Inflammatory Bowel Disease: A Genomic Picture Predicts a Changing Response From the Laboratory–Part II

Eric B. Staros, MD
Quest Diagnostics, Schaumburg, IL

DOI: 10.1309/16CRQ017R8A8GG3T

After decades of stagnation, molecular discoveries have altered our view of inflammatory bowel disease (IBD).

This recent change in our understanding of the IBD's pathogenesis combined with emerging biologic therapeutics will benefit the laboratory, the clinician, and the patient.

Crohn's disease (CD) and ulcerative colitis (UC) constitute IBD. Population studies indicate that the combined prevalence of these diseases measure 1 per 1,000 persons. Part I of this series examined the genomic picture of IBD [F1]. Part II will continue this examination as well as look at the molecular side and its implications for the clinical laboratory.

IBD-1 Linkage Analysis

Hugot and colleagues1 described the first genome screen for CD in 1996. The investigators initially screened 2 consecutive and independent panels of families comprising 40 affected sibling pairs using 270 markers throughout the genome. An analysis of allele sharing by affected sibling pairs showed that only 4 markers achieved statistical significance. Using data obtained from a second set of affected families localized the susceptibility locus to a pericentric region of chromosome 16.

The initial work by Hugot was greeted with concern and mild skepticism. Prior to this time, susceptibility loci for other complex diseases had rarely been replicated. Examples of unimpressive attempts to duplicate these sorts of findings can be found in multiple sclerosis, affective disorders, and type II diabetes. At the same time, success had been achieved with high penetrance Mendelian syndromes, most notably the breast cancer-1 gene (BRCA1) which had been discovered using a similar linkage analysis process. The contrasting levels of success diverge according to inheritance patterns. Mendelian inheritance offers hope for finding susceptibility loci, while non-Mendelian (ie, complex) diseases usually failed critical reanalysis.

Refinements in the localization of IBD1 proceeded immediately. By employing a positional-cloning strategy, based on linkage analysis followed by linkage disequilibrium mapping and construction of bacterial artificial chromosomes, Hugot and colleagues identified 3 independent single nucleotide polymorphisms (SNPs) associated with CD.2 A frameshift polymorphism and 2 missense variants of NOD2 were found within the LRR domain encoded by the gene. Genotyping of patients with UC did not show an association with these SNPs. This observation agrees with earlier data showing the lack of linkage between UC and the IBD1 susceptibility locus.3

A connection to immune dysregulation is readily apparent. NOD2 belongs to the Apaf-1/Ced-4 superfamily...
of apoptosis regulators that is expressed in monocytes. From its amino terminus to its carboxy terminus, NOD2 is 2 apoptotic recruitment domains: a nucleotide-binding domain and the LRR region. Furthermore, NOD2 modulates activation of nuclear factor kappa-B (NF-kB) via the carboxy-terminal LRR domain. The same region of NOD2 also acts as an intracellular receptor for microbial bacterial flora. These findings indicate that the NOD2 gene product confers susceptibility to CD by modifying the recognition of microbial flora and/or overactivating NF-kB in intestinal monocytes.

Other clinical and experimental observations provide credibility to NOD2’s role in CD pathogenesis. First, antibiotic therapy causes transient improvements in CD patients, offering support that enteric bacteria may have an etiologic role in the pathogenesis of CD. Second, susceptibility to spontaneous IBD in mice has been associated with the mutations of Toll-like receptor 4, a member of a family of NF-kB activators that binds to lipopolysaccharides via its LRR domain. Third, NF-kB has an important role in IBD via its activation of mononuclear cells of the intestinal mucosa. Furthermore, sulfasalazine and glucocorticoids, the 2 most common therapies for CD, are known to inhibit NF-kB.6

Interestingly, NOD2 is implicated in another familial chronic granulomatous disorder. Blau syndrome (MIM 186580) was first described in a 3-generation kindred and is an autosomal dominantly inherited disease characterized by multiorgan, tissue-specific granulomatous inflammation. Its clinical findings include granulomatous arthritis, skin rash, uveitis, and camptodactyly. Blau syndrome is distinct from both juvenile- and adult-onset rheumatoid arthritis, and it is the only human model of a Mendelian inherited disorder exhibiting a multisystem inflammatory pathology.

