Genetics and epigenetics of inflammatory bowel disease

Wawrzyniak Marcin, Scharl Michael
Department of Gastroenterology and Hepatology, University Hospital Zurich, University of Zurich, Switzerland

Summary
The relevance of genetic and epigenetic alterations in the pathogenesis of inflammatory bowel disease (IBD) is still poorly understood. So far, 240 risk gene loci have been associated with IBD. They are mainly involved in regulating innate and adaptive immunity, as well as maintaining intestinal epithelial barrier function. However, the functional consequences of the identified genetic polymorphisms for IBD pathogenesis in vivo are often unknown. Even less is known about the role for epigenetic modifications in IBD pathogenesis. Though a number of epigenetic events seem to be causatively involved IBD pathogenesis, our knowledge about the functional relevance of those epigenetic modifications is scanty. This opens up a broad research field that generates novel insights into the pathophysiology of intestinal and chronic inflammatory disease. Patterns of DNA methylation and histone modifications might serve not only as biomarkers of disease activity or disease course, but also as new targets in therapeutic interventions in IBD patients.

Keywords: inflammatory bowel disease, DNA methylation, histone acetylation, pathophysiology, genetics, epigenetics, susceptibility genes

Introduction
Crohn’s disease and ulcerative colitis are the main subtypes of inflammatory bowel disease (IBD). From a clinical perspective, they represent a chronic intestinal inflammation that often begins in young adulthood and is frequently relapsing. Crohn’s disease represents a discontinuous, transmural inflammation that can occur anywhere in the gastrointestinal tract, whereas ulcerative colitis is a continuous inflammation of the mucosal layer of the colon that always starts in the rectum. In addition to the gastrointestinal tract inflammation, so-called extraintestinal symptoms are common, affecting the joints, eyes, skin and liver [1]. In Crohn’s disease, fistulas and stenosis are a severe clinical problem that often require surgery [2, 3]. Since IBD is a life-long burden in many patients, it obviously impacts the quality of life of the affected patients and has severe socioeconomic consequences [1]. Importantly, the incidence of IBD is rising in Switzerland and worldwide [4, 5].

IBD develops as a result of a combination of complex interactions between the individual’s genetic background, alterations in the composition of the intestinal microbiota on a qualitative as well as quantitative level, a dysregulated innate and adaptive immune system and environmental factors, such as diet, drugs and smoking [6]. To date, 240 susceptibility loci have been identified by genome-wide association studies (GWAS); however, only a minor part of disease risk and heritability can be explained by genetic factors alone [7–9]. It was proposed more than a decade ago that epigenetic regulation of gene expression might play a role in the development and regulation of IBD [10]. Today, several publications suggest that epigenetic mechanisms might help us to classify and diagnose patients, improve our understanding of IBD and, more importantly, provide new treatment opportunities.

Genetic factors contributing to IBD pathogenesis
Genetic factors have been widely considered as important risk factors for the onset of inflammatory bowel disease. GWAS performed in the last few years have been extremely successful in identifying genes that contribute to IBD susceptibility. NOD2 (located within the IBD1 locus) was the first gene to be associated with Crohn’s disease [11, 12]. Since then, several additional genes implicated in IBD have been identified. The strongest genetic effects were IL23R in IBD (odds ratio [OR] 2.01), NOD2 in Crohn’s disease (OR 3.01) and HLA in ulcerative colitis (OR 1.44). Most gene loci showed the same direction of effect in Crohn’s disease and ulcerative colitis, but there were some exceptions. For example, NOD2 and PTPN22 exhibited a significant protective effect in ulcerative colitis, but were risk factors for Crohn’s disease. Nevertheless, GWAS have helped us gain a better understanding of the genetic basis and their contribution with the pathogenesis of IBD. It is important to mention that many IBD risk loci are shared with other autoimmune or chronic inflammatory diseases, such as rheumatoid arthritis, systemic lupus erythematosus, multiple sclerosis, type 1 diabetes, vitiligo or psoriasis. Further, there are IBD risk loci that are associated with both IBD subforms and some that are only associated with either Crohn’s disease or ulcerative colitis. This is particularly true for genes associated with epithelial barrier func-
tion (associated with ulcerative colitis) and genes involved in cellular innate immunity (Crohn’s disease) [8, 13–15]. For example, some genetic variations have been linked to particular pathways and certain disease phenotypes. Variations in genes associated with autophagy, phagocytosis or paneth cell failure, such as NOD2, IRGM, ATG16L1 or NCF4/NCF2 have been associated with segmental, early-onset disease or structuring disease [16]. Mutations in genes of the adaptive immune system, such as within the interleukin (IL)-10 / IL-10 receptor signalling pathway have been associated with a severe form of very early onset IBD [17]. A further challenge nowadays is to bring increasing knowledge about the role of intestinal microbiota in IBD pathogenesis together with the observed genetic alterations in IBD patients on a functional level.

