CHAPTER 2

ROLE OF DEFECTIVE AUTOPHAGIA AND THE INTESTINAL FLORA IN CROHN’S DISEASE

Anouk Regeling¹, Rajesh Somasundaram², Colin de Haar², Janneke van der Woude², Henri Braat² and Maikel P. Peppelenbosch²

¹University of Groningen, University Medical Center Groningen, Department of Gastroenterology and Hepatology, Groningen, The Netherlands,
²Erasmus MC, Erasmus University of Rotterdam, Department of Gastroenterology and Hepatology, Rotterdam, The Netherlands

ABSTRACT

The precise mechanisms underlying the development of Crohn's disease (CD) remain controversial, but sufficient data have been collected to suggest that an uncontrolled immune response within the intestinal mucosa leads to inflammation in a genetically susceptible host. Although lack of mucosal regulatory T-cells causes colitis in humans and experimental rodents, patients with CD have more rather as less regulatory activity in the intestine, apparently excluding defects in tolerance as the cause of CD. Genome-wide association studies (GWAS) have identified many gene variants that confer susceptibility and which seem associated to diminished functioning of especially innate immunity. In apparent agreement, CD patients are impaired with respect to innate immune responses and in controlling bacterial flora in the intestine. Furthermore, severe genetic deficiencies in innate immunity, like e.g. lack of NADP oxidase activity or diminished function of the Wiskott Aldrich syndrome protein, are associated with colitis in mice and men, and are often mistakenly diagnosed as CD. Thus we favor the view that the primary defect in CD is a lack in innate immunity, causing second tier immunological defenses to combat otherwise easily controlled bacterial breaches of the mucosal barrier.
The primary function of the intestinal tract is absorption of nutrients and secretion of waste products, both active at the innermost layer of the gastrointestinal tract (GI) mucosal lining, which consist of a thin permeable epithelium directly exposed to the external environment, called the lumen. The intestinal tract can be considered the largest surface in humans that constantly is exposed to a variety of environmental antigens, pathogenic microbes and houses a large community of commensals. Its function in food and water absorption necessitates that the major part of the intestine consists of a single layer epithelium, ill-adapted at withstanding the mechanical wear and tear associated with the passage of food and thus despite the fact that the intestinal tract is composed of numerous barriers to protect it against pathogenic invasion. Bacteria constantly pass the epithelial and the entire intestine is in a constant state of low-grade inflammation. Nevertheless, the regulatory mechanisms present usually limit this inflammation to a subclinical state. In Inflammatory Bowel Disease (IBD), however, the intestine is characterized by chronic episodal gastrointestinal inflammation, interspaced by periods of remission. The set of disorders grouped under this common denominator consist of Crohn’s disease (CD) and ulcerative colitis (UC) as well as a group of more indeterminate disorders. Despite the overlapping pathological and clinical characteristics of CD and UC, they also show several distinctive pathological features. CD can be distinguished from UC by distinct clinical phenotypes with respect to location and nature of the inflammation. Most often, in about 50% of the patients, the terminal ileum is involved, although CD may affect the entire gastrointestinal tract from the mouth to the perianal area. In about 30% of the patients the disease is located in both the ileum and colon and in approximately 20% the disease is limited to the colon. Areas of inflammation typically reveal discontinuous transmural involvement that often leads to development of complications such as microperforations and fistulas, abdominal abscesses or granulomas, depending upon their location and severity. UC is even as CD a relapsing inflammatory disease, but is in adults characterized by the presence of continuous inflammation limited to the mucosal layers and occasionally the submucosa of the colon, although in pediatric patients the entire tract may be involved. Typically, in adult UC, the disease often involves the rectum and extends proximally, but remains restricted in the colon.

The prevalence and incidence is the lowest in southern climates and underdeveloped countries, such as South America, southeast Asia and Africa, but increases fast in incidence in these regions as well, especially in Brazil and China. This variation in
incidence rates significantly depends on geographic location and may be a result of environmental factors, such as industrialization, sanitation, hygiene and access to specialized health care. However, the prevalence and incidence rates also differ in different racial (e.g. African Americans, Asians, Hispanics, Caucasians) and ethnic status (e.g. Jewish vs. non-Jewish), implicating an important role for environmental factors as well as genetic influences. IBD often begins in adolescence and early adulthood, although a second peak is seen between ages of 50-80, without any gender specificity. Although, the etiology of IBD has not yet been fully defined, several factors have been hypothesized to be risk factors for IBD. Researchers are in general agreement that IBD rises from a combination of factors, which include the external environmental (e.g. smoking, diet, geography), genetic susceptibility, the intestinal microbial flora and the responses in innate and adaptive immunity.

