Genetics of Inflammatory Bowel Disease

Amir Karban MD1,2, Rami Eliakim MD1 and Steven R. Brant MD1

1 Department of Gastroenterology, Rambam Medical Center, Haifa, Israel
2 Harvey M. and Lyn P. Meyerhoff Inflammatory Bowel Disease Center, and Department of Medicine, the Johns Hopkins University School of Medicine, Baltimore, MD, USA

Key words: inflammatory bowel disease, Crohn’s disease, ulcerative colitis, genetics, NOD2 gene, genetic counseling

Abstract
The etiology of inflammatory bowel diseases, Crohn’s disease and ulcerative colitis, is uncertain. Studies of specific environmental factors and immune dysfunction have provided limited insight into disease pathogenesis. There is ample evidence that these diseases are in part the result of genetic predisposition. The early search for candidate genes focused on genes involved in the regulation of immune function. Recent genome-wide searches reported several susceptibility loci for Crohn’s disease and ulcerative colitis. The recent identification of the IBD1 gene (NOD2) with mutations that are associated with susceptibility to Crohn’s disease will have a major impact on the understanding of the genetics of this disease.

IMA 2002:4:798–802

For Editorial see page 815

Crohn’s disease and ulcerative colitis, the chronic inflammatory bowel diseases, are common causes of gastrointestinal illness in young people in the western world. CD may affect any part of the gastrointestinal tract, most commonly the ileum and colon, and is characterized by the presence of discontinuous areas of transmural inflammation. UC inflammation is confined to the colonic mucosa, continuous from the rectum to a variable extent proximally. Both diseases may have extra-intestinal manifestations.

On a global scale, the incidence distribution of UC and CD follows distinct patterns. Northern countries, such as England, Norway, Sweden and the United States, have the highest rates of the diseases. The prevalence of IBD in the USA is 200–300/100,000 with a similar prevalence of CD and UC. By contrast, countries in southern Europe (including Israel), South Africa and Australia have lower incidence rates. The geographic variation in incidence rates may be in flux. It appears that Israel and the southern countries of Europe are ‘catching up’ to their northern neighbors [1]. The maximum prevalence of CD and UC was observed in the kibbutz population in Israel in 1997: 167.2/100,000 and 65.1/100,000 respectively [2,3]. Several studies have suggested that within specified geographic areas, the incidence rate of IBD is consistently two to fourfold higher in Ashkenazi Jews (eastern European origin) than in other ethnic groups [4]. In addition to them having a greater population risk of IBD, Ashkenazi Jews also have a greater familial risk. IBD is uncommon in developing countries. The fact that the incidence in immigrants to western developed countries is approaching that of the natives suggests the presence of a risk factor in developed countries.

One of the difficulties that patients have in accepting their diagnosis of IBD is that a specific etiology has not been identified. With regard to environmental factors in western industrial countries, the only consistently measurable risk factor for CD has been smoking, which appears to be protective for UC. This association was not found among Jewish Israeli CD patients [5]. Several other risk factors have been postulated but remain unproven, notably measles and the use of birth control pills. While a gut luminal bacterium has been implicated in human IBD studies and animal models of IBD, the evidence indicating a specific microbial agent is inconclusive [6]. Furthermore, immunologic studies have shown an increased presence of autoantibodies and antibodies against common dietary products [7], but the role of dietary products in the etiology of IBD is unknown.

Genetic etiology of IBD
Within the environment of western industrial countries, the greatest risk for developing IBD is genetic. There is both direct and indirect evidence of IBD having a genetic etiology, as described below.

Twin studies
Epidemiologic data on twins constitute the most compelling evidence. A carefully controlled twin study, based on the Swedish twin registry, found that in the CD group of 18 unselected monozygotic pairs 8 were concordant for CD versus only 1 of 26 dizygotic pairs. In the UC group, 1 of 16 monozygotic pairs was discordant for the disease, but all the other 20 pairs (dizygotic or unknown zygosity) were discordant. The proband pair-wise concordance rates among monozygotic twins were 58.3% for CD and 6.3% for UC [8]. Similar results were reported from the Danish twin registry [9]: the proband pair-wise concordance rates among monozygotic pairs were 58.3% for CD and 18.2% for UC. The dramatic increase in monozygotic twin concordance rates as compared to dizygotic twin concordance rates for CD (and the similar dizygotic and non-twin sibling concordance rates) can only be explained by hypothesizing a strong genetic component underlying CD [10].

