Genetics and Innate and Adaptive Immunity in IBD

Galliano Zanello · David Kevans · Ashleigh Goethel · Mark Silverberg · Andrea Tyler · Kenneth Croitoru
Division of Gastroenterology, Department of Medicine, Zane Cohen Center for Digestive Diseases, Mount Sinai Hospital, University of Toronto, Toronto, ON, Canada

Abstract
Inflammatory bowel disease (IBD) is an abnormal inflammatory response within the gut to a trigger that has yet to be identified. The family history in many patients, especially those with Crohn’s disease, suggests a genetic predisposition. It has been hypothesized that the abnormal inflammatory response is due in part to genetic alterations in the normal homeostatic processes that regulate host interactions with the normal gut microbes. Genetic studies have identified increasing numbers of genetic risk factors that involve a diverse series of pathways such as receptors of innate immune response, defects in epithelial barrier function, immune- and cytokine-related genes and genes involved in autophagy. Studies further suggest that abnormal immune responses in IBD patients are directed against the intestinal microbiota, with activation of both innate and adaptive immune responses. Indeed, studies have shown bacterial-derived antigen as drivers of T cell immune responses. More recently, Th17, regulatory T cells and unconventional innate-like T cells have been implicated in the induction and regulation of intestinal inflammation. The seminal discoveries of pathogen recognition receptors including Toll-like receptors and nucleotide-binding oligomerization domain receptors have changed our understanding of how immune cells respond to microbes and the role this may play in IBD pathogenesis. Understanding these mechanism will lead to new strategies in the treatment and prevention of IBD.

Introduction
Inflammatory bowel disease (IBD) is thought to be due to an abnormal inflammatory response within the gut to a trigger that has yet to be identified. The strong family history in many patients, especially those with Crohn’s dis-
ease (CD), suggests a genetic predisposition. It has been hypothesized that the abnormal inflammatory response is due in part to genetic alterations in the normal homeostatic processes at play in the gut mucosa that normally serve to ensure a symbiotic relationship between the host and the normal gut microbes. Genetic studies in patient cohorts have identified increasing numbers of genetic risk factors that involve a diverse series of molecular and cellular pathways that may be altered in these individuals. These pathways include receptors of innate immune response, defects in epithelial barrier function, immune- and cytokine-related genes and genes involved in autophagy – a cellular pathway not previously considered in studies of IBD. Animal and human studies further suggest that abnormal immune responses in IBD patients are directed against the intestinal microbiota, with evidence of a loss of the normal homeostatic mechanisms leading to activation of both innate and adaptive immune responses. The adaptive immune response has long been the focus of study in understanding the pathogenesis of IBD. Th1 and Th2 cell subsets have been identified as mediators of immune inflammation in early studies and antigen specific responses key to their activation. Indeed, studies have shown bacterial-derived antigens as initiators and drivers of T cell immune responses. More recently, Th17, regulatory T cells (Treg) and unconventional innate-like T cells have been identified as important in the induction and regulation of intestinal inflammation, and efforts are in place to define their role in IBD. The seminal discoveries of pathogen recognition receptors including Toll-like receptors and nucleotide-binding oligomerization domain receptors have changed our understanding of how immune cells without antigen-specific receptors respond to microbes and secrete inflammatory cytokines in the intestine.

The Genetics of IBD

Epidemiological Evidence for a Genetic Contribution to IBD Pathogenesis

Epidemiological studies provide compelling evidence that genetic factors contribute to the pathogenesis of IBD. First-degree relatives of patients with IBD have a 3–20 times greater likelihood of developing the disease than the general population [1–3]. Siblings of CD patients have a relative risk of developing CD up to 35 times background population [3]. Finally, twin studies show that the heritability of IBD is high with CD and ulcerative colitis (UC) monozygotic-dizygotic concordance rates of 30 versus 4 and 16 versus 4%, respectively [4–6].
Progress in Characterizing the Human Genome

