UiT The Arctic University of Norway

Faculty of Health Science

Department of Clinical Medicine

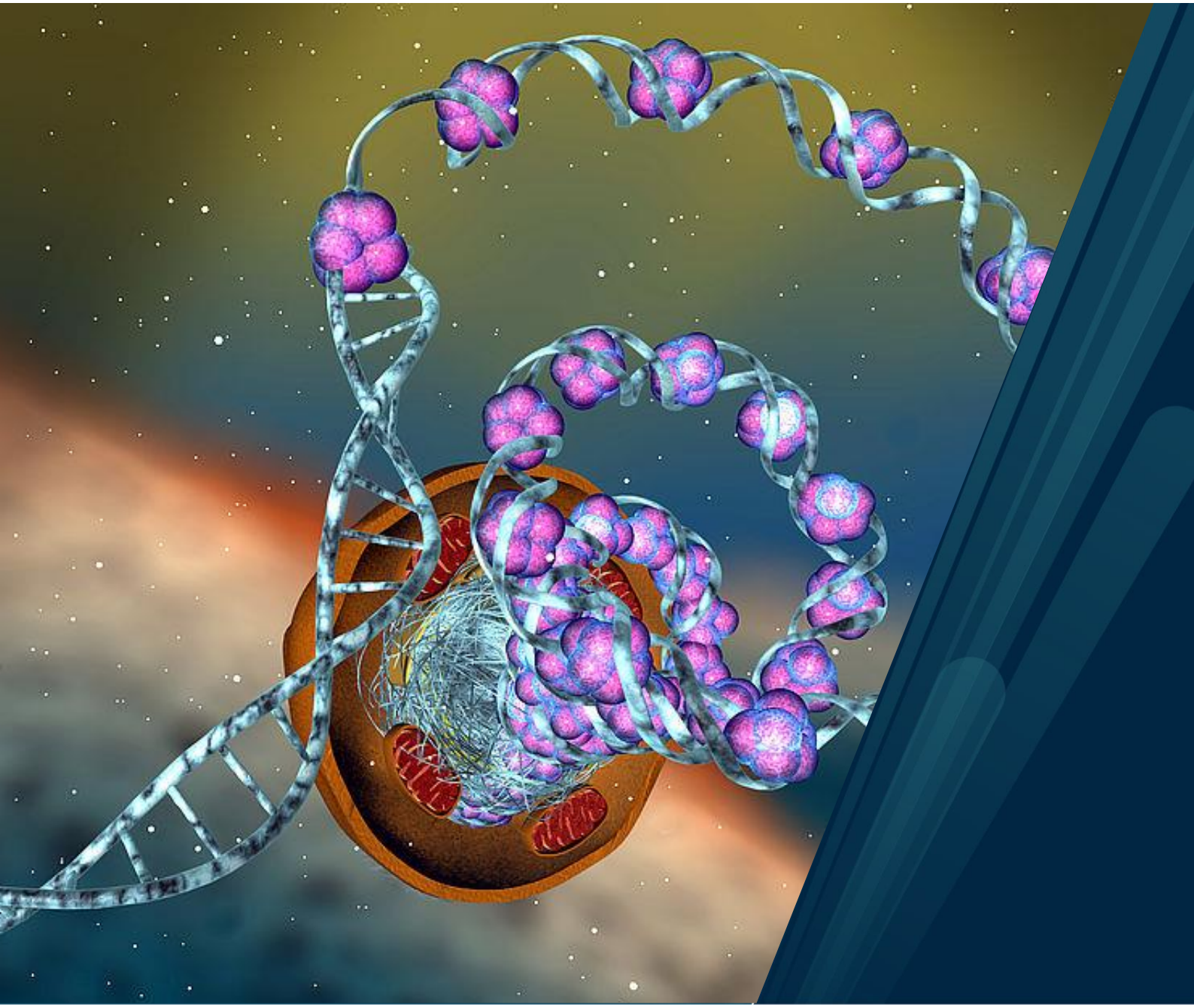
Epigenetics in Inflammatory Bowel Disease

Contribution of DNA methylation to Ulcerative Colitis pathogenesis

Hagar Taman

A dissertation for the degree of Philosophiae Doctor (Ph.D.)

December 2021



Cover picture: https://www.viennabiocenter.org/research/key-discoveries/epigenetics/

Figure 9: https://www.ebi.ac.uk/training/online/courses/functional-genomics-i-introduction-and-design/what-is-functional-genomics/

Figure 11: https://www.researchgate.net/figure/Basic-workflow-for-NGS-library-preparation-RNA-or-DNA-is-extracted-from-sample_fig1_260119313

Figure 13: https://media.springernature.com/original/springer-static/image/chp%3A10.1007%2F978-1-4939-7481-8_20/MediaObjects/316169_3_En_20_Fig2_HTML.gif

Table of Contents

Acknowledgements

The work presented in this thesis was conducted at the Genomic Support Centre Tromsø (GSCT), Department of Clinical Medicine, Faculty of Health Sciences, UiT- The Arctic University of Norway. My sincerest gratitude to Helse Nord for financially supporting this work.

The decision to start my Ph.D. journey was not easy to make. I knew it would be difficult sometimes, exciting at others and challenging for most of it. I appreciate everyone who stood by me during this time and made this journey a bit easier. I am grateful to all of you who have been involved in making my PhD successful.

My deepest and sincerest gratitude to my main supervisor Prof. Ruth H. Paulssen for believing in me and giving me the chance to perform this work. It would have not been possible without your support, encouragement, patience and not to mention your excellent knowledge, guidance, and supervision. Thank you for all our talks, both the professional and personal, the scientific and non-scientific. Thank you for giving me the right push when I need it the most. Your contributions and guidance throughout this journey in writing the papers and this thesis are deeply appreciated.

I would like to thank my co-supervisor Prof. Jon Florholmen, the coordinator of the Advanced Study of Inflammatory Bowel disease (ASIB), which this work is part of for providing the patients material used in the studies. Your comments and feedback on all my papers were of great value.

Thanks to my co-authors Christopher G. Fenton, Endre Anderssen and Inga V. Hensel. Endre, thank you for data visualization. A special thanks to Chris for all the bioinformatic help I got and providing me with the basic knowledge in bioinformatics. Your knowledge within bioinformatics is vast and it has been of great value for this project, and for that I am deeply grateful. Your comments and feedback on my papers and this thesis were of great importance. Thanks to Inga for helping with experiment performance.

I would like to thank all my colleagues at GSCT for providing a friendly and supportive atmosphere to carry out this work. Thanks to Lars Daniel Håland for taking some of the workload

off me during the last two years. Amrinder, Wei, Mithlesh and Line, thank you for the talks and discussions during our PhD journey.

To Åse B., Gro, Ingrid, Åse V., and Edel, thank you for your friendly wishes and continuous support during this journey. A special thanks to Ingrid for performing all the qPCR experiments.

Elizabeth “Busy Lizzie” A. Fredheim, your friendship means a lot to me. Thank you for all our talks both professional and non-professional, and encouragement during this time.

Elisabeth Pedersen, we got to know each other in the beginning of our Ph.D. journey, and here we are at the finish line. It is a joy knowing you, thank for all the coffee breaks and the talks.

To all my friends in Sweden, Natalia, Denise, Faduma, Janna, Sara, Kristina “Tina” and Ulrika, thank you for your love and continuous support and good wishes.

Finally, yet importantly, special thanks to my family for their unrelenting love and support. My deepest gratitude to my mum Zeinab for believing and encouraging me to keep pursuing my dreams and setting the bar high, not only during the PhD but during my whole life. Without you I would not be the person I am today.

To my husband Moustafa, you came into my life in the middle of this journey, and from day one you have been nothing but loving, supporting, and encouraging. You have pushed me to continue even at times when I did not want to. I am grateful for your support. Lastly, a billion thank you to the apple of my eye, my son Ryan (1) who joined our lives during this journey. You are my pride and joy; your beautiful smile lightens my world.

Hagar Taman

Tromsø, December 2021

List of papers

Paper I

Transcriptomic Landscape of Treatment-Naïve Ulcerative Colitis

Hagar Taman, Christopher G. Fenton, Inga V. Hensel, Endre Anderssen, Jon Florholmen, Ruth H. Paulssen. (2018)

Journal of Crohn's and Colitis, 12(3), 327-336. doi:10.1093/ecco-jcc/jjx139

Paper II

Genome-wide DNA Methylation in Treatment-naïve Ulcerative Colitis

Hagar Taman, Christopher G. Fenton, Inga V. Hensel, Endre Anderssen, Jon Florholmen, Ruth H. Paulssen. (2018)

Journal of Crohn's and Colitis, 12(11), 1338-1347 doi.10.1093/ecco-jcc/jjy117

Paper III

DNA hypo-methylation facilitates anti-inflammatory responses in severe ulcerative colitis

Hagar Taman, Christopher G. Fenton, Endre Anderssen, Jon Florholmen, Ruth H. Paulssen (2021)

PLoS ONE, 16(4 April). doi: 10.1371/journal.pone.0248905

Summary

Ulcerative colitis (UC) is a chronic inflammatory disorder in the gastrointestinal tract, and along with Crohn's disease (CD) comprise the two common forms of inflammatory bowel disease (IBD). Only approximately 20 % of all IBD cases can be explained by known genetic variants, suggesting a more complex pathogenesis which is still not fully understood. An interplay of environmental factors, composition of the intestinal microbiome, nutrition, immune response and genetic variation have been suggested as contributing to UC. It has been implied that epigenetic mechanisms might play a role in IBD disease development.

DNA methylation is an epigenetic process which regulates gene expression via structural modifications of DNA. The overall goal of this work was to explore if DNA methylation contributes to UC pathogenesis. Newly diagnosed treatment-naïve UC patients with different degrees of disease severity as well as healthy controls were included in this study. Next generation sequencing (NGS) technology was applied to obtain transcriptomic and DNA methylation profiles of UC.

The results show that DNA methylation profiles differ according to disease severity and gender. For mild to moderate UC, the transcriptomic profiles revealed genes regulating tissue-specific pathophysiological properties of tight junctions in the mucosa. A gender-dependent pathogenesis of UC could be observed. Increased expression of genes related to preservation of mucosal integrity and detoxification of microbial-derived metabolites were found in females, whereas genes related to anti-microbial, and cytotoxicity were found in males, indicating a higher risk for developing colorectal cancer (CRC). Correlations of the transcriptomic and DNA methylation profiles revealed a prominent hypermethylation of genes related to homeostasis and defence, and hypomethylation for genes related to immune response in mild to moderate UC. Surprisingly, the DNA methylation profile for severe UC revealed that hypomethylation were prominent in genes related to anti-inflammatory responses, indicating that hypomethylation might mitigate inflammation during severe UC.

The obtained molecular signatures can be potentially useful for developing epigenetic drugs and allow new treatment strategies for UC patients in the future.

1. Introduction

1.1 Inflammatory Bowel Disease

Inflammatory bowel disease (IBD) is a chronic autoimmune inflammatory disease, which affects the gastrointestinal (GI) track.[1,2] IBD has two major subtypes, Crohn's disease (CD) and ulcerative colitis (UC).[3,4] UC and CD share similar clinical manifestations; however, they are distinct pathophysiological entities.[5] It is estimated that 6.8 million people suffer from IBD worldwide, and the prevalence rate is increasing (Figure 1).[6] The global incidence of IBD has also been rising during the last few decades. In 2017 a total of 1.3 million people across Europe were estimated to suffer from IBD.[7] The classification of UC and CD depends on where the inflammation first manifest itself. CD can affect any part of the GI track, however, the most affected regions are the terminal ileum or perianal region.[8] UC affects the large intestinal and rectum, the inflammation in UC is located within the innermost lining or mucosa. UC manifest itself as continuous areas of inflammation and ulceration, with no segments of normal tissue.[5] UC is less prone to complications relative to CD as most UC patients have a mild manifestation.[9] The focus of this thesis is on UC pathogenesis.

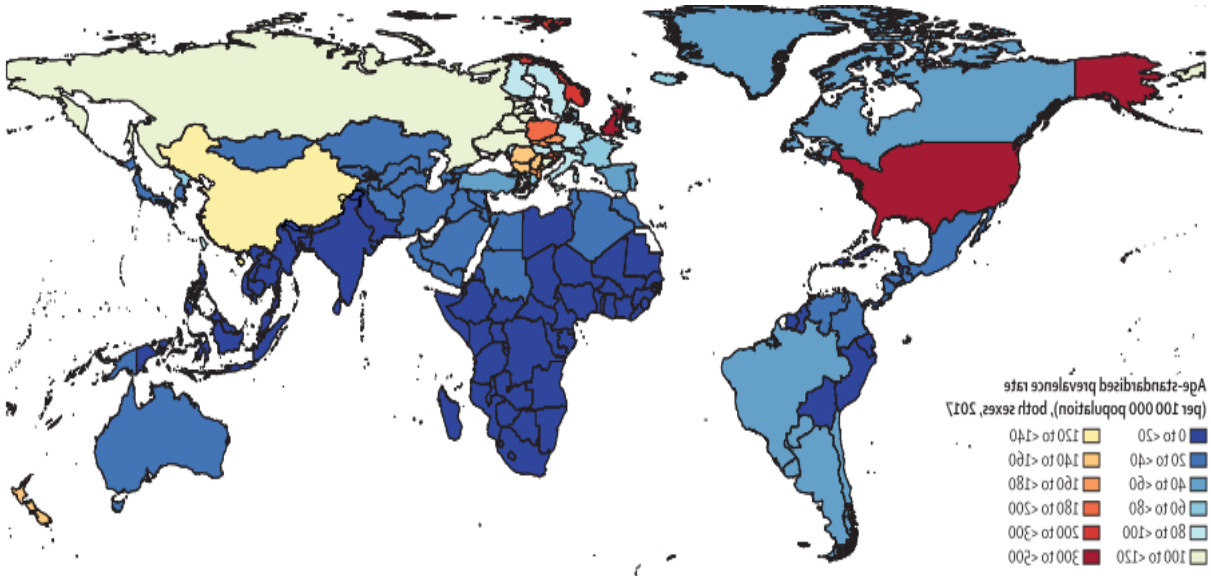


Figure 1. Worldwide prevalence rate for IBD. With permission from [2]

1.1.2 Ulcerative Colitis

The first case of ulcerative colitis was described in 1859 by Samuel Wilks.[10] UC is characterized by constant and diffuse inflammation that is restricted to the colonic mucosa, and mucosal surface. The condition of UC is defined by a relapsing and remitting course. The condition begins in the rectum and generally extends proximally in a ceaseless approach throughout the whole colon.[10–12] There are several varieties of UC depending on where in the colon the inflammation occurs. When inflammation is located in the distal part of the colon and rectum, it is termed ulcerative proctitis, if it is located to the descending colon, it is referred to as limited or distal colitis, and if the inflammation involves the entire colon, it is termed pancolitis (Figure 2).[5,13] The most common symptoms for UC are, diarrhoea, rectal bleeding, faecal urgency and/or tenesmus. Depending on the disease location some patients may experience abdominal pain. In severe inflammation cases patients may also experience fever, weight loss or perforation. Between 10-30 % of UC patients experience extraintestinal manifestations (EIM), which usually precede the onset of gastrointestinal symptoms.[9–12] There is no golden standard for diagnosing UC.[12] A full investigation consisting of endoscopic studies, histological evaluation and laboratory test should be performed to confirm the diagnosis.[9,12,14] For patients with suspected UC, stool cultures *for Clostridium difficile*, campylobacter species and *Escherichia coli* are recommended to rule out infective causes.[9,15]

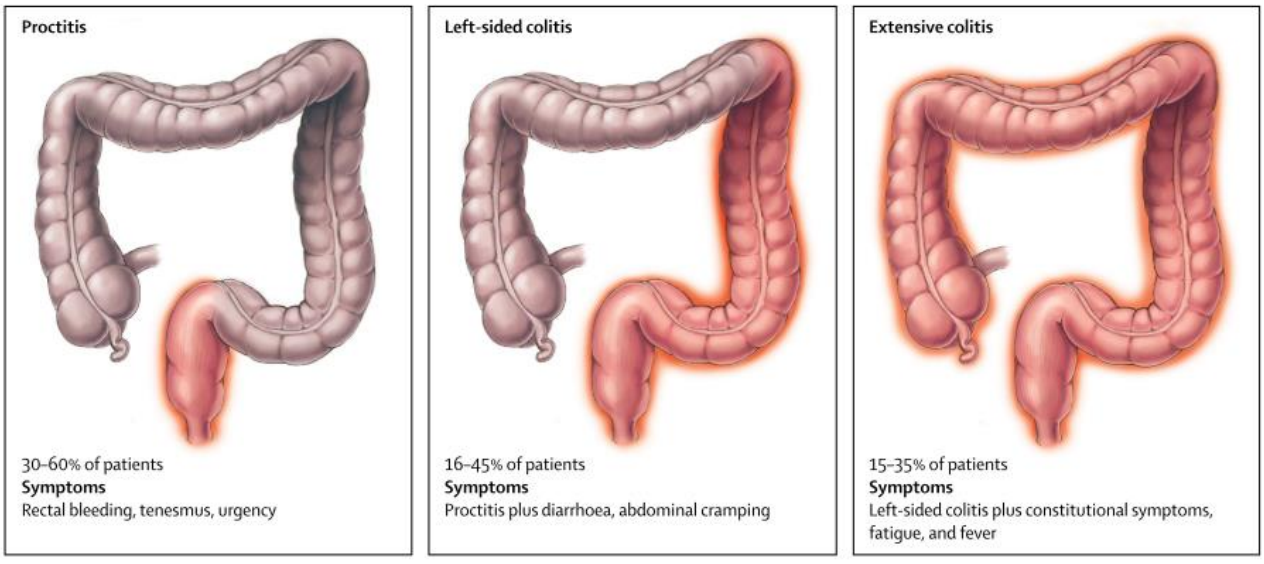


Figure 2. UC's different varieties. With permission from [28]

1.1.2 Management of UC

This thesis is based on studies performed with treatment- naïve UC patient's material. Nevertheless, treatment of UC will be mentioned below.

There is no definitive cure for UC, however, the existing medications facilitates disease management. The main goals of the treatment is to reduce inflammation, induce and maintain remission in combination with mucosal healing.[10,14,16] Medications included in UC treatment are; 5- aminosalicylic acid (5-ASA), corticosteroids, immunosuppressive drugs, and biologicals.[10,11] The most common used biological agents are TNF-α antagonists which include infliximab, adalimumab and golimumab, and anti-integrin agents such as vedolizumab.[17,18] The consensus for treating active UC has been a "step-up" approach, where patients start with 5-ASA, if they do not respond other medication are initiated (Figure 3).[10,15,19] However, new treatment strategies focus on choosing treatment based on disease activity, the degree of colonic involvement, age of onset and disease duration.[10,11,20] For acute severe ulcerative colitis intravenous corticosteroids are the first line of treatment. If patients do not respond to corticosteroids, then intravenous cyclosporine or infliximab should be considered. For patients with an inadequate response to drug treatment surgical colectomy could be an option.[21] During the last years faecal microbiota transplantation (FMT) has emerged as a new approach for correcting the dysbiosis underling IBD pathogenesis.[22,23] Emerging data have suggested that FMT can help improve the disease condition. However, data also show that FMT can have severe side effects.[24,25]

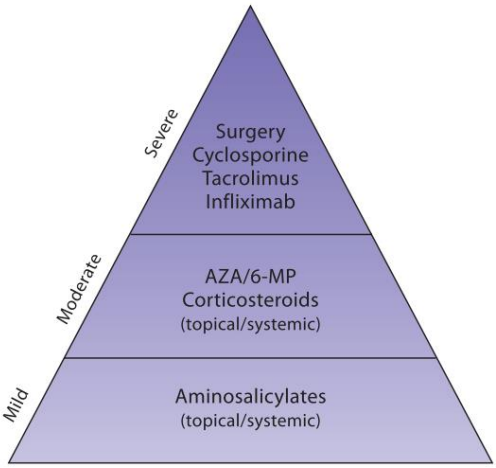


Figure 3. Step-up treatment approach for UC. With permission from [19]

1.1.3 Classification and Diagnosis of UC

Because of the inherent heterogeneity in disease progression, an accurate classification of UC is critical for clinical management. The Montreal classification is the most used subclassification system for UC.[26] It classifies UC based on both the extent and severity of the disease.[20] Based on disease extent UC is divided into three subgroups: ulcerative proctitis, left-sided UC and extensive UC (Figure 2). UC severity classification is divided into four disease activity: clinical remission, mild UC, moderate UC, and severe UC (Table 1).[27] Other scoring system beside Montreal classification exists. The Mayo scoring system for UC severity uses clinical features, endoscopic features, and physician assessment to create a score. The score system ranges from 0-12, higher score than 12 is an indicator of a very severe disease state.[20]

Disease severity	**Symptoms**
S0 (clinical remission)	No symptoms
S1 (Mild UC)	Four or less stools per day (with or without blood), absence of systemic symptoms, normal inflammatory markers.
S2 (Moderate UC)	More than four stools per day, minimal signs of systemic symptoms.
S3 (Severe UC)	Six or more bloody stools per day, pulse rate of ≥90/min, Temp ≥37.5 °C, haemoglobin <10.5g/dl, ESR ≥30mm/h

Table 1. Montreal classification of UC, based on severity of disease. Adapted from [19]

UC diagnosis is a combination of clinical presentation, endoscopic findings, histological assessment, and the elimination of alternative diagnoses.[20,28] A precise diagnosis of UC involves identifying the extent and severity of inflammation, which is crucial for correct treatment selection and for predicting patient's prognosis.[9] Therefore, an endoscopic biopsy is required to confirm diagnosis. Classic endoscopic findings for mild UC include erythema, granular mucosa, loss of the vascular pattern. In moderate UC erosions or microulcerations are apparent. In severe UC shallow ulcerations with spontaneous bleeding are observed.[9,28]

Histological evaluation is important in UC diagnosis to assess disease severity. Histological findings include deformation of crypt architecture, increased lymphocytes and plasma cells in the lamina propria, mucin reduction and Paneth cell metaplasia.[20,28] To refine severity assessment,

several scoring systems have been developed and used during the last decades. The three most used histological scoring systems are, Geboes Score (GS) which is the most used, Nancy Index (NI) and Robarts Histopathology Index (RHI).[29]

Even though laboratory tests are not diagnostic they can help assess and monitor disease activity.[9] Full blood counts, erythrocyte sedimentation rate (ESR), C-reactive protein (CRP), faecal lactoferrin and faecal calprotectin (FC) levels can be helpful in determining inflammation severity. Hypoalbuminemia is an indicator for severe disease, and is a predictor of colectomy and weak response to biological drugs.[20]

1.2 Epidemiology and risk factors

1.2.1 Pathogenesis of UC

Despite the profound discoveries made in the last century, understanding disease pathogenesis and molecular mechanisms still present a challenge for many diseases. UC represents that challenge. The inflammation in UC is caused by a complex interplay between environmental factors, gut microbiota, immune response, and nutrition in a genetically susceptible host (Figure 4).[30–35]

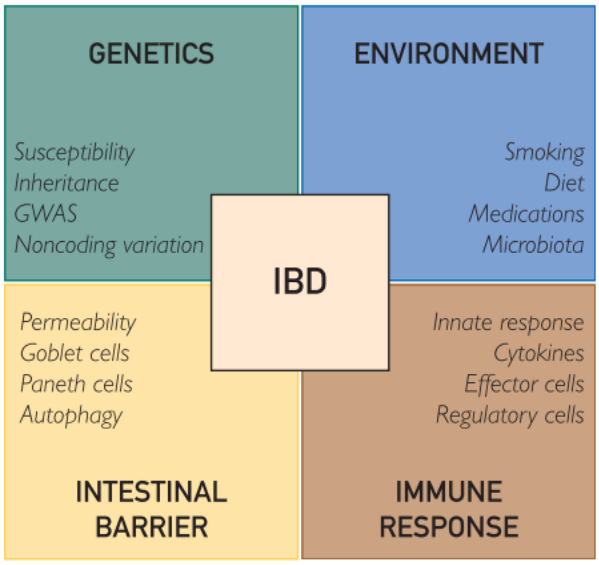


Figure 4. The different factors involved in IBD pathogenesis. With permission from [30]

1.2.2 Genetics

Genetics is a major and important risk factor for the onset of IBD. Twin studies have confirmed their role in disease susceptibility.[36,37] Even with our current knowledge the exact cause of IBD is not fully understood. IBD is multifactorial disease, where both environmental factors and genetics play a role.[38] Almost four decades ago, epidemiological studies indicated that IBD is a polygenic disorder, where several susceptibility loci contribute to the overall risk of disease.[3,39] Genome-wide association studies (GWAS) have successfully identified genes that contribute to IBD susceptibility. Nucleotide-binding oligomerization domain-containing protein 2 (NOD2) is the first susceptibility gene identified for CD.[40,41] The discovery of NOD2 indicated that innate immunity and pattern recognition receptors have a role in IBD pathogenesis.[39] Autophagy as a pathogenic pathway in IBD via autophagy related 16-like 1 (ATG16L1) and immunity-related GTPase family protein (IRGM) was detected by GWAS. GWAS studies have emphasized the crucial role inflammatory signalling pathways such as IL-23 driven T helper cell responses play in IBD.[36,39,42–44] Two meta-analysis have been performed for UC, revealing a total of 47 UC susceptibility loci. The analysis highlighted the role of defective barrier function in UC pathogenesis.[45] GWAS for CD revealed 71 susceptibility loci.[46,47] Many of the identified risk loci for IBD are shared by other autoimmune or chronic inflammatory diseases, such as rheumatoid arthritis, multiple sclerosis, type 1 diabetes, and psoriasis. This suggests an overlap in functional pathways, as well as pleiotropy (one gene influences multiple phenotypes).[44,48] Technological advances in GWAS, whole exome sequencing (WES), and whole genome sequencing (WGS) have now identified more than 240 susceptibility genes for IBD.[36,42,45,48] The majority of them are shared by both subtypes, while others are specifically associated to either UC (23 loci) or CD (30 loci).[36,39] This is especially true for genes associated with epithelial barrier function (UC), and genes engaged in cellular innate immunity (CD).[48] Despite the identification of these susceptibility genes, only a small part of disease risk and heritability can be explained by genetics alone.[3,36,44,48] Studies indicates that around 20 % of all IBD cases, can be explained by genetic variance (13.1 % for CD and 8.2 % for UC).[43,49]

1.2.3 Environmental factors

IBD is considered a disease of the modern world. As the western world industrialized, chronic immune-mediated diseases such as IBD emerged.[50] Environmental factors has been implicated to play a role in IBD pathogenesis.[3] These factors can facilitate IBD pathogenesis by exerting an effect on the gut microbiome.[30] Improvements in personal hygiene and sanitary condition in developed countries have reduced exposure to microbial stimulation, leading to a decreased incidence of infectious diseases. A hygienic environment has led to a reduction in microflora alpha diversity, which can lead to an abnormal immune response such as autoimmunity once an infection does occur.[32,51] Other environmental factors such as diet, vaccination, mode of child delivery (vaginal vs caesarean), breastfeeding, and stress have been associated with IBD.[52–56] Another environmental factor that regularly has been associated with IBD is smoking. Smoking increases the risk for developing CD amongst current smokers. Additionally smoking increases the risk for a more complicated disease course, treatment course, and it is correlated with higher relapse rate in CD. In contrast, smoking has a protective role against UC and is associated with a less severe disease course, better long term prognosis, as well as lower relapse rate.[30,35,50–52] Use of medication such as non-steroidal anti-inflammatory drugs (NSAIDs) and antibiotics can alter the composition of intestinal microbiome, consequently triggering IBD development.

1.2.3.1 Gut microbiota

Bacterial microbiota is the most well-studied element of gut microorganisms. In the gastrointestinal tract bacterial microbiota can reach a level of 10^{11}-10^{12} organisms per gram of luminal content in the colon.[57,58] Alterations in the composition of gut microbiota is known as dysbiosis.[57,59–61] Bacterial dysbiosis may contribute to pathogenesis in IBD.[60–62] There are three major bacterial phyla within the human gut, *Firmicutes, Bacteroidetes* and *Actinobacteria*.[57,59,60,63] *Proteobacteria* and *Verrucomicrobia* are also present in the human gut microbiome.[64,65] One main common feature in IBD patients is reduction in biodiversity of the gut microbiota, known as α diversity (diversity of species). A reduction in specific taxa including *Firmicutes* (*Faecalibacterium prausnitzii*)[66,67] and *Bacteroidetes* is present in IBD patients.[59,68,69] An increase in the *Proteobacteria* phylum is observed in IBD patients including *Enterobacteriaceae* (*E. coli* and

Fusobacterium) and *Gammaproteobacteria*.[57,62,68] There are certain species of gut microbiota that may have a protective role against IBD, including *Lactobacillus*, *Bifidobacterium* and *Faecalibacterium*.[66] Short chain fatty acid producing bacteria within the phyla *Firmicutes* and *Bacteroidetes* could each have a possible role in protecting against intestinal inflammation.[59,68,70]

1.2.3.2 Nutrition

Westernized life style and diet may be a key contributor to the increased incidence of IBD in Europe and Asia.[71,72] Shifting to an animal-source diet is a major change in the developed "western" world, resulting in an altered gut microbiota and microbiome.[71] The impact of different fat types in IBD pathogenesis have been investigated, specifically ω-3 and ω-6 polyunsaturated fatty acids (PUFA). Research shows that a balance between the anti-inflammatory ω-3 PUFA and the pro-inflammatory ω-6 PUFA is essential to maintain homeostasis.[73] Western diet contains high ω-6 to ω-3 ratio, resulting in a greater likelihood for developing IBD.[74,75] Higher protein intake has been associated with increased IBD incidence, and with increased relapse rate.[76] Interestingly, other studies could not prove an association between high protein consumption and increased incidence of UC.[77,78] Carbohydrates as a dietary risk factor for IBD was presented in 1977.[76] Since, several studies have associated high sugar and low dietary fibre intake with IBD development.[75] A correlation between fructose and lactose malabsorption and IBD severity has been established.[79] Increased intake of fermentable carbohydrates such as fructose, sucrose, lactose and glucose had been shown to overwhelm the intestine's absorptive capacity. Resulting in dysbiosis and gut permeability, therefore promoting inflammation.[80] Western style diet includes refined carbohydrates such as starch and sucrose, which can have an impact on the gut microbiota.[81] Low consumption of dietary fibre is associated with increased IBD incidence.[75,82] Dietary fibre maintains mucosal barriers and promote bacterial diversity, both of which have a positive impact on the intestinal homeostasis. Dietary fibre induces production of short-chain fatty acids, which can reduce inflammation and promote mucosal homeostasis.[60,71,83]

1.3 Inflammation in UC

To maintain tissue homeostasis, the immune system responds to harmful stimuli by inducing inflammation. The triggers responsible for initiating inflammation in UC are not the same as in other diseases.[84] Most inflammatory reactions are acute and self-limiting, however the

inflammation in UC is chronic and restricted to the intestinal mucosa (Figure5).[84] Granulocytes are recruited to the inflammatory site during acute inflammation. They neutralize and eliminate damage-associated molecular patterns (DAMPs) and pathogen-associated molecular patterns (PAMPs). In addition, granulocytes defend against bacterial, viral, and fungal invasion. Chemokine signalling is involved in recruiting granulocytes. Neutrophils secretes an antibacterial protein myeloperoxidase (MPO), which catalyses the production of reactive oxygen species to dismantle microbes and release tissue degrading enzymes. Initiating these processes eliminates pathogens, but can further inflammation, eventually leading to epithelial damage.[84]

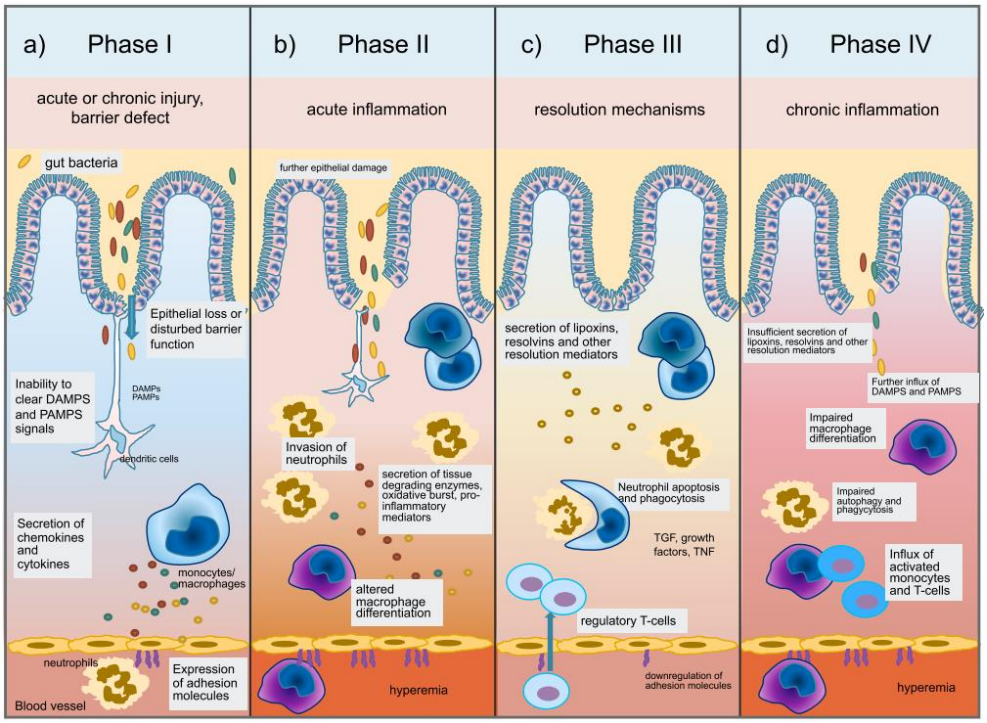


Figure 5. Different phases of inflammation. With permission from [84]

In UC, the response of the immune system plays a crucial role in initiating, magnifying, and expanding inflammation. Inflammation is induced by an increase in several pro-inflammatory mediators released from different cell types. Cells which promote proliferation of antigen-specific effectors to initiate the adaptive immune response consequently leading to local and systemic inflammation.[85] Cytokines can drive inflammation through the production of inflammatory mediators and activation of inflammatory pathways. Several cytokines and inflammatory

mediators have been shown to play a role in UC, such as tumour necrosis factor (TNF-α), interleukin (IL)-1 and IL-6.[85]

1.3.1 Inflammatory mediators

Antigen presenting cells (APCs) are essential in initiating UC immunopathogenesis as they release cytokines which actively regulate inflammation. Dendritic cells (DCs) belong to the APCs family. During an innate immune response DCs produce pattern-recognition receptors (PRRs), which include Toll-like receptors (TLRs) and NOD-like receptors (NLRs). TLRs and NLRs respond to common intracellular and extracellular pathogens.[85] DCs can also initiate an adaptive immune response by activating naïve T cells.[85,86] Naïve T-cells can differentiate into T helper cells T_h1, T_h2 or T_h17, and T regulatory cells (Tregs). T_h2 response mediates the pathogenesis of UC.[87–89] However, recent findings show an atypical T_h2 immune response in UC. In an atypical T_h2 response, UC tissue with undetectable levels of IL-14 showed an upregulation of IL-13 and INF-γ.[85] Inflamed mucosa from IBD patients showed increased levels of T_h17.[90,91] During inflammation T_h17 can have a pro-inflammatory role. It enables the secretion of pro-inflammatory cytokines such as IL-17A, TNF-α and INF-γ.[92] During inflammation intestinal macrophages (APCs) have a pro-inflammatory role, with increased production of IL-23, IL-1β, TNF, IL-1, IL-6, IL-8, IL-12, IL-18, and chemokines (chemokine (C-X-C motif)) ligand 9 (CXCL9) and CXCL10.[85,93]

The classic pro-inflammatory cytokines such as IL-1, IL-6 and TNF-α are well documented mediators in UC pathogenesis. However, other T_h2 driven cytokines such as IL-10 and IL-13 also plays a role in UC.[85,94,95] TNF-α is a multifunctional cytokine which play a prominent role in promoting inflammation in UC, by enhancing IL-6 and IL-1β production and thereby initiating an immune response.[85] In addition, TNF-α stimulates several pro-inflammatory effects in chronic intestinal inflammation, including initiation of Paneth cell death via necroptosis, activation of macrophages and T cells, and epithelial cell damage.[96]

IL-1 has both regulatory and proinflammatory roles in UC, it produces IL-1α and IL-1β which induces fever and acute phase proteins.[97] UC patients show increased levels of IL-1 receptor

antagonist (IL-1Ra), and increased levels of IL-1Ra/ IL-1 with increased disease activity. It has been suggested that IL-1Ra could play a role in downregulating the inflammatory mechanism.[85]

Another typical pro-inflammatory cytokine is IL-6, which is secreted from macrophages during acute phase in the inflammatory response. IL-6 binds to IL-6 soluble receptor (sIL-6R), then the IL-6/sIL-6R complex binds to the gp130 surface molecule (the IL-6R subunit-β), consequently activating intestinal target cells.[95] In addition, activation of gp130 leads to activation of signal transducer and activator of transcription-3 (STAT3). Activation of STAT-3 can result in induction of the anti-apoptotic factors Bcl-2 and Bcl-xl which lead to apoptosis resistance in T-cells. By not undergoing apoptosis T-cells are accumulated in the lamina propria leading to chronic inflammation seen in UC.[85]

IL-9 has been repeatedly found in inflamed mucosa of UC patients, which is produced from T_h9 cells. IL-9 is over-expressed in inflamed mucosa, and increased IL-9 levels in serum correlates with severe prognosis.[95] Studies have shown that mucosal T_h9 cells in colitis are stimulated via epithelial IL-33.[85,96] IL-9 repress the proliferation of intestinal epithelial cells (IEC) and damage the expression of tight-junction proteins, therefore altering the barrier function. It is suggested that this alteration occurs through regulation of claudin-2 (CLDN2).[85]

IL-33 is a member of the IL-1 family, which are produced by IEC in UC patients, and can initiate innate an immune response.[94] IL-33 represses the T_h1 type cytokine response in favour of inducing T_h2 cytokine response.[95] During chronic colitis IL-33 display pro-inflammatory functions, however it can provide protection in more chronic phases of inflammation by inducing proliferation of $Foxp3^+$ regulatory T cells, consequently counteracting the pro-inflammatory functions mediated by IL-23.[94,96]

Not all cytokines display a pro-inflammatory role in UC, anti-inflammatory cytokines such as IL-10 and transforming growth factor-β (TGFβ) play a protective role against colitis. Induction and production of IL-10 by intestinal macrophages impair production of the pro-inflammatory IL-1β. [96] IL-10 is produced by T lymphocytes, such as $Foxp3^-$ type 1 T_{reg} cells and $Foxp3^+$ T_{reg}, these cells are crucial for colitis resolution, in addition they repress the pro-inflammatory immune response of $CX3CR1^+$ macrophages and T cells. Additionally, IL-10 signalling via STAT3 is important

for suppressing intestinal inflammation.[96] Treg cells produces TGFβ, consequently suppressing the pro-inflammatory roles of mucosal macrophages and effector T cells. TGFβ plays a key role in intestinal homeostasis, by regulating the Foxp3^{+} T_{reg} cell pool.[94,95] Despite the existing research, the precise mechanism of the immune response in IBD is not fully understood. This is due to the intricacy of the immune system and the complexity of the innate and adaptive cytokine network.[87,88,98] Figure 6 summarises the cytokines network involved in UC.