The odds of individual single nucleotide polymorphisms (SNPs) on the risk of developing IBD have been characterised in some detail in recent years, largely with odds ratios only slightly over 1, indicating a rather mild clinical effect. However, there is still very little knowledge on the numbers and potential interplay of these mutations, and ultimately their impact on course of disease in patients with IBD. Although there is robust knowledge on the prevalence of individual SNPs in patients with Crohn’s disease and ulcerative colitis (or IBD overall) as compared with healthy subjects, we know less about the quantity and composition of SNPs in patients with IBD, including the frequency of patients carrying more than one SNP risk-allele. However, understanding of the functional and clinical consequences of the associated alleles is still an ongoing process. Some association signals correspond to non synonymous coding variations, but the majority of signals do not. They involve noncoding genetic variations mainly related to changes in gene expression. Moreover, it has been shown that many effects seem to be highly cell-type specific. The integration of genetic, transcriptomic and epigenetic studies should lead to more insight into IBD pathogenesis and new future treatment options [18].

A recent study by Cleynen et al. revealed that the genetic risk score representing all known risk alleles for IBD showed a strong association with the disease subphenotypes defined in the Montreal classification system for IBD. Said classification distinguishes IBD into three subphenotypes: ulcerative colitis, colonic Crohn’s disease and ileal Crohn’s disease. Furthermore, they found that disease location is an intrinsic aspect of a patient’s disease, in part genetically determined, and the major driver for changes in disease behaviour over time [19].

Interestingly, in twin pair studies only a 40 to 50% concordance in the onset of IBD was detected [20, 21]. On the one hand, this clearly underlines the importance of genetic susceptibility in the disease development. However, on the other hand, this also represents a clear limitation in the concept of a genetic cause of IBD and clearly points to the (additional) involvement of other factors, such as the intestinal microbiota and environmental factors, in IBD pathogenesis. This latter aspect is also supported by the observation that similar genetic risk factors or risk gene profiles for Crohn’s disease, ulcerative colitis and other chronic inflammatory or autoimmune diseases in one individual result in the development of Crohn’s disease or ulcerative colitis and in the other individual may result in the development of another disease. Further, the fact that the risk increase, namely the odds ratio, which is associated with a large number of these genetic variations, is only about 1.2, which means that having such a variant means the chance of developing this disease is only 20% higher than for anybody who is not carrying this particular risk variant. One the one hand, this demonstrates that many people who are carrying those risk genes are healthy, or at least not affected by such a disease. On the other hand, it means also that genetic testing is not likely to be useful for diagnosis of IBD, even though there is an increasing market for this approach.

Epigenetics in inflammatory bowel disease

Principles of epigenetic mechanisms

The term “epigenetics” was introduced in 1942 by Waddington to explain how a phenotype might be produced by interaction between genes and their environment [22]. The modern definition refers to heritable alternations of gene expression events that are caused independently of genetic information carried by the primary DNA sequence [23]. The main epigenetic mechanisms controlling gene expression include DNA methylation, histone modification that modulates chromatin structure, micro RNA (miRNA) interference that regulates posttranscriptional steps, and positioning of nucleosomes [24]. By controlling patterns of gene expression, epigenetic mechanisms are involved in correct cell development, differentiation, function and homeostasis. In addition, all mentioned mechanisms are influenced by exposure to environmental factors, persists through mitosis and meiosis, and, more importantly, can be reversed [25, 26]. It was proposed more than a decade ago that epigenetic regulation of gene expression might play a role in development and regulation of IBD [10]. Today, several publications suggest that epigenetic mechanisms might help us to classify and diagnose patients, improve our understanding of IBD and, more importantly, provide new treatment opportunities.