Tromp and colleagues8 undertook a genome-wide search for the Blau syndrome susceptibility locus. Sixty-two patients from a 74 member pedigree were genotyped with dinucleotide-repeat markers. Linkage analyses then proceeded under a dominant mode of inheritance. A marker in the pericentric region of chromosome 16 demonstrated a maximum LOD score. At the same time, HLA markers failed to show any association with disease inheritance. The genetic connection is further clarified by Miceli-Richard and colleagues.9 These investigators performed a mutation screen of 4 families with Blau syndrome and identified 3 different missense mutations within the apoptotic region of NOD2.

The significance of mutations within an apoptotic region draws attention to other pathologic and genetic findings. First, granulomas are the pathologic hallmark of both Blau syndrome and CD. A connection is apparent; NOD2 is expressed primarily in monocytes, a precursor cell to the epithelial cells and multinucleated cells in granulomas. In addition, numerous inherited autoinflammatory diseases, including familial Mediterranean fever (MIM 249100), are traced to mutations in apoptotic or death domain regions. Clearly, the pathogenesis of autoinflammatory disorders, including Blau syndrome and CD, relies in part upon the modulation of cell death genetic programming.10

Important research leads to new questions: (1) What is the contribution of these genetic variations in a diverse patient population? (2) What are the spectrum alleles for this gene? (3) How might these alleles correlate with clinical phenotypes? (4) And, by what cellular mechanisms do these variants provide for disease susceptibility?

In order to answer some of these questions, Vermeire and colleagues11 collected a population of 231 patients with CD from the province of Quebec. The prevalence of the 3 previously described mutations (Arg702Trp, 12.9%; Gyc908Arg, 5.2%; and Leu1007fsinsC, 10.3%) was found to be similar to the findings reported by Hugot and colleagues. In addition, the NOD2 mutations showed a distinct predilection for CD ileitis and ileocolitis, while CD limited to the colon was not associated with these 3 genetic variants.

Vermeire and colleagues also examined the SNPs in genomic regions surrounding the NOD2 gene. They aimed to determine the haplotype structure of this genomic region and mark the extent of linkage disequilibrium.
Similar to the findings of Daly and colleagues, tightly linked regions were separated into blocks, with a limited number of haplotypes per block. In addition, all 3 causal alleles occurred in a single block, thereby yielding 3 sub-haplotypes. As would be expected, this haplotype without any of the 3 causal variants was not associated with CD.

The underlying mechanism by which NOD2 variants confer risk for CD is not completely understood. Ogura and colleagues investigated the functional effect of the Leu1007fsinC variant and found that disease susceptibility is potentially related to a decreased ability to sense bacteria in the gastrointestinal tract. Given the adaptive nature of the immune system, a defect in sensing bacteria would likely result in an exaggerated inflammatory response.

Unfortunately, NOD2 and its 3 described mutations do not predict a response to infliximab. To discover this, Vermeire and colleagues studied 245 CD patients (86 fistulizing and 159 luminal) receiving infliximab [T1]. These patients were genotyped for the 3 variants of NOD2, without prior knowledge of treatment response. In total, 32.6% of CD patients carried mutations in NOD2, as compared to 15% of a control population (P<0.001). Despite differences in TNF-alpha production in mucosal biopsy specimens, there was no relationship between the mutations in NOD2 and short-term infliximab response.

A decrease in NF-kB function appears to underlie the NOD2 mutations. The Leu1007fts mutation truncates the NOD2 protein causing the greatest decrease NF-kB activity. In contrast, Arg702Trp and Gly908Arg mutations led to a NF-kB response that was greater than the frameshift mutation, but less than the wild-type gene. All of these mutations are found in the LRR domain, a region required for the recognition of the LPS in bacterial cell membranes.

Interestingly, the mutations associated with Blau syndrome alter a different domain of the NOD2 gene, and result in alterations in the nucleotide binding region of the protein. This suggests that Blau syndrome is related to a gain of function mutation. This is sharp contrast to CD where the bacterial inflammatory recognition functions are decreased.

One could speculate that the NOD2 mutations associated with CD would result in a decreased elimination of intracellular pathogens, resulting in a compensatory increase of other inflammatory processes. In a sense, granulomatous inflammation, a hallmark of CD and a backup for more potent inflammatory processes, indicates the failure of the NOD2 dependent system for eliminating pathogens.

**IBD2 is Strongly Associated with UC**

After the initial report of IBD1, Satsangi and colleagues reported a susceptibility locus for both CD and UC. Having studied 186 sibling pairs from 160 families, the strongest linkage was found in a region spanning 41 centimorgans on the long arm of chromosome 12. In addition, 2 independent groups have reported a positive transmission disequilibrium test for the marker D12S83 in separate patient populations. Since the first report, however, replication has been seen in a number of studies, but usually not to the degree originally observed. Though a number of factors may be involved, Parkes and colleagues argue that the strongest association to IBD2 has been seen in data in which there are a significant number of ulcerative colitis families. In fact, analysis of the combined Oxford/Pittsburgh data has shown significant heterogeneity at the IBD2 locus between UC and CD.

