and case-control samples for SNPs, including the three principal DPMs (R702W, G908R, and 1007*fs*), one background SNP (P268S), and a new SNP (IVS8+158 [C→T], JW1), identified by us after sequencing (see below), by the *Taq*man MGB bialleic discrimination system using an ABI 7900 instrument (Applied Biosystems).

All the sequenced individuals ($N = 12$) were of Ashkenazi Jewish origin. DNA sequences of *NOD2/CARD15* exons, including the 5′ UTR and splicing signal regions, were analyzed, according to the protocol of the Big Dye Terminator Ready Reaction kits (Applied Biosystems). Sequencing primers were the same as PCR primers except for exon 9, in which nested primers were used (table 1). Sequence data were analyzed on an ABI 377 DNA ana-

Table 1

Primers Used for PCR and Sequencing *NOD2/CARD15*

Exon	Forward Primer	Reverse Primer	Size (bp)	Annealing Temperature (°C)
Ex1	TCTCCTCCCCAGATGTTTAAGATG	CCAGCCAAGGATGCCACAGC	856	63
Ex2	TGCCTTCTCTGGGTCTCAAT	ATGGACCAAGTTACCCCACA	751	53
Ex3	GACTGCCCTTCCCTTTCTG	ACATTGCTCCATCAGCCTTC	200	55
Ex4D	CAGAGCCCCTTCCCGTCATC	AGCACAGTGTCCGCATCGTCATTG	625	65
Ex4C	CTGGAGGAGCTCTTCAGCAC	AACAGTTCCTGGTGGCATTT	668	60
Ex4B	CCTGCTCCAAGAGACCTCAG	TCAGATGTCTGGCACTCAGC	678	60
Ex4A	AGATCACAGCAGCCTTCCTG	ATCTGGGCAGTGTTGCAAAG	582	60
Ex5&6	TTTTGGGGGATTTGTAGATT	CTGGGGAGATCACAGCATTAGAGA	631	54
Ex7	ACTCTCTCCCTGGCTTGTC	CGTCCCGCTGCCCCTTTC	435	55
Ex8	GAGGCCACTCTGGGATTGAG	CCTGATCCAGCCCAATATCTT	463	57
Ex9	TGCCAGGCACTATATTAAGGT	GGGCTGGATCAGGTACATT	878	53
Ex10	CTTTATTGGTTACCTTCACTTC	GCTGCAATGGAGAGTGGG	654	55
Ex11	GATGGCACGGGTACTCTT	ACTGAGGTTCGGAGAGCTAAA	511	56
Ex12	GAGGGCACCAGGGTTTGCTCA	GATCAGCAGAGGCCAGTCCCATAC	698	56
Ex9 nested primer	CCCCAGAGCAGAGAATCC	CTTTCCCTGCTCTGACATAC		55

NOTE.—Exon 4 was divided into five overlapping areas and individually sequenced.

lyzer with associated software, and alignments were performed using BioEdit.

Statistical Analysis

Two-point nonparametric linkage analyses were performed separately by use of SIBPAL from the Statistical Analysis for Genetic Epidemiology (SAGE) version 3.1 package on the Jewish and non-Jewish multiplex families. Family-based association was evaluated by a transmission/disequilibrium test (TDT) (Spielman et al. 1993) in all families with CD. Since only 29% of the families were multiplex families, and there was no difference in the TDT results between simplex and multiplex families, we report the results from all families combined. To perform the TDT, GENEHUNTER2 (Kruglyak et al. 1996) was used for four-locus SNP haplotypes, and SIMWALK2 (Sobel and Lange 1996) was used to construct haplotypes for five SNPs. In the case-control samples, PHASE was used to construct haplotypes (Stephens et al. 2001). Haplotypic genotype frequencies were compared between case and control individuals, using the χ^2 test. Odds ratio (OR) and its 95% CI were calculated for each risk haplotype, referenced to the same low-risk haplotype by the Mantel-Haenszel method (Schlesselman 1982). Population-attributable risk (PAR) was estimated under the assumption that the frequency of a risk haplotype in the control group can be regarded as approximately representative of the target population and the OR as an approximation of the relative risk. Thus, PAR = Pe(OR − 1)/[Pe(OR − 1) + 1], where Pe = frequency of a risk haplotype in the control group (Schlesselman 1982).

Results

Stratified Linkage Analysis on the NOD2/CARD15 Genotype

When the families were analyzed in each ethnic group, we found a trend toward increased mean allele sharing (MAS) at markers D16S403 (31 Mb from *NOD2/CARD15,* MAS 0.55, $P = .11$) and D16S411 (1.8 Mb from *NOD2/CARD15,* MAS 0.54, $P = .20$) in the Jewish families with CD. In the non-Jewish families with CD, increased MASs (>0.50) were observed at all marker positions, with the peak at marker D16S411 (MAS 0.56, $P = .094$), but they did not attain statistical significance in this sample size.

When we divided the Jewish families into *NOD2/CARD15* DPMs+ and *CARD15* DPMs− subgroups on the basis of the possession of any of the three principal DPMs in the family, the evidence for linkage in families DPM− for *NOD2/CARD15* increased and reached a significant level, with the two peaks at marker D16S403 (MAS = 0.70, $P = .0008$) at proximal 16p and at marker D16S411 (MAS = 0.59, $P = .1$) at proximal 16q near *NOD2/CARD15* (fig. 1). In contrast, in the non-Jewish families with CD, the significantly increased MAS was observed only in families DPM+ for *NOD2/CARD15* (the highest MAS = 0.61 at D16S411, $P = .0075$), whereas there was no evidence for any MAS increase in families DPM− for *NOD2/CARD15* (MAS = 0.47–0.52) (fig. 1). These results indicated a different contribution of the three known DPMs to the IBD1 linkage of CD between the two ethnic groups and raised the possibility that additional predisposing gene(s) or additional *NOD2/CARD15* DPM(s) exist at the *IBD1* locus and play a more important role in the sus-

ceptibility to CD in the Jewish population when compared with the non-Jewish population.

