Genetic susceptibility for inflammatory bowel disease across ethnicities and diseases
van Sommeren, Suzanne

DOI:
10.33612/diss.100597247

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2019

Link to publication in University of Groningen/UMCG research database

Citation for published version (APA):

Copyright
Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

Take-down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): http://www.rug.nl/research/portal. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.
CHAPTER 1

Introduction and outline of the thesis
The inflammatory bowel diseases (IBD) are characterised by chronic, relapsing immune mediated inflammation of the gastro-intestinal tract. Although IBD only has two main entities, Crohn’s disease (CD) and ulcerative colitis (UC), it encompasses a heterogeneous clinical disease spectrum with a complex underlying pathogenesis.

Clinical aspects

Patients presenting with idiopathic inflammation in the gastro-intestinal tract are strived to be classified as UC or CD patients mainly according to clinical and endoscopic features.

Inflammation in UC is restricted to the colon, in which the rectum is practically always affected and can extend proximally in a continuous fashion. The main classification is based on disease location: at diagnosis the disease is confined to the rectum (proctitis) in 30–60% of patients, extended inflammation till the flexura lienalis (left sided disease) is seen in 16–45% of patients and extended inflammation beyond the flexura lienalis resulting in pancolitis is seen in 15–35% of patients.1,2 Symptoms of ulcerative colitis can be related directly to the inflammation of the colon or colon with (bloody) diarrhoea as a key symptom. Accompanying abdominal pain occurs more often in colitis, urgency occurs more often in proctitis. Because the inflammation is restricted to the mucosa, UC behaves in a non-stricturing and non-penetrating fashion; however, the severity of inflammation and thereby the severity of symptoms can vary greatly. In severe cases systemic inflammatory reactions (fever, tachycardia) are seen. The most severe form of UC, acute severe ulcerative colitis, is associated with complications such as gross bleeding, toxic megacolon, in which the colon dilatates, and (micro)perforation, and has a mortality of 1%.3 Furthermore, longstanding inflammation of the colon increases the risk off colorectal cancer (CRC), which necessitates surveillance colonoscopies.4 Next to CRC, the main indications for surgery are medication refractory disease and refractory acute severe ulcerative colitis, in which case most often a (procto)colectomy is performed. Despite a range of available medical therapies up to 15.6% of patients undergo surgery within 10 years of the diagnosis.5

A typical aspect of CD is the widespread and discontinuous disease location (mouth to anus), mainly affecting the terminal ileum (47%), the colon (28%) or both (21%). In only 3% of the patients the upper gastrointestinal tract is affected (e.g. above the terminal ileum).6 Moreover disease behaviour in CD is more varied compared to UC. A distinction is made between non-stricturing and non-penetrating (or inflammatory) disease, stricturing disease and penetrating disease. The main classification is based on location and disease behaviour and is made with the Montreal classification.7 The variance in disease location and behaviour explains the variety of symptoms. Inflammatory disease affecting the colon resembles UC. Inflammation in the terminal ileum can result in vague abdominal pain and loss of weight. Stricturing inflammation results in stenosis with accompanying symptoms of nausea, abdominal pain and weight loss. Penetrating disease results through (micro)perforation in abscesses or fistula with skin or other organs. CD often presents with the inflammatory form, the stricturing and penetrating forms are relatively rare at diagnosis (only 17% for stricturing disease and 13% for penetrating disease respectively);6 however, through prolonged inflammation strictures, or penetrating complications occur frequently in the course of the disease and they are frequent indications for surgery, with around 50% of patients needing surgery within 10 years of the diagnosis. Unfortunately, due to the susceptibility of the entire gastrointestinal tract to developing inflammation, CD tends to reoccur, as a result of which one third of the patients require multiple surgeries.8 After multiple resections and with extensive disease, the remaining small intestines might not be sufficient to sustain nutritional uptake, resulting in intestinal failure.9

Although the characteristics of UC and CD described seem to be clearly distinctive, it is not always possible to distinguish between the two
entities. If a clear distinction cannot be made, the disease is called undetermined IBD (IBD-U). Sometimes a definite diagnosis can be made through the course of the disease; however, in some patients this remains uncertain. Furthermore, rectal sparing or the presence of backwash ileitis in UC might be misinterpreted as CD. On the other hand, 5–10% of the patients diagnosed with UC eventually turn out to have CD.

The relapsing nature is shared between CD and UC. Although clinical risk factors have been identified to predict a severe disease course or complications (for example young age at diagnosis, smoking and extensive small intestine disease for CD and pancolitis and deep ulcerations in UC) no individual prognosis can be made.