ROLE OF REDUCED TOLERANCE
Classically, IBD is viewed as diminished tolerance against the commensal flora. Evidence for this comes from observations that rodents reconstituted with immune systems that lack regulatory capacity, develop intestinal flora-dependent colitis so called transfer colitis and inhibition of regulatory T-cell activity causes CD-like colitis in men. Also, genetically abolished signaling of the tolerogenic hormone IL-10 is a well established model for colitis, whereas exogenous application of IL-10 using genetically modified bacteria shows promise in treating severe CD in the clinic (a phase I trial with transgenic bacteria expressing IL-10 in CD, all pointing to the idea that lack of tolerogenic capacity is the problem in this disease (Figure 1)). However, patients with CD have supernormal levels of regulatory T-cells, especially in the inflamed lesions and thus although reduced tolerance can cause colitis, it is not the root cause of CD and alternative explanations are called for.

GENETICS OF IBD: DIMINISHED FUNCTIONALITY OF THE INNATE IMMUNE SYSTEM CAUSING SUSCEPTIBILITY TO CD?
Although all diseases seem to be ‘genetic’, the contribution of differences gene product structure or expression pattern to the development of disease is highly dependent on the interplay between environmental and genetic influences. In addition to the contribution of environmental factors in the pathogenesis of IBD, strong evidence from epidemiological studies that examined the occurrence of CD or UC within different familial aggregations, implicate the importance of genetic influences in IBD.
Family studies provided strong evidence for genetic influence towards the susceptibility in IBD. Up to 5-10% of affected individuals have at least one first-degree family member with IBD with the relative risk to siblings ranging from 5-35 for CD and 10-15 for UC, suggesting that positive family history is more common in CD patients than in UC patients\textsuperscript{11,12}. This notion was confirmed in studies of twins and those persons having a twin with IBD had the highest risk of developing IBD in comparison to the general population\textsuperscript{12,13}. In monozygotic twins, the concordance rate for CD is reported to range between 40-60% compared to 4-12% in dizygotic twins, which is almost comparable to all siblings\textsuperscript{11,12,14}. The concordance rate for UC is less frequent in monozygotic twins as well as in dizygotic twins and ranges from 6-17% and 0-5% respectively\textsuperscript{15}. Collectively, these family data suggest a stronger effect of genetic factors in CD compared to UC and the heritability in CD seems to be more important and led to the notion that knowledge of the risk genes involved could provide important insights into the pathogenesis of IBD.