TDT

We performed the family-based association test (using the TDT method) to investigate any ethnic differences of the CD association with the three known *NOD2/CARD15* DPMs. We observed that all the 702W, 908R, or 1007*fs* variants almost always (98%) occur on a common background haplotype, which contains the 268S variant, and that the 702W, 908R, and 1007*fs* variants were never found on the same haplotype in both family groups.

In the non-Jewish families, there were very significant associations with 702W (transmitted/not transmitted [T/NT]: 42/11, $P = .000021$) and 1007*fs* (T/NT: 34/12, $P = .0012$), and there was also an increased frequency of transmission with 908R (T/NT: 10/6), but this was not statistically significant (table 2). Although the number of families was moderate, we did observe a significant association between CD and the 908R haplotype (T/NT: 13/2, $P = .0045$) in the Jewish families. Meanwhile, the association of the 1007*fs* and of the 702W variants to CD was not observed. Furthermore, there appeared to be a preferential transmission of the haplotype with only the background variant (268S) in Jewish (T/NT: 8/5) but not in non-Jewish families (T/NT: 31/38), though this excess in Jews did not attain statistical significance. Although the lack of association with certain DPMs in Jewish families and preferential transmission of 268S alone could be due to the small sample size, the apparent ethnic difference in allele transmission led us to test these DPMs further in a larger ethnically matched case-control sample.

Table 2

TDT of the *NOD2/CARD15* Haplotypes in Ashkenazi-Jewish and Non-Jewish White Families with CD

	TDT in Ashkenazi Jews		TDT in Non-Jewish Whites	
Haplotype[a]	T/NT	*P* Value	T/NT	*P* Value
702W	2/4		42/11	.000021
908R	13/2	.0045	10/6	
1007*fs*	5/4		34/12	.0012
268S alone	8/5		31/38	
No variant	16/27		36/85	

[a] The haplotypes 702W, 908R, and 1007*fs* represent the haplotype that has the rare variant of each mutation. All 702W, 908R, and 1007*fs* haplotypes also possess the rare variant of the background SNP (268S).

Case-Control Study

We examined haplotypic genotypes in an independent Jewish and non-Jewish case-control panel. Because of small numbers in rare allele homozygotes and compound heterozygotes, we combined the rare allele homozygotes

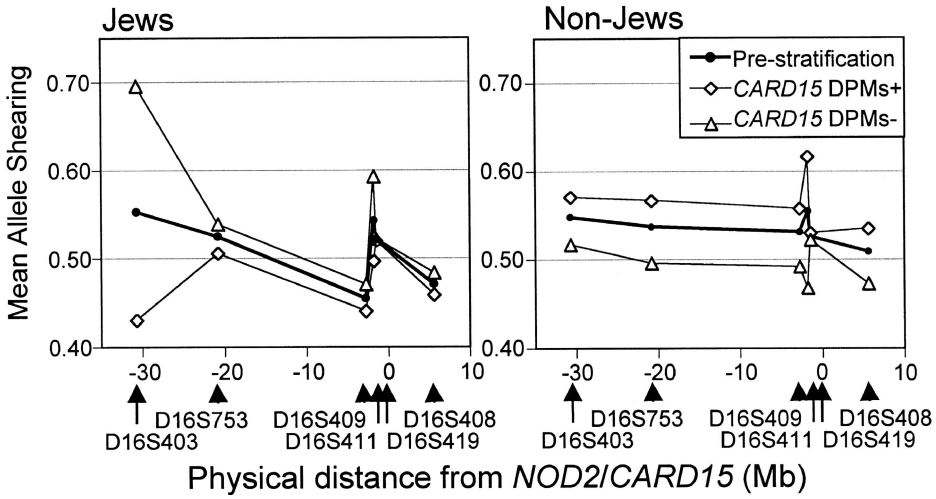


Figure 1 MAS in multiplex families with Crohn disease before and after stratification on presence or absence of three principal *NOD2/CARD15* DPMs (702W, 908R, and 1007*fs*). Six microsatellite markers spanning *NOD2/CARD15* locus on chromosome 16 were genotyped (from left to right, D16S403, D16S753, D16S409, D16S411, D16S419, and D16S408, respectively). *Left panel,* Jewish families. *Right panel,* non-Jewish families. Black circles represent the MAS from entire families (prestratification) in each ethnic group. White triangles represent the MASs in the families where there are no principal *NOD2/CARD15* DPMs, and white squares represent the MASs in families with principal *NOD2/CARD15* DPMs, after stratification.

with heterozygotes and also lumped all other rare haplotypes together so that six main haplotypic groups were formed (table 3).

In comparable sample sizes, we observed a significant association between the haplotype with the frameshift mutation (1007*fs*) and CD in both the Jewish (OR = 7.50, $P = .0041$) and non-Jewish samples (OR = 3.54, $P = .024$) (combined $P < .001$). Increases of the 908R haplotype were seen in both ethnic groups, although the association did not attain statistical significance in the non-Jewish group, probably because of low frequency of this haplotype in the non-Jewish population. With regard to the 702W haplotype, we observed a significant association with CD (OR = 2.50, $P = .022$) in the non-Jewish population only. In the Jewish population, the association of the 702W haplotype did not reach statistical significance, although an increased OR was observed (OR = 2.00, $P = .24$).