Adding to the heterogeneity of the disease spectrum are the extraintestinal manifestations (EIM) of IBD, which can affect multiple organ systems and can impact quality of life more than IBD itself. 63.4% of IBD patients suffer from at least one EIM during their lifetime. Some EIMs like erythema nodosum, peripheral arthritis and oral aphthous ulcers occur with active intestinal disease whereas other EIMs like axial spondylarthropathy/ankylosing spondylitis, uveitis and primary sclerosing cholangitis follow their own clinical course. It is believed that there is a shared pathogenic mechanism between IBD and its EIMs; however, this link is not yet clarified. One theory suggests activation of the immune response by translocated bacteria through the inflamed and leaky intestinal barrier. Therefore, a crucial component of the treatment of EIMs is successful treatment of intestinal inflammation. Furthermore, some conditions are frequently seen in IBD patients as a consequence of metabolic abnormalities due to intestinal inflammation. Examples are osteopathy, nephrolithiasis or thromboembolic events. A third category of diseases occurs more often in patients with IBD, but is not IBD-specific like the EIM. These associated diseases like insulin-dependent diabetes mellitus, thyroid disease and vitiligo probably share immune mediated pathways with IBD, possibly due to a shared genetic background.

Epidemiology

IBD is a ‘disease of the young’, with the peak age of onset in the second to fourth decade of life. The incidence has been rising in recent years; in the Netherlands in 2010 the annual incidence for IBD was 40.36/100,000, for CD 17.49/100,000 and for UC 21.47/100,000. The prevalence of IBD in the Netherlands is high, with 830/100,000 inhabitants (CD: 331/100,000; UC: 475/100,000), and is comparable to the prevalence in Europe. IBD has always been considered a disease of industrialised countries, with the highest prevalence in Europe and North America. Recently, and simultaneously with the industrialization of developing countries the incidence of IBD is rising in, for example, Asia. Since IBD occurs more often in urban areas compared to rural areas and the fact that children of migrants from low-incidence to high-incidence countries show an increase in incidence this suggests environmental risk factors are driving factors.

Treatment

Because IBD is an immune mediated disease treatment is focused on suppression of the immune system. The first goal is inducing remission and because IBD tends to flare up, the next goal is to maintain remission in order to prevent long-term damage and complications. This means that patients might have to take lifelong medication. There are multiple definitions of remission and treatment strategies and therapy regimes to achieve this. It falls beyond the scope of this introduction to go into great detail, but a key feature of the medication used is that they are not specific in suppressing the immune system, resulting in significant infectious and oncological risks. In order to expand the therapeutic arsenal and make medication more specific great effort is put into development of new medication. Examples of new biologicals are Vedolizumab, which is a monoclonal antibody that blocks the gut specific α4β7 integrin, and Ustekinumab which targets interleukin 12 and
interleukin 23. Based on new insights in disease pathogenesis, new therapy targets and new drugs are still being developed.

One of the biggest clinical challenges in improving the therapeutic strategy is to select the right treatment for individual patients (personalised treatment). Till this day it remains impossible to predict the effect of a drug on the disease, the duration of this effect and occurrence of side effects in individual patients. This challenge is further complicated by the unpredictable development of anti-drug-antibodies that greatly diminish the clinical effect of monoclonal antibodies like Infliximab. Furthermore, because the disease course is unpredictable patients often develop a significant flare of disease activity before treatment escalation has taken place. Ideally, clinicians would like to offer individual patients a treatment strategy with a predicted good effectiveness and minimal side-effects at the correct time during the disease course.

**Impact**

Considering that IBD is a chronic disease, with the impact of symptoms described above, the intense treatment and the erratic prognosis, its influence on daily life can be tremendous, diminishing patients’ quality of life, their ability to participate in society, sexuality and fertility. The financial burden of IBD is significant, which is not surprising because IBD often affects young people and their ability to work, and taking into account the considerable costs of medication. The costs per patient even exceeds the costs for other chronic diseases like COPD, hypertension and diabetes.

**Pathogenesis**

Similar to the complexity of the disease spectrum is the complexity of the disease pathogenesis. It is believed that IBD is caused by an inappropriate reaction of the immune system in response to commensal gut bacteria in a genetic susceptible host. Besides genetic factors, disease susceptibility is thought to be influenced by the gut microbiome and several environmental factors. Furthermore it is believed that there are complex interactions between the different factors.

**Environmental factors and the microbiome**

The influence of environmental factors (the exposome) has been well established for years, but has been highlighted by the recent increase in incidence in developing countries that adopt the Western lifestyle. Smoking increases the risk of developing CD and is associated with a more severe disease course. On the other hand UC occurs more often in former smokers. Diet can influence the risk of IBD through dietary macronutrients (fibres, saturated fat), micronutrients (vitamin D, zinc) or by altering the microbiome. It is possible that factors like (childhood) antibiotic use and hygienic factors during childhood also mainly influence disease risk through adjusting the microbiome.