CD = Crohn’s disease
UC = ulcerative colitis
IBD = Inflammatory bowel disease
However, the twin data also suggest that IBD is not completely explained by genetics. The lack of complete concordance in monozygotic twins may partly be attributable to the unaffected twin not being exposed to an environmental trigger or risk factor or, alternatively, having been exposed to an unidentified protective environmental effect.

Associated genetic syndromes
The association of CD with other known genetically determined disorders has been used as indirect evidence for a genetic component. Among these associated diseases are: Turner’s syndrome [11], as well as the complex autoimmune genetic diseases ankylosing spondylitis (especially in conjunction with the HLA-B27 haplotype), psoriasis [12], and multiple sclerosis [13]. Other rare genetic syndromes such as Hemansky-Pudlak syndrome [14], glycogen storage disease type Ib [15] and pachydermoperiostosis [16] have also been associated with an increased risk of IBD. These genetic syndromes should raise suspicion for IBD in symptomatic individuals.

Familial prevalence of IBD
Many studies have demonstrated an increased prevalence of IBD among relatives of patients with CD and UC. Overall, 10–20% of individuals with IBD report one or more additional relatives with IBD [17]. IBD is found 5–12 times more frequently in relatives of people who have CD or UC than in relatives of normal controls. The family IBD history for patients with CD is mostly CD, and for patients with UC it is mostly UC. However, the cross-disease association may be considerable [18]. For example, the Meyerhoff IBD Center at Johns Hopkins in Baltimore (USA) reported the IBD family history in 540 consecutive CD patients. They found that 17%, 19%, and 5.4% had first-degree relatives with IBD, CD, and UC respectively [17]. Otholm et al. [19] reported similar results in the Danish population, with the risk for UC and CD among offspring of patients with IBD being 2–13 times higher than the risk within the general population. Consistent with findings in a more genetic subset, in the Johns Hopkins study those with a family history of IBD were significantly more likely to have a younger age at diagnosis, to be of Jewish ethnicity, and to have small bowel disease [20]. Investigators in Israel also found CD clustering in families, but less than in European or American studies. In a familial study of 189 Jewish CD patients conducted in Israel by Zlotogora and co-workers [21], the IBD prevalence of first-degree relatives was only 6.6%. The prevalence was similar in the families of Ashkenazi and non-Ashkenazi origin.

Although familial clustering could possibly be due to shared familial exposure to an environmental risk factor (such as an infectious agent), no evidence for an increased risk of CD to spouses or adopted family members was found. One frequently used measure of genetic clustering, λ (prevalence in siblings divided by population prevalence), has been estimated at 36.5 for CD, 16.6 for UC and 24.7 for IBD [22]. In comparison with other complex genetic diseases the λ and thus genetic clustering for IBD is relatively high. For example, in insulin-dependent diabetes mellitus and schizophrenia, λ is 15 and 8.6, respectively.

**IBD – a complex genetic disorder**
IBD is considered a complex genetic disorder predicted to involve multiple genes of relatively low penetrance, since the familial patterns of inheritance do not conform to simple Mendelian models. Clearly, a recessive model cannot fit the majority of families as the risk to siblings is not much greater than the risk to offspring. In a Danish study, an autosomal recessive model fit CD inheritance in 7% of families, and a dominant model with low penetrance fit UC inheritance in 9–13% of families [23]. The results from this study were not significantly different from a multifactorial model. A complex mode of inheritance is also predicted by the finding that the sibling risk of a Jewish proband is consistently greater than that of a non-Jewish proband [4]; any simple Mendelian inherited trait would have the same sibling risk independent of the incidence in the ethnic subpopulation (e.g., cystic fibrosis, Tay-Sachs disease).

Genetic complexity can result from gene-gene and gene-environment interactions. One gene polymorphism may require the concurrent inheritance of a second gene polymorphism elsewhere in the genome, inherited independently, for full disease expression to occur, or exposure to environmental triggers may be necessary.