It was intriguing that we observed a highly significant association between the haplotype carrying only the background variant (268S alone haplotype) and CD in the Jewish sample (OR = 3.13, $P = .0023$) but no evidence for association of this haplotype in non-Jews (OR = 1.06, $P = .834$). These data indicated that the 268S alone haplotype predisposes to risk for CD in Jews and might contain an as-yet-unrecognized predisposing mutation(s).

Sequence Analysis

To search for possible unrecognized mutation(s) on the 268S alone haplotype in Jews, we sequenced the NOD2/CARD15 exons, including the 5′ UTR and splicing-signal regions in 12 Jewish individuals. These consisted of seven patients with CD with the 268S-alone haplotype (CD1–CD7; fig. 2), three patients (CD8–CD10), and two normal control individuals (NC1 and NC2) without the 268S variant. We found 12 sequence variants, 10 identified elsewhere and 2 newly identified variants, namely intervening sequence (IVS) 8 + 158 (a variant of this SNP is denoted as "JW1") and R791Q. The IVS8 + 158 is a C→T mutation in the palindrome sequence in the intron 8 splicing region, and R791Q is a G→A nonsynonymous mutation in exon 4, causing an amino acid substitution of arginine to glutamine. We did not find any nonsynonymous mutation commonly possessed by Jewish patients with CD with the 268S-alone haplotype. Figure 2 shows the position of these variants in the 12 sequenced individuals. Nine of the 12 variants were found in the Jewish patients with the 268S-alone haplotype. The most common variant observed in those patients was IVS10–133 (SNP9). In addition, 5′ UTR–59 (rs2076752), S178S (rs2067085), R459R (SNP6), R587R (SNP7), R791Q, IVS8+158 (JW1), V955I, and IVS10+64 (rs1077861) were also observed.

Genotyping of the IVS8+158 (JW1)

Of the nine variants observed in the patients with the 268S-alone haplotype, the only common variant not documented elsewhere was the newly identified IVS8+158 (JW1). We therefore genotyped this JW1 variant as a possible DPM or as a marker identifying a disease-predisposing haplotype. Since the available sample size in the Jewish families was limited, we utilized the case-control panel to investigate the role of the JW1 variant.

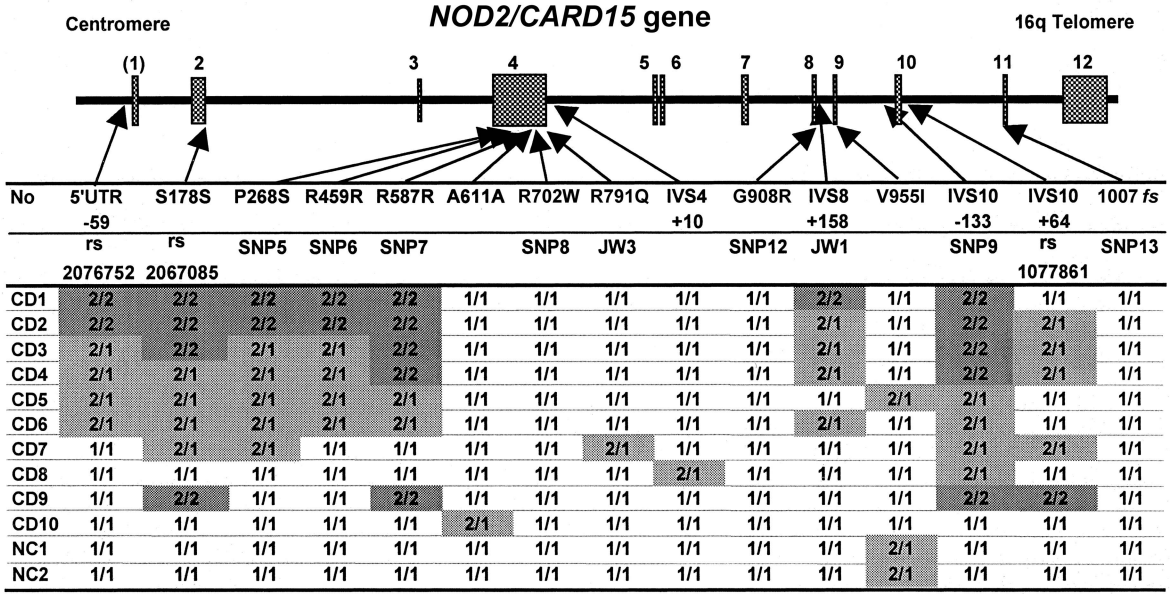
This JW1 variant also showed linkage disequilibrium (LD) with other reported DPMs. In the Jewish population, all the 1007*fs* or 702W variants occurred on the same haplotype as JW1. In contrast, the 908R variant occurred on the haplotype that did not possess JW1 (fig. 3). This trend was similarly observed in the non-Jewish groups. Meanwhile, the 268S-alone haplotype was divided into two haplotype groups on the basis of the possession of

Table 3

Association Results of *NOD2/CARD15* Haplotype Groups Defined by Four SNPs in Ashkenazi-Jewish and Non-Jewish White Patients with CD

	Association in Ashkenazi Jews					Association in Non-Jewish Whites				
Four-SNP-Locus-Haplotype[a]	CD (%)	Control (%)	*P* Value	OR	95% CI	CD (%)	Control (%)	*P* Value	OR	95% CI
702W (2211)	7.1	6.3	.25	2.00	0.6–6.6	13.8	6.9	.022	2.50	1.1–5.6
908R (2121)	12.5	8.8	.067	2.50	0.9–6.8	5.4	3.5	.24	1.96	.6–6.1
1007*fs* (2112)	10.7	2.5	.0041	7.50	1.6–35.5	7.8	2.8	.025	3.54	1.1–11.3
268S alone (2111)	31.2	17.7	.0023	3.13	1.5–6.6	22.2	26.5	.83	1.06	.6–1.8
Other	2.6	1.2	.23	3.75	0.4–37.4	3.0	0	.022		
Reference	35.7	63.2				47.5	60.1			

[a] Each haplotype group consists of individuals with the following haplotypic genotype. The four-SNP positions are P268S, R702W, G908R, and 1007*fs*, respectively. 1 = wild type; 2 = the rare variant of the SNP.; 702W (2211) = heterozygote or homozygote of 702W; 908R (2121) = heterozygote or homozygote of 908R; 1007*fs* (2112) = heterozygote or homozygote of 1007*fs*; 268S alone (2111) = heterozygote or homozygote of 268S without any of other three DPMs; Other = compound heterozygotes; Reference = homozygotes of wild type on all SNPs.



