UiT The Arctic University of Norway

Faculty of Health Science

Novel Molecular Characteristics of Ulcerative Colitis

Wei Meng

A dissertation for the degree of Doctor of Philosophy (PhD) - February 2024

Table of Contents

Acknowledgements

The project was conducted in the Clinical Bioinformatics Research Group at the Department of Clinical Medicine, Faculty of Health Sciences at UiT-The Arctic University of Norway. This PhD project was carried out from October 2019 to November 2023.

This project would not have been finished without the extraordinary guidance of my main supervisor, Prof. Ruth Paulssen. I thank her for her support and understanding during the PhD period. Her supervisions and instructions were priceless when I was lost in the project. The words from her were very warm, straightforward, and inspiring when I missed the target. Without her care and concern, my PhD journey would be drowning in the pool. I am also grateful for the extraordinary help from my co-supervisor, Christopher Graham Fenton. His practical experience helped immensely during my PhD journey. I want to give a special thanks to Kay-Martin Johnsen as my co-supervisor. Without his data on patients and professional view of the clinics, this project would not have been possible.

I would like to thank all my friends and colleagues in the group, Endre Anderssen, Rafi Ahmad, Amrinder Singh, Mithlesh Ray, Hagar Taman, Lars Daniel Håland, and Line Strand Karlstrøm for helping me with the experiments and developing research ideas. I benefited a lot from their academic discussions and life experiences.

At last, I thank my family. They provided support to my life to let me reach the goal of finishing this journey sweetly. My wife Wei Li, and my daughter Ao-Lei always cheered me up when I was down. Thanks to my parents, who gave me the support and encouragement to finish my PhD. Thanks to all the friends that I made and kept me active.

Summary

Inflammatory bowel disease (IBD) is a chronic autoimmune inflammatory disease consisting of two major types: Ulcerative colitis (UC) and Crohn's disease (CD). UC is limited to the colon in contrast to Crohn's disease affecting any area in the gastrointestinal tract. Both UC and CD are characterized by a relapsing and remitting time course. The pathogenesis of UC involves a complex interplay of genetic susceptibility, environmental factors, and immune system dysregulation. The complex relationship between UC remission and relapse in terms of genetic alterations, epigenetic modifications, and non-coding RNAs (ncRNAs) is not fully understood.

Current treatment strategies are hampered by the lack of the ability to predict relapse in UC patients. Clinical diagnostic tools are frequently limited by visible disease manifestations, which are often absent in remission patients. Therefore, the molecular characterization of remission is needed.

Whole transcriptome sequencing (RNA-seq) and whole genome bisulfite sequencing (WGBS) data from mucosal biopsies have been used to explore the molecular landscape in remission and active UC. A group of mitochondrial RNAs and snoRNAs was identified that may be able to predict the duration of remission in UC. Several remission-specific genes were identified involving pro- and anti-inflammatory pathways whose expression may be under the control of methylation. Several long non-coding RNAs (lncRNAs) were discovered that may be under the regulation of methylation and involved in the inflammatory immune response.

The results may provide an understanding of UC pathogenesis and potential diagnostic markers that may contribute to disease management for remission patients.

Abbreviations

5-ASA: 5-Aminosalicylic Acid

6-MP: 6-Mercaptopurine

AMPs: Antimicrobial Peptides

APC: Antigen-presenting cells

BS-seq: Bisulfite Conversion sequencing

CD: Crohn's disease

cDNA: complementary DNA

CRP: C-Reactive Protein

CRC: Colorectal Cancer

DC: Dendritic Cells

DEGs: Differentially Expressed Genes

DMRs: Differentially Methylated Regions

ECCO: European Crohn's and Colitis Organisation

EWAS: Epigenome-wide association study

FMT: Faecal Microbiota Transplantation

GO: Gene Ontology

GWAS: Genome-wide association study

HLA: Human Leucocyte Antigen

IBD: Inflammatory Bowel Disease

IFNγ: Interferon Gamma

IL: Interleukin

IM: Immunomodulators

IOIBD: International Organization for the Study of Inflammatory Bowel Disease

lncRNAs: Long non-coding Ribonucleic Acid

MES: Mayo Endoscopic Subscore

MTX: Methotrexate

NN: Controls

PLS: Partial Least Square regression

PGE-MUM: Prostaglandin E-major Urinary Metabolite

PRO: Patient-Reported Outcome

PUCAI: Paediatric Ulcerative Colitis Activity Index

RNA-seq: RNA sequencing

RM: Remission patients

RRBS: Reduced Representation Bisulfite Sequencing

SNP: Single Nucleotide Polymorphism

TF: Transcription Factor

TJs: Tight Junctions

TNF-α: Tumour Necrosis Factor-alpha

Tregs: Regulatory T cells

UC: Ulcerative colitis

WGBS: Whole Genome Bisulfite Sequencing

List of papers

Paper-I

Anti-apoptotic genes and non-coding RNAs are potential outcome predictors for ulcerative colitis

Wei Meng, Kay-Martin Johnsen, Christopher G Fenton, Jon Florholmen, Ruth H Paulssen

Functional Integrative Genomics, 2023 May 18;23(2):165.

https://doi.org/10.1007/s10142-023-01099-9

Paper-II

DNA methylation fine-tunes pro-and anti-inflammatory signalling pathways in inactive ulcerative colitis

Wei Meng, Christopher G Fenton, Kay-Martin Johnsen, Hagar Taman, Jon Florholmen, and Ruth H Paulssen

Scientific reports, 2024 March, 14, 6789

https://doi.org/10.1038/s41598-024-57440-0

Paper-III

Methylation-Regulated Long Non-Coding RNA Expression in Ulcerative Colitis

Christopher G. Fenton, Mithlesh Kumar Ray, Wei Meng, and Ruth H. Paulssen

International Journal of Molecular Sciences. 2023, 24(13), 10500.

https://doi.org/10.3390/ijms241310500

1 Introduction

1.1 Inflammatory Bowel Disease

Inflammatory Bowel disease (IBD) is a chronic inflammatory disease with a relapsing-remitting course. IBD comprises a group of disorders that cause inflammation in the gastrointestinal tract [1]. There are two major types of IBD: Ulcerative colitis (UC) and Crohn's disease (CD), with some differences in their pathophysiologies [2]. The exact causes of IBD have not been fully understood. UC and CD have similarities in clinical manifestations. UC is characterized by chronic inflammation primarily in the colon while CD affects the whole gastrointestinal tract. Symptoms associated with IBD include diarrhoea, rectal bleeding, weight loss, and abdominal pain. IBD affected 6.8 million people in 2017, and the number is predicted to keep rising due to an increasing prevalence [3,4]. Although the global age-standardized death rate from IBD decreased from 1990 to 2017 due to recent treatment advances [5], UC patients still have an increased risk of developing colorectal cancer (CRC) [6,7].

1.1.1 Ulcerative Colitis

UC is limited to the mucosal layer of the colon and the rectum [8]. The cause of UC is believed to be a result of genetic susceptibility and environmental risk factors including lifestyle choices, stress, and gut microbiota alterations. The interplay between these factors can compromise the immune response in the gastrointestinal tract and eventually lead to chronic inflammation.

1.1.2 Management of UC

After diagnosis of UC, medical treatment and clinical management are often required. The treatment decision follows a step-by-step procedure as depicted in **Figure 1**, with the consideration of the treatment response and severity. 5-aminosalicylic (5-ASA) is commonly used as the first-line treatment for UC patients with mild to moderate disease. 5-ASA treatment may reduce the risk of CRC [9]. However, only 40%-70% of patients respond to 5-ASA treatment [10]. Patients that do not respond to 5-ASA often receive steroids. For mild to

moderate UC, a combination of 5-ASA and corticosteroids is used, especially for left-sided and extensive UC [11]. 5-ASA is also commonly used for maintenance therapy for remission patients [12].

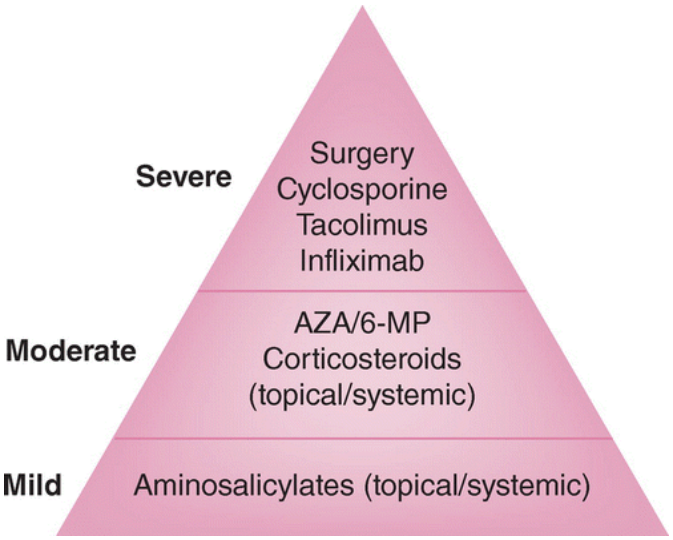


Figure 1. *The treatment pyramid after diagnosis of UC. With permission from [13].*

When the treatment with corticosteroids is not successful, immunomodulators (IMs) are used. IMs including azathioprine, 6-mercaptopurine(6-MP), and methotrexate suppress the immune system [14]. However, the effect of IMs has not been carefully evaluated. Evidence indicates that prolonged low doses of methotrexate could lead to chronic liver disease, fibrosis, and cirrhosis [15,16]. It has been reported that immune suppression may be associated with a higher rate of adverse events such as increased risk of cancers [17]. With sparse data available, the efficacy of IMs as a treatment option is still elusive [18].

Recently, the biological agent anti-TNFα (e.g., infliximab) has been used for patients with moderate-to-severe UC or patients who do not respond to conventional therapies like 5-ASA, corticosteroids, or IMs [19]. 40% of UC patients achieved remission through the combination of anti-TNF-α with IMs [11]. When patients do not respond to treatments or have severe complications, surgery such as colectomy or proctocolectomy is considered [20].

1.2 Remission

Remission is a term to describe UC in the absence of inflammation in the colonic mucosa. Three definitions of remission are widely used, namely clinical remission, endoscopic remission, and histological remission. These definitions are based on different criteria and may be used in combination to determine whether a patient has achieved remission.

Clinical remission includes the cessation of rectal bleeding and normal stool frequency [21]. Diagnosis of clinical remission is often relying on patient-reported outcomes and endoscopy scores [22]. Endoscopic remission is defined as the absence of inflammation in the colon during endoscopy. The most widely used clinical scoring system is the Mayo score [23]. The evaluation of histological remission is controversial in terms of criteria and methods [23]. Histological activity such as crypt structure for defining remission was used in studies focused on incorporating more objective measures of disease activity [24]. Currently, the Geboes score is a widely used histological method for evaluating inflammation in UC [25,26]. Achieving histological remission is challenging in clinical studies because of invasive nature of biopsies taking [23]. The term “mucosal healing” is commonly used for the remission assessment with only endoscopic results [27]. In some studies, the combination of mucosal healing and clinical remission is described as deep remission, which is associated with better clinical outcomes for UC patients [28].

1.2.1 Classification and Diagnosis of UC

To ensure effective treatment strategies for UC, it is important to have a detailed system for disease classification. Anatomically, UC is located in various parts of the colon (**Figure 2**). Based on disease severity, UC is classified into mild, moderate, or severe disease.

Additional endoscopic, histologic and laboratory tests are needed to diagnose UC and to define the inflammatory status [23,29]. Biomarkers including calprotectin, faecal C-reactive protein (CRP), and prostaglandin E-major urinary metabolite (PGE-MUM), are extensively used as diagnostic indicators but are not sensitive or specific to UC [30,31]. Currently, it is not possible to diagnose UC using a single biomarker [30].