To identify the gene variants conferring increased lifetime risk for contracting IBD, two statistical genetic approaches have been applied. Genome-wide linkage mapping strategies analyses a relative limited number (300-5000) of known genetic markers (e.g. microsatellites, restriction fragment length polymorphism (RFLP)
associated with a certain phenotype of the disease) to identify genomic regions of a chromosome shared between affected individuals as candidate disease loci. Although this technique was successful in identifying over ten shared chromosomal regions for disease risk\textsuperscript{16-18}, the usefulness of this approach in complex diseases such as IBD is limited, because of the involvement of multiple different genetic interacting risk factors and the non-genetic risk factors. The identified candidate disease loci are often large and contain multiple genes, often without any apparent relation between the gene products in these regions and the actual disease. Nevertheless, using this approach three polymorphisms (or single nucleotide polymorphisms [SNPs]) in \textit{NOD2} (nucleotide-binding oligomerization domain containing 2), previously known as \textit{CARD15} (caspase activated recruitment domain protein 15), were identified as alleles associated with increased propensity of contracting CD and thus \textit{NOD2} was the first example of an IBD susceptibility gene\textsuperscript{19,20}. The cytosolic NOD2 protein plays an important role in the innate immunity and is mainly expressed on epithelial cells, Paneth cells, which are located at the base of the intestinal crypts and antigen presenting cells (APCs), such as macrophages, monocytes and dendritic cells\textsuperscript{20,21}. It functions as an intracellular pattern recognition receptor (PRR) for invading pathogenic bacteria, including commensals residing in the lumen of the intestinal tract that have entered the mucosa. Binding to its major ligand, N-acetyl muramyl dipeptide (MDP), a degradation product of peptidoglycan, the structural component of the cell wall in Gram-positive bacteria\textsuperscript{22-24}, leads to activation of the nuclear transcription factor NF-κB pathway\textsuperscript{20} and mitogen-activated protein kinase (MAPK) pathway, resulting in pro-inflammatory mediators, such as TNF-α, IL-1β and IL-6\textsuperscript{25}. Recent studies have revealed an impaired mucosal clearance of bacteria in \textit{NOD2}-deficient mice\textsuperscript{26}, implicating that NOD2 may play a central role in mucosal immunity. Interestingly, \textit{NOD2} alleles associated with CD displayed reduced rather as enhanced capacity to activate the pro-inflammatory transcription factor NF-κB\textsuperscript{20}. This finding, together with the observation that patients with CD are often defective in functionality of their innate immune system\textsuperscript{27}, led to the suggestion that CD should be considered as an (innate) immune deficiency\textsuperscript{28}. Furthermore, \textit{in vitro} experiments showed that monocytes isolated from CD patients had lower phagocytic activity towards \textit{Candida albicans} than those obtained from healthy individuals\textsuperscript{29}. Moreover, both in humans as well as in animals, genetic defects that provoke reduced innate immunity, e.g. deficiency of NADPH oxidase or Wiskott-Aldrich syndrome, cause CD-like colitis. Taken together, the body of contemporary biomedical literature strongly supports the concept that monocyte dysfunction is to be associated with
the pathogenesis of CD-like and other autoimmunity\textsuperscript{30} (Figure 2). Support from this notion also comes from analysis of other risk genes.

**INNATE IMMUNITY AND AUTOPHAGY**

Recent advances in high-throughput genotyping techniques and increased knowledge about the HapMap Project enabled researchers to perform genome-wide association studies (GWAS) for several complex diseases, with CD leading the way. In these hypothesis-free methods of genome scanning, up to 500,000-1000,000 SNPs across the human genome are examined in both individuals with the disease as in healthy controls. The frequencies of these genetic variants found are statistically compared between the two groups to identify any association with the SNP and disease.

In comparison to genome-wide linkage and association studies describes above, which are restricted to study a relative small number of well-phenotyped patients with a limited number of genetic variants in a few selected genes of suspected involvement in the disease pathogenesis and the lack of power to identify genes with a weak effect, GWAS make use of moderately sized cohorts, thereby increasing the
homogeneity within the studied population and significantly reducing the number of false positives\textsuperscript{31-33}. The GWAS carried out so far, has led to an increased number of known genetic risk factors and these discoveries reveal novel insight regarding to pathways or mechanisms involved in the disease pathogenesis. Because of the genetic contribution is higher for CD then UC, early GWAS focused on CD and led to the identification of more than 30 loci that are associated with CD, an amount that account for approximately 20% of the genetic susceptibility to CD\textsuperscript{34}. Most of the gene variants found support; broadly speaking the concept that reduced function of the innate immune system contributes to the susceptibility to CD.