No	5'UTR -59 rs 2076752	S178S rs 2067085	P268S SNP5	R459R SNP6	R587R SNP7	A611A	R702W SNP8	R791Q JW3	IVS4 +10	G908R SNP12	IVS8 +158 JW1	V955I	IVS10 -133 SNP9	IVS10 +64 rs 1077861	1007 *fs* SNP13
CD1	2/2	2/2	2/2	2/2	2/2	1/1	1/1	1/1	1/1	1/1	2/2	1/1	2/2	1/1	1/1
CD2	2/2	2/2	2/2	2/2	2/2	1/1	1/1	1/1	1/1	1/1	2/1	1/1	2/2	2/1	1/1
CD3	2/1	2/2	2/1	2/1	2/2	1/1	1/1	1/1	1/1	1/1	2/1	1/1	2/2	2/1	1/1
CD4	2/1	2/1	2/1	2/1	2/2	1/1	1/1	1/1	1/1	1/1	2/1	1/1	2/2	2/1	1/1
CD5	2/1	2/1	2/1	2/1	2/1	1/1	1/1	1/1	1/1	1/1	1/1	2/1	2/1	1/1	1/1
CD6	2/1	2/1	2/1	2/1	2/1	1/1	1/1	1/1	1/1	1/1	2/1	1/1	2/1	1/1	1/1
CD7	1/1	2/1	2/1	1/1	1/1	1/1	1/1	2/1	1/1	1/1	1/1	1/1	2/1	2/1	1/1
CD8	1/1	1/1	1/1	1/1	1/1	1/1	1/1	1/1	2/1	1/1	1/1	1/1	2/1	1/1	1/1
CD9	1/1	2/2	1/1	1/1	2/2	1/1	1/1	1/1	1/1	1/1	1/1	1/1	2/2	2/2	1/1
CD10	1/1	1/1	1/1	1/1	1/1	2/1	1/1	1/1	1/1	1/1	1/1	1/1	1/1	1/1	1/1
NC1	1/1	1/1	1/1	1/1	1/1	1/1	1/1	1/1	1/1	1/1	1/1	2/1	1/1	1/1	1/1
NC2	1/1	1/1	1/1	1/1	1/1	1/1	1/1	1/1	1/1	1/1	1/1	2/1	1/1	1/1	1/1

Figure 2 *NOD2/CARD15* polymorphisms observed in P268S-alone haplotypes in Ashkenazi Jewish patients with CD. CD1 and CD2 are homozygous patients with the P268S rare allele without known three DPMs (702W, 908R, and 1007*fs*). CD3–CD7 are heterozygous patients with 268S without the three DPMs. CD8–CD10 are patients and NC1 and NC2 are normal control individuals without any of four variants (268S and three DPMs). IVS = Intervening sequences; 1 = Common allele; 2 = Rare allele.

the JW1 variant in both ethnic groups. The new JW1 variant is located telomeric to P268S SNP and is almost always accompanied by the 268S variant (272/283 = 96%) similar to the other three principal DPMs.

In this haplotypic-genotype analysis, we observed a strong association of the 268S-JW1 haplotype (without the three known DPMs) with CD in Jews, independent of the other DPM haplotype groups (table 4). Comparing within the 268S-alone haplotype group, the 268S-JW1 haplotype showed a remarkably increased association (OR = 5.75), and the most significant *P* value

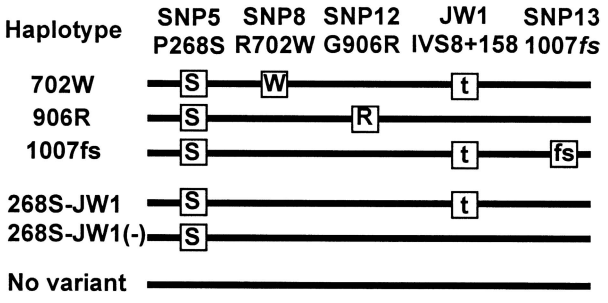


Figure 3 Haplotype structure of the five *NOD2/CARD15* variants genotyped in this study. The letters in white squares indicate the variant (rare allele) of each SNP. The capital letters S, W, and R represent the amino acid variants serine, tryptophan, and arginine, respectively. The small letter t represents the nucleotide variant thymine at the JW1 SNP, and fs represents the frameshift mutation at the 1007*fs* SNP.

(*P* = .0005) among all DPM haplotype groups in the Jewish population. Furthermore, the 268S-JW1 group in the Ashkenazi Jews showed the highest population-attributable risk (PAR, 15.1%) among all DPM groups (fig. 4). When the 268S-JW1 haplotype was taken into account, the population-attributable risk of *NOD2/CARD15* risk variants in Ashkenazi Jews became 28.3%, which was double the PAR of the three DPMs observed in non-Jewish whites (14.6%).