Investigators have not yet identified the gene responsible for IBD2. Based on IBD2’s position and biological plausibility, a number of genes may be placed on a list of likely candidates. Some of these candidate genes have already been excluded from consideration.

One interesting gene in this region is the AVIL gene. AVIL encodes a protein (advillin) that belongs to a family of genes involved in the morphogenesis of microvilli. Tumer and colleagues evaluated 24 unrelated patients with linkage to IBD2, as well as 91 individuals from 19 affected IBD families for putative SNPs. Similar to other candidate genes to date, no association to this gene was found. The investigators concluded that AVIL can probably be excluded as the gene underlying the IBD2 locus.

Also within this region is the gene for beta-7 integrin. It encodes a protein that is involved in lymphocyte homing to the gut and the retention of intra-epithelial lymphocytes. The 16 exons and the promoter region of this gene where screened for polymorphisms. Although 14 SNPs were found, data from 464 affected IBD families did not show an association with the disease phenotype. Further investigation of this region warrants attention due to the success of antiadhesion molecule therapy in patients with IBD.

---

**NOD2/CARD Allele Frequencies for Arg702Trp, Gly908Arg, and Leu1007fsinC in a Quebec Population**

<table>
<thead>
<tr>
<th>Mutation</th>
<th>Patients with CD/IBD (n = 231)</th>
<th>Patients with Sporadic IBD (n = 135)</th>
<th>Patients with Familial IBD (n = 96)</th>
<th>Control Individuals (n = 71)</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arg702Trp</td>
<td>12.9</td>
<td>12.3</td>
<td>13.9</td>
<td>4.2</td>
<td>.0012</td>
</tr>
<tr>
<td>Gly908Arg</td>
<td>5.2</td>
<td>5.3</td>
<td>5.1</td>
<td>0.7</td>
<td>.0022</td>
</tr>
<tr>
<td>Leu1007fsinC</td>
<td>10.3</td>
<td>11.7</td>
<td>7.7</td>
<td>0.7</td>
<td>5.6 x 10^-8</td>
</tr>
<tr>
<td>Overall</td>
<td>45.0</td>
<td>45.3</td>
<td>44.3</td>
<td>9.0</td>
<td>2.6 x 10^-8</td>
</tr>
</tbody>
</table>
IBD3 Embraces the HLA Region and TNF-alpha Gene

The IBD3 susceptibility locus on chromosome 6 includes the major HLA genes and the genes encoding TNF-alpha. Four independent linkage studies have confirmed the location. Although HLA class I, II, and III associations with IBD have been documented, there remains uncertainty whether some of these associations are the results of linkage disequilibrium within this region of the genome.

Nevertheless, studies have shown that certain extraintestinal manifestations of IBD are associated with particular HLA antigens. In some investigations, HLA antigenic differences have also been shown between UC and CD. In 1 study, DR and DQ molecules separate UC and CD on genetic grounds, indicating that the contribution of the HLA class II genes to disease susceptibility is quite different for the 2 disorders. At the present time, both single HLA determinants or tightly linked haplotypes offer little toward understanding the pathogenesis of IBD.

In the same region, the TNF-alpha gene offers a more interesting story. It has been shown that TNF-alpha is increased in the serum, stool, and intestinal mucosa of patients with IBD. Due to its position and function, the TNF-alpha gene is a strong candidate for contributing to the pathogenesis of IBD. Moreover, the overproduction of TNF-alpha and the efficacy of anti-TNF-alpha medications, most notably infliximab (Remicade), offer an appealing connection between the molecular genetics and clinical practice.

Animal studies also point to an important role for TNF-alpha. A deletion of 3-prime regulatory elements from the TNF-alpha transcript in mice results in increased production of TNF-alpha and a CD-like phenotype. In other studies, TNF-alpha knockout mice show a marked reduction in chemically induced intestinal inflammation.

Similar to NOD2, TNF-alpha is influenced by NF-kB. In monocytes, the predominant source of TNF-alpha, NF-kB is required for the activation and production of TNF-alpha. Increased levels of NF-kB translocation to the nucleus and TNF-alpha production have been shown in lamina propria monocytes of patients with IBD, especially CD.