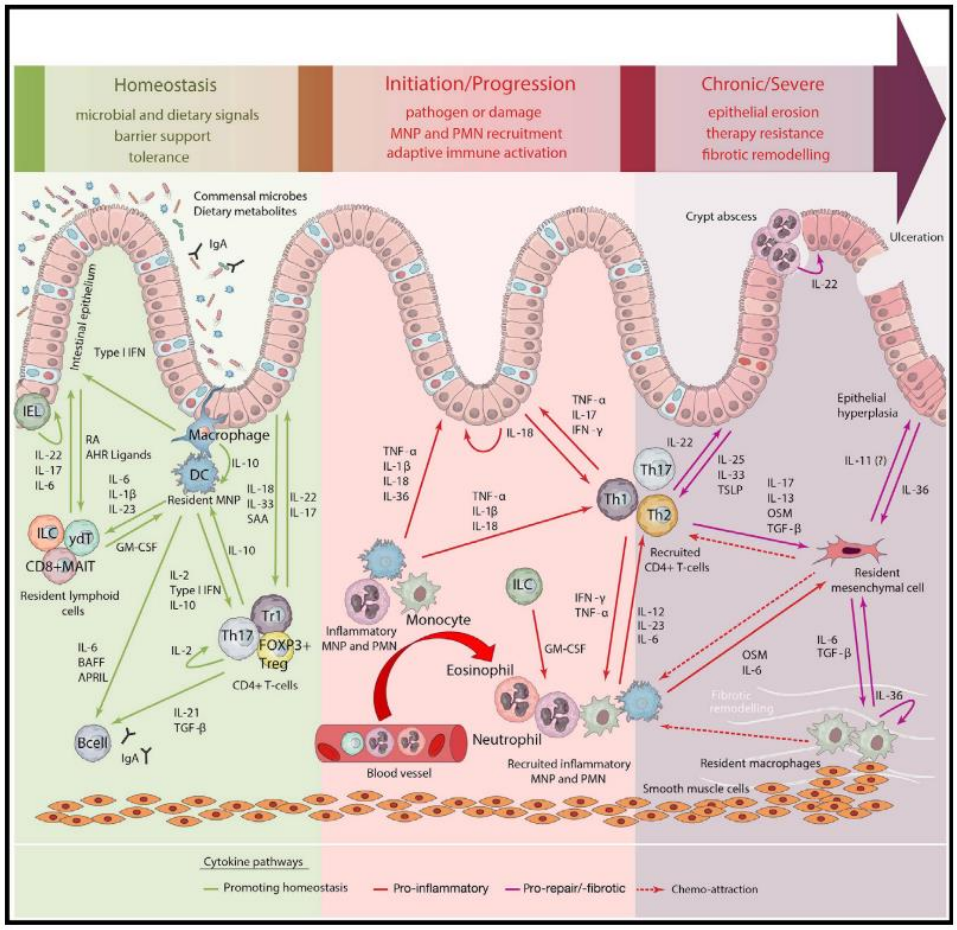


Figure 6. An overview of cytokine network in UC. With permission from [94]

1.4 Current status for biomarker in UC

To date, no biomarker exists that can routinely be used for diagnosing UC.[99] The diagnosis is based on a combination of clinical aspects, laboratory tests, endoscopy, radiology, and pathology aspects.[100] The ultimate IBD biomarker should be non-invasive, rapid, and

reproducible. Discriminating between active and non-active disease conditions is the most important criterion.[99,101] Unfortunately, as no biomarker has fulfilled the above criteria, a combination of different tests can be used as an alternative to classic inflammatory markers.[99,101] The existing biomarkers measure the ongoing inflammation and disease activity. They serve as a support for diagnosis and initial severity assessment.[99] High levels of the serological biomarker C-reactive protein (CRP) are associated with UC and CD, although poor correlation with UC is reported.[99,100] High CRP levels combined with high stool frequency on day 3 is predictive for colectomy in 85 % of hospitalized patients with acute severe UC and can predict clinical relapse and therapy failure in both UC and CD.[99] The best serological markers to distinguish between CD and UC are ASCA (anti-saccharomyces cerevisiae antibody) and ANCA (antineutrophil cytoplasmic antibodies) respectively. Levels of atypical perinuclear ANCAs (pANCAs) are increased considerably in UC. Faecal and serum calprotectin (FC) is an inflammatory biomarker which helps to assess diagnosis and treatment response. Patients with active IBD have increased levels of FC which are strongly associated with endoscopic activity. FC can predict clinical relapse in both UC and CD as well as endoscopic and histologic activity at certain cut-offs.[99] An emerging biomarker for IBD is miRNAs, it has been demonstrated that miRNAs were differentially expressed in colonic mucosa of UC patients. miRNAs levels differ between active and inactive UC in both colonic tissue and serum samples. miRNAs are considered a tool for assessing treatment response due to different miRNAs levels before and after treatment.[100]

1.5 Epigenetics

Epigenetics is the study of how environment and behaviours impact gene activity, causing phenotypic changes. The epigenetic concept was first described by Waddington in 1942.[102,103] Epigenetic changes alter the phenotypic gene expression without changing the DNA sequence itself, however, they are reversible.[103–106] Epigenetic processes occurs naturally and they are crucial for many organism functions, nevertheless, if they occur incorrectly they can have major consequences on health. Several illnesses and diseases are linked to epigenetic changes, such as cancers, respiratory diseases, and autoimmune illnesses.[107]

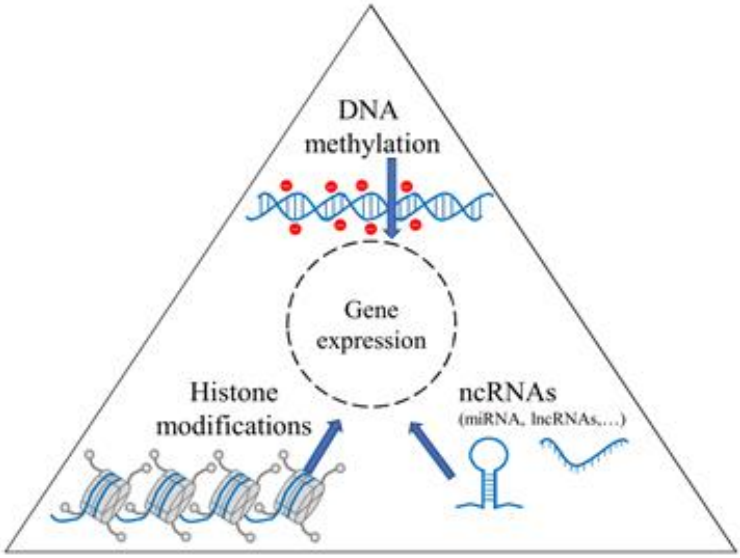


Figure 7. Epigenetic mechanisms. Adapted from (112)

1.5.1 Epigenetic modifications

Every cell in an organism exhibits different phenotypes and functions.[108,109] This is due to different quantitative and qualitative gene expression, thus gene expression is important for the differentiation and development process in an organism.[109] The epigenome can enable these differences.[108] Epigenetic mechanisms have an impact on chromatin state, therefore regulating gene expression.[103,108] Epigenetic changes are easily and rapidly affected by environmental changes, whereas our genetic code is constant throughout life.[103] To generate various combination of different phenotypes from the same genotype, epigenetic marks are passed on from the mother cell to its offspring.[110] The main epigenetic mechanisms controlling gene expression are 1) histone modifications, 2) DNA methylation and 3) non-coding RNAs (Figure 7).[36,48,108,110–112]

1.5.2 DNA methylation mechanism

The concept of DNA methylation was first introduced by Hotchkiss in 1948, when he detected a modified cytosine during an experiment. He hypothesized that it was 5-methylcytosine (5mC), and further indicated that this modified cytosine occurred naturally in DNA.[113] Many researchers suggested that DNA methylation might control gene expression, however it was in

1980s that several studies proved DNA methylation was directly involved in gene regulation and cell differentiation.[114,115] Now it is established that DNA methylation is a key epigenetic factor affecting gene activities.[48,113]

DNA methylation is a chemical modification of the DNA. It is a covalent addition of a methyl group from S-adenyl methionine (methyl donor) to the fifth carbon of a cytosine nucleotide, resulting in a 5-methylcytosine formation (5mC) (Figure 8).[36,108,111,113]

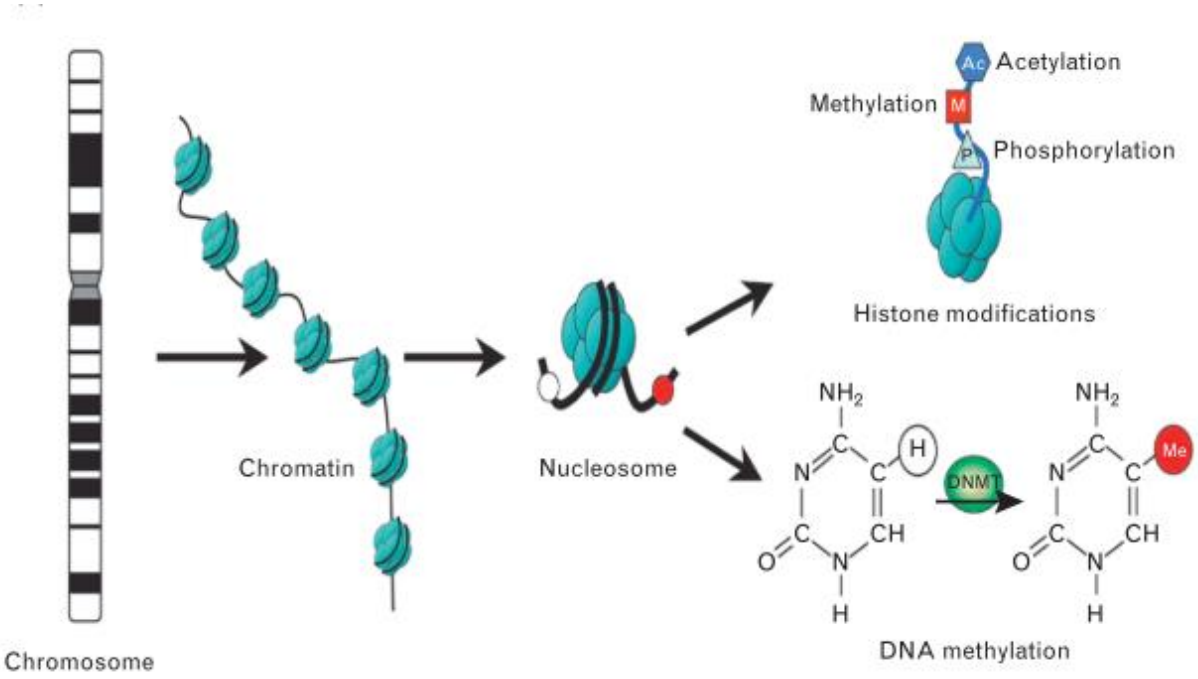


Figure 8. Epigenetic modification, DNA methylation occurs at the 5th position of a cytosine residue. With permission from [103]

For full methylation to occur, cytosine have to be methylated on both DNA strands.[105] This occurs mostly at a cytosine phosphate guanine (CpG) dinucleotides sequence.[36,37,105] CpG dinucleotides have a tendency to cluster in regions called CpG islands. CpG islands are classified as regions with more than 200 bases with a C+G content of at least 50 %.[105,110,111] CpG sequences have a low frequency in the human genome about 1-2 % and they show a reduced transcriptional activity[36,37,110] The major part (~60 %) of CpG islands in the human genome are associated with gene promoters or first exons of all genes.[48,110,116,117]

DNA methylation is catalysed by a family of DNA methyltransferase (DNMTs).[113] There are five members in the DNMT family; DNMT1, DNMT2, DNMT3A, DNMT3B and DNMT3L.[48,105,118] Only DNMT1, DNMT3A and DNMT3B possess methyltransferase activity.[108,110] These three are further classified into maintenance DNMTs which include DNMT1, and de novo DNMTs which include DNMT3A and DNMT3B. [110,111,113,119] Each of these DNMTs are vital for several biological processes, for instance DNMT1 is crucial for genomic imprinting, heterochromatin formation and gene silencing. Whereas DNMT3A and DNMT3B are crucial for embryonic development and have a vital role in de novo methylation in the genome.[106] DNA methylation is reversible, a process referred to as DNA demethylation. DNA demethylation occurs either passively or actively.[113,120]

Passive demethylation occurs in dividing cells when DNMT1 activity is inhibited or dysfunctional. Active demethylation takes place in both dividing and non-dividing cells. However, the process needs enzymatic reactions to process the 5mC to restore it back to its original state. This occurs with the help of the ten-eleven translocation (TET) family of dioxygenases. The TET proteins perform an iterative process where they first oxidize 5mC to 5-hydroxymethylcytosine (5hmC), then to 5-formylcytosine (5fC) and finally to 5-carboxylcytosine (5caC).[113,120,121] 5hmC play a key role in DNA demethylation, however it is implied that 5hmC can also regulate gene expression.[113,122]

1.5.3 Biological effect of DNA methylation

The biological consequence of DNA methylation is linked to its involvement in regulatory regions of genes (promoters or transcription start sites) and consequent effects on gene expression.[123] DNA methylation plays an important role in various biological processes, such as embryonic development, silencing of transposable elements, X-chromosome inactivation, genomic imprinting and controlling gene expression.[124–126] DNA methylation can regulate gene expression by recruiting proteins that inhibit gene expression, or by inhibiting transcription factor from binding to the DNA. In the first case it results in repress gene expression, and altering gene expression pattern in the second case.[113] Hypomethylation is when methylation occurs less than the reference. Hypermethylation can reduce gene expression and can lead to gene inactivation.[127]

In foetal samples, many of the genes involved in intestinal epithelial defence are hypermethylated, resulting in loss or decreased expression levels, causing incomplete or impairment of intestinal epithelial cell/barrier function in paediatric IBD.[125] A recent study has shown that anti-inflammatory genes such as interleukin 10 (IL-10) and CXADR-like membrane protein (CLMP) are hypomethylated in severe UC.[128]

1.5.4 DNA methylation and inflammation

The inflammatory response is complex, involving a refined regulatory network that carry out functions at signal-specific, gene-specific, and cell-specific levels. Activation of antimicrobial defence genes, tissue repair/remodelling genes, and immune response genes are part of this network.[111] Transcription factors of the NF-κB, FOXP3, IRF and STAT families as well as DNA methylation have been demonstrated to play a vital role in regulating inflammatory genes.[108,129]

During chronic inflammation, proinflammatory genes are upregulated, while anti-inflammatory genes are downregulated. This is presumably due to epigenetic regulation.[108] Protease- activated receptor (PAR2) exhibit both anti- and proinflammatory effects on colon, by stimulating the production of T-helper cell type 1 (T_h1), and cytokines such as tumour necrosis factor alpha (TNF-α), interleukin 1 (IL-1) and interferon gamma (IFN-γ). Increased methylation levels of PAR2 are linked to severe phenotypes of UC.[36,130] In inflammatory disorders such as psoriasis DNA methylation levels has been reported to be reduced across the genome. Recent studies demonstrated reduced DNA methylation at the distal enhancer element of the interferon-γ (INFG) gene in effector $CD4^{-}CD8^{-}CD3^{+}TCR^{+}$ ("double negative " DN) T cells.[131,132] In systematic lupus erythematosus (SLE) a monozygotic twin study uncovered 49 hypomethylated regions in SLE patients.[133] Another study demonstrated that increased expression of cytokines was associated with DNA hypomethylation in the promoter regions of interleukin 10 (IL10) and interleukin 13 (IL13) genes in $CD4^{+}$ T cells.[134]

1.5.5 DNA methylation and IBD

Recent epigenetic studies including DNA methylation reveal potential pathogenic explanations for development of IBD. They present promising gene candidates for exposing the process of disease manifestation beyond the known risk loci. Alterations in the methylation status in IBD-associated genes could drastically change the levels of gene expression, subsequently impacting disease onset and progression.[135,136] The first study associating DNA methylation with UC pathogenesis was reported by Gloria et al. the study showed that incorporation of the 3^{H}-methyl group into DNA was 10-fold higher in UC patients compared to controls, and considerably higher in histologically active than nonactive disease.[137] Other studies have shown that genes such as E-cadherin (cell-cell adhesion molecule), P16INK4a (a cyclin- dependent kinase inhibitor), cadherin 1 (CDH1), glial cell derived neurotrophic factor (GDNF), and ATP binding cassette subfamily B member (MDR1) have higher methylation levels in the promoter regions in UC patients.[136,138,139] The hypermethylation in CDH1 is linked to long standing inflammation, indicating a positive association between DNA methylation and inflammation.[140]

Furthermore, DNA methylation has been related to various clinical aspects, such as disease severity, disease phenotype, active inflammation and dysplasia, and disease duration.[136] As for CD, the current data concerning the involvement of DNA methylation in CD pathogenesis is limited. In a study by Cooke et al, they showed that there is differential methylation between inflamed mucosa from CD/ UC patients and controls for a substantial number of genes.[141] Another study demonstrated DNA promotor hypermethylation in CD patients' serum using the cancer specific and highly methylated gene transcription elongation regulator 1 like (TCERGIL).[142]

1.6 Functional Genomics

Functional genomics refers to the study of how genes and intergenic areas of the genome contribute to various biological processes. Functional genomics determines how separate elements of a biological system work together to generate a distinct phenotype. It focuses on the dynamic expression of gene products during a certain circumstance, e.g., during a disease development. The methods included in functional genomics are: genomics and epigenomics

(DNA), transcriptomics (RNA), proteomics (protein) and metabolomics (metabolite) (Figure 9).[143] Epigenomics and transcriptomics were applied in the current work, and will be shortly discussed below.

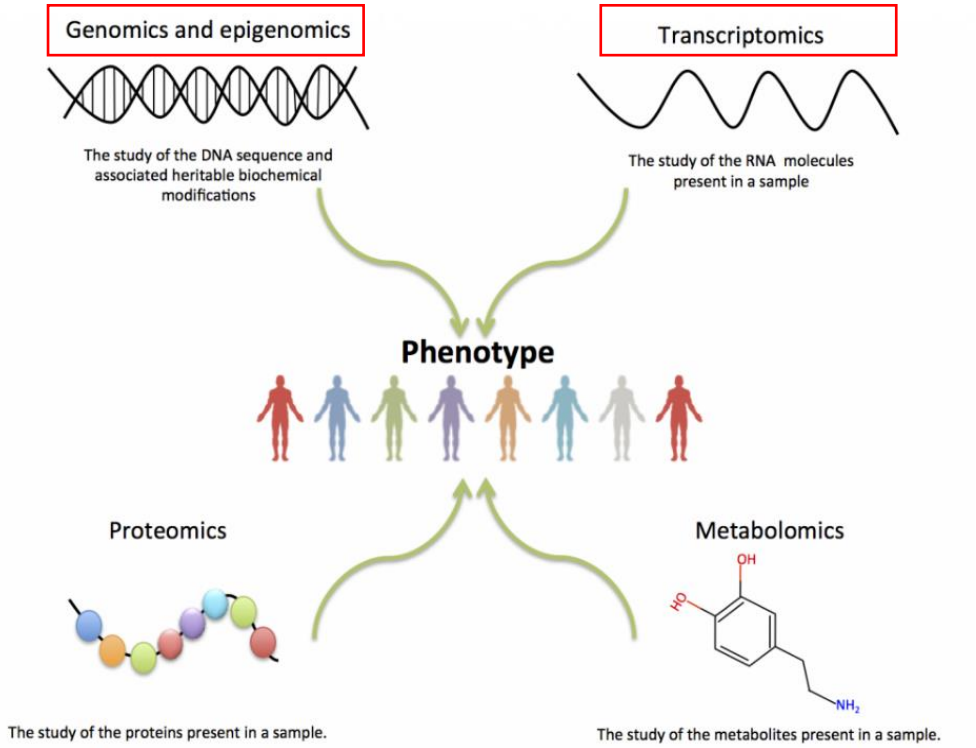


Figure 9. The different omics applied in functional genomics.

1.6.1 Transcriptomics and epigenomics

In a biological system the genetic information is passed in four possible ways: (a) DNA → DNA (DNA replication), (b) DNA → RNA (first step in protein synthesis), (c) RNA → protein (second step in protein synthesis), (d) RNA → RNA (viruses coping themselves).[144] To understand how gene expression changes and how these changes may contribute to human diseases transcriptomic (RNA) studies have been used. Transcriptomics uses the total transcripts in a cell for gene expression analysis.[145] Transcriptomic data is generated by two methods: sequencing of separate transcripts (RNA-Seq) or hybridization of transcripts to an ordered array of nucleotide probes (microarray).[145]

To study epigenetic modification's effect on the genetic material in a cell is referred to as epigenomics. In this field, identifying the location and understanding the functions of all the

chemical alterations that marks the genome is investigated. Epigenomic maintenance is a vital process for genome stability and the robustness of phenotypic attributes.[146,147] The most studied epigenetic modifications are DNA methylation and histone modifications.[148] High throughput assays have revolutionized the field of epigenetic studies during the past decade.[149]

The important steps involved in generating epigenetic and transcriptomic data involves (1) experimental design, (2) laboratory performance, and (3) data analysis. Figure 10 shows the steps involved in generating DNA methylation and transcriptomic data and the data analysis steps applied in this thesis.

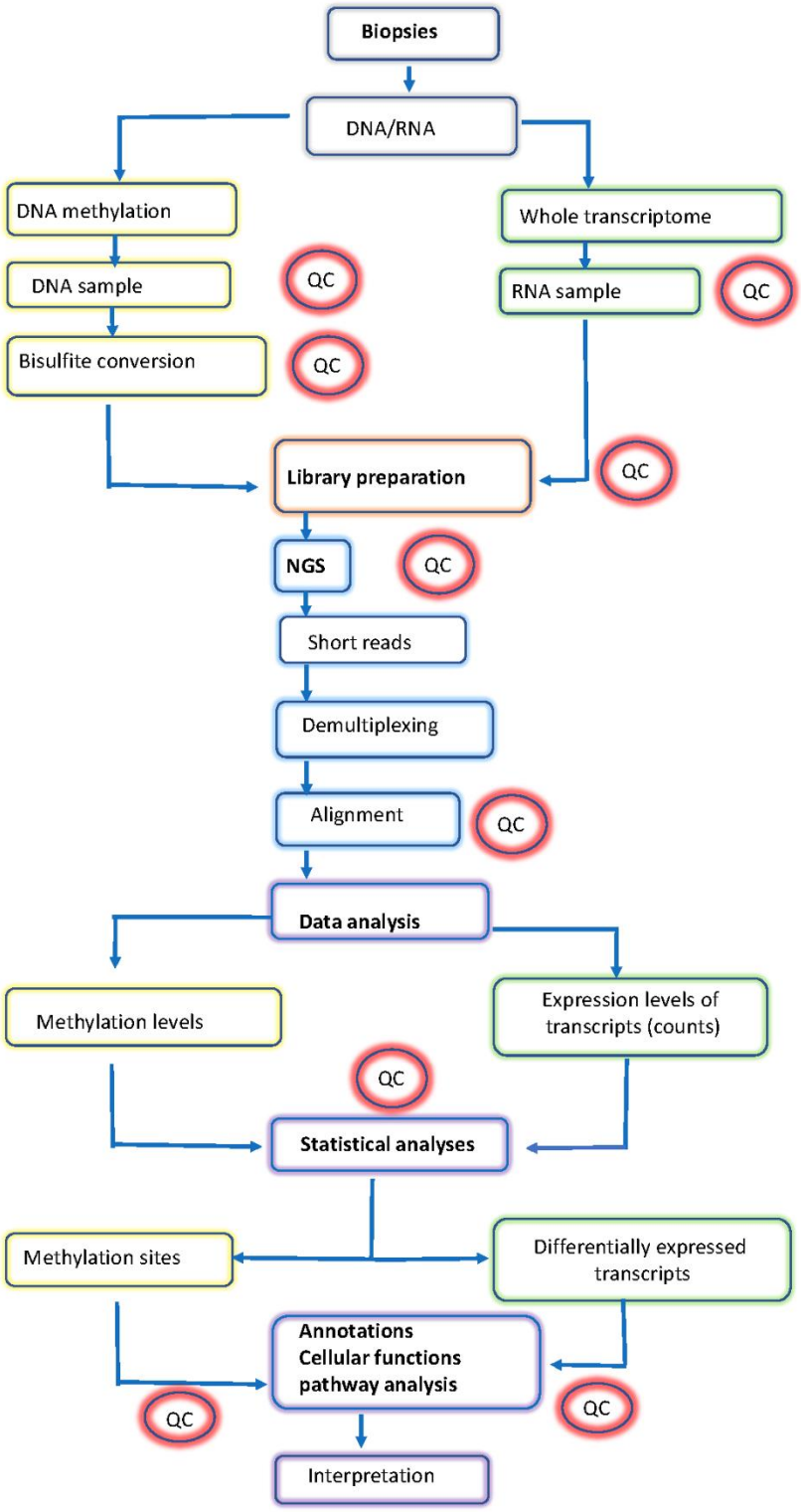


Figure 10. Flow chart of the steps included in the experimental design.

1.7 Clinical outcome

The early induction of biological therapies such as anti-TNFα in combination with immunomodulators have shown substantial improvement in response and remission rates in both UC and CD.[150] Biological therapies have induced long lasting mucosal healing and deep remission as well as reduce the risk for future complications.[150] However, as low as 30 % of IBD patients respond to treatment, and a large percentage discontinue the treatment due to variety of unpleasant events.[151] Over time, patients can lose this response which results in clinical relapse and disease progression.[150] Similarly, discontinuation of treatment leads to the same outcome.[151] The risk of relapse after anti-TNF discontinuation is between 30- 40 % after one year, and more than 50 % beyond two years.[151,152] UC patients with a confirmed "deep" remission condition before discontinuations of biological therapies have a greater chance of maintaining remission at 12 months of follow-up.[151] UC outcome varies over time, and it is dependent on disease extension, severity and activity.[153] Due to this variation a predictive biomarker for disease outcome is needed. Biomarkers will improve the treatment strategy in the context of personalized medcine.[154]

2. Objectives of the thesis

The pathogenesis of IBD is still not completely understood, and genetics only explain approximately 20 % of IBD cases, suggesting a more complex pathogenesis. Studies are indicating that DNA methylation have an impact on the transcriptome in IBD pathogenesis. The overall objectives of this work are to understand the role of epigenetic modifications specifically DNA methylation in the initiation and progression of UC. The correlation of subject gene expression and DNA methylation data provide the opportunity to characterize changes observed within individual subjects. Thus, providing insight into the susceptibility of patients with different clinical UC pathologies. In addition, identifying new biomarkers would be of great value for determining new personalized treatment strategies for UC patients. Therefore, the objectives of this work are as follows:

- To characterize and describe the complete transcriptomic landscape in treatment-naïve UC patients.
- To explore the genome-wide DNA-methylation status, and to examine the association between DNA methylation and gene expression levels in treatment-naïve UC patients.
- To identify the role of DNA methylation in severe UC and to what degree it correlates with disease severity.

3. Materials and Methods

3.1 Experimental Design

For the experiments performed in this work, colon biopsies from treatment-naïve UC patients were obtained from the ASIB study biobank. DNA and RNA were extracted, genome-wide bisulfite library and whole transcriptome library were prepared and sequenced. Various statistical and bioinformatic analysis methods were applied to the DNA methylation and transcriptome data to extract relevant data (Figure 10).

3.1.1 Biopsy selection

Well stratified treatment-naïve UC biopsies were achieved from an established biobank approved by the Norwegian Board of Health (the ASIB study biobank at the University Hospital of North Norway (UNN)). This study, collection and storage of biological material were approved by the Regional Ethics Committee of North Norway and the Norwegian Social Data Services (REK Nord 2012/1349). All enrolled participants signed an informed and written consent form. The enrolled participants for this study were newly diagnosed, treatment-naïve UC patients with mild, moderate, and severe disease activity. UC diagnosis was established based on clinical endoscopic and histological criteria as defined by the ECCO guidelines.[12] Assessment of the inflammation degree was established during colonoscopy using the UC Disease Activity Index (UCDAI).[155,156] Subjects undergoing endoscopy for cancer screening with normal colonoscopy and normal histological examination, served as controls. To assess the degree of UC activity tumour necrosis factor alpha (TNF-α) mRNA expression levels were quantified by real-time polymerase chain reaction (qPCR).[157] For mild to moderate UC, biopsies were obtained from the sigmoid part of the colon, whereas for severe UC the biopsies were obtained from the recto-sigmoid part of the colon. Biopsies were taken from an active inflammation site.

3.1.2 DNA/ RNA isolation and assessment

Total RNA and genomic DNA were simultaneously isolated from the biopsies using the AllPrep DNA/ RNA Mini Kit (Qiagen), the isolation was performed by the QiaCube Instrument (Qiagen) according to the manufacturer's protocol. In brief, the biopsies were homogenized, the

supernatant was transferred to tubes and placed in the QiaCube instrument, DNA and RNA were isolated according to the protocol. RNA and DNA were assessed for quantity and purity using the Nanodrop ND-1000 spectrophotometer (ThermoFisher Scientific). RNA purity and RIN value were assessed with the Experion Automated Electrophoresis System (Bio-Rad). The used RNA samples had a RIN value between 8.0 -10.0. All RNA and DNA samples were kept at -80° C until further use.

3.1.3 Library preparation

Preparing libraries for sequencing requires producing a collection of cDNA/DNA fragments. The principle is that DNA or RNA sequence is fragmented to smaller parts and an adapter is ligated to both fragments ends. Thereafter the constructed libraries were sequenced. Figure 11 shows an overview of library preparation principle.[158]

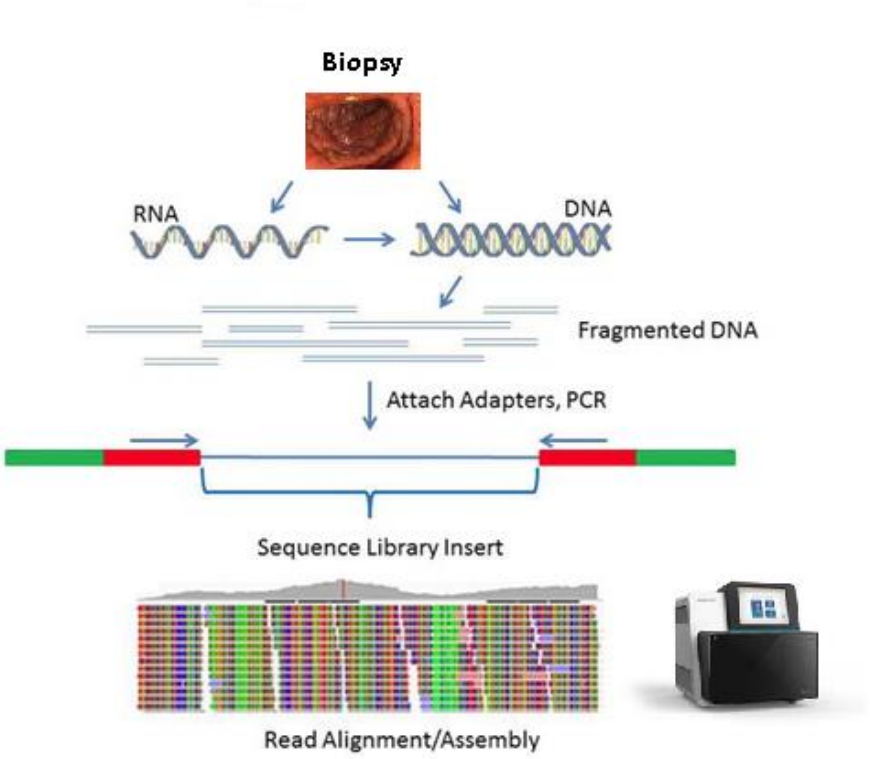


Figure 11. Basic workflow for NGS library preparation.

3.1.3.1 whole transcriptome library preparation

Whole transcriptome libraries were prepared using the TruSeq Stranded Total RNA LT Sample Prep Kit by Illumina. Material input was 1 µg of total RNA. In brief, RNA was depleted for rRNA and fragmented then followed by cDNA synthesis. An A-tail and adapters were ligated to the cDNA. Samples were cleaned-up and libraries were amplified by PCR. Finally, libraries were validated, and fragment size was assessed. The generated libraries comprised fragments with an average size of 307 base pairs. Libraries were normalized and pooled prior to sequencing.

3.1.3.2 Genome-wide library preparation

The DNA methylation libraries were prepared using the SeqCap Epi CpGiant Enrichment kit (Roche). This kit targets genomic regions from bisulfite-treated genomic DNA. The bisulfite treatment converts all unmethylated cytosines to uracil and leaving methylated cytosines unaffected.

Briefly, 1 µg of genomic DNA was fragmented, an A-tail and adapters were ligated to the DNA fragments. A dual size selection was performed, and DNA was bisulfite treated (Zymo Research). A PCR amplification step was performed, products were cleaned up and hybridized to the SeqCap Epi libraries. The samples were hybridized at 47°C for 72 hours. Samples were washed and cleaned up, followed up by a PCR amplification step. Finally, the PCR products were cleaned up and the DNA libraries were eluted (Figure 12). DNA libraries were validated, and fragment size was assessed. The generated libraries had an average fragment size of 322-329 base pairs. Libraries were diluted and pooled prior to sequencing.

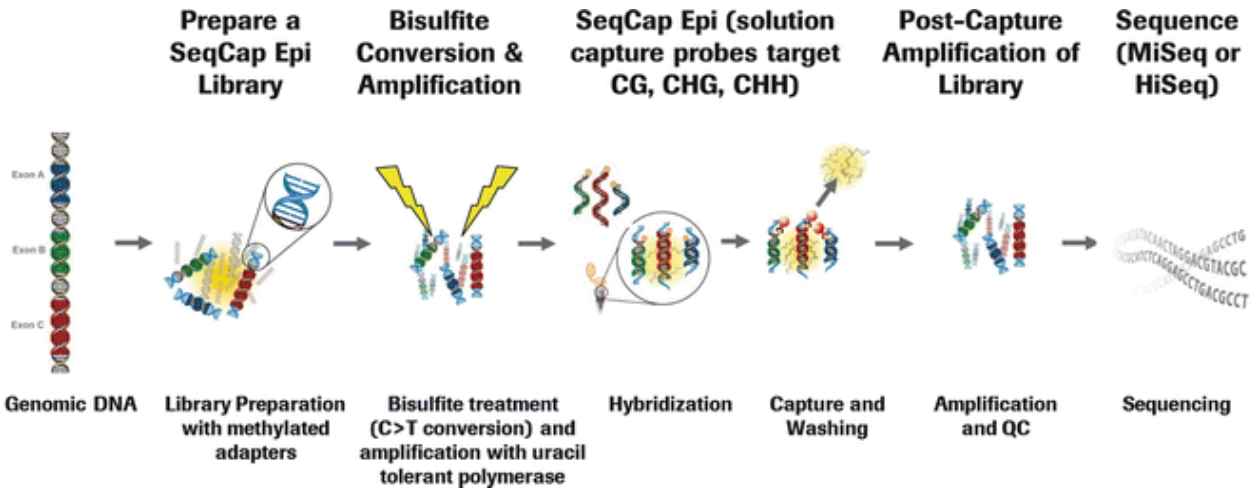


Figure 12. Workflow of the SeqCap Epi CpGiant Enrichment library preparation. With permission from Roche.

3.2 Next generation sequencing

Both the SeqCap Epi- and whole transcriptome libraries were sequenced on the NextSeq 550 system platform. Both libraries were diluted prior to pooling, thereafter denaturized, and diluted to loading concentration. A spike in control (PhiX control) was added to the libraries before sequencing. The PhiX control was diluted to the same final concentration as libraries. Denatured and diluted libraries were combined with PhiX. The final percentage of PhiX in the libraries were 1 %. The libraries were sequenced using a high output flow cell 150 cycles, and according to the manufacturer's instructions. The libraries were sequenced using pair-end mode.

3.3 Data analysis and Bioinformatics

To extract the biological relevance in both the transcriptome and epigenome data, the sequences underwent several statistical steps. At first transcriptome and epigenome data were analysed separately. Methylation and transcriptome were then correlated for each to understand the relationship between gene expression and DNA methylation.

3.3.1 RNA-Seq data analysis

Data analysis for RNA-Seq are divided into three main sections; 1) data pre-processing, which includes normalization, quality control, read alignment quantification and common statistical methods.[145,159] 2) Data analysis and visualization, e.g. performing principal component analysis (PCA), generating heatmaps, and performing hierarchical clustering.[160] 3) Data validation, this can be done by performing qPCR, using in vivo and in vitro models or by applying in silico applications (Figure 11).[161,162] A successful data analysis will generate gene list with possible candidate genes that differ between treatments or populations. To know the function of generated genes and their products Gene Ontology (GO) or Protein Analysis Through Evolutionary Relationships (PANTHER) can be applied.[159]

The first step in analysing the transcriptome data performed in this thesis were base calling and quality scoring check on the on-board computer of the NextSeq 550. The algorithm package STAR- 2.5.2b and the htseq-counts were used for down-stream analysis.[163] htseq-counts are used for pre-processing RNA-Seq alignments for differential expression calling. The human genome assembly GRCH38p.11 was applied for the alignment of the transcripts [*https://www.ncbi.nlm.nih.gov/grc/ human/data*]. Read-counts were normalized using the DESeq2-Rlog variance stabilised transformation. A data set of significantly differentially expressed transcripts were established by incorporating transcripts with read counts of > 30 and fold change > 1 compared to controls.

3.3.2 DNA methylation data analysis

DNA methylation data analysis includes the same steps as for RNA-Seq analysis (Figure 11). High quality sequencing data is critical to achieve good alignment and obtain correct methylation scores.[164] Visualizing methylation data and obtaining a DNA methylation profile can be done through several genome browsers.[164] After data visualization differentially methylated positions (DMPs) and differentially methylated regions (DMRs) are determined and a list of DMRs is produced.[165] Due to their close location to gene promoter or body, DMPs and DMRs are often associated with genes, and to determine their function gene set enrichment analysis (GSEA) and

pathway analysis can be applied. Another approach to interpret the function and significance of the discovered DNA methylation is to correlate it with gene expression.[165]

Like transcriptome analysis, epigenome data analysis starts with base calling and quality scoring. For the DNA methylation analysis, the Bismark Bisulfite Mapper v0.16.0 [*www.bioinformatics.bbsrc.ac.uk/projects/bismark/*] was applied to align reads to the same genome build applied in the transcriptome analysis to generate methylation counts. Relative methylation is expressed as a number between 0-1, where 0 means 0 % of Cs are methylated at that position and 1 means 100 % or all Cs are methylated. The global methylation analysis mapped included more than 9 million cytosine sites genome wide. To obtain significantly differentially methylated regions (DMRs) the globalTest function from the BiSeq package was applied [*https://bioconductor.org/packages/release/bioc/html/BiSeq.html*] on promoter regions. Promoter regions with a global test p-value < 0.05 were kept. The promoter region was defined as 2000 bp upstream and 200 bp downstream of the transcription start site (TSS).

3.3.3 Correlating transcriptome and epigenome

The DNA methylation analysis generated a list of differentially methylated promoter regions which was correlated to the differentially expressed genes (DEGs) generated from the transcriptome data. Inclusion criteria for promoter DNA methylation was average read coverage> 10 and adjusted p-value < 0.05. Whereas for the transcriptome data, the following criteria were applied; read coverage > 30, fold change > 1 and adjusted p-value < 0.05.

A negative correlation between methylation and transcription was used to select genes of interest. Upregulated DEGs were correlated with low methylation status (hypomethylation), and downregulated DEGs were correlated with high methylation status (hypermethylation) were selected for further analysis.

3.3.4 Common visualization and statistical methods

Both transcriptome and epigenome sequencing data were subjects to several visualization and statistical methods, some of which will be mentioned below.

Principal component analysis (PCA) is applied to visualize the data. The idea behind PCA is to reduce the dimensionality of a dataset, while retaining as much "variability" (i.e. statistical information) as possible.[166] It is an unsupervised method, which means that it can detect patterns without reference to prior information. A normalization step is crucial prior to PCA, this is to make sure that each variable in the dataset contributes equally to the analysis.[167]

Gene expression studies and a small number of biological samples present a statistical challenge, therefore the Limma package was used to reduce potential false positive results. Limma borrows information across genes variability to produce a more stable analysis.[168,169]

Multiple hypothesis testing is common in medical research and is used to avoid drawing a false-positive conclusion during hypotheses testing. The Bonferroni adjustment is the simplest and most used method for multiple testing correction.[170] It is done by dividing the significance level by the number of tests thar are being performed.[171,172]

The significance of a result must be quantified regarding the null hypothesis. To reject the null hypothesis a cut-off of 0.05 is commonly applied, this is referred to as p-value. With thousands of observations the likelihood of false-positive rate increases. Benjamini and Hochberg was as a multiple test adjustment method.[173]

Partial least squares (PLS) analysis is a multivariate approach that allows comparison between multiple response variables and the rest of the data. PLS helps visualize the relationship between the data and known variables, therefore it is a supervised analysis.[174,175]

Fisher's exact test is applied to determine if there are non-random associations between two categorical variables. It is widely applied in analysis with small sample size.[176,177]

3.3.5 Functional annotations and pathway analysis

Functional annotation is the process of collecting gene information. For example, gene name aliases, molecular and biological pathway membership, subcellular position, domains, etc.[178] Here, functional annotations and pathway analysis have been applied on genes that were negatively correlated to DNA methylation to help understand the biological relevance.

Pathway analysis was done by over-representation analysis with the Kyoto Encyclopaedia of Genes and Genomes (KEGG) [*www.genome.jp/Kegg/*], and the PANTHER classification system [*http://pantherdb.org/*]. The human gene database GeneCards [*http://www.genecards.org/*] was used for single gene annotations.

3.3.6 Cell deconvolution

Mucosal biopsies contain various cell types, this presents a problem when comparing these samples. In differential gene expression this introduces a confounding issue. Changes in gene expression may reflect changes in sample cell composition. To solve this issue cell deconvolution method can be applied. This method can estimate cell type fraction in a sample.[179] *In silico* cell deconvolution was applied to the transcriptome dataset to obtain cell type-specific expression profiles.

4. Summary of results

4.1 Paper I

Transcriptomic Landscape of Treatment-Naïve Ulcerative Colitis

Hagar Taman, Christopher G. Fenton, Inga V. Hensel, Endre Anderssen, Jon Florholmen, Ruth H. Paulssen. (2017)

Journal of Crohn's and Colitis, 12(3), 327-336. doi:10.1093/ecco-jcc/jjx139

This study describes and characterize the transcriptomic landscape in UC. For this purpose, well- stratified, treatment-naïve UC patient samples were included representing mucosal biopsies from treatment-naïve UC patients (n=14) and healthy controls (n=16). Total RNA was subjected to RNA sequencing and data analysis.

Comparing treatment-naïve UC to controls revealed 1480 significantly DEGs. The fraction of epithelial cells was decreased in UC, while other cell population such as monocytes, neutrophils, T cells and B cells/lymphoid cells were increased during inflammation. Among the significantly DEGs, 79 DEGs were identified as IBD susceptibility genes, and 58 DEGs were expressed in a gender-specific manner. UC patients have a higher risk for developing CRC, the data identified four genes that could be considered CRC risk factors, especially for male UC patients. Additionally, the data revealed that AQP9 might have a functional role in the synthesis and/ or the function of mucus. AQP9 along with CLDN2, might regulate tissue specific physiological features in tight junctions of UC patients.

4.2 Paper II

Genome-wide DNA Methylation in Treatment-naïve Ulcerative Colitis

Hagar Taman, Christopher G. Fenton, Inga V. Hensel, Endre Anderssen, Jon Florholmen, Ruth H. Paulssen. (2018)

Journal of Crohn's and Colitis, 12(11), 1338-1347 doi.10.1093/ecco-jcc/jjy117

This study has two main aims; the first aim is to explore the genome-wide DNA methylation status in treatment-naïve UC, the second is to correlate the DNA methylation pattern to the gene expression levels in a well stratified treatment-naïve UC patient group. Therefore, mucosal biopsies from treatment-naïve UC patients (n= 10), and healthy controls (n=11) underwent genome-wide DNA bisulfite sequencing. To obtain a dataset with differentially methylated genes, the generated data was subjected to several statistical methods.

