

University of Groningen

Down the line from genome-wide association studies in inflammatory bowel disease

Spekhorst, Lieke; Visschedijk, Marijn C.; Weersma, Rinse; Festen, Eleonora

Published in:
Expert review of clinical immunology

DOI:
[10.1586/1744666X.2015.990439](https://doi.org/10.1586/1744666X.2015.990439)

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
Final author's version (accepted by publisher, after peer review)

Publication date:
2015

[Link to publication in University of Groningen/UMCG research database](#)

Citation for published version (APA):

Spekhorst, L. M., Visschedijk, M. C., Weersma, R. K., & Festen, E. A. (2015). Down the line from genome-wide association studies in inflammatory bowel disease: the resulting clinical benefits and the outlook for the future. *Expert review of clinical immunology*, 11(1), 33-44. [25493555]. DOI: 10.1586/1744666X.2015.990439

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.

EXPERT
REVIEWS

Down the line from genome-wide association studies in inflammatory bowel disease: the resulting clinical benefits and the outlook for the future

Expert Rev. Clin. Immunol. Early online, 1–12 (2014)

Lieke M Spekhorst^{1,2,‡},
Marijn C Visschedijk^{1,2,‡},
Rinse K Weersma*¹ and
Eleonora Anna Festen^{1,2}

¹Department of Gastroenterology and Hepatology, University of Groningen and University Medical Centre Groningen, 9700 RB Groningen, The Netherlands

²Department of Genetics, University of Groningen and University Medical Centre Groningen, 9700 RB Groningen, The Netherlands

*Author for correspondence:
r.k.weersma@umcg.nl

‡Authors contributed equally

Inflammatory bowel disease (IBD), consisting of Crohn's disease and ulcerative colitis, is a chronic inflammatory disease of the gut. The etiology of IBD is complex, involving genetic as well as environmental factors. Genetic studies have identified 163 genetic risk loci for IBD, which have led to new insights into the biological mechanisms of the disease. The currently known IBD risk loci show an almost 75% overlap with genetic risk loci for other immune mediated diseases. Current studies are focused on the translation of the identified risk loci to clinical practice. The first steps towards this translation are being taken with the identification of genetic risk factors for drugs toxicity, specific disease course and response to therapy. In this review we will discuss how the IBD genetic risk loci were identified and how this knowledge can be translated towards clinical practice.

KEYWORDS: crohn's disease • drug targets • GWAS • immune mediated diseases • immunochip • inflammatory bowel disease • risk models • ulcerative colitis

Inflammatory bowel disease (IBD) is a chronic immune mediated disease affecting the gastro-intestinal tract. The prevalence of IBD in the Western World, is approximately 1 in 1000 individuals and there is an increasing trend in incidence and prevalence in developing countries [1]. IBD consists of two distinct diseases; Crohn's disease (CD) and ulcerative colitis (UC), which have some overlapping clinical and pathological features. In CD the inflammation can occur throughout the entire gastro-intestinal tract, the inflammation can affect all mucosal layers and can be complicated by strictures, and formation of abscesses and fistula. In UC the inflammation is limited to the colon and only affects the upper mucosal layer. Formation of abscesses and fistula does generally not occur in UC and stenosis is a very rare complication in UC [2,3].

IBD is a complex disease, meaning that its aetiology is multifactorial: genetic, epigenetic

and environmental factors interact and give rise to the disease. Environmental factors like smoking, medication, appendectomy, exposure to pollution, and diet have been implicated to play an essential role in the pathogenesis of IBD [4]. Smoking and prior appendectomy have been proven to be protective in UC, but in contrast smoking can aggravate inflammation in CD and increase the risk for CD [5]. The composition of the gut microbiota, is also likely to be a major factor in IBD disease pathogenesis [6–9].

Over the preceding decades, co-occurrence and familial aggregation of IBD was observed indicating that IBD has a strong hereditary component [10]. As the field of genetic research evolved rapidly it revealed new insights into pathogenesis of IBD. Initially genetic linkage studies were performed in families with an extremely high prevalence of IBD, which led to the identification of the first genetic risk

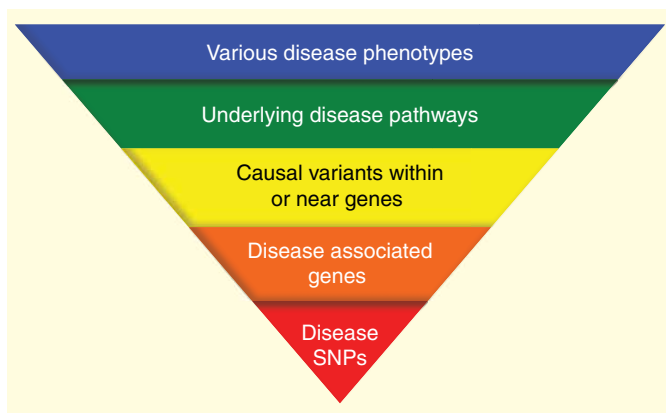


Figure 1. The steps leading from findings of genome-wide association studies (GWAS) to clinically significant outcomes.

From bottom to top: In red: GWAS identify single nucleotide polymorphisms (SNPs) that are associated to the disease. In orange: the next challenge is the identification of the gene that correlates with the associated SNP. In yellow: then causal variants can be identified that influence the disease associated genes. In green: once it is known which genes are involved in the disease underlying disease pathways can be constructed through co-expression and protein-protein interaction analysis. In blue: finally different genetic backgrounds can be identified that lead to slightly different disease mechanisms resulting in diverse disease phenotypes.

locus for IBD, containing the gene *NOD2* [11–13]. As genome-wide-associations studies (GWAS) became available, hypothesis-free testing of common genetic variants for association to disease became possible [14]. GWAS were very successful in IBD: within four years 99 genetic risk loci for IBD had been identified, making it the most successfully GWA studied immune-mediated disease [15]. Subsequently a new genotyping chip was designed which targeted areas on the genome likely to play a role in immune mediated disease: the Immunochip. Again this effort has been very successful in IBD: 163 genetic risk loci have been identified to date [16].

The genetic risk loci identified for IBD so far have shed new light on the biological pathways underlying the disease. The translation of all of this knowledge on the genetic background of IBD towards clinical practice is still difficult. The steps from GWAS identified genetic disease risk variants to clinical practice are outlined in FIGURE 1. Currently projects are being undertaken to clarify the exact effect of genetic risk variants on the phenotypic variation seen in IBD. A better understanding of how genetic risk variants lead to different disease phenotypes can help us predict disease course per individual, which in turn can help us determine the most adequate treatment for each individual.

In this review we will first consider the clinical aspects of IBD. We will then highlight the new insights that GWAS and the Immunochip have brought us for IBD. We will discuss the overlap between IBD genetic risk loci and their underlying biological pathways and the genetic risk loci and pathways for other immune mediated diseases. Finally, we will discuss how these pathways can be used for therapeutic targeting and how this can help pave the way towards ‘personalized medicine’.

The clinical presentation of IBD

IBD is a chronic mucosal inflammatory disease of the gastrointestinal tract, characterized by periods of remission and relapse. Because of this dynamical character of the disease, patients can experience severe symptoms during an exacerbation, while symptoms can be mild or absent during remission. Active IBD can cause symptoms like abdominal discomfort, diarrhoea, weight loss, rectal bleeding and fatigue. In addition to inflammation of the gut, 25% of the patients have extra-intestinal manifestations (EIM). Arthralgia is the most common EIM, but ophthalmological and primary mucocutaneous EIMs are also common [2,17,18]. Direct disease symptoms and EIMs influence psychosocial functioning and might result in a significantly lower quality of life and loss of work productivity [19,20].

Disease remission can be induced and sustained by medical treatment such as mesalazine, corticosteroids, and immunosuppressants like azathioprine and anti-TNF antibody therapies [21]. Nevertheless, up to 20% of the UC patients and almost 50% of the CD patients will need surgery within 10 years after diagnosis because of refractory disease, fibrostenotic disease, complications or development of colorectal carcinoma [22].