Remarkably, however, was the detection of the association with CD in multiple GWAS of two genes involved in autophagia, \textit{ATG16L1} (autophagy-related 16-like 1) and \textit{IRGM} (immunity-related GTPase) of CD\textsuperscript{34-37}. Autophagy is a process by which cells encapsulate cytosolic debris, invaded pathogens, or old cellular organelles destined for degradation and fuse these with the lysosomal apparatus\textsuperscript{38}. Defined by the marker rs2241880, a non-synonymous amino acid change (threonine to alanine (T300A)) at position 300 was found that carried all the disease risk for the \textit{ATG16L1} locus and has been replicated in several independent cohorts\textsuperscript{34,36,37,39-41}. Several groups have provided evidence that this genetic association is highly associated with ileal CD\textsuperscript{35,37,42-44}. Interestingly, this SNP resides in an evolutionary conserved domain of the \textit{ATG16L1} protein, located in exon 9 and translated into all known splice variants of \textit{ATG16L1}\textsuperscript{45}. This mutation seems to have a role in the protein stability and its interaction with other member proteins from the autophagic machinery\textsuperscript{46}. \textit{ATG16L1} seems to be broadly expressed in intestinal epithelial cells, lymphocytes and macrophages\textsuperscript{35,37}, although downregulation in \textit{ATG16L1} mRNA expression in colonic CD biopsies\textsuperscript{47}, no significant differences in the levels of protein expression has been observed in intestinal tissue of CD patients versus healthy controls\textsuperscript{35} and the expression of \textit{ATG16L1} was independent of the amino acid substitution T300A\textsuperscript{45}. It is thus reasonable to assume that the susceptibility to CD conferred by change of residue 300 in \textit{ATG16L1} is consequence of altered function of the protein.

A variant for a second autophagy-related gene, \textit{IRGM}, was detected for producing CD susceptibility in a WTCCC (Wellcome Trust Case Control Consortium) study\textsuperscript{36}. In contrast to \textit{ATG16L1}, no causative mutations associated with CD were detected in the coding region of \textit{IRGM}, but a strong non-coding SNP (rs13361189) was found to be in perfect linkage disequilibrium with a 20-kb deletion polymorphism immediately upstream of the gene\textsuperscript{36,37,48}, implicating the involvement of regulatory sequences that control protein expression or post-transcriptional events, such...
as splicing. Compared to unaffected individuals of the reference population, the deletion allele showed an increased frequency in IBD patients, 10% versus 15% and 14% respectively, including association to CD and UC\textsuperscript{48}. IRGM belongs to the INF-γ induced p47-immunity related GTPase family and these IRG-related autophagy genes are not highly conserved, there are only two IRG members found in humans, in contrast to the structural diversity of IRG proteins found in mice\textsuperscript{49}. Studies have shown the importance of IRGM in eliminating intracellular pathogens, such as \textit{Mycobacterium tuberculosis} by INF-γ mediated autophagy\textsuperscript{50,51} and several knockdown and overexpression experiments with IRGM show both an altered efficiency of antibacterial autophagy. IRGM is differentially expressed in several tissues, including colon, small intestine, macrophages and monocytes\textsuperscript{36}, although these expression levels are low and it is difficult to detect endogenous IRGM\textsuperscript{52}. In humans, \textit{IRGM} has a high variation in the genomic upstream sequences, leading to at least 5 splice variants, which could explain the varying expression patterns in different cells types, suggesting that regulation of IRGM is tissue specific and allows spatio-temporal-dependent fine-tuning of IRGM functionality. The functional significance and regulation of the different IRGM isoforms have not yet been elucidated and defining these functions will substantially aid understanding the mechanistic basis of apparent importance of the rs13361189 polymorphism in CD.

Impaired innate immunity predisposes to CD, suggesting that the polymorphisms in autophagia-related genes that confer increased susceptibility to CD may somehow be related to changes in innate immune functionality. The ability of cells to maintain a constant internal environment is dependent on the balance between their synthesis and degradation processes. In eukaryotes, the evolutionary conserved autophagy response is used to keep this homeostasis. Autophagy is a process that enables cells to recycle unnecessary or damaged components in a highly regulated fashion. The purpose of this process is dependent of the environmental conditions the cell resides in. During nutrient-rich conditions, autophagy is simply activated to degrade long-lived or misfolded proteins and to dispose damaged cytosolic organelles, such as leaky mitochondria, thus preventing unwanted apoptosis and even potentially toxic aggregates\textsuperscript{53}. Beside this quality control system as it patrols among the cytoplasm in search of damaged components and promoting protein turnover, it also functions in scaling down certain organelles, which can be present in excessive amounts within the cells, for example peroxisomes as well as the endoplasmic reticulum (ER). In response to cellular stress events, such as nutrient deficiency, autophagy becomes strongly induced in order to supply the cell of nutrients (e.g. amino acids and energy
(ATP)) through catabolism of the cells own constituents\textsuperscript{54,55}. In addition to these physiological described functions of autophagy, maintaining homeostasis and cell survival, it also plays a role in host defense responses by promoting elimination of intracellular pathogens, including viruses, parasites and bacteria in a more selective manner, a process also referred to xenophagy. This defense mechanism has been widely studied in several pathological processes in eukaryotic organisms\textsuperscript{55,56} and is now implicated in a wide range of human diseases, including autoimmunity and inflammatory disorders with a direct relevance to the regulation of innate immune responses\textsuperscript{57-59}.