Of all the significantly differentially expressed genes (DEGs) obtained in Paper I, 25 % correlated with DNA methylation patterns. Thirty percent of the methylation in these genes occurred at CpG sites or near their transcription start site (TSS). It was clearly observed that genes involved in homeostasis and defence were hypermethylated, whereas genes involved in immune response were hypomethylated. The data also showed that 25 differentially methylated genes were identified as IBD-susceptibility genes. In addition, four genes showed methylation patterns despite the absence of known CpG islands.

In treatment-naïve UC, distinctive functional patterns for hyper- and hypomethylation were uncovered, which are of significance for the development and pathogenesis in UC.

4.3 Paper III

DNA hypo-methylation facilitates anti-inflammatory responses in severe ulcerative colitis

Hagar Taman, Christopher G. Fenton, Endre Anderssen, Jon Florholmen, Ruth H. Paulssen (2021) PLoS ONE, 16(4 April). doi: 10.1371/journal.pone.0248905

The main intention of this study is to define specific epigenetic changes that may possibly be responsible for the degree of disease severity. This was achieved by examining the relationship between transcriptomic and genome-wide DNA methylation profiles in a well-stratified, treatment-naïve severe UC patient population.

Three groups are compared in this study, therefore mucosal biopsies from 1) treatment-naïve severe UC patients (n=8), 2) treatment-naïve mild UC patients (n=8), and 3), healthy controls (n= 8) were collected and subjected to both whole transcriptome sequencing (RNA-Seq) and genome-wide DNA bisulfite sequencing. The generated data went through several statistical methods to gain a dataset with significantly differentially expressed genes (DEGs), which was correlated to DNA methylation in severe UC.

In comparing severe to mild UC, the data showed that DNA hypomethylation were observed in 80 % of all the correlated DEGs. Pathway annotations of the hypomethylated genes revealed neutrophil degranulation and immune-regulatory interactions of the lymphoid system. Anti-inflammatory genes were found to be hypomethylated in severe UC, such as IL-10, SIGLEC5, CD86, CLMP and members of inflammasomes NLRP3 and NLRC4. Hypomethylation of these genes indicate an interaction between the epithelium and lamina propria to diminish inflammation in the gut.

5. Discussion

Experimental considerations

Why study treatment-naïve UC? Numerous studies have tried to describe both gene expression and DNA methylation profiles in UC to solve the molecular mechanisms behind disease pathogenesis and development. Pharmaceutical drugs in general can lead to alterations in gene expression and epigenetic homeostasis.[180,181] Therefore, treatment strategies used for inducing remission in UC introduce systematic biases in data analysis. Previous and current studies have shown that using different treatment strategies for managing UC, such as immunosuppressant and biologics have short- and long-term side effect on the immune response.[182,183] Whereas other medications have the ability to inhibit the DNA methylation process.[184,185] The aim of this study is to investigate the DNA methylation status in UC at disease onset, without medication interference. To remove any effects of pharmaceuticals, it was essential to only include treatment-naïve material in this study. In addition, this approach may help identify new and important factors for predicting clinical outcome and crucial targets for future therapeutic treatment, which will be useful in customising and personalizing treatment strategies for UC patients. The patient's conditions included in this study are mild, moderate, and severe disease activity, which has been established based on the ECCO guidelines and the UC Disease Activity Index (UCDAI).[12,155,156] The control group consists of samples taken from subjects that underwent colon cancer screening examination.

The effects of sample size and selection on results. The study was designed as a pilot study with only a limited number of patient samples included. The sample size was considered sufficient to perform experiments with sufficient statistical power. Some samples did not meet the criteria required to continue with library preparation due to poor quality or low RNA/DNA content which reduced the sample size. In addition, only 15-30 % of the UC patient population will experience at least one episode of severe colitis during the disease course[186–188], which made it challenging to obtain a sufficient sample size (**Paper III**). Ideally, it would have been more favourable to have a large sample size to work with as larger sample sizes give more robust statistics.[189,190]

The sample size was completely dependent on the biopsy selection available in the biobank. During biopsy selection the main goal was to include a well-documented stratified treatment-naïve UC patients to allow for proper stratification of sample material. Thus, avoiding the mixture of UC disease state in downstream analysis. The study included participants with similar characteristics in respect to age and gender distribution. Nevertheless, the age distribution in the control group was higher compared to the UC group. This is due to colon cancer screening being more common among the elder population. The age distribution in mild and moderate UC patients are similar, however, the age distribution in severe UC patients is slightly higher. This is due to the increased risk for acute severe UC as patients get older.[191] Additionally, even though the gender ratio male/female is similar in the study groups, the number of male subjects is higher than the female subjects. Preferably an equal distribution between subjects in respect to gender and age would have been advantageous. The samples were taken from the ASIB biobank. They should reflect the lifestyle, diet, and ethnicity of the Norwegian population. One could argue that the results generated from this study is not applicable to other populations. Disease condition accounted for most of the PCA explained variability, whereas age and other characteristics added very little if any except for gender. The PCA did reveal a gender-specific separation between the samples in disease condition (**Paper I**).

Intestinal biopsies are heterogenous, it is composed of different cell types such as stromal, epithelial, and immune cells[192], which makes data analysis more challenging. In complex samples it is difficult to determine whether changes in gene expression represent changes in cell composition or genetic regulation. The biopsies were collected from an active inflammation site in the sigmoid (**Paper I & II**) and the recto-sigmoid (**Paper III**) part of the colon. The amount of different cell types in these tissues will have an impact on the expression profile.[192] Therefore, to determine the relative quantities of the different cell types within these samples, a cell deconvolution method was applied. This method provides an estimation of the relative quantities of different cell types present in the samples. To achieve accurate measurements the emerging technology of single-cell RNA-Seq (scRNA-Seq) should be considered in future studies to overcome the heterogeneity in tissue biopsies.[193,194]

Due to the hard conditions during bisulfite conversion genomic DNA is degraded and fragmented, resulting in small sequences that are about 500 nucleotides in length.[195] In addition, for the deamination to be efficient, the DNA are required to be protein free.[196] Therefore, isolating an adequate amount of DNA that is free of contamination is crucial for a successful bisulfite treatment. Another disadvantage with bisulfite treatment is the high read duplication rate caused by the enrichment of CpG-rich parts of the genome.[197–199] Unfortunately, due to high duplication rates a few samples were eliminated from the dataset after the preliminary data analysis. Despite all these disadvantages, bisulfite treatment is still considered the golden standard when measuring DNA methylation. This is mainly because of their single base resolution, and the low input requirements.[197]

The transcriptomic profile of treatment naïve UC

The main objective of this work is to investigate how epigenetic modification like DNA methylation contributes to the pathogenesis of UC. One would expect heavily methylated promoters to express less transcripts, and vice versa. Therefore, it was essential to establish a gene expression profile per sample from which to compare the sample DNA methylation profile. DNA methylation can regulate gene expression by supressing the transcriptional process, and it has been linked to disease development.[113,126,200] Altered methylation status in IBD associated genes has been implicated as transcriptional activity modifiers, and therefore contributing to disease development and progression.[36,48]

The transcriptomic profile in mild to moderate UC revealed that the inflammatory response in both the innate and adaptive immune system plays a crucial role in UC development. This is reflected in the increased fractions of different immune cells (**Paper I**). Furthermore, we observed that the fraction of epithelial cells is diminished as a function of disease severity (**Paper I & III**). Especially in severe UC epithelial cell fractions were further reduced when compared to mild UC (**Paper III**). Intestinal epithelial cells are part of the defence mechanisms utilized by the intestinal mucosal immune system to protect itself from harmful pathogens.[201] They build a physical barrier that protects against damaging luminal microenvironment, while offering

selective permeability for nutrition absorption. Intestinal epithelial cells are crucial for maintaining intestinal homeostasis.[202,203] A reduction in epithelial cells results in an impaired intestinal permeability and dysregulation of homeostasis. [204–207] Perhaps the observed hypermethylation and downregulation of genes involved in epithelial cells homeostasis, such as proline-rich acidic protein 1 (PRAP1), and members of the solute carrier protein family (SLC6A19 and SLC3A1) (**Paper II**) helps explain this reduction. In inflammatory gut conditions TNF-α levels together with other inflammatory cytokines are increased, which impacts the composition of tight junction within intestinal epithelium cells, resulting in increased intestinal permeability and oedema.[208–210] This dysregulation of water and solute homeostasis has been implied in UC development previously.[58,204,211,212]

Aquaporin 9 (AQP9) a previously unreported gene, and claudin 2 (CLDN2) are two of the most prominent differentially expressed genes in our mild to moderate UC patient material (**Paper I**). AQP9 is a water channel found in the membrane of intestinal tight junctions, whereas claudins are one of the membrane proteins that compile tight junctions in the intestinal epithelium.[210,213] CLDN2 is defined as a "pore-forming" protein, which facilitates para-cellular water transport across tight junctions in impaired epithelium.[214,215] The upregulation of AQP9 and CLDN2 observed in mild UC could be a response to inflammation, as they fight the disruptions within the epithelial barrier of the colon. Considering this, one might suggest that CLDN2 and AQP9 share similar properties regarding water transport, and our results may indicate that CLDN2 together with AQP9 regulate tissue-specific pathophysiological properties of tight junctions in UC. AQP9 might have another function, it was recently found to be expressed in a subset of mucin-producing goblet cells in colon and small intestine, indicating the potential role it has in mucus secretion.[216,217]

Normally, goblet cells are depleted during intestinal infection and a reduction in mucin synthesis and secretion occurs. Consequently, this allows pathogens access to the underlying epithelium.[218] However, several mucins displayed upregulated expressions in mild UC such as mucin 1 (MUC1), mucin 16 (MUC16), mucin 4 (MUC4), mucin 5ac (MUC5AC), and mucin 5b (MUC5B). Goblet cells are specialized intestinal epithelial cells, they are responsible for the protection of the mucus layer through mucin production, and they constitute a vital part of the

innate immune and antimicrobial defence.[203,219] MUC1 and MUC4 are transmembrane mucins, they serve to protect the epithelium,[216] increased levels of MUC1 and MUC4 have been confirmed previously.[220,221] In our dataset MUC5B shows increased expression in mild UC, this is contradictory to other studies where MUC5B expression is increased in colorectal cancer (CRC), and not in UC.[222]

Excitingly, our result shows that gender differences do occur at a molecular level between treatment-naïve UC patients, which has not been reported previously. The most prominent gender differences for UC were noted for the following genes: mucin 5B (MUC5B), regenerating family member 3 alpha (REG3A), defensin A5 (DEFA5) and interleukin 33 (IL33). IL33, REG3A and DEF5A showed increased expression in male UC compared to females. Gender-specific differences in IBD have been reported for some aspects such as disease presentation, disease course and complications.[223] In UC specifically, the incidence in both genders is similar before the age of 45, after that the incidence is higher among males.[224] Nevertheless, gender differences data are limited and with contradictory findings.[223,225] Male IBD patients are associated with a more severe disease course, and higher risk for developing CRC.[225,226] In UC males have higher risk of developing CRC.[217,227–229] IL33 contributes to CRC progression, an upregulation of REG3A is established in CRC tissue.[230,231] Single nucleotide polymorphism in human DEF5A may confer susceptibility to IBD.[232] Taken together these genes can be considered as CRC risk factors in UC.

Hypermethylation in treatment naïve UC

Several studies have demonstrated that hypermethylation of certain genes can result in a higher risk of developing colitis associated CRC.[233–236] The long lasting inflammation is linked to hypermethylation of gene promoter sites, and therefore, DNA methylation levels can be used as a biomarker for detecting patients at high risk for developing CRC.[138]

Interestingly, UDP glucuronosyltransferase family 1 member A8 (UGT1A8) was hypermethylated in both mild to moderate and severe UC (**Paper II, Paper III**). It is expressed in the colon and participates in drug metabolism.[237] A recent study have demonstrated that UGT1A8 expression is significantly downregulated in colorectal tumour tissue.[238] Taken together one could

suggest that hypermethylation and downregulation of UGT1A8 in UC might increase the risk for developing CRC in UC patients. UTG1A8 could potentially be used as a biomarker for CRC.

For mild UC, butyrophilin-like 3 (BTNL3) was observed to be hypermethylated with downregulated gene expression (**Paper II**). BTNL3 shares structural features with B7-molecules, and therefore, it can regulate T-cell mediated immune responses. BTNL3 has been associated with CRC, and it might have a role in CRC progression and development.[239] In inflamed mucosa the fraction of T lymphocytes is increased (**Paper I & III**), perhaps because of the hypermethylation of BTNL3. It is interesting that the hypermethylation of BTNL3 did not occur at CpG sites nor at any of the known regulatory transcriptional *cis*-acting elements like DNAse1 and enhancers, indicating novel regulatory features of DNA methylation in UC. It is known that DNA methylation usually occurs in areas rich in CpG sites, however it is not limited to them. The impact of methylation in non CpG rich areas on gene expression is gaining importance.[240] Interestingly, we observed that DNA methylation could occur at other areas than super enhancers, DNAse1 accessible, CpG island, and transcription factor binding sites, and still have an impact on gene expression (**Paper II**). Non-CpG site methylation has been previously reported in tissue of mouse brain.[241] In treatment-naïve UC, the genes affected by non-CpG methylation are involved in pro-inflammatory response and possible antimicrobial activities.

Two genes which are related to gut mucosal defence mechanisms were hypermethylated in mild UC, intestinal alkaline phosphatase 1 (ALPI) and defensin B1 (DEFB1) (**Paper II**). ALPI maintain gut homeostasis by hydrolysing lipopolysaccharides (LPS) in cell walls of gram-negative bacteria, which might contribute to systemic inflammation. Inactivation of LPS inhibits the stimulation of pro-inflammatory cytokines.[242] Reduced expression of ALPI has been associated with elevated intestinal inflammation, bacterial translocation and dysbiosis.[242] In mild UC this reduction might be explained by the hypermethylation and downregulation of ALPI.

Defensins are expressed in epithelial cells and are involved in innate immune response. DEFB1 encodes the human β-defensin 1 (hDB-1) which is an antimicrobial compound expressed in human colonic and epithelial cells. It is suggested that hDB-1 plays a key role in intestinal microbe regulation.[243] hDB-1 reduction in the intestinal epithelia is facilitated by the protein

thioredoxin, which have been found to be reduced in inflammation sites in CD. It is implicated that reduction in thioredoxin impairs the defence mechanism against intestinal microbiota by decreasing the antimicrobial activity of hBD-1.[243] The same mechanism could be implied for UC, in addition to the observed hypermethylation.

Hypomethylation in treatment naïve UC

Few studies have characterized DNA hypomethylation's role in the pathogenesis of UC. We noted that promoters of the interleukin family such as IL-10 and IL17A, and members of the C-X-C motif chemokine ligands (CXCL) were hypomethylated and the genes upregulated (**Paper III and II**). This is in concordance with previous studies.[53] It is noted though that the upregulation of these genes differed according to disease activity. At the transcriptomic level these genes were upregulated (**Paper I**), confirming the negative correlation between hypomethylation and upregulation of genes. Chemokines are chemotaxis proteins that play a key role in immune response, CXCL binds to CXC chemokine receptors (CXCR) to initiate and recruit leukocytes to inflammatory sites, different CXCL binds to different CXCR.[244] In mild UC, CXCL5, CXCL6, CXCR1 and CXCR2 were hypomethylated and upregulated (**Paper II)**. In UC the CXCR1/2- CXCL8 complex plays a pleiotropic role by recruiting neutrophils to the epithelium and lamina propria in inflammation site which result in elimination of microbial pathogens.[244] On the other hand, uncontrolled expression of pro-inflammatory cytokines, and exaggerated neutrophil recruitment are detrimental. This could lead to severe mucosal damage, compromising intestinal integrity and therefore promoting IBD-related processes, such as fibrosis, tissue injury, tumorigenesis and tumour progression.[244] High fractions of neutrophils are observed in both mild to moderate and severe UC (**Paper I & III**). In severe UC specifically, we noticed that two patients have significantly high fractions of neutrophils which could indicate a subtype of UC. A recent study has demonstrated that the expression of CXCL6 chemokine which binds to CXCR1, was upregulated in UC animal model.[245] This might suggest that the hypomethylation and upregulation observed for CXCL6 and CXCR1/2 in **Paper II** contributes to the excessive neutrophil recruitment, leading to further mucosal and tissue damage.

IL10 was hypomethylated and upregulated in severe UC versus mild UC (**Paper III**). During inflammation IL-10 exert an anti-inflammatory role where it limits secretion of pro-inflammatory cytokines such as IL-1, and TNF-α, which are increased during UC inflammation[85,97,246,247] The possible function of IL-10 during severe UC might be to initiate a fast and effective response against microbial invasion in UC.[87,248] Even though IL-10 mainly function as an anti-inflammatory cytokine, it can exert pro-inflammatory properties. It generates cytotoxic lymphocytes and activates pro-inflammatory B cells.[248,249] GWAS studies have shown that single nucleotide polymorphisms (SNPs) in IL-10 and its receptors IL-10Rα and IL-10Rβ are correlated with very early-onset of colitis.[246,247] IL-10 signalling is vital for maintaining gastrointestinal homeostasis, macrophages will downregulate the effect of pro-inflammatory mechanisms in the lamina propria as a response to IL-10 signalling.[249,250] Perhaps one could hypothesize that the observed hypomethylation and upregulation of IL-10 expression (**Paper III**) reduces the inflammation in severe UC.

Members of the NLR family pyrin domain containing NOD-like receptor family, such as NLRP3 and NLRP12, both were hypomethylated in severe UC (**Paper III**), which are members of the inflammasome family.[251] Inflammasomes detect intra- or extracellular stimuli. They respond by activating caspase-1, producing IL-1β, and IL-18, and start the inflammatory process.[251–253] Inflammasomes are expressed in colonic epithelial cells, where they regulate intestinal homeostasis.[251,252] NLRP3 and NLRP12 has been implicated to have a role in intestinal inflammation in murine models of colitis, and studies shows that NLRP3 and NLRP12 deficient mice: lose epithelial integrity, develop severe inflammation, and the risk of tumorigenesis increases.[251,254] This indicates the protective role they have against colitis. It is conceivable that the observed hypomethylation and upregulation of these genes could induce anti-inflammatory signals, which in turn counteracts severe inflammation and prevent further damage and consequently maintaining intestinal homeostasis and mitigating inflammation in severe UC. In summary, hypomethylation of anti-inflammatory genes may reduce the degree of inflammation in UC.

6. Conclusion

Combining well stratified, treatment-naïve UC patient material with high-throughput sequencing technology and correlating the transcript expression to the DNA methylation profile provided us with an insight into what occurs during UC disease initiation.

On the transcriptomic level, genes involved in water transport across tight junctions were upregulated, indicating their role in regulating tissue-specific pathophysiological features in the mucosa. Mucins, which protects the epithelium against harmful pathogens, were also upregulated, therefore confirming their protective role. Additionally, gender-specific differences between males and females have been observed which has not been reported previously. Upregulation of CRC related genes were observed in male UC patients, and they can be considered a CRC risk factor in UC.

Correlating DNA methylation to gene expression revealed hypermethylation of gene promoters involved in homeostasis of epithelial cells, resulting in downregulation of their gene expression. Hypermethylation was also observed in gen promoters involved in gut mucosal defence system. Hypomethylation on the other hand was detected in gene promoters involved in innate immune response and anti-inflammatory response. The observed hypomethylation of anti-inflammatory genes could have a potential role in reducing the degree of inflammation. Additionally, DNA methylation patterns were found in non CpG sites or known regulatory transcriptional cis-acting elements, which could indicate new regulatory features of DNA methylation in UC.

These findings should be considered when establishing new treatment strategies and personalized treatment for UC patients in the future.

7. Future perspectives

The quest to find the characterize mechanism causing UC continues. We would like to continue contributing to the field with our future studies. Studies that include single cell sequencing in combination with spatial transcriptomics is something we would like to pursue. Applying these two approaches will provide a deeper understanding of disease initiating mechanisms. We would like to investigate the effects of histone modification in UC. Additionally, we would like to explore the interplay between DNA methylation and histone modification in UC.

References

1. Burisch, J., Jess, T., Martinato, M., Lakatos, P. L. & EpiCom, E. The burden of inflammatory bowel disease in Europe. *J Crohns Colitis* **7**, 322–337 (2013).

2. Taman, H. *et al.* Transcriptomic Landscape of Treatment-Naive Ulcerative Colitis. *J Crohns Colitis* **12**, 327–336 (2018).

3. Loddo, I. & Romano, C. Inflammatory bowel disease: genetics, epigenetics, and pathogenesis. *Front. Immunol.* **6**, 551 (2015).

4. Taman, H. *et al.* Genome-wide DNA methylation intreatment-naïve ulcerative colitis. *J. Crohn's Colitis* **12**, 1338–1347 (2018).

5. Head, K. A., Jurenka, J. S. & Ascp, M. T. Inflammatory Bowel Disease Part I: Pathophysiology and Conventional and Alternative Treatment Options. *Altern Med Rev* **8**, 247–283 (2003).

6. Alatab, S. *et al.* The global, regional, and national burden of inflammatory bowel disease in 195 countries and territories, 1990-2017: a systematic analysis for the Global Burden of Disease Study 2017. *Lancet Gastroenterol Hepatol* **5**, 17–30 (2020).

7. Zhao, M. *et al.* The burden of inflammatory bowel disease in Europe in 2020. *J Crohns Colitis* (2021) doi:10.1093/ecco-jcc/jjab029.

8. Khor, B., Gardet, A. & Xavier, R. J. Genetics and pathogenesis of inflammatory bowel disease. *Nature* **474**, 307–317 (2011).

9. Danese, S. & Fiocchi, C. Ulcerative colitis. *N Engl J Med* **365**, 1713–1725 (2011).

10. Feuerstein, J. D., Moss, A. C. & Farraye, F. A. Ulcerative Colitis. *Mayo Clin. Proc.* **94**, 1357–1373 (2019).

11. Ordas, I., Eckmann, L., Talamini, M., Baumgart, D. C. & Sandborn, W. J. Ulcerative colitis. *Lancet* **380**, 1606–1619 (2012).

12. Magro, F. *et al.* Third European Evidence-based Consensus on Diagnosis and Management of Ulcerative Colitis. Part 1: Definitions, Diagnosis, Extra-intestinal Manifestations, Pregnancy, Cancer Surveillance, Surgery, and Ileo-anal Pouch Disorders for the European Crohn's and. *J. Crohn's Colitis* **11**, 1–39 (2017).

13. Kayal, M. & Shah, S. Ulcerative Colitis: Current and Emerging Treatment Strategies. *J. Clin. Med. 2020, Vol. 9, Page 94* **9**, 94 (2020).

14. Collins, P. & Rhodes, J. Ulcerative colitis: diagnosis and management What is it, and who gets it? *BMJ* **333**, 340–343 (2006).

15. Lamb, C. A. *et al.* British Society of Gastroenterology consensus guidelines on the management of inflammatory bowel disease in adults. *Gut* **68**, s1–s106 (2019).

16. Berends, S. E. *et al.* Clinical Pharmacokinetic and Pharmacodynamic Considerations in the Treatment of Ulcerative Colitis. *Clin Pharmacokinet* **58**, 15–37 (2019).

17. Arora, Z. & Shen, B. Biological therapy for ulcerative colitis. *Gastroenterol. Rep.* **3**, 103–109 (2015).

18. Seo, G. S. & Chae, S.-C. Biological therapy for ulcerative colitis: An update. *World J. Gastroenterol.*

20, 13234–13238 (2014).

19. Rogler, G. Medical management of ulcerative colitis. *Dig. Dis.* **27**, 542–549 (2009).

20. Gajendran, M. *et al.* A comprehensive review and update on ulcerative colitis. *Dis Mon* **65**, 100851 (2019).

21. Van Assche, G., Vermeire, S. & Rutgeerts, P. Management of acute severe ulcerative colitis. *Gut* **60**, 130–133 (2011).

22. Haifer, C., Leong, R. W. & Paramsothy, S. The role of faecal microbiota transplantation in the treatment of inflammatory bowel disease. *Curr Opin Pharmacol* **55**, 8–16 (2020).

23. Tan, P., Li, X., Shen, J. & Feng, Q. Fecal Microbiota Transplantation for the Treatment of Inflammatory Bowel Disease: An Update. *Front. Pharmacol.* **11**, 1409 (2020).

24. Imdad, A. *et al.* Fecal transplantation for treatment of inflammatory bowel disease. *Cochrane Database Syst. Rev.* **2018**, (2018).

25. Singh, A., Mahajan, R., Kao, D., Midha, V. & Sood, A. Long term management of ulcerative colitis with Faecal Microbiota Transplantation. *Med. Microecol.* **6**, 100026 (2020).

26. Lai, L., Li, H., Feng, Q., Shen, J. & Ran, Z. Multi-factor mediated functional modules identify novel classification of ulcerative colitis and functional gene panel. *Sci. Reports* | **11**, 5669 (2021).

27. Satsangi, J., Silverberg, M. S., Vermeire, S. & Colombel, J.-F. The Montreal classification of inflammatory bowel disease: controversies, consensus, and implications. *Gut* **55**, 749–753 (2006).

28. Ungaro, R., Mehandru, S., Allen, P. B., Peyrin-Biroulet, L. & Colombel, J. F. Ulcerative colitis. *Lancet* **389**, 1756–1770 (2017).

29. Magro, F. *et al.* Comparison of different histological indexes in the assessment of UC activity and their accuracy regarding endoscopic outcomes and faecal calprotectin levels. *Gut* **68**, 594–603 (2019).

30. Ramos, G. P. & Papadakis, K. A. Mechanisms of Disease: Inflammatory Bowel Diseases. *Mayo Clin. Proc.* **94**, 155–165 (2019).

31. De Souza, H. S. P. & Fiocchi, C. Immunopathogenesis of IBD: current state of the art. *Nat. Rev. Gastroenterol. Hepatol.* **13**, 13–27 (2016).

32. De Souza, H. S. P., Fiocchi, C. & Iliopoulos, D. The IBD interactome: An integrated view of aetiology, pathogenesis and therapy. *Nature Reviews Gastroenterology and Hepatology* vol. 14 739–749 (2017).

33. de Souza, H. S. P. Etiopathogenesis of inflammatory bowel disease: today and tomorrow. *Curr Opin Gastroenterol* **33**, 222–229 (2017).

34. Lee, S. H., Kwon, J. eun & Cho, M. La. Immunological pathogenesis of inflammatory bowel disease. *Intestinal Research* vol. 16 26–42 (2018).

35. Guan, Q. A Comprehensive Review and Update on the Pathogenesis of Inflammatory Bowel Disease. *J. Immunol. Res.* **2019**, (2019).

36. Zeng, Z., Mukherjee, A. & Zhang, H. From Genetics to Epigenetics, Roles of Epigenetics in

Inflammatory Bowel Disease. *Front Genet* **10**, 1017 (2019).

37. Annese, V. Genetics and epigenetics of IBD. *Pharmacological Research* vol. 159 (2020).

38. Xavier, R. J. & Podolsky, D. K. Unravelling the pathogenesis of inflammatory bowel disease. *Nature* **448**, 427–434 (2007).

39. Cleynen, I. & Vermeire, S. The genetic architecture of inflammatory bowel disease: Past, present and future. *Curr. Opin. Gastroenterol.* **31**, 456–463 (2015).

40. Ogura, Y. *et al.* A frameshift mutation in NOD2 associated with susceptibility to Crohn's disease. *Nature* **411**, 603–606 (2001).

41. Hugot, J. P. *et al.* Association of NOD2 leucine-rich repeat variants with susceptibility to Crohn's disease. *Nature* **411**, 599–603 (2001).

42. de Lange, K. M. *et al.* Genome-wide association study implicates immune activation of multiple integrin genes in inflammatory bowel disease. *Nat. Publ. Gr.* **49**, 256–261 (2017).

43. Jostins, L. *et al.* Host-microbe interactions have shaped the genetic architecture of inflammatory bowel disease. *Nature* **49**, 119–124 (2012).

44. Uhlig, H. H. & Muise, A. M. Clinical Genomics in Inflammatory Bowel Disease. *Trends in Genetics* vol. 33 (2017).

45. Anderson, C. A. *et al.* Meta-analysis identifies 29 additional ulcerative colitis risk loci, increasing the number of confirmed associations to 47. *Nat Genet* **43**, 246–252 (2011).

46. Barrett, J. C. *et al.* Genome-wide association defines more than 30 distinct susceptibility loci for Crohn's disease. *Nat. Genet.* **40**, 955–962 (2008).

47. Franke, A. *et al.* Meta-Analysis Increases to 71 the Tally of confirmed Crohn's Disease Susceptibility Loci. *Nat Genet* **42**, 1118–1125 (2010).

48. Marcin, W. & Michael, S. Genetics and epigenetics of inflammatory bowel disease. *Swiss Medical Weekly* vol. 148 (2018).

49. Liu, J. Z. *et al.* Association analyses identify 38 susceptibility loci for inflammatory bowel disease and highlight shared genetic risk across populations Europe PMC Funders Group. *Nat. Genet.* **47**, 979–986 (2015).

50. Kaplan, G. G. The global burden of IBD: from 2015 to 2025. *Nat Rev Gastroenterol Hepatol* **12**, 720–727 (2015).

51. Abegunde, A. T. *et al.* Environmental risk factors for inflammatory bowel diseases: Evidence based literature review EVIDENCE-BASED MEDICINE. *World J Gastroenterol* **22**, 6296–6317 (2016).

52. Carbonnel, F., Jantchou, P., Monnet, E. & Cosnes, J. Environmental risk factors in Crohn's disease and ulcerative colitis: an update. *Gastroentérologie Clin. Biol.* **33 Suppl 3**, S145–S157 (2009).

53. Legaki, E. & Gazouli, M. Influence of environmental factors in the development of inflammatory bowel diseases. *World J. Gastrointest. Pharmacol. Ther.* **7**, 112–125 (2016).

54. Ananthakrishnan, A. N. *et al.* Environmental triggers in IBD: A review of progress and evidence. *Nat. Rev. Gastroenterol. Hepatol.* **15**, 39–49 (2018).

55. Klement, E., Cohen, R. V., Boxman, J., Joseph, A. & Reif, S. Breastfeeding and risk of inflammatory bowel disease: a systematic review with meta-analysis. *The American journal of clinical nutrition* vol. 80 1342–1352 (2004).

56. Sun, Y. *et al.* Stress Triggers Flare of Inflammatory Bowel Disease in Children and Adults. *Frontiers in Pediatrics* vol. 7 432 (2019).

57. Zuo, T. & Ng, S. C. The Gut Microbiota in the Pathogenesis and Therapeutics of Inflammatory bowel disease. *Frontiers in Microbiology* vol. 9 2247 (2018).

58. Antoni, L., Nuding, S., Wehkamp, J. & Stange, E. F. Intestinal barrier in inflammatory bowel disease. *World J. Gastroenterol.* **20**, 1165–1179 (2014).

59. Nishida, A. *et al.* Gut microbiota in the pathogenesis of inflammatory bowel disease. *Clin. J. Gastroenterol.* **11**, 1–10 (2018).

60. Mentella, M. C., Scaldaferri, F., Pizzoferrato, M., Gasbarrini, A. & Miggiano, G. A. D. Nutrition, IBD and Gut Microbiota: A Review. *Nutrients* vol. 12 (2020).

61. Casén, C. *et al.* Deviations in human gut microbiota: A novel diagnostic test for determining dysbiosis in patients with IBS or IBD. *Aliment. Pharmacol. Ther.* **42**, 71–83 (2015).

62. Sokol, H. *et al.* Fungal microbiota dysbiosis in IBD. *Gut* **66**, 1039–1048 (2017).

63. Matsuoka, K. & Kanai, T. The gut microbiota and inflammatory bowel disease. *Seminars in Immunopathology* vol. 37 47–55 (2015).

64. Frank, D. N. *et al.* Molecular-phylogenetic characterization of microbial community imbalances in human inflammatory bowel diseases. *Proc. Natl. Acad. Sci. U. S. A.* **104**, 13780–13785 (2007).

65. Jandhyala, S. M. *et al.* Role of the normal gut microbiota. *World J. Gastroenterol.* **21**, 8836–8847 (2015).

66. Sokol, H. *et al.* Faecalibacterium prausnitzii is an anti-inflammatory commensal bacterium identified by gut microbiota analysis of Crohn disease patients. *Proc. Natl. Acad. Sci. U. S. A.* **105**, 16731–16736 (2008).

67. Wang, W. *et al.* Increased proportions of Bifidobacterium and the Lactobacillus group and loss of butyrate-producing bacteria in inflammatory bowel disease. *J. Clin. Microbiol.* **52**, 398–406 (2014).

68. Lane, E. R., Zisman, T. L. & Suskind, D. L. The microbiota in inflammatory bowel disease: current and therapeutic insights. *J. Inflamm. Res.* **10**, 63–73 (2017).

69. Gevers, D. *et al.* The treatment-naive microbiome in new-onset Crohn's disease. *Cell Host Microbe* **15**, 382–392 (2014).

70. Huda-Faujan, N. *et al.* The Impact of the Level of the Intestinal Short Chain Fatty Acids in Inflammatory Bowel Disease Patients Versus Healthy Subjects. *Open Biochem. J.* **4**, 53–58 (2010).

71. Rizzello, F. *et al.* Implications of the westernized diet in the onset and progression of IBD. *Nutrients* vol. 11 (2019).

72. Weber, A. T., Shah, N. D., Sauk, J. & Limketkai, B. N. Popular Diet Trends for Inflammatory Bowel Diseases: Claims and Evidence. *Curr Treat Options Gastro* **17**, 564–576 (2019).

73. Raphael, W. & Sordillo, L. M. Dietary polyunsaturated fatty acids and inflammation: The role of phospholipid biosynthesis. *International Journal of Molecular Sciences* vol. 14 21167–21188 (2013).

74. Shoda, R., Matsueda, K., Yamato, S. & Umeda, N. Epidemiologic analysis of Crohn disease in Japan: Increased dietary intake of n-6 polyunsaturated fatty acids and animal protein relates to the increased incidence of Crohn disease in Japan. *Am. J. Clin. Nutr.* **63**, 741–745 (1996).

75. Hou, J. K., Abraham, B. & El-Serag, H. Dietary intake and risk of developing inflammatory bowel disease: A systematic review of the literature. *Am. J. Gastroenterol.* **106**, 563–573 (2011).

76. Dixon, L. J., Kabi, A., Nickerson, K. P. & McDonald, C. Combinatorial effects of diet and genetics on inflammatory bowel disease pathogenesis. *Inflammatory Bowel Diseases* vol. 21 912–922 (2015).

77. Jowett, S. L. *et al.* Influence of dietary factors on the clinical course of ulcerative colitis: A prospective cohort study. *Gut* **53**, 1479–1484 (2004).

78. Hart, A. R. *et al.* Diet in the aetiology of ulcerative colitis: A European prospective cohort study. *Digestion* **77**, 57–64 (2008).

79. Limdi, J. K. Dietary practices and inflammatory bowel disease. *Indian Journal of Gastroenterology* vol. 37 284–292 (2018).

80. Martinez-Medina, M. *et al.* Western diet induces dysbiosis with increased e coli in CEABAC10 mice, alters host barrier function favouring AIEC colonisation. *Gut* **63**, 116–124 (2014).

81. Statovci, D., Aguilera, M., MacSharry, J. & Melgar, S. The impact of western diet and nutrients on the microbiota and immune response at mucosal interfaces. *Frontiers in Immunology* vol. 8 838 (2017).

82. Sonnenberg, A. Geograhic and temporal variations of sugar and margarine consumption in relation to Crohn's disease. *Digestion* **41**, 161–171 (1988).

83. Looijer-van Langen, M. A. C. & Dieleman, L. A. Prebiotics in chronic intestinal inflammation. *Inflammatory Bowel Diseases* vol. 15 454–462 (2009).

84. Rogler, G. Resolution of inflammation in inflammatory bowel disease. *Lancet Gastroenterol. Hepatol.* **2**, 521–530 (2017).

85. Tatiya-Aphiradee, N., Chatuphonprasert, W. & Jarukamjorn, K. Immune response and inflammatory pathway of ulcerative colitis. *J. Basic Clin. Physiol. Pharmacol.* **30**, 1–10 (2019).

86. Patente, T. A. *et al.* Human Dendritic Cells: Their Heterogeneity and Clinical Application Potential in Cancer Immunotherapy. *Fron Immunol* **0**, 3176 (2019).

87. Geremia, A., Biancheri, P., Allan, P., Corazza, G. R. & Di Sabatino, A. Innate and adaptive immunity in inflammatory bowel disease. *Autoimmunity Reviews* vol. 13 3–10 (2014).

88. Wallace, K. L., Zheng, L. B., Kanazawa, Y. & Shih, D. Q. Immunopathology of inflammatory bowel disease. *World J. Gastroenterol.* **20**, 6–21 (2014).

89. Pennock, N. D. *et al.* T cell responses: naïve to memory and everything in between. *Adv Physiol Educ* **37**, 273–283 (2013).

90. Fujino, S. *et al.* Increased expression of interleukin 17 in inflammatory bowel disease. *Gut* **52**, 65–70 (2003).

91. Baumgart, D. C. & Sandborn, W. J. Crohn's disease. in *The Lancet* vol. 380 1590–1605 (Elsevier B.V., 2012).

92. Kempski, J., Brockmann, L., Gagliani, N. & Huber, S. TH17 cell and epithelial cell crosstalk during inflammatory bowel disease and carcinogenesis. *Front. Immunol.* **8**, 1373 (2017).

93. Castro-Dopico, T. & Clatworthy, M. R. IgG and Fcgamma Receptors in Intestinal Immunity and Inflammation. *Front Immunol* **10**, 805 (2019).

94. Friedrich, M., Pohin, M. & Powrie, F. Cytokine Networks in the Pathophysiology of Inflammatory Bowel Disease. *Immunity* **50**, 992–1006 (2019).

95. Neurath, M. F. *Cytokines in inflammatory bowel disease*. *Nature Reviews Immunology* vol. 14 329–342 (Nature Publishing Group, 2014).

96. Neurath, M. Targeting immune cell circuits and trafficking in inflammatory bowel disease. *Nat. Immunol.* **20**, 970–979 (2019).

97. Dinarello, C. A. Overview of the IL-1 family in innate inflammation and acquired immunity. *Immunol Rev* **281**, 8–27 (2018).

98. Katsanos, K. H. & Papadakis, K. A. Inflammatory Bowel Disease: Updates on Molecular Targets for Biologics. *Gut Liver* **11**, 455–463 (2017).

99. Dragoni, G., Innocenti, T. & Galli, A. Biomarkers of Inflammation in Inflammatory Bowel Disease: How Long before Abandoning Single-Marker Approaches? *Dig. Dis.* **39**, 190–203 (2021).

100. Norouzinia, M., Chaleshi, V., Alizadeh, A. H. M. & Zali, M. R. Biomarkers in inflammatory bowel diseases: insight into diagnosis, prognosis and treatment. *Gastroenterol. Hepatol. From Bed to Bench* **10**, 155 (2017).

101. Fengming, Y. & Jianbing, W. Biomarkers of inflammatory bowel disease. *Dis. Markers* **2014**, (2014).

102. Waddington, C. H. The epigenotype. 1942. *Int. J. Epidemiol.* **41**, 10–13 (2012).

103. Jenke, A. C. & Zilbauer, M. Epigenetics in inflammatory bowel disease. *Curr Opin Gastroenterol* **28**, 577–584 (2012).

104. Bonasio, R., Tu, S. & Reinberg, D. Molecular signals of epigenetic states. *Science (80-.).* **330**, 612–616 (2010).

105. Ventham, N. T., Kennedy, N. A., Nimmo, E. R. & Satsangi, J. Beyond gene discovery in inflammatory bowel disease: The emerging role of epigenetics. *Gastroenterology* **145**, 293–308 (2013).

106. Arora, I. & Tollefsbol, T. O. Computational methods and next-generation sequencing approaches to analyze epigenetics data: Profiling of methods and applications. *Methods* **187**, 92–103 (2021).

107. Weinhold, B. Epigenetics: The Science of Change. *Environ. Health Perspect.* **114**, A160 (2006).

108. Scarpa, M. & Stylianou, E. Epigenetics: Concepts and relevance to IBD pathogenesis. *Inflamm Bowel Dis* **18**, 1982–1996 (2012).

109. Gibney, E. R. & Nolan, C. M. Epigenetics and gene expression. *Hered.* **105**, 4–13 (2010).

110. Portela, A. & Esteller, M. Epigenetic modifications and human disease. *Nat Biotechnol* **28**, 1057–1068 (2010).

111. Bayarsaihan, D. Epigenetic mechanisms in inflammation. *J Dent Res* **90**, 9–17 (2011).

112. Roberti, A., Valdes, A. F., Torrecillas, R., Fraga, M. F. & Fernandez, A. F. Epigenetics in cancer therapy and nanomedicine. *Clin. Epigenetics* **11**, 1–18 (2019).

113. Moore, L. D., Le, T. & Fan, G. DNA Methylation and Its Basic Function. *Neuropsychopharmacology* **38**, 23–38 (2012).

114. Compere, S. J. & Palmiter, R. D. DNA methylation controls the inducibility of the mouse metallothionein-I gene in lymphoid cells. *Cell* **25**, 233–240 (1981).

115. Holliday, R. & Pugh, J. DNA modification mechanisms and gene activity during development. *Science (80-.).* **187**, 226–232 (1975).

116. Fogel, O., Richard-Miceli, C. & Tost, J. Epigenetic Changes in Chronic Inflammatory Diseases. in *Advances in Protein Chemistry and Structural Biology* vol. 106 139–189 (Academic Press Inc., 2017).

117. Däbritz, J. & Menheniott, T. R. Linking immunity, epigenetics, and cancer in inflammatory bowel disease. *Inflammatory Bowel Diseases* vol. 20 1638–1654 (2014).