IBD therapy is complicated by the fact that IBD is a heterogeneous disease with a variety of clinical phenotypes, each of which require specific treatment. Currently the treatment paradigm of IBD is shifting from treating symptomatic patients (the ‘step-up’ approach) to starting intensified treatment regimens early in the disease to prevent complicated disease behaviour and flares of the disease (the ‘top-down’ approach) [23]. Before starting this top-down treatment it is essential to select those patients at risk for severe disease, to avoid high costs, ‘over treating’ and potentially severe side effects. At this moment one of the major issues that clinicians face is the selection of these patients at risk for a severe disease course. There are no clinical parameters or biomarkers which predict how the disease will develop, so selecting these high-risk patients is complex. Some known risk factors like early age of onset, familial occurrence of IBD and extensive disease at presentation are considered to be predictive for severe disease [24,25]. However, these factors offer a relatively slim foundation for aggressive treatment with expensive drugs with potentially severe side-effects [26]. Better understanding of the influence of environmental, genetic, and microbiomic factors on IBD phenotype will not only improve choice of treatment for IBD patients but will also enhance our understanding of the underlying disease mechanisms.

Genetic studies prior to the genome wide association scans

As mentioned previously, family studies showed that genetic factors play an important role in IBD risk: the occurrence of CD and UC in first-degree family members is respectively 10-fold and 4-fold higher compared to the general population and in monozygotic twins the concordance of CD and UC is respectively 30 and 15% [10].

The first genetic studies in IBD were linkage and candidate gene association studies. Linkage studies map genetic risk loci

by testing a series of marker alleles for co-segregation (linkage) with disease status through different generations in a family. In candidate gene association studies a candidate gene is selected based on what is known about the disease mechanism. This gene is then tagged with common genetic variants, and tested for association to a trait. A total of 10 linkage and association studies for IBD were performed in the period from 1996 to 2004 [27]. Compared to similar studies in Mendelian diseases the yield of these studies in complex disease might be considered disappointing. This is mainly caused by the fact that the effect size of genetic risk variants for complex disease tends to be much smaller than that for Mendelian disease. Nonetheless, linkage studies identified the first risk gene for CD: *NOD2* on chromosome 16 [11–13]. In 2001 three low frequency coding mutations (R702W, G908R and 3020insC) in the *NOD2* gene were found to be independently associated with CD in Caucasian patients [12,13,28]. These three variants lead to odds ratios (ORs) for CD between 2 and 4 in individuals heterozygote for the variants, and to ORs between 20 and 40 in individuals homozygote for the variants. Nowadays *NOD2* variants remain the strongest genetic risk variants for CD.

Alongside *NOD2*, linkage studies suggested a link between IBD and three other genetic loci with a lower effect size: the IBD3 locus (in the *HLA*-region), the IBD5 locus (containing *SLC22A4*, *SLC22A5* and other genes), and a locus on chromosome 5q31 (containing no genes; a 'gene desert') [27].

Genome wide association studies

In the early 2000s tremendous technological advances and the progress of the Human Genome Project provided the opportunity to perform GWAS [29,30]. GWAS typically focus on associations between complex traits and 100,000–500,000 single-nucleotide polymorphisms (SNPs), selected to tag a maximum of genetic variation over the whole genome. A SNP is a DNA sequence variation, which occurs commonly in a population. A GWAS looks for statistically significant differences in allele frequencies of these SNPs between a large number of individuals with disease status (cases) and healthy controls. The associated SNPs mark genomic loci (which can contain several genes) in the human genome, which influence the risk of disease. Unlike linkage studies, GWAS are not restricted to sibling pairs and families, and thus have more statistical power to detect genetic risk loci of small to moderate effect sizes. Due to correction for multiple testing, there is a strict protocol for replication and the genome-wide statistical significance association for true positives was set to a P-value $<5 \times 10^{-8}$ [31].

The first GWAS study in European ancestry CD patients confirmed *NOD2* as a CD risk gene. Moreover, it identified the association between CD and a locus containing *IL23R*, encoding a pro-inflammatory cytokine, which stimulates T-cells towards chronic inflammation [32]. The most remarkable pathway discovered through early CD GWAS studies is the autophagy pathway, which was discovered through associations at loci containing *ATG16L1* and *IRGM* [33–35]. The early GWAS studies in UC showed substantial overlap of genetic risk

background with CD (*IL23R*, *IL12B*, *NKX2-3* en *MST1*), but also some UC specific loci (*IL10*, *HLA*). *NOD2*, *ATG16L1* and *IRGM* remain CD specific loci [36–39].

A GWAS typically uses approximately 500–2000 cases and a similar number of controls genotyped at 100,000–500,000 SNPs. To increase the power of GWAS to detect more genetic risk variants a meta-analysis of all previously published CD GWAS was performed, which included 3000 cases and 5000 controls. This meta-analysis confirmed 11 previously putative loci and helped discover 21 new CD risk loci, including loci containing *STAT3* and *JAK2* [40].

The appreciation of the need to further increase sample sizes led to increased collaboration through the International IBD Genetics Consortium (IIBDGC) [41] to bring together investigators and GWAS datasets from IBD genetics groups around the world. The IIBDGC published three GWAS meta-analyses between 2008 and 2011. A CD meta-analysis of six GWAS with a total sample size of ~50,000 individuals identified 30 new loci, bringing the total count of CD genetic risk loci to 71 [42].

Similarly a meta-analysis of six GWAS studies in UC with a sample size of ~17,000 individuals identified 29 new UC loci. This increased the number of UC genetic risk loci to 47 [15].

These GWAS meta-analyses increased the total number of confirmed IBD risk loci to 99, including at least 28 shared association signals between CD and UC.

ImmunoChip

As GWAS results for other immune mediated diseases followed, it became clear that some risk variants were disease-specific, but that most risk variants were shared between IBD and various other immune mediated diseases. In 2011 approximately 51 of the thus far identified 99 IBD risk loci turned out to be shared with 23 different diseases, most of which immune mediated diseases [43,44]. This concept formed the basis for the development of the ImmunoChip: a chip composed of all genetic variants correlated to immune mediated disease [45,46]. The ImmunoChip covers almost 200,000 SNPs for 12 distinct immune mediated diseases (CD, UC, autoimmune thyroid disease, ankylosing spondylitis, celiac disease, IgA deficiency, multiple sclerosis [MS], primary sclerosing cholangitis [PSC], primary biliary cirrhosis [PBC], psoriasis, rheumatoid arthritis [RA], systemic lupus erythematosus [SLE] and Type 1 diabetes [T1D]). The ImmunoChip was designed to densely genotype immune mediated disease risk loci with common genetic variants [45]. The ImmunoChip project had two main goals. The first goal was to validate the already identified disease risk loci and to establish previously putative genetic risk loci as definite genetic risk loci by testing them in a large number of new cases and controls, a process termed 'deep replication'. To achieve this goal the top ~3000 associated SNPs for each disease known from GWAS and GWAS meta-analysis were tested in a large number of case and control samples that had not been tested in previous GWAS. The second goal of the project was to fine map each risk locus to identify the most likely focus of the genetic variant that is actually causal to the disease

association. To achieve this goal each known genetic risk locus was densely covered with SNPs on the Immunochip [46].

The Immunochip, though purposefully designed, has a few limitations that should be taken into consideration.

First of all, not all loci are densely covered; putative genetic risk loci are generally only covered by a handful of SNPs. This means that, especially in the putative loci, the SNP showing the strongest association to disease is unlikely to be causal, and more likely to be in 'linkage disequilibrium' with the causal SNP. Linkage disequilibrium means that the most strongly associated SNP inherits with the causal variant because they are on a stretch of DNA that does not break during cell replication. A second limitation of the Immunochip is that it is less sensitive in non-European ethnic groups, because the SNPs have been selected from genome reference sets based on individuals of European origin. A third limitation is that the Immunochip only includes relatively common genetic variants (minor allele frequency >0.5%), whereas more rare variants can confer larger effects on disease risk. A final limitation, inherent to the design of the Immunochip, is that it does not cover the whole genome, but only focuses on known immune disease genetic risk loci [46].