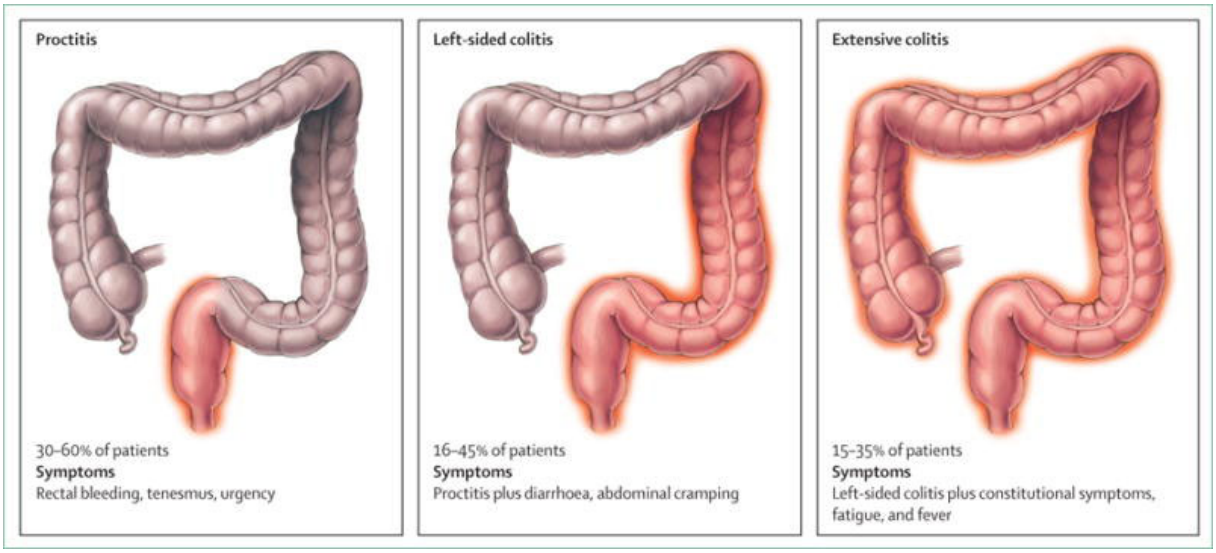


Figure 2. *Ulcerative Colitis phenotypes. With permission from [32].*

1.2.2 Clinical Outcome

The efficacy of the different treatments is not satisfactory as up to 40% of severe UC patients do not respond to anti-TNF-α treatment [33]. Anti-TNF-α agents in conjunction with IMs have demonstrated marked improvements in both response and remission rates in UC [34]. Of note, these therapies have not only facilitated prolonged mucosal healing and deep remission but also minimized the potential for future complications [35]. However, despite a general decline in the risk of colectomy, clinical data show that emergency colectomy rates have not changed over time [36]. Patients in remission often relapse upon discontinuation of treatment due to side effects [28,37]. Discontinuation of anti-TNF-α treatment results in relapse rates of 30-40% within the first year and exceeds 50% beyond two years [34]. Patients in deep remission have a higher likelihood of maintaining remission at a 12-month follow-up [28]. Patients achieving histological remission have a minimal risk of relapse [38]. The European Crohn's and Colitis Organisation (ECCO) is proposing recommendations for treatment options to achieve remission [39].

1.3 Pathogenesis of UC

The pathogenesis of UC includes a complex interplay between genetic, environmental, microbial, and immunological factors. These factors collectively contribute to the development and progression of UC [40].

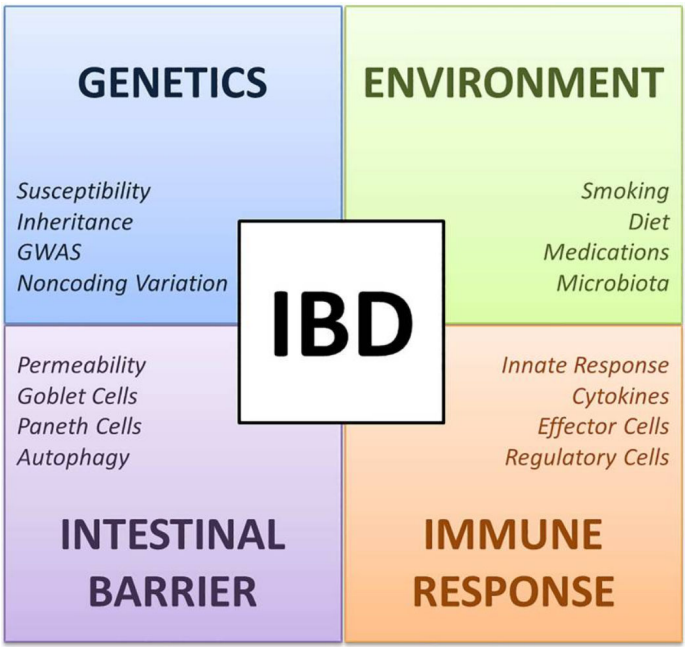


***Figure 3**. The factors that contribute to IBD pathogenesis. With permission from [1].*

1.3.1 Genetics

Individuals with a family history of UC have an increased risk of developing the disease compared to the general population [41]. Yet, genetic risk factors explain only a small fraction (7.5–22%) of UC prevalence [42,43]. Genome-wide association studies (GWAS) identified 163 loci associated with IBD, including 23 loci specific to UC and 110 loci shared between UC and CD [44,45]. IBD risk genes include interleukin-10 (IL10), CD40 molecule (CD40), signal transducers and activators of transcription 1 and 4 (STAT1 and STAT4), interleukin-2 (IL2), and major histocompatibility complex, class II DR beta 1 (HLA-DRB1). However, the exact influence of these genes in IBD pathogenesis is not fully understood. Susceptibility genes are involved in immune regulation, inflammatory responses, and gut barrier function. Single nucleotide polymorphisms (SNPs) in genes such as caspase recruitment domain family member 15 (CARD15), nucleotide-binding oligomerization domain containing 2 (NOD2) and several Toll-like receptors (TLRs) are involved in innate immunity and associated with UC [46,47]. Genes

involved in adaptive immunity associated with an increased risk of UC include the human leukocyte antigens (HLAs), particularly HLA-DRB1 [48]. Interleukins (ILs) are cytokines that are key molecular modulators involved in UC [49]. Several reports suggested that the interleukin 23 receptor (IL23R) and signal transducer and activator of transcription (STAT3) are associated with the Th17 pathway [50,51]. Other IBD risk genes are involved in the maintenance of epithelial barrier integrity, such as extracellular matrix protein 1 (ECM1), cadherin-1 (CDH1), hepatocyte nuclear factor 4 alpha (HNF4A), laminin beta 1 (LAMB1), prostaglandin EP4 receptor (PTGER4), solute carrier family 22, member 4/5 (SLC22A4/SLC22A5), myosin 9B (MYO9B), and multidrug resistance mutation 1 (MDR1) [52].

1.3.2 Environmental Factors

UC is influenced by a multitude of environmental factors, including smoking, hygiene, lifestyle, microbiota, body weight, stress, antibiotic usage, and diet **(Figure 4)** [53,54]. These factors can directly or indirectly affect microbial diversity and modulate immune responses [55,56].

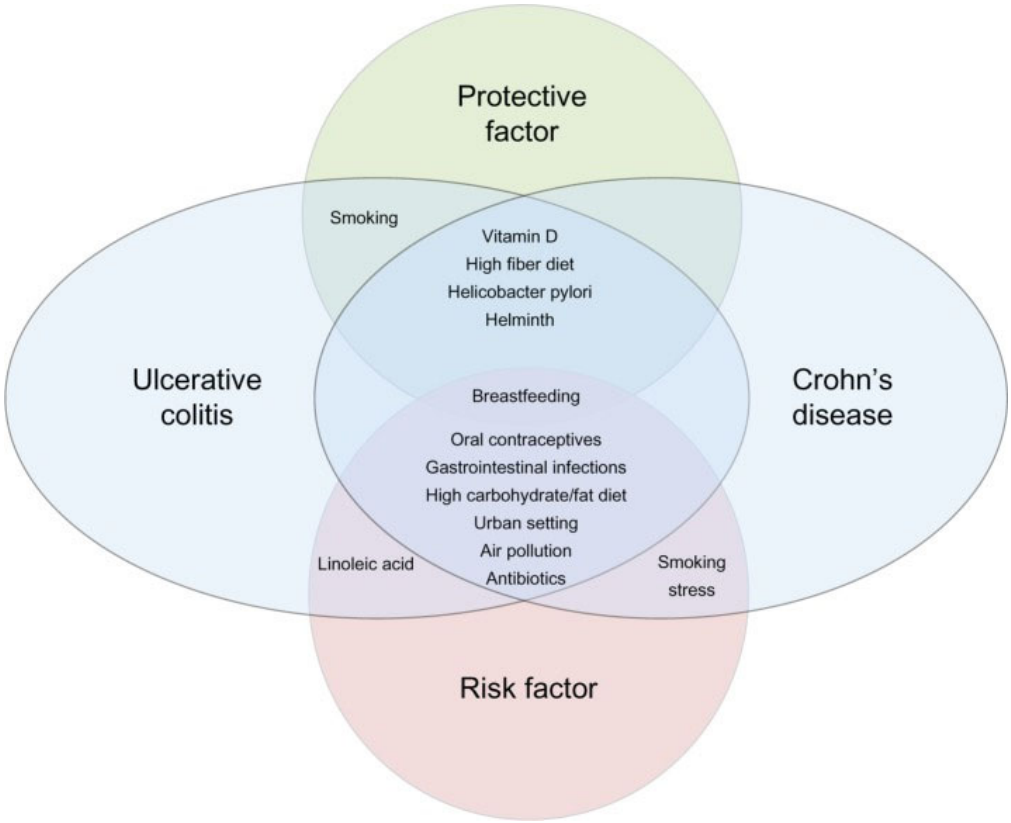


Figure 4. *The environmental factors of IBD. With permission from [57].*

1.3.3 Gut Microbiota

Dysbiosis characterized by an altered composition of microbiota has been associated with heightened intestinal inflammation in UC [58,59]. The bacteria *B. vulgatus* is associated with UC severity [60]. To address dysbiosis in the gut microbiota, faecal microbiota transplantation (FMT) has been used as a treatment option for UC patients. However, FMT is not always an effective treatment [61,62].

1.3.4 Nutrition

Diet varies significantly between regions and cultures [63]. The precise mechanisms of how diet affects UC development are unknown. Epidemiological evidence suggests that children of immigrants moving from undeveloped to developed countries show a higher incidence of IBD [64]. The Mediterranean diet, which includes substantial amounts of whole grains, fruits, and vegetables, has been shown to help alleviate UC [56].

1.4 Inflammation in UC

Dysbiosis in the gut has been implicated to be important in the development and progression of UC [65]. The dysbiosis disrupts the normal function and composition of the epithelial layer, which may be involved in the development and progression of UC through regulation of gut permeability [66]. Epithelium consists of several types of cells including stem cells, goblet cells, Paneth cells, glial cells, and epithelial cells **(Figure 5)**. Stem cells in the crypt can differentiate into different cell types including goblet cells and epithelial cells. Paneth cells are secretory cells that produce antimicrobial peptides [67]. Glial cells provide nutrition to neurons and communicate with neurons [68]. Goblet cells located in the epithelium produce and secrete mucus. Depletion of goblet cells results in a weakened mucosal layer [69]. Structural weakening of the mucus barrier is an early event in UC pathogenesis. The weakening of the mucosal barrier exposes the underlying epithelium to potential damage [70]. Epithelial cells are a physical barrier that maintains the permeability of nutrients but limits the entry of pathogens. Disrupted tight junctions (TJs) in the epithelial layer increase epithelial

permeability, which allows bacteria, toxins, and other antigens to penetrate the intestinal tract [71]. Pathogens go through disruptive TJs and then trigger an immune response resulting in inflammation [72].