The progress in understanding the genetic basis of IBD has been possible by rapid changes in technology of genetic sequencing and by large-scale multicentered collaborative studies of a very heterogeneous and polygenic disorder. The ‘completion’ of sequencing of the human genome in 2001 set the stage [7, 8]. Early drafts of the human genome showed previously unrecognized variability, with regions characterized by high gene density and others more sparsely populated. Further, the number of protein-coding genes (30–40,000) was much lower than expected and more complex than those described in other organisms [7]. Projects such as the HapMap and 1,000 Genomes Project, now provide accurate haplotype information on human DNA polymorphisms in multiple populations [9, 10]. The identification of millions of single nucleotide polymorphisms (SNPs), allowed for quantification of genetic variability across populations [11]. Despite these advances in understanding the human genome, the role of genetic factors in diseases such as diabetes, obesity and IBD demonstrates complex heritability, with numerous loci modulating disease risk. In addition, these diseases show incomplete penetrance, polygenicity and differential epigenetic regulation. Genetic heterogeneity, where multiple genes may cause similar phenotypes, and copy number variation, could lead to alterations in phenotype, further complicating genetic analyses [12, 13]. In spite of these complexities, major advances have been made in the genetic study of IBD.

Techniques for Analyzing the Human Genome

Early genetic studies used existing knowledge of either a disease or basic cellular process to identify candidate genes and then attempted to determine whether alterations to that gene were associated with disease pathogenesis. When candidate gene approach was not possible, positional cloning and linkage analysis were developed to identify loci involved in disease [12]. Recent studies more commonly use SNP markers, with substitutions (nucleotide switch e.g. from A to G) as the most common form of polymorphism measured in genotyping studies, and insertions and deletions less common. These studies assume that there are linkages between such markers and disease-specific polymorphisms [12]. Identifying transmission disequilibrium defined how often a specific marker allele is passed from a heterozygote parent to an affected offspring [14]. These approaches are both cost and labor intensive, and require a large number of families to provide statistical power. Current approaches to identifying polymorphisms or loci of interest involve the use of
case-control study designs or genome-wide association studies (GWAS), with large cohorts of cases and controls in conjunction with high-throughput technologies [15].

**SNP Genotyping Platforms and Analysis**

Several genotyping platforms have been developed which allow the high-throughput identification of loci of interest for genetic association studies. These can be used for both candidate SNP-based analyses and larger-scale GWAS. SNP markers are selected from HapMap datasets and must be detectable in a sufficient proportion of the population (detection in ≥1% of the population) to be measured in experiments [9] or the 1,000 Genomes Project [16]. Current GWA chips can provide information on up to 4.5 million SNP markers (Illumina), thus rapidly providing a huge amount of information.

**Key IBD Genetic Studies**

Major histocompatibility complex class II alleles were among the first candidate genes tested and shown to be associated with IBD [17]. A significant breakthrough utilizing linkage mapping studies was the discovery in 1996 of the IBD 1 locus, a pericentromeric region on chromosome 16 as a susceptibility locus for CD [18, 19]. Hugot et al. [20] further investigated the IBD 1 susceptibility locus for CD which had been mapped to chromosome 16. Using a positional cloning strategy, based on linkage analysis followed by linkage disequilibrium mapping, a frameshift variant and two missense variants of *NOD2* were identified as being associated with CD [20, 21]. The discovery of *NOD2* as an important genetic risk was a major advance in understanding CD pathophysiology as it highlighted the importance of innate immunity in CD.

The first GWAS in IBD was performed in 2005 with a modest-sized Japanese CD case-control study and a genome-wide panel of SNPs of relatively limited density. This study mapped the region containing tumor necrosis factor superfamily ligand member 15 as conferring the highest risk for CD among Pacific-Asian populations [22]. More recent GWAS in North American and European Caucasian populations have utilized larger case and control populations along with several hundred thousand SNP genotyping platforms [23–25]. Meta-analyses of these datasets have now identified loci with even smaller effects sizes, increasing to 163 the number of described IBD-associated loci [26–28].
Further important insight into IBD pathogenesis came from a whole-genome linkage study of consanguineous pedigrees with severe, infantile-onset, autosomal recessive mendelian CD, followed by sequencing candidate genes in these regions. By this approach, loss-of-function mutations in the interleukin (IL)-10 receptor subunit genes IL10RA and IL10RB and later the IL-10 cytokine were identified as causes of CD or CD-like disease. IL-10 and its receptors primarily function to downregulate inflammation [29].