The number of new loci that can be identified with genome wide significance, that is with confidence that the finding is not a false positive, depends in part on the sample size and in part on the allele frequency of the genetic variants tested [47]. The Immunochip project identified 64 new risk loci for IBD increasing the number of known IBD genetic risk loci to an impressive number of 163 risk loci. Of these 163 risk loci, 30 are specific for CD and 23 for UC. The other 110 loci are shared by both diseases, implying that a shared biological mechanism plays a role in both phenotypes [16]. Whereas the Immunochip substantially increased the number of known genetic risk loci for IBD, these risk loci explain only a minority of the genetic variance in disease risk: 13–13.5% in CD and 7.5–9% in UC [16,48]. The fact that we can only explain such a small amount of risk variance in IBD suggests that other factors, like rare risk variants not tagged by the chips used so far, or interactions between genetic variants and environmental factors, also play a substantial role in the risk for IBD. Interactions within and between genetics and environment modulate the risk for a disease. These interactions are still poorly understood, which makes, for example, predicting disease risk with genetic risk variants complex.

Besides identifying new shared and unshared risk loci for several immune mediated diseases, the Immunochip also detected some highly interesting discordant associations, that is, instances in which a risk locus is shared for two diseases but where the associated variant conveys risk in one disease and is protective in the other disease. Some interesting cases of discordant associations are seen between CD and UC. Although the clinical phenotypes of the diseases are clearly overlapping, and shared risk loci are thus to be expected, a number of shared loci were found that showed a risk effect in opposite directions for each disease. An example is the locus containing the

PTPN22 gene (encoding the protein tyrosine phosphatase) [49]. The main risk variant in this locus, Arg620Trp, increases risk for UC, but is protective for CD [16,50]. The mechanism underlying these discordant associations has not yet been clarified.

Immune mediated diseases

IBD is, as described earlier, a chronic inflammatory disease caused by an excessive inflammatory reaction to the host's own gut microbiome. This makes IBD one of a large range of what can be termed immune mediated disorders: chronic inflammatory diseases caused by an inflammatory reaction to an antigen that, in healthy individuals, is tolerated or quickly disposed of. For some of these diseases the antigen is known, such as the gluten protein in celiac disease. Celiac disease represents a special case of an immune mediated disorder, because the antigen can be avoided with a gluten-free diet, thereby more or less 'curing' the disease. IBD also forms a special case among the immune mediated diseases: the instigative antigen is generally assumed to be the commensal flora of the gut. However, unlike gluten, the commensal flora of the gut is not something one can avoid. For most other immune mediated disorders, the instigative antigen is unknown, complicating the unravelling of the disease mechanism. For RA, ankylosing spondylitis, SLE, T1D and autoimmune inflammatory disease of the thyroid, antigens have been proposed but not conclusively proven.

Although each of these diseases have different phenotypes and might have different instigative antigens they share common inflammatory pathways. Immune mediated disorders are known to co-occur within families or even within individuals, suggesting they also share a common genetic background.

In 2008, early in the GWAS era, results of the different immune mediated disorders already revealed 23 genetic risk loci to be shared by two or more immune mediated diseases (ankylosing spondylitis, asthma, auto-immune thyroid disease, coeliac disease, MS, psoriasis, RA, SLE, T1D, CD and UC) [44]. As data accumulated, in 2011 approximately 51 IBD genes were found to be shared with 23 different diseases, including immune mediated disorders, infectious diseases and other gastro-intestinal disorders [43,51]. Noteworthy among these shared genetic loci are the risk loci that encode proteins from the adaptive immune system (*IL23R*, *IL10*, *IL12B*, *IL27*, *IL18RAP*), these loci seem to play a role in a shared pathway for CD, UC and several other immune mediated diseases [43].

The Immunochip once again showed that shared genetic risk loci might not have the same effect in each disease. Besides the previously observed correlated and discordant associations (the same haplotype is protective in one disease but increases the risk for another disease), it showed association signals that are non-correlated (different risk haplotypes are seen in shared risk loci), and association signals that are correlated and concordant (a risk variant increases the risk for more than one immune mediated disease) [50]. This phenomenon is called pleiotropy, meaning that seemingly unrelated phenotypes can be derived from the same risk variants and that seemingly related phenotypes can derive from divergent risk variants. The

mechanism underlying this pleiotropy might be that the combination of different genetic risk variants and environmental factors determine the immune mediated disease that a patient will develop.

Almost three-quarters of the IBD loci were found to be overlapping with other immune mediated diseases [43,50] and 71 loci were associated with two or more immune mediated diseases (IBD, ankylosing spondylitis, coeliac, psoriasis, RA, Type 1 Diabetes). Of these 71 loci 45% resulted in an increased risk, 14% had opposites effects (protective versus risk variant) and 42% shared the same loci but with different risk haplotypes [50].

Pathogenetic pathways in IBD

GWAS and Immunochip meta-analysis have revealed 163 risk loci for IBD, most of which are shared for CD and UC and which identify key regulating pathways for both diseases. To advance our understanding of the mechanisms underlying disease we have to carefully select and study candidate gene(s) within each locus to see how these contribute to disease susceptibility. The most replicated and confirmed loci have been extensively studied and their candidate genes and their pathways reveal possible disease mechanisms for CD and UC also in functional studies and mouse models.

One of the strongest susceptibility loci in CD is the locus containing *NOD2* [12,13]. *NOD2* is located on chromosome 16 and encodes an intracellular pattern-recognition receptor of the innate immune system [52,53]. This receptor recognizes viral and microbial components; maintaining the gut mucosal barrier through regulation of microbiome homeostasis and activation of the innate immune response [54]. One of the mechanisms through which *NOD2* regulates microbiome homeostasis is by the production of antibacterial defensins [55]. *NOD2* receptor signalling against microbial components depends on its intracellular localization. *NOD2* risk variants cause a disrupted receptor, which make the receptor unable to recognise intracellular bacteria. This probably leads to dysbiosis of the intraluminal contents, thereby causing an inappropriate immune response [13,56,57]. Carriage of two *NOD2* risk variants also increases the risk of a severe CD disease phenotype resulting in penetrating and/or stricturing disease [58].

Another important pathway in the disease pathogenesis of CD that has been discovered through genetic studies is autophagy (*ATG16L1*, *SMURF1*, *LRRK2* and *IRGM*) [33,59,60]. Autophagy describes a cellular pathway in which organelles and foreign proteins are being delivered to the lysosome in the cell for degradation, which makes it a crucial immune defence mechanism [61,62]. Moreover there is a functional link between *ATG16L1* and *NOD2*, as *NOD2* recruits *ATG16L1* to the plasma membrane at the bacterial site of invasion to initiate autophagy. Several risk variants (SNPs), associated with *NOD2* and *ATG16L1* have been found to affect bacterial autophagy. This implies that deficient bacterial autophagy plays a key role in the disease pathogenesis of CD [59,63,64]. The autophagy related *NOD2* and *ATG16L1* variants are specific to CD, but

other autophagy related genes are associated with both CD and UC (*SMURF1*, *LRRK2* and *IRGM*) [16,65].