**Identifying IBD genes**
**Candidate gene studies**
The early search for candidate IBD genes focused on regulatory genes of immune function; particularly genes of the major histocompatibility complex. Human leukocyte antigen DRB1*1502 (DR2) is associated with increased risk of UC in Japanese and Ashkenazi Jews, but the allele is rare in non-Jewish Europeans with UC. An association between CD and HLA association is less tentative, with the strongest evidence coming from HLA-DRB1*0101 [24]. More recently, genes outside the HLA system have come under scrutiny. A number of investigators have sought a genetic association of IBD with cytokine genes. A recent large UK study and meta-analysis of previous studies provides strong evidence that allele 2 of the interleukin-1 receptor antagonist (IL-1ra) gene is associated with UC [25]. Anti-tumor necrosis factor-alpha antibodies have been effective in treating CD. Early reports suggested that specific TNF alleles can predict response to the medication; a Japanese study showed evidence of an association with TNF promoter polymorphisms and CD [26]. These polymorphisms were related to high transcriptional promoter activity.

Other examined candidate genes have included the transporter antigen processing genes TAP1 and TAP2, the third component of complement C3, T cell receptor genes and intracellular adhesion molecules, and most recently the natural resistance-associated macrophage protein 1, which modifies intracellular adhesion (27). No significant association was found between either polymorphisms of these genes and patients with IBD. Very recently groups have been evaluating the genetic association of genes that, when disrupted in rodents, produced IBD-like phenotypes generating animal models of IBD, such as interleukin-10 [28].

![HLA = human leukocyte antigen](image)
![TNF = tumor necrosis factor](image)
Currently, the only clinically established genetic markers relevant to IBD are HLA-B27, helpful in establishing coexisting ankylosing spondylitis, and HLA-B8 and DR3 that are associated with 60–80% of primary sclerosing cholangitis, a disease found in 3% of IBD patients.

<table>
<thead>
<tr>
<th>Research group</th>
<th>Reference</th>
<th>No. of pedigrees studied</th>
<th>Loci with evidence of linkage to IBD</th>
</tr>
</thead>
<tbody>
<tr>
<td>France</td>
<td>29</td>
<td>25–53 CD (see text)</td>
<td>16 centromere (IBD1)*</td>
</tr>
<tr>
<td>Oxford, England</td>
<td>30</td>
<td>89–97 (CD: UC or mixed)</td>
<td>3p, 7q, 12q (IBD2)</td>
</tr>
<tr>
<td>Chicago/Johns Hopkins</td>
<td>31</td>
<td>174 (CD: UC or mixed)</td>
<td>1q (IBD7), 3q, 4q, 16q (IBD1)*</td>
</tr>
<tr>
<td>Germany/UK</td>
<td>32</td>
<td>268 (CD: UC or mixed)</td>
<td>1q, 4q, 6p (IBD3), 10q, 12q, 16(IBD1)*, 22</td>
</tr>
<tr>
<td>Los Angeles, USA</td>
<td>33</td>
<td>46 CD</td>
<td>5q*, 14q (IBD4)*</td>
</tr>
<tr>
<td>Pittsburgh, USA</td>
<td>34</td>
<td>62 CD</td>
<td>14q (IBD4)*</td>
</tr>
<tr>
<td>Toronto, Canada</td>
<td>35</td>
<td>158 (CD: UC or mixed)</td>
<td>3p, 5q*, 6p (IBD3), 19p</td>
</tr>
</tbody>
</table>

* Linkage evidence reported for CD only

### Genome-wide scanning

Recent investigative efforts have focused on performing whole-genome linkage studies in multiply affected (multiplex) IBD pedigrees to identify human chromosomal regions (loci) that have genetic evidence for (or linkage to) encoding IBD susceptibility genes. The hypothesis is that IBD clustering within families will be the result of the co-inheritance of specific disease-causing gene polymorphisms or mutations. Loci can be mapped by identifying genetic polymorphic markers (DNA sequences of known chromosomal location containing frequently found sequence variations or polymorphisms), whose polymorphisms are co-inherited among affected relatives significantly more than that expected by chance alone. In a whole-genome linkage study, polymorphic markers located throughout the human genome are determined (genotyped) in multi-affected pedigrees. If there is significantly increased co-inheritance of polymorphisms in the affected relatives for markers in a given region, that region is considered ‘linked’ to the IBD trait. Once linkage is replicated by additional studies and a susceptibility locus is established, genes that are encoded within such loci can be evaluated for mutations or functional polymorphisms maximally associated with cases compared to controls. In contrast to candidate gene studies, this ‘positional cloning’ strategy does not require any a priori functional knowledge of the genes, other than that they are located within a susceptibility locus. One of the first great successes of positional cloning was the identification of the cystic fibrosis gene (CFTR).