The 268S-JW1(−) haplotype did not show any disease association in the Jewish population. In addition, in the non-Jewish population, we did not find any association with either the 268S-JW1(+) haplotype (CD 15.6% vs. control 15.3%) or the 268S-JW1(−) haplotype (6.6% vs. 11.1%).

A phylogenetic tree of *NOD2/CARD15* haplotypes comprising five genotyped SNPs (P268S, R702W, G908R, 1007*fs,* and IVS8+158[JW1]) was constructed with the use of CLUSTALW (Thompson et al. 1994; fig. 5). In this tree, the 908R haplotype forms a separate branch, whereas haplotypes of 268S-JW1, 1007*fs*, and 702W compose a related group characterized by the possession of the JW1 variant. However, in the JW1 group, the 702W haplotype hardly contributed to the risk of CD in the Jewish patients (PAR = 0.8%); conversely, the 268S-JW1 haplotype hardly contributed to the risk of CD in the non-Jewish patients (PAR = 0.3%). Therefore, we did not delineate a common ancestral haplotype representing the susceptibility to CD in both ethnic groups.

This seems to suggest that an additional causative factor exists on the 268S-JW1 haplotype in the Jewish population.

Discussion

We report the delineation of a novel haplotype (268S-JW1) for the susceptibility to CD in the Ashkenazi Jewish population at the *IBD1* locus on chromosome 16. Although we did not find a new common nonsynonymous variant of *NOD2/CARD15* on this haplotype, 268S-JW1 appeared to contribute predominantly to the risk of CD in Jews, as it presented the highest population-attributable risk (15.1%) among *NOD2/CARD15* variants in the Jewish population. This observation indicates that JW1 is a variant that identifies an unrecognized predisposing factor(s) to CD in the Jewish population. This factor(s) may account for some of the increased susceptibility to CD in the Jewish population.

The JW1 variant reported here is located in the palindrome sequence in intron 8; we could not identify any new common coding mutation in *NOD2/CARD15* on this 268S-JW1 haplotype. Hence, there is a possibility that this variant may affect the expression and function of the NOD2/CARD15 molecule. But we found neither signal sequences nor promoter sequence homology with the JW1 sequence, and there was no preserved sequence between human and mouse species around the JW1 variation (WWW Signal Scan; WWW Promoter Scan; UCSC Genome Bioinformatics). We also checked the splicing of intron 8 in JW1 homozygous mononuclear cells, comparing with control cells without the JW1 mutation, yet we did not find any change of splicing length in intron 8 (data not shown). Further, in the non-Jewish population, we did not observe an association of the 268S-JW1 haplotype with CD. Therefore, these available data argue that the JW1 variant alone does not have a pathologic effect directly, but it more likely occurs in strong LD with a population-specific causative factor in Jews.

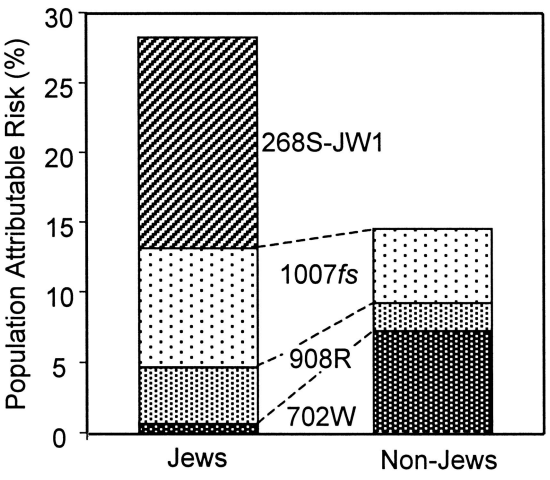


Figure 4 PAR of each *CARD15*-predisposing haplotype in Ashkenazi Jews and non-Jewish whites. The variant names in this figure represent each risk haplotype defined in table 3 and table 4.

Additional strategies will be required to identify the Jewish-specific causative factor to elucidate the mechanisms of the high incidence and prevalence of CD in the Jewish population.

In the Jewish families with CD who did not carry any known *NOD2/CARD15* DPMs, we observed the strongest evidence for linkage at marker D16S403 and an increased mean allele sharing at marker D16S411, close to *NOD2/CARD15*. This phenomenon was not observed in the non-Jewish population. These results indicate that the relative extent of contribution of the three known principal DPMs to the IBD1 linkage to CD is different among ethnic groups, and imply the existence of additional responsible haplotype(s) specific for CD in the Jewish population.

The possible existence of population-specific responsible factors and/or haplotypes other than known principal DPMs in *NOD2/CARD15* has been recently re-

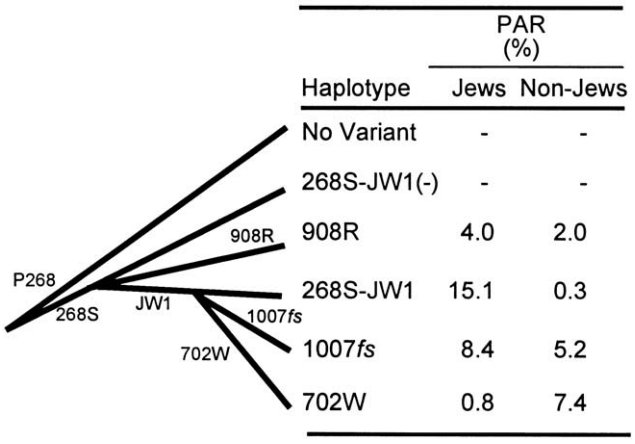
Table 4

Association Results of *NOD2/CARD15* Haplotype Groups Defined by JW1 in Ashkenazi-Jewish and Non-Jewish White Patients with CD

	Association in Ashkenazi Jews					Association in Non-Jewish Whites				
Five-SNP-Locus Haplotype [a]	CD (%)	Control (%)	*P* Value	OR	95% CI	CD (%)	Control (%)	*P* Value	OR	95% CI
268S-JW1	20.5	6.3	.0005	5.75	2.0–16.4	15.6	15.3	.4431	1.29	.7–2.5
268S-JW1(−)	10.7	11.3	.29	1.67	.6–4.3	6.6	11.1	.4908	.75	.3–1.7

Note.—The 268S-JW1 and 268S-JW1(−) groups are the two subgroups of the "268S Alone" group in table 3, defined by presence or absence of the JW1 rare variant. The association was tested in comparison with the "Reference" group, as defined in table 3.