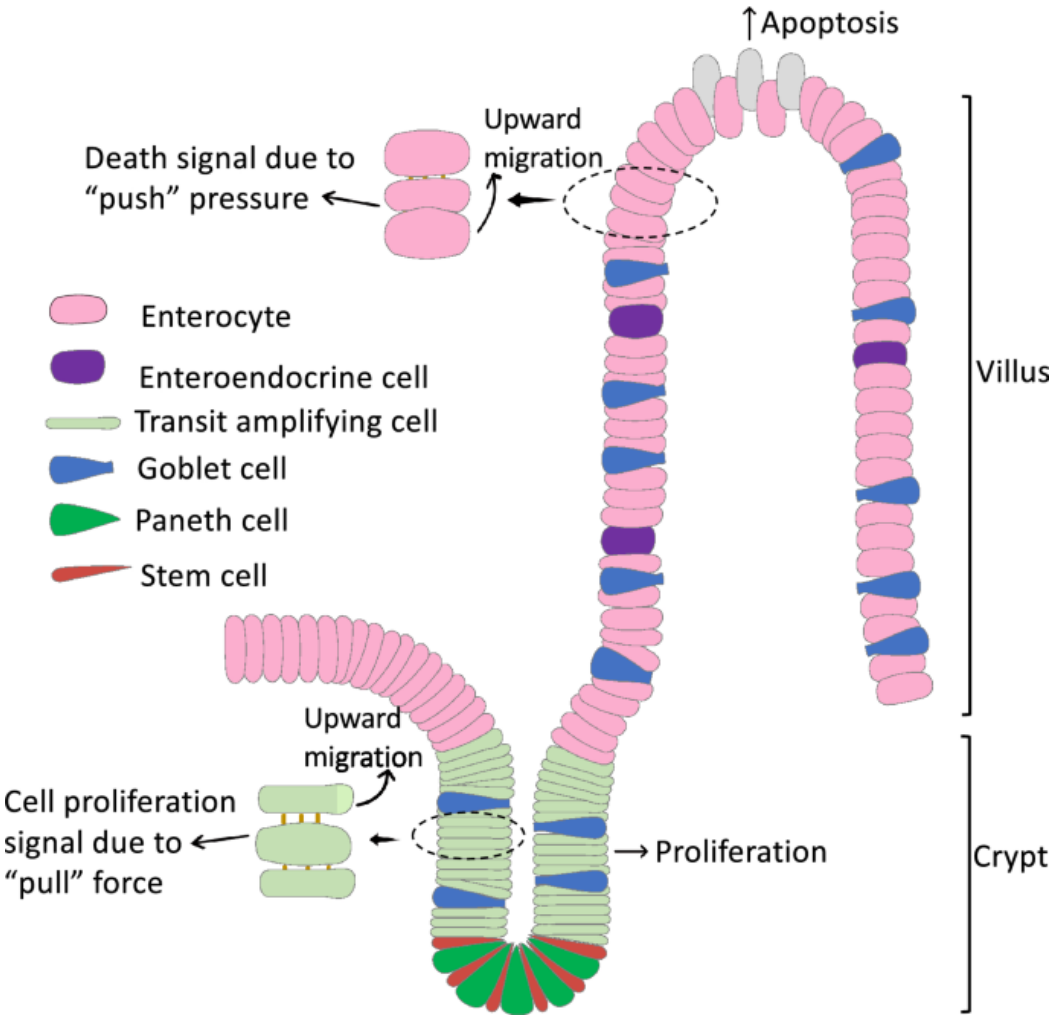


***Figure 5**. The crypt contains various cell types. With permission from [73].*

Under the epithelium, the lamina propria harbours numerous immune-derived cells, including macrophages and lymphocytes **(Figure 6)** [74,75]. In the event of injury or infection, immune cells within lamina propria migrate to the site of damage and initiate an immune response [76].

The immune response is classified into two branches: the innate and the adaptive immune responses. The innate immune system serves as the first line of defence against microbial

pathogens and is characterized by a rapid non-specific immune response [77]. Dendritic cells (DCs), macrophages, and B cells are antigen-presenting cells (APCs) communicating between innate and adaptive immune responses [78]. APCs can initiate an immune response to combat threats [79]. These cells express pattern recognition receptors (PPRs), such as Toll-like receptors (TLRs) and nucleotide-binding oligomerization domain-coding proteins (NODs), to recognize specific molecular patterns [71]. After recognition, TLRs stimulate downstream signals by secreting pro-inflammatory mediators [80]. These mediators, including cytokines and chemokines, are key molecules that connect the innate immune response and adaptive immune response.

The adaptive immune system generates specific and long-lasting responses to pathogens. This system is composed of T cells and B cells. T cells can differentiate into cytotoxic CD8+ T cells, and T helper cells (CD4+ T cells) after being stimulated by APCs [81]. CD8+ T cells participate in apoptosis [82]. T helper cells release various cytokines to regulate the immune response [83]. Among T helper cells, Th1, Th2, Th17, and regulatory T cells (Tregs) are pivotal in an adaptive immune response [84]. Chronic inflammation in UC may be partly due to an imbalance between pro-inflammatory T cell subsets (Th1/Th2 and Th17) and anti-inflammatory Tregs [85]. B cells have antigen-binding receptors and recognize antigens directly [86].

The immune system in active UC is activated, resulting in an elevated production of pro-inflammatory cytokines and chemokines [87]. The complex network of immune response and interactions is depicted in **Figure 6**.

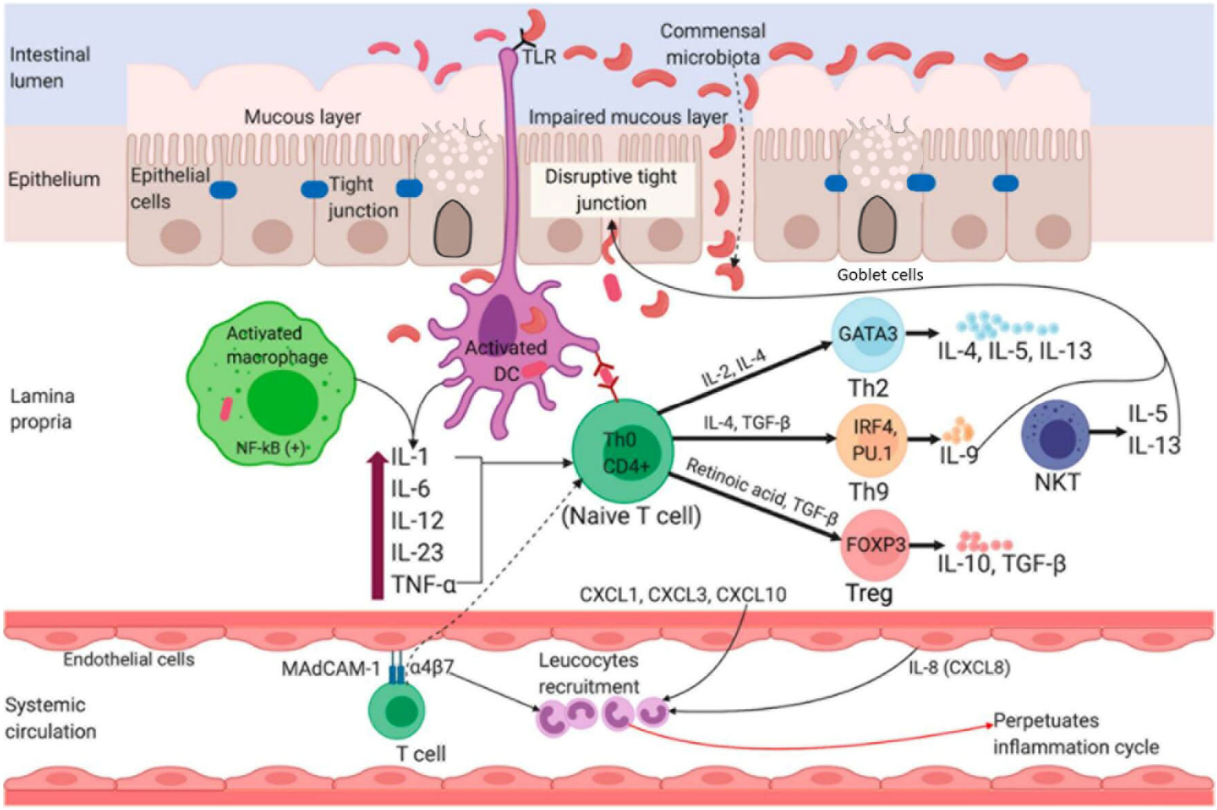


Figure 6. *Pathogenesis of UC. Adapted and with permission from [88].*

1.4.1 Inflammatory Mediators

Many pro-inflammatory molecules and signalling pathways contribute to the immune response in UC. Key inflammatory mediators released in various immune cells include Tumour Necrosis Factor-alpha (TNF-α), interleukin-6 (IL6), and interleukin-1 beta (IL1B). The dysregulation of pro-inflammatory cytokines such as TNF-α can contribute to immune cell recruitment and activation, leading to chronic inflammation [89]. Cytokines can alter the balance of anti-inflammatory and pro-inflammatory responses. This may induce tissue damage and subsequentially lead to an increased exposure to antigens and sustained immune responses [32,90].

IL6 is pro-inflammatory in chronic inflammation and anti-inflammatory in acute inflammation [91]. IL6 is highly expressed in active UC and involved in the differentiation of Th17 cells and Treg cells [92,93]. An increased expression of IL6 and its consequent interaction with CD4+ T cells lead to the induction of the expression of the anti-apoptotic genes, such as STAT3 and B-cell lymphoma-extra-large (BCLxL). Anti-apoptosis elevates the population of CD4+ T cells in

the lamina propria, which leads to perpetuating chronic intestinal inflammation [94,95]. The gut microbiome may also stimulate macrophages and DCs to release IL6 [96].

IL10 is a regulatory cytokine that inhibits antigen presentation and pro-inflammatory cytokine release [97]. Treg produces IL10 to impair the production of IL1B thereby suppressing intestinal inflammation [98]. IL10 may also inhibit the production of pro-inflammatory cytokines such as TNF-α and interferon-gamma (IFN-γ) in Th1 cells and macrophages [99].

1.5 Epigenetics

The term "epigenetics" refers to changes in gene expression patterns that are brought on by changes in the DNA structure or its related proteins rather than changes in the DNA sequence [100]. Epigenetic modifications include DNA methylation, histone modification and non-coding RNAs, depicted in **Figure 7** [101–104]. Epigenetic regulation of gene expression levels is heavily influenced by the environment [105,106]. The interactions between the genome and the environment have the potential to modify gene expression and impact an individual's susceptibility to certain diseases [107]. The expression of many genes is regulated by epigenetics during embryogenesis and ageing [108].

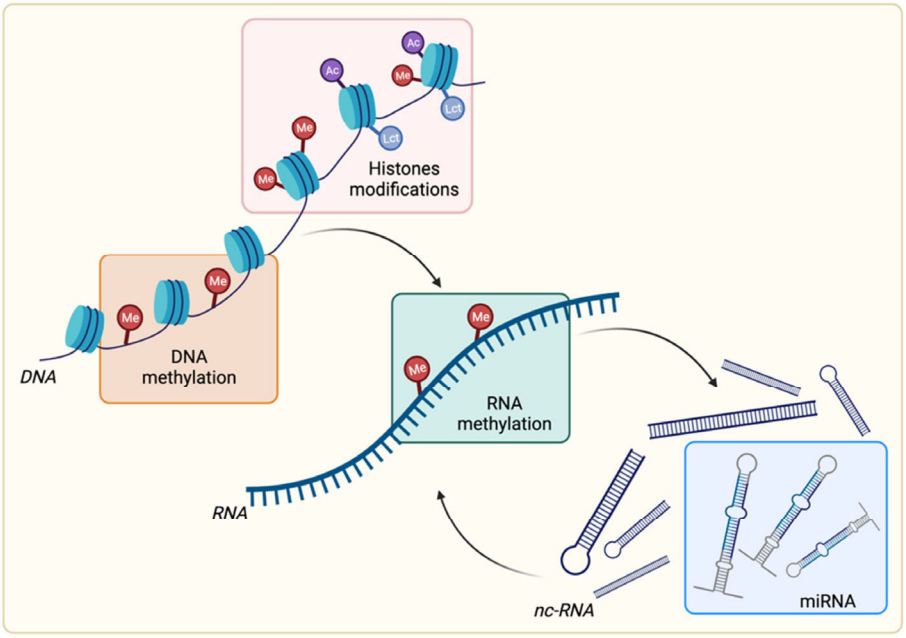


Figure 7. *Overview of epigenetic modifications. With permission from [109].*

1.5.1 DNA Methylation Mechanism and Biological Effects

DNA methylation is the addition or removal of a methyl group on cytosine nucleotides. Methylation events occur typically on the CpG sites where a cytosine is followed by a guanine nucleotide. Large stretches of CpG sites are called CpG islands [110]. CpG islands are located mostly in promoter regions of transcripts [111]. The methylating of CpG sites can regulate the expression of a transcript through activation or inactivation of the transcription **(Figure 8)**.