118. Bestor, T. H. The DNA methyltransferases of mammals. *Human Molecular Genetics* vol. 9 2395–2402 (2000).

119. Reik, W., Dean, W. & Walter, J. Epigenetic reprogramming in mammalian development. *Science* vol. 293 1089–1093 (2001).

120. Wu, H. & Zhang, Y. Reversing DNA Methylation: Mechanisms, Genomics, and Biological Functions. *Cell* **156**, 45 (2014).

121. Wu, X. & Zhang, Y. TET-mediated active DNA demethylation: mechanism, function and beyond. *Nat Rev Genet* **18**, 517–534 (2017).

122. Valinluck, V. *et al.* Oxidative damage to methyl-CpG sequences inhibits the binding of the methyl-CpG binding domain (MBD) of methyl-CpG binding protein 2 (MeCP2). *Nucleic Acids Res.* **32**, 4100–4108 (2004).

123. Ventham, N. T. *et al.* Integrative epigenome-wide analysis demonstrates that DNA methylation may mediate genetic risk in inflammatory bowel disease. *Nat. Commun.* **7**, (2016).

124. Kumar, A. *et al.* Epigenetic modulation of intestinal Na(+)/H(+) exchanger-3 expression. *Am J Physiol Gastrointest Liver Physiol* **314**, G309-g318 (2018).

125. Kraiczy, J. *et al.* Assessing DNA methylation in the developing human intestinal epithelium: potential link to inflammatory bowel disease. *Mucosal Immunol* **9**, 647–658 (2016).

126. Li, E. & Zhang, Y. DNA methylation in mammals. *Cold Spring Harb Perspect Biol* **6**, a019133 (2014).

127. Hamilton, J. P. Epigenetics: principles and practice. *Dig Dis* **29**, 130–135 (2011).

128. Taman, H., Fenton, C. G., Anderssen, E., Florholmen, J. & Paulssen, R. H. DNA hypo-methylation facilitates anti-inflammatory responses in severe ulcerative colitis. *PLoS One* **16**, 1–17 (2021).

129. Medzhitov, R. & Horng, T. Transcriptional control of the inflammatory response. *Nature Reviews Immunology* vol. 9 692–703 (2009).

130. Tahara, T. *et al.* Promoter methylation of protease-activated receptor (PAR2) is associated with

severe clinical phenotypes of ulcerative colitis (UC). *Clin Exp Med* **9**, 125–130 (2009).

131. Surace, A. E. A. & Hedrich, C. M. The Role of Epigenetics in Autoimmune/Inflammatory Disease. *Front. Immunol.* **10**, 1525 (2019).

132. Brandt, D., Sergon, M., Abraham, S., Mäbert, K. & Hedrich, C. M. TCR+ CD3+ CD4− CD8− effector T cells in psoriasis. *Clin. Immunol.* **181**, 51–59 (2017).

133. BM, J. *et al.* Changes in the pattern of DNA methylation associate with twin discordance in systemic lupus erythematosus. *Genome Res.* **20**, 170–179 (2010).

134. Zhao, M. *et al.* Hypomethylation of IL10 and IL13 pomoters in CD4+T clls of ptients with sastemic lpus erythematosus. *J. Biomed. Biotechnol.* **2010**, (2010).

135. Denizot, J. *et al.* Diet-induced hypoxia responsive element demethylation increases CEACAM6 expression, favouring Crohn's disease-associated Escherichia coli colonisation. *Gut* **64**, 428–437 (2015).

136. Yi, M. J., Kim, T. O., Joo, M. Y. & Tae, O. K. Epigenetic Alterations in Inflammatory Bowel Disease and Cancer. *Intest. Res.* **13**, 112–121 (2015).

137. Glória, L. *et al.* DNA hypomethylation and proliferative activity are increased in the rectal mucosa of patients with long-standing ulcerative colitis. *Cancer* **78**, 2300–2306 (1996).

138. Karatzas, P. S., Gazouli, M., Safioleas, M. & Mantzaris, G. J. DNA methylation changes in inflammatory bowel disease. *Ann. Gastroenterol.* **27**, 125–132 (2014).

139. Hsieh, C.-J. *et al.* Hypermethylation of the p16INK4a Promoter in Colectomy Specimens of Patients with Long-standing and Extensive Ulcerative Colitis. *Cancer Res.* **58**, (1998).

140. Azarschab, P., Porschen, R., Gregor, M., Blin, N. & Holzmann, K. Epigenetic control of the E-cadherin gene (CDH1) by CpG methylation in colectomy samples of patients with ulcerative colitis. *Genes Chromosom. Cancer* **35**, 121–126 (2002).

141. Cooke, J. *et al.* Mucosal Genome-wide Methylation Changes in Inflammatory Bowel Disease. *Inflamm. Bowel Dis. Bowel Dis* **18**, 2128–2137 (2012).

142. Bae, J., Park, J., Yang, K., Kim, T. & Yi, J. Detection of DNA hypermethylation in sera of patients with Crohn's disease. *Mol. Med. Rep.* **9**, 725–729 (2014).

143. Selzer, P. M., Marhöfer, R. J. & Koch, O. The Functional Analysis of Genomes. in *Applied Bioinformatics* 91–122 (Springer, Cham, 2018). doi:10.1007/978-3-319-68301-0_6.

144. Cobb, M. 60 years ago, Francis Crick changed the logic of biology. *PloS Biol* **15**, (2017).

145. Lowe, R., Shirley, N., Bleackley, M., Dolan, S. & Shafee, T. Transcriptomics technologies. *PLoS Comput. Biol.* **13**, (2017).

146. Alabert, C. & Groth, A. Chromatin replication and epigenome maintenance. *Nat Rev Mol Cell Biol* **13**, 153–167 (2012).

147. Dabin, J., Fortuny, A. & Polo, S. E. Epigenome maintenance in response to DNA damage. *Mol. Cell* **62**, 712–727 (2016).

148. Zhu, J. *et al.* Genome-wide chromatin state transitions associated with developmental and

environmental cues. *Cell* **152**, 642–654 (2013).

149. Laird, P. W. Principles and challenges of genomewide DNA methylation analysis. *Nat. Rev. Genet.* **11**, 191–203 (2010).

150. Tighe, D. *et al.* One-Year Clinical Outcomes in an IBD Cohort Who Have Previously Had Anti-TNFa Trough and Antibody Levels Assessed. *Inflamm. Bowel Dis.* **23**, 1154–1159 (2017).

151. Shivaji, U. N. *et al.* Clinical outcomes, predictors of prognosis and health economics consequences in IBD patients after discontinuation of the first biological therapy: *Ther Adv Gastroenterol* **13**, 1–16 (2020).

152. Doherty, D. *et al.* European Crohn's and Colitis Organisation Topical Review on Treatment Withdrawal ['Exit Strategies'] in Inflammatory Bowel Disease. *J. Crohns. Colitis* **12**, 17–31 (2018).

153. Magro, F. *et al.* Review of the disease course among adult ulcerative colitis population-based longitudinal cohorts. *Inflamm. Bowel Dis.* **18**, 573–583 (2012).

154. Siegel, C. Refocusing IBD Patient Management: Personalized, Proactive, and Patient-Centered Care. *Am. J. Gastroenterol.* **113**, 1440–1443 (2018).

155. McNally, P. R. Budesonide for the Induction of Remission for Mild- Moderate Ulcerative Colitis. *Visible Hum. J. Endoscpy* **13**, 2–7 (2014).

156. Tursi, A. *et al.* Treatment of relapsing mild-to-moderate ulcerative colitis with the probiotic VSL#3 as adjunctive to a standard pharmaceutical treatment: a double-blind, randomized, placebo-controlled study. *Am J Gastroenterol* **105**, 2218–2227 (2010).

157. Olsen, T. *et al.* Tissue levels of tumor necrosis factor-alpha correlates with grade of inflammation in untreated ulcerative colitis. *Scand. J. Gastroenterol.* **42**, 1312–1320 (2007).

158. Kappelmann-Fenzl, M. Library Construction for NGS. in 39–45 (Springer, Cham, 2021). doi:10.1007/978-3-030-62490-3_3.

159. Wolf, J. B. W. Principles of transcriptome analysis and gene expression quantification: An RNA-seq tutorial. *Mol. Ecol. Resour.* **13**, 559–572 (2013).

160. Marini, F. & Binder, H. pcaExplorer: an R/Bioconductor package for interacting with RNA-seq principal components. *BMC Bioinformatics* **20**, (2019).

161. Rajkumar, A. P. *et al.* Experimental validation of methods for differential gene expression analysis and sample pooling in RNA-seq. (2011) doi:10.1186/s12864-015-1767-y.

162. Xiao, W. H. *et al.* Identification of commonly dysregulated genes in colorectal cancer by integrating analysis of RNA-Seq data and qRT-PCR validation. *Cancer Gene Ther.* **22**, 278–284 (2015).

163. Dobin, A. & Gingeras, T. R. Mapping RNA-seq Reads with STAR. *Curr. Protoc. Bioinforma.* **51**, 11.14.1-11.14.19 (2015).

164. Bock, C. Analysing and interpreting DNA methylation data. *Nat. Rev. Genet. 2012 1310* **13**, 705–719 (2012).

165. Rauluseviciute, I., Drabløs, F. & Rye, M. B. DNA methylation data by sequencing: experimental approaches and recommendations for tools and pipelines for data analysis. *Clin Epigenetics* **11**, (2019).

166. Jolliffe, I. T. & Cadima, J. Principal component analysis: a review and recent developments. *Phil. Trans. R. Soc. A* **374**, (2016).

167. Lever, J., Krzywinski, M. & Altman, N. Principal component analysis. *Nat. Publ. Gr.* **14**, (2017).

168. Ritchie, M. E. *et al.* limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* **43**, e47 (2015).

169. Smyth, G. K. *et al. limma: Linear Models for Microarray and RNA-Seq Data User's Guide*. (2002).

170. Bonferroni, C. E. Teoria statistica delle classi e calcolo delle probabilita. *Libr. internazionale Seeber* **62**, (1936).

171. Noble, W. S. How does multiple testing correction work? *Nat. Biotechnol.* **27**, (2009).

172. Groenwold, R. H. H., Goeman, J. J., Cessie, S. Le & Dekkers, O. M. Multiple testing: when is many too much? *Eur. J. Endocrinol.* **184**, E11–E14 (2021).

173. Benjamini, Y. & Hochberg, Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to MultipleTesting. *J. R. Stat. Soc. B* **57**, (1995).

174. Abdi, H. & Williams, L. J. Partial Least Squares Methods: Partial Least Squares Correlation and Partial Least Square Regression. *Methods Mol. Biol.* **930**, 549–579 (2013).

175. Wold, S., Sjöström, M. & Eriksson, L. PLS-regression: a basic tool of chemometrics. *Chemom. Intell. Lab. Syst.* **58**, 109–130 (2001).

176. Kim, H.-Y. Statistical notes for clinical researchers: Chi-squared test and Fisher's exact test Open lecture on statistics. *Restor Dent Endod* **42**, 152–155 (2017).

177. Bind, M.-A. C. & Rubin, D. B. When possible, report a Fisher-exact P value and display its underlying null randomization distribution. *PNAS* **117**, 19151–19158 (2020).

178. Berardini, T. Z. *et al.* Functional Annotation of the Arabidopsis Genome Using Controlled Vocabularies. *Plant Physiol.* **135**, 745–755 (2004).

179. Jaakkola, M. K. & Elo, L. L. Computational deconvolution to estimate cell type-specific gene expression from bulk data. *NAR Genomics Bioinforma.* **3**, (2021).

180. Csoka, A. B. & Szyf, M. Epigenetic side-effects of common pharmaceuticals: a potential new field in medicine and pharmacology. *Med Hypotheses* **73**, 770–780 (2009).

181. Garcia-Calzon, S. *et al.* Diabetes medication associates with DNA methylation of metformin transporter genes in the human liver. *Clin Epigenetics* **9**, 102 (2017).

182. D'Haens, G. Risks And Benefit Of Biologic Therapy For Inflammatory Bowel Disease. *Gut* **56**, 725–732 (2007).

183. Rawla, P., Sunkara, T. & Raj, J. P. Role of biologics and biosimilars in inflammatory bowel disease: current trends and future perspectives. *J. Inflamm. Res.* **11**, 215 (2018).

184. Ciechomska, M., Roszkowski, L. & Maslinski, W. DNA Methylation as a Future Therapeutic and Diagnostic Target in Rheumatoid Arthritis. *Cells* **8**, (2019).

185. Brandt, B., Rashidiani, S., Bán, Á. & Rauch, T. A. DNA Methylation-Governed Gene Expression in Autoimmune Arthritis. *Int. J. Mol. Sci Sci.* **20**, (2019).

186. Jain, S., Ahuja, V. & Limdi, J. K. Optimal management of acute severe ulcerative colitis. *Postgr. Med J* (2019) doi:10.1136/postgradmedj-2018-136072.

187. Dulai, P. S. & Jairath, V. Acute severe ulcerative colitis: latest evidence and therapeutic implications. *Ther Adv Chronic Dis* **9**, 65–72 (2018).

188. Hindryckx, P., Jairath, V. & D'Haens, G. Acute severe ulcerative colitis: from pathophysiology to clinical management. *Nat Rev Gastroenterol Hepatol* **13**, 654–664 (2016).

189. Biau, D. J., Kernéis, S. & Porcher, R. Statistics in Brief: The Importance of Sample Size in the Planning and Interpretation of Medical Research. *Clin Ortho Relat Res* **466**, 2282–2288 (2008).

190. Faber, J. & Fonseca, L. M. How sample size influences research outcomes. *Dent. Press J Ortho* **19**, 27–29 (2014).

191. Dorofeyev, A., Rassokhina, O. & Dorofeyeva, A. Late-Onset of Acute Severe Ulcerative Colitis: Clinical Case. *Dig Dis* **35**, 45–49 (2017).

192. Boland, B. S. *et al.* Heterogeneity and clonal relationships of adaptive immune cells in ulcerative colitis revealed by single-cell analyses. *Sci. Immunol.* **5**, 4432 (2020).

193. Corridoni, D., Chapman, T., Antanaviciute, A., Satsangi, J. & Simmons, A. Inflammatory Bowel Disease Through the Lens of Single-cell RNA-seq Technologies. *Inflamm. Bowel Dis.* **26**, 1658–1668 (2020).

194. Li, G. *et al.* Identification of Novel Population-Specific Cell Subsets in Chinese Ulcerative Colitis Patients Using Single-Cell RNA Sequencing. *Cell. Mol. Gastroenterol. Hepatol.* **12**, 99–117 (2021).

195. Kumar, A., Dalan, E. & Carless, M. A. Chapter Three - Analysis of DNA methylation using pyrosequencing. in *Epigenetics Methods* (ed. Tollefsbol, T.) vol. 18 37–62 (Academic Press, 2020).

196. Darst, R. P., Pardo, C. E., Ai, L., Brown, K. D. & Kladde, M. P. Bisulfite sequencing of DNA. in *Curr Protoc Mol Biol* vol. Chapter 7 1–17 (2010).

197. Singer, B. D. A Practical Guide to the Measurement and Analysis of DNA Methylation. *Am J Respir Cell Mol Biol.* **61**, 417–428 (2019).

198. Feng, S., Zhong, Z., Wang, M. & Jacobsen, S. E. Efficient and accurate determination of genome-wide DNA methylation patterns in Arabidopsis thaliana with enzymatic methyl sequencing. *Epigenetics Chromatin* **13**, (2020).

199. Zhou, L. *et al.* systematic evaluation of library preparation methods and sequencing platforms for high-throughput whole genome bisulfite sequencing. *Sci Rep* **9**, (2019).

200. Spainhour, J. C., Lim, H. S., Yi, S. V & Qiu, P. Correlation Patterns Between DNA Methylation and Gene Expression in The Cancer Genome Atlas. *Cancer Inform.* **18**, (2019).

201. Shi, N., Li, N., Duan, X. & Niu, H. Interaction between the gut microbiome and mucosal immune system. *Mil. Med. Res. BioMed Cent. Ltd* **4**, (2017).

202. Garcia-Carbonell, R., Yao, S. J., Das, S. & Guma, M. Dysregulation of intestinal epithelial cell RIPK pathways promotes chronic inflammation in the IBD gut. *Front. Immunol* **10**, 1094 (2019).

203. Martini, E., Krug, S. M., Siegmund, B., Neurath, M. F. & Becker, C. Mend Your Fences: The Epithelial Barrier and its Relationship With Mucosal Immunity in Inflammatory Bowel Disease. *Cell Mol*

Gastroenterol Hepatolenterol Hepatol **4**, 33–46 (2017).

204. Mankertz, J. & Schulzke, J. R.-D. D. Altered permeability in inflammatory bowel disease: pathophysiology and clinical implications. *Curr Opin Gastroenterol* **23**, 379–383 (2007).

205. Zhang, J. *et al.* The proline-rich acidic protein is epigenetically regulated and inhibits growth of cancer cell lines. *Cancer Res.* **63**, 6658–6665 (2003).

206. Ghishan, F. K. & Kiela, P. R. Epithelial transport in inflammatory bowel diseases. *Inflamm. Bowel Dis.* **20**, 1099–1109 (2014).

207. Pérez-Torras, S. *et al.* Transportome profiling identifies profound alterations in crohn's disease partially restored by commensal bacteria. *J. Crohn's Colitis* **10**, 850–859 (2016).

208. He, F. *et al.* Mechanisms of tumor necrosis factor-alpha-induced leaks in intestine epithelial barrier. *Cytokine* **59**, 264–272 (2012).

209. Al-Sadi, R., Guo, S., Ye, D. & Ma, T. Y. TNF-a Modulation of Intestinal Epithelial Tight Junction Barrier Is Regulated by ERK1/2 Activation of Elk-1. *Am. J. Pathol.* **183**, 1871–1884 (2013).

210. Watari, A. *et al.* Rebeccamycin Attenuates TNF-α-Induced Intestinal Epithelial Barrier Dysfunction by Inhibiting Myosin Light Chain Kinase Production. *Cell. Physiol. Biochem.* **41**, 1924–1934 (2017).

211. Katinios, G. *et al.* Increased Colonic Epithelial Permeability and Mucosal Eosinophilia in Ulcerative Colitis in Remission Compared With Irritable Bowel Syndrome and Health. *Inflamm Bowel Dis* • **26**, (2020).

212. Michielan, A. & D'incà, R. Clinical Evaluation, and Therapy of Leaky Gut. *Mediat. Inflamm* **2015**, (2015).

213. Okada, S., Misaka, T., Matsumoto, I., Watanabe, H. & Abe, K. Aquaporin-9 is expressed in a mucus-secreting goblet cell subset in the small intestine. *FEBS Lett.* **540**, 157–162 (2003).

214. Landy, J. *et al.* Tight junctions in inflammatory bowel diseases and inflammatory bowel disease associated colorectal cancer. *World J. Gastroenterol.* **22**, 3117–3126 (2016).

215. Rosenthal, R. *et al.* Claudin-2, a component of the tight junction, forms a paracellular water channel. *J. Cell Sci.* **123**, 1913–1921 (2010).

216. Yang, S. & Yu, M. Role of Goblet Cells in Intestinal Barrier and Mucosal Immunity. *J Inflamm Res* **14**, 3171–3183 (2021).

217. Xie, J. & Itzkowitz, S. Cancer in inflammatory bowel disease. *World J Gastroenterol* **14**, 378–389 (2008).

218. Dorofeyev, A. E., Vasilenko, I. V, Rassokhina, O. A. & Kondratiuk, R. B. Mucosal barrier in ulcerative colitis and Crohn's disease. *Gastroenterol Res Pr.* **2013**, 431231 (2013).

219. Grondin, J. A., Kwon, Y. H., Far, P. M., Haq, S. & Khan, W. I. Mucins in Intestinal Mucosal Defense and Inflammation: Learning From Clinical and Experimental Studies. *Front. Immunol* **11**, (2020).

220. Bankole, E., Read, E., Curtis, M. A., Neves, J. F. & Garnett, J. A. The Relationship between Mucins and Ulcerative Colitis: A Systematic Review. *J. Clin. Med* **10**, (2021).

221. Vancamelbeke, M. *et al.* Genetic and Transcriptomic Bases of Intestinal Epithelial Barrier

Dysfunction in Inflammatory Bowel Disease. *Inflamm Bowel Dis* **0**, 1–12 (2017).

222. Walsh, M. D. *et al.* Expression of MUC2, MUC5AC, MUC5B, and MUC6 mucins in colorectal cancers and their association with the CpG island methylator phenotype. *Mod. Pathol.* **26**, 1642–1656 (2013).

223. Greuter, T. *et al.* Gender Differences in Inflammatory Bowel Disease. *Digestion* **101**, 98–104 (2020).

224. Rustgi, S. D., Kayal, M. & Shah, S. C. Sex-based differences in inflammatory bowel diseases: a review: *Ther Adv Gastroenterol* **13**, 1–11 (2020).

225. Cargiolli, M., Miranda, A. & Bottiglieri, M. E. Inflammatory bowel disease: Gender difference. *Ital. J. Gender-Specific Med.* **3**, 23–28 (2017).

226. Zelinkova, Z., Janneke Van Der Woude, C. & Zuzana Zelinkova, C. J. van der W. Gender and Inflammatory Bowel Disease. *J. Clin. Cell. Immunol.* **5**, 245–250 (2014).

227. Axelrad, J. E., Lichtiger, S. & Yajnik, V. Inflammatory bowel disease and cancer: The role of inflammation, immunosuppression, and cancer treatment. *World J Gastroenterol* **22**, 4794–4801 (2016).

228. Goodman, W. A., Erkkila, I. P. & Pizarro, T. T. Sex matters: impact on pathogenesis, presentation and treatment of inflammatory bowel disease. *Nat Rev Gastroenterol Hepatol* **17**, 740–754 (2020).

229. Hovde, Ø. *et al.* Malignancies in Patients with Inflammatory Bowel Disease: Results from 20 Years of Follow-up in the IBSEN Study. *J. Crohns. Colitis* **11**, 571–577 (2017).

230. Mertz, K. D. *et al.* The IL-33/ST2 pathway contributes to intestinal tumorigenesis in humans and mice. *Oncoimmunology* **5**, e1062966 (2016).

231. Ye, Y. *et al.* Up-regulation of REG3A in colorectal cancer cells confers proliferation and correlates with colorectal cancer risk. *Oncotarget* **7**, 3921–3933 (2015).

232. Ferguson, L. R. *et al.* Single nucleotide polymorphisms in human Paneth cell defensin A5 may confer susceptibility to inflammatory bowel disease in a New Zealand Caucasian population. *Dig. Liver Dis.* **40**, 723–730 (2008).

233. Kim, T. O. *et al.* DNA hypermethylation of a selective gene panel as a risk marker for colon cancer in patients with ulcerative colitis. *Int. J. Mol. Med.* **31**, 1255–1261 (2013).

234. KOBAYASHI, M. *et al.* Hypermethylation of Corticotropin Releasing Hormone Receptor-2 Gene in Ulcerative Colitis Associated Colorectal Cancer. *In Vivo (Brooklyn).* **34**, 57–63 (2020).

235. Planell, N. *et al.* Transcriptional analysis of the intestinal mucosa of patients with ulcerative colitis in remission reveals lasting epithelial cell alterations. *Gut* **62**, 967–976 (2013).

236. Choi, J., Kim, D.-W., Shin, S.-Y., Park, E.-C. & Kang, J.-G. Effect of Ulcerative Colitis on Incidence of Colorectal Cancer: Results from the Nationwide Population-Based Cohort Study (2003-2013). *J. Cancer* **7**, 681–686 (2016).

237. Huang, Y. H. *et al.* Identification and functional characterization of UDP-glucuronosyltransferases UGT1A8*1, UGT1A8*2 and UGT1A8*3. *Pharmacogenetics* **12**, 287–297 (2002).

238. Beyerle, J. *et al.* Expression Patterns of Xenobiotic-Metabolizing Enzymes in Tumor and Adjacent Normal Mucosa Tissues among Patients with Colorectal Cancer: The ColoCare Study. *Cancer*

Epidemiol. Biomarkers Prev. **29**, 460–469 (2020).

239. Lebrero-Fernandez, C. *et al.* Altered expression of Butyrophilin (BTN) and BTN-like (BTNL) genes in intestinal inflammation and colon cancer. *Immunity, Inflamm. Dis.* **4**, 191–200 (2016).

240. Arechederra, M. *et al.* Hypermethylation of gene body CpG islands predicts high dosage of functional oncogenes in liver cancer. *Nat Commun* **9**, 1–16 (2018).

241. Xie, W. *et al.* Base-resolution analyses of sequence and parent-of-origin dependent DNA methylation in the mouse genome. *Cell* **148**, 816–831 (2012).

242. Fawley, J. & Gourlay, D. M. Intestinal alkaline phosphatase: A summary of its role in clinical disease. *Journal of Surgical Research* vol. 202 225–234 (2016).

243. Kim, J. M. Antimicrobial Proteins in Intestine and Inflammatory Bowel Diseases. *Intest. Res.* **12**, 20–33 (2014).

244. Zhu, Y. *et al.* CXCL8 chemokine in ulcerative colitis. *Biomed. Pharmacother.* **138**, 111427 (2021).

245. Boshagh, M. A. *et al.* ELR positive CXCL chemokines are highly expressed in an animal model of ulcerative colitis. *J Inflamm Res* **12**, 167–174 (2019).

246. Engelhardt, K. & Grimbacher, B. IL-10 in humans: lessons from the gut, IL-10/IL-10 receptor deficiencies, and IL-10 polymorphisms. *Curr Top Microbiol Immunol* **380**, 1–18 (2014).

247. Wei, H.-X., Wang, B. & Li, B. IL-10 and IL-22 in Mucosal Immunity: Driving Protection and Pathology. *Front. Immunol.* **11**, (2020).

248. Trinchieri, G. Interleukin-10 production by effector T cells: Th1 cells show self control. *J Exp Med* **204**, 239–243 (2007).

249. Bedke, T., Muscate, F., Soukou, S., Gagliani, N. & Huber, S. IL-10-producing T cells and their dual functions. *Semin. Immunol.* **44**, 101335 (2019).

250. Li, B., Alli, R., Vogel, P. & Geiger, T. L. IL-10 modulates DSS-induced colitis through a macrophage–ROS–NO axis. *Mucosal Immunol.* **7**, 869–878 (2014).

251. Liu, L. & Li, X. NLRP3 Inflammasome in Inflammatory Bowel Disease: Friend or Foe? *Dig Dis Sci* **62**, 2211–2214 (2017).

252. Wagatsuma, K. & Nakase, H. Contradictory Effects of NLRP3 Inflammasome Regulatory Mechanisms in Colitis. *Int J Mol Sci* **21**, 1–24 (2020).

253. Zhen, Y. & Zhang, H. NLRP3 Inflammasome and Inflammatory Bowel Disease. *Front. Immunol.* **10**, (2019).

254. Davis, B. K. *et al.* Emerging Significance of NLRs in Inflammatory Bowel Disease. *Inflamm. Bowel Dis.* **20**, 2412 (2014)

Paper I

Journal of Crohn's and Colitis, 2018, 327–336
doi:10.1093/ecco-jcc/jjx139
Advance Access publication October 10, 2017
Original Article

Original Article

Transcriptomic Landscape of Treatment—Naïve Ulcerative Colitis

Hagar Taman,[a,*] Christopher G. Fenton,[a,*] Inga V. Hensel,[a,b] Endre Anderssen,[a] Jon Florholmen,[b,c] Ruth H. Paulssen[a,b]

[a]Genomic Support Centre Tromsø [GSCT], Department of Clinical Medicine, UiT—The Arctic University of Norway, Tromsø, Norway [b]Gastroenterology and Nutrition Research Group, Department of Clinical Medicine, UiT—The Artic University of Norway, Tromsø, Norway [c]University Hospital of North Norway, Tromsø, Norway

*Both authors contributed equally to this work.

Corresponding author: Ruth H. Paulssen, PhD, Department of Clinical Medicine, Gastroenterology and Nutrition Research Group, UiT—The Arctic University of Norway, Faculty of Health, Sykehusveien 38, N-9038 Tromsø, Norway. Tel.: +47 77 64 54 80; email: ruth.h.paulssen@uit.no

Abstract

Background and Aims: Ulcerative colitis [UC] is a chronic inflammatory disease that effects the gastrointestinal tract and is considered one of the most prominent and common forms of inflammatory bowel disease [IBD]. This study aimed to define and describe the entire transcriptomic landscape in a well-stratified, treatment-naïve UC patient population compared with control patients by using next-generation technology, RNA-Seq.

Methods: Mucosal biopsies from treatment-naïve UC patients [n = 14], and healthy controls [n = 16] underwent RNA-Seq. Principal component analysis [PCA], cell deconvolution methods, and diverse statistical methods were applied to obtain and characterise a dataset of significantly differentially expressed genes [DEGs].

Results: Analyses revealed 1480 significantly DEGs in treatment-naïve UC when compared with controls. Cell populations of monocytes, T cells, neutrophils, B cells/ lymphoid cells, and myeloid cells were increased during inflammation, whereas the fraction of epithelial cells were reduced in UC, which is reflected by the DEGs; 79 DEGs were identified as IBD susceptibility genes, and 58 DEGs were expressed in a gender-specific manner. MUC5B, REG3A, DEFA5, and IL33 might be considered as colorectal cancer [CRC] risk factors following UC in males. AQP9 together with CLDN2 may have a role regulating tissue-specific physiological properties in tight junctions in UC. An additional functional role for AQP9 in the synthesis and/or the function of mucus can be implied.

Conclusions: This study reveals new potential players in UC pathogenesis in general, and provides evidence for a gender-dependent pathogenesis for UC. These results can be useful for the development of personalised treatment strategies for UC in the future.

Key Words: Gender; RNA-Seq; ulcerative colitis [UC]

1. Introduction

Ulcerative colitis [UC] is a chronic inflammatory disorder of the gastrointestinal tract, and along with Crohn's disease [CD] comprise two of the common forms of inflammatory bowel disease [IBD]. In contrast to CD, the inflammations in UC are limited to the mucosa and submucosa of the colon, and the rectum.[1] Chronic inflammations

have been shown to increase the risk for the development of colorectal cancer [CRC].[2] Just 10–20% of IBD cases can be explained by genetic variances,[3,4] suggesting a much more complex pathogenesis, perhaps an underlying interplay between environmental factors, the intestinal microbiome, nutrition, and genetic factors. Susceptibility genes are responsible for triggering a dysregulation of the immune system and affecting the gut barrier function.[5,6] Nevertheless, despite decades of investigation, the complex pathogenesis of IBD is not fully understood.

Many attempts have been made to describe transcriptional levels in UC,[7–9] usually by using hybridisation-based methods like microarrays.[10] These methods are restricted to predefined and often well-annotated RNA species. Next-generation sequencing [NGS] techniques [RNA-Seq] have no such restrictions. One recent report used RNA-Seq for gene expression profiling in patients with IBD including patients with different treatment strategies.[11] For the present study, a clearly defined group of newly diagnosed [treatment-naïve] UC patients was used. It is believed that RNA-Seq, together with well-stratified UC patient material, can potentially provide a more comprehensive and correct transcriptomic profile of UC. The results of this study do not only reveal new potential players in treatment-naïve UC pathogenesis in general, but in addition provide evidence of a gender-dependent pathogenesis for UC. These results can be useful for the development of general and/or gender-dependent treatment strategies for UC in the future.

2. Material and Methods

2.1. Patient material

A standardised sampling method was used to collect mucosal biopsies from the colon of 14 newly diagnosed, treatment-naïve UC patients with mild to moderate disease activity and from 16 controls. Samples from subjects performing a cancer screening, with normal colonoscopy and normal colonic histological examination, served as controls. UC was diagnosed based upon established clinical endoscopic and histological criteria as defined by the ECCO guidelines.[12] The grade of inflammation was assessed during colonoscopy using the UC Disease Activity Index [UCDAI] endoscopic sub-score with 3 to 10 for mild to moderate disease.[13] Apart from one rectal control sample, all biopsies were taken from the sigmoid part of the colon. Tumour necrosis factor alpha [TNF-α] mRNA expression levels were measured by real-time polymerase chain reaction [PCR], thereby indicating the grade of UC activity.[14] The samples were taken from an established Biobank approved by the Norwegian Board of Health. All patient characteristics are depicted in Table 1. The participants signed an informed and written consent form. The study was approved by the Regional Ethics Committee of North Norway and the Norwegian Social Science Data Services [REK Nord 2012/1349].

Table 1. Demographic information of patient samples.

Characteristics	Control group [*n* = 16]	Ulcerative colitis [*n* = 14]
Male/female	11/5	9/5
Age mean ± SD, years	52.9 ± 16.9	39.57 ± 15.24
Endo score mean ± SD	0	1.93 ± 0.27
Clinical score ± SD	0	7.23 ± 2.45
TNF-α level ± SD	3663 ± 1973	15907 ± 9623

SD, standard deviation; TNF, tumour necrosis factor.

2.2. DNA and RNA isolation

Total RNA was isolated using the Allprep DNA/RNA Mini Kit from Qiagen [Cat. No.: 80204] and the QIAcube instrument [Qiagen], according to the manufacturer's protocol. RNA quantity and purity were assessed by using the NanoDrop ND-1000 spectrophotometer [ThermoFisher Scientific, Wilmington, DE, USA]. The Experion Automated Electrophoresis System [Bio-Rad, Hercules, CA, USA] and the RNA StdSens Analysis Kit [Bio-Rad, cat. No.: 700–7103] was used to evaluate RNA integrity, according to the instruction manual. RNA samples were kept at –70°C until further use. All RNA samples used for analyses had a RIN value between 8.0 and 10.0.

2.3. Quantitative polymerase chain reaction

The TNF-α levels in biopsies were measured by using quantitative polymerase chain reaction [qPCR]. RNA quantity was assessed with NanoVue Plus [GE Healthcare, UK]. Synthesis of cDNA was performed using the QuantiTect Reverse Transcription Kit [Qiagen, cat. No.: 205314], and the QuantiNova Probe PCR Kit [Qiagen, cat. No.: 208256]. Beta-actin [β-actin] was used as housekeeping gene. For the detection, a CFX Connect Real Time PCR Detection System [Bio-Rad, USA] was used. The results were measured in copies/µg. Values < 7000 copies/µg protein are considered as non-inflamed tissues, and values > 7000 copies/µg protein are considered inflamed tissues.[14]

2.4. Library preparation & Next generation sequencing

Transcriptome libraries were prepared with the TruSeq Stranded Total RNA LT Sample Prep Kit from Illumina [Cat. No.: RS-122–2203]. The amount of input material was 1 µg of total RNA. The Bioanalyzer 2100 [Agilent Technologies, Santa Clara, USA], and the Agilent DNA 1000 kit [Cat. No.: 5067-1504] were used to assess RNA libraries quality, according to the instruction manual. The RNA libraries comprised fragments with an average size of 307 base pairs. The libraries were normalised to 10 nM and subsequently sequenced with the NextSeq 550 instrument [Illumina, USA] according to the manufacturer's instructions. The average number of uniquely mapped reads per sequencing run was 88 million reads per sample.

2.5. Data analysis

Base calling and quality scoring were performed as a first step including quality check on the on-board computer of the NextSeq 550. The algorithm packages STAR-2.5.2b and the htseq-count were used for downstream analysis.[15] Transcripts were aligned to human genome assembly GRCH38p.11 [https://www.ncbi.nlm.nih.gov/grc/human/data]. Read counts were transformed using the DESeq2-Rlog variance stabilised transform, and significantly differentially expressed transcripts were identified by including transcripts with a read count of > 30 and fold change > 2 as compared with controls. Additional annotation was added using the PANTHER classification system [http://pantherdb.org/], the Kyoto Encyclopaedia of Genes and Genomes [KEGG] [www.genome.jp/kegg/], and the human gene database GeneCards [http://www.genecards.org/]. For principal component analysis [PCA], the top 5000 most variable of the DESeq2-Rlog variance stabilised genes were used. For the estimation of specific cell populations in patient samples, all DESeq2-Rlog normalised transcripts with a log2 average mean > 5 were included. The analysis was performed using the R/Bioconductor CellMix manual [http://web.cbio.uct.ac.za/~renaud/CRAN/web/CellMix/] with the

IRIS [Immune Response In Silico] weighted marker list characteristic for the different cell types.[16] Epithelial markers cadherin 1 [CDH1], epithelial cell adhesion molecule [EPCAM], phosphatidylinositol glycan anchor biosynthesis class F [PIGF], L1 cell adhesion molecule [L1CAM], and laminin subunit alpha 1 [LAMA1] were added to the IRIS marker list and weighted strongly to give an estimate of epithelial presence in samples. The contrast matrix [[N_{female} - UC_{female}] - [N_{male} - UC_{male}]] was used in DESeq2 to determine differentially regulated genes between UC and controls which differed significantly between male and female. Results were limited to adjusted p-value < 0.05 and a log2 fold change > 1.0. Genes associated with the risk of IBD were downloaded from the genome-wide association studies [GWAS] catalogue using the search term IBD on 29 November 2016.[17] Gene expression data for the 295 genes associated with IBD single nucleotide polymorphism [SNPs] was analysed by k-means clustering and three primary groups were identified. For interpretation, genes were further subdivided using the gene ontology according to roles in inflammatory processes. The inflammatory properties of the genes were classified with the gene ontology as part of an inflammatory response [GO: 0006954] with subcategories for negative [GO: 0050728] and positive [GO: 0050729] regulators of inflammation. Heat maps were produced for each cluster to visualise the gene expression patterns of the GWAS associated genes in conjunction with their regulatory and inflammatory roles.

3. Results

3.1. Patients

A standardised sampling method was used to collect mucosal biopsies from treatment-naïve UC patients [n = 14] and control samples [n = 16], as described above. The biopsies from UC patients showed mild to moderate disease activity [as defined by UCDAI], with clinical scores 7.23 ± standard deviation [SD] 2.45, and endo scores of 1.93 ± SD 0.27, estimated according to established clinical endoscopic and histological criteria, and as defined by the ECCO guidelines.[12] The control group consisted of biopsies with normal colonoscopy, colon histology, and immunohistochemistry, and clinical and endo scores = 0. The biopsies of UC patients were taken from the sigmoid part of the colon. The gender distribution for both groups was almost equal, with nine males in the UC group and 11 males in the control group, and five females in each group. The age distribution differed between the two groups, with 39.57 ± SD 15.24 in the UC group, and 52.9 ± SD 16.9 in the control group. In order to obtain information about the inflammatory status of UC, TNF-α mRNA expression levels were measured by qPCR.[14] Levels of TNF-α in control samples were estimated at 3663 ± SD 1973, and for UC samples 15907 ± SD 9623. A summary of all patient characteristics is depicted in Table 1.

3.2. Characterization of the whole transcriptome in treatment-naïve UC

The entire transcriptome representing treatment-naïve UC was established by RNA-Seq. Pre-processing of the sequencing data revealed expression of approximately 22 000 transcripts. Initial principal component analysis [PCA] of the 5000 most variable transcripts revealed a clear distinction between UC and control samples along the first principal component with a 59.6% explained variance [Figure 1].

In order to estimate specific cell populations in UC and control tissue samples, a cell deconvolution method was applied as described. The deconvolutions were restricted to the following cell types: epithelial cells, monocytes, T cells, neutrophils, B cells/lymphoid cells, and myeloid cells. The results show a clear difference of cell fractions present in UC and control samples. An enrichment of monocytes, neutrophils, myeloid cells, T cells, and B/Lymphoid cells was observed in all UC samples, whereas the epithelial cell fraction was decreased in almost all UC samples when compared with control samples. The results of the deconvolution experiments are depicted in Figure 2.

To further describe and analyse the entire transcriptome, significantly differentially expressed transcripts were adjusted to p-value < 0.05 and a cut-off of log2 fold-change > 1.0 [n = 1480] was used for downstream analyses [Supplementary Data 1, available at *ECCO-JCC* online] whereof the top 30 differentially expressed genes [DEGs] with log2 fold-change > 3.5 are shown in Table 2. The differentially expressed gene transcripts were related to currently known 295 IBD susceptibility genes [Supplementary Figure 1, available at *ECCO-JCC* online] as revealed by GWAS [see Figure 3].[18,19] The identified gene transcripts have been annotated to different inflammatory processes and their transcriptional levels. Transcripts with unchanged expression are omitted in Figure 3. However, for a complete overview see Supplementary Figure 1. The data depicted in Figure 3 show 71 upregulated [24%] and eight downregulated transcripts [2.7%] linked with susceptibility to IBD. The annotation of the upregulated transcripts revealed genes involved in inflammatory responses, like chemokine receptors [CCR5, CCR3, CXCR1, CXCR2], chemokine ligands [CXCL5, CXCL1, CXCL2, CXCL3, CXCL4 [PF4], and CCL20], tumour necrosis factor receptor superfamily members (TNFRSF5 [CD40] and TNFRSF9), interleukin 19 [IL19], solute carrier family 11 member 1 [SLC11A1], intercellular adhesion molecule 1 [ICAM1], and signal transducer and activator of transcription 3 [STAT3], T cell specific antigens [CD28 and CD6], chemokine ligands [CCL2 and CCL11], oncostatin M [OSM] and its receptor oncostatin M receptor [OSMR], inflammatory bowel disease Protein 1 (NOD2 [IBD1]), and cyclooxygenase-2 (COX2 [PTGS2]). Only one of the downregulated transcripts is involved in inflammatory response, epoxide hydrolase 2 [EPHX2]. In addition,

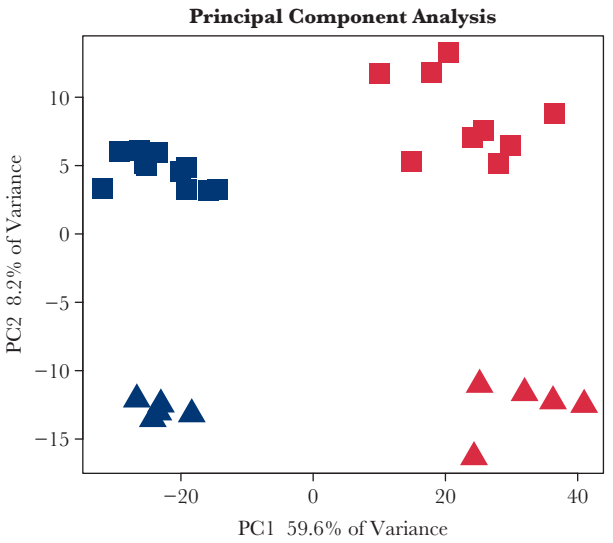


Figure 1. Principal component analysis [PCA]. Unsupervised PCA analysis showing the difference between UC [red] and control [blue] as well as gender, male control [blue square], female control [blue triangle], male UC [red square] and female UC [red triangle]. There is a 59.6% variance between UC and control samples, and an 8.2% variance between male and female samples. UC, ulcerative colitis.