The first published genome-wide screen in IBD concentrated exclusively on CD. Hugot et al. [29] studied 41 affected sibpairs in 29 French families with CD. To extend evidence of linkage, a second panel of 71 sibpairs in 53 European families was genotyped for four markers on chromosomes 16 and 1p. The combined data set analysis demonstrated evidence of linkage over a 40 cM region in the pericentromeric region of chromosome 16. This first positionally identified IBD locus was designated IBD1. Interestingly, loci for psoriasis, lupus and the autosomal dominant granulomatous autoimmune disorder, Bluw syndrome, also map to pericentromeric chromosome 16. However, the degree of sharing of this region in relatives showed that IBD1 could explain only a minor proportion of the overall familial CD risk.

Since Hugot's study, there have been six additional published genome-wide screens for IBD. They are summarized in chronological order in Table 1.

### Identification of the first gene for CD: NOD2

Apoptotic activating factor 1 and NOD1, also called CARD4, are members of a family of intracellular proteins that contain an N-terminal caspase recruitment domain, a centrally located nucleotide-binding domain, and a C-terminal regulatory domain. NOD1 promotes apoptosis, but unlike APAF1, it does so through nuclear factor kappa-B activation. NOD1 has striking structural similarity to a class of disease-resistance genes in plants that induce localized cell death at the site of pathogen invasion. By searching a genomic database for NOD1 homologs, Ogura et al. [36] obtained cDNAs encoding NOD2. Northern blot analysis detected 7.0 and 5.5 kb NOD2 transcripts in peripheral blood leukocytes (primarily in monocytes), with little or no expression in other tissues. It was shown that like NOD1, exposure of NOD2 to bacterial lipopolysaccharides results in NFkB activation. NOD2 was mapped to 16q12 (within the linkage region IBD1). All the coding exons and the flanking introns of the NOD2 gene were sequenced in 12 CD-affected individuals from CD families with increased evidence of linkage to IBD1, and in 4 case controls. Ogura and team [37] identified a 1 basepair insertion (C) at nucleotide 3020 (3020insC) in exon 11 of the NOD2 gene. This resulted in a frameshift mutation followed by a premature stop codon. The predicted truncated NOD2 protein contained 1,007 amino acids instead of the 1,040 amino acids of the wild-type protein. The allele frequency of the 3020insC mutation was 8.1% among non-Jewish Caucasians and 8.4% among Ashkenazi Jews. The frequency among control Caucasians was 4%, and among UC patients 3%. Cells transfected with NOD2 carrying the 3020insC mutation had no significant response to bacterial lipopolysaccharide.

Using a positional cloning strategy based on linkage analysis, Hugot and colleagues [38] identified three independent mutations in the NOD2 gene that were strongly associated with CD. In addition to the 3020insC mutation they identified two non-conserved missense mutations, R908W and A1007R. The relative risk of CD for individuals who were heterozygous, homozygous or compound heterozygous (carrying two different mutations) for the identified NOD2 mutations was 3-fold, 38-fold, and 44-fold higher than for normal controls, respectively. Hampe et al. [39] recently evaluated the 3020insC mutation in German and British IBD

**APA** = apoptotic activating factor 1
**NFkB** = nuclear factor kappa-B
patients and controls. They observed an increased risk of 2.6 and 42.1 for heterozygous and homozygous carriers, respectively. Interestingly, they noted that because the number of unaffected persons in the overall population carrying a NOD2 mutation exceeds the number of persons with CD, and especially those with CD and NOD2 mutations, the extrapolated proportion (penetrance) of a heterozygous or homozygous carrier actually developing CD was only 0.03% and 1.7%, respectively! This suggests that other genetic and environmental factors are also necessary for developing CD, even with increased risk from NOD2.