Many biological processes are regulated by methylation such as embryonic development, suppression or activation of transcription, and X-chromosome inactivation [112,113]. Notably, 60%-80% of CpG sites in the human genome are methylated [114]. Hypermethylation is an increase in methylation, which is the default for most CpG islands [115]. Hypermethylated DNA may have limited access for transcription factor binding, which may further reduce the gene expression level [116].

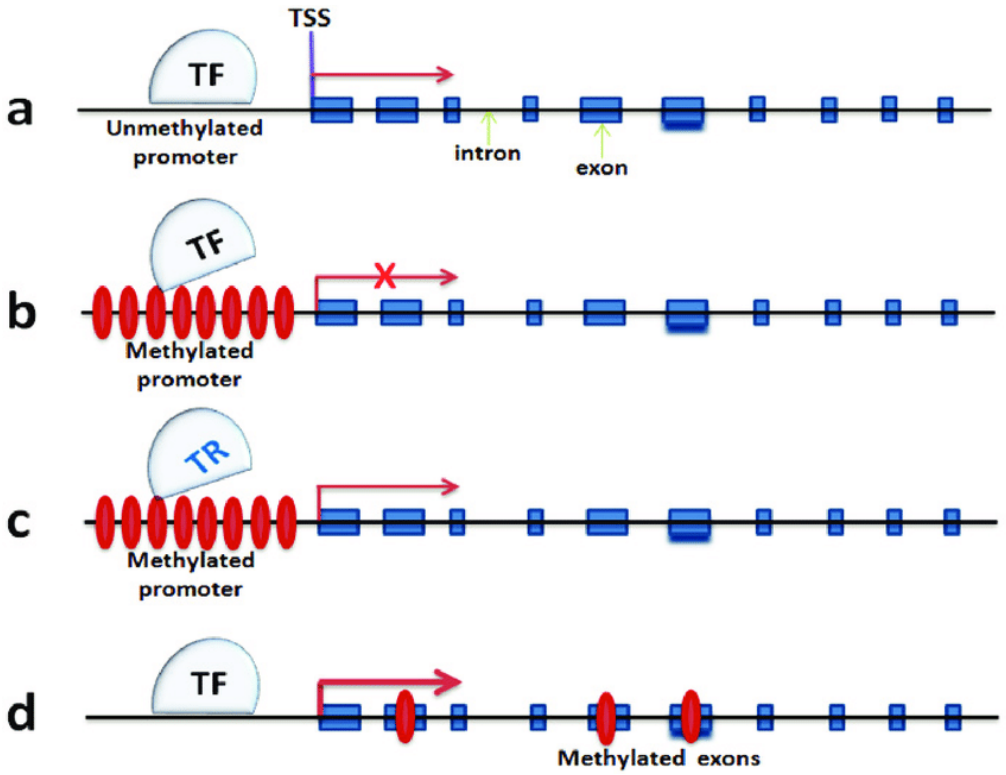


***Figure 8**. DNA methylation mechanisms in regulating gene expression. With permission from [117].*

1.5.2 DNA Methylation in UC

The influence of DNA methylation on the UC progression is still unclear. DNA methylation has been shown to be involved in the dysregulation of T-cell development and differentiation [118–121]. A Change in methylation profiles of CD8+ T cells and B cells has been reported [122,123]. Studies have shown that several pro-inflammatory genes such as toll-like receptors 2 (TLR2), toll-like receptors 4 (TLR4), and TNF-α are under epigenetic regulation [124–126]. It is hereby noted that DNA methylation can also occur in the absence of known CpG islands [127]. Previous research has revealed a global hypomethylation of DNA in UC patients [128]. Hypomethylation of sialic acid binding Ig like lectin 5 (SIGLEC5), cluster of differentiation 86 (CD86), and CXADR like membrane protein (CLMP) is related to severe UC [129]. Protease-activated receptor 2 (PAR2) hypermethylation is observed related to the severity of UC [130].

DNA hypomethylation is related to increased blood CRP levels and expression of NOD2 [131]. Transcription factors that are methylation regulated and associated with inflammation are the nuclear factor kappa-light-chain-enhancer of activated B cells (NF-κB), forkhead box P3 (FOXP3), interferon regulatory factors (IRFs), and STAT3 [132].

1.6 Functional Genomics

Functional genomics is a field that studies transcriptomics, genomics, proteomics, and metabolomics to define distinct phenotypes **(Figure 9)**.

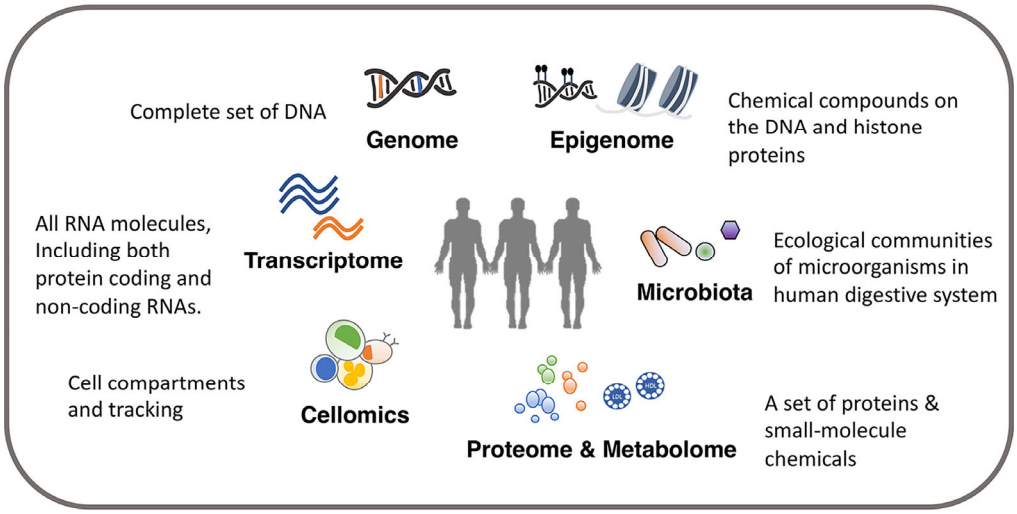


***Figure 9**. Overview of functional genomics. With permission from [133].*

1.6.1 Transcriptomics and Epigenomics

Transcriptomics focuses on the study of all RNA molecules present in a cell or tissue at a given time, collectively known as the transcriptome. The transcriptome includes ribosomal RNA (rRNA), transfer RNA (tRNA), messenger RNA (mRNA), non-coding RNA, and other RNA species [134]. There are three common methods to measure the transcriptome: PCR arrays, microarray, and next-generation sequencing (NGS).

PCR arrays are a sensitive technique used to quantify gene expression levels [135]. This method uses reverse transcription to create complementary DNA (cDNA) followed by amplification of specific transcripts and housekeeping genes (controls) [136]. PCR arrays have advantages offering precise and rapid quantification of known genes [137]. PCR arrays are not able to discover novel transcripts or provide a comprehensive view of the transcriptome [137].

Microarray uses hybridization of cDNA transcripts to a pre-designed array of nucleotide probes. This allows for the relative quantification of known transcripts [138]. They are suitable for large-scale studies, though this method is non-explorative and limited to the probes designed. The limited dynamic range between signal background levels and saturation levels may be unfavourable for downstream analysis [139,140].

RNA sequencing is a NGS approach, which sequences cDNA from reversed transcribed RNA, providing quantitative data [141]. RNA-seq can be used for gene expression or whole transcriptome including novel transcripts and isoforms in the genome with higher resolution and dynamic range than microarrays [142].

Beyond transcriptome analysis, sequencing can be used to study epigenetic modifications. DNA methylation can be detected by bisulfite conversion sequencing (BS-seq), which uses bisulfite treatment to change unmethylated cytosine to uracil [143,144] **(Figure 10)**. DNA methylation status may aid in the functional interpretation of the genome [145].

Several methods are available for the detection of methylation events, including PCR-based methods, microarrays, reduced representation bisulfite sequencing (RRBS) and NGS.

PCR-based methods such as methylation-specific polymerase chain reaction (MSP) use specific primers to detect the methylation status of targeted sequences. After treating samples with bisulfite, two sets of primer targets both methylated DNA and unmethylated DNA sequence [146]. After amplification, sequences from the methylation-specific primers could provide qualitative result with high sensitivity [147]. This PCR-based method is not suitable for genome-wide analysis because it requires predefined regions [146].

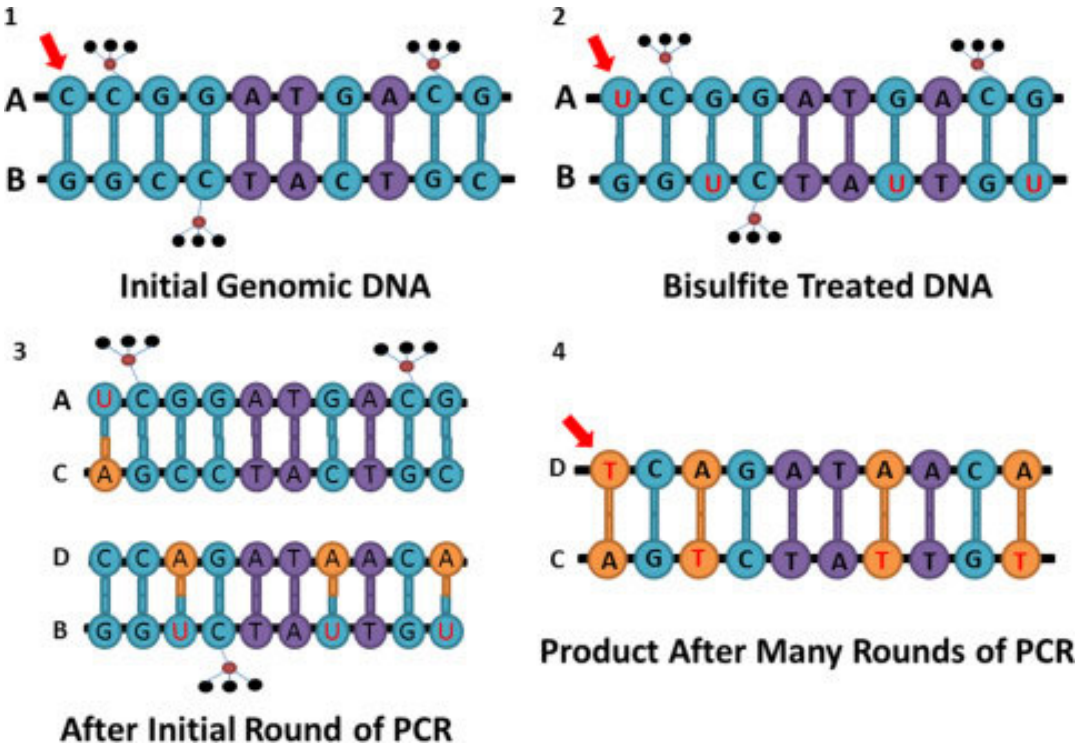


***Figure 10**. The bisulfite conversion in bisulfite sequencing. With permission from [148].*

Array-based methods use probes to hybridize bisulfite-converted DNA that enable analysis on selected CpG sites or regions [149]. Illumina Infinium array platform uses two types of probes for each CpG site: One for the methylated sequence and another for the unmethylated. After hybridization, fluorescent signals can be detected with different intensities to determine the methylation status at each CpG site [150]. Though the number of detected CpG sites is limited, array-based methods are commonly used due to the established standardised processing methods [151].

Another method for detecting methylation status is reduced representation bisulfite sequencing (RRBS). This method provides the opportunity for large-scale investigation of methylation patterns across the entire genome [152]. RRBS uses the restriction enzyme MspI which recognizes the DNA sequence CCGG [153] **(Figure 11)**. Then, the adapter-ligated DNA is selected for the proper size, treated with bisulfite, and PCR amplified [154]. The treated DNA is sequenced. RRBS mostly targets CpG rich regions, which leaves CpG-poor regions and intergenic areas with relevant methylation changes ignored [155]. Furthermore, the restriction enzyme is also limited to the genomic regions containing recognition sites, leaving the sequence lacking those sites undetected [156]. The advantage of RRBS is that it is cost-effective compared to WGBS. RRBS allows for single-nucleotide resolution methylation analysis [155].