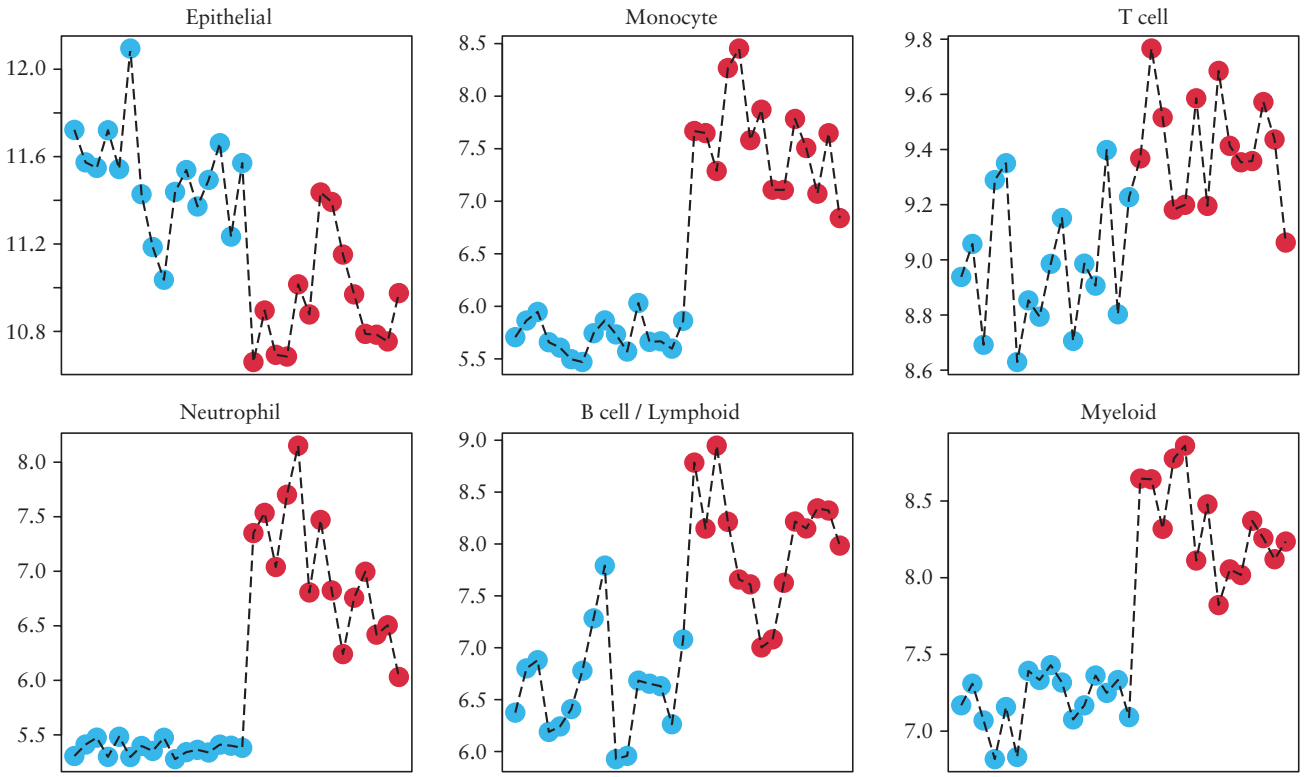


Figure 2. Estimation of cell population between samples using cell deconvolution methods using the Bioconductor CellMix package. Epithelial markers [CDH1, EPCAM, L1CAM, and LAMA1] were added to the Immune Response In Silico [IRIS] marker list and weighted heavily to help estimate epithelial contribution. The blue dots indicate control samples and red dots indicate ulcerative colitis [UC] samples, respectively.

three significantly differentially expressed microRNAs [MIR155HG, MIR3936 and MIR4435-2HG] and 17 long non-coding RNAs [lncRNAs] have been identified [Supplementary Data 1].

The current study was compared with microarray- based studies carried out on two different Affymetrix-based microarray platforms.[20,21] Only well-characterised probes with Entrez IDs and official gene symbols were included in the comparison. Overall agreement between the studies when comparing UC patients and normal controls are shown in the Venn diagram [Supplementary Figure 2, available at *ECCO-JCC* online]. Only one gene, transmembrane 4 L6 family member 20 [TM4SF20], was found to be significantly [log fold-change > 1 and $p < 0.05$] downregulated in our study but significantly upregulated in one of the others.[21]

3.3. Gender-specific transcription

PCA analysis did not only distinguished between UC and control samples [Figure 1] but could also distinguish transcripts to a lesser extent [variance 8.2%] and in a gender- specific manner. The contrast matrix [[N_{female} - UC_{female}] - [N_{male} - UC_{male}]] was applied to determine differentially regulated genes with gender-specific effects; 58 significantly differentially expressed genes with adjusted p-value < 0.05 and log 2 FC > 1.0 were identified [see Supplementary Data 2, available at *ECCO-JCC* online]. A boxplot of eight selected genes show gender-specific differences [Figure 4: Paneth cell specific defensin alpha 5 [DEFA5], serine protease 8 [PRSS8], mucin 5B [MUC5B], phospholipase A2 group IIA [PLA2G2A], Fc fragment of IgG binding protein [FCGBP], regenerating family member 3 alpha [REG3A], interleukin 33 [IL33], and interferon alpha inducible protein 27 [IFI27]. The expression levels of the selected genes are higher in the UC group compared with the control group. However, the expression levels do not only show a difference between UC samples and control samples, but also show significant differences between UC samples and control samples for males and females. It is noted that thiosulphate sulphurtransferase [TST], mercaptopyruvate sulphurtransferase [MPST], and fucosyltransferase 2 [FUT2] are the only GWAS IBD susceptibility genes that are found to be expressed in a gender-specific manner [Supplementary Data 2; Figure 3].

4. Discussion

This study provides a unique, comprehensive, and quantitative record of high-resolution gene expression in a treatment-naïve UC patient population using next-generation RNA-Seq technology. Previous gene expression studies with UC patient material mostly used hybridisation-based methodologies like microarrays,[7,22–25] and only one recent study reported RNA-Seq of human IBD patient material.[11] RNA-Seq technology has some advantages over microarray technology, such as the ability of impartial detection of new transcripts and the easy detection of rare and low abundance transcripts.[10] In addition, RNA-Seq does not rely on pre-designed complement-sequence detection probes, and is therefore free of issues associated with probe redundancy and annotation.[26] Attempts were made in order to decipher if genes found in other studies[20,21] including UC patients behaved differently in the present study [Supplementary Figure 2, available at *ECCO-JCC* online]. However, it is difficult to assign the variability between experiments to either effects of the technical platform or different severity of disease that was mild to moderate in our dataset and involved patients resistant to standard treatment in the above-mentioned studies.

Table 2. Top 30 differentially expressed genes [DEGs] in treatment-naïve ulcerative colitis [UC].

Gene symbol	Log2 FC	*p*-adjusted
ABCA12	4,46	4,98E-35
AQ9	5,37	3,56E-32
CHI3L1	5,36	8,62E-57
CLDN2	4,49	1,82E-38
CXCL1	4,30	7,06E-34
CXCL5	4,63	7,17E-20
CXCL6	4,06	5,46E-26
CXCR1	4,64	3,93E-23
DEFA5	4,09	1,1E-15
DUOX2	5,73	8,62E-57
DUOXA2	7,21	1,93E-77
FAM83A	4,11	2,94E-22
FCGR3B	4,55	6,03E-30
FPR2	4,23	1,78E-21
HCAR3	4,32	5,01E-22
IL17A	5,03	1,5E-33
KCNJ15	3,95	1,22E-17
LCN2	4,95	2,25E-53
MMP10	4,99	1,44E-34
MMP3	5,02	3,31E-44
MMP7	5,32	2,15E-29
NOS2	4,03	9,9E-40
PI3	4,11	2,73E-31
REG1A	5,05	1,28E-22
REG3A	4,19	2,63E-15
SAA1	6,40	5,16E-46
SAA2	6,23	7,31E-43
SLC6A14	5,66	1,73E-48
TNIP3	5,60	3,49E-58
TREM1	4,13	8,56E-26

In addition, recent UC transcriptome studies were performed by using non-stratified patient populations and material including both treated and treatment-naïve UC,[11] which might have resulted in biased gene expression profiles. Therefore, in this study, the UC patient material was thoroughly stratified and only treatment-naïve patients were included [see Table 1 and Material and Methods]. This approach should provide an opportunity to investigate the transcriptional profile of UC without the interference of any medications. Recent reports have shown that medication given to UC patients, such as immunosuppressant drugs, have short- and long-term side effects on the immune response.[27] This treatment might introduce unwanted bias to experiments aiming to investigate the prior to medication status of UC patients. Furthermore, treatment-naïve UC transcriptomic signatures might also become important in order to decide individual treatment options for patients in the future. Our RNA-Seq revealed 1480 significantly DEGs [see Supplementary Data 1]. In the future, these DEGs could be used as a fingerprint for disease outcome and a beneficial treatment choice.

Human tissue samples are highly heterogeneous, and the amount of different cell types in a biopsy will have a certain impact on gene expression profiles. Therefore, and in order to determine the relative quantities of different immune cell types in the heterogeneous biopsies, cell deconvolution methods were applied [Figure 2]. As expected, the results clearly show and confirm that the innate and adaptive immune systems are triggered by inflammation as the fractions of monocytes, neutrophils, myeloid cells, T cells, and B/lymphoid cells are found to be increased in UC [Figure 2]. These differences could very well due to inflammatory infiltrates; as for which cells contribute the most, we will need a much larger sample size and most likely confirmatory or parallel experiments using a different technique such as single cell sequencing. One interesting aspect with these data is that one could potentially identify various patterns. For example, UC patient no. 5 from the left [Figure 2] is rather high in monocytes and neutrophils but relatively low in T cells and B/lymphoid cells, but to decipher these types of patterns would require larger study populations. In concordance with these observations, the different inflammatory responses are reflected by increased expression of genes like leucocyte immunoglobulin-like receptors, cytokines/chemokines and their respective receptors and ligands, and T cell specific antigens [see Table 2; and Supplementary Data 1]. In particular, many of the DEGs that are involved in the control of bacterial proliferation showed increased expression during mucosal inflammation, like regenerating family members [REGs] and defensins, which is in concordance with previous studies.[7,28–32] It is interesting to note that the here observed increased expression of factors predictive of response to anti-TNF-α therapy, oncostatin [OSM] and its receptor [OSMR], is of particular relevance for anti-TNF-α resistant patients.[33]

Not surprisingly, the fraction of epithelial cells was lowered in UC [Figure 2]. Once the mucosal epithelium is compromised by inflammation, the fraction of functional epithelial cells diminishes, which then leads to a 'leaky' intestinal epithelium.[34] The impaired ion transport and dysfunctional tight junctions in the epithelium are followed by chronic diarrhoea. In addition, increased levels of circulating TNF-α, and other cytokines lead to a rise of intestinal permeability, thereby causing oedema.[35] This dysregulation of water and solute homeostasis has been suggested to play a role in UC. In concordance with our results, the expression levels of water channels like aquaporin 7 [AQP7] and aquaporin 8 [AQP8], have been shown to be reduced in the human intestinal mucosa in early stage IBD.[36] Interestingly, aquaporin 9 [AQP9] was one of the most prominent expressed genes in our UC patient material, which has not been reported before [Table 2]. AQP9 plays a role in specialised leukocyte functions such as immunological response and bactericidal activity, and is located in the membranes of tight junctions in the intestine.[37] In addition claudins, that are exclusively localised at tight junctions, were differentially expressed [Table 2], with claudin 2 [CLDN2] levels being increased in UC samples, which is in accordance with former results.[38–40] The 'pore-forming'CLDN2, as a component of the tight junction, forms a water channel and thus mediates para-cellular water transport across the tight junctions in impaired epithelium.[41] In addition, the observed increase of CLDN2 correlates with UC severity on both protein and transcriptional levels.[38] This might indicate that CLND2 and AQP9 might share similar properties regarding water transport. It is possible that like epithelial cells, other cells contribute to the elevated levels of CLDN and/or AQP9. However, no such information is available today. Cell sorting followed by single cell sequencing might shed a light on this. The increased expres-sion of CLDN and AQP9 might be a response to inflammation, fighting the disturbances in the epithelial barrier of the colon. Taken these results together, it could be hypothesised that CLDN2 together with AQP9 regulate tissue-specific pathophysiological properties of tight junctions in UC.

Intestinal infection generally leads to depletion of goblet cells and reduction in mucin synthesis and secretion, allowing pathogens to access the underlying epithelium.[42] Interestingly, AQP9 could play another role for the synthesis and/or function of mucus, as recent immuno-histological studies have shown that AQP9 is expressed in a subset of mucin-producing goblet cells in the small intestine

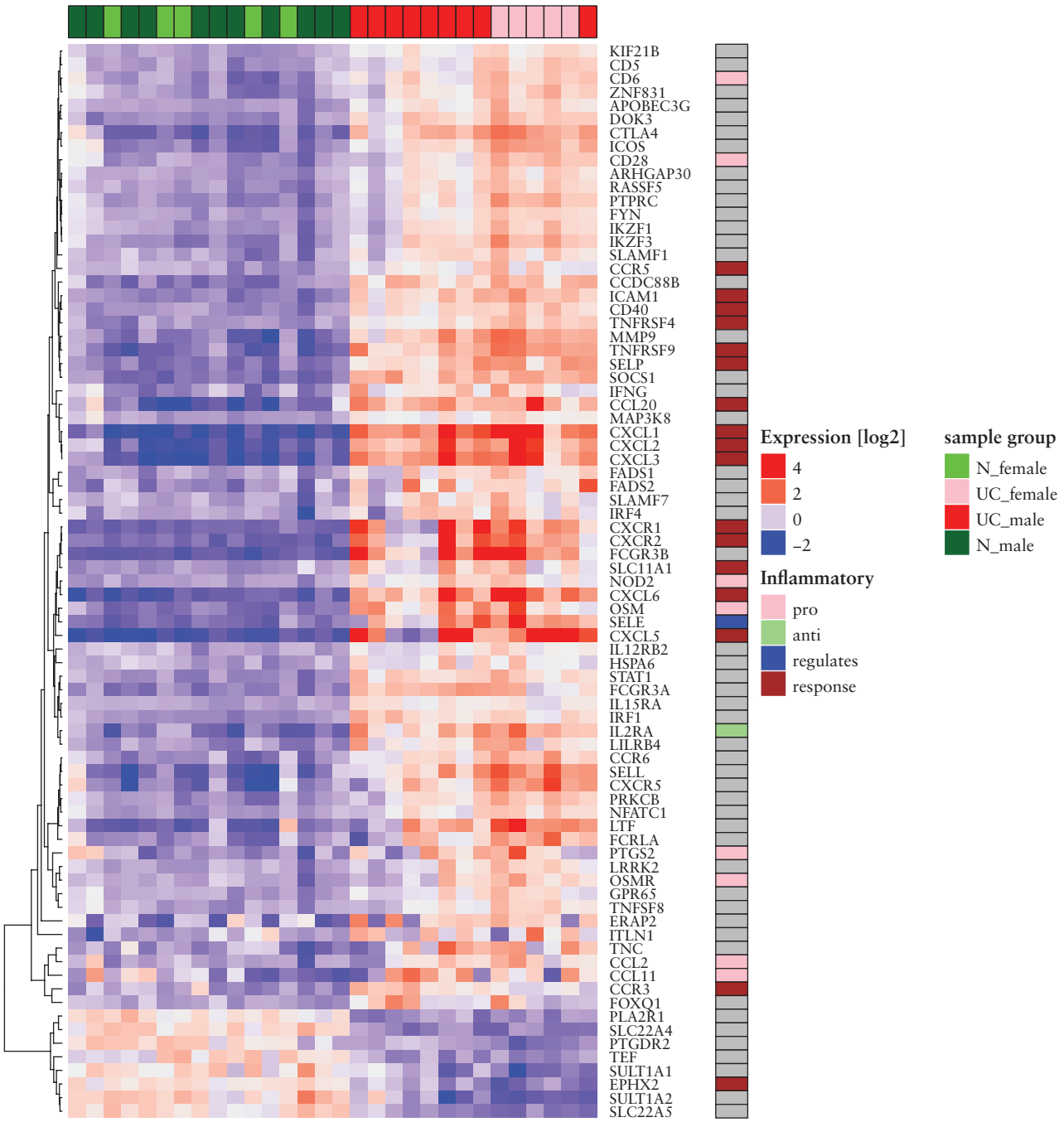


Figure 3. Gene expression pattern across controls and UC patients of differentially expressed genes that are also reported as associated with GWAS risk SNPs. Colours indicate deviation from the mean expression level for each gene. Samples are colour-coded in the top bar according to gender and diagnosis [control, UC] and genes are colour-coded in the sidebar according to their inflammation-related gene ontology annotations. Transcripts with unchanged differential expression are omitted. UC, ulcerative colitis; GWAS, genome-wide association studies; SNPs. Single nucleotide polymorphisms.

and colon.[37] However, mucin 2 [MUC2], which is the major contributor to healthy lubrication of the mucosa, was not differentially expressed, which is in concordance with previous findings.[43] In contrast, several other mucins showed increased expression in UC: mucin 1 [MUC1], mucin 16 [MUC16], mucin 4 [MUC4], mucin 5Ac [MUC5AC], and mucin 5B [MUC5B] [Table 2; and Supplementary Data 1]. Elevated levels of MUC1 and MUC4 in UC have been demonstrated before.[44,45] Contrary to results reported by others who have shown that expression of MUC5B is increased in colorectal cancer [CRC] but not in UC,[46] we here report increased expression of MUC5B in UC. It is hereby noted that in normal colon, MUC5B has been shown to be secreted by colonic goblet cells; however it is expressed in minor quantities.[47] In addition, we can report not only an increased but also gender-specific differences in the expression of MUC5B in UC, as discussed below [see also Figure 4; and Supplementary Data 2]. A cytokine-induced mucin hypersecretion has been reported for MUC5AC in an *in vitro* model where expression increased in a TNF-α dose-dependent manner.[48] This might be also the situation in our UC material, as elevated TNF-α levels have been one criterion for patient stratification [see Materials

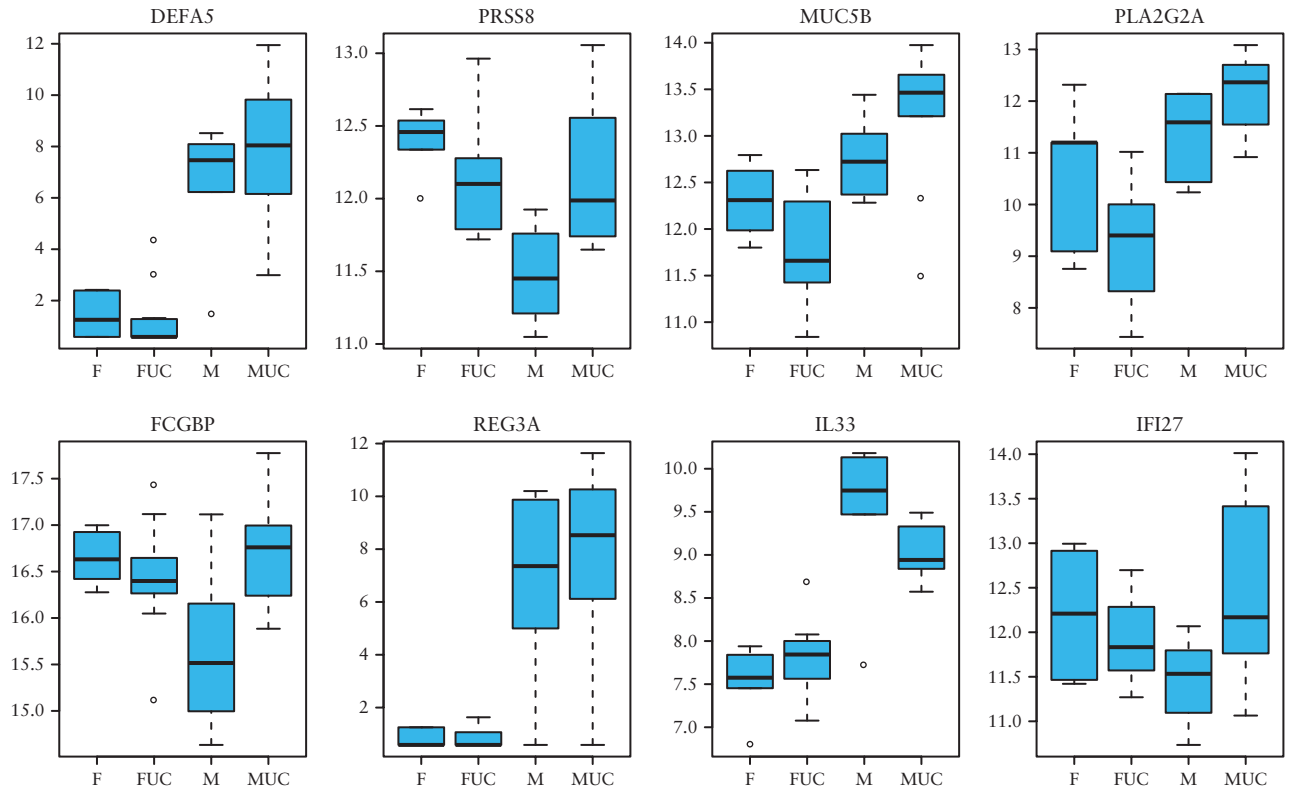


Figure 4. A boxplot of selected genes that show gender-specific differences within a greater subset of UC versus control significant differentially expressed genes. F and M indicate control female and male samples, FUC and MUC indicate female and male UC samples.

and Methods]. A role for MUC16 in UC, which showed the most prominent elevated expression levels [log2 FC 2.16] has not been established yet but has been recently proposed to be a biomarker for epithelial ovarian cancer.[49]

Recently, micro RNAs [miRNAs] and long non-coding RNAs [lncRNAs] have been addressed as having a role in IBD pathogenesis.[50,51] Three microRNAs have been identified in our data: MIR155HG, MIR3936, and MIR4435-2HG. It has been reported that MIR155 is involved in intestinal inflammation and immunity of UC by playing a role in the differentiation of B and T cells and dendritic cells, thereby contributing to the development of regulatory T cells.[52–54] In addition, we observed 17 significantly differentially expressed lncRNAs [Supplementary Data 1]. One functional relationship between UC and a particular lncRNA has been recently demonstrated.[55] However, the relationship between UC and miRNAs and/or lncRNAs still remains unclear.

In order to characterise gene expression data from RNA-Seq further, one approach was to investigate how many differentially expressed genes are associated with IBD. Until today, approximately 300 SNPs associated with IBD have been discovered through GWAS.[17] Although it is expected that this RNA-Seq study cannot confirm GWAS findings, 79 of the significantly differentially expressed genes might be associated with IBD [see Table 2]. It is hereby noted that the GWAS susceptibility genes did not distinguish between UC and CD. However, although one should be cautious about assigning specificity to SNPs, since most GWAS studies are done comparing a patient population with a population of disease-free controls and not between subtypes of the disease, some SNPs are indeed more strongly related to UC than CD and vice versa.[56] Among our significantly differentially expressed genes, SP140 nuclear body protein [SP140] and strawberry notch homolog 2 [SBNO2] are located close to SNPs that are more strongly related to CD risk, and both of these genes have shown upregulated gene expression in UC [Supplementary Data 1].

Recently, gender-dependent differences in IBD pathologies have been proposed.[57,58] Sex-stratified analysis of long-term complications of IBD show consistently higher risk of CRC in male IBD patients.[59] In addition, a recent population study reported that patients with UC are the high-risk group in incidence of CRC and that the risk is found to be higher in male than in female UC patients.[60] The current available information about gender-specific differences in UC is sparse and contradictory to some extent.[61] It is a common belief that understanding gender differences in any disease is important for recognising the factors contributing to the disease expression and to determine its prognosis so that clinicians can offer an appropriate medical therapy. However, molecular manifestations of a gender specificity for UC has not been established. We can for the first time show that gender differences on a molecular level occur not only between treatment-naïve UC patients [Figure 1], but also between control patients [Figure 4]. The most pronounced gender differences for UC were observed for mucin 5B [MUC5B], regenerating family member 3 alpha [REG3A], defensin A5 [DEFA5], and interleukin 33 [IL33] [Figure 4]. In UC, IL33 expression is specifically increased and has been shown to be involved in the inflammatory tumour microenvironment and to contribute to the progression of CRC.[62–64] In this context, it is interesting to note that IL33 is found to be more increased in male UC than in female UC. In addition, SNPs in human DEF5A may confer susceptibility to IBD.[32] Furthermore, an upregulation of REG3A in colorectal cancer cells confers proliferation and correlates with colorectal cancer risk.[65] All these above-mentioned genes were more elevated in male samples than in female samples and can therefore be considered as CRC risk factors in UC.

It is interesting to note that DEFA5, IL33, and REG3A showed also increased expression levels in control male samples when compared with control female samples. Taken all these observations together, a possible link between the expression of these genes and higher risk for the development of CRC following UC can be proposed for males. For some gender-specific genes, differential gene expression was observed between control male and female samples. Genes like serine protease 8 [PRSS8], Fc fragment of IgG binding protein [FCGBP], and interferon alpha-inducible protein 27 [IFAN27] showed increased expression in control female compared with control male samples. Interestingly, all these genes have been shown to convey protective properties in the mucosa: FCGBP has been shown to be involved in the maintenance of the mucosal structure as a gel-like component of the mucosa; PRSS8 preserves colonic integrity and protects against inflammation, as has been demonstrated in DSS-induced inflammation of mice; and IFN-α conveys antiviral activities.[66–68] In addition, three gender-specific GWAS IBD susceptibility genes showed increased expression in female UC [Supplementary Data 2]. Two belonged to the family of sulphurtransferases, and both genes are involved in detoxification processes.[69] The here observed expression of thiosulphate [TST]—and mercaptopyruvate [MPST] —sulphurtransferase is essential for sulphide detoxification in order to preserve healthy mucosa. Dysregulation of expression and/or activity of these enzymes may accompany development of UC.[70] Here, fucosyltransferase 2 [FUT2] also might play a role in this regard, since this enzyme is involved in host-microbe interactions and has been shown to mediate interaction with intestinal microbiota, thereby influencing its composition.[71–73]

In conclusion, this study shows for the first time that the use of well-stratified treatment-naïve UC patient samples in combination with high-throughput RNA-Seq technology can reveal new molecular players that might be important in UC pathogenesis. Potentially significant might be the regulation of tissue-specific pathophysiological properties of tight junctions in the mucosa as reflected by increased expression of AQP9 and CLDN2 and the expression of different mucins, particularly MUC5B and MUC16. In addition, a gender-dependent molecular manifestation could be established. The molecular patterns of UC revealed increased expression of genes involved in preserving mucosal integrity and detoxification of microbial-derived metabolites in females. The expression of antimicrobial and cytotoxic genes in male UC patients may contribute to the higher risk for the development of CRC observed in males. These results can be useful for the development of new treatment and patient stratification strategies for UC. In addition, these expression patterns can be extremely useful if combined with UC remission data in the future.

Funding

This work was supported by the Northern Norway Regional Health Authority Helse-Nord [SPF-1209–14].

Conflict of Interest

The authors declare no conflict of interests regarding the publication of this paper.

Acknowledgments

The authors thank Rasmus Goll and Mona Dixon Gundersen for providing clinical evaluation of some patients, Ingrid Christiansen for the technical help performing the TNF-α levels measurements, and Odd-Sverre Moen for administrating the patient samples.

Author Contributions

HT performed most of the experiments and wrote parts of the manuscript. CGF performed most statistical analyses and revised the manuscript. IVH performed a part of the experiments and revised the manuscript. EA performed a part of the data analysis and revised the manuscript. JF was involved in evaluating and providing clinical samples from patients and healthy controls and revised the manuscript. RHP was involved in project inception, design, supervision, and manuscript writing and revision.

Supplementary Data

Supplementary data are available at *ECCO-JCC* online.

References

1. Khor B, Gardet A, Xavier RJ. Genetics and pathogenesis of inflammatory bowel disease. *Nature* 2011;**474**:307–17.
2. Planell N, Lozano JJ, Mora-Buch R, *et al.* Transcriptional analysis of the intestinal mucosa of patients with ulcerative colitis in remission reveals lasting epithelial cell alterations. *Gut* 2013;**62**:967–76.
3. Jostins L, Ripke S, Weersma RK, *et al.* Host-microbe interactions have shaped the genetic architecture of inflammatory bowel disease. *Nature* 2012;**49**:119–24.
4. Liu JZ, van Sommeren S, Huang H, *et al.*; International Multiple Sclerosis Genetics Consortium; International IBD Genetics Consortium. Association analyses identify 38 susceptibility loci for inflammatory bowel disease and highlight shared genetic risk across populations. *Nat Genet* 2015;**47**:979–86.
5. Loddo I, Romano C. Inflammatory bowel disease: genetics, epigenetics, and pathogenesis. *Front Immunol* 2015;**6**:551.
6. Xavier RJ, Podolsky DK. Unravelling the pathogenesis of inflammatory bowel disease. *Nature* 2007;**448**:427–34.
7. Lawrance IC, Fiocchi C, Chakravarti S. Ulcerative colitis and Crohn's disease: distinctive gene expression profiles and novel susceptibility candidate genes. *Hum Mol Genet* 2001;**10**:445–56.
8. Burczynski ME, Peterson RL, Twine NC, *et al.* Molecular classification of Crohn's disease and ulcerative colitis patients using transcriptional profiles in peripheral blood mononuclear cells. *J Mol Diagn* 2006;**8**:51–61.
9. Watanabe T, Kobunai T, Yamamoto Y, *et al.* Predicting ulcerative colitis-associated colorectal cancer using reverse-transcription polymerase chain reaction analysis. *Clin Colorectal Cancer* 2011;**10**:134–41.
10. Wang Z, Gerstein M, Snyder M. RNA-Seq: a revolutionary tool for transcriptomics. *Nat Rev Genet* 2009;**10**:57–63.
11. Holgersen K, Kutlu B, Fox B, *et al.* High-resolution gene expression profiling using RNA sequencing in patients with inflammatory bowel disease and in mouse models of colitis. *J Crohns Colitis* 2015;**9**:492–506.
12. Magro F, Gionchetti P, Eliakim R, *et al.* Third European evidence-based consensus on diagnosis and management of ulcerative colitis. Part 1: definitions, diagnosis, extra-intestinal manifestations, pregnancy, cancer surveillance, surgery, and ileo-anal pouch disorders. *J Crohn's Colitis* 2017;**11**:649–70.
13. Mcnally PR. Budesonide for the induction of remission for mild- moderate ulcerative colitis. *Visible Hum J Endoscpy* 2014;**13**:2–7.
14. Olsen T, Goll R, Cui G, *et al.* Tissue levels of tumor necrosis factor-alpha correlate with grade of inflammation in untreated ulcerative colitis. *Scand J Gastroenterol* 2007;**42**:1312–20.
15. Dobin A, Gingeras TR. Mapping RNA-seq Reads with STAR. *Curr Protoc Bioinformatics* 2015;**51**:11.14.1–19.
16. Abbas AR, Baldwin D, Ma Y, *et al.* Immune response in silico [IRIS]: immune-specific genes identified from a compendium of microarray expression data. *Genes Immun* 2005;**6**:319–31.
17. Welter D, MacArthur J, Morales J, *et al.* The NHGRI GWAS Catalog, a curated resource of SNP-trait associations. *Nucleic Acids Res* 2014;**42**:D1001–6.
18. 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**:246–52.

19. de Lange KM, Moutsianas L, Lee JC, *et al.* Genome-wide association study implicates immune activation of multiple integrin genes in inflammatory bowel disease. *Nat Genet* 2017;**49**:256–61.
20. Arijs I, De Hertogh G, Lemaire K, *et al.* Mucosal gene expression of antimicrobial peptides in inflammatory bowel disease before and after first infliximab treatment. *PLoS One* 2009;**4**:e7984.
21. Vanhove W, Peeters PM, Staelens D, *et al.* Strong upregulation of AIM2 and IFI16 inflammasomes in the mucosa of patients with active inflammatory bowel disease. *Inflamm Bowel Dis* 2015;**21**:2673–82.
22. Flach CF, Eriksson A, Jennische E, Lange S, Gunnerek C, Lönnroth I. Detection of elafin as a candidate biomarker for ulcerative colitis by whole-genome microarray screening. *Inflamm Bowel Dis* 2006;**12**:837–42.
23. Noble CL, Abbas AR, Cornelius J, *et al.* Regional variation in gene expression in the healthy colon is dysregulated in ulcerative colitis. *Gut* 2008;**57**:1398–405.
24. Watanabe T, Kobunai T, Toda E, *et al.* Gene expression signature and the prediction of ulcerative colitis-associated colorectal cancer by DNA microarray. *Clin Cancer Res* 2007;**13**:415–20.
25. Olsen J, Gerds TA, Seidelin JB, *et al.* Diagnosis of ulcerative colitis before onset of inflammation by multivariate modeling of genome-wide gene expression data. *Inflamm Bowel Dis* 2009;**15**:1032–8.
26. Zhao S, Fung-Leung WP, Bittner A, Ngo K, Liu X. Comparison of RNA-Seq and microarray in transcriptome profiling of activated T cells. *PLoS One* 2014;**9**:e78644.
27. D'Haens G. Risks and benefits of biologic therapy for inflammatory bowel diseases. *Gut* 2007;**56**:725–32.
28. Lai Y, Li D, Li C, *et al.* The antimicrobial protein REG3A regulates keratinocyte proliferation and differentiation after skin injury. *Immunity* 2012;**37**:74–84.
29. van Beelen Granlund A, Østvik AE, Brenna Ø, Torp SH, Gustafsson BI, Sandvik AK. REG gene expression in inflamed and healthy colon mucosa explored by in situ hybridisation. *Cell Tissue Res* 2013;**352**:639–46.
30. Wu F, Dassopoulos T, Cope L, *et al.* Genome-wide gene expression differences in Crohn's disease and ulcerative colitis from endoscopic pinch biopsies: insights into distinctive pathogenesis. *Inflamm Bowel Dis* 2007;**13**:807–21.
31. Ufer M, Häsler R, Jacobs G, *et al.* Decreased sigmoidal ABCB1 [P-glycoprotein] expression in ulcerative colitis is associated with disease activity. *Pharmacogenomics* 2009;**10**:1941–53.
32. Ferguson LR, Browning BL, Huebner C, *et al.* Single nucleotide polymorphisms in human Paneth cell defensin A5 may confer susceptibility to inflammatory bowel disease in a New Zealand Caucasian population. *Dig Liver Dis* 2008;**40**:723–30.
33. West NR, Hegazy AN, Owens BMJ, *et al.* Oncostatin M drives intestinal inflammation and predicts response to tumor necrosis factor-neutralizing therapy in patients with inflammatory bowel disease. *Nat Med* 2017;**23**:579–89.
34. Mankertz J, Schulzke JD. Altered permeability in inflammatory bowel disease: pathophysiology and clinical implications. *Curr Opin Gastroenterol* 2007;**23**:379–83.
35. He F, Peng J, Deng XL, *et al.* Mechanisms of tumor necrosis factor-alpha-induced leaks in intestine epithelial barrier. *Cytokine* 2012;**59**:264–72.
36. Ricanek P, Lunde LK, Frye SA, *et al.* Reduced expression of aquaporins in human intestinal mucosa in early stage inflammatory bowel disease. *Clin Exp Gastroenterol* 2015;**8**:49–67.
37. Okada S, Misaka T, Matsumoto I, Watanabe H, Abe K. Aquaporin-9 is expressed in a mucus-secreting goblet cell subset in the small intestine. *FEBS Lett* 2003;**540**:157–62.
38. Landy J, Ronde E, English N, *et al.* Tight junctions in inflammatory bowel diseases and inflammatory bowel disease associated colorectal cancer. *World J Gastroenterol* 2016;**22**:3117–26.
39. Prasad S, Mingrino R, Kaukinen K, *et al.* Inflammatory processes have differential effects on claudins 2, 3 and 4 in colonic epithelial cells. *Lab Invest* 2005;**85**:1139–62.
40. Oshima T, Miwa H, Joh T. Changes in the expression of claudins in active ulcerative colitis. *J Gastroenterol Hepatol* 2008;**23**[Suppl 2]:S146–50.
41. Rosenthal R, Milatz S, Krug SM, *et al.* Claudin-2, a component of the tight junction, forms a paracellular water channel. *J Cell Sci* 2010;**123**:1913–21.
42. Dorofeyev AE, Vasilenko IV, Rassokhina OA, Kondratiuk RB. Mucosal barrier in ulcerative colitis and Crohn's disease. *Gastroenterol Res Pract* 2013;**2013**:431231.
43. Allen A, Hutton DA, Pearson JP. The MUC2 gene product: a human intestinal mucin. *Int J Biochem Cell Biol* 1998;**30**:797–801.
44. Kim YS, Ho SB. Intestinal goblet cells and mucins in health and disease: recent insights and progress. *Curr Gastroenterol Rep* 2010;**12**:319–30.
45. Shirazi T, Longman RJ, Corfield AP, Probert CS. Mucins and inflammatory bowel disease. *Postgrad Med J* 2000;**76**:473–8.
46. Walsh MD, Clendenning M, Williamson E, *et al.* Expression of MUC2, MUC5AC, MUC5B, and MUC6 mucins in colorectal cancers and their association with the CpG island methylator phenotype. *Mod Pathol* 2013;**26**:1642–56.
47. Einerhand AW, Renes IB, Makkink MK, van der Sluis M, Büller HA, Dekker J. Role of mucins in inflammatory bowel disease: important lessons from experimental models. *Eur J Gastroenterol Hepatol* 2002;**14**:757–65.
48. Smirnova MG, Birchall JP, Pearson JP. TNF-alpha in the regulation of MUC5AC secretion: some aspects of cytokine-induced mucin hypersecretion on the in vitro model. *Cytokine* 2000;**12**:1732–6.
49. Felder M, Kapur A, Gonzalez-Bosquet J, *et al.* MUC16 [CA125]: tumor biomarker to cancer therapy, a work in progress. *Mol Cancer* 2014;**13**:129.
50. Chapman CG, Pekow J. The emerging role of miRNAs in inflammatory bowel disease: a review. *Ther Adv Gastroenterol* 2015;**8**:4–22.
51. Mirza AH, Berthelsen CH, Seemann SE, *et al.* Transcriptomic landscape of lncRNAs in inflammatory bowel disease. *Genome Med* 2015;**7**:39.
52. Takagi T, Naito Y, Mizushima K, *et al.* Increased expression of microRNA in the inflamed colonic mucosa of patients with active ulcerative colitis. *J Gastroenterol Hepatol* 2010;**25**[Suppl 1]:S129–33.
53. Kohlhaas S, Garden OA, Scudamore C, Turner M, Okkenhaug K, Vigorito E. Cutting edge: the Foxp3 target miR-155 contributes to the development of regulatory T cells. *J Immunol* 2009;**182**:2578–82.
54. O'Connell RM, Taganov KD, Boldin MP, Cheng G, Baltimore D. MicroRNA-155 is induced during the macrophage inflammatory response. *Proc Natl Acad Sci U S A* 2007;**104**:1604–9.
55. Padua D, Mahurkar-Joshi S, Law IK, *et al.* A long noncoding RNA signature for ulcerative colitis identifies IFNG-AS1 as an enhancer of inflammation. *Am J Physiol Gastrointest Liver Physiol* 2016;**311**:G446–57.
56. Peloquin JM, Goel G, Kong L, *et al.* Characterization of candidate genes in inflammatory bowel disease-associated risk loci. *JCI Insight* 2016;**1**:e87899.
57. Betteridge JD, Armbruster SP, Maydonovitch C, Veerappan GR. Inflammatory bowel disease prevalence by age, gender, race, and geographic location in the U.S. military health care population. *Inflamm Bowel Dis* 2013;**19**:1421–7.
58. Cosnes J, Gower-Rousseau C, Seksik P, Cortot A. Epidemiology and natural history of inflammatory bowel diseases. *Gastroenterology* 2011;**140**:1785–94.
59. Zelinkova Z, Janneke Van Der Woude C. Gender and inflammatory bowel disease. *J Clin Cell Immunol* 2014;**5**:245–50.
60. Choi J, Park T. Multivariate generalized multifactor dimensionality reduction to detect gene-gene interactions. *BMC Syst Biol* 2013;**7**[Suppl 6]:S15.
61. Lesuis N, Befrits R, Nyberg F, van Vollenhoven RF. Gender and the treatment of immune-mediated chronic inflammatory diseases: rheumatoid arthritis, inflammatory bowel disease and psoriasis: an observational study. *BMC Med* 2012;**10**:82.
62. Kobori A, Yagi Y, Imaeda H, *et al.* Interleukin-33 expression is specifically enhanced in inflamed mucosa of ulcerative colitis. *J Gastroenterol* 2010;**45**:999–1007.
63. Seidelin JB, Bjerrum JT, Coskun M, Widjaya B, Vainer B, Nielsen OH. IL-33 is upregulated in colonocytes of ulcerative colitis. *Immunol Lett* 2010;**128**:80–5.
64. Mertz KD, Mager LF, Wasmer MH, *et al.* The IL-33/ST2 pathway contributes to intestinal tumorigenesis in humans and mice. *Oncoimmunology* 2016;**5**:e1062966.