In the process of DNA methylation a methyl group is covalently added to 5’ carbon of cytosines that are part of cytosine-guanine dinucleotides (CpG) [27]. Full methylation occurs when cytosine residues on both DNA strands are methylated [28]. Representing less than 1% of all dinucleotides, CpG dinucleotides are rather rare in the genome, but are often concentrated in particular regions of the genome called “CpG islands”. CpG islands are typically, but not exclusively, associated with gene promoters or first exons of approximately two thirds of all genes [29]. CpG islands are mainly protected from methylation and remain unmethylated, whereas in other regions of the genome, CpGs are hypermethylated [30]. DNA methylation is catalysed by so-called DNA methyltransferases (DNMTs). Based on protein sequence homology, the DNMT protein family consisted initially of five members: DNMT1, DNMT2, DNMT3A, DNMT3B, DNMT3L [31]. After it was recognised that DNMT2 methylates RNA and DNMT3L lacks 5-cytosine-methyl-transferase activity, DNMT’s were subdivided into those maintaining DNA methylation patterns (DNMT1) and de novo methylating DNMTs (DNMT3A, DNMT3B) [32]. In general, hypermethylation of CpG islands in regulatory genetic elements such as promoters is transcriptionally repressive and leads to gene si-
lencing [33]. Recently, the so-called ten-eleven translocation (TET) family of enzymes have been identified and proven to oxidise 5-hydroxymethylcytosine, which is a crucial step in the demethylation of previously methylated DNA regions [34].

An additional level of epigenetic regulation of gene expression is achieved with the process of posttranslational histone modifications. In eukaryotic cells, genomic DNA is wrapped around eight core histone proteins (two molecules of each H2A, H2B, H3B and H4 histone) to form a nucleosome, a basic subunit of chromatin. DNA compaction in chromatin is one of the most important mechanisms regulating gene expression. Chromatin that is loosely packed with tightly attached DNA, which favours active transcription of genes, is called euchromatin. In contrast, highly compacted heterochromatin is transcriptionally silent owing to limited accessibility by transcription factors. Posttranslational histone modifications occur mainly in histone tails and include acetylation, methylation, ubiquitination, phosphorylation, sumoylation and citrullination [35].

Histone acetylation, the addition of acetyl groups to lysine residues of histone is catalysed by histone acetyltransferases (HATs). Removal of acetyl groups is performed by histone deacetylases (HDACs). In general, chromatin opening during histone acetylation is associated with transcriptional gene activity, whereas increased activity of HDACs and histone deacetylation causes hypoacetylation, chromatin compacting and gene silencing (fig. 1) [36].

In contrast to methylation of cytosine residues in DNA, which leads to transcriptional repression, histone methylation by histone methyltransferases (HMTs) can be associated with either repression or active transcription [37]. The outcome depends on the position of targeted amino acid, the type of residues involved (arginine vs lysine) and the degree of methylation [30]. For example, methylation of lysine (K) in position 4, 36 and 79 of H3 is found in open chromatin whereas closed heterochromatin is enriched in methylation of H3K9 and H3K27.

In summary, it is becoming more and more clear that the epigenetic mechanisms that regulate gene expression are not separate but rather connected in such a way that DNA methylation may influence histone modification and vice versa [38].

Epigenetic modifications in IBD

In the first study trying to understand epigenetic mechanisms in IBD, Gloria et al. showed that, as compared with healthy controls, rectal mucosa from ulcerative colitis patients was characterised by global hypomethylation. A similar DNA global hypomethylation profile was observed in patients with active, inflamed ulcerative colitis as compared with patients with inactive ulcerative colitis [39]. In IBD, it has been demonstrated that local inflammation increases colonic epithelial cell turnover and accelerates DNA methylation changes [40]. Increased DNA methylation may result in genetic instability that leads to cancer development. As candidate genes referred to carcinogenesis are known, epigenetic changes could be easily investigated. That is the reason why the first DNA methylation studies were mainly focused on IBD related cancer development.

Genes that were already associated with carcinogenesis, namely CDKN2a/p16INK4A, CDKN2a/p14ARF, CDH1, MLH1, HPP1 and MYOD1 have been found differentially methylated in colonic mucosa from ulcerative colitis patients with dysplasia and/or carcinoma as compared with quiescent mucosa from the same patients [41–45].

Figure 1: Epigenetic regulation of gene expression – histone acetylation. Cells in the gastrointestinal tract are exposed to several environmental factors, dietary antigens, intestinal microbiota and microbiota derived-metabolites that may modulate epigenetic machinery. In the steady state, there is a balance between acetylation and deacetylation. Increased activity of histone deacetylases (HDACs) leads to higher histone deacetylation, which causes formation of heterochromatin and silences gene expression. On the other hand, treatment with HDAC inhibitors decreases HDAC activity and restores balance between acetylation and deacetylation in the cell, allowing genes to be transcribed.