[a] Each haplotype group under "five-SNP-locus haplotype" consists of individuals with the following haplotypic genotype. Each SNP position is P268S, R702W, G908R, JW1, and 1007*fs*, respectively. "1" means wild type; "2" is the rare variant of the SNP. 268S-JW1 (without other three DPMs): 21121/11111, 21121/21111, 21121/21121, 21121/11121. 268S-JW1(−) (without other three DPMs): 11111/21111, 21111/21111.



Haplotype	PAR (%) Jews	Non-Jews
No Variant	-	-
268S-JW1(-)	-	-
908R	4.0	2.0
268S-JW1	15.1	0.3
1007*fs*	8.4	5.2
702W	0.8	7.4

Figure 5 The phylogenetic tree of five haplotypes generated by using CLUSTALW (a multiple alignments and tree-making program). This tree indicates the relationships between the CD-associated haplotypes in both Jewish and non-Jewish populations and a possible haplotype development consisting of five genotyped SNPs (P268S, R702W, G908R, 1007*fs,* and IVS8+158 [JW1]).

ported by Hampe and his colleagues (2002*a*). After excluding the influence of the known *NOD2/CARD15* DPMs in German families with CD, they still observed strong linkage for CD at marker D16S3068 in proximal 16p and at marker D16S3019 in central 16q and weak linkage at marker D16S409 in proximal 16q. The marker D16S3068 is located between D16S403 and D16S753 in our study, and marker D16S409 is located 2.8 Mb centromeric to *NOD2/CARD15*. In addition, Hampe et al. also observed an ethnic difference between German and British families, as there was no remaining linkage in British families with CD after the influence of known DPMs was excluded. Our linkage results in Jewish families were quite in accord with their observations and support the existence of additional population-specific responsible factors and/or haplotypes on chromosome 16.

The *NOD2/CARD15* association with CD has been observed in many populations (but not in all), such as lack of association of the 908R variant in Norwegians (haplotype frequency: CD 0.86%, control 1.2%) (Hampe et al. 2002*b*) and the 908R variant in the Dutch population (allele frequency: CD 4.3%, control 3.0%) (Murillo et al. 2002). We did not observe a significant association between 702W and CD in both family and case-control studies in the Jewish population. These ethnic differences in the association of *NOD2/CARD15* DPMs may be explained by the limited sample size or by the low haplotype and allele frequencies in these populations. Or this difference could imply that the positive association and linkage of these DPMs (908R and/or 702W) to CD is based on LD between a common unknown causative factor on their haplotypes and these variants. However, as we did not identify a common ancestral haplotype between Jewish and non-Jewish populations (fig. 5), an additional causative factor seems to exist on the 268S-JW1 haplotype in the Jewish population.

In conclusion, we found a novel *NOD2/CARD15* haplotype (268S-JW1) in Ashkenazi Jews to be strongly associated with CD, and this haplotype appeared to contribute to the higher incidence and prevalence of CD in this population. Furthermore, we observed evidence for the possible existence of a population-specific disease causative factor(s) at the *IBD1* locus and the possibility of the ethnic-specific additional linkage on the short arm of chromosome 16 in Ashkenazi Jewish patients with CD. Further studies on this novel susceptible *NOD2/CARD15* haplotype (268S-JW1) should lead to the identification of additional as-yet-unrecognized predisposing factor(s) to CD and contribute to our understanding of the underlying mechanism of the linkage of IBD1 to CD.

Acknowledgments

The authors thank patients, their families, and referring physicians for their support to our ongoing genetic studies of IBD. This study was supported by the National Institutes of Health grants DK54967 and DK46763, General Clinical Research Center grant RR00425, and the Cedars-Sinai Board of Governors' Chair in Medical Genetics. Some of the results of this article were obtained by using the SAGE program, which is supported by a United States Public Health Service Resource grant (1 P41 RR03655) from the National Center for Research Resources.

Electronic-Database Information

Accession numbers and URLs for data presented herein are as follows:

BioEdit Sequence Alignment Editor, http://www.mbio.ncsu.edu/BioEdit/bioedit.html (for Windows 95/98/NT)

BioInformatics & Molecular Analysis Section (BIMAS), http://bimas.dcrt.nih.gov/molbio/signal/ (for WWW Signal Scan) and http://bimas.dcrt.nih.gov/molbio/proscan/ (for WWW Promoter Scan)

CLUSTALW, http://www.ddbj.nig.ac.jp/E-mail/clustalw-e.html (for multiple alignments and tree-making)

GENEHUNTER2, http://linkage.rockefeller.edu/soft/gh/ (for statistical genetics computer application)

Online Mendelian Inheritance in Man (OMIM), http://www.ncbi.nlm.nih.gov/Omim/ (for Crohn disease [MIM 266600])

PHASE, http://www.stats.ox.ac.uk/mathgen/software.html (for haplotype reconstruction)

SIMWALK2, http://watson.hgen.pitt.edu/docs/simwalk2.html (for haplotype reconstruction)