65. Ye Y, Xiao L, Wang SJ, *et al.* Upregulation of REG3A in colorectal cancer cells confers proliferation and correlates with colorectal cancer risk. *Oncotarget* 2016;**7**:3921–33.
66. Harada N, Iijima S, Kobayashi K, *et al.* Human IgGFc binding protein [FcgammaBP] in colonic epithelial cells exhibits mucin-like structure. *J Biol Chem* 1997;**272**:15232–41.
67. Keppner A, Malsure S, Nobile A, Auberson M, Bonny O, Hummler E. Altered Prostasin [CAP1/Prss8] expression favors inflammation and tissue remodeling in DSS-induced colitis. *Inflamm Bowel Dis* 2016;**22**:2824–39.
68. Katakura K, Lee J, Rachmilewitz D, Li G, Eckmann L, Raz E. Toll-like receptor 9-induced type I IFN protects mice from experimental colitis. *J Clin Invest* 2005;**115**:695–702.
69. Billaut-Laden I, Allorge D, Crunelle-Thibaut A, *et al.* Evidence for a functional genetic polymorphism of the human thiosulphate sulfurtransferase [Rhodanese], a cyanide and H2S detoxification enzyme. *Toxicology* 2006;**225**:1–11.
70. Ramasamy S, Singh S, Taniere P, Langman MJ, Eggo MC. Sulfide-detoxifying enzymes in the human colon are decreased in cancer and upregulated in differentiation. *Am J Physiol Gastrointest Liver Physiol* 2006;**291**:G288–96.
71. Wacklin P, Mäkivuokko H, Alakulppi N, *et al.* Secretor genotype [FUT2 gene] is strongly associated with the composition of Bifidobacteria in the human intestine. *PLoS One* 2011;**6**:e20113.
72. Wacklin P, Tuimala J, Nikkilä J, *et al.* Faecal microbiota composition in adults is associated with the FUT2 gene determining the secretor status. *PLoS One* 2014;**9**:e94863.
73. Rausch P, Rehman A, Künzel S, *et al.* Colonic mucosa-associated microbiota is influenced by an interaction of Crohn disease and FUT2 [Secretor] genotype. *Proc Natl Acad Sci U S A* 2011;**108**:19030–5.

Paper II

Journal of Crohn's and Colitis, 2018, 1–10
doi:10.1093/ecco-jcc/jjy117
Advance Access publication August 22, 2018
Original Article

Original Article

Genome-wide DNA Methylation in Treatment-naïve Ulcerative Colitis

Hagar Taman,[a] Christopher G. Fenton,[a] Inga V. Hensel,[a,b] Endre Anderssen,[a] Jon Florholmen,[b,c] Ruth H. Paulssen[a,b]

[a]Genomics Support Centre Tromsø [GSCT], Department of Clinical Medicine, Arctic University of Norway, Tromsø, Norway [b]Gastroenterology and Nutrition Research Group, Department of Clinical Medicine, Arctic University of Norway, Tromsø, Norway [c]Department of Gastroenterology, University Hospital of North Norway, Tromsø, Norway

Corresponding author: Ruth H. Paulssen, PhD, Department of Clinical Medicine, Gastroenterology and Nutrition Research Group, UiT The Arctic University of Norway, Faculty of Health, Sykehusveien 38, N-9038 Tromsø, Norway/ Tel.: +47 77 64 54 80; Email: ruth.h.paulssen@uit.no

Abstract

Background and Aims: The aim of this study was to investigate the genome-wide DNA methylation status in treatment-naïve ulcerative colitis [UC], and to explore the relationship between DNA methylation patterns and gene expression levels in tissue biopsies from a well-stratified treatment-naïve UC patient group.

Methods: Mucosal biopsies from treatment-naïve patients [*n* = 10], and a healthy control group [*n* = 11] underwent genome-wide DNA bisulfite sequencing. Principal component analysis [PCA] and diverse statistical methods were applied to obtain a dataset of differentially methylated genes. DNA methylation annotation was investigated using the UCSC Genome Browser. Gene set enrichments were obtained using the Kyoto Encyclopaedia of Genes and Genomes [KEGG] and PANTHER.

Results: Of all significantly differentially expressed genes [DEGs], 25% correlated with DNA methylation patterns; 30% of these genes were methylated at CpG sites near their transcription start site [TSS]. Hyper-methylation was observed for genes involved in homeostasis and defence, whereas hypo-methylation was observed for genes playing a role in immune response [i.e. chemokines and interleukins]. Of the differentially DNA methylated genes, 25 were identified as inflammatory bowel disease [IBD] susceptibility genes. Four genes [DEFFA6, REG1B, BTNL3, OLFM4] showed DNA methylation in the absence of known CpG islands.

Conclusions: Genome-wide DNA methylation analysis revealed distinctive functional patterns for hyper-and hypo-methylation in treatment-naïve UC. These distinct patterns could be of importance in the development and pathogenesis of UC. Further investigation of DNA methylation patterns may be useful in the development of the targeting of epigenetic processes, and may allow new treatment and target strategies for UC patients.

Keywords: Genome-wide DNA methylation; ulcerative colitis [UC]

1. Introduction

Ulcerative colitis [UC] is one of the two most common conditions that constitute inflammatory bowel disease [IBD] in addition to Crohn's disease [CD]. Whereas CD can affect any area in the gastrointestinal tract, UC affects the mucosa and submucosa of the colon and rectum.[1,2] Due to chronic inflammation, patients affected by UC

have a higher risk of developing colorectal cancer.[3] Approximately 20% of IBD cases can be explained by known genetic variants, suggesting a more complex pathogenesis.[4–6] The current knowledge of the underlying causes of UC is still incomplete. A complex interplay between genetic variation, host immune system, environmental factors and intestinal microbiota has been suggested.[4] Therefore, it has been implied that epigenetic mechanisms may play an important role in disease development of UC.[7–11] Epigenetic processes regulate gene expression via modifications of DNA, histone proteins and chromatin, and are known to play a role in complex disease phenotypes.[12] Epigenetic modifications, such as DNA methylation, are believed to have a role in the immune dysfunction associated with IBD.[13] They are influenced by several environmental factors such as diet[14,15] and smoking[16] which are known to be associated with inflammatory diseases.[17,18] Gene-specific changes in DNA methylation in the pathogenesis of IBD have been recently reported.[19–22] DNA methylation plays a regulatory role in gene transcription, either by activation of proteins that interfere with the suppression of gene transcription, or by inhibiting transcription factors from binding to DNA.[23–25] That is why it isimportant to examine the interaction between gene expression and DNA methylation.

Therefore, in the present study we applied genome-wide methylation profiling by using bisulfite sequencing in order to obtain DNA methylation patterns at a single base-pair resolution.[26,27] This method is a more quantitative approach in producing data with genome-wide coverage than other technologies.[28,29] In addition, DNA methylation has been correlated to transcriptional levels of genes in order to demonstrate possible regulatory DNA methylation features of relevance for UC.

2. Materials and Methods

2.1. Patient material

A standardised sampling method was used to collect mucosal biopsies from the colon of newly diagnosed, treatment-naïve UC patients with mild to moderate disease activity [n = 10] and controls [n = 11]. For controls, biopsies from subjects undergoing cancer screening with normal colonoscopy and normal colonic histological examination were used. UC was diagnosed based upon established clinical, endoscopic and histological criteria as defined by the ECCO guidelines.[30] The grade of inflammation was assessed during colonoscopy using the UC disease activity index [UCDAI] endoscopic sub-score, with 3 to 10 for mild to moderate disease.[31] All biopsies were taken from the sigmoid part of the colon and the case biopsies from a site of active inflammation. Tumour necrosis factor alpha [TNF-α] mRNA expression levels were measured by real-time quantitative polymerase chain reaction PCR [qPCR], thereby indicating the grade of UC activity.[32] This study is part of a larger, already published study, where the gene expression using transcriptome data was assessed.[33] The samples were taken from an established biobank approved by the Norwegian Board of Health. The participants signed an informed and written consent form. The study was approved by the Regional Ethics Committee of North Norway and the Norwegian Social Data Services [REK Nord 2012/1349].

2.2. DNA and RNA isolation

Genomic DNA was isolated using the Allprep DNA/RNA Mini Kit from Qiagen and the QIAcube instrument [Qiagen, Hilden, Germany], according to the manufacturer's protocol. DNA quantity and purity were assessed by using the NanoDrop ND-1000 spectrophotometer [Thermo Fisher Scientific, Wilmington, DE, USA]. The DNA samples were kept at -80°C until further handling.

2.3. Quantitative polymerase chain reaction [qPCR]

The TNF-α levels in biopsies were measured using qPCR. RNA quantity was assessed with NanoVue Plus [GE Healthcare, UK]. Synthesis of cDNA was performed using the QuantiTect Reverse Transcription Kit [Qiagen], and the QuantiNova Probe PCR Kit [Qiagen, Hilden, Germany]. Beta-actin [β-actin] was used as housekeeping gene. For the detection, a CFX Connect Real Time PCR Detection System [Bio-Rad, USA] was used. The results were measured in copies/µg protein. Values <7000 copies/µg protein are considered as non-inflamed tissues, and values >7000 copies/µg protein are considered as inflamed tissues.[32]

2.4. Library preparation and next-generation sequencing

DNA libraries were prepared with the SeqCap Epi CpGiant Enrichment kit [Roche, Switzerland]. The DNA was bisulfite-converted using the EZ DNA Methylation-lightning Kit [Zymo Research, USA] before the hybridisation step and according to the manufacturer's instructions. Transcriptome libraries were prepared as described previously.[33] The amount of input material was 1060 ng of genomic DNA per sample. The Bioanalyzer 2100 and the Agilent DNA 1000 kit [Agilent Technologies, Santa Clara, USA] were used to assess the quality of DNA libraries. DNA libraries with an average fragment size of 329 bp were generated, then diluted to 2 nM, and subsequently sequenced with the NextSeq 550 instrument [Illumina, USA] according to the manufacturer's instructions.

2.5. Data analysis

The algorithm package STAR-2.5.2b [https://github.com/alexdobin/STAR] was used for down-stream analysis of the transcriptome.[34] Transcripts were aligned to UCSC GRCh38/hg38 [http://hgdownload.cse.ucsc.edu/goldenPath/hg38/]. The count matrix was generated by HTSeq-count [https://htseq.readthedocs.io/en/release_0.9.1/], normalised by DESeq2 [https://bioconductor.org/packages/release/bioc/html/DESeq2.html]. Principal component analysis [PCA],[35] Limma,[36] and p-value adjustment methods[37] in Bioconductor R [https://www.bioconductor.org/] were used to obtain and characterise a dataset of significant DEGs and for analysis of relative methylation in patient samples.

For DNA methylation analyses, the Bismark Bisulfite Mapper v0.16.0 [www.bioinformatics.bbsrc.ac.uk/projects/bismark/] was used and the same genome build as the transcriptome was used to generate methylation counts. The globalTest function from the BiSeq package [https://bioconductor.org/packages/release/bioc/html/BiSeq.html] was used to find significant differentially methylated regions between UC and normal samples. As whole-genome bisulfite sequencing [WGBS] data are extremely computationally expensive and rather large, areas of interest were reduced to the promoter regions of expressed transcripts [DESeq2] and transcripts whose normalised log2 [counts] were greater than 5. To find differentially methylated regions, Goeman's Global test was used.[38,39] A modified algorithm of the Goemans's Global test in the BiSeq package [https://bioconductor.org/packages/release/bioc/vignettes/BiSeq/inst/doc/BiSeq.pdf] works with relative methylation data and score tests, with methylation levels as independent variables. Promoters with Goeman's test p-values < 0.05 were kept for further use. Significant regions with low coverage [few methylation sites] or a poor Goeman's statistic were removed. Differentially methylated regions [DMR] of interest were restricted to 200 bp downstream and 2000 bp upstream of a transcription start site [TSS]. DMR regions were further restricted to those containing a minimum of four methylation events. DMRs were investigated with the UCSC Genome Browser [https://genome.ucsc.edu/].

Genes associated with the risk of IBD were downloaded from the genome-wide association studies [GWAS] catalogue,[40] using the search term IBD [www.ebi.ac.uk/gwas]. Gene set enrichments were performed by using the PANTHER classification system [https:// pantherdb.org/], the Kyoto Encyclopaedia of Genes and Genomes [KEGG; www.genome.jp/kegg/]. For principal component analysis [PCA] of the transcription data, the top 5000 most variable of the DESeq2-Rlog variance stabilised transcripts were used. For the estimation of specific cell populations in patient transcription samples, all DESeq2-Rlog normalised transcripts with a log2 average mean >5 were included. The analysis was performed using the R/Bioconductor CellMix manual [http://web.cbio.uct.ac.za/~renaud/CRAN/web/CellMix/] with the IRIS weighted marker list characteristic for the different cell types[41] and as described previously.[33] To investigate the correlation between the cell deconvolution PCA and the methylation PCA scores, partial least squares regression [PLSR] [https://cran.r-project.org/web/packages/pls/index.html] was used.[42] In order to visualise co-variation, the procrustes algorithm [https://www.rdocumentation.org/packages/vegan/versions/2.4-2/topics/procrustes] was applied.[43]

Fisher's exact test[44] was used to compare if the gene list depicted in Supplementary Data 2 [available as Supplementary data at *ECCO-JCC* online] associated with significant methylation changes to gene lists from previous microarray-based methylation studies.[22,45,46]

3. Results

3.1. Patients

Mucosal biopsies from treatment-naïve UC patients [n = 10] and controls [n = 11] were collected according to a standardised sampling method, as described in Materials and Methods [section 2]. The disease activity within UC patients was classified[30] as mild to moderate as described by the UCDAI; the biopsies showed clinical scores of 7 ± standard deviation [SD] 2.6 and endoscopy scores of 1.9 ± SD 0.3. The biopsies from the control group showed normal colonoscopy, colon histology, and immunohistochemistry, with clinical and endoscopy scores = 0. All biopsies were taken from the sigmoid part of the colon. Gender distribution within both groups was almost identical, with seven males in the UC group, eight males in the control group, and three females in each group. The age distribution differed between the groups, at 37 ± SD 12 years in the UC group, and 52 ± SD 14 years in the control group. TNF-α mRNA expression levels were measured by qPCR to estimate the inflammatory status of UC.[32] TNF-α measurements in UC group were estimated as 13,240 ± SD 6056, and for control group as 4291 ± SD 1878. A summary of all patient characteristics is listed in Table 1.

Table 1. Patient characteristics.

Characteristics	Control group [n = 11]	Ulcerative colitis [n = 10]
Male/female	8/3	7/3
Age mean ± SD	52 ± 14	37 ± 12
Endo score mean ± SD	0	1.9 ± 0.3
Clinical score ± SD	0	7 ± 2.6
TNF-α level ± SD	4291 ± 1878	13,240 ± 6056

SD, standard deviation; TNF, tumour necrosis factor.

3.2. Characterisation of DNA methylation in treatment-naïve UC

Pre-processing of the initial RNA and bisulfite sequencing data revealed expression of about 22 000 transcripts which were used for initial principal component analysis [PCA] depicting relative methylation counts [0–100%] for over 9 million cytosine positions for the whole genome of all patient samples, both treatment-naïve UC and normal controls. PCA revealed a clear distinction between UC samples and normal control samples along the first component with a 13.5% explained variance [Figure 1]. Only one patient sample could not be distinguished from normal samples by this method [Figure 1]. A PCA plot indicating the age of participants showed no evidence for age clustering [Supplementary Data 1, available as Supplementary data at *ECCO-JCC* online].

The whole transcriptome of treatment-naïve UC has been recently established, and has been used as a basis for the interpretation of DNA methylation patterns in UC.[33] Results show that 25% of the significantly DEGs [P-value < 0.05; log2 fold-change > 1.0; n = 357] correlated with the observed differential DNA methylation, which resulted in 30% hyper-methylated [n = 87] and 70% hypo-methylated [n = 270] genes [Supplementary Data 2]. Further analysis revealed that approximately 30% of the genes showed DNA methylation at CpG sites in the neighbourhood of their transcription start site [TSS], whereas the remaining 70% showed DNA methylation events at *cis*-acting elements like DNAse1 and enhancers. The relationship between the raw methylation data per sample, average difference between groups, and relationship to the TSS and transcript expression for all differentially methylated genes is depicted in Supplementary Data 3 [available as Supplementary data at *ECCO-JCC* online].

Of the differentially methylated genes, 25 have been related to the currently known 295 IBD susceptibility genes, of which 23 were hypo-methylated and two were hyper-methylated, and correlated with their direction of transcription [Table 2]. Gene annotation revealed their involvement in pathways for cell adhesion, intracellular

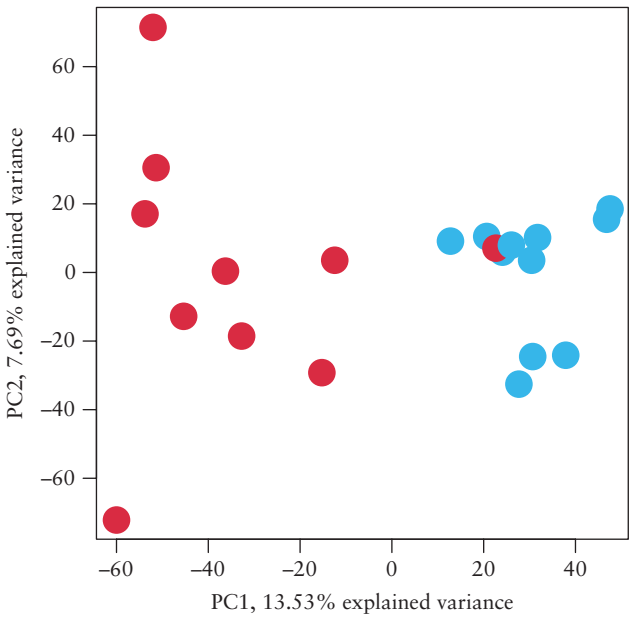


Figure 1. Unsupervised principal component analysis [PCA] depicting relative methylation counts [0–100%] for over nine M cytosine positions, including treatment-naïve ulcerative colitis [UC] [red; n = 10] and normal control [blue; n = 11] patient tissue samples with a variance of 13.5%.

signalling, metabolism, and transport. For six of the IBD susceptibility genes, neutrophil-activating protein 78 [CXCL5], fatty acid desaturase 1 [FADS1], intercellular adhesion molecule 1 [ICAM], solute carrier family 22 member 4 [SLC22A4], tumour necrosis factor receptor superfamily, member 5 [CD40; TNFRSF5], and TNF receptor superfamily member 4 [TNFRSF4] DNA methylation occurred at the transcription start site [TSS] and as indicated in Table 2.[47,48]

The most down-regulated [$P < 0.05$ and log2FC > 1.5] and hyper-methylated genes in treatment-naïve UC are depicted in Table 3. Annotations revealed genes with the most possible relevance for UC: six members of the solute carrier family [SLC17A8, SLC22A4, SLC25A34, SLC30A10, SLC3A1, and SLC6A19], two guanylate cyclase activators [GUCA2A and GUCA2B], defensin B1 [DEFB1], intestinal alkaline phosphatase 1 [ALPI], two UDP-glucuronosyltransferases [UGTA8 and UGTA10], bone morphogenic protein/retinoic acid inducible neural-specific 3 [BRNP3], and proline rich acidic protein [PRAP1].

Hypo-methylated genes in treatment-naïve UC are listed in Table 4. DEGs with log2FC > 3.0 and $P < 0.05$ and their corresponding methylation status have been included in the list. The annotation of the hypo-methylated transcripts revealed genes involved in inflammatory responses, like chemokine receptors [CXCR1, CXCR2], chemokine ligands [CXCL5, CXCL6], interleukins [IL 17A, IL1B], defensins [REG3A, DEFA6], and genes involved in cytokine signalling [SAA1, SAA2, LCN2]. Other hypo-methylations relevant for UC are observed for transporters like aquaporin 9 [AQP9], members of the solute carrier family [SLC6A4, SLC6A14], oncostatin [OSM], and olfactomedin [OLFM4]. For four genes, the observed DNA methylation occurred in the absence of CpG islands or other well-known *cis*-acting regulatory domains; these genes are defensin A6 [DEFA6], olfactomedin 4 [OLFM4], regenerating protein beta 1 [REG1B], and butyrophilin like protein3 [BTNL3], as shown in Figure 2 and Table 5.

In order to explain if cell type populations in biopsies can explain some of the variation in DNA methylation profiles, previously reported cellndeconvolution results of the transcriptome[33] and DNA methylation data were patient-matched and underwent PCA analysis. Further, partial least squares regression [PLSR][42] between the cell deconvolution PCA and the methylation PCA scores showed a strong correlation [Supplementary Data 4, available as Supplementary data at *ECCO-JCC* online].

To visualise co-variation between cell type contributions to methylation data, a biplot of the initial deconvolution PCA for the transcriptome was used in order to display information on both samples and variables [cell types] of the PCA result graphically. The procrustes algorithm[43] was then used to overlay the methylation PCA sample scores onto the cell deconvolution biplot. The result shows that cell type is a significant determinant in methylation profile [Supplementary Data 5, available as Supplementary data at *ECCO-JCC* online].

Table 2. DNA methylated inflammatory bowel disease [IBD] susceptibility genes in treatment-naïve ulcerative colitis [UC].

Gene symbol & annotation	% methyl	$p < 0.05$ methyl	#c	$p < 0.05$ transcript	Log2 FC > 1.0 transcript
Cell adhesion & intercellular signalling					
CD40[a]	6,81	0,0426	94	1,12E-20	1,64
CXCL5[a]	19,21	1,38E-05	30	7,17E-20	4,62
CXCL6	10,20	0,0041	18	5,46E-26	4,06
CXCR5	16,18	0,0004	36	1,36E-07	2,06
FCGR3A	2,69	0,0036	30	2,02E-20	2,13
ICAM1[a]	2,65	0,0034	18	6,61E-28	1,96
IFNG	8,55	0,0057	10	1,54E-06	1,54
IL12RB2	1,97	0,0035	39	7,15E-10	1,03
IL2RA	25,53	0,0001	24	2,51E-17	2,45
ITGAL	15,90	0,0002	10	2,64E-12	1,29
OSM	22,59	0,0002	36	2.00E-17	3,09
SLAMF1	26,79	3,12E-06	11	1,72E-11	1,34
SLAMF7	11,29	0,0062	7	1,11E-07	1,04
TNFRSF9	0,40	0,0075	6	5,27E-24	2,91
TNFRSF4[a]	3,65	0,0348	76	3,46E-18	1,73
TNFSF8	28,03	0,0002	23	1,71E-06	1,01
Intracellular signalling					
APOBEC3G	36,88	0,0241	55	1,68E-15	1,35
CCDC88B	17,29	5,89E-06	118	1,18E-23	1,95
CD6	19,40	3,13E-05	48	2,34E-10	1,44
DOK3	10,80	0,0082	30	5,03E-22	1,77
UBASH3A	5,39	0,0381	8	1,29E-06	1,05
Metabolism					
ARHGAP30	36,89	2,22E-07	17	1,19E-12	1,13
FADS1[a]	0,18	0,0141	30	7,93E-12	1,30
SULT1A2	-8,85	0,0089	13	7,56E-09	-1,67
Transport					
SLC22A4[a]	-0,37	0,0092	172	1,24E-27	-1,72

#c indicates number of methylated cytosines; % methyl indicates % difference of DNA methylation normal [N]-UC.

[a]CpG sites at their transcription start site [TSS].

Table 3. TOP down-regulated and hyper-methylated genes in treatment-naïve ulcerative colitis [UC]

Gene symbol	% methyl	$p < 0.05$ methyl	#c	$p < 0.05$ transcript	Log2 FC > 1.5 transcript
ADIRF	-5,72	0,004	36	4,34E-18	-1,88
AGMO	-10,86	0,022	7	9,44E-09	-1,57
ALPI	-3,05	0,0102	34	1,10E-14	-1,91
ANKRD62	-0,19	0,0011	91	2,47E-14	-2,25
BCHE	-4,17	0,0442	10	4,86E-12	-1,69
BRINP3	-0,22	0,0407	29	6,76E-11	-1,74
CLDN8	-0,45	0,0005	8	2,17E-05	-1,9
CYP3A4	-15,45	0,0009	2	3,87E-21	-3,58
DEFB1	-7,09	0,0334	16	2,16E-11	-1,94
FABP1	-16,56	1,52E-05	18	1,36E-07	-1,66
FAM151A	-3,59	0,0091	16	4,28E-12	-1,59
FRMD1	-2,58	0,0076	109	4,49E-14	-2,26
GBA3	-14,5	0,0033	7	1,15E-07	-1,98
GUCA2A	-12,84	0,0134	18	8,25E-10	-2,26
GUCA2B	-3,08	0,0345	41	2,67E-08	-2,18
HAVCR1	-12,55	3,60E-05	6	4,58E-12	-2,28
HMGCS2	-13,13	6,56E-05	18	2,91E-08	-2,41
HSD17B2	-2,69	0,0456	15	1,71E-23	-2,11
MEP1A	-11,49	0,009	18	1,76E-16	-2,15
OTC	-6,32	0,0359	28	7,09E-10	-1,68
PCK1	-6,73	0,0212	18	1,52E-07	-2,18
PNLIPRP2	-15,2	5,99E-05	23	1,63E-05	-1,74
PRAP1	-1,32	0,0025	260	1,97E-09	-2,49
SLC17A8	-2,89	0,0082	45	2,95E-06	-1,6
SLC22A4	-0,35	0,0174	172	1,25E-27	-1,72
SLC25A34	-12,01	0,0099	40	2,17E-15	-1,76
SLC30A10	-7,67	0,0006	14	4,89E-10	-1,89
SLC3A1	-15.79	0,0002	10	1,29E-17	-2,43
SLC6A19	-12,17	0,0057	68	1,53E-09	-2,82
SULT1A2	-11,71	0,003	13	7,56E-09	-1,67
TINCR	-0,66	0,0107	72	1,58E-12	-1,98
TMIGD1	-9,19	0,0135	6	1,16E-09	-2,32
UGT1A10	-8,16	0,0005	34	4,04E-15	-1,79
UGT1A8	-5,96	0,0091	23	1,67E-10	-1,85

#c indicates number of methylated cytosines. % methyl indicates % difference of DNA methylation normal [N]-UC.

Comparison of methylated genes [Supplementary data 2] showed significant overlaps with previous comparisons between methylation status peripheral blood mononuclear cells [PBMCs] from UC cases and controls,[45] intestinal biopsies from controls and UC patients,[46] and rectal biopsies[22] [Supplementary Data 6, available as Supplementary data at *ECCO-JCC* online].

4. Discussion

It is generally accepted that epigenetic mechanisms like DNA methylation are contributing factors in the pathogenesis of IBD.[13,45] The present study is the first comprehensive study giving a truly genome-wide description of DNA methylation of treatment-naïve UC using next-generation sequencing [NGS]-based bisulfite-sequencing. Furthermore, this study provides an interpretation of DNA methylation status in treatment-naïve UC with correlation to transcriptional levels of genes.[33]

Genome-wide DNA methylation changes in UC have been usually investigated by applying microarray technologies.[19,45,49,50] A correlation between DNA methylation and gene transcription has not been established, except for two recent publication where a few gene candidates have been confirmed by pyro-sequencing.[22,46] However, the evaluated degree of overlap between the present gene list [Supplementary Data 2] and previous genome-wide analyses of methylation using microarray-based technologies[22,45,46] is much larger than those expected by chance [Supplementary Data 6, available as Supplementary data at *ECCO-JCC* online]. The use of microarray technology has several limitations, one of which is that attached array oligo probes might include single nucleotide polymorphisms [SNPs] or repetitive elements which can affect the outcome of the methylation analysis.[19] In addition, pre-defined oligo probes do not cover all regions in the genome where methylation could occur,[28,29] leaving possible methylation events undetected and/or resulting in compromised DNA methylation patterns. This may be the reason for contradictory results regarding the DNA methylation of neutrophil-activating peptide 78 [CXCL5] in UC, where hyper-methylation has been recently reported which is in contrast to the hypo-methylation observed in this study.[20] All these limitations are bypassed with next-generation sequencing technology where methylation detection occurs at a single base-pair, thereby providing methylation profiles with full nucleotide level resolution.[26]

The use of a thoroughly stratified patient group representing only treatment-naïve UC for DNA methylation analysis offered a unique opportunity to investigate the DNA methylation state before

Table 4. TOP up-regulated and hypo-methylated genes in treatment-naïve ulcerative colitis [UC].

Gene symbol	% methyl	$p < 0.05$ methyl	#c	$p < 0.05$ transcript	Log2 FC > 3.0 transcript
AQ9	6,39	0,0013	17	3,562E-32	5,37
C2CD4A	2,43	0,0007	64	1,99E-35	3,72
CD300E	13,58	0,0036	11	4,45E-22	3,44
CHI3L2	11,96	6,26E-06	38	8,53E-17	3,32
CHRDL2	10,62	0,0403	65	5,82E-20	3,34
CXCL5	18,85	1,20E-05	30	7,17E-20	4,63
CXCL6	11,03	0,0020	18	5,46E-26	4,06
CXCR1	6,08	0,0013	6	3,92E-23	4,64
CXCR2	6,03	0,0159	20	2,37E-17	3,44
DEFA6	10,44	0,0002	12	1,74E-15	3,75
DMBT1	11,24	5,84E-06	16	4,19E-17	3,49
FCN1	8,55	0,0459	12	3,57E-28	3,25
FFAR2	2,16	9,25E-05	54	1,91E-20	3,02
GABRP	7,43	0,0257	16	8,36E-35	3,72
GZMB	12,88	0,0005	6	3,37E-24	3,16
HCAR2	9,48	0,0012	26	8,69E-22	3,58
HCAR3	14,16	0,0010	18	5,01E-22	4,32
IL17A	9,08	7,82E-06	32	1,50E-33	5,03
IL1B	9,15	2,73E-08	13	1,09E-32	3,28
LCN2	5,23	0,0243	29	2,24E-53	4,95
LYPD5	3,93	0,0254	40	5,55E-23	3,07
OLFM4	10,75	0,0251	24	1,02E-13	3,49
OSM	22,21	0,0008	36	2,01E-17	3,10
PI3	2,26	0,0153	14	2,73E-31	4,11
REG1B	6,63	0,0283	7	3,76E-12	3,73
REG3A	13,27	0,0299	2	2,63E-15	4,19
S100A9	8,26	0,0002	40	1,00E-17	3,46
SAA1	10,55	5,19E-06	26	5,16E-46	6,40
SAA2	11,89	0,0075	6	7,31E-43	6,23
SLC26A4	5,22	8,27E-07	54	3,26E-15	3,28
SLC6A14	3,99	0,0432	29	1,73E-48	5,66

#c indicates number of methylated cytosines. % methyl indicates % difference of DNA methylation normal [N]-UC.

prescription of any medication. This is of importance, since recent reports implied that medication, such as various non-prescription, over-the-counter non-steroidal anti-inflammatory drugs [NSAIDs] and immunosuppressive drugs can have short- and long-term effects on the immune response.[51] For example, aspirin has been shown to result in hypo-methylation of cadherin 1 [CDH1] in the gastric mucosa.[52] In addition, immunosuppressant therapy and long-standing disease might inaugurate unwanted bias in experiments aiming to investigate the treatment-naïve status of UC.[53–55] In addition, age may affect DNA methylation.[56,57] However, the results from the patient population of this study indicated that age does not seem to play a significant role [Supplementary Data 1, available as Supplementary data at *ECCO-JCC* online]. This might be due to the small number of patients aged over 60 in both control and patient groups.

Many studies have characterised DNA methylation in UC without relating the obtained data to transcriptional levels of genes.[19,20,22] The correlation of DNA methylation status with transcription levels of genes are of importance in order to define physiological implications of the DNA methylation event. We have previously characterised the whole transcriptome of treatment-naïve UC and used these data in order to relate DNA methylation to gene expression.[33] The results revealed that only 72% of DNA methylation events correspond with differential gene transcription levels [see Supplementary Data 2]. Annotations of DNA methylation sites covered regions of 2000 bp upstream and 200 bp downstream of the transcription start site [TSS] of genes, and were found to be correlated with transcription levels, thereby revealing possible disease-specific methylation patterns. It is noted that CpG islands are associated with the control of gene expression, and it would be expected that CpG islands might display tissue-specific patterns of DNA methylation.[23,58,59] However, it has been shown that CpG islands associated with TSS rarely show tissue specific methylation patterns[24,25,60–62] Instead, CpG regions located as far as 2 kb from CpG islands have highly conserved patterns of tissue-specific methylation, and methylation is highly correlated with reduced gene expression.[63] Taking this into consideration, our data revealed 90 genes containing CpG sites in this region whereof 58 were at the transcription start site [TSS] of genes; 34 genes showed DNA methylation upstream TSS and might be considered as tissue-specific DNA methylation sites [Supplementary Data 3]. A number of hyper-methylations with corresponding gene expression levels have been found in this study [Table 3]. For example, bone morphogenic/retinoic acid inducible neural-specific protein 3 [BRINP3] has been reported to be usually under-expressed in UC.[64] BRINP3 expression is influenced by DNA hyper-methylation within its promoter, as has been reported recently.[65]

Cell type populations present in tissue biopsies might also explain some of the variation observed in DNA methylation profiles [Supplementary Data 4 and 5]. During inflammation of the mucosa, the fraction of epithelial cells is diminished, which results in impaired intestinal permeability and a dysregulation of homeostasis.[66,67] This might be reflected by the hyper-methylation and

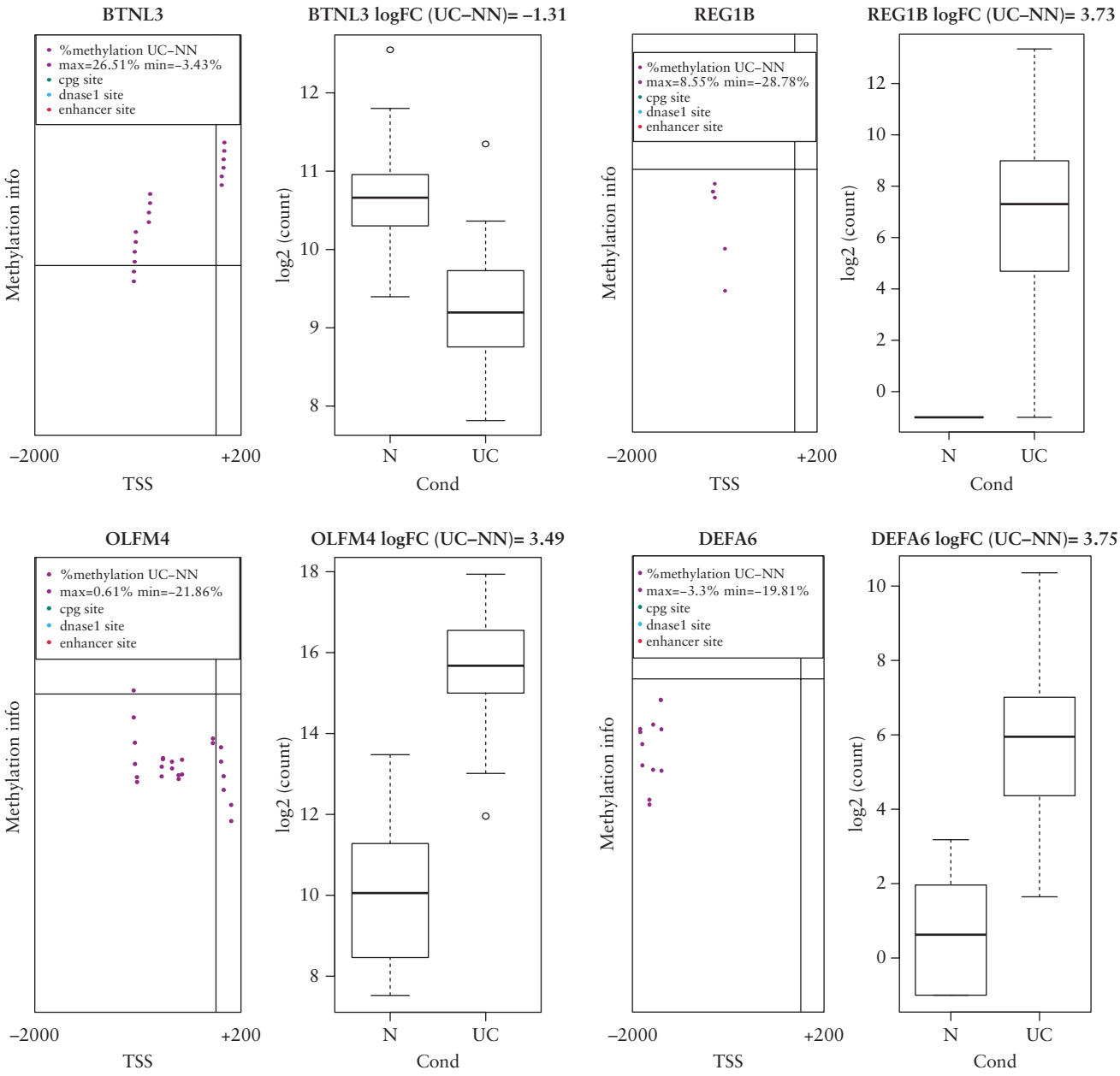


Figure 2. Genes with novel DNA methylation features in treatment-naïve ulcerative colitis [UC]. The left side of individual illustrations shows the difference in relative methylation level between UC and normal samples [N]. Transcription start site [TSS] is indicated as a vertical line. The horizontal line shows where UC methylation equals N methylation. Each black circle represents a methylation event. Black circles over the horizontal line represent an increase in UC sample methylation compared with N methylation at that site. Black circles under the horizontal line represent an increase in N sample methylation compared with UC methylation at that site. The region between 2000 bp upstream and 200 bp downstream of the TSS is shown. The right side of individual illustrations shows boxplots of DESEQ2 log2 normalised values for gene of interest, normal control [N] versus ulcerative colitis [UC]. Genes are indicated: defensin A6 [DEFA6]; olfactomedin 4 [OLFM4]; butyrophilin like 3 [BTNL3]; regenerating protein 1B [REG1B].

Table 5. Genes with novel DNA methylation features.

Gene symbol	% methyl	$p < 0.05$ methyl	#c	$p < 0.05$ transcript	log2 FC > 1.0 transcript
BTNL3	-13,20	0.0012	16	4,97E-07	-1.31
DEFA6	9,20	0.0002	12	4,56E-17	+3.75
OLFM4	9,95	0.0251	24	3,83E-15	+3.49
REG1B	6,68	0.0282	7	1,95E-13	+3.73

#c indicates number of methylated cytosines. % methyl indicates % difference of DNA methylation normal-ulcerative colitis [N-UC].

down-regulation of proline-rich acidic protein 1 [PRAP1] and members of the solute carrier protein family [SLC6A19 and SLC3A1] which are involved in the maintenance of homeostasis in epithelial cells.[68–70] In concordance, hyper-methylation of genes that are involved in the gut mucosal defence system could also be detected, such as intestinal alkaline phosphatase 1 [ALPI] which is involved in the prevention of bacterial translocation in the gut,[71] and defensin B1 [DEFB1] which is predominately expressed in neutrophils and is implicated in the resistance of epithelial surfaces to microbial colonisation.[72] Members of the UDP glucuronosyltransferase family [UGT1A10 and UGT1A8] are located primarily in gastrointestinal [GI] mucosa from the duodenum to through the colon,[73–75] and are involved in detoxification in order to restrict GI absorption of damaging chemicals via the cytochrome P450 system [CYP3A4]; all are hyper-methylated in UC [Table 3]. In addition, the reported down-regulation of guanylate cyclase activators GUCA2A and GUCA2B, which are involved in gastrointestinal fluid and electrolyte balance, is most likely linked to the hyper-methylation of both genes during inflammation observed here.[76]

On the other hand, hypo-methylation with corresponding up-regulation of genes relevant in UC has been observed. This include genes that are up-regulated due to the response to inflammation, and represent mostly genes with association to the innate immune system like chemokines, chemokine receptors, cytokines, interleukins, and transporters [Table 4].[33]

For further characterisation, the differentially methylated genes were related to currently known IBD susceptibility genes, as revealed by genome-wide association studies [GWAS].[47,48] Of the significantly differentially DNA methylated genes, 25 are associated with IBD, [Table 2] of which six genes have CpG islands located at their transcription start site [TSS].

The majority of DNA methylation occurs on cytosines that precede a guanine nucleotide or CpG sites. However, this study reveals previously unknown DNA methylation patterns of genes in treatment-naïve UC, which are not dependent on CpG sites or known regulatory transcriptional *cis*-acting elements like DNAse1 and enhancers. This has been also reported in the tissue of adult mouse brain, where a significant percentage of methylated non-CpG sites have been identified.[77] This phenomenon implies novel regulatory features of DNA methylation in UC, with genes involved in pro-inflammatory responses and possible antimicrobial activities. These would involve defensin A6 [DEFA6] and regenerating protein 1B [REG1B], facilitation of cell adhesion through interaction with lectins and cadherins (olfactomedin 4 [OLFM4]), and lipid metabolism (butyrophilin-like 3 [BTLN3]) [Figure 2 and Table 5].[78] All four genes have been associated with colorectal cancer [CRC] and/or have shown to play a role in CRC progression and development.[79–83] However, the role of non-CpG methylation is still unclear.