HAT = histone acetyltransferase

HDACs inhibitors decrease activity of HDACs gene “switched on”

HDACs inhibitors increase activity of HDACs gene “switched off”

HAT = histone acetyltransferase

Environmental factors

Intestinal microbiota

Food antigens

Normal cell

HDACs

HATs

Euchromatin

Heterochromatin

HATs

HATs

HATs

HDACs

HDACs

HDACs

HDACs
estingly, as a consequence of hypermethylation lower levels of CDH1, encoding cell adhesion molecule E-cadherin, and MLHI were observed in immunohistochemical stains of gut tissue [46].

Later, the research focus was directed to a potential role of epigenetic regulation events in IBD pathogenesis itself. GWAS analysis revealed an association between IBD and polymorphisms in a gene encoding an enzyme responsible for establishment of DNA methylation – DNMT3a [47]. Moreover, as compared with uninfamed paired samples, in inflamed mucosa from ulcerative colitis patients, higher levels of expression of DNMT1 and DNMT3b were reported [48]. Increased interest in the role of DNA methylation in the pathogenesis of IBD was followed with the development of advanced platform-based DNA methylation array technologies, which shifted interest from single, candidate gene approaches to broad, general methylation analysis.

Epigenome-wide methylation association studies (EWAS) were initially performed using peripheral blood. Nimmo et al. reported a methylation profile that is characteristic for Crohn’s disease, with 50 significantly altered methylation sites in Crohn’s disease patients compared with controls [49]. Differences in methylation were observed in genes important for immune responses, such as IL21R, S100A13, FASLG, MAPK13, RIPK3 or PRF1. Use of peripheral blood as a material to investigate IBD-related methylation changes was questioned by a study where peripheral blood mononuclear cells (PBMCs) from monoyzotic twins and IBD patients did not showed differentially methylated genes with exception of hypermethylation of a locus of TEP3 gene (which has no clear relevance in IBD pathogenesis) [50].

Obviously more relevant for IBD research are tissue-specifie variations in DNA methylation. In an EWAS study using whole tissue intestinal biopsy specimens from monoyzotic twins, Hasler et al. were able to identify 61 differently methylated loci, including several loci responsible for regulation of immune responses. Interestingly, differently methylated loci were later validated in other cohorts and showed differentially expressed transcripts (CFI, FLNA, HKDC1, IGHG1, MT1H, PTN, SLC7A7, SPINK4, THY1, TK1) [51]. A nice example of a gene whose promoter region is hypermethylated and whose corresponding transcript is downregulated in the rectal mucosa of ulcerative colitis patients without differences in methylation profile in circulating leucocytes, is BRIP3 [52]. As BRIP3 was never identified in GWAS studies, it also serves as perfect example that epigenetic studies can identify new genes relevant for IBD pathogenesis.

In a very interesting study, Cooke et al. compared DNA methylation profiles in isolated intestinal epithelial cells from inflamed and uninfamed rectal biopsies from ulcerative colitis and Crohn’s disease patients [53]. Differentially methylated genes identified during this study had already been reported in other EWAS [49] and as Crohn’s disease-associated (TAP1, IL8R, PCKL, PTFR) or ulcerative colitis-associated (ICAMS, CDH1, CARD9, IL8RB, IL8RA) susceptibility genes in GWAS as well [54].

The comparison between different studies investigating the role of epigenetic mechanisms in IBD is difficult, as the main problem is reproducibility and lack of consistency regarding the type of tissue analysed (PBMC’s, epithelial cells, biopsies), controls (healthy controls, unaffected tissue from the same patient) and heterogeneity of the analysed population. As the methylome signature is specific for a given cell type, changes in cell proportions in tissues due to inflammation might mimic true epigenetic changes and lead to a false understanding of the whole process. Despite the use of statistical algorithms for estimating cell proportions in tissues, methylation profiles should ideally be studied in sorted cell populations to allow proper conclusions about real epigenetic changes.

Methylation and acetylation events of histones have been studied in IBD to a lesser extent. In dextran sulphate sodium (DSS) and 2,4-trinitrobenzene sulfonic acid (TNBS) induced experimental rat models of colitis, histone acetylation was observed in colonic tissues [55]. In this study, an increase in histone 4 acetylation on lysine (K8 and K12) was reported in inflamed mucosa as compared with uninfamed mucosa. Identical pattern of acetylation were confirmed in biopsies from patients with Crohn’s disease [56]. However, most of our understanding of histone modifications and their influence on IBD pathogenesis come from the use of HDAC inhibitors and might be therefore somehow artificial.