The *IL23R* genetic IBD risk locus is part of an important disease pathway for IBD: T-helper 17 (Th17) signalling. The *IL23R* gene in this locus encodes an IL23 receptor whose subunit interacts with interleukin 23 (IL23) a pro-inflammatory cytokine. IL23 regulates the immune response against exogenous antigens in the gut, by inflammation through the production and differentiation of Th17 lymphocyte cells [66,67]. Mutations or variation in or near the *IL23R* gene are hypothesized to cause an inappropriate immune response to the commensal flora of the gut [32,68,69]. Loci with genes encoding proteins with functions downstream in the IL23-Th17 pathway have also been identified as IBD risk loci. These genes affect components of the IL23 pathway that are expressed in Th17, Th1 and other innate lymphoid cells. Among these IBD risk loci are loci containing *JAK2* (Janus kinase 2) and *STAT3* (signal transducer and activator of transcription 3). The activation of the IL23R at the cell surface activates a secondary intracellular signalling pathway JAK2/STAT3. This *JAK/STAT* pathway plays a role in the innate and adaptive immunity, and particularly in the progression of inflammation in the Th17 cell pathway [70]. Several genes encoding pro-inflammatory cytokines in the Th17 pathway, which have been shown to be overexpressed in IBD, have also been implicated in the genetic background of IBD: *IL22*, *IL21* and *IL26* [71-76].

'Ten-year' review: genetic research translated to the clinic

In this post-GWAS era, our knowledge on the molecular background of IBD has progressed to such a high level, that an inevitable and crucial question arises: How do we translate this knowledge to clinical practice to improve treatment of IBD or even to detect IBD early and prevent it from progressing to a full blown inflammatory disease (FIGURE 2) [77]?

Screening individuals at risk

Could we screen the general population for risk of IBD with the genetic risk variants that have been identified? The answer is that we could, in theory, but the predictive value of such a genetic screening test would be very low, because, as mentioned earlier, the genetic variants identified so far explain less than 20% of the genetic risk for IBD [39,78]. To improve the predictive power of this predictive test by enriching it with environmental risk factors is also not yet feasible, since our knowledge on environmental risk factors for IBD is still relatively poor. So screening the general population for IBD risk is not yet possible. Moreover, one could wonder whether we really should screen the general population for a disease that we cannot (yet) prevent. Genetic screening of individuals at high risk for IBD, for example children from families with a high prevalence of IBD with the currently known IBD risk loci, is similarly unlikely to be successful. The currently known genetic risk variants for IBD are genetic variants that are common (minor allele frequency >1%) in the general population, and hence will

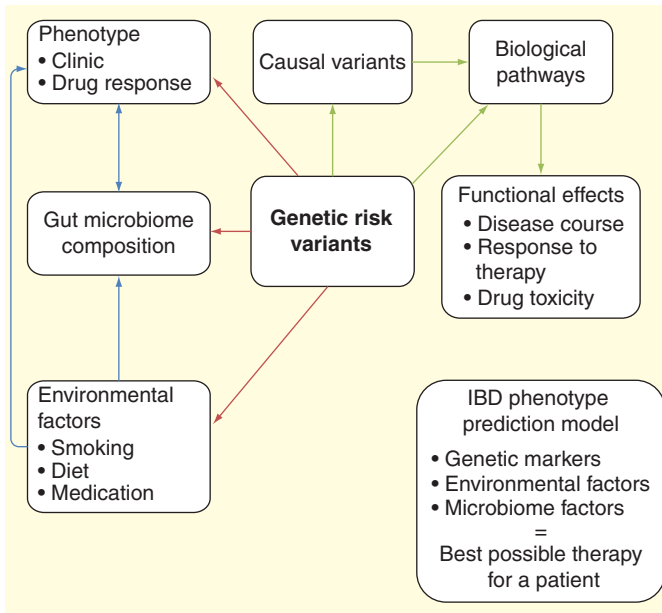


Figure 2. From genetic risk variants to the clinic. The red arrows show the interaction of genetic risk variants with other factors leading to inflammatory bowel disease (IBD) phenotypes. De blue arrows show the interaction of the non-genetic IBD risk factors with IBD phenotype. De green arrows show the route research will have to take down the line from the identification of genetic risk variants to clinical practice. The square in the lower right corner of the figure shows how our current knowledge could lead to 'personalized medicine' i.e., the best possible therapy for a patient.

not provide a test with a high specificity. Moreover, in families in which IBD is highly prevalent the underlying genetic risk variants are likely to be rare variants that may well have not yet been identified. Finally screening in high-risk individuals would only be useful if minimally invasive or preventative therapy such as dietary intervention would be available. Unfortunately, dietary interventions have not been proven effective in IBD, so early detection of IBD would only lead to earlier recognition of disease, followed by standard treatment [79].

Drug development

While our knowledge on the genetic background of IBD cannot be used for risk prediction, it does provide important insights into the disease mechanism of IBD. Each genetic risk locus is a potential drug target for IBD therapy. Before such targeted drugs can be developed, the target loci have to meet several important criteria.

The first criterion for the development of targeted drugs is that the causal variant in the genetic locus has to be known and the effect of this causal variant on the biological function of the gene it affects has to be understood [80]. Currently, we cannot yet meet this criterion: the genetic variants that have been identified by GWAS are variants with a high prevalence in the general population. Hence these variants are very suitable for mapping genetic variance between cases and healthy controls but they are unlikely to have a direct effect on the

function of a gene. This means that the 163 genetic variants that are associated to IBD are not the true causal genetic variants that affect the genes in these loci but that these 163 variants are correlated to the causal genetic variants in these loci. Fine-mapping of these genetic risk loci and targeted re-sequencing of genetic risk loci in large cohorts of IBD cases and healthy controls is currently being performed and will lead to better understanding of the causal link between the IBD associated common genetic variants and the disease mechanism (FIGURE 1). The translation from common genetic risk variant to clinically significant output initially involves the identification of the gene that the variant is likely to be correlated to by gene prioritizing. The causal gene is likely to have a differential expression as a result of the genetic variant, i.e., by identifying expression Quantitative Trait Loci (eQTL). The causal gene might also be identified because it encodes proteins that are known to interact with proteins known to be involved in IBD mechanisms, i.e., by protein-protein interaction (PPI) or through Gene Relationships Across Implicated Loci (GRAIL).

The second criterion for developing targeted drugs is that the causal genetic variant, or its effect on gene function, is druggable, and that targeting this mechanism does not lead to adverse events [80]. Some CD risk variants might for example cause an inadequate innate immune response, but upregulating this innate immune response with a drug might lead to increased inflammation caused by an exaggerated innate immune response. Other CD risk variants are located in gene deserts where the identification of a druggable candidate gene might be impossible.

In short, the road from the currently identified IBD genetic risk variants to the development of new IBD drugs seems long and hardly cost-effective. In spite of this there are several examples of approved drugs for which genetic studies identified the drug target to be associated with IBD, while the drug had already been developed independently from genetic knowledge (TABLE 1). Ustekinumab (Janssen-Cilag), a human antibody against IL-12 and IL-23, has for example been shown to be effective in the treatment of CD [81,82]. Ustekinumab blocks binding of IL-12 and IL-23 and thereby blocks the inflammatory cascade downstream of these interleukins. The importance of this inflammatory pathway had already been observed in the genetic background of IBD since *IL23R* and *IL12B* are important risk loci for IBD.

GWAS identified genetic loci containing *JAK2*, *STAT3* and *TYK2* genes as risk loci for IBD. The proteins encoded by these genes, JAK2, STAT3 and TYK2, play crucial roles in secondary pro-inflammatory signalling after activation of the IL23 pathway, as described earlier. Tofacitinib (Pfizer) a selective JAK-inhibitor seems to be effective in the induction and maintenance of remission of IBD and is currently being tested in Phase III trials [83,84].

The autophagy pathway was identified as an important disease mechanism for CD through the recognition of *ATG16L1*, *NOD2* and *IRGM* as CD risk genes. The genetic risk variants in these loci seem to lead to less effective autophagy and

Table 1. New inflammatory bowel disease drugs based on known genetic risk factors.

Potential drugs	IBD risk genes targeted	Effect	Study status
Ustekinumab	Blocking binding of IL12-IL23	Blocks the inflammatory cascades down stream	Effective in CD 82,83
Tofacitinib	Blocking selective JAK2 inhibitor	Blocking secondary pro-inflammatory signalling after activation of the IL23 pathway	Phase III trial 84,85
mTor inhibitor	ATG16L1, NOD2 and IRGM (autophagy)	Up-regulation of the autophagy pathway	Showed no efficacy 86,87
Denosumab	Apoptosis regulatory gene TNFSF11	Regulation T-cell/dendritic communication	Candidate for testing therapy 88,89

This table lists new Inflammatory Bowel Disease (IBD) drugs, the genetic risk locus (indicated with their most likely candidate gene) that they target, their presumed effect and the current status of their studies.