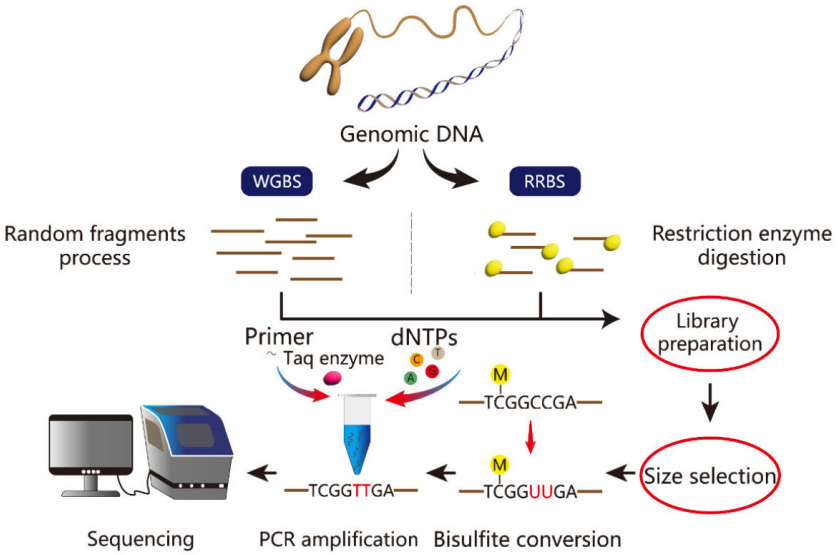


***Figure 11**. Overview of WGBS and RRBS. Adapted and with permission from [157].*

Next-generation Sequencing (NGS) can be used to sequence the whole genome, or regions of interest by targeted capture or amplification. There are two methods of NGS for methylation detection, which can be classified into indirect and direct methods. Indirect methods such as whole-genome bisulfite sequencing (WGBS) which requires amplification [158]. Direct method such as Oxford Nanopore sequencing can detect all methylation events in real-time without the need for amplification [159].

WGBS provides information across the entire genome including low density CpG areas [160,161]. Due to the substantial number of potential methylated CpG sites, and the large amount generated data, WGBS requires significant financial investment and computational resources to process and analyse [160]. Another limitation is bisulfite treatment, which converts both 5-hydroxymethylcytosine (5hmC) and 5-methylcytosine (5mC) to uracil [162]. Bisulfite treatment is harsh, it is performed at a low pH and temperatures up to 90°C [163]. This condition may lead up to 90% DNA loss [158].

In contrast, nanopore sequencing can be applied as a WGBS method, which provide direct detection of methylation marks from native DNA. Nanopore sequencing uses tiny pores to measure the current in real-time when molecule passes. It could provide insights into both CpG and non-CpG methylation with reduced preparation steps and lower costs [159]. The long-read capacity of Nanopore also improves the interpretation of genome structural variation, which is challenging with short-read WGBS [164].

1.6.2 Non-coding RNAs

Non-coding RNAs (ncRNAs) do not appear to have any functions in protein-coding, including small nucleolar RNAs (snoRNAs) and long non-coding RNAs (lncRNAs) [165]. Studies show that snoRNAs are involved in rRNA modification [166]. LncRNAs are transcripts longer than 200 nucleotides including miRNAs, siRNAs, and some snoRNAs [167]. LncRNAs regulate gene expression by various mechanisms including transcription, post-transcriptional modifications, translation, and post-translational modifications [168–170]. In UC, lncRNAs can participate in the innate and adaptive immune responses [171]. Dysregulation of specific lncRNAs has been implicated in UC pathogenesis [172]. LncRNAs have been shown to be involved in regulation of lipid and atherosclerosis, and T-cell receptor signalling [173]. For example, interferon gamma-antisense RNA 1 (IFNG-AS1) was reported as a pro-inflammatory lncRNA that regulates the expression of IFNG [174]. Metastasis-associated lung adenocarcinoma transcript 1 (MALAT1) is an example of a newly discovered lncRNA associated with UC [175].

2 Aims

The pathogenesis of UC is complex and defined by the interplay between genetic and environmental factors. The overall research goal of this work was to study the interplay between gene expression and DNA methylation in UC. The correlation between gene expression and DNA methylation is believed to provide novel specific molecular signatures in UC.

Aim 1: To identify molecular signatures that could predict relapse in UC remission.

Aim 2: To characterize DNA methylation patterns in UC remission patients.

Aim 3: To investigate the interplay between methylation and lncRNA expression associated with UC pathogenesis.

3 Material and Methods

3.1 Biopsy collection and study populations

UC biopsies were taken from an established biobank, approved by the Norwegian Board of Health, the Advanced Study of Inflammatory Bowel Disease (ASIB) study biobank at the University Hospital of North Norway (UNN) during the period 2004-2014. A standardised sampling method was used to collect mucosal biopsies. All studies were approved by the Regional Committee of Medical Ethics of Northern Norway, Ref no: 14/2004, 1349/2012 and 29895/2020. The study participants signed informed and written consent forms. The study participant consisted of treatment-naïve UC patients with mild to moderate disease activity, UC patients in remission, and control groups **(Table 1)**. The diagnosis of UC was established based on the clinical histology and endoscopic criteria as given by the ECCO guidelines [176]. The degree of inflammation was established during endoscopy using the UC Disease Activity Index (UCDAI) [177]. Control subjects underwent cancer screening, showing normal colonoscopy and normal histological examination. The exclusion criteria included serious medical conditions, immunological disorders, irritable bowel syndrome (IBS), polyps, cancer, or abnormal histology in colonic biopsies. To investigate the degree of UC activity, tumour necrosis factor-alpha (TNF-α) mRNA expression levels were quantified by real-time polymerase chain reaction (qPCR) [178]. Biopsies from patients with active UC were obtained from the active inflammation site and from the sigmoid part of the colon. Biopsies obtained from UC patients in remission with zero endoscopic score were included in the study. Additional patient information is protected by patient confidentiality.

Table 1. The patient cohort.

	Included		Disease severity
Paper I	Normal controls	n = 16	-
	Remission	n = 26	Remission
	Treatment-naïve UC	n = 14	Mild - moderate
Paper II	Normal controls	n = 11	-
	Remission	n = 20	Remission
	Treatment-naïve UC	n = 12	Mild - moderate
Paper III	Normal controls	n = 11	-
	Treatment-naïve UC	n = 13	Mild - moderate

3.1.1 Tissue handling

Biopsy samples were taken from the sigmoid colon and were preserved in RNA stabilization reagent RNAlater (Ambion/ Qiagen). The biopsies in RNAlater were kept at room temperature for maximum overnight, then at -20 °C until RNA isolation.

3.1.2 Quantitative polymerase chain reaction (qPCR)

Quantitative polymerase chain reaction (qPCR) was used to quantify mRNA to measure the transcription of TNF-α [178]. The number of copies of a gene reflects the activity of gene expression.

3.1.3 Histology

For histological examination, biopsies were immediately preserved in 10% formalin and further evaluated by experienced pathologists. The assessment involved scoring based on the Geboes Score [26].

3.2 Library preparation

The Allprep DNA/ RNA Mini Kit and the QIAcube instrument (Qiagen) were used to isolate total RNA and DNA from the same patient sample. The RNA quantity and purity were assessed by using the NanoDrop ND-1000 spectrophotometer (ThermoFisher Scientific). The Experion Automated Electrophoresis System (Bio-Rad) and the RNA StdSens Analysis Kit (Bio-Rad) were used to evaluate RNA integrity. The RNA samples were kept at −70°C until further use. All RNA samples used for analyses showed an RNA integrity number (RIN) value between 8.0 and 10.0.

The DNA methylation libraries were prepared following the methods described by Taman et al [127]. The SeqCap Epi CpGiant Enrichment kit (Roche) was used to prepare the DNA methylation libraries which has been discontinued in 2022. Briefly, 1 µg of genomic DNA was fragmented and ligated with an A-tail and adapters. DNA underwent bisulfite treatment (Zymo Research) and a dual-size selection. After a PCR amplification, the products were cleaned and then hybridized to the SeqCap Epi libraries. The samples underwent a 72-hour hybridization

process at 47°C. Then the samples were cleaned and washed, followed up by a PCR amplification step. Finally, the PCR products were cleaned up and the DNA libraries were eluted **(Figure 12)**. DNA fragment size was measured, and DNA libraries were assessed with Bioanalyzer 2100 and the Agilent DNA 1000 kit (Agilent Technologies). The average fragment size of the established libraries was between 322 and 329 base pairs. Before sequencing, libraries were diluted prior to pooling. Then libraries were denaturized and diluted to loading concentration. PhiX control was diluted to the same concentration as libraries. Libraries were combined with PhiX. The final percentage of PhiX was 1%.

RNA libraries were prepared with the TruSeq Stranded Total RNA LT Sample Prep Kit from Illumina. 1µg of total RNA was used as the input material. The quality of the RNA libraries was assessed by the Bioanalyzer 2100 and the Agilent DNA 1000 kit (Agilent Technologies). The average size of the fragments generated by RNA libraries was 301 bp. Libraries were normalized to 10 nM and diluted to 4 nM prior to sequencing.

DNA and RNA libraries were sequenced on the NextSeq 550 instrument, using a high output flow cell 150 cycles (Illumina), according to the manufacturer's instruction. The libraries were sequenced using paired-end mode.

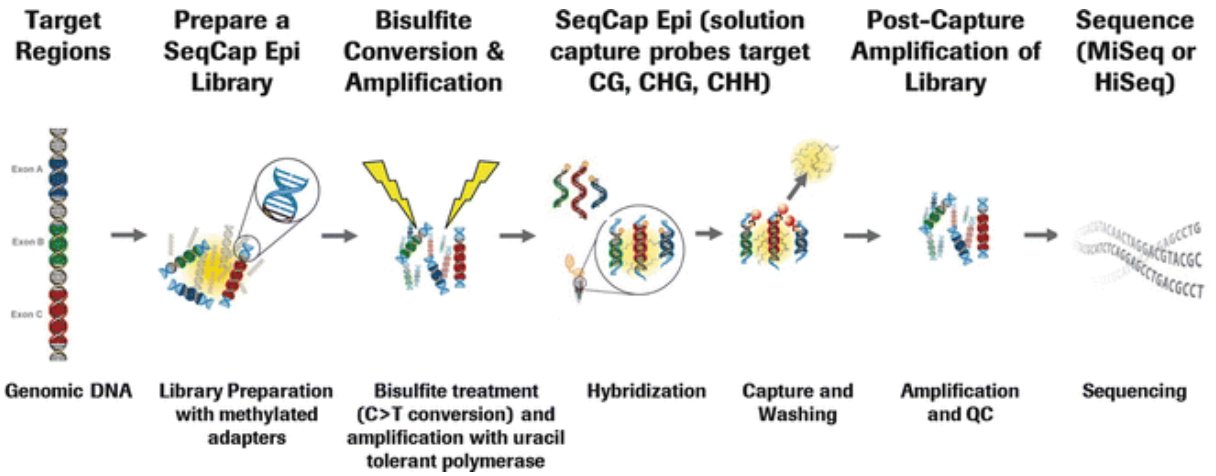


***Figure 12**. SeqCap Epi workflow. With permission from [179].*

3.3 Preprocessing of data

The sequenced data went through a quality scoring check on the on-board computer of the Illumina NextSeq 550 instrument, and only sequences with a q-score greater than 30 were kept. Adapter trimming was done by the Illumina NextSeq 550 instrument. Illumina bcl2fastq software v2.20 was used to demultiplex the output data. The read quality was also controlled by multiQC [180]. The RNA-seq FastQ files were aligned to the reference genome (GENCODE Human Release 33, Human Genome Assembly GRCh38.p13, https://www.ncbi.nlm.nih.gov/grc/human/data) by STAR (Version 2.7.3a) [181]. DESeq2 was used to normalize the gene counts [182]. No additional filtering for low counts or low variable genes was performed prior to data analysis.

For methylation data, Illumina NextSeq 550 was used to sequence the bisulfite-converted DNA with the same preprocessing steps above. The resulting FastQ files were aligned to the same reference genome mentioned above with bowtie 2 [183]. Bismark was used to generate methylation counts [184]. Bismark cytosine reports were generated for methylation coverage.