In 2 independent Caucasian cohorts, van Heel and colleagues showed an association of a novel TNF-857C promoter polymorphism with IBD (P=0.001). By further subdividing the IBD patients, additional associations were seen within the subphenotypes of the IBD population, but only in patients who were not carrying the common NOD2 mutations.

Van Heel and colleagues further explored the transcriptional control of the TNF-alpha gene. The TNF-857C polymorphism prevented the binding of the OCT1 transcription regulation factor, thereby allowing for greater production of TNF-alpha. Adjacent to OCT1 binding region is the NF-kB binding region. Studies show that OCT1 can physically interact with NF-kB and inhibit its activating effects. Although not proven, it appears likely that IBD pathogenesis is partially dependent upon the TNF-857C polymorphism's ablation of the OCT1 binding site and subsequent enhancement of the NF-kB mediated inflammatory response.

IBD4

Evidence suggests that a region on chromosome 14q11-12 as being involved in the susceptibility to CD. Four groups of investigators have demonstrated linkage to this region, albeit with varying degrees of statistical significance. All of these patient datasets have been heavily composed of CD families, notably from studies in Pittsburg, Los Angeles, Belgium, and Chicago. To date, efforts have narrowed the region of susceptibility linkage, however, no specific gene has been identified as contributing to disease susceptibility.

IBD5 Hastens the Disease Onset

Rioux and colleagues studied 256 father-mother-child trios where the child had CD and at least 1 parent was unaffected. This population offered an examination of a young population with CD. A linkage disequilibrium mapping process was used due to its ability to localize the fine structure of rare disease genes, and a hierarchical
strategy searched for microsatellite polymorphisms in a region of chromosome 5q31.

Similar to the non-random association of alleles at different loci found within the HLA region of the IBD3 susceptibility locus, the IBD5 region is also complicated by tightly linked alleles. Mapping revealed a common haplotype with 11 SNPs that were in almost complete linkage disequilibrium.

Despite this challenge, investigators attempted to identify the causal allele by examining all known genes within the critical region for allelic variants that might confer an increased susceptibility to CD. Attention was focused on a cytokine gene cluster which includes a number of plausible candidate genes, including the IL4, IL5, and IL13 genes.

The IBD5 haplotype is increased in unrelated individuals with CD who had 1 or 2 NOD2 mutations. It is not elevated in CD patients who do not have the 3 common NOD2 mutations. Hence, investigators have suggested that IBD5 appears to hasten the onset of the disease only in patients with recognized NOD2 mutations.

Difficulties notwithstanding, some investigators have suggested that the OCT3 gene may be the source of the IBD5 susceptibility locus. OCT3 produces a gene product that is a member of the cation/carnitine transporter family of proteins. These proteins are essential for intestinal fatty acid oxidation, and the inhibition of this metabolic pathway has been shown to cause experimental colonic ulceration in an animal model.

One important avenue of inquiry has already appeared. DNA arrays provide a wide-reaching and comprehensive approach to the study of IBD. By evaluating a repertoire of mRNA transcripts, researchers hope to identify expression patterns that will differentiate UC from CD, reveal new mechanisms of disease pathogenesis, and focus attention toward molecular candidates for therapeutic intervention.

To date, expression array data supports many widely held concepts about IBD. In particular, distinctions were seen between UC and CD by examining a number of gene expression groupings. The cancer-related gene cluster showed altered expression in UC but not in CD or normal epithelium. As an example, the gene Mxil, an inhibitor of myc and putative tumor suppressor gene was downregulated in UC. The association with myc oncogene offers a molecular mechanism for a long-standing clinical understanding. Specifically, patients with UC are at risk for colonic glandular dysplasia and transformation to adenocarcinoma. For patients with CD, the risk for neoplasia is not as great.

In contrast to UC, CD overexpressed a number of inducible natural antimicrobial genes, and genes associated with cellular stress. For example, the gene for defensin was the most overexpressed in a comprehensive array. Compared to the UC profile, defensin showed double the expression in CD patients. In a manner somewhat reminiscent of NOD2, defensin 5 is an antimicrobial peptide which is highly expressed in the Paneth cells of the small intestines. Like NOD2, a conserved plant disease-resistance gene and regulator of the response to LPS, defensin is an early responder to cellular stress caused by microorganisms.