ATG16L1: Autophagy related 16-like 1; CD: Crohn's disease; IL12: Interleukin 12; IL23: Interleukin 23; IRGM: Immunity-related GTPase family M protein; JAK2: Janus kinase 2; NOD2: Nucleotide-binding oligomerization domain-containing protein 2; TNFSF11: Tumor necrosis factor (ligand) superfamily, member 11.

consequently lead to less efficient disposal of invasive microbes. Everolimus (Novartis) and Sirolimus (Pfizer) are both mammalian target of rapamycin (mTOR) inhibitors and up-regulate the autophagy pathway [85,86]. Both drugs are registered for immune suppression after solid organ transplantation. Because of their known immunosuppressant effects and their specific effect on the autophagy pathway the drugs were tested in CD. Although case-reports of the treatment of CD with Sirolimus seemed promising, a randomized case-control study with Everolimus was terminated early because the drug showed no efficacy [85,86]. One could speculate that in the future such trials should be repeated but then including only cases with impaired autophagy e.g., carrying the *ATG16L1*, *NOD2* or *IRGM* risk variants.

Finally the homing of leukocytes to the gut seems to be an interesting disease mechanism since genetic risk loci containing *CCR6* and *CXCR5*, genes encoding proteins involved in this process, are associated to IBD [16]. Before this genetic knowledge was available drugs that impede leukocyte migration to the gut had already been developed. Initial trials of Natalizumab (Biogen), an α -4 integrin inhibitor, showed promising results for inducing and maintaining remission in CD [87,88]. However serious side-effects were observed: the drug also prevents migration of leukocytes to the central nervous system, which increases the risk of severe infections of the central nervous system [89]. Vedolizumab (Millennium Pharmaceuticals, Inc) is a more specific integrin inhibitor: it specifically inhibits α -4- β -7 receptors, which makes this drug a specific inhibitor of leukocyte migration to the gut. Vedolizumab was shown to be effective for induction and short-term maintenance of remission of disease in UC patients and CD patients with prior failure on anti-TNF therapy [90–92].

Drug repositioning

As mentioned earlier, developing new drugs targeted on IBD genetic risk loci will be a long and difficult process. The case of the testing of the mTOR inhibitors Everolimus and Sirolimus in CD however perfectly illustrates an alternative scenario: identifying alternative or refined indications for existing drugs approved for other indications. This process, called drug repositioning, is likely to become an important alternative to classic

drug development because of the increasing costs of the development of new drugs and our increasing understanding of the biological background of diseases. For drug repositioning one considers the genetic background and biological pathways involved in a disease and identifies registered drugs or known investigative new drugs that target these biological pathways. A new and interesting candidate for drug repositioning in IBD is Denosumab. Denosumab (Amgen/GlaxoSmithKline) is registered for prevention of fractures caused by osteoporosis in postmenopausal women [93,94]. The drug acts through TNFSF11, encoded by the *TNFSF11* gene that has previously been identified as a risk locus for CD and bone mineral density [42,95].

Predicting drug toxicity

Already important discoveries have been made in identifying genetic variants that predict drug toxicity. Variants in the Human Leukocyte Antigen (HLA) region were found to be associated with thiopurine induced pancreatitis. Patients being homozygous for the HLA-DQA1*02:01–HLA-DRB1*07:01 haplotype have a risk of 17% for developing pancreatitis after thiopurine administration. Another genetic variant that confers susceptibility to drug toxicity lies in the *NUDT15* gene, which is associated with thiopurine-induced leukopenia [96,97]. As the genetic risk variants for severe side effects often have a large effect size it is relatively easy to gain enough power, i.e., to collect a dataset large enough, to identify them. Also the clinical significance of these side effects is so big that testing for the genetic risk variant before starting a drug might be feasible. The International Serious Adverse Events Consortium (iSAEC) together with the IIBDGC are leading the research into the genetic background of the major side effects of IBD medical therapy: pancreatitis through thiopurines, kidney damage through mesalazine and neurological side-effects in anti-TNF-alpha therapy [98].

Five-year view & expert commentary

In recent years the treatment paradigm in IBD has changed from treating symptomatic patients to starting intensified treatment regimens early in the disease to prevent complicated disease behaviour and flares of the disease [23]. However, as

mentioned before one major clinical issue in IBD is that there are no clinical parameters or biomarkers to predict how the disease will develop, so it is extremely difficult to identify patients who will benefit from aggressive treatment. Another important factor, which makes the selection of patients at risk for a severe disease course difficult, is the extreme heterogeneity of the disease. IBD has a variety of disease subphenotypes of which severity is difficult to classify (mild, moderate or severe disease). It has already been established that CD patients with a severe disease course (operations, age of onset below 40 years) carry more genetic risk variants than CD patient with a mild disease course [78]. But to identify genetic variants that are associated with specific disease phenotypes or behaviour is essential to collect clinical characteristics in a uniform and reproducible manner. So far, studies for genetic associations to subphenotypes have been performed in very small cohorts of well-phenotyped individuals or larger cohorts of poorly phenotyped individuals. Hence, only genetic risk variants with a strong risk effect have been reported to be associated to specific disease phenotypes. The *NOD2* genetic risk variants have been reported to predict severe disease course in CD and it has been suggested that the heterogeneity of the association signal from the *HLA*-region might be caused by specific *HLA* variants being associated to specific subphenotypes of IBD [16,42,99,100].

The Dutch government has funded a very large and exhaustively phenotyped prospective cohort of IBD patients initiated by the Parelsoer Institute and the Initiative on Crohn and Colitis (ICC). The ICC is an initiative formed by gastroenterologists from all eight University Medical Centres of the Netherlands. The Parelsoer Institute, financed by the Dutch government, is developing a biobank in which biomaterial phenotypical data is being collected in a uniform matter [101]. This biobank has been funded in 2007 and already contains a large collection of both phenotype data and biomaterials. Currently, phenotype and genotype data are being integrated in an attempt to translate the genetic findings to the clinic. We hope that this integration of genetic and clinical data will expand our current knowledge on biological pathways and reveal new clinical insights.

Besides the extensively growing knowledge on the genetics of IBD, developments in gene-sequencing technologies, as well as increased availability of computational biology, have lead to novel insights into the microbial composition of the human gut microbiota. Profiling studies of the intestinal microbiome have shown that IBD is associated with characteristic shifts in the composition of the intestinal microbiota, reinforcing the view that IBD results from altered interactions between intestinal microbes and the mucosal immune system [102,103]. Decreased complexity of the gut microbial ecosystem is a common feature in patients with CD or UC [9,104]. The human microbiome is a very dynamical and interactive system. Future studies with a multifaceted approach to the microbiome in IBD are essential. From a clinical perspective the increased understanding of the

microbiome will hopefully lead to new treatment options like anti- and probiotics [9,93].

Once we have established which factors predict severe disease outcomes in IBD we might be able to start intensive treatment early on in disease in patients with predicted severe disease and spare patients with predicted mild disease excess treatment and unnecessary side-effects [23].

Conclusion

In this review we have outlined the development of genetic studies in IBD from linkage studies, via GWAS, to the Immuno-chip study. We have discussed the shared genetic and biological background of IBD with other immune mediated diseases. We have outlined what genetic studies have taught us about the disease mechanisms underlying IBD. Finally we have discussed how these findings can be translated to clinical practice.