Statistical Analysis for Genetic Epidemiology (SAGE), http://darwin.cwru.edu/octane/sage/sage.php (for release 3.1)

UCSC Genome Bioinformatics, http://genome.ucsc.edu/ (to

compare the Human *NOD2/CARD15* gene sequence with mouse genome sequence)

References

Abreu MT, Taylor KD, Lin Y-C, Hang T, Gaiennie J, Landers CJ, Vasiliauskas EA, Kam LY, Rojany M, Papadakis K, Rotter JI, Targan SR, Yang H (2002) Mutations in NOD2 are associated with fibrostenosing disease in patients with Crohn's disease. Gastroenterology 123:679–688

Ahmad T, Armuzzi A, Bunce M, Mulcahy-Hawes K, Marshall SE, Orchard TR, Crawshaw J, Large O, de Silva A, Cook JT, Barnardo M, Cullen S, Welsh KI, Jewell DP (2002) The molecular classification of the clinical manifestations of Crohn's disease. Gastroenterology 122:854–866

Akolkar PN, Gulwani-Akolkar B, Lin XY, Zhou Z, Daly M, Katz S, Levine J, Present D, Gelb B, Desnick R, Mayer L, Silver J (2001) The IBD1 locus for susceptibility to Crohn's disease has a greater impact in Ashkenazi Jews with early onset disease. Am J Gastroenterol 96:1127–1132

Annese V, Latiano A, Bovio P, Forabosco P, Piepoli A, Lombardi G, Andreoli A, Astegiano M, Gionchetti P, Riegler G, Sturniolo GC, Clementi M, Rappaport E, Fortina P, Devoto M, Gasparini P, Andriulli A (1999) Genetic analysis in Italian families with inflammatory bowel disease supports linkage to the IBD1 locus: a GISC study. Eur J Hum Genet 7: 567–573

Brant SR, Fu Y, Fields CT, Baltazar R, Ravenhill G, Pickles MR, Rohal PM, Mann J, Kirschner BS, Jabs EW, Bayless TM, Hanauer SB, Cho JH (1998) American families with Crohn's disease have strong evidence for linkage to chromosome 16 but not chromosome 12. Gastroenterology 115: 1056–1061

Cavanaugh JA, Callen DF, Wilson SR, Stanford PM, Sraml ME, Gorska M, Crawford J, Whitmore SA, Shlegel C, Foote S, Kohonen-Corish M, Pavli P (1998) Analysis of Australian Crohn's disease pedigrees refines the localization for susceptibility to inflammatory bowel disease on chromosome 16. Ann Hum Genet 62:291–298

Cho JH, Fu Y, Kirschner BS, Hanauer SB (1997) Confirmation of a susceptibility locus for Crohn's disease on chromosome 16. Inflamm Bowel Dis 3:186–190

Curran ME, Lau KF, Hampe J, Schreiber S, Bridger S, Macpherson AJ, Cardon LR, Sakul H, Harris TJ, Stokkers P, Van Deventer SJ, Mirza M, Raedler A, Kruis W, Meckler U, Theuer D, Herrmann T, Gionchetti P, Lee J, Mathew C, Lennard-Jones J (1998) Genetic analysis of inflammatory bowel disease in a large European cohort supports linkage to chromosomes 12 and 16. Gastroenterology 115:1066–1071

Cuthbert AP, Fisher SA, Mirza MM, King K, Hampe J, Croucher PJ, Mascheretti S, Sanderson J, Forbes A, Mansfield J, Schreiber S, Lewis CM, Mathew CG (2002) The contribution of NOD2 gene mutations to the risk and site of disease in inflammatory bowel disease. Gastroenterology 122:867–874

Diamond J, Rotter JI (2002) The evolution of human genetic diseases. In: King RA, Rotter JI, Motulsky AG (eds) Genetic Basis of Common Disease, 2nd ed. Oxford University Press, New York, pp 50–64

Hampe J, Cuthbert A, Croucher PJ, Mirza MM, Mascheretti S, Fisher S, Frenzel H, King K, Hasselmeyer A, MacPherson AJ, Bridger S, van Deventer S, Forbes A, Nikolaus S, Lennard-Jones JE, Foelsch UR, Krawczak M, Lewis C, Schreiber S, Mathew CG (2001) Association between insertion mutation in NOD2 gene and Crohn's disease in German and British populations. Lancet 357:1925–1928

Hampe J, Frenzel H, Mirza MM, Croucher PJ, Cuthbert A, Mascheretti S, Huse K, Platzer M, Bridger S, Meyer B, Nurnberg P, Stokkers P, Krawczak M, Mathew CG, Curran M, Schreiber S (2002*a*) Evidence for a *NOD2*-independent susceptibility locus for inflammatory bowel disease on chromosome 16p. Proc Natl Acad Sci USA 99:321–326

Hampe J, Grebe J, Nikolaus S, Solberg C, Croucher PJ, Mascheretti S, Jahnsen J, Moum B, Klump B, Krawczak M, Mirza MM, Foelsch UR, Vatn M, Schreiber S (2002*b*) Association of *NOD2* (*CARD 15*) genotype with clinical course of Crohn's disease: a cohort study. Lancet 359:1661–1665

Hampe J, Schreiber S, Shaw SH, Lau KF, Bridger S, Macpherson AJ, Cardon LR, Sakul H, Harris TJ, Buckler A, Hall J, Stokkers P, van Deventer SJ, Nurnberg P, Mirza MM, Lee JC, Lennard-Jones JE, Mathew CG, Curran ME (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