The implication of NOD2 mutations on CD pathogenesis is still unknown. Several mechanisms can explain the susceptibility to CD in individuals carrying the NOD2 mutations. A deficit in sensing bacteria in monocytes might result in an exaggerated inflammatory response by the adaptive immune system. It is also possible that NOD2 variants might be hyper-responsive to lipopolysaccharide from certain bacterial strains.

**Genetic counseling**

There are no established guidelines for CD risk to children of affected parents. In the USA, the lifetime risk to children of one parent with IB is 7–11%, and half of the risk will be reached during the third decade of life. The risk is likely to be higher if the parent has a family history of IB, developed IB at an early age, or is of Ashkenazi Jewish ancestry. The findings of Zlotogora et al. [21] suggest that because IB can be less familial in Israeli Jews, the risk to offspring may be lower than in the USA. There is only one large study of couples in which both parents had IB. Among 19 couples, 12 of the 23 children who were 20 years of age or older (52%) developed IB, usually CD (40). Presently there are insufficient data to estimate risks for counseling of patients with IB in non-white populations.

Clinically, patients with CD and a family history of IB are more likely to be diagnosed at a younger age, to have small bowel rather than colonic disease, and perhaps more extra-intestinal manifestations and more complications of perforation, and abscesses [20].

When siblings are concerned about developing IB, they should be counseled accordingly, taking into account their age and subtracting the past years of risk. Guidance should also be given to avoid potential risk factors (e.g., smoking). Risk to parents because their child developed IB and risk to more distant relatives of a proband with IB are less than the sibling and offspring risk.

Currently, there are no established tests (such as the UC and CD associated antibodies, antineutrophil cytoplasmic antibodies or anti-Saccharomyces cerevisiae antibodies, respectively) for potential identification of relatives at greater risk for developing IB that are beyond the investigative stages. The implications of a positive ANCA or ASCA test in an asymptomatic relative of a proband with IB are unknown. Use of these tests is not indicated in asymptomatic individuals.

The identification of the association between NOD2 mutations and susceptibility to CD has prompted meaningful progress towards understanding the genetics of CD. We and others are seeking additional gene mutations and variations to understand the genetic propensity for developing IBD. As the genetic arm of IBD pathophysiology is dissected, more exact genetic counseling, directed therapeutic modalities and preventive strategies will be possible. Yet, it is agreed that the knowledge to date is not sufficient for genetic counseling based on the NOD2 genotype. There are many variants of the NOD2 gene but their relevance to CD susceptibility is not yet known. As mentioned before, most people who are heterozygous and homozygous for the NOD2 mutations will never develop CD. It is postulated that there are associated environmental factors or other gene mutations that have to coexist with the NOD2 mutations for expressing CD.

**References**


---

ANCA = antineutrophil cytoplasmic antibodies

ASCA = anti-Saccharomyces cerevisiae antibodies
Reviews

40. Bennett RA, Rubin PH. Present DH. Frequency of inflammatory bowel disease in offspring of couples both presenting with inflammatory bowel disease. Gastroenterology 1991;100:638

Correspondence: Dr. A. Karban, Institute of Gastroenterology, Rambam Medical Center, Haifa 31096, Israel.
Phone: (972-4) 854-2954; Cellular: (054) 256-776
Fax: (972-4) 854-3058
Email: akarban@hotmail.com

Show me a sane man and I will cure him for you

C.G. Jung (1875-1961), Swiss psychiatrist, who studied religion and dream symbolism, saw the unconscious as a source of spiritual insight, and distinguished between introversion and extroversion.

Capsule

Searching for schizophrenia genes

The search for genes involved in schizophrenia has been a long and confusing one. Levinson et al. present results of an ongoing effort by a multicenter collaboration to use a sample of about 800 families with schizophrenia to evaluate evidence of genetic linkage in chromosomal regions that have been “nominated” by more than one study. Although previous reports (including one published in Science) had strongly implicated region 1q as being involved, this study did not find any evidence of linkage. Nonreplication suggests that the affects of 1q on schizophrenia are, if present, likely to be very small at the population level.