Regarding the heterogeneity of tissues, it is clear that the methylation events could occur in different cell subtypes present in the tissue samples from UC patients. This might be the situation for DEFA6, which is a Paneth cell-specific protein and which is predominantly abundant in the epithelia of the intestinal mucosal surface and in the granules of neutrophils.[84] Epithelial cells are impaired and less abundant in the inflamed mucosa,[66] and it is therefore believed that the observed hypo-methylation of DEFA6 most likely occurs in the neutrophils with elevated fractions in inflamed mucosal tissue.[85] The same might be the situation for BTNL3 which modulates T-cell mediated immune response. The observed hyper-methylation of BTLN3 may take place in the increased fractions of T lymphocytes present in inflamed mucosa. However, with isolated cell fractions or single cell sequencing approaches, one would be able to confirm these results.

In conclusion, this comprehensive study shows for the first time that the use of well-stratified treatment- naïve UC patient samples in combination with genome-wide bisulfite-sequencing technology can reveal DNA methylation patterns of importance for UC pathogenesis. Potentially significant might be the differential DNA methylation patterns, with observed hyper-methylation of genes involved in homeostasis and defence, and hypo-methylation of genes involved in immune response with representative members of the innate immune system. Further investigation of such players may be useful for the development of epigenetic drugs and may allow new treatment strategies for UC patients in the future.

Funding

This work was supported by the Northern Norway Regional Health Authority [SPF-1209-14].

Conflict of Interest

The authors declare no conflict of interest regarding the publication of this paper.

Acknowledgements

The authors thank Ingrid Christiansen for the technical help performing the TNF-α levels measurements, and Odd-Sverre Moen for administrating the patient samples.

Author Contributions

HT performed most of the experiments and wrote parts of the manuscript. CGF performed most statistical analyses and revised the manuscript. IVH performed a part of the experiments and revised the manuscript. EA performed a part of data analysis and revised the manuscript. JF was involved in evaluating and providing clinical samples from patients and healthy controls and revised the manuscript. RHP was involved in projects inception, design, analysis, supervision, manuscript writing, and revision.

Supplementary Data

Supplementary data are available at *ECCO-JCC* online.

References

1. Khor B, Gardet A, Xavier RJ. Genetics and pathogenesis of inflammatory bowel disease. *Nature* 2011;**474**:307–17.
2. Head KA, Jurenka JS. Inflammatory bowel disease. Part 1: ulcerative colitis–pathophysiology and conventional and alternative treatment options. *Altern Med Rev* 2003;**8**:247–83.
3. Planell N, Lozano JJ, Mora-Buch R, *et al.* Transcriptional analysis of the intestinal mucosa of patients with ulcerative colitis in remission reveals lasting epithelial cell alterations. *Gut* 2013;**62**:967–76.
4. Jostins L, Ripke S, Weersma RK, *et al.*; International IBD Genetics Consortium [IIBDGC]. Host-microbe interactions have shaped the genetic architecture of inflammatory bowel disease. *Nature* 2012;**491**:119–24.
5. Liu JZ, van Sommeren S, Huang H, *et al.*; International Multiple Sclerosis Genetics Consortium; International IBD Genetics Consortium. Association analyses identify 38 susceptibility loci for inflammatory bowel disease and highlight shared genetic risk across populations. *Nat Genet* 2015;**47**:979–86.
6. Shaw KA, Cutler DJ, Okou D, *et al*. Genetic variants and pathways implicated in a pediatric inflammatory bowel disease cohort. *Genes Immun* 2018, March 28. doi: 10.1038/s41435-018-0015-2. [Epub ahead of print.]
7. Loddo I, Romano C. Inflammatory bowel disease: genetics, epigenetics, and pathogenesis. *Front Immunol* 2015;**6**:551.

8. Karatzas PS, Gazouli M, Safioleas M, Mantzaris GJ. DNA methylation changes in inflammatory bowel disease. *Ann Gastroenterol* 2014;**27**:125–32.
9. Saito S, Kato J, Hiraoka S, *et al.* DNA methylation of colon mucosa in ulcerative colitis patients: correlation with inflammatory status. *Inflamm Bowel Dis* 2011;**17**:1955–65.
10. Toyota M, Itoh F, Kikuchi T, *et al.* DNA methylation changes in gastrointestinal disease. *J Gastroenterol* 2002;**37**[Suppl 14]:97–101.
11. Mohn F, Weber M, Rebhan M, *et al.* Lineage-specific polycomb targets and de novo DNA methylation define restriction and potential of neuronal progenitors. *Mol Cell* 2008;**30**:755–66.
12. Murphy TM, Mill J. Epigenetics in health and disease: heralding the EWAS era. *Lancet* 2014;**383**:1952–4.
13. Scarpa M, Stylianou E. Epigenetics: concepts and relevance to IBD pathogenesis. *Inflamm Bowel Dis* 2012;**18**:1982–96.
14. Lim U, Song MA. Dietary and lifestyle factors of DNA methylation. *Methods Mol Biol* 2012;**863**:359–76.
15. Dick KJ, Nelson CP, Tsaprouni L, *et al.* DNA methylation and body-mass index: a genome-wide analysis. *Lancet* 2014;**383**:1990–8.
16. Tsaprouni LG, Yang TP, Bell J, *et al.* Cigarette smoking reduces DNA methylation levels at multiple genomic loci but the effect is partially reversible upon cessation. *Epigenetics* 2014;**9**:1382–96.
17. Mishkin S. Dairy sensitivity, lactose malabsorption, and elimination diets in inflammatory bowel disease. *Am J Clin Nutr* 1997;**65**:564–7.
18. Rubin DT, Hanauer SB. Smoking and inflammatory bowel disease. *Eur J Gastroenterol Hepatol* 2000;**12**:855–62.
19. Lin Z, Hegarty JP, Cappel JA, *et al.* Identification of disease-associated DNA methylation in intestinal tissues from patients with inflammatory bowel disease. *Clin Genet* 2011;**80**:59–67.
20. Karatzas PS, Mantzaris GJ, Safioleas M, Gazouli M. DNA methylation profile of genes involved in inflammation and autoimmunity in inflammatory bowel disease. *Medicine [Baltimore]* 2014;**93**:e309.
21. Nimmo ER, Prendergast JG, Aldhous MC, *et al.* Genome-wide methylation profiling in Crohn's disease identifies altered epigenetic regulation of key host defense mechanisms including the Th17 pathway. *Inflamm Bowel Dis* 2012;**18**:889–99.
22. Cooke J, Zhang H, Greger L, *et al.* Mucosal genome-wide methylation changes in inflammatory bowel disease. *Inflamm Bowel Dis* 2012;**18**:2128–37.
23. Moore LD, Le T, Fan G. DNA methylation and its basic function. *Neuropsychopharmacology* 2013;**38**:23–38.
24. Eckhardt F, Lewin J, Cortese R, *et al.* DNA methylation profiling of human chromosomes 6, 20 and 22. *Nat Genet* 2006;**38**:1378–85.
25. Maunakea AK, Nagarajan RP, Bilenky M, *et al.* Conserved role of intragenic DNA methylation in regulating alternative promoters. *Nature* 2010;**466**:253–7.
26. Darst RP, Pardo CE, Ai L, Brown KD, Kladde MP. Bisulfite sequencing of DNA. *Curr Protoc Mol Biol* 2010;**Chapter 7**:Unit 7.9.1–17.
27. Li Y, Tollefsbol TO. DNA methylation detection: bisulfite genomic sequencing analysis. *Methods Mol Biol* 2011;**791**:11–21.
28. Teh AL, Pan H, Lin X, *et al.* Comparison of methyl-capture sequencing vs. infinium 450K methylation array for methylome analysis in clinical samples. *Epigenetics* 2016;**11**:36–48.
29. Harris RA, Wang T, Coarfa C, *et al.* Comparison of sequencing-based methods to profile DNA methylation and identification of monoallelic epigenetic modifications. *Nat Biotechnol* 2010;**28**:1097–105.
30. Magro F, Gionchetti P, Eliakim R, *et al.*; European Crohn's and Colitis Organisation [ECCO]. Third European evidence-based consensus on diagnosis and management of ulcerative colitis. Part 1: definitions, diagnosis, extra-intestinal manifestations, pregnancy, cancer surveillance, surgery, and ileo-anal pouch disorders. *J Crohns Colitis* 2017;**11**:649–70.
31. McNally PR. Literature review: Core I & II: colonic release budesonide for the induction of remission for mild-moderate ulcerative colitis. *VHJOE* 2014;**13**:1.
32. Olsen T, Goll R, Cui G, *et al.* Tissue levels of tumor necrosis factor-alpha correlates with grade of inflammation in untreated ulcerative colitis. *Scand J Gastroenterol* 2007;**42**:1312–20.
33. Taman H, Fenton CG, Hensel IV, *et al.* Transcriptomic landscape of treatment - naive ulcerative colitis. *J Crohns Colitis* 2018;**12**:327–36.
34. Dobin A, Davis CA, Schlesinger F, *et al.* STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* 2013;**29**:15–21.
35. Pearson K. On lines and planes of closest fit to systems of points in space. *Philosophical Magazine* 1901;**2**:14.
36. Ritchie ME, Phipson B, Wu D, *et al.* limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res* 2015;**43**:e47.
37. Benjamini Y, Hochberg Y. Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J R Stat Soc* 1995;**57**:12.
38. Goeman JJ, Van de Geer SA, Van Houwelingen HC. Testing against a high dimensional alternative. *J R Stat Soc B* 2006;**68**:477–93.
39. Klein HU, Hebestreit K. An evaluation of methods to test predefined genomic regions for differential methylation in bisulfite sequencing data. *Brief Bioinform* 2016;**17**:796–807.
40. MacArthur J, Bowler E, Cerezo M, *et al.* The new NHGRI-EBI Catalog of published genome-wide association studies [GWAS Catalog]. *Nucleic Acids Res* 2017;**45**:D896–901.
41. Abbas AR, Baldwin D, Ma Y, *et al.* Immune response in silico [IRIS]: immune-specific genes identified from a compendium of microarray expression data. *Genes Immun* 2005;**6**:319–31.
42. Mevik B, Wehrens R. The pls package: principle component and partial least squares regression in r. *J Stat Softw* 2007;**18**:23.
43. Peres-Neto PR, Jackson DA. How well do multivariate data sets match? The advantages of a Procrustean superimposition approach over the Mantel test. *Oecologia* 2001;**129**:169–78.
44. Fisher RA. On the interpretation of $\chi 2$ from contingency tables, and the calculation of p. *J R Stat Soc* 1922;**85**:7.
45. McDermott E, Ryan EJ, Tosetto M, *et al.* DNA methylation profiling in inflammatory bowel disease provides new insights into disease pathogenesis. *J Crohns Colitis* 2016;**10**:77–86.
46. Häsler R, Feng Z, Bäckdahl L, *et al.* A functional methylome map of ulcerative colitis. *Genome Res* 2012;**22**:2130–7.
47. 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**:246–52.
48. de Lange KM, Moutsianas L, Lee JC, *et al.* Genome-wide association study implicates immune activation of multiple integrin genes in inflammatory bowel disease. *Nat Genet* 2017;**49**:256–61.
49. Harris RA, Nagy-Szakal D, Pedersen N, *et al.* Genome-wide peripheral blood leukocyte DNA methylation microarrays identified a single association with inflammatory bowel diseases. *Inflamm Bowel Dis* 2012;**18**:2334–41.
50. Kang K, Bae JH, Han K, *et al.* A genome-wide methylation approach identifies a new hypermethylated gene panel in ulcerative colitis. *Int J Mol Sci* 2016;**17**. doi: 10.3390/ijms17081291.
51. Tahara T, Shibata T, Yamashita H, *et al.* Chronic nonsteroidal anti-inflammatory drug [NSAID] use suppresses multiple CpG islands hyper methylation [CIHM] of tumor suppressor genes in the human gastric mucosa. *Cancer Sci* 2009;**100**:1192–7.
52. Tahara T, Shibata T, Nakamura M, *et al.* Chronic aspirin use suppresses CDH1 methylation in human gastric mucosa. *Dig Dis Sci* 2010;**55**:54–9.
53. Harris RA, Nagy-Szakal D, Mir SA, *et al.* DNA methylation-associated colonic mucosal immune and defense responses in treatment-naïve pediatric ulcerative colitis. *Epigenetics* 2014;**9**:1131–7.
54. Axelrad JE, Lichtiger S, Yajnik V. Inflammatory bowel disease and cancer: The role of inflammation, immunosuppression, and cancer treatment. *World J Gastroenterol* 2016;**22**:4794–801.
55. Herfarth HH, Kappelman MD, Long MD, Isaacs KL. Use of methotrexate in the treatment of inflammatory bowel diseases. *Inflamm Bowel Dis* 2016;**22**:224–33.
56. Horvath S, Zhang Y, Langfelder P, *et al.* Aging effects on DNA methylation modules in human brain and blood tissue. *Genome Biol* 2012;**13**:R97.
57. Ciccarone F, Tagliatesta S, Caiafa P, Zampieri M. DNA methylation dynamics in aging: how far are we from understanding the mechanisms? *Mech Ageing Dev* 2018;**174**:3–17.

58. Laird PW. Principles and challenges of genomewide DNA methylation analysis. *Nat Rev Genet* 2010;**11**:191–203.
59. Zhang G, Pradhan S. Mammalian epigenetic mechanisms. *IUBMB Life* 2014;**66**:240–56.
60. Rakyan VK, Hildmann T, Novik KL, *et al.* DNA methylation profiling of the human major histocompatibility complex: a pilot study for the human epigenome project. *PLoS Biol* 2004;**2**:e405.
61. Meissner A, Mikkelsen TS, Gu H, *et al.* Genome-scale DNA methylation maps of pluripotent and differentiated cells. *Nature* 2008;**454**:766–70.
62. Illingworth RS, Gruenewald-Schneider U, Webb S, *et al.* Orphan CpG islands identify numerous conserved promoters in the mammalian genome. *PLoS Genet* 2010;**6**:e1001134.
63. Irizarry RA, Ladd-Acosta C, Wen B, *et al.* The human colon cancer methylome shows similar hypo- and hypermethylation at conserved tissue-specific CpG island shores. *Nat Genet* 2009;**41**:178–86.
64. Chen L, Su L, Li J, *et al.* Hypermethylated FAM5C and MYLK in serum as diagnosis and pre-warning markers for gastric cancer. *Dis Markers* 2012;**32**:195–202.
65. Smith PJ, Levine AP, Dunne J, *et al.* Mucosal transcriptomics implicates under expression of BRINP3 in the pathogenesis of ulcerative colitis. *Inflamm Bowel Dis* 2014;**20**:1802–12.
66. Mankertz J, Schulzke JD. Altered permeability in inflammatory bowel disease: pathophysiology and clinical implications. *Curr Opin Gastroenterol* 2007;**23**:379–83.
67. He F, Peng J, Deng XL, *et al.* Mechanisms of tumor necrosis factor-alpha-induced leaks in intestine epithelial barrier. *Cytokine* 2012;**59**:264–72.
68. Zhang J, Wong H, Ramanan S, Cheong D, Leong A, Hooi SC. The proline-rich acidic protein is epigenetically regulated and inhibits growth of cancer cell lines. *Cancer Res* 2003;**63**:6658–65.
69. Ghishan FK, Kiela PR. Epithelial transport in inflammatory bowel diseases. *Inflamm Bowel Dis* 2014;**20**:1099–109.
70. Pérez-Torras S, Iglesias I, Llopis M, *et al.* Transportome profiling identifies profound alterations in Crohn's disease partially restored by commensal bacteria. *J Crohns Colitis* 2016;**10**:850–9.
71. Fawley J, Gourlay DM. Intestinal alkaline phosphatase: a summary of its role in clinical disease. *J Surg Res* 2016;**202**:225–34.
72. Müller CA, Autenrieth IB, Peschel A. Innate defenses of the intestinal epithelial barrier. *Cell Mol Life Sci* 2005;**62**:1297–307.
73. Basu NK, Kole L, Kubota S, Owens IS. Human UDP-glucuronosyltransferases show atypical metabolism of mycophenolic acid and inhibition by curcumin. *Drug Metab Dispos* 2004;**32**:768–73.
74. Basu NK, Kubota S, Meselhy MR, *et al.* Gastrointestinally distributed UDP-glucuronosyltransferase 1A10, which metabolizes estrogens and nonsteroidal anti-inflammatory drugs, depends upon phosphorylation. *J Biol Chem* 2004;**279**:28320–9.
75. Dellinger RW, Fang JL, Chen G, Weinberg R, Lazarus P. Importance of UDP-glucuronosyltransferase 1A10 [UGT1A10] in the detoxification of polycyclic aromatic hydrocarbons: decreased glucuronidative activity of the UGT1A10139Lys isoform. *Drug Metab Dispos* 2006;**34**:943–9.
76. Brenna Ø, Bruland T, Furnes MW, *et al.* The guanylate cyclase-C signaling pathway is down-regulated in inflammatory bowel disease. *Scand J Gastroenterol* 2015;**50**:1241–52.
77. Xie W, Barr CL, Kim A, *et al.* Base-resolution analyses of sequence and parent-of-origin dependent DNA methylation in the mouse genome. *Cell* 2012;**148**:816–31.
78. Paulssen RH, Taman H, Fenton CG, *et al.* Novel regulatory DNA methylation features of relevance for ulcerative colitis. *Am J Gastroenterol* 2018;**113**:P-050.
79. Park JM, Han NY, Han YM, *et al.* Predictive proteomic biomarkers for inflammatory bowel disease-associated cancer: where are we now in the era of the next generation proteomics? *World J Gastroenterol* 2014;**20**:13466–2476.
80. van der Flier LG, Haegebarth A, Stange DE, van de Wetering M, Clevers H. OLFM4 is a robust marker for stem cells in human intestine and marks a subset of colorectal cancer cells. *Gastroenterology* 2009;**137**:15–7.
81. Lebrero-Fernández C, Wenzel UA, Akeus P, *et al.* Altered expression of Butyrophilin [BTN] and BTN-like [BTNL] genes in intestinal inflammation and colon cancer. *Immun Inflamm Dis* 2016;**4**:191–200.
82. van Beelen Granlund A, Østvik AE, Brenna Ø, Torp SH, Gustafsson BI, Sandvik AK. REG gene expression in inflamed and healthy colon mucosa explored by in situ hybridisation. *Cell Tissue Res* 2013;**352**:639–46.
83. Gersemann M, Becker S, Nuding S, *et al.* Olfactomedin-4 is a glycoprotein secreted into mucus in active IBD. *J Crohns Colitis* 2012;**6**:425–34.
84. Elphick DA, Mahida YR. Paneth cells: their role in innate immunity and inflammatory disease. *Gut* 2005;**54**:1802–9.
85. Arijs I, De Hertogh G, Lemaire K, *et al.* Mucosal gene expression of antimicrobial peptides in inflammatory bowel disease before and after first infliximab treatment. *PLoS One* 2009;**4**:e7984.

Paper III

PLOS ONE

RESEARCH ARTICLE

DNA hypo-methylation facilitates anti-inflammatory responses in severe ulcerative colitis

Hagar Taman[1,2], Christopher G. Fenton[2], Endre Anderssen[2], Jon Florholmen[3], Ruth H. Paulssen[1,2]*

1 Clinical Bioinformatics Research Group, Department of Clinical Medicine, UiT- The Arctic University of Norway, Tromsø, Norway, **2** Genomics Support Centre Tromsø (GSCT), Department of Clinical Medicine, UiT- The Arctic University of Norway, Tromsø, Norway, **3** Gastroenterology and Nutrition Research Group, Department of Clinical Medicine, UiT- The Arctic University of Norway, Tromsø, Norway

* ruth.h.paulssen@uit.no

Citation: Taman H, Fenton CG, Anderssen E, Florholmen J, Paulssen RH (2021) DNA hypo-methylation facilitates anti-inflammatory responses in severe ulcerative colitis. PLoS ONE 16(4): e0248905. https://doi.org/10.1371/journal.pone.0248905

Editor: Hiroyasu Nakano, Toho University Graduate School of Medicine, JAPAN

Received: December 14, 2020

Accepted: March 6, 2021

Published: April 1, 2021

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Funding: THis work was supported by the Northern Norway Regional Health Authority Helse-Nord (SFP-1209-14).s

Competing interests: The authors have no competing interests.

Abstract

Severe ulcerative colitis (UC) is a potentially life-threatening disease with a potential colorectal cancer (CRC) risk. The aim of this study was to explore the relationship between transcriptomic and genome-wide DNA methylation profiles in a well-stratified, treatment-naïve severe UC patient population in order to define specific epigenetic changes that could be responsible for the grade of disease severity. Mucosal biopsies from treatment-naïve severe UC patients (n = 8), treatment-naïve mild UC (n = 8), and healthy controls (n = 8) underwent both whole transcriptome RNA-Seq and genome-wide DNA bisulfite- sequencing, and principal component analysis (PCA), cell deconvolutions and diverse statistical methods were applied to obtain a dataset of significantly differentially expressed genes (DEGs) with correlation to DNA methylation for severe UC. DNA hypo-methylation correlated with approximately 80% of all DEGs in severe UC when compared to mild UC. Enriched pathways of annotated hypo-methylated genes revealed neutrophil degranulation, and immuno-regulatory interactions of the lymphoid system. Specifically, hypo-methylated anti-inflammatory genes found for severe UC were IL10, SIGLEC5, CD86, CLMP and members of inflammasomes NLRP3 and NLRC4. Hypo-methylation of anti-inflammatory genes during severe UC implies an interplay between the epithelium and lamina propria in order to mitigate inflammation in the gut. The specifically DNA hypo-methylated genes found for severe UC can potentially be useful biomarkers for determining disease severity and in the development of new targeted treatment strategies for severe UC patients.

Introduction

Ulcerative colitis (UC) is an inflammatory disorder that affects the mucosa and submucosa of the colon and rectum and is a chronic disease with a relapsing course [1]. Disease severity is wide ranging with most UC patients manifesting a mild to moderate disease activity [2, 3]. However, between 15–30% of UC patients will experience at least one incident of acute severe

colitis during the disease course, requiring hospitalization for immediate medical treatment [2, 4, 5]. Patients whose inflammation is more severe and more extensive are more likely to develop CRC [6]. In cases where medical therapy fails, colectomy is considered [2–5, 7]. The underlying causes of UC are still not completely understood. It has been suggested that UC is the result of a dysregulated immune response to environmental factors and commensal pathogens in a genetically predisposed host [8, 9]. Therefore, epigenetic mechanisms, such as DNA methylation have been implied to play a key role in disease development of UC [10–13]. Methylation of cytosine groups in DNA molecules can change the structure and interactions of a DNA sequence without changing the sequence [14]. In mammals, methylation primarily occurs in CpG dinucleotides and when occurring in CpG rich areas of promoters is linked to lasting stable repression of gene expression [15].

Epigenetic modifications, such as DNA hyper-methylation are believed to have a role in the immune dysfunction associated with IBD [12, 13]. However, less attention has been devoted to the role of DNA hypo-methylation for UC which represents one of the major DNA methylation states that refers to a relative decrease from an ordinary methylation level. UC by itself might induce hypo-methylation of DNA and a decrease in DNA methylation can have an impact on the predisposition to pathological states and UC development. Global DNA hypo-methylation has been suggested to contribute to neoplastic transformation which suggest that DNA hypo-methylation plays a previously unappreciated role in intestinal adenoma initiation [16].

Recently, whole transcriptomic and genome-wide DNA methylation profiles for treatment-naïve UC have been established for mild and moderate disease [17, 18]. This study focuses on the role of DNA hypo-methylation in a severe UC phenotype in comparison to a mild UC phenotype with the aim to identify DNA hypo-methylation patterns that might correlate with disease severity. This attempt makes it possible to identify biomarker groups that can help determine new potential personalized treatment targets for patients with severe UC and might improve the clinical outcome for this patient group.

Materials and methods

Patient material

Twenty-four mucosal biopsies were collected with a standardized sampling method from three patient groups, newly diagnosed treatment-naïve UC patients with severe disease activity (n = 8), newly diagnosed treatment-naïve UC patients with mild disease activity (n = 8), and normal control patients (n = 8). The biopsies were taken from the recto-sigmoid part of the colon. Subjects which underwent cancer screening, and showed normal colonoscopy and normal colonic histological examination, served as controls. Diagnosis of UC disease activity was based on established clinical, endoscopic and histological criteria as defined by the ECCO guidelines [19]. The inflammation grade was evaluated during colonoscopy using the UC disease activity index (UCDAI) [20]. Control biopsies showed normal colonoscopy, normal colon histology and immunohistochemistry, with a clinical and an endoscopic score of 0. TNF-α mRNA expression was detected by quantitative real-time polymerase chain reaction (qPCR) [21]. All patient characteristics are depicted in Table 1. The samples were taken from an established Biobank approved by the Norwegian Board of Health. The study was approved by the Regional Ethics Committee of North Norway and the Norwegian Social Science Data Services (REK Nord 2012/1349).

DNA and RNA isolation

Genomic DNA and total RNA were isolated with the Allprep DNA/RNA Mini Kit from Qiagen (Cat no: 80204) and the QIAcube instrument (QIAGEN, Hilden, Germany), according to

Table 1. Patients characteristics.

Characteristics	Control (n = 8)	UC mild (n = 8)	UC severe (n = 8)
Male/Female	5/3	6/2	6/2
Age mean ± SD	54.1 ± 22.3	39.6 ± 15.2	45.1 ± 24.4
TNF-α Level ± SD	4246 ± 1973	8400 ± 3280	31350 ± 26916
Endo Score mean ± SD	0	1.75 ± 0.46	2.38 ± 0.52
Clinical Score ± SD	0	7.75± 1.48	9.75 ± 2.12

SD, standard deviation; TNF, tumour necrosis factor

https://doi.org/10.1371/journal.pone.0248905.t001

the manufacturer's instructions. The quantity and quality of both DNA and RNA were assessed with Qubit 3 and Nanodrop One (Thermo Fisher Scientific, Wilmington, Delaware, USA), respectively. RNA integrity was evaluated with the Experion Automated Electrophoresis System (Bio-Rad, Hercules, CA, USA) and the RNA StdSens Analysis Kit (Bio-Rad, cat no: 700–7103), according to the manufacturer's protocol. All RNA samples used for this analysis had a RIN value between 8.0–10.0. Both DNA and RNA were kept at -70˚C until further use.

Quantitative polymerase chain reaction (qPCR)

Quantitative polymerase chain reaction (qPCR) were used to measure TNF-α mRNA levels in all biopsies. RNA quantity was assessed with NanoVue Plus (GE Healthcare, UK). cDNA synthesis was performed with QuantiTect Reverse Transcription Kit (Qiagen, cat no: 205314), and the QuantiNova Probe PCR Kit (Qiagen, cat no: 208256). CFX Connect Real Time PCR Detection System (Bio-Rad, Hercules, CA, USA) was used for detection. The results were measured in copies/μg. Tissue samples with values <7000 copies/μg are considered non-inflamed, while tissue samples with >7000 copies/μg are considered inflamed [21].

Library preparation and next generation sequencing

DNA libraries were prepared with the SeqCap Epi CpGiant Enrichment Kit (Roche, Switzerland). DNA was bisulfite converted using the EZ DNA Methylation-lightning Kit (Zymo Research, USA, cat no: D5030) prior to the hybridization step and according to the manufacturer's instructions. The amount of input material was 1060 ng of genomic DNA per sample. DNA libraries quality were assessed using the Bioanalyzer 2100, and the Agilent DNA 1000 kit (cat no: 5067–1504, Agilent Technologies, Santa Clara, USA), according to the manufacturer's instructions. DNA libraries generated fragments with an average size of 322 bp. DNA libraries were diluted to 2 nM prior to sequencing. Whole transcriptome libraries were prepared with the TruSeq Stranded Total RNA LT Sample Prep Kit from Illumina (cat no: RS-122-2203). The amount of input material was 1μg of total RNA. The Bioanalyzer 2100 and the Agilent DNA 1000 kit (cat no: 5067–1504, Agilent Technologies, Santa Clara, USA) were used to assess the quality of the RNA libraries. RNA libraries generated fragments with an average size of 301 bp, libraries were normalized to 10 nM and diluted to 4 nM prior to sequencing. Both DNA and RNA libraries were sequenced on the NextSeq 550 instrument, using a high output flow cell 150 cycles (cat no: FC-404-2002, Illumina, USA) and according to the manufacturer's instruction. The libraries were sequenced using paired-end mode.

Data analysis

Base calling, quality scoring and quality check were performed as a first step including quality check on the on-board computer of the NextSeq 550. The data analysis was carried out in the

Bioconductor R framework (www.bioconductor.org). STAR-2.5.2b (https://github.com/alexdobin/STAR) was used to align raw Illumina reads to UCSC genome browser GRCH38p.11 (https://www.ncbi.nlm.nih.gov/grc/human/data). Htseq-count was used for generating the raw gene count matrix [22]. DESeq2 was used to Vst-normalize the gene count matrix [23], and compare severe UC vs mild UC in R (3.5.3) (https://doi.org/10.18129/B9.bioc.DESeq) [24]. Differentially expressed genes (DEGs) between severe UC vs mild UC transcripts were filtered with a read count > 30 and a corrected $p < 0.05$. P-values were corrected for multiple testing using the method of Benjamini and Hochberg [25].

Pathway enrichment was performed using ReactomePA bioconductor packages hypergeometric model (http://bioconductor.org/packages/release/bioc/html/ReactomePA.html). ReactomePA hypergeometric model assesses whether the number of selected genes associated with a reactome pathway is significantly larger than expected. P-values were corrected for multiple testing using the method of Benjamini and Hochberg [25]. Principal component analysis (PCA) of the transcriptome data was performed using the 1000 most variable genes [26]. Genes associated with the risk of IBD were downloaded from the genome-wide association studies (GWAS) catalogue, using the search term IBD (www.ebi.ac.uk/gwas) [27].

For DNA methylation analyses, the Bismark Bisulfite Mapper v0.16.0 (www.bioinformatics.bbsrc.ac.uk/projects/bismark/) was used to align reads to the same aforementioned genome build and calculate methylated and un-methylated DNA positional count matrices. Relative methylation is expressed as a number between 0–1 where 0 means 0% of C's are methylated at that position and 1 means 100% or all C's are methylated. The global methylation analysis mapped included more than 9 million cytosine sites genome- wide. In order to improve interpretation of the dataset, further analysis was restricted to genomic regions within the promoter regions of severe UC compared to mild UC DEGs. Significant differential methylation patterns from above DEGs were found using the globalTest function of the BiSeq Bioconductor package (https://www.bioconductor.org/packages/release/bioc/html/BiSeq.html). Only promoters with a global test p value less than 0.05 where kept. The promoter region was defined as 2000 bp upstream and 200 bp downstream of the transcription start site (TSS). Note that the same patients were used to generate both the methylation and the gene expression data. We could therefore correlate the average promoter relative methylation to the corresponding gene expression. Those promoter/gene pairs with correlations less than -0.6 were kept. A negative correlation occurs when methylation is high, and expression is low or vice versa. Global relative methylation patterns were analysed by principal component analysis (PCA).

Cell populations were estimated by absolute cell deconvolution using the RNA-Seq data. Samples raw counts per million were submitted to the absolute procedure of Monaco. This is a procedure specifically developed for deconvolution of human immune cell types from RNAseq data. Results were merged for T-cells, neutrophils, monocytes, and B-cell types to obtain four main types of immune cell populations [28]. The epithelial and stromal cell fractions were subsequently estimated based on the epithelial cell markers, epithelial cell adhesion molecule (EPCAM), cadherin 17 (CDH17), cadherin 1 (CDH1) and cadherin 18 (CDH18), and the stromal cell markers, endoglin (ENG), thy-1 cell surface antigen (THY1), actin alpha 2, smooth muscle (ACTA2) and collagen type II alpha 1 chain (COL2A1). Cell populations estimates were compared using ANOVA and Tukey's range test [29].

Results and discussion

In this study an integrative epigenome data set, combining genome-wide methylation data and whole-transcriptome data was established in order to gain insight into the molecular mechanisms of severe UC and to explore the epigenetic variation induced by severe

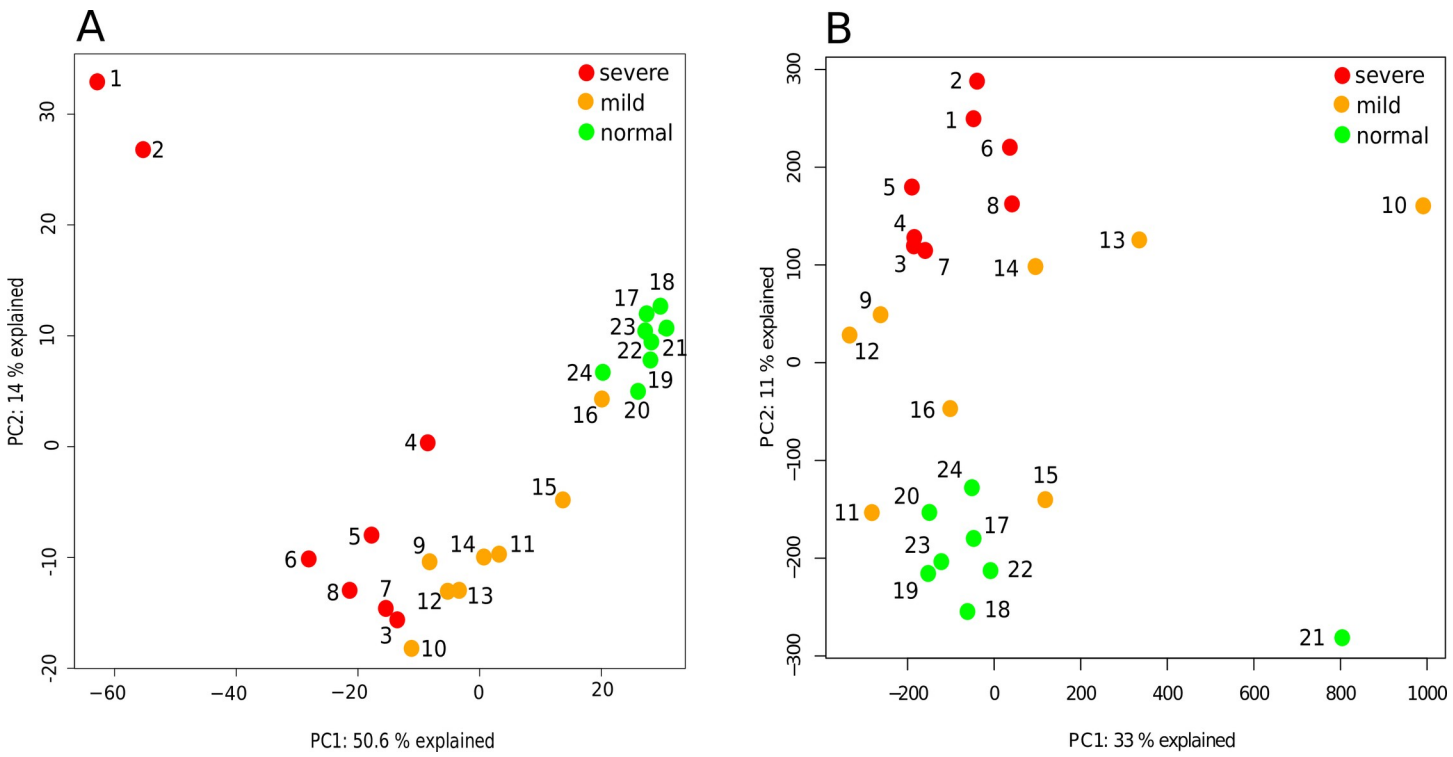


Fig 1. Principal component analysis (PCA). (A) PCA of gene expression data of the thousand most variable genes (transPCA). Unsupervised PCA analysis presenting the difference between severe UC (red, n = 8), mild UC (orange, n = 8) and control (green, n = 8). The first two components explain 51% and 13.5% of the variability in the gene expression data. (B) PCA depicting the global methylation (methPCA) of relative methylation counts (0–100%) for over 9 million cytosine positions including normal (green, n = 8), treatment-naïve mild UC (orange, n = 8) and severe UC (red, n = 8) patient tissue samples. The first two components explain 33% and 11% of the variability in the methylation data.

https://doi.org/10.1371/journal.pone.0248905.g001

inflammation of the colon. The chosen experimental design used in this study was to compare joined transcriptomic and DNA methylation data from each individual patient. This allows for rigorous analysis of the transcriptomic and DNA methylation status of UC patients irrespective of inter-individual differences in environmental or genetic background. In addition, the use of a thoroughly stratified patient group representing only treatment-naïve patients with severe UC for DNA methylation analysis offered a unique opportunity to investigate the DNA methylation state prior to prescription of any medication (Table 1). This is of importance, since UC medications such as immunosuppressive drugs have been shown to have short- and long-term side effects on immune response and can change DNA methylation status [30–33]. Genome-wide DNA methylation in treatment-naïve mild and moderate ulcerative colitis has been reported previously [18, 34]. In this study, we report specific DNA methylation patterns found for treatment-naïve severe UC.

Initial principal component analysis (PCA) revealed a clear separation of severe and mild UC patient phenotypes on both, the transcriptomic- and DNA methylation level (Fig 1). To prevent confusion, the different PCAs discussed in this study are designated transPCA representing transcriptomic data, and methPCA representing DNA methylation data. TransPCA of top thousand most variable differentially expressed genes resulted in a separation of severe and mild UC and control samples along the first principal component (PC1) with 50.6% explained variance, and 14% explained variance along the second principal component (PC2). A complete list of all DEGs is depicted in S1 Table. Differentially expressed IBD susceptibility genes (n = 47) are listed in S2 Table. Two of the UC samples (#1 and #2) in the transPCA separated

from the severe patient sample group, probably indicating a different phenotype of severe UC (Fig 1A). Indeed, these extreme gene expressions may be related to high fractions of neutrophils and monocytes or loss of epithelial cells in these samples (Fig 2 and S4 Table).

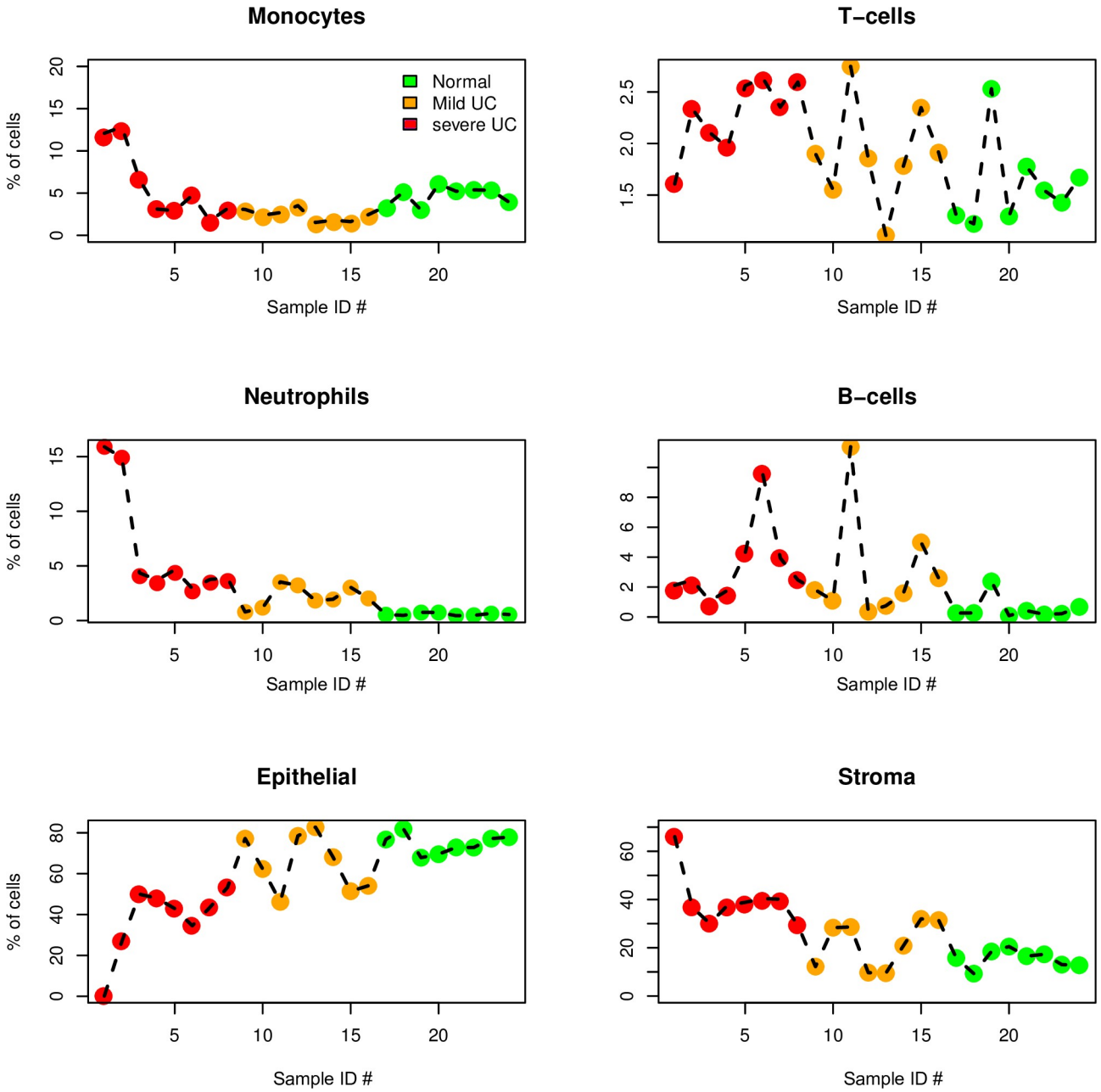


Fig 2. Cell fraction estimation between samples using cell deconvolution algorithm absolute deconvolution of human immune cell types. The fractions of different cell populations in severe and mild UC and control tissue samples were estimated from gene expression data, using absolute cell deconvolution as described in Materials and Methods. The deconvolutions were solved for the following cell types: epithelial cells, monocytes, T cells, neutrophils, B cells, and stroma cells. Each panel shows the estimated percentage of the indicated cell types (y-axis) across all 24 samples sorted according to sample ID numbers (y-axis). For ease of comparison, sample ID numbers are identical to those shown in PCA of methylation and gene expression data (Fig 1). Plot markers are colour coded according to sample group. The fractions of epithelial and stromal cells were estimated from the non-immune cell remainder and the expression levels of the stromal and epithelial marker genes. The epithelial markers (EPCAM, CDH1, CDH17 and CDH18) and stroma markers (ENG, THY1, ACTA2 and COL2A1) were used. Severe UC is indicated by red dots, mild UC is indicated by orange dots, and control is indicated by green dots. Statistical comparison of cell population estimates can be found in S4 Table.

https://doi.org/10.1371/journal.pone.0248905.g002

PCA of global DNA methylation data (methPCA) depicts relative methylation counts [1–100%] for over 9 million cytosine positions for the whole genome of all patient samples, severe UC, mild UC, and normal controls (Fig 1B). The methPCA revealed a distinction between the patient groups along the first component with 33% explained variance. Severe UC samples showed a clear separation from both mild UC and control samples along the second component with an explained variance of 11%. Sample (#10) representing a mild phenotype of UC in the methPCA appeared to be an outlier, as a lower sequencing coverage was observed compared to all the other samples (Fig 1B). This was also the case for one normal sample (#21). These outliers were not removed from the dataset since the transcriptomic data of these samples did not show the same tendency. Further analysis revealed that 34, 8% of all significantly DEGs correlated with DNA methylation. Limiting correlation to $r < -0.6$ resulted in a total of 79 genes of which 77, 2% were hypo-methylated (n = 61) (Table 2 and S3 Table) and 22, 8%, were hyper-methylated (n = 18) (Table 3 and S3 Table). Approximately, 9% of the correlating genes showed DNA methylation at CpG sites in the neighbourhood of the transcription start site (TSS), whereas the remaining 91% of the genes showed methylation at *cis*-acting elements like enhancers and DNAse1 (Fig 3). This is somewhat different for mild UC, where approximately 30% of the genes showed DNA methylation at CpG sites and the remaining 70% of genes showed methylation at *cis*-acting elements [18]. There exists no common opinion on how many methylation sites are necessary for transcription regulation. That's why correlation analysis was applied in this study. These observed changes nevertheless correlated well with expression changes but cannot explain the underlying molecular events that may cause the transcriptional changes [35]. Complete lists of methylated DEGs correlating with transcription and the respective profiles are depicted in S3 Table, S1 Fig and Fig 3.