The role of the microbiome has only been recognised since a few years, driving a tremendous effort in research in this field. The key feature is a reduction of the diversity of the microbiome in IBD, which is more pronounced in CD than in UC. As mentioned above, the microbiome is sensitive to influences of numerous external and internal factors, amongst which is the genetic background.

**Genetic background**

**Locus identification**

The heritability of IBD was demonstrated by segregation in families and a higher disease concordance in monozygotic twins than in dizygotic twins. Concordance is higher in CD compared to UC, suggesting that the heritable component is higher for CD. Moreover, it was clear that the heritability of IBD doesn’t behave in a Mendelian fashion and, as is typical for a complex disease, multiple genetic factors underlie...
disease pathogenesis. Hereafter, tremendous efforts were undertaken to unravel the genetic background of IBD.

**Linkage and candidate gene studies**

The first IBD gene to be identified in 1996 was NOD2\(^44\), after a locus (called IBD1) was found to be associated with Crohn’s disease. It was identified by a linkage study, in which genetic markers are tested for cosegregation (linkage) with disease status in affected families. In 2001 three causal mutations (R702W, G908R and L1007fs) in NOD2 were found.\(^{45,46}\) The NOD2 locus still remains the locus with the greatest risk in causing disease (or effect size): heterozygotes have an odds ratio (OR) of 2–4, homozygotes have an OR of 20–40.\(^{45,46}\) By means of linkage studies multiple loci were identified; however many of them failed replication. Besides NOD2 IBD5/5q31 (harboring SLC22A4, SLC22A5), IBD3/6p21 (HLA region) showed consistent replication.\(^{47}\)

Results from candidate gene studies also suffered from replication issues. In candidate gene studies markers in genes with suspected involvement in disease pathogenesis are selected. In case of the IBD studies candidate genes have roles in the innate or adaptive immune system or the intestinal mucosal barrier. Only a few associations, like CARD9 and IL18RAP, were consistently replicated,\(^{48}\) and some tested loci that would later turn out to be associated, like the IL10 locus, were missed.\(^{49}\)

**Genome-wide association studies**

With the development of genotyping arrays, the research of IBD genetics took a new turn. In genotyping arrays genetic markers (by means of single nucleotide polymorphisms – SNPs) are strategically selected based on linkage disequilibrium (LD) patterns, enabling the coverage of the majority of the genome by a relatively small selection of SNPs. In a genome-wide association study (GWAS) these SNPs are tested in a large selection of cases and controls to find SNPs that are significantly associated. This study design enables hypothesis-free testing and testing genetic variants with modest to small effect sizes, which wasn’t possible with the linkage and candidate gene studies. In order to correct for the multiple statistical tests done in a GWAS stringent criteria have to be met to claim an association.

Interestingly, the first GWAS in IBD was performed in a Japanese CD population.\(^{50}\) Later studies were primarily focused on European (derived) populations. Therefore, there has remained a relative gap of knowledge in the genetic background of IBD in non-European populations. At the start of the research presented in this thesis in 2010 17 GWAS were published, nine in CD,\(^{50–58}\) six in UC\(^59–64\) and two in early-onset IBD.\(^{65,66}\) As could be expected, many of the candidate genes in associated loci play roles in the innate (CARD9, MST1) or adaptive immune system (PTPN22, ICOSLG) or the intestinal mucosal barrier (CDH1, HNF4α, LAMB1, MUC19). Interestingly, candidate genes in pathways that were not implicated in the pathogenesis of IBD before turned out to be associated, like IRGM and ATG16L1 in CD, which together with NOD2 play crucial roles in autophagy. Autophagy is the process of degradation of damaged cytoplasmic components by the lysosome and has been studied in the context of CD after genetic association became apparent.\(^{67,68}\) The importance of certain pathways is stressed by the large enrichment of candidate genes found to be associated. The Th17/IL23 pathway is such an example with many key genes residing in IBD associated loci (IL23R, IL12B, JAK2, STAT3, IL17REL, CCR6).