Administration of HDAC inhibitors in DSS and TNBS-induced experimental colitis reduces disease severity and expression of pro-inflammatory cytokines [56]. Additionally, inhibition of HDAC9 prevents colitis in mice as a result of increased development and suppressive T regulatory cell (Treg) function [57]. HATs and HDACs do not act exclusively on histones, but can modulate acetylation of non-histone proteins including p53, STAT3 (signal transducer and activator of transcription-3) or NFκB (nuclear factor kappa B) [58].

Interestingly, short-chain fatty acids (SCFAs), bacterial metabolites that are formed as result of anaerobic fermentation of dietary fibre, possess HDAC inhibitory activity [59, 60]. Many bacteria from the Firmicutes and Bacteroides genera secrete SCFAs (acetate, propionate, butyrate) at high concentration [61] and reduced numbers of bacteria that produce SCFAs have been reported in patients with IBD [62]. Conversely, application of Roseburia, a bacterium that is able to produce butyrate, showed positive effects on ulcerative colitis treatment [63, 64]. Possible mechanisms of action involve generation of Tregs from naïve CD4+ T cells. In an experimental setup, butyrate led to increased histone H3 acetylation within Fosph3 loci, the key transcription factor required for Treg cell differentiation [59, 60]. In addition, butyrate might modulate the function of intestinal macrophages [65]. Lipopolysaccharide-induced secretion of proinflammatory mediators such as IL-12 and IL-6, but not of tumour necrosis factor-alpha or monocyte chemoattractant protein-1 (MCP-1), was downregulated after treatment of macrophages with butyrate. As intestinal macrophages are the most abundant cells in lamina propria, bacterial-derived butyrate induces macrophages hyporesponsiveness and maintains tolerance.

The potential role of gut microbiota in the development of IBD, as well as of dysbiosis in gut microbiota composition in IBD patients, have been reported [66]. However, there is more and more evidence suggesting that, via epigenetic regulation of gene expression, commensal microbiota may play a beneficial role in IBD treatment.
Conclusion

Current knowledge about genetic and epigenetic involvement in IBD pathogenesis is still poor. Since the identification of the first IBD risk gene, NOD2 in 2001, GWAS have unravelled 240 risk gene loci involved in IBD pathogenesis, and many of them are involved in regulating innate and adaptive immune responses or intestinal barrier function. This strongly emphasises that, in addition to genetic alterations, the intestinal microbiota and environmental factors also play a critical role in IBD pathogenesis. However, there is still only a little knowledge about the direct consequences of the identified SNPs for IBD pathogenesis and human physiology overall, since the functional consequences of those genetic variations in vivo are often still unknown.

Even less information is available about the role for epigenetic modifications in IBD patients and their impact on IBD pathogenesis. EWAS and other approaches detected a number of epigenetic events that might be causatively involved in the onset of IBD, but our knowledge about the functional relevance of those epigenetic modifications is still scarce. This however opens up a broad research field that might help to obtain crucial novel insights into the pathophysiology of intestinal and chronic inflammatory disease. Specific patterns of DNA methylation and histone modifications might serve not only as biomarkers for disease activity or disease course, but also as new targets for therapeutic interventions in IBD patients. So it will be important to further unravel and elucidate the exact functional consequences of genetic and epigenetic alterations in IBD pathogenesis to pave the road for the development of novel therapeutic strategies.

Search strategy and selection criteria

References for this review were identified through searches of PubMed with the search terms genetics, epigenetics, IBD susceptibility genes, IBD risk genes as well as IBD and methylation from 1938 until July, 2018. Articles were also identified through searches of the authors’ own files. Only papers published in English were reviewed. The final reference list was generated on the basis of originality and relevance to the broad scope of this review.

Disclosure statement

No study sponsors had any involvement in study design, data collection, interpretation or writing of the manuscript. MS has received research grants from Calypso, Pharmabioime, Roche, Takeda and Vifor and speaker’s honoraria from FALK. MW has no conflict of interest to declare.

References


61 Louis P, Flint HJ. Diversity, metabolism and microbial ecology of bu


63 Bemi Canani R, De Costanzo M, Leone L. The epigenetic effects of bu