GWAS and the Immuno-chip have shown us that the underlying predisposition for the IBD phenotype is diminished innate immune response followed by an exaggerated inflammatory reaction to the commensal flora of the gut. However, genetic variants by themselves explain only a small portion of disease risk. We have to explore their interaction with environmental factors and their association to specific disease phenotypes in order to gain a comprehensive understanding of disease mechanisms. Currently large prospective and retrospective studies are being performed focused on identifying genetic risk variants that can predict a specific disease course, response to therapy, or development of drug toxicity. Only through such large well-phenotyped multi-omics studies, will we be able to translate our knowledge on the genetic background of IBD to clinical practice. At the outset the main benefit of translating genetic risk variants to clinical practice will be the prediction of drug toxicity or severe side effects from IBD therapy. The first genetic variants predicting drug toxicity are currently being published. In the near future we hope that our knowledge on the genetic background of IBD can be the basis for the development of targeted drug therapies. The process of drug repositioning based on genetic knowledge on IBD could lead to quick wins in the development of targeted therapy, and this venue should be pursued. A final important promise that our knowledge on the genetic background of IBD holds is that of personalized medicine: adapting treatment to each individual patient based on his or her genetic profile. Large studies with multi-omics data on each patient are being performed to realise the data integration needed in order to achieve personalized medicine.

We hope that in a few years we can predict disease course and best possible treatment for each patient at diagnosis with a predictive test constructed of genetic markers, microbial markers, protein markers and environmental factors. By this time the range of possible IBD treatments should have increased, and we should have a wide choice in targeted treatments for severe disease, but also for the treatment of very mild disease.

Financial & competing interests disclosure

LMS, MCV, RKW and EAF worked on the content of the paper. The first draft of the manuscript was written by LMS, MCV and EAF. All authors contributed to the final manuscript.

Weersma RK is supported by a VIDI grant (016.136.308) from the Netherlands Organization for Scientific Research (NWO). Visschedijk MC is supported by an AGIKO grant (92.003.577) from The

Netherlands Organization for Scientific Research (NWO). Festen EAM is supported with a Mandema stipend from the University Medical Centre Groningen (UMCG), University of Groningen. The authors have no other relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript apart from those disclosed.

Key issues

- Inflammatory bowel diseases (IBD), consisting of ulcerative colitis (UC) and Crohn's disease (CD), are chronic inflammatory diseases with a complex pathogenesis, originating from an aberrant immune response to the commensal intestinal bacterial flora in a genetically susceptible host.
- Genome-wide association studies identified 163 genetic risk variants for IBD, which account for 4,1-13,5% of disease risk variance.
- Genetic studies have led to the identification of *ATG16L1*, *IRGM* and *NOD2* as CD risk genes, thus showing the importance of the process of autophagy in CD pathogenesis.
- Almost three-quarters of the IBD genetic risk loci were found to be overlapping with other immune mediated diseases.
- The IL23 pathway has been the most extensively studied risk pathway for inflammatory disease; variants around this gene seem to give rise to a range of different immune mediated phenotypes.
- Fine-mapping of genetic risk loci and sequencing of genetic risk loci in IBD cases and healthy controls is currently being performed and will lead to better understanding of the causal link between the IBD associated common genetic variants and the disease mechanism.
- The newly identified IBD risk loci and biological mechanisms are suitable targets for drug therapy and several new drugs are emerging targeting these mechanisms
- The new insights in IBD pathogenesis gained through molecular research will lead to the realization of personalized medicine: adapting treatment to the individual patient based on his or her genetic profile.

References

Papers of special note have been highlighted as:

- of interest
- of considerable interest

- Molodecky N, Soon I, Rabi D, et al. Increasing incidence and prevalence of the inflammatory bowel diseases with time, based on systematic review. *Gastroenterology* 2012;142(1):46-54
- Bernstein CN, Fried M, Krabshuis JH, et al. World Gastroenterology Organization Practice Guidelines for the diagnosis and management of IBD in 2010. *Inflamm Bowel Dis* 2010;16(1):112-24
- Sands BE. From symptom to diagnosis: clinical distinctions among various forms of intestinal inflammation. *Gastroenterology* 2004;126(6):1518-32
- Molodecky NA, Soon IS, Rabi DM, et al. Increasing incidence and prevalence of the inflammatory bowel diseases with time, based on systematic review. *Gastroenterology* 2012;142(1):46-54.e42. quiz e30
- O'Toole A, Korzenik J. Environmental triggers for IBD. *Curr Gastroenterol Rep* 2014;16(7):396
- Nagalingam NA, Lynch S V. Role of the microbiota in inflammatory bowel diseases. *Inflamm Bowel Dis* 2012;18(5):968-84
- Manichanh C, Borruel N, Casellas F, Guarner F. The gut microbiota in IBD. *Nat Rev Gastroenterol Hepatol* 2012;9(10):599-608
- Leone V, Chang EB, Devkota S. Diet, microbes, and host genetics: the perfect storm in inflammatory bowel diseases. *J Gastroenterol* 2013;48(3):315-21
- Kostic AD, Xavier RJ, Gevers D. The microbiome in inflammatory bowel disease: current status and the future ahead. *Gastroenterology* 2014;146(6):1489-99
- Brant SR. Update on the heritability of inflammatory bowel disease: the importance of twin studies. *Inflamm Bowel Dis* 2011;17(1):1-5
- Hugot JP, Laurent-Puig P, Gower-Rousseau C, et al. Mapping of a susceptibility locus for Crohn's disease on chromosome 16. *Nature* 1996;379(6568):821-3
- Hugot JP, Chamaillard M, Zouali H, et al. Association of NOD2 leucine-rich repeat variants with susceptibility to Crohn's disease. *Nature* 2001;411(6837):599-603
- Ogura Y, Bonen DK, Inohara N, et al. A frameshift mutation in NOD2 associated with susceptibility to Crohn's disease. *Nature* 2001;411(6837):603-6
- Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature* 2007;447(7145):661-78
- Anderson CA, Boucher G, Lees CW, et al. Meta-analysis identifies 29 additional ulcerative colitis risk loci, increasing the number of confirmed associations to 47. *Nat Genet* 2011;43(3):246-52
- Jostins L, Ripke S, Weersma RK, et al. Host-microbe interactions have shaped the genetic architecture of inflammatory bowel disease. *Nature* 2012;491(7422):119-24
- Lakatos PL, Lakatos L, Kiss LS, et al. Treatment of extraintestinal manifestations in inflammatory bowel disease. *Digestion* 2012;86(Suppl 1):28-35
- Ott C, Schölmerich J. Extraintestinal manifestations and complications in IBD. *Nat Rev Gastroenterol Hepatol* 2013;10(10):585-95
- Turnbull GK, Vallis TM. Quality of life in inflammatory bowel disease: the interaction of disease activity with psychosocial