The major DNA methylation event in treatment-naïve severe UC seems to be hypo-methylation. It is intriguing that approximately 80% of all significant DEGs which correlated to DNA methylation were hypo-methylated in severe UC compared to mild UC (S3 Table). A global hypo-methylation of mucosal DNA in UC compared to normal controls has been reported earlier and it has been suggested that these epigenetic changes in the mucosa might contribute to cancer development [36]. It is well-known that severe inflammation results in an impairment of the epithelial mucosal layer which is followed by diffusion of commensal bacteria and significantly increase of leukocyte infiltration into the gut [37]. This is confirmed by hypo-methylation of leukocyte–specific transcript 1 (LST1), leukocyte associated immunoglobulin-like receptor 1 (LAIR1), sialic acid binding Ig-like lectin 5 (SIGLEC5), and leukocyte surface antigen CD53 (CD53) (Table 1), decreased fractions of epithelial cells and increased fractions of neutrophils, T cells, and monocytes during severe UC compared to mild UC (Fig 2). Fractions of immune cell subtypes, stroma, and epithelial cells on the basis of the gene expression data using cell deconvolution, showed that severe UC differed from mild UC by increased proportions of monocytes ($p = 0.03$) and neutrophiles ($p = 0.02$) and a loss of stroma ($p = 0.001$) and epithelial cells ($p = 0.001$). No significant differences were found between mild UC and normal controls (S4 Table). In addition, pathway enrichment of significantly and differentially DNA methylated genes revealed their involvement in two pathways, neutrophil degranulation, and immuno-regulatory interaction between lymphoid and non-lymphoid cell (Table 4). For both pathways only hypo-methylated genes could be annotated.

Seven IBD susceptibility genes were identified, B-lymphocyte activation marker BLAST1 (CD48), interleukin 10 (IL10), protein tyrosine phosphatase receptor type C (PTPRC), Slam family members (SLAMF7 & SLAMF1), TNF superfamily member 8 (TNFSF8), and docking protein 3 (DOK3) which all were hypo-methylated and up-regulated in severe UC (S2 Table). The hypo-methylation of CD48, IL-10 and PTPRC has not been observed for mild UC [18] and seem to be a specific feature of severe UC. It is interesting to note that only four genes of

Table 2. Hypo-methylated genes in treatment-naïve severe ulcerative colitis (UC).

Gene symbol	log2 FC >1.0 transcription	#c	% methyl	SD methyl
ADGRE3	1,55	82	9,92	25,04
ANGPTL2	1,02	114	12,71	18,31
C3AR1	1,91	27	12,06	23,4
CARD6	1,19	20	25,09	4,86
CASS6	1,13	72	14,72	15,65
CD300A	1,29	94	9,54	23,5
CD300E	3,15	30	12,09	19,97
CD48*	1,06	36	15,17	8,37
CD53	1,41	37	12,03	15,32
CD86	1,21	27	14,53	12,22
CD93	1,60	102	14,18	22,72
CFP	1,37	69	13,67	17,7
CLMP	1,35	213	11,13	16,77
CSF2RB	1,44	59	7,06	25,37
CSF3R	3,42	89	13,10	19,6
CST7	1,79	39	11,59	8,4
CTSK	1,59	41	8,98	17,65
CXCR2	2,42	48	8,67	22,98
DNAH17	1,35	90	6,79	15,57
DOK3*	1,31	107	7,22	21,64
FAM124B	1,15	71	19,16	6,2
GNAI2	1,08	109	5,02	18,99
GPSM3	1,18	54	4,29	35,51
IL10*	1,77	27	8,18	9,56
IL18R1*	1,19	36	6,91	15,51
IL1RN	2,98	67	7,36	18,99
ITGB2	1,23	98	20,99	10,13
ITPRIP	1,34	47	26,95	12,64
LAIR1	1,01	116	9,22	25,66
LILRA1	2,52	85	9,66	13,85
LILRB1	1,59	87	8,00	24,64
LILRB2	1,38	74	25,59	23,24
LINC00877	1,06	94	6,15	16,06
LST1	1,55	40	14,70	21,09
MYO1G	1,31	69	19,29	12,31
NFE2	2,93	40	14,94	9,4
NKG7	1,06	61	13,37	28,59
NLRC4	1,74	27	11,95	12,08
NLRP12	3,45	67	18,19	14,84
NLRP3	1,16	70	17,31	21,24
P2RY13	1,25	15	8,88	21,45
PLEKHO1	1,05	50	21,61	6,5
PPP1R18	1,43	147	11,27	13,71
PTPRC*	1,08	18	13,83	17,35
RHOH	1,10	45	21,01	4,74
SCARF1	1,45	97	19,1	23,68
SELPLG	1,38	78	11,57	6,3

(*Continued*)

Table 2. (Continued)

Gene symbol	log2 FC >1.0 transcription	#c	% methyl	SD methyl
SEMA4A	1,23	47	20,45	8,21
SIGLEC5	2,27	54	8,02	30,56
SLA	1,29	35	20,31	6,25
SLAMF1*	1,29	44	15,47	9,16
SLAMF7*	1,09	21	9,93	31,92
SLAMF8	1,60	46	11,27	20,1
SNX20	1,05	48	22,87	19,76
SPARC	1,76	72	13,47	23,71
SPI1	1,43	41	10,81	26,34
TIE1	1,58	52	9,44	14,43
TNFSF14	1,02	122	13,04	21,64
TNFSF8*	1,16	71	12,16	19,28
TREML2	1,72	48	7,52	25,03
WARS	1,25	31	3,28	9,94

#c indicates number of methylated cytosines; % methyl indicates % difference of DNA methylation severe UC vs. mild UC; SD indicates standard deviation; all results shown with $p < 0.05$.

*indicates IBD susceptibility genes.

https://doi.org/10.1371/journal.pone.0248905.t002

the top DEGs were differentially methylated, colony stimulating factor 3 receptor (CSF3R) and NLR family pyrin domain containing 12 (NLRP12) which are hypo-methylated (Table 2), and transmembrane protein 72 (TMEM72) and UDP glucuronosyltransferase family 1 member A8 (UGT1A8) which are hyper-methylated (Table 3). UGT1A8 has been found to be hyper-

Table 3. Hyper-methylated genes in treatment- naïve severe ulcerative colitis (UC).

Gene symbol	log2 FC >1.0 transcription	#c	% methyl	SD methyl
C2orf82	-1,24	54	-21,84	14,36
C2orf88	-1,22	13	-5,89	12,73
CES2	-1,08	293	-4,39	12,17
DRAIC	-1,43	45	-20,47	8,71
ENTPD5	-1,13	28	-22,12	7,49
MAGIX	-1,13	107	-19,44	12,01
MMP28	-1,32	122	-3,99	20,58
NGEF	-1,11	42	-15,9	14,13
P3H2	-1,26	21	-2,41	17,45
PFKFB2	-1,02	163	-4,59	29,1
PPARGC1A	-1,51	42	-6,09	15,36
PRKG2	-1,42	100	-21,46	11,761
PVRL3	-1,04	30	-22,71	10,86
SLC22A18AS	-1,45	115	-15,49	16,66
SLC51B	-1,17	33	-19,92	8,45
TMEM72	-1,61	58	-14,90	13,3
TRPM4	-1,07	109	-30,69	9,94
UGT1A8	-1,77	31	-13,65	16,15

#c indicates number of methylated cytosines; % methyl indicates % difference of DNA methylation severe UC vs. mild UC; SD indicates standard deviation; all results shown with $p < 0.05$.

https://doi.org/10.1371/journal.pone.0248905.t003

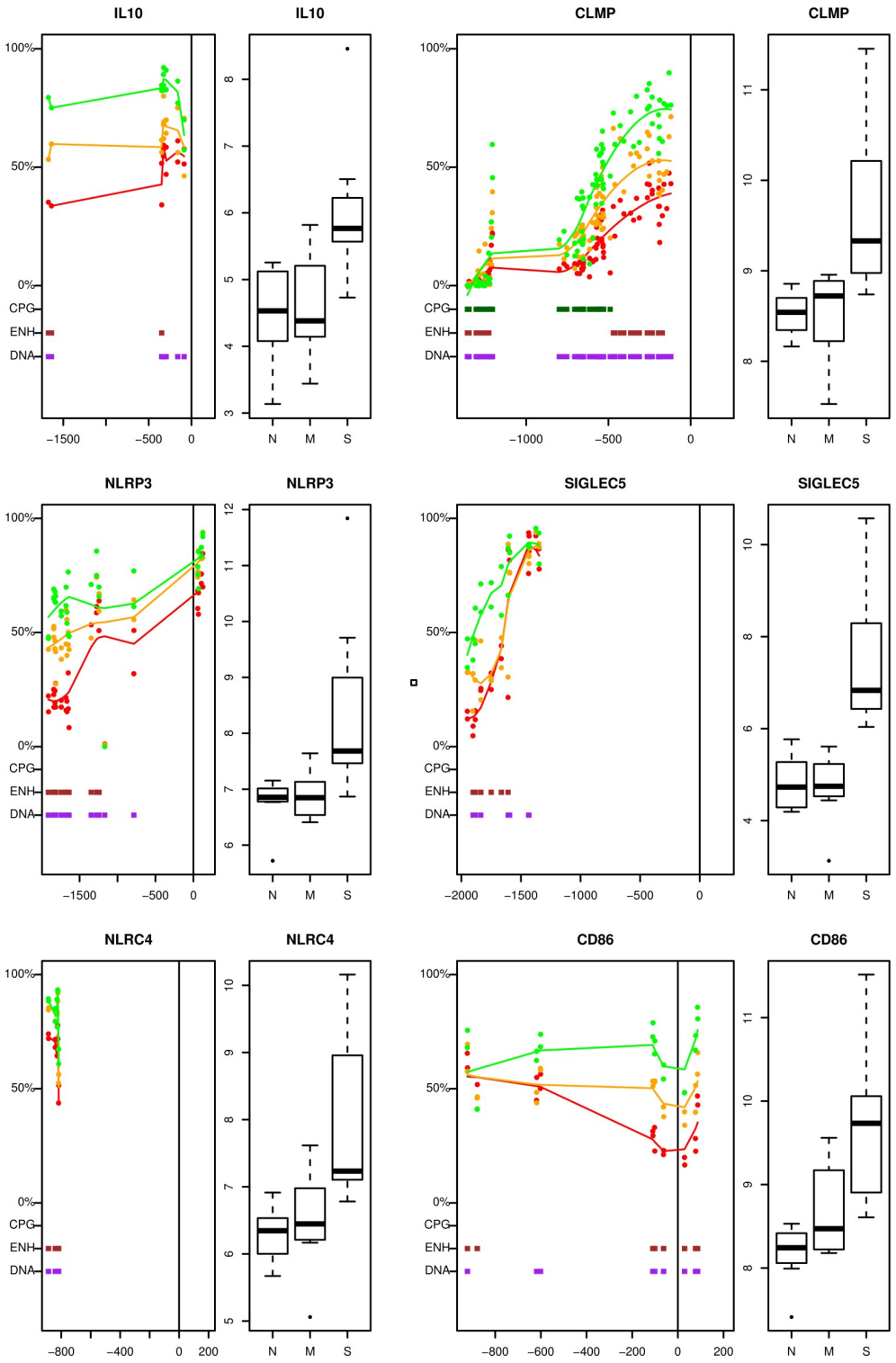


Fig 3. Selection of the most specifically expressed and hypo-methylated genes in severe ulcerative colitis. On the left of each individual illustration the differences in relative methylation levels between normal samples (green), mild UC (orange) and severe UC (red) is shown. Red, green, and orange lines represent the mean relative methylation for severe UC, mild UC and normal samples. The transcription start-site (TSS) is indicated as a vertical line. The x axis is numbered relative to the transcription start site, where minus indicated number of base pairs downstream for TSS (200 bp), and positive number of base pairs upstream from TSS (up to 2000 bp). UCSC genome browser mapped CPG sites (CPG) indicated in dark green, enhancer sites (ENH) indicated in brown, and DNAse1 sites (DNA) indicated in

purple. On the right, boxplots of DESEQ2 log2 normalised values for the gene of interest in normal control (N), mild UC (M) and severe UC (S) are shown. Genes are indicated: interleukin 10 (IL10), CXADR- like membrane protein (CLMP), NLR family pyrin domain containing 3 (NLRP3), sialic acid binding Ig like lectin 5 (SIGLEC5), NLR family CARD domain containing 4 (NLRC4) and T-lymphocyte activation antigen CD86 (CD86).

https://doi.org/10.1371/journal.pone.0248905.g003

methylated in mild UC in an earlier report [18]. A selection of the most specifically methylated DEGs in severe UC clearly show the differences in relative methylation levels for severe UC (S), mild UC (M), and normal samples (N) upstream from the transcription start site (TSS) and indicate UCSC genome browser mapped CpG sites and cis- elements, like enhancers and Dnase1 sites (Fig 3) The boxplots of DESEQ2 log2 normalised values for interleukin 10 (IL10), CXADR- like membrane protein (CLMP), NLR family pyrin domain containing 3 (NLRP3), sialic acid binding Ig like lectin 5 (SIGLEC5), NLR family CARD domain containing 4 (NLRC4) and T-lymphocyte activation antigen CD86 (CD86) showed a clear correlation of DNA methylation status and transcription, thereby clearly indicate specific alterations through hypo-methylation of these genes in severe UC.

The observed hypo-methylation of the NLR family pyrin domain containing NOD-like receptor family members (NLRP3 and NLRP12) in severe UC may maintain intestinal homeostasis and adapt responses against multiple intestinal insults [37–39]. In response to inflammation, hypo-methylation of NRLP inflammasomes may confer anti-inflammatory signals in order to improve severe colitis and to prevent further damage, thereby acting as a defence mechanism to mitigate inflammation. The NLRP3 inflammasome is expressed in both, gut epithelial (IEC) and immune cells (DCs, macrophages, B cells) and may therefore governing the balance of intestinal homeostasis depending on specific cell populations [40–42]. Hypo-methylation of NLRP12 and NLRC4 may regulate gut microbiota in order to supress intestinal inflammation and subsequent intestinal damage in severe UC [43–47]. It is interesting to note that the cassette of NLRs in severe UC is different from those found in mild UC, and that PRRs like Toll-receptors (TLR1, TRL2, TRL4, TRL6, TLR8 and TLR9) are all up-regulated, but not hypo-methylated in severe UC (S1 Table) [17, 18].

A similar interplay between the innate and adaptive immune system can be implied for IL10, a cytokine which has pleiotropic effects in immuno-regulation and inflammation which is expressed and hypo-methylated in severe UC but not in mild UC [17, 18, 48]. IL10 expression during severe UC might counteract excessive inflammatory immune responses by down-regulating the function of antigen presenting cells (APCs), thus providing feedback regulation for pro-inflammatory T cells [49–52]. The increased expression of IL10 produced by T cells may also play a role in mediating tolerance against commensal bacteria, whereas the expression of IL10 in peripheral tissues may lead to down-modulation of the immune response. It has been recently shown that macrophages in the lamina propria preferentially induce IL10 producing cells while DCs promote the generation of Th17 cells [53–55]. It can be therefore believed that hypo-methylation and increased expression of IL10 counteracts severe

Table 4. Reactome enriched pathways of methylated genes in severe ulcerative colitis (UC).

Enriched pathways for severe UC vs. mild UC, $p_{adj.} < 0.05$	Gene symbol
Neutrophil degranulation (innate immune system)	ADGRE3, C3AR1, CD53, CD93, CD300A, CFP, CXCR2, DOK3*, ITGB2, LAIR1, LILRB2, PTPRC*, SIGLEC5
Immuno-regulatory interactions between a Lymphoid and a non-Lymphoid cell (adaptive immune system)	CD300A, CD300E, ITGB2, LAIR1, LILRA1, LILRB1, LILRB2, SIGLEC5, SLAMF7*, TREML2

*indicates IBD susceptibility genes.

https://doi.org/10.1371/journal.pone.0248905.t004

inflammatory signals and aims to dampen severe intestinal inflammation. During severe UC, hypo-methylation of IL10 might also induce tolerogenic DCs that exhibit high expression of co-stimulatory molecules combined with highly expressed inhibitory leucocytes immunoglobulin like receptors (LILRs) and secrete IL10 resulting in the induction of T cells with regulatory capacities (Tregs) [56]. Many of these receptors (LILRA1, LILRB1 and LILRB2) are more hypo-methylated in severe UC than in mild UC (Table 2) [18]. LILRB receptors expressed on immune cells bind to MHC class I molecules on antigen-presenting cells (APCs, DCs) and transduces a negative signal that inhibits stimulation of an immune response. This suggests a role of these receptors in balancing the inflammatory response in face of bacterial infection in severe UC. Although other cells such as macrophages and B cells are also able to present antigens via MHC, DCs are the only cell type to activate naïve T cells and to induce antigen specific immune responses in all adaptive immune cells [57, 58]. An increase of cell fractions in monocytes, neutrophils, T-cells, B-cells and stroma cells were observed in all severe UC samples, whereas the cell fraction of epithelial cells was significantly decreased in all severe UC samples compared with mild UC (Fig 2 and S4 Table).

In concordance with enhanced fractions of T cells in severe UC (Fig 2) increased expression of CD86, a coactivator DC marker involved in T cell activation during microbial infection was observed in severe UC [59]. Other hypo-methylated genes of relevance for the defence of severe inflammation is CXADR-like membrane protein (CLMP) which stabilizes the gut vascular barrier localized between endothelial and epithelial cells in junctional complex involved in cell adhesion and which is required for normal intestinal homeostasis and development (Fig 2 and Table 2) [60].

All the above discussed defence mechanisms might prevent a complete collapse of a functional mucosal barrier during severe inflammation. It is therefore believed that the increase of protective genes and anti-inflammatory pathways induced by hypo-methylation are defence mechanisms, thereby counteracting and alleviating severe inflammation in the gut. Nonetheless, the study is not without limitations, the sample size used here can be considered low due to low number of patients with a severe UC phenotype, but still show sufficiently separation in the PCA (Fig 1). In addition, due to the heterogeneity of the tissue biopsies it is difficult to account NLRP inflammasomes to specific and distinct cell type and single-cell sequencing might overcome this problem. However, the strength of this study lies within the study design where a treatment-naïve patient group with severe UC have been used in order to compare joint transcriptomic and DNA methylation data from each individual patient. This matching of data reduces the chances of introducing influential variable and inter-individual differences and avoid confounding effects of prior medications while highlighting lasting changes to the regulatory patterns underlying the disease that may be of clinical utility.

Conclusion

Hypo-methylation of genes with anti-inflammatory character during severe UC implies a functional interplay between the epithelium and lamina propria to mitigate inflammation in the gut. The specifically DNA hypo-methylated genes found for severe UC can potentially be useful biomarkers for determining disease severity and in the development of new targeted treatment strategies for patients with severe UC.

Supporting information

S1 Fig. Hypo-methylated genes in severe ulcerative colitis. On the left of each individual illustration the differences in relative methylation levels between normal samples (green), mild UC (orange) and severe UC (red) is shown. Red, green and orange lines represent the mean

relative methylation for severe UC, mild UC and normal samples. The transcription start site (TSS) is indicated as a vertical line. The x axis is numbered relative to the transcription start site, where minus indicated number of base pairs downstream for TSS (200 bp), and positive number of base pairs upstream from TSS The regions upstream (up to 2000 bp). UCSC genome browser mapped CPG sites (CPG) indicated in dark green, enhancer sites (ENH) indicated in brown, and DNAse1 sites (DNA) indicated in purple. On the right, boxplots of DESEQ2 log2 normalised values for the gene of interest in normal control (N), mild UC (M) and severe UC (S) are shown. Genes are indicated by the respective gene symbol.
(PDF)

S1 Table. List of significantly differentially expressed genes (DEGs).
(XLSX)

S2 Table. Differentially expressed IBD susceptibility genes in severe UC.
(DOCX)

S3 Table. List of DEGs correlating to DNA methylation with r > - 0.6.
(XLSX)

S4 Table. Comparison of cell population estimates.
(XLSX)

Acknowledgments

The authors would like to thank Ingrid Christiansen for the technical help performing the TNF-α levels measurements and Renate W. Meyer for administrating patient samples.

Author Contributions

Conceptualization: Ruth H. Paulssen.

Data curation: Christopher G. Fenton, Endre Anderssen.

Formal analysis: Hagar Taman.

Funding acquisition: Jon Florholmen, Ruth H. Paulssen.

Investigation: Hagar Taman, Jon Florholmen.

Methodology: Hagar Taman, Christopher G. Fenton, Endre Anderssen, Jon Florholmen, Ruth H. Paulssen.

Project administration: Ruth H. Paulssen.

Resources: Ruth H. Paulssen.

Software: Christopher G. Fenton, Endre Anderssen.

Supervision: Ruth H. Paulssen.

Validation: Ruth H. Paulssen.

Visualization: Christopher G. Fenton, Endre Anderssen.

Writing – original draft: Hagar Taman, Christopher G. Fenton, Ruth H. Paulssen.

Writing – review & editing: Christopher G. Fenton, Endre Anderssen, Jon Florholmen, Ruth H. Paulssen.

References

1. Khor B, Gardet A, Xavier RJ. Genetics and pathogenesis of inflammatory bowel disease. Nature. 2011; 474: 307–317. https://doi.org/10.1038/nature10209 PMID: 21677747
2. Jain S, Ahuja V, Limdi JK. Optimal management of acute severe ulcerative colitis. Postgrad Med J. 2019/01/14. 2019; 95: 32–40. https://doi.org/10.1136/postgradmedj-2018-136072 PMID: 30636193
3. Dulai PS, Jairath V. Acute severe ulcerative colitis: latest evidence and therapeutic implications. Ther Adv Chronic Dis. 2018/02/02. 2018; 9: 65–72. https://doi.org/10.1177/2040622317742095 PMID: 29387331
4. Fernandes SR, Santos P, Miguel Moura C, Marques da Costa P, Carvalho JR, Isabel Valente A, et al. The use of a segmental endoscopic score may improve the prediction of clinical outcomes in acute severe ulcerative colitis. Rev Esp Enferm Dig. 2016/10/06. 2016; 108: 697–702. https://doi.org/10.17235/reed.2016.4470/2016 PMID: 27701881
5. Hindryckx P, Jairath V, D'Haens G. Acute severe ulcerative colitis: from pathophysiology to clinical management. Nat Rev Gastroenterol Hepatol. 2016/10/26. 2016; 13: 654–664. https://doi.org/10.1038/nrgastro.2016.116 PMID: 27580687
6. Choi JK, Kim DW, Shin SY, Park EC, Kang JG. Effect of Ulcerative Colitis on Incidence of Colorectal Cancer: Results from the Nationwide Population-Based Cohort Study (2003–2013). J Cancer. 2016/04/15. 2016; 7: 681–686. https://doi.org/10.7150/jca.14264 PMID: 27076849
7. Choy MC, Seah D, Faleck DM, Shah SC, Chao CY, An YK, et al. Systematic Review and Meta-analysis: Optimal Salvage Therapy in Acute Severe Ulcerative Colitis. Inflamm Bowel Dis. 2019/01/04. 2019. https://doi.org/10.1093/ibd/izy383 PMID: 30605549
8. Moustafa A, Li W, Anderson EL, Wong EHM, Dulai PS, Sandborn WJ, et al. Genetic risk, dysbiosis, and treatment stratification using host genome and gut microbiome in inflammatory bowel disease. Clin Transl Gastroenterol. 2018/01/19. 2018; 9: e132. https://doi.org/10.1038/ctg.2017.58 PMID: 29345635
9. Jostins L, Ripke S, Weersma RK, Duerr RH, McGovern DP, Hui KY, et al. Host–microbe interactions have shaped the genetic architecture of inflammatory bowel disease. Nature. 2012; 491: 119–124. https://doi.org/10.1038/nature11582 PMID: 23128233
10. Harris RA, Nagy-Szakal D, Pedersen N, Opekun A, Bronsky J, Munkholm P, et al. Genome-wide peripheral blood leukocyte DNA methylation microarrays identified a single association with inflammatory bowel diseases. Inflamm Bowel Dis. 2012; 18: 2334–2341. https://doi.org/10.1002/ibd.22956 PMID: 22467598
11. Hasler R, Feng Z, Backdahl L, Spehlmann ME, Franke A, Teschendorff A, et al. A functional methylome map of ulcerative colitis. Genome Res. 2012/07/25. 2012; 22: 2130–2137. https://doi.org/10.1101/gr.138347.112 PMID: 22826509
12. Karatzas PS, Gazouli M, Safioleas M, Mantzaris GJ. DNA methylation changes in inflammatory bowel disease. Ann Gastroenterol. 2014/04/16. 2014; 27: 125–132. Available: https://www.ncbi.nlm.nih.gov/pubmed/24733658 PMID: 24733658
13. McDermott E, Ryan EJ, Tosetto M, Gibson D, Burrage J, Keegan D, et al. DNA Methylation Profiling in Inflammatory Bowel Disease Provides New Insights into Disease Pathogenesis. J Crohns Colitis. 2016; 10: 77–86. https://doi.org/10.1093/ecco-jcc/jjv176 PMID: 26419460
14. Harris RA, Wang T, Coarfa C, Nagarajan RP, Hong C, Downey SL, et al. Comparison of sequencing-based methods to profile DNA methylation and identification of monoallelic epigenetic modifications. Nat Biotechnol. 2010/09/21. 2010; 28: 1097–1105. https://doi.org/10.1038/nbt.1682 PMID: 20852635
15. Deaton AM, Bird A. CpG islands and the regulation of transcription. Genes Dev. 2011/05/18. 2011; 25: 1010–1022. https://doi.org/10.1101/gad.2037511 PMID: 21576262
16. Sheaffer KL, Elliott EN, Kaestner KH. DNA hypomethylation contributes to genomic instability and intestinal cancer initiation. Cancer Prevention Research. 2016; 9: 534–546. https://doi.org/10.1158/1940-6207.CAPR-15-0349 PMID: 26883721
17. Taman H, Fenton CG, Hensel I V., Anderssen E, Florholmen J, Paulssen RH. Transcriptomic landscape of treatment-naïve ulcerative colitis. Journal of Crohn's and Colitis. 2018; 12: 327–336. https://doi.org/10.1093/ecco-jcc/jjx139 PMID: 29040430
18. Taman H, Fenton CG, Hensel I V., Anderssen E, Florholmen J, Paulssen RH. Genome-wide DNA methylation intreatment-naïve ulcerative colitis. Journal of Crohn's and Colitis. 2018; 12: 1338–1347. https://doi.org/10.1093/ecco-jcc/jjy117 PMID: 30137272
19. Magro F, Gionchetti P, Eliakim R, Ardizzone S, Armuzzi A, Barreiro-de Acosta M, et al. Third European Evidence-based Consensus on Diagnosis and Management of Ulcerative Colitis. Part 1: Definitions, Diagnosis, Extra-intestinal Manifestations, Pregnancy, Cancer Surveillance, Surgery, and Ileo-anal Pouch Disorders. J Crohns Colitis. 2017/02/06. 2017; 11: 649–670. https://doi.org/10.1093/ecco-jcc/jjx008 PMID: 28158501

20. Tursi A, Brandimarte G, Papa A, Giglio A, Elisei W, Giorgetti GM, et al. Treatment of Relapsing Mild-to-Moderate Ulcerative Colitis With the Probiotic VSL#3 as Adjunctive to a Standard Pharmaceutical Treatment: A Double-Blind, Randomized, Placebo-Controlled Study. American Journal of Gastroenterology. 2010; 105: 2218–2227. https://doi.org/10.1038/ajg.2010.218 PMID: 20517305

21. Olsen T, Goll R, Cui G, Husebekk A, Vonen B, Birketvedt GS, et al. Tissue levels of tumor necrosis factor-alpha correlates with grade of inflammation in untreated ulcerative colitis. Scand J Gastroenterol. 2007; 42: 1312–1320. https://doi.org/10.1080/00365520701409035 PMID: 17852866

22. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. STAR: ultrafast universal RNA-seq aligner. Bioinformatics. 2012/10/30. 2013; 29: 15–21. https://doi.org/10.1093/bioinformatics/bts635 PMID: 23104886

23. Anders S, Huber W. Differential expression analysis for sequence count data. Genome Biology. 2010; 11. https://doi.org/10.1186/gb-2010-11-10-r106 PMID: 20979621

24. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biology. 2014; 15. https://doi.org/10.1186/s13059-014-0550-8 PMID: 25516281

25. Benjamini Y, Hochberg Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. Journal of the Royal Statistical Society. 1995; 57: 12.

26. Pearson K. On Lines and Planes of Closest Fit to Systems of Points in Space. Philosophical Magazine. 1901; 2: 14. https://doi.org/10.1080/14786440109462720

27. MacArthur J, Bowler E, Cerezo M, Gil L, Hall P, Hastings E, et al. The new NHGRI-EBI Catalog of published genome-wide association studies (GWAS Catalog). Nucleic Acids Res. 2016/12/03. 2017; 45: D896–D901. https://doi.org/10.1093/nar/gkw1133 PMID: 27899670

28. Monaco G, Lee B, Xu W, Mustafah S, Hwang YY, Carré C, et al. RNA-Seq Signatures Normalized by mRNA Abundance Allow Absolute Deconvolution of Human Immune Cell Types. Cell Reports. 2019; 26: 1627–1640.e7. https://doi.org/10.1016/j.celrep.2019.01.041 PMID: 30726743

29. Tukey JW. Comparing Individual Means in the Analysis of Variance. Biometrics. 1949; 5: 99. https://doi.org/10.2307/3001913 PMID: 18151955

30. Tahara T, Shibata T, Nakamura M, Yamashita H, Yoshioka D, Okubo M, et al. Chronic aspirin use suppresses CDH1 methylation in human gastric mucosa. Dig Dis Sci. 2009/02/03. 2010; 55: 54–59. https://doi.org/10.1007/s10620-008-0701-4 PMID: 19184424

31. Tahara T, Shibata T, Yamashita H, Nakamura M, Yoshioka D, Okubo M, et al. Chronic nonsteroidal anti-inflammatory drug (NSAID) use suppresses multiple CpG islands hyper methylation (CIHM) of tumor suppressor genes in the human gastric mucosa. Cancer Sci. 2009/05/13. 2009; 100: 1192–1197. https://doi.org/10.1111/j.1349-7006.2009.01175.x PMID: 19432886

32. Peters FS, Peeters AMA, Mandaviya PR, van Meurs JBJ, Hofland LJ, van de Wetering J, et al. Differentially methylated regions in T cells identify kidney transplant patients at risk for de novo skin cancer. Clin Epigenetics. 2018/06/28. 2018; 10: 81. https://doi.org/10.1186/s13148-018-0519-7 PMID: 29946375

33. De Vries LCS, Wildenberg ME, De Jonge WJ, D'Haens GR. The Future of Janus Kinase Inhibitors in Inflammatory Bowel Disease. J Crohns Colitis. 2017/02/06. 2017; 11: 885–893. https://doi.org/10.1093/ecco-jcc/jjx003 PMID: 28158411

34. Paulssen RH, Taman H, Fenton CG, Hensel I V, Anderssen E, Florholmen J. Novel Regulatory DNA Methylation Features of Relevance for Ulcerative Colitis. American Journal of Gastroenterology. 2018; 113: P-050. Available: https://www.nature.com/articles/ajg2017452#p050-novel-regulatory-dna-methylation-features-of-relevance-for-ulcerative-colitis

35. Schultz MD, He Y, Whitaker JW, Hariharan M, Mukamel EA, Leung D, et al. Human body epigenome maps reveal noncanonical DNA methylation variation. Nature. 2015; 523: 212–216. https://doi.org/10.1038/nature14465 PMID: 26030523

36. Glória L, Cravo M, Pinto A, de Sousa LS, Chaves P, Leitão CN, et al. DNA hypomethylation and proliferative activity are increased in the rectal mucosa of patients with long-standing ulcerative colitis. Cancer. 1996/12/01. 1996; 78: 2300–2306. https://doi.org/10.1002/(sici)1097-0142(19961201)78:11<2300::aid-cncr5>3.0.co;2-q PMID: 8940998

37. Zmora N, Levy M, Pevsner-Fishcer M, Elinav E. Inflammasomes and intestinal inflammation. Mucosal Immunol. 2017/04/13. 2017; 10: 865–883. https://doi.org/10.1038/mi.2017.19 PMID: 28401932

38. Hirota SA, Ng J, Lueng A, Khajah M, Parhar K, Li Y, et al. NLRP3 inflammasome plays a key role in the regulation of intestinal homeostasis. Inflamm Bowel Dis. 2010/09/28. 2011; 17: 1359–1372. https://doi.org/10.1002/ibd.21478 PMID: 20872834

39. Davis BK, Philipson C, Hontecillas R, Eden K, Bassaganya-Riera J, Allen IC. Emerging significance of NLRs in inflammatory bowel disease. Inflamm Bowel Dis. 2014/08/26. 2014; 20: 2412–2432. https://doi.org/10.1097/MIB.0000000000000151 PMID: 25153506

40. Lissner D, Siegmund B. The multifaceted role of the inflammasome in inflammatory bowel diseases. ScientificWorldJournal. 2011/08/02. 2011; 11: 1536–1547. https://doi.org/10.1100/tsw.2011.139 PMID: 21805022

41. Ringel-Scaia VM, McDaniel DK, Allen IC. The Goldilocks Conundrum: NLR Inflammasome Modulation of Gastrointestinal Inflammation during Inflammatory Bowel Disease. Crit Rev Immunol. 2016/01/01. 2016; 36: 283–314. https://doi.org/10.1615/CritRevImmunol.2017019158 PMID: 28322135

42. Zhen Y, Zhang H. NLRP3 inflammasome and inflammatory bowel disease. Frontiers in Immunology. 2019/03/16. 2019; 10: 276. https://doi.org/10.3389/fimmu.2019.00276 PMID: 30873162

43. Nordlander S, Pott J, Maloy KJ. NLRC4 expression in intestinal epithelial cells mediates protection against an enteric pathogen. Mucosal Immunol. 2013/11/28. 2014; 7: 775–785. https://doi.org/10.1038/mi.2013.95 PMID: 24280936

44. Chen L, Wilson JE, Koenigsknecht MJ, Chou WC, Montgomery SA, Truax AD, et al. NLRP12 attenuates colon inflammation by maintaining colonic microbial diversity and promoting protective commensal bacterial growth. Nat Immunol. 2017/03/14. 2017; 18: 541–551. https://doi.org/10.1038/ni.3690 PMID: 28288099

45. Franchi L, Munoz-Planillo R, Nunez G. Sensing and reacting to microbes through the inflammasomes. Nat Immunol. 2012/03/21. 2012; 13: 325–332. https://doi.org/10.1038/ni.2231 PMID: 22430785

46. Miao EA, Mao DP, Yudkovsky N, Bonneau R, Lorang CG, Warren SE, et al. Innate immune detection of the type III secretion apparatus through the NLRC4 inflammasome. Proc Natl Acad Sci U S A. 2010/02/06. 2010; 107: 3076–3080. https://doi.org/10.1073/pnas.0913087107 PMID: 20133635

47. Miao EA, Alpuche-Aranda CM, Dors M, Clark AE, Bader MW, Miller SI, et al. Cytoplasmic flagellin activates caspase-1 and secretion of interleukin 1beta via Ipaf. Nat Immunol. 2006/05/02. 2006; 7: 569–575. https://doi.org/10.1038/ni1344 PMID: 16648853

48. Engelhardt KR, Grimbacher B. IL-10 in humans: lessons from the gut, IL-10/IL-10 receptor deficiencies, and IL-10 polymorphisms. Curr Top Microbiol Immunol. 2014/07/10. 2014; 380: 1–18. https://doi.org/10.1007/978-3-662-43492-5_1 PMID: 25004811

49. Moore KW, de Waal Malefyt R, Coffman RL, O'Garra A. Interleukin-10 and the interleukin-10 receptor. Annu Rev Immunol. 2001/03/13. 2001; 19: 683–765. https://doi.org/10.1146/annurev.immunol.19.1.683 PMID: 11244051

50. Huber S, Gagliani N, Esplugues E, O'Connor W Jr., Huber FJ, Chaudhry A, et al. Th17 cells express interleukin-10 receptor and are controlled by Foxp3(-) and Foxp3+ regulatory CD4+ T cells in an interleukin-10-dependent manner. Immunity. 2011/04/23. 2011; 34: 554–565. https://doi.org/10.1016/j.immuni.2011.01.020 PMID: 21511184

51. Chaudhry A, Samstein RM, Treuting P, Liang Y, Pils MC, Heinrich JM, et al. Interleukin-10 signaling in regulatory T cells is required for suppression of Th17 cell-mediated inflammation. Immunity. 2011/04/23. 2011; 34: 566–578. https://doi.org/10.1016/j.immuni.2011.03.018 PMID: 21511185

52. O'Garra A, Murphy KM. From IL-10 to IL-12: how pathogens and their products stimulate APCs to induce T(H)1 development. Nat Immunol. 2009/08/21. 2009; 10: 929–932. https://doi.org/10.1038/ni0909-929 PMID: 19692989

53. Izcue A, Hue S, Buonocore S, Arancibia-Carcamo C V, Ahern PP, Iwakura Y, et al. Interleukin-23 restrains regulatory T cell activity to drive T cell-dependent colitis. Immunity. 2008/04/11. 2008; 28: 559–570. https://doi.org/10.1016/j.immuni.2008.02.019 PMID: 18400195

54. Makita S, Kanai T, Oshima S, Uraushihara K, Totsuka T, Sawada T, et al. CD4+CD25bright T cells in human intestinal lamina propria as regulatory cells. J Immunol. 2004/08/24. 2004; 173: 3119–3130. https://doi.org/10.4049/jimmunol.173.5.3119 PMID: 15322172

55. Mowat AM, Bain CC. Mucosal macrophages in intestinal homeostasis and inflammation. J Innate Immun. 2011/10/26. 2011; 3: 550–564. https://doi.org/10.1159/000329099 PMID: 22025201

56. Brown DP, Jones DC, Anderson KJ, Lapaque N, Buerki RA, Trowsdale J, et al. The inhibitory receptor LILRB4 (ILT3) modulates antigen presenting cell phenotype and, along with LILRB2 (ILT4), is upregulated in response to Salmonella infection. BMC Immunol. 2009/10/29. 2009; 10: 56. https://doi.org/10.1186/1471-2172-10-56 PMID: 19860908

57. Wilson NS, El-Sukkari D, Villadangos JA. Dendritic cells constitutively present self antigens in their immature state in vivo and regulate antigen presentation by controlling the rates of MHC class II synthesis and endocytosis. Blood. 2003/11/08. 2004; 103: 2187–2195. https://doi.org/10.1182/blood-2003-08-2729 PMID: 14604956

58. Lei-Leston AC, Murphy AG, Maloy KJ. Epithelial Cell Inflammasomes in Intestinal Immunity and Inflammation. Front Immunol. 2017/10/06. 2017; 8: 1168. https://doi.org/10.3389/fimmu.2017.01168 PMID: 28979266

59. Corinti S, Albanesi C, la Sala A, Pastore S, Girolomoni G. Regulatory activity of autocrine IL-10 on dendritic cell functions. J Immunol. 2001/03/20. 2001; 166: 4312–4318. https://doi.org/10.4049/jimmunol.166.7.4312 PMID: 11254683

60. López-Posadas R, Stürzl M, Atreya I, Neurath MF, Britzen-Laurent N. Interplay of GTPases and Cytoskeleton in Cellular Barrier Defects during Gut Inflammation. Frontiers in immunology. 2017; 8: 1240. https://doi.org/10.3389/fimmu.2017.01240 PMID: 29051760