Without meaningful therapeutic choices, the molecular findings, including mutations of NOD2, will be of little practical importance. Luckily, many new drugs are in the latter stages of development [T2]. An example of this is natalizumab, a monoclonal directed against a leukocyte cell adhesion molecule referred to as an integrin. In a placebo-controlled, double-blind study, 18 patients with active CD received a 3-mg/kg infusion of natalizumab and 12 patients received a placebo. At 2 weeks, the CD activity decreased significantly in the natalizumab group, but not in the placebo group. Similar findings have been seen in another study. Other medications including recombinant interleukin receptor antagonist and epidermal growth factor also show promise.

Many of these medications are directed at the heart of the innate immune systems, targeting the mechanisms that underpin the pathogenesis of IBD. Without being overly optimistic, it is reasonable to predict that numerous therapeutic choices will soon exist for IBD. The expense of new biologic agents will require an individualized approach be taken to prescribing these drugs. As an example, one-third of CD patients do not respond to infliximab. Yet as of this date, no test can predict a patient’s response to therapy. For the clinical laboratory, the need for predictive molecular markers heralds a challenge that is moving toward us. Shortly, we will be routinely asked to use molecular markers to individualize therapy for IBD patients.

New Biologic Agents with Importance to IBD

<table>
<thead>
<tr>
<th>Medication, generic name (company)</th>
<th>Trade Name</th>
<th>FDA approval and areas of interest</th>
<th>Activity</th>
<th>Comments on Therapy and Potential Complications</th>
</tr>
</thead>
<tbody>
<tr>
<td>infliximab (Centocor)</td>
<td>Remicade</td>
<td>FDA approved for rheumatoid arthritis and for severe Crohn’s disease. Also of interest in TRAPS and UC</td>
<td>Anti-TNF-alpha monoclonal which eliminates TNF from the circulation and causes apoptosis of monocytes and macrophages</td>
<td>Tuberculosis and anti-dsDNA as a laboratory finding, and is rarely associated with a lupus-like syndrome. Infusion site reactions</td>
</tr>
<tr>
<td>adalimumab (Abbott Laboratories)</td>
<td>Humira</td>
<td>FDA approved for rheumatoid arthritis</td>
<td>Humanized anti-TNF monoclonal which eliminated TNF from the circulation and causes apoptosis of macrophages</td>
<td>Similar infectious complications as infliximab. The positive anti-ds DNA are anticipated to be fewer than with infliximab due to the medication’s humanized monoclonal nature</td>
</tr>
<tr>
<td>etanercept (Immunex)</td>
<td>Entrel</td>
<td>FDA approved for rheumatoid arthritis</td>
<td>A TNF-receptor-Fc fusion protein that competes with TNF</td>
<td>Possibly tuberculosis. Injection site reactions</td>
</tr>
<tr>
<td>anakinra (Amgen)</td>
<td>Kineret</td>
<td>For sepsis and rheumatoid arthritis</td>
<td>Binds IL-1 receptor as an antagonist</td>
<td>Injection site reactions</td>
</tr>
<tr>
<td>Anti-alpha 4 beta7 integrin antibody (Millennium Pharmaceuticals)</td>
<td>LDP-2</td>
<td>For UC and CD</td>
<td>Specifically binds to a cell adhesion molecule and inhibits WBC migration</td>
<td>Phase II trials</td>
</tr>
<tr>
<td>Bortezomib MLN-341 (formerly known as LDP-341, Millennium Pharmaceuticals)</td>
<td>Velcade</td>
<td>FDA approved for multiple myeloma</td>
<td>Inhibits the proteasome via decreased degradation of Ikappa B and decreased activity of NF-kappa B</td>
<td>Rare side effects including alopecia, gastrointestinal disturbance, and myelosuppression have been seen in clinical trial</td>
</tr>
<tr>
<td>Interleukin-10, rIUL-10 (Schering Plough)</td>
<td>NA</td>
<td>Interleukin 10 is a major cytokine inflammatory inhibitor</td>
<td>Non-specifically decreases the inflammatory process via suppression of IL-2 and IFN-gamma</td>
<td>CD Phase III trial, failed</td>
</tr>
<tr>
<td>Natalizumab (Elan/Biogen)</td>
<td>Antegren</td>
<td>Not FDA approved, best studied in multiple sclerosis</td>
<td>A humanized monoclonal antibody against alpha-4 integrin, a cell adhesion molecule. Inhibits WBC migration</td>
<td>Phase II studies studying multiple sclerosis are completed. Phase III clinical trials for multiple sclerosis and CD began in December 2001. Complications have not been documented at this time</td>
</tr>
</tbody>
</table>