It is beyond the scope of this article to review all published GWAS data; however, several important studies are highlighted. In 2006, a GWAS performed in Caucasian cases with ileal CD and controls found a highly significant association between CD and the IL23R gene on chromosome 1p31, a gene that encodes a subunit of the receptor for the proinflammatory cytokine IL-23. An uncommon coding variant (rs11209026) was shown to confer a strong protective effect against CD, and additional noncoding IL23R variants are independently associated [23].

In addition, a strong association was identified in the genomic region encoding ATG16L1, a component of a large protein complex essential for the process of autophagy. This was the first identification of an association between autophagy and CD pathogenesis. Autophagy describes the process whereby targeted cytoplasmic constituents are isolated from the rest of the cell within autophagosomes, which are then fused with lysosomes and degraded. This pathway is also recognized as an immune mechanism due to its role in the destruction of intracellular pathogens. The Wellcome Trust Case Control Consortium (WTCCC) performed a GWAS in the British population examining approximately 2,000 individuals for each of 7 major diseases and a shared set of approximately 3,000 controls. Case-control comparisons for CD identified 9 independent association signals for CD. Association signals were found at a number of previously identified loci including NOD2 and IL23R providing replication of these previous findings [30]. A further GWAS utilizing the WTCCC cohort identified an association between IRGM, another gene in the autophagy pathway and CD [25].

To identify loci that have even smaller effects on IBD risk, large datasets from individual GWAS have been combined in meta-analyses. A meta-analysis published in 2010 used data from 6 index CD GWAS providing a discovery cohort of more than 6,000 CD cases and 15,000 controls. Approximately 900,000 SNPs were tested for CD associations. This study identified 30 new CD susceptibility loci which in combination with previously confirmed loci increased to 71 the number of loci with genome-wide significance associated with CD [26]. Similarly, a meta-analysis including 6 UC GWAS with a large sample size (e.g. discovery cohort of 6,333 UC cases and 15,056 controls) identified 29 additional
UC risk loci, increasing the number of UC-associated loci to 47 and the total number of confirmed IBD risk loci to 99 [27].

The recently published International IBD Genetics Consortium Immunochip Study is the most comprehensive GWAS in IBD to date, compiling data from 7 CD and 8 UC GWAS in Caucasian cohorts with replication of these GWAS signals using independent CD, UC and healthy control cohorts and genotyping with the Illumina Immunochip (Illumina Inc., San Diego, Calif., USA). The Immunochip contains more than 200,000 autoimmune disease-associated SNPs. The Immunochip Study evaluated a combined sample size total of more than 75,000 subjects with greater than 20,000 CD, 15,000 UC and 25,000 healthy controls [28]. As a result, 71 new IBD associations were discovered, resulting in a total of 163 IBD-associated loci, which meet genome-wide significance thresholds.

**Insights into IBD Biology Derived from Genetic Studies**

Rare cases of early-onset IBD may be single-gene, mendelian disorders, e.g. mutations in IL10RA or IL10RB; however, the genetic risk in the majority of IBD patients is mediated by relatively common alleles (allele frequencies greater than 5%) with modest or low effects sizes (odds ratios less than 1.5). A small number of lower frequency risk alleles with larger effects on IBD susceptibility have also been identified, including \( \text{NOD2} \) and \( \text{IL23R} \) variants; however, none are sufficient alone to cause IBD. These findings suggest that most IBD cases are multifactorial in etiology, reflecting an interaction of multiple genetic risk alleles, microbial, immune and environmental factors [31].

It is important to point out that IBD genetic risk loci are not specific for IBD alone in that meta-analyses of GWAS and the recent Immunochip Study showed that some IBD risk loci are also associated with other immune-mediated disorders (66 with other immune-mediated diseases) [28, 32].