- function. *Am J Gastroenterol* 1995;90(9): 1450-4
20. Blumberg RS, Dittel B, Hafler D, et al. Unraveling the autoimmune translational research process layer by layer. *Nat Med* 2012;18(1):35-41
- **Presents general model for translational research from genetics to clinical practice for immune mediated disease.**
21. Mowat C, Cole A, Windsor A, et al. Guidelines for the management of inflammatory bowel disease in adults. *Gut* 2011;60(5):571-607
22. Frolkis AD, Dykeman J, Negrón ME, et al. Risk of surgery for inflammatory bowel diseases has decreased over time: a systematic review and meta-analysis of population-based studies. *Gastroenterology* 2013;145(5):996-1006
23. D'Haens G, Baert F, van Assche G, et al. Early combined immunosuppression or conventional management in patients with newly diagnosed Crohn's disease: an open randomised trial. *Lancet* 2008;371(9613): 660-7
24. Imielinski M, Baldassano RN, Griffiths A, et al. Common variants at five new loci associated with early-onset inflammatory bowel disease. *Nat Genet* 2009;41(12): 1335-40
25. Kugathasan S, Baldassano RN, Bradfield JP, et al. Loci on 20q13 and 21q22 are associated with pediatric-onset inflammatory bowel disease. *Nat Genet* 2008;40(10): 1211-15
26. Golovics PA, Mandel MD, Lovasz BD, Lakatos PL. Inflammatory bowel disease course in Crohn's disease: is the natural history changing? *World J Gastroenterol* 2014;20(12):3198-207
27. Mathew CG, Lewis CM. Genetics of inflammatory bowel disease: progress and prospects. *Hum Mol Genet* 2004;13 Spec No:R161-8
28. Cuthbert AP, Fisher SA, Mirza MM, et al. The contribution of NOD2 gene mutations to the risk and site of disease in inflammatory bowel disease. *Gastroenterology* 2002;122(4):867-74
29. Sachidanandam R, Weissman D, Schmidt SC, et al. A map of human genome sequence variation containing 1.42 million single nucleotide polymorphisms. *Nature* 2001;409(6822): 928-33
30. Lander ES, Linton LM, Birren B, et al. Initial sequencing and analysis of the human genome. *Nature* 2001;409(6822):860-921
31. Rothman KJ, Greenland S, Lash TL. *Modern Epidemiology*. Khoury MJ, Millikan R, Gwinn M, editors, 3rd Edition. Genetic and Molecular Epidemiology. Lippincott, Williams & Wilkins; Philadelphia, PA: 2008
32. Duerr RH, Taylor KD, Brant SR, et al. A genome-wide association study identifies IL23R as an inflammatory bowel disease gene. *Science* 2006;314(5804):1461-3
33. Hampe J, Franke A, Rosenstiel P, et al. A genome-wide association scan of nonsynonymous SNPs identifies a susceptibility variant for Crohn disease in ATG16L1. *Nat Genet* 2007;39(2):207-11
34. Parkes M, Barrett JC, Prescott NJ, et al. Sequence variants in the autophagy gene IRGM and multiple other replicating loci contribute to Crohn's disease susceptibility. *Nat Genet* 2007;39(7):830-2
35. WTCCC. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature* 2007; 447(7145):661-78
36. Fisher SA, Tremelling M, Anderson CA, et al. Genetic determinants of ulcerative colitis include the ECM1 locus and five loci implicated in Crohn's disease. *Nat Genet* 2008;40(6):710-12
37. Silverberg MS, Cho JH, Rioux JD, et al. Ulcerative colitis-risk loci on chromosomes 1p36 and 12q15 found by genome-wide association study. *Nat Genet* 2009;41(2): 216-20
38. Franke A, Balschun T, Karlsen TH, et al. Sequence variants in IL10, ARPC2 and multiple other loci contribute to ulcerative colitis susceptibility. *Nat Genet* 2008; 40(11):1319-23
39. Festen EAM, Stokkers PCF, van Diemen CC, et al. Genetic analysis in a Dutch study sample identifies more ulcerative colitis susceptibility loci and shows their additive role in disease risk. *Am J Gastroenterol* 2010;105(2):395-402
40. Barrett JC, Hansoul S, Nicolae DL, et al. Genome-wide association defines more than 30 distinct susceptibility loci for Crohn's disease. *Nat Genet* 2008;40(8):955-62
41. IIBDGC. Available from: www.ibdgenetics.org/
42. Franke A, McGovern DPB, Barrett JC, et al. Genome-wide meta-analysis increases to 71 the number of confirmed Crohn's disease susceptibility loci. *Nat Genet* 2010; 42(12):1118-25
43. Lees CW, Barrett JC, Parkes M, Satsangi J. New IBD genetics: common pathways with other diseases. *Gut* 2011;60(12):1739-53
44. Zhernakova A, van Diemen CC, Wijmenga C. Detecting shared pathogenesis from the shared genetics of immune-related diseases. *Nat Rev Genet* 2009;10(1):43-55
45. Trynka G, Hunt KA, Bockett NA, et al. Dense genotyping identifies and localizes multiple common and rare variant association signals in celiac disease. *Nat Genet* 2011;43(12):1193-201
46. Cortes A, Brown MA. Promise and pitfalls of the Immunochip. *Arthritis Res Ther* 2011;13(1):101
47. Visscher PM, Brown MA, McCarthy MI, Yang J. Five years of GWAS discovery. *Am J Hum Genet* 2012;90(1):7-24
48. Chen G-B, Lee SH, Brion M-JA, et al. Estimation and partitioning of (co) heritability of inflammatory bowel disease from GWAS and immunochip data. *Hum Mol Genet* 2014;23(17):4710-20
49. Diaz-Gallo L-M, Espino-Paisán L, Fransen K, et al. Differential association of two PTPN22 coding variants with Crohn's disease and ulcerative colitis. *Inflamm Bowel Dis* 2011;17(11):2287-94
50. Parkes M, Cortes A, van Heel DA, Brown MA. Genetic insights into common pathways and complex relationships among immune-mediated diseases. *Nat Rev Genet* 2013;14(9):661-73
- **Indepth discussion of shared disease pathways between inflammatory diseases.**
51. Festen EAM, Goyette P, Green T, et al. A meta-analysis of genome-wide association scans identifies IL18RAP, PTPN22, TAGAP, and PUS10 as shared risk loci for Crohn's disease and celiac disease. *PLoS Genet* 2011;7(1):e1001283
52. Van Limbergen J, Radford-Smith G, Satsangi J. Advances in IBD genetics. *Nat Rev Gastroenterol Hepatol* 2014;11(6): 372-85
53. Girardin SE, Boneca IG, Viala J, et al. Nod2 is a general sensor of peptidoglycan through muramyl dipeptide (MDP) detection. *J Biol Chem* 2003;278(11): 8869-72
54. Couturier-Maillard A, Secher T, Rehman A, et al. NOD2-mediated dysbiosis predisposes mice to transmissible colitis and colorectal cancer. *J Clin Invest* 2013;123(2):700-11
55. Kobayashi KS, Chamaillard M, Ogura Y, et al. Nod2-dependent regulation of innate and adaptive immunity in the intestinal tract. *Science* 2005;307(5710):731-4
56. Philpott DJ, Sorbara MT, Robertson SJ, et al. NOD proteins: regulators of