3.4 Data Analysis

A brief description of the analysis is described below, depicted in **Figure 13**.

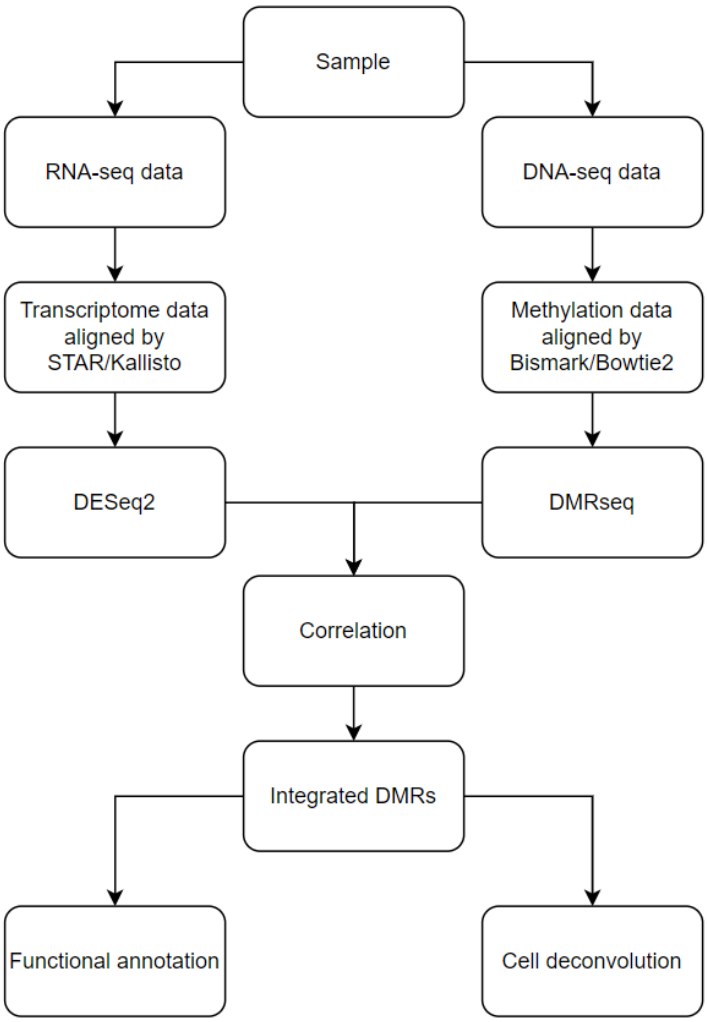


Figure 13. *The workflow for analysis of DMRs.*

The left side of **Figure 13** describes the transcriptomic part of the analysis. DESeq2 is used for the differential expression analysis of normalized gene count data. Genes that were significantly differentially expressed (DEGs) with p adjust< 0.05 were kept.

The right side of **Figure 13** describes the DNA bisulfite sequencing of the analysis. The R package dmrseq was used for the processed methylation data to identify differential methylation regions (DMRs) [185].

3.4.1 Correlating Transcriptome and Epigenome

The correlation analysis was done by correlating average methylation levels to proximal gene expression levels. For each DMR located within the 2 kb upstream region of the transcription start site (TSS), sample DMR and DEG were correlated using average DMR methylation and DEG expression levels by Kendall correlation [186]. Reference of TSS and promoter regions was used from R package TxDb.Hsapiens.UCSC.hg38.knownGene and UCSC known genes (GRCh38) [187,188]. Other regulatory features like enhancer and CpG flanking regions used ENSEMBL homo sapiens regulatory features release 104 (https://ftp.ensembl.org/pub/release-104/regulation/homo_sapiens). A negative correlation between DMRs and DEGs with a p-value less than 0.1 was kept. A negative correlation means that the expression level increases when the methylation level decreases, and *vice versa*.

For correlation between methylation and lncRNAs, DMRs located within the 20 kb region of the DEGs were considered. Pearson correlation was used to correlate average DMR methylation and differentially expressed lncRNA expression levels. To find a correlation between lncRNA expression and protein-coding gene expression, lncRNAs were correlated to 500 kb upstream or downstream of differentially expressed protein-coding genes.

3.4.2 Principal Component Analysis (PCA)

PCA is a method for analysing complex datasets. This method provides an approximation of the original data using several principal components (PCs). Normalization is crucial to PCA for ensuring each variable contributes equally to the components. The result of PCA is visualized with the first two principal components which are the most variant of all components. PCA is unsupervised and without reference to prior information.

3.4.3 COX Proportional Hazard Model

COX proportional hazard model is used to study the relationship between the covariates and the accumulative survival rate [189]. It can estimate the effects of several covariates (genes) on the time of event occurrence. The COX model's key assumption is that covariates' impact on the hazard ratio is constant over time [190].

3.4.4 Partial Least Square (PLS)

PLS is like PCA as an analysis that reduces data dimensions to increase readability. PLS finds a linear regression model of two variables and projects them to a new space instead of PCA maximizing the variance [188]. PLS is a supervised method [191].

3.4.5 Gene Annotations

To get functional annotations, genes were enriched through over-representation analysis with ReactomePA, PANTHER (www.pantherDB.org) and GO [192,193]. Over-representation analysis is a method that determines if the genes are from the predefined gene sets (the pathways). This process used R package ReactomePA, and clusterProfiler [194,195]. PANTHER was performed through PANTHER with Fisher's exact test on Reactome pathways.

3.4.6 Cell Deconvolution

Cell fractions were deconvoluted from epigenetic data by EpiDISH [196]. EpiDISH only supports Illumina probe IDs. The methylation positions were overlapped to Illumina EPIC identifier by genome location. The average methylation levels were converted for each overlap with probe intensities (beta values) for each sample. Thus, making it possible for EpiDISH to estimate the different fractions in a tissue sample by methylation levels.

3.4.7 Statistical Tests

Fisher's exact test is a statistical significance test used in the analysis of tables. It is used for the investigation of association on categorical data [197].

The T-test is a statistical test that compares the averages of two groups to determine if there is a significant difference between the two groups [198].

Tukey honestly significant difference test is a one-step multiple comparison procedure based on studentized range distribution [199]. This test is used for comparison between three groups of cells.

A p-value less than 0.05 is usually considered significant to reject the null hypothesis. However, even a 5% false discovery rate in a large dataset is high enough to draw a false conclusion. To avoid this, the p-value needs to be adjusted by multiple correction. Benjamini-Hochberg for the p adjust method was used in this study [200].

4 Summary of the results

4.1 Publication 1

Anti-apoptotic genes and non-coding RNAs are potential outcome predictors for ulcerative colitis

Wei Meng, Kay-Martin Johnsen, Christopher G Fenton, Jon Florholmen, Ruth H Paulssen

Functional Integrative Genomics, 2023 May 18;23(2):165.

https://doi.org/10.1007/s10142-023-01099-9

Currently, there are no known clinical, immunologic, genetic, or laboratory markers that can accurately predict relapse. Therefore, there is no clear recommendation for discontinuing treatment. The aim of this study was to investigate the possibility of identifying genetic markers associated with the duration and outcome of remission by using transcriptional analysis combined with COX survival analysis. Data was analysed from whole transcriptome sequencing (RNA-seq) on mucosal biopsy samples taken from patients with active treatment-naïve UC (UC, n = 14), remission (RM, inactive UC, n = 26), and healthy controls (NN, n = 16). COX proportional hazard regression analysis was applied to the remission samples using remission duration as an event. The COX analysis could distinguish between two different patient groups within remission, and the subgroups differed in duration. To ensure the COX results were associated with UC, a PCA was performed that included both controls and active UC. The patient group with the longest remission duration and no relapse showed an increased expression of a set of snoRNAs and anti-apoptotic factors belonging to the MTRNR2-like gene family. Seven random remission samples not used in the analysis were used to validate the results.

4.2 Publication 2

DNA methylation fine-tunes pro-and anti-inflammatory signalling pathways in inactive ulcerative colitis

Wei Meng, Christopher G Fenton, Kay-Martin Johnsen, Hagar Taman, Jon Florholmen, and Ruth H Paulssen

Scientific reports, 2024 March, 14, 6789

https://doi.org/10.1038/s41598-024-57440-0

Immunological dysfunction associated with inflammatory bowel disease (IBD) and ulcerative colitis (UC) have been linked to DNA methylation. In patient samples with inactive UC (remission), variations in DNA methylation and related gene expression may point to possible remission-specific regulatory mechanisms. Data obtained by targeted bisulfite sequencing and whole transcriptome sequencing from mucosal biopsies of healthy controls (NN, n = 11), inactive UC patients (RM, remission, n = 20), and patients with treatment-naïve UC (UC, n = 14) were analysed. The differentially methylated regions (DMRs) were found using DMRseq. Correlation analysis was used to compare DMR methylation levels with the nearest differentially expressed genes (DEGs). The correlated genes were then visualized by principal component analysis (PCA) to ensure their relevance to UC. Next, DMR-regulated genes were annotated functionally. The correlational comparisons revealed 38 genes (correlation p-value < 0.1) that have a remission-specific methylation and expression profile. These genes include IL1B and STAT3. Both IL1B and STAT3 are implicated in the cytokine and IL-10 signalling pathway. DNA methylation events appear to fine-tune both pro- and anti-inflammatory processes to maintain a prolonged healing process in inactive UC.

4.3 Publication 3

Methylation-Regulated Long Non-Coding RNA Expression in Ulcerative Colitis

Christopher G. Fenton, Mithlesh Kumar Ray, Wei Meng, and Ruth H. Paulssen

International Journal of Molecular Sciences. 2023, 24(13), 10500.

https://doi.org/10.3390/ijms241310500

It has been demonstrated that long non-coding RNAs (lncRNAs) are involved in the pathophysiology of ulcerative colitis (UC). While lncRNA expression and other epigenetic processes like DNA methylation have been extensively researched in UC, the significance of their interaction has not yet been thoroughly investigated. It is well known that UC is influenced by an interplay between environmental factors and epigenetic mechanisms. In this investigation, mucosal biopsies from healthy controls (NN, n = 13) and treatment-naïve UC patients (UC, n = 11) were used. Whole-genome bisulfite sequencing (WGBS) and lncRNA expression data were evaluated from each sample. To find lncRNAs that may be regulated by upstream differentially methylated regions (DMRs), a correlation analysis between lncRNA expression and DMRs was performed. Additionally, by comparing their expression, the nearest protein-coding genes linked to DMR-regulated lncRNAs were identified. The investigation discovered UC-associated lncRNAs that may be controlled by DMRs, including MIR4435-2HG, ZFAS1, IL6-AS1, and Pvt1. Genes implicated in inflammatory immunological responses, such as SLC15A4, CCL18, and SERPINB1, are located downstream from DMR-regulated lncRNAs.

5 Discussion

UC is a chronic inflammatory disease with a remission-remitting course thereby placing a heavy burden on the healthcare system [32]. Inflammation is necessary for maintaining colon homeostasis, which is a balance between pro- and anti-inflammatory processes [201]. Homeostasis is significantly affected by environmental, immunological, genetic, and treatment factors [73,202]. UC is a chronic disease and difficult to treat. The goal of treatment for UC patients is to relieve the symptoms and thereby achieve and maintain a remission state. UC patients in remission often face the risk of relapse switching to a pro-inflammatory state [203]. There is currently no method to predict if or when a patient will relapse. Therefore, characterizing remission is highly recommended.

Remission is defined as the absence of symptoms. However, current UC assessment methodologies have their limitations. Laboratory markers such as TNF-α and C-reaction protein (CRP) assess systemic inflammation and are not specific to UC [204]. Clinical, histological, and endoscopic scores offer limited insight into the molecular characteristics of UC [205]. For example, quiescent inflammation is not detectable by endoscopy but has been shown in UC remission patients at a molecular level [206]. This study investigated both gene expression and methylation data from active, remission and control samples. Comparisons between these three groups allowed the identification of molecular signatures believed to be specific to remission.