**Challenges for Genetics**

Of the estimated heritability of IBD derived from twin and family studies in IBD, less than 30% is explained by genetic variants discovered in GWAS [31]. This finding may be due to the possibility that heritability assessments have overestimated the heritability of IBD. There are however a number of other potential explanations for this ‘missing heritability’ which need to be explored in future studies. GWAS are not optimally designed to discover rare variants with large
effect sizes, and further linkage studies complemented by candidate gene approaches may uncover some of these variants. In addition, the issues of copy number variants, gene-gene interactions, gene-microbiota interactions and epigenetic phenomena may each contribute directly to or modulate the genetic risk of IBD.

Most of the identified 163 IBD-associated SNPs are in linkage with other proximal SNPs resulting in the implication of multiple genes per locus. Potentially, any SNP in a haplotype block may be a true causative polymorphism. Only a few loci thus far have been shown to be functional, i.e. an IBD-associated SNP can be implicated in the dysfunction of a particular gene. Determining which IBD-associated SNPs in coding regions result in functional consequences and which noncoding SNP variants regulate gene expression will be important future endeavors [32]. As knowledge of IBD genetics becomes further refined, follow-up studies to define the association of IBD-associated variants and disease behavior, risk of complications and response to medical therapy are required before genetic discovery can be translated into routine clinical practice.

Immune Basis of IBD

Innate Immunity in IBD

The innate immune response is an evolutionary conserved antigen nonspecific system of defense against microorganisms. Innate immune cells such as epithelial cells, stromal cells, dendritic cells (DCs), macrophages and innate lymphoid cells (ILCs) express pattern recognition receptors that can sense specific pathogen-associated molecular patterns expressed on microorganisms, and initiate the development of a rapid inflammatory response characterized by the secretion of cytokines and chemokines and the recruitment of immune cells. Innate myeloid cells can also activate the inflammasome, a multiprotein complex promoting the maturation of the inflammatory cytokines IL-1β and IL-18. In addition, DCs are antigen-presenting cells responsible for T cell activation, and RORγt+ ILCs have been shown recently to present antigen and regulate CD4+ T cell response to commensal bacteria [33].

Epithelial Barrier Function

The epithelial barrier provides anatomical containment of the microbiota and ensures a tolerant response to commensal microorganisms. Epithelial cells can sense intestinal microbiota through pattern recognition receptors and trigger
the first immune signals. In the small intestine, epithelial cells form a barrier covered by a discontinuous mucous layer secreted by goblet cells. Paneth cells, granule-containing specialized cells found in the epithelial crypts of the small intestine, secrete antimicrobial peptides with a bactericidal effect. In the large intestine, epithelial cells are covered by a thick, continuous, double mucus layer, which insulates the epithelium from the high microbial load. Mutations in genes involved in tight-junction structure or Paneth cell function have been described in IBD [34]. Animal studies have shown that a defective barrier can be associated with an increase in mucosal immune activation leading to an inflammatory environment in the gut [reviewed in reference 34].

**Autophagy**

Autophagy is a conserved pathway induced by cellular starvation, stress or infection. This process involves the formation of double membrane vesicles that surround and degrade cytoplasmic material, organelles and microbes that have invaded the cell. CD is associated with mutations in *ATG16L1*, *IRGM* and *NOD2*, which are all involved in autophagy; however, it remains unclear how alterations in this cellular process lead to a defective bacterial handling or antigen presentation by DCs and activation of mucosal inflammation [35, 36].

**Innate Lymphoid Cells and Innate-Like T Cells**

ILCs are emerging as fundamental effectors of innate immunity and tissue remodeling. Several subpopulations of ILCs have been identified and the heterogeneity within these populations continues to be an issue for discovery. The definitions for the different subsets are based on the profiling of their cytokine-producing ability and specific transcription factors they express, along with differences in cell surface molecules [reviewed in reference 37]. Although for some, their cytokine production profiles are similar to those of helper T cell subsets, they lack antigen-specific T cell receptors. How and if these ILCs are involved in the induction or perpetuation of IBD is still unknown. In the absence of adaptive immunity, these cells can trigger innate colitis in *Rag1*⁻/⁻ mice infected with *Helicobacter hepaticus* [38]. More recently, human studies have shown that ILCs secreting IL-17 [39] and IFN-γ [40] accumulate in the mucosa of CD patients indicating that these cells may play an important role in IBD.