Several morphologically and functionally distinct forms of autophagy are described, including macroautophagy, microautophagy and chaperone-mediated autophagy, in which its activation is dependent on cargo selection and size and specific components involved. Macroautophagy (hereafter called autophagy) seems to be the process to which IBD susceptibility is linked. This form of autophagy can be divided into three phases; initiation, elongation and maturation. In the elongation phase of autophagy, ATG16L1 complexes with ATG5 and ATG12, and seems to have a central position in the complex\textsuperscript{60}. The complex is located on the outside of the autophagosome double membrane, and is essential for autophagosome formation\textsuperscript{61}. Excess or reduced amounts of the ATG5–ATG12 conjugate seem to inhibit the conversion of LC3-I to LC3-II, which is a necessity and one step further in autophagosome formation and altering the balance between the levels of the ATG5–ATG12 conjugate on one hand and the levels of the ATG16L1-CD susceptibility gene product on the other hand, may cause reduced autophagia and perhaps aberrant innate immunity. The other CD susceptibility gene associated with autophagia is \textit{IRGM}, short for immunity-related GTPase family M. IRGM is not a part of the autophagy process itself, but of the initiation of the pathway. IFN-γ stimulates autophagy through IRGM, providing a direct link between this autophagy regulator and functioning of the immune system\textsuperscript{58,61}.

Specific evidence linking CD-susceptibility polymorphisms in autophagic genes to diminished autoimmunity comes from experiments with siRNA's directed at ATG16L1, which prevented autophagy of \textit{Salmonella typhimurium} in the HeLa cell-line, suggesting that diminished function of ATG16L1 can indeed reduce innate immunity\textsuperscript{37}. Insight into the possible functioning of ATG16L1 was gained from experiments in which mice were engineered to express a hypomorphic allele of \textit{ATG16L1}. These animals exhibited a Paneth cell phenotype\textsuperscript{62}. In general, secreting cell types seem vulnerable to disbalances in autophagocytic machinery. Already
in 1992 it was shown that defective phagocytosis in the adipokinetic hormone secreting cells in locusts cause gross structural and functional problems in this cell type\(^3\) and the same holds true for Paneth cells. Importantly, however, Paneth cells control the small intestine bacterial flora through the production of defensins and thus impaired Paneth cell function may be considered a specialized innate immunodeficiency. Although extrapolating these observations from a highly artificial murine model to the human situation is not completely straight-forward, Cadwell et al. do present evidence that patients homozygous for the ATG16L1 risk allele have a Paneth cell phenotype and also the ileal localization of disease linked to this allele (the colon does not contain Paneth cells) would support such an extrapolation. Indeed, CD is characterized by a specific decrease in ileal Paneth cell α-defensins. Hence, autophagia-related CD might constitute a specialized form of reduced innate immunity-dependent IBD.

Immunity in general, but especially innate immunity, is common denominator for a hotchpotch of processes related to controlling bacterial load. Genetic deficiencies in these processes would be expected to manifest themselves at spatiotemporal and also foreign body-specific fashions. As we get to know better the functioning of the different risk genes for IBD in physiology, we shall also better understand the specific clinical phenotypes associated with such disease and hopefully such insight will contribute to improved therapy.

REFERENCES


43. Fowler, EV, Doecke, J, Simms, LA, Zhao, ZZ, Webb, PM, Hayward, NK and Radford-Smith, GL. ATG16L1 T300A shows strong associations with disease subgroups in a large Australian IBD population: further support for significant disease heterogeneity. *Am J Gastroenterol* 2008;103:2519-2526.


48. McCarroll, SA, Huett, A, Kuballa, P, Chilewski, SD, Landry, A, Goyette, P and Xavier, RJ. Deletion polymorphism upstream of IRGM associated with altered IRGM expression and