- inflammation in health and disease. *Nat Rev Immunol* 2014;14(1):9-23
57. Lipinski S, Grabe N, Jacobs G, et al. RNAi screening identifies mediators of NOD2 signaling: implications for spatial specificity of MDP recognition. *Proc Natl Acad Sci USA* 2012;109(52):21426-31
 58. Adler J, Rangwalla SC, Dwamena BA, Higgins PDR. The prognostic power of the NOD2 genotype for complicated Crohn's disease: a meta-analysis. *Am J Gastroenterol* 2011;106(4):699-712
 59. Van Limbergen J, Kabakchiev B, Stempak JM, et al. Hypothesis-free analysis of ATG16L1 demonstrates gene-wide extent of association with Crohn's disease susceptibility. *Gut* 2013;62(2):331-3
 60. Rioux JD, Xavier RJ, Taylor KD, et al. Genome-wide association study identifies new susceptibility loci for Crohn disease and implicates autophagy in disease pathogenesis. *Nat Genet* 2007;39(5):596-604
 61. Hussey S, Travassos LH, Jones NL. Autophagy as an emerging dimension to adaptive and innate immunity. *Semin Immunol* 2009;21(4):233-41
 62. Mizushima N, Levine B, Cuervo AM, Klionsky DJ. Autophagy fights disease through cellular self-digestion. *Nature* 2008;451(7182):1069-75
 63. Travassos LH, Carneiro LAM, Ramjeet M, et al. Nod1 and Nod2 direct autophagy by recruiting ATG16L1 to the plasma membrane at the site of bacterial entry. *Nat Immunol* 2010;11(1):55-62
 64. Cooney R, Baker J, Brain O, et al. NOD2 stimulation induces autophagy in dendritic cells influencing bacterial handling and antigen presentation. *Nat Med* 2010;16(1):90-7
 65. Orvedahl A, Sumpter R, Xiao G, et al. Image-based genome-wide siRNA screen identifies selective autophagy factors. *Nature* 2011;480(7375):113-17
 66. McGovern D, Powrie F. The IL23 axis plays a key role in the pathogenesis of IBD. *Gut* 2007;56(10):1333-6
 67. Hue S, Ahern P, Buonocore S, et al. Interleukin-23 drives innate and T cell-mediated intestinal inflammation. *J Exp Med* 2006;203(11):2473-83
 68. Brand S. Crohn's disease: th1, Th17 or both? The change of a paradigm: new immunological and genetic insights implicate Th17 cells in the pathogenesis of Crohn's disease. *Gut* 2009;58(8):1152-67
 69. Yu RY, Gallagher G. A naturally occurring, soluble antagonist of human IL-23 inhibits the development and in vitro function of human Th17 cells. *J Immunol* 2010;185(12):7302-8
 70. Coskun M, Salem M, Pedersen J, Nielsen OH. Involvement of JAK/STAT signaling in the pathogenesis of inflammatory bowel disease. *Pharmacol Res* 2013;76:1-8
 71. Brand S, Beigel F, Olszak T, et al. IL-22 is increased in active Crohn's disease and promotes proinflammatory gene expression and intestinal epithelial cell migration. *Am J Physiol Gastrointest Liver Physiol* 2006;290(4):G827-38
 72. Fujino S, Andoh A, Bamba S, et al. Increased expression of interleukin 17 in inflammatory bowel disease. *Gut* 2003;52(1):65-70
 73. Annunziato F, Cosmi L, Santarlasci V, et al. Phenotypic and functional features of human Th17 cells. *J Exp Med* 2007;204(8):1849-61
 74. Seiderer J, Elben I, Diegelmann J, et al. Role of the novel Th17 cytokine IL-17F in inflammatory bowel disease (IBD): upregulated colonic IL-17F expression in active Crohn's disease and analysis of the IL17F p.His161Arg polymorphism in IBD. *Inflamm Bowel Dis* 2008;14(4):437-45
 75. Dambacher J, Beigel F, Zitzmann K, et al. The role of the novel Th17 cytokine IL-26 in intestinal inflammation. *Gut* 2009;58(9):1207-17
 76. Christophi GP, Rong R, Holtzapfel PG, et al. Immune markers and differential signaling networks in ulcerative colitis and Crohn's disease. *Inflamm Bowel Dis* 2012;18(12):2342-56
 77. Fransen K, Mitrovic M, van Diemen CC, Weersma RK. The quest for genetic risk factors for Crohn's disease in the post-GWAS era. *Genome Med* 2011;3(2):13
 78. Weersma RK, Stokkers PCF, van Bodegraven AA, et al. Molecular prediction of disease risk and severity in a large Dutch Crohn's disease cohort. *Gut* 2009;58(3):388-95
 79. Hou JK, Lee D, Lewis J. Diet and Inflammatory Bowel Disease: review of Patient-Targeted Recommendations. *Clin Gastroenterol Hepatol* 2014;12(10):1592-600
 80. Plenge RM, Scolnick EM, Altshuler D. Validating therapeutic targets through human genetics. *Nat Rev Drug Discov* 2013;12(8):581-94
 - **Presents an overview of the steps that have to be taken to go from genetic findings to drug targets.**
 81. Sandborn WJ, Feagan BG, Fedorak RN, et al. A randomized trial of Ustekinumab, a human interleukin-12/23 monoclonal antibody, in patients with moderate-to-severe Crohn's disease. *Gastroenterology* 2008;135(4):1130-41
 82. Sandborn WJ, Gasink C, Gao L-L, et al. Ustekinumab induction and maintenance therapy in refractory Crohn's disease. *N Engl J Med* 2012;367(16):1519-28
 83. Sandborn WJ, Ghosh S, Panes J, et al. Tofacitinib, an oral Janus kinase inhibitor, in active ulcerative colitis. *N Engl J Med* 2012;367(7):616-24
 84. Sandborn WJ, Ghosh S, Panes J, et al. A Phase 2 Study of Tofacitinib, an Oral Janus Kinase Inhibitor, in Patients with Crohn's Disease. *Clin Gastroenterol Hepatol* 2014;12(9):1485-93.e2
 85. Reinisch W, Panes J, Lémann M, et al. A multicenter, randomized, double-blind trial of everolimus versus azathioprine and placebo to maintain steroid-induced remission in patients with moderate-to-severe active Crohn's disease. *Am J Gastroenterol* 2008;103(9):2284-92
 86. Massey DCO, Bredin F, Parkes M. Use of sirolimus (rapamycin) to treat refractory Crohn's disease. *Gut* 2008;57(9):1294-6
 87. Sakuraba A, Annunziata ML, Cohen RD, et al. Mucosal healing is associated with improved long-term outcome of maintenance therapy with natalizumab in Crohn's disease. *Inflamm Bowel Dis* 2013;19(12):2577-83
 88. Juillerat P, Wasan SK, Fowler SA, et al. Efficacy and safety of natalizumab in Crohn's disease patients treated at 6 Boston academic hospitals. *Inflamm Bowel Dis* 2013;19(11):2457-63
 89. Yousry TA, Major EO, Ryschkevitich C, et al. Evaluation of patients treated with natalizumab for progressive multifocal leukoencephalopathy. *N Engl J Med* 2006;354(9):924-33
 90. Sandborn WJ, Feagan BG, Rutgeerts P, et al. Vedolizumab as induction and maintenance therapy for Crohn's disease. *N Engl J Med* 2013;369(8):711-21
 91. Feagan BG, Rutgeerts P, Sands BE, et al. Vedolizumab as induction and maintenance therapy for ulcerative colitis. *N Engl J Med* 2013;369(8):699-710

92. Thomas S, Baumgart DC. Targeting leukocyte migration and adhesion in Crohn's disease and ulcerative colitis. *Inflammopharmacology* 2012;20(1):1-18
93. Lacey DL, Boyle WJ, Simonet WS, et al. Bench to bedside: elucidation of the OPG-RANK-RANKL pathway and the development of denosumab. *Nat Rev Drug Discov* 2012;11(5):401-19
94. Häusler KD, Horwood NJ, Chuman Y, et al. Secreted frizzled-related protein-1 inhibits RANKL-dependent osteoclast formation. *J Bone Miner Res* 2004;19(11):1873-81
95. Tilg H, Moschen AR, Kaser A, et al. Gut, inflammation and osteoporosis: basic and clinical concepts. *Gut* 2008;57(5):684-94
96. Yang S-K, Hong M, Baek J, et al. A common missense variant in NUDT15 confers susceptibility to thiopurine-induced leukopenia. *Nat Genet* 2014;46(9):1017-20
97. Heap GA, Weedon MN, Bewshea CM, et al. HLA-DQA1-HLA-DRB1 variants confer susceptibility to pancreatitis induced by thiopurine immunosuppressants. *Nat Genet* 2014;46(10):1131-4
- **The identification of a genetic variant that predicts a severe side-effect of thiopurines. Predicting severe side-effects will be one of the first clinical applications of genetic tests in complex disease in the clinic.**
98. International SAE Consortium. Available from: www.saeconsortium.org
99. Cleynen I, González JR, Figueroa C, et al. Genetic factors conferring an increased susceptibility to develop Crohn's disease also influence disease phenotype: results from the IBDchip European Project. *Gut* 2013; 62(11):1556-65
100. Cleynen I, Mahachie John JM, Henckaerts L, et al. Molecular reclassification of Crohn's disease by cluster analysis of genetic variants. *PLoS One* 2010; 5(9):e12952
101. PSI. Available from: www.string-of-pearls.org
102. Morgan XC, Tickle TL, Sokol H, et al. Dysfunction of the intestinal microbiome in inflammatory bowel disease and treatment. *Genome Biol* 2012;13(9):R79
103. Gevers D, Kugathasan S, Denson LA, et al. The treatment-naive microbiome in new-onset Crohn's disease. *Cell Host Microbe* 2014;15(3):382-92
104. Huttenhower C, Kostic AD, Xavier RJ. Inflammatory Bowel Disease as a Model for Translating the Microbiome. *Immunity* 2014;40(6):843-54