Remission status in UC can vary widely between individuals. The variability in remission patients includes genetic factors, clinical manifestations, and choice of treatment. All remission patients in this study received initially 5-ASA. Some of the patients received immunosuppressives in addition to 5-ASA. 5-ASA might influence methylation by increasing the expression of DNA methyltransferase 1 (DNMT1), which catalyses the transfer of methyl groups to specific CpG sites in DNA [207]. As the active UC patients were all treatment-naïve, differences in methylation between remission samples versus active UC may be due to the effect of 5-ASA treatment **(Paper II)**. Many UC studies lack proper patient stratification and annotation. Therefore, treatment-naïve patients were used to compare active UC and UC patients in remission. This allowed us to compare methylation changes without possible

treatment effects in active UC as treatments like 5-ASA, steroids and biologics may affect DNA methylation [208].

The overall approach for this study was to relate the methylome and the transcriptome within the individual patients. The first step was to investigate the gene expression profiles of active UC, remission, and control samples using PCA. The initial PCA showed a clear separation of the samples **(Paper I, Figure 2),** which was consistent with previously published results [206,209]. Interestingly, the remission samples separated into two clusters. To help explain this separation, the clinical metadata was further investigated. Among the clinical data, duration was the most variable **(Paper I, Table 1)**. An initial PLS-COX analysis was performed to see if duration could help explain the separation between the remission samples. This PLS-COX model revealed that there was enough power needed to study remission expression data in respect to duration (likelihood test $p < 0.05$, **Paper I, Table S2)**. PLS reduced the dimensionality of the data for simplifying interpretation. A second PLS-COX model was used to study the contribution of each gene to the duration of remission. A total of 287 genes were shown to have a considerable influence on duration.

The COX models were only performed on remission samples. A PCA was performed including remission COX results. The PCA included both UC and control samples. This ensured that the remission COX results were indeed related to UC. The clear separation of UC and control samples indicated that these genes are related to UC **(Paper I, Figure 3)**. Furthermore, the PCA revealed the separation of remission samples into two clusters. The clusters differed in the duration of remission **(Paper I, Figure 3)**. To validate the COX model results, seven samples were randomly excluded before the analysis. Of the 287 genes, four mitochondrial genes from the excluded samples showed a significant association between their expression and duration of remission **(Paper I**, **Figure 8)**.

DNA methylation is a dynamic process that can be affected by environmental factors [210]. Differences in gene expression between remission, control and active UC may be due to different DNA methylation patterns [211]. To study methylation, whole genome bisulfite sequencing (WGBS) was used which is currently the golden standard for studying genome-wide DNA methylation [212]. However, the interpretation of methylation data is difficult as there is no consensus on the number or the methylation level of CpG sites needed to have a

functional effect [213,214]. Another difficulty is variability due to differences in cell type population, and individual characteristics such as age, gender, etc. [215]. Therefore, it is important to correlate methylation data with expression data from the same patient.

There exist two approaches for differential methylation analysis. One approach is to compare the methylation levels at individual CpG sites. A second approach is to compare the methylation levels of methylated regions (DMRs). Approaches such as methylKit and MethylSig can estimate methylation changes at individual CpG sites, although they may be limited to predefined regions [216,217]. The difficulty in studying individual CpG sites is the sheer number which can lead to an overly harsh multiple correction. Aggregation of CpG methylation levels into regions can help reduce the effect of multiple corrections. Biologically, there is a strong correlation between the methylation levels of nearby CpGs [218].

A common problem for bisulfite-treated methylation data is low or missing coverage of CpG sites [158]. There are several methods to help overcome low coverage challenges [219,220]. These methods may reduce variations between groups through the estimation of missing or low coverage values based on neighbouring methylation sites [221]. The analysis of methylated regions may exclude the contribution of isolated CpG sites. In this study, differential methylation analysis was performed at the region level by dmrseq. Dmrseq uses a smoothing method to detect DMRs and control the false discovery rate without the requirement of predefined regions [185].

Most methylation analysis packages do not support multiple comparisons [217]. Therefore, three comparisons were performed using dmrseq. Each comparison gave DMRs differing in position and length. DMRs from all three comparisons were merged by overlapping genomic coordinates. For each overlapped DMR, the average methylation level of control samples or active UC samples was compared with remission samples using a t-test. Any DMR that showed a significant difference in methylation levels when compared to both control and active UC samples were considered specific to remission **(Paper II)**.

To further assess the possible biological functions of remission-specific DMRs, a paired correlation analysis was conducted. The analysis correlated DMR methylation levels with the expression levels of the closest differentially expressed genes (DEGs) across all samples. The

correlation was limited to DMRs located within the promoter region of the neighbouring DEG gene or 2kb upstream of the gene transcription start site **(Paper II)**. DNA methylation within the promoter region has been shown to affect gene expression [222,223]. In our study, only a negative correlation between promoter DMR methylation level and gene expression level was considered. The assumption is that promoter hypermethylation will decrease gene expression, and hypomethylation will increase gene expression. Several genes whose expression was negatively correlated with upstream DMR methylation levels were identified **(Figure 2, Paper II)**. A PCA showing separation between normal, remission and active UC samples suggests that the expression of these DMRs regulated genes is related to UC **(Figure 4, Paper II)**.

Gene expression and methylation profiles are specific in remission patients compared to active UC and controls. The analysis of expression profiles across UC, remission and controls revealed a distinct separation between the three groups **(Figure 2, Paper I)**. In remission, the methylation patterns of several genes related to inflammation were different from the controls **(Paper II)**. This finding indicates that the micro-inflammation status may be regulated by methylation-regulated gene expression. For instance, short-term remission has an elevated expression of inflammatory genes such as carbamoyl-phosphate synthetase 2 (CAD) and coiled-coil domain containing 134 (CCDC134). CAD could inhibit intestinal epithelial cells' antibacterial activity [224] CCDC134 promotes CD8+ T-cell activation [225]. The inflammation status is also influenced by DNA methylation. Hypermethylated STAT3 and IL1B showed decreased expression in remission compared to active UC **(Paper II)**. Yet STAT3 and IL1B expression levels could not be restored to control levels. STAT3 could interact with protein disulfide-isomerase A3 (PDIA3), which could potentially increase the expression of interleukin enhancer binding factor (ILF3) [226]. ILF3 has been identified as a high relapse-risk gene in remission **(Paper I)**.

Mucosal healing is essential in maintaining a functional epithelial and mucosal barrier in the colon. Colonic tissue cell types are derived from intestinal stem cells (ISCs) located at the bottom of colonic crypts. A hallmark of UC is the destruction of colonic crypts, reducing the number of intestinal stem cells (ISCs) [227]. In remission, crypt architecture is believed to be restored [228]. In this study, stem cell marker PROM1 was upregulated in the remission group as compared to active UC **(Figure 6, Paper I)**, suggesting stem cell differentiation and

proliferation are enhanced in the remission group. Stem cell differentiation requires a considerable amount of energy [229]. A hypothesis suggests that aetiology and progression of UC are related to the altered mitochondrial function [230]. Mitochondria genes of the MTRNR2-like family were identified as having a protective effect in remission **(Figure 6, Paper I)**. MTRNR2-like genes were not found to be regulated by methylation in remission **(Paper II)**. Mitochondrially encoded 16S rRNA (MTRNR2) encodes Humanin, which is a peptide reducing inflammatory cytokines such as IL6, IL1B, and TNF-α [231]. These cytokines have an anti-apoptotic effect. In active UC, an increase in apoptosis may lead to a damaged colonic epithelium [232]. Two genes found specifically hypermethylated in remission which are associated with the fine-tuning cell apoptosis include: competing endogenous lncRNA 2 or microRNA let-7b (CERNA2) and citron rho-Interacting serine/threonine kinase (CIT). MTRNR2-like genes, along with methylation-regulated CERNA2 and CIT, are involved in reducing inflammation, energy production, and apoptosis, all essential processes for epithelial healing and maintenance [233,234].

In addition to the specific protein-coding genes found in remission, an unexpectedly considerable number of small ncRNAs (snoRNAs) were found with increased expression in long-term remission. Several snoRNAs were found to have a protective effect in remission **(Figure S2, Paper I)**. The exception was small nucleolar RNA, C/D Box 101 (SNORD101). A function of snoRNAs in remission is the polarization of macrophages which promotes cell proliferation and tissue repair through the phosphatidylinositol-3-kinase/protein kinase B (PI3K/AKT) pathway [235–237]. This indicates a role for snoRNAs in mucosal healing. However, further studies are needed to understand snoRNAs' specific roles and regulatory effects in remission.

MTRNR2-like genes are long non-coding RNAs, and their expression may be important for remission duration. Therefore, the study endeavoured to further investigate the effects of methylation on lncRNAs, specifically lncRNAs in UC **(Paper III)**. It has been shown that lncRNAs may exert their functions by regulating gene transcription and protein expression [238]. Methylation events were found associated with lncRNAs, such as AC007750.1 (lnc-SLC4A10-7), which is considered to regulate inflammation [239]. This implies that methylation-

influenced lncRNA expression contributes to the regulatory mechanisms in UC pathogenesis, particularly in inflammatory immune responses **(Paper III)**.

Recent research has identified several biomarkers for monitoring the progression of UC. Among the most widely studied biomarkers are calprotectin, which has been related to deep remission and to the severity of inflammation [240,241]. Those biomarkers are useful to observe treatment outcome but do not reflect the disease status at the molecular level. Previous transcriptomic sequencing studies have primarily emphasized on differences between active UC patients and controls [242–245]. A description of the remission status is often missing from these studies. However, three studies concerning remission in UC have revealed different gene expression results [206,209,246]. A microarray study which identified differentially expressed genes in remission including regenerating family member 4 (REG4), S100 calcium-binding protein P (S100P), serpin family B member 5 (SERPINB5), and regenerating family member 1 alpha (REG1A) [209]. Mitochondrial-related genes including PPARG coactivator 1 alpha (PPARGC1A) were identified by whole transcriptome sequencing in a remission study [246]. In another transcriptome study IL1B, TNF-α, STAT3, mucin 5AC (MUC5AC) and ADAMTS like 5 (ADAMTSL5) were identified [206]. These findings have shown that remission patients have distinct transcriptional profiles with inflammatory signatures. The existing of signatures may represent the quiescent inflammation in remission patients. These transcriptome profiles implicate several immune-related pathways which are active in maintaining remission. In **Paper I**, we tried to explore the relationship between the transcriptomic signatures and remission duration. Longer remission duration is associated with changes in the expression of mitochondrial and non-coding genes. The decreased mitochondrial gene expression in active UC is consistent with Haberman's findings [246], suggesting a role of restored mitochondrial function in remission. For non-coding genes, the results indicated snoRNAs were related to remission. The snoRNAs have been reported in the regulation of immune responses and anti-tumour immunity [247,248].

Previous remission studies have obtained various -omics data from different sources, such as blood and/or faecal samples. In blood samples, proteomic analyses have identified IL-10 as a potential predictor of relapse [249]. IL-10 is a cytokine with pro- and anti-inflammatory roles in IBD [99]. The results presented in **Paper II** showed that the IL10 pathway may be regulated

by methylation and thereby related to remission. Previous studies have shown that long-term UC patients have a higher risk of developing CRC [250–252]. This may be due to methylation changes and subsequent changes in gene expression, such as APC Regulator of WNT signalling pathway (APC), cadherin 13 (CDH13), alanyl aminopeptidase (ANPEP), and CBY1 interacting BAR domain containing 1 (CIBAR1/FAM92A1) [253]. Our findings indicate methylation regulated genes in remission patients such as AFAP1-AS1 might contribute to CRC development [254]. These results suggest maintaining remission may not fully mitigate the risk of CRC in UC patients.

It is hereby noted that the microbiome could regulate methylation through their metabolites especially short-chain fatty acids [255]. For example, propionate could induce hypermethylation and inhibit expression of DAB Adaptor Protein 1 (DAB1) [256]. In faecal samples, the changes in the abundance of *Akkermansia muciniphila* might serve as a microbial marker for predicting relapse [257,258]. Methylation changes in remission observed in **Paper II** may be related to the changes in gut microbiome composition. Therefore, collectively integrating multiple -omics data, such as metabolomics, proteomics, and genomics may present new opportunities for monitoring disease status.

5.1 Limitations

Colonic mucosal biopsies contain many different cell types that may have different expression profiles [259]. Whole-genome and transcriptome sequencing are not able to measure the contribution of gene expression and methylation in cell compositions. Therefore, the different cell populations heavily affect the expression level and methylation level. The relationship between DNA methylation and gene expression is likely to be cell specific [260]Click or tap here to enter text..