Because the first studies included relatively small sample sizes, they first picked up the ‘low-hanging fruit’: the loci with relatively large effect sizes like TNFSF15 in the Japanese population (OR 2.17)\(^{50}\) and IL23R in a combined Jewish and non-Jewish European collection (OR 2.2–3.8).\(^{51}\) As it became clear that most genetic variants had much smaller effect sizes, with ORs below 1.5, it became more important to increase statistical power by increasing sample sizes. Therefore, the next effort was combining GWAS data in meta-analyses. Shortly after starting this research, two large meta-analyses in CD and UC were published, increasing the total number of IBD associated loci to 99.\(^{69,70}\) It turned out that
many loci associated with CD or UC were shared, implicating large overlap in disease pathogenesis. Also, many genetic loci are shared with other immune mediated diseases. This led with the combined effort of multiple research groups to the development of a customised genotyping platform that primarily targets genetic loci involved in 11 immune mediated diseases, the Immunochip. In 2012, a large meta-analysis of previously reported GWAS data, combined with an extra cohort of individuals genotyped on the Immunochip was published.\textsuperscript{71} In total, 75,000 individuals were analysed, which gave enough statistical power to increase the number of IBD associated loci with 71 (to a total of 163 loci). Two-third of these loci are shared between CD and UC, 30 seem to be CD specific and 23 seem to be UC specific.

**Uncovering functional implications**

Although the GWASs and meta-analyses where very successful in identifying disease risk loci, many challenges remained. GWAS platforms were designed to cover large stretches of the genome, the SNPs that act as genetic markers are (mostly) not causal variants but tag causal variants through LD. For some risk loci causal variants have been found, for example the NOD\textsubscript{2} and IRGM loci;\textsuperscript{45,46,72,73} however, for most of the loci the causal variants remain unknown. The dense coverage of loci on the Immunochip aids in solving this problem. Another possible solution is sequencing the locus.

The next challenge is to determine the effect of causal variants and which genes are influenced by genetic variation. The most obvious explanation through which genetic risk variants can influence disease pathogenesis is by influencing gene expression. Many associated SNPs have a strong correlation with gene expression, the so-called eQTL effect.\textsuperscript{74} And although this particular SNP is not likely to cause a disruption in gene expression itself, the eQTL effect can point out the candidate gene for a locus and we can learn more about a genetic risk factor if we understand how gene expression is influenced by that risk factor. Other approaches use network analyses (based on for example gene ontology or protein-protein interaction) to find enrichment of certain pathways in the entire set of genes residing within associated loci.\textsuperscript{75,76} These methods were used mostly to select candidate genes in the published GWAS.

Finally a correlation will have to be made between genetic risk factor, the gene(s) influenced by genetic variation, it’s role in a signalling pathway and the influence of this pathway on disease pathogenesis.

**Aim and outline of the thesis**

The complexity of the clinical disease spectrum and the response to treatment are probably a reflection of (subtle) differences in disease pathogenesis. In order to gain insight into the disease development and disease course and improve quality of life by realising personalised treatment, we need to gain full insight in the pathogenesis of IBD.

In this thesis we focus on expanding knowledge about the genetic basis of disease pathogenesis.

Part I starts by describing our research focused on identification of genetic risk loci and candidate genes and explains how the effect of genetic variation on gene expression plays a role. Chapter 2 is a replication study of three previously reported associated loci resulting from a genome-wide association study (GWAS) performed in ulcerative colitis (UC). The candidate genes residing in the loci (CDH\textsubscript{1}, HNF\textsubscript{4}A and LAMB\textsubscript{1}) play roles in the intestinal mucosal barrier integrity. We analysed these loci in an independent Dutch UC cohort. Chapter 3 describes the largest meta-analysis published at the time, containing nearly 100,000 individuals including a substantial proportion of individuals from East-Asian, Indian and Iranian descent. This study plays two roles: identifying new IBD risk loci and investigating the overlap in genetic background of IBD across populations.

Next we focus on the role of gene expression. In the study presented in chapter 4 we selected SNPs with known influence on gene expression from GWAS results that did not reach the
genome wide significant threshold before. We hypothesised that these genetic loci were more likely to be truly associated to IBD because they have functional consequences. We performed a replication study in an independent Crohn’s disease (CD) cohort. In the research presented in Chapter 5, we investigated the effect of genetic risk on gene expression of genes in the Th17/IL23 pathway. This pathway is very important in IBD from both a biological and genetic perspective.

In part II, we focus on shared genetics with other diseases. The clinical spectrum of IBD is broad and includes several extraintestinal manifestations. Chapter 6 presents the overlap in genetic factors and sharing of biological pathways between IBD and its extraintestinal manifestations. In Chapter 7 we focus on a disease occurring after allogeneic hematopoietic cell transplantation and clinically resembling Crohn’s disease: gastro-intestinal graft-vs.-host disease. We identify risk loci and examine whether there is overlap with IBD.

Chapter 8 gives an overview of the results of the work presented and provides a general discussion and future perspectives.

REFERENCES


58. The 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.


68. Kuballa P, Huett A, Rioux JD, et al. Impaired autophagy of an intracellular pathogen induced by


Part 1

Identification of genetic risk factors for Inflammatory Bowel Disease and the influence on gene expression