Hugot JP, Chamaillard M, Zouali H, Lesage S, Cezard JP, Belaiche J, Almer S, Tysk C, O'Morain CA, Gassull M, Binder V, Finkel Y, Cortot A, Modigliani R, Laurent-Puig P, Gower-Rousseau C, Macry J, Colombel JF, Sahbatou M, Thomas G (2001) Association of NOD2 leucine-rich repeat variants with susceptibility to Crohn's disease. Nature 411: 599–603

Hugot JP, Laurent-Puig P, Gower-Rousseau C, Olson JM, Lee JC, Beaugerie L, Naom I, Dupas JL, Van Gossum A, Orholm M, Bonaiti-Pellie C, Weissenbach J, Mathew CG, Lennard-Jones JE, Cortot A, Colombel JF, Thomas G (1996) Mapping of a susceptibility locus for Crohn's disease on chromosome 16. Nature 379:821–823

The IBD International Genetics Consortium (2001) International collaboration provides convincing linkage replication in complex disease through analysis of a large pooled data set: Crohn disease and chromosome 16. Am J Hum Genet 68:1165–1171

Kruglyak L, Daly MJ, Reeve-Daly MP, Lander ES (1996) Parametric and nonparametric linkage analysis: a unified multipoint approach. Am J Hum Genet 58:1347–1363

Lesage S, Zouali H, Cezard JP and EPWG-IBD Group, Colombel J-F and EPIMAD Group, Belaiche J and GETAID Group, Almer S, Tysk C, O'Morain C, Gassull M, Binder V, Finkel Y, Modigliani R, Gower-Rousseau C, Macry J, Merlin F, Chamaillard M, Jannot A-S, Thomas G, Hugot J-P (2002) *CARD15/NOD2* mutational analysis and genotype-phenotype correlation in 612 patients with inflammatory bowel disease. Am J Hum Genet 70:845–857

Moslehi R, Chu W, Karlan B, Fishman D, Risch H, Fields A, Smotkin D, Ben-David Y, Rosenblatt J, Russo D, Schwartz P, Tung N, Warner E, Rosen B, Friedman J, Brunet J-S, Narod SA (2000) *BRCA1* and *BRCA2* mutation analysis of 208 Ashkenazi Jewish women with ovarian cancer. Am J Hum Genet 66:1259–1272

Motulsky AG (1979) Possible selective effects of urbanization

on Ashkenazi Jews. In: Goodman RM, Motulsky AG (eds) Genetic disease among Ashkenazi Jews. Raven Press, New York, pp 301–312

——— (1995) Jewish diseases and origins. Nat Genet 9: 99–101

Murillo L, Crusius JB, van Bodegraven AA, Alizadeh BZ, Pena AS (2002) CARD15 gene and the classification of Crohn's disease. Immunogenetics 54:59–61

Ogura Y, Bonen DK, Inohara N, Nicolae DL, Chen FF, Ramos R, Britton H, Moran T, Karaliuskas R, Duerr RH, Achkar JP, Brant SR, Bayless TM, Kirschner BS, Hanauer SB, Nunez G, Cho JH (2001*a*) A frameshift mutation in NOD2 associated with susceptibility to Crohn's disease. Nature 411: 603–606

Ogura Y, Inohara N, Benito A, Chen FF, Yamaoka S, Nunez G (2001*b*) Nod2, a Nod1/Apaf-1 family member that is restricted to monocytes and activates NF-κB. J Biol Chem 276:4812–4818

Ohmen JD, Yang HY, Yamamoto KK, Zhao HY, Ma Y, Bentley LG, Huang Z, Gerwehr S, Pressman S, McElree C, Targan S, Rotter JI, Fischel-Ghodsian N (1996) Susceptibility locus for inflammatory bowel disease on chromosome 16 has a role in Crohn's disease, but not in ulcerative colitis. Hum Mol Genet 5:1679–1683

Parkes M, Satsangi J, Lathrop GM, Bell JI, Jewell DP (1996) Susceptibility loci in inflammatory bowel disease. Lancet 348:1588

Rotter JI, Yang H, Shohat T (1992) Genetic complexities of inflammatory bowel disease and its distribution among the Jewish people. In: Bonne-Tamir B, Adam A (eds) Genetic diversity among Jews: diseases and markers at the DNA level. Oxford University Press, New York, pp 395–411

Schlesselman JJ (1982) Case-control studies design, conduct, analysis. Oxford University Press, New York, pp 183-190

Sobel E, Lange K (1996) Descent graphs in pedigree analysis: applications to haplotyping, location scores, and marker-sharing statistics. Am J Hum Genet 58:1323–1337

Spielman RS, McGinnis RE, Ewens WJ (1993) Transmission test for linkage disequilibrium: the insulin gene region and insulin-dependent diabetes mellitus. Am J Hum Genet 52: 506–516

Stephens M, Smith NJ, Donnelly P (2001) A new statistical method for haplotype reconstruction from population data. Am J Hum Genet 68:978–989

Struewing JP, Hartge P, Wacholder S, Baker SM, Berlin M, McAdams M, Timmerman MM, Brody LC, Tucker MA (1997) The risk of cancer associated with specific mutations of BRCA1 and BRCA2 among Ashkenazi Jews. N Engl J Med 336:1401–1408

Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res 22: 4673–4680

Vermeire S, Wild G, Kocher K, Cousineau J, Dufresne L, Bitton A, Langelier D, Pare P, Lapointe G, Cohen A, Daly MJ, Rioux JD (2002) *CARD15* genetic variation in a Quebec population: prevalence, genotype-phenotype relationship, and haplotype structure. Am J Hum Genet 71:74–83

Yang H, Rotter JI (2000) The genetics of ulcerative colitis and Crohn's disease. In: Kirsner JB (ed) Inflammatory bowel diseases, 5th ed. W. B. Saunders, Philadelphia, pp 250–279