Unconventional innate-like T cells include γδ T cells, invariant natural killer T cells and mucosal-associated invariant T cells. These cells are highly represented in mucosal sites, and they respond to IL-23 stimulation. How polymorphisms in the *IL23R* gene affect the function of these cells in IBD is still unknown.
Adaptive immunity involves antigen-specific responses mediated by T and B cells. These lymphoid populations produce and express antigen-specific receptors (T cell receptors and immunoglobulin, respectively). The ability to mount an antigen-specific response requires that these cells see and respond to foreign antigen and undergo antigen-induced expansion. This process takes time, and so adaptive immunity is seen as the induction of an immune response that is in place to protect against reexposure to a pathogen. The process of T and B cell education and elimination of self-reactive lymphocytes is an important mechanism in order to maintain a healthy and appropriate immune response. Much of this regulation involves lymphocytes that can turn off immune and inflammatory responses.

CD4+ T cell subsets are essential mediators of immune homeostasis and inflammation. Using a murine T cell transfer model, Powrie et al. [41] showed that transfer of naïve CD4+ T cell subsets to lymphopenic mice leads to the development of colitis. This animal model also led to the formal demonstration that colitis development requires antigen-specific T cell activation, while its prevention is dependent on Treg cell function.

Th1 and Th17 Cells

Human studies have described an increase in Th1-Th17 effector T cell activity in CD and Th2 effector T cells in UC, although this Th1-Th17/Th2 paradigm is being reconsidered [42, 43]. Intestinal Th1 cells respond to intracellular bacteria and viruses and develop in response to IL-12 secreted by antigen-presenting cells. Th1 cells express the transcription factor T-bet and secrete IFN-γ. In the murine T cell transfer model, treatment with anti-IFN-γ antibody can prevent the colitis induced by naïve CD4+CD45RBhigh T cells [41]. However, anti-IFNγ antibody is ineffective in the treatment of humans with active CD [44].

Intestinal Th17 cells respond to extracellular pathogens and develop in response of TGF-β, IL-6 and IL-23. In IBD, IL23-induced Th17 cells are an important mediator of mucosal inflammation. Th17 cells express the transcription factor RORyt, and their function is mediated by different cytokines including IL-17A, IL-17F, IL-21, IL-22 and GM-CSF. Experimental studies have shown that IL-17 expression can induce inflammation and mucosal damage or intestinal protection. In the T cell transfer colitis, transfer of RORy-null T cells failed to increase mucosal IL-17 cytokine levels and did not induce colitis [45], while in a Citrobacter rodentium infection model, IL-17A and IL-17F exert protective effects [46]. In the DSS model of colitis, IL-17F deficiency resulted in reduced colitis, whereas IL-17 knockout mice developed more severe disease, thus dem-
onstrating opposing activities of IL-17F and IL-17A [47]. The complexity of IL-17 biology has also been highlighted in clinical studies showing that inhibition of IL-17 can exacerbate CD [48].

IL-21 is an important factor for Th17 differentiation. However, in active human IBD, the majority of IL21-producing cells are IFN-γ+ Th1 cells and not IL-17A+ [49], thus raising the issue of the requirement of Th17-derived IL21 in disease pathogenesis.

IL-22 produced by Th17 cells appears to have an IL-23 dependency [50]. In CD, serum IL22 levels are increased in active disease and may correlate with disease-associated IL23R polymorphisms [51]. Interestingly, absence of IL-22-producing cells is observed in active UC [52].

Regulatory T Cells
Numerous studies using mouse models support a role for Treg cells in IBD. In the T cell colitis model, transfer of CD4+CD45RBlow T cell subset containing a regulatory population prevents the development of colitis [41]. The majority of Treg cells express the transcription factor Foxp3, although some can be Foxp3+, a population termed Tr1 cells. Foxp3 is important in Treg development and function, as observed in humans carrying a mutation of Foxp3 who develop IPEX syndrome [53, 54]. Treg cells can exert their regulatory mechanism through secreted soluble cytokines such as IL-10, TGF-β1, IL-35 and cell-cell interaction mediated by surface proteins (CTLA-4, GITR, and OX40) [55, 56]. Treg isolated from the inflamed mucosa of IBD patients show normal suppressive function in vivo [57]; however, the failure of regulatory mechanisms in IBD may be due to a Smad7-dependent T effector resistance to suppression by TGF-β1 [58], T effector resistance to apoptosis [59], Treg sensitivity to apoptosis [60], or by the plastic conversion of Treg cells into Th1 like and Th17 like as a result of the inflammatory intestinal milieu [61, 62].

Effect of Microbiota on T Cell Development and IBD
The human gut is home to more than a trillion bacteria, representing about 500–1,000 different species [63]. In a healthy individual, these bacteria are separated from the immune system by mucus and the epithelial barrier; however, this intense load of microbes can pose a burden on the immune system; an insult resulting in damage to the epithelium can lead to microbial translocation into the mucosa, which will induce an immune response. Without the presence of these bacteria, the immune system does not develop appropriately. In ‘germ-free’ or gnotobiotic mice, there is a reduction in the frequency and size of Peyer’s patches, decreased intestinal IgA production, decreased CD4+ T cell numbers within the lamina propria and the intraepithelial compartment [64]. Therefore,
the microbiota is key to the development of a normal mucosal immune system. Indeed, Ivanov et al. [65] identified segmented filamentous bacteria as responsible for the induction of Th17 cells within the murine mucosa. Furthermore, work done by Mazmanian et al. [66] identified that a factor expressed on the surface of Bacteroides fragilis, a common gut microbe, has the ability to induce CD4+ T cells to express Foxp3 and convert into IL-10-producing Treg cells. Polysaccharide A, the factor capable of inducing these Foxp3+ Treg cells, has been shown to protect against colitis development and EAE. Most recently, work from Atarashi et al. [67] isolated and identified a cocktail of 17 Clostridia species from human feces capable of inducing and maintaining Treg development. The gut microbiota has an obvious role in inducing intestinal inflammation, but the extent to which the microbiota is involved in IBD pathogenesis is still unclear.

**Manipulation of Intestinal T Cells in the Treatment of IBD**

The understanding of how the adaptive immune system via Treg induction by antigen can regulate the development of colitis is important to developing new strategies for treatment of IBD. Indeed, feeding antigen to mice receiving effector T cell transfer can lead to the induction of Treg cells and prevention of colitis [68]. Building on the observation that feeding the monoclonal anti-CD3 antibody can prevent the induction of EAE, we have shown that feeding anti-CD3 antibody can prevent the induction of T cell-induced colitis [69]. Clinical trials are now evaluating approaches to induce Treg cell activity in patients with IBD. Indeed, in vitro-expanding Treg cells from humans were shown to provide some benefit in treating established mucosal inflammation [70].

**Conclusions**

In less than two decades, the field of IBD genetics has made huge advances in our understanding of the pathogenesis of IBD. The 163 IBD-associated loci identified thus far have identified and implicated numerous important molecular and cellular pathways which may contribute to IBD. The importance of host-microbe interactions in particular has been emphasized, and further study of the interaction of intestinal microbiota and host immune system is likely to further advance IBD knowledge. A focus on the definition of IBD risk variants associated with actual alterations in gene function will also be important, so that these alterations can be targeted by therapeutic interventions.

Numerous IBD risk variants are within critical immunological pathways, such as Paneth cell function, autophagy, bacterial sensing and production of cytokines. It has become evident that IBD is a multifactorial disease, involving
both innate and adaptive aspects of the immune system. Adding to the complexity is the enormous variation of the human microbiota and the effect it has on both immune development and disease pathogenesis. By studying patients with IBD and animal models of colitis, we can begin to piece together the complex network of immune pathways, signals and cells that lead to IBD development. Hopefully, this will lead to new strategies for treatment and prevention.

Acknowledgements

Funding provided by Crohn’s and Colitis Foundation of Canada, Canadian Institutes of Health Research, Mount Sinai Hospital Department of Medicine Research Fund and Canadian Association of Gastroenterology.

Disclosure Statement

The authors declare that no financial or other conflict of interest exists in relation to the content of the chapter.

References

30 Wellcome Trust Case Control Consortium: Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. Nature 2007;447:661–678.