Genome is a dynamic system. A cross-sectional study could be seen as a snapshot of remission and UC, which may not be representative of patients at different time points. The cross-sectional nature of this study may limit the ability to understand the temporal aspects of gene expression and epigenetic modifications associated with the transition from active disease to remission. Furthermore, all results are from *in silico* analysis and need experimental validation.

5.2 Future Perspectives

This study identified potential molecular signatures associated with UC. Future studies are necessary to validate our results, particularly non-coding RNAs and methylation profiles, by using larger patient cohorts. The validation studies should aim to investigate the clinical implications of the observed molecular markers. UC is highly heterogeneous disease, where single markers may not provide enough information. A system biology approach by integrating multi-omics data would provide a more comprehensive view of the UC remission status.

More research should focus on explaining the functional roles of mitochondrial genes, snoRNAs, and lncRNAs in modulating immune responses and epithelial repair processes. Tools such as CRISPR/Cas9 could help study the function of the above-mentioned genes. The understanding of gene function and networks could identify new therapeutic targets for improving treatment strategies.

Recent new technologies like spatial transcriptomics and single-cell RNA sequencing could be used to investigate the micro-environment and cellular heterogeneity in mucosal biopsies [261]. Data obtained by these technologies could help in the understanding of this complex disease.

6 Conclusion

This study investigates molecular signatures by analysing transcriptomic and methylation profiles in UC. Transcriptomic analysis revealed that mitochondrial genes might potentially predict the risk of relapse. Several genes and lncRNAs were identified that may be methylation-regulated and involved in inflammatory immune responses. Further investigations of molecular signatures in UC are necessary to define clinical outcomes and to develop treatment strategies and management for UC patients.

References

1. Ramos GP, Papadakis KA. Mechanisms of Disease: Inflammatory Bowel Diseases. Mayo Clinic Proceedings. 2019. pp. 155–165. doi:10.1016/j.mayocp.2018.09.013

2. Head K, Jurenka J. Inflammatory bowel disease part II: Crohn's disease - Pathophysiology and conventional and alternative treatment options. Alternative Medicine Review. 2004. pp. 360–401.

3. Alatab S, Sepanlou SG, Ikuta K, Vahedi H, Bisignano C, Safiri 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. 2020;5. doi:10.1016/S2468-1253(19)30333-4

4. Ananthakrishnan AN. Epidemiology and risk factors for IBD. Nature Reviews Gastroenterology and Hepatology. 2015. pp. 205–217. doi:10.1038/nrgastro.2015.34

5. Wang R, Li Z, Liu S, Zhang D. Global, regional and national burden of inflammatory bowel disease in 204 countries and territories from 1990 to 2019: A systematic analysis based on the Global Burden of Disease Study 2019. BMJ Open. 2023;13. doi:10.1136/bmjopen-2022-065186

6. Ellen Kuenzig M, Manuel DG, Donelle J, Benchimol EI. Life expectancy and health-adjusted life expectancy in people with inflammatory bowel disease. CMAJ. 2020;192. doi:10.1503/cmaj.190976

7. Lakatos PL, Lakatos L. Risk for colorectal cancer in ulcerative colitis: Changes, causes and management strategies. World Journal of Gastroenterology. 2008. pp. 3937–3947. doi:10.3748/wjg.14.3937

8. Conrad K, Roggenbuck D, Laass MW. Diagnosis and classification of ulcerative colitis. Autoimmunity Reviews. 2014. pp. 463–466. doi:10.1016/j.autrev.2014.01.028

9. Stolfi C, Pallone F, Monteleone G. Colorectal cancer chemoprevention by mesalazine and its derivatives. Journal of Biomedicine and Biotechnology. 2012. doi:10.1155/2012/980458

10. Karagozian R, Burakoff R. The role of mesalamine in the treatment of ulcerative colitis. Therapeutics and Clinical Risk Management. 2007. doi:10.24292/01.mf.0321.7

11. Burri E, Maillard MH, Schoepfer AM, Seibold F, Van Assche G, Rivière P, et al. Treatment Algorithm for Mild and Moderate-to-Severe Ulcerative Colitis: An Update. Digestion. 2020. pp. 2–15. doi:10.1159/000504092

12. Murray A, Nguyen TM, Parker CE, Feagan BG, MacDonald JK. Oral 5-aminosalicylic acid for maintenance of remission in ulcerative colitis. Cochrane Database of Systematic Reviews. 2020;2020. doi:10.1002/14651858.CD000544.pub5

13. Lightner AL, Strong SA. Medical Management of Ulcerative Colitis. The ASCRS Textbook of Colon and Rectal Surgery: Fourth Edition. 2021. pp. 765–779. doi:10.1007/978-3-030-66049-9_45

14. Bascones-Martinez A, Mattila R, Gomez-Font R, Meurman JH. Immunomodulatory drugs: Oral and systemic adverse effects. Med Oral Patol Oral Cir Bucal. 2014;19. doi:10.4317/medoral.19087

15. Gaies E, Jebabli N. Methotrexate Side Effects: Review Article. J Drug Metab Toxicol. 2012;3. doi:10.4172/2157-7609.1000125

16. Toscano E, Cotta J, Robles M, Lucena MI, Andrade RJ. Toxicidad hepática inducida por los nuevos fármacos inmunosupresores. Gastroenterol Hepatol. 2010;33: 54–65. doi:10.1016/j.gastrohep.2009.07.003

17. Ardizzone S, Cassinotti A, Manes G, Porro GB. Immunomodulators for all patients with inflammatory bowel disease? Therapeutic Advances in Gastroenterology. 2010. doi:10.1177/1756283X09354136

18. Chhibba T, Ma C. Is there room for immunomodulators in ulcerative colitis? Expert Opinion on Biological Therapy. 2020. pp. 379–390. doi:10.1080/14712598.2020.1708896

19. Awan H, Fatima U, Eaw R, Knox N, Alrubaiy L. The Efficacy of Currently Licensed Biologics for Treatment of Ulcerative Colitis: A Literature Review. Cureus. 2023. doi:10.7759/cureus.37609

20. Cima RR, Pemberton JH. Medical and surgical management of chronic ulcerative colitis. Archives of Surgery. 2005. pp. 300–310. doi:10.1001/archsurg.140.3.300

21. Travis SPL, Higgins PDR, Orchard T, Van Der Woude CJ, Panaccione R, Bitton A, et al. Review article: Defining remission in ulcerative colitis. Alimentary Pharmacology and Therapeutics. 2011. pp. 113–124. doi:10.1111/j.1365-2036.2011.04701.x

22. Yoon H, Jangi S, Dulai PS, Boland BS, Prokop LJ, Jairath V, et al. Incremental Benefit of Achieving Endoscopic and Histologic Remission in Patients With Ulcerative Colitis: A Systematic Review and Meta-Analysis. Gastroenterology. 2020;159. doi:10.1053/j.gastro.2020.06.043

23. Peyrin-Biroulet L, Bressenot A, Kampman W. Histologic Remission: The Ultimate Therapeutic Goal in Ulcerative Colitis? Clinical Gastroenterology and Hepatology. 2013/08/06. 2014;12: 929–34 e2. doi:10.1016/j.cgh.2013.07.022

24. Narang V, Kaur R, Garg B, Mahajan R, Midha V, Sood N, et al. Association of endoscopic and histological remission with clinical course in patients of ulcerative colitis. Intest Res. 2018;16: 55–61. doi:10.5217/ir.2018.16.1.55

25. Kruis W, Kiudelis G, Rácz I, Gorelov A, Pokrotnieks J, Horynski M, et al. Once daily versus three times daily mesalazine granules in active ulcerative colitis: A double-blind, double-dummy, randomised, non-inferiority trial. Gut. 2009;58. doi:10.1136/gut.2008.154302

26. Geboes K, Riddell R, Öst A, Jensfelt B, Persson T, Löfberg R. A reproducible grading scale for histological assessment of inflammation in ulcerative colitis. Gut. 2000/08/15. 2000;47: 404–409. doi:10.1136/gut.47.3.404

27. D'Haens G, Sandborn WJ, Feagan BG, Geboes K, Hanauer SB, Irvine EJ, et al. A Review of Activity Indices and Efficacy End Points for Clinical Trials of Medical Therapy in Adults With Ulcerative Colitis. Gastroenterology. 2007;132: 763–786. doi:10.1053/j.gastro.2006.12.038

28. Felice C, Pugliese D, Marzo M, Andrisani G, Nardone OM, Papa A, et al. Deep remission improves clinical outcomes after infliximab discontinuation in inflammatory bowel diseases. United European Gastroenterol J. 2014;2: A221. Available: http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=emed15&NEWS=N&AN=72261373

29. Rogler G, Vavricka S, Schoepfer A, Lakatos PL. Mucosal healing and deep remission: What does it mean? World J Gastroenterol. 2013/11/28. 2013;19: 7552–7560. doi:10.3748/wjg.v19.i43.7552

30. Norouzinia M, Chaleshi V, Alizadeh AHM, Zali MR. Biomarkers in inflammatory bowel diseases: Insight into diagnosis, prognosis and treatment. Gastroenterology and Hepatology from Bed to Bench. 2017. pp. 155–167. doi:10.22037/ghfbb.v0i0.1163

31. Argmann C, Hou R, Ungaro RC, Irizar H, Al-Taie Z, Huang R, et al. Biopsy and blood-based molecular biomarker of inflammation in IBD. Gut. 2023;72: 1271–1287. doi:10.1136/gutjnl-2021-326451

32. Ungaro R, Mehandru S, Allen PB, Peyrin-Biroulet L, Colombel JF. Ulcerative colitis. The Lancet. Lancet Publishing Group; 2017. pp. 1756–1770. doi:10.1016/S0140-6736(16)32126-2

33. Derakhshan Nazari MH, Shahrokh S, Ghanbari-Maman L, Maleknia S, Ghorbaninejad M, Meyfour A. Prediction of anti-TNF therapy failure in ulcerative colitis patients by

ensemble machine learning: A prospective study. Heliyon. 2023;9: e21154. doi:10.1016/j.heliyon.2023.e21154

34. Tighe D, Hall B, Jeyarajah SK, Smith S, Breslin N, Ryan B, et al. One-Year Clinical Outcomes in an IBD Cohort Who Have Previously Had Anti-TNFa Trough and Antibody Levels Assessed. Inflamm Bowel Dis. 2017;23: 1154–1159. doi:10.1097/MIB.0000000000001093

35. Cai Z, Wang S, Li J. Treatment of Inflammatory Bowel Disease: A Comprehensive Review. Frontiers in Medicine. 2021. doi:10.3389/fmed.2021.765474

36. Monstad I, Hovde Ø, Solberg IC, Moum BA. Clinical course and prognosis in ulcerative colitis: Results from population-based and observational studies. Annals of Gastroenterology. 2014. pp. 95–104.

37. Shivaji UN, Bazarova A, Critchlow T, Smith SCL, Nardone OM, Love M, et al. Clinical outcomes, predictors of prognosis and health economics consequences in IBD patients after discontinuation of the first biological therapy. Therap Adv Gastroenterol. 2020;13. doi:10.1177/1756284820981216

38. Jangi S, Holmer AK, Dulai PS, Boland BS, Collins AE, Pham L, et al. Risk of Relapse in Patients with Ulcerative Colitis with Persistent Endoscopic Healing: A Durable Treatment Endpoint. J Crohns Colitis. 2021;15: 567–574. doi:10.1093/ecco-jcc/jjaa184

39. Raine T, Bonovas S, Burisch J, Kucharzik T, Adamina M, Annese V, et al. ECCO Guidelines on Therapeutics in Ulcerative Colitis: Medical Treatment. J Crohns Colitis. 2022;16. doi:10.1093/ecco-jcc/jjab178

40. Singh N, Bernstein CN. Environmental risk factors for inflammatory bowel disease. United European Gastroenterology Journal. 2022. pp. 1047–1053. doi:10.1002/ueg2.12319

41. Childers RE, Eluri S, Vazquez C, Weise RM, Bayless TM, Hutfless S. Family history of inflammatory bowel disease among patients with ulcerative colitis: A systematic review

and meta-analysis. J Crohns Colitis. 2014;8: 1480–1497. doi:10.1016/j.crohns.2014.05.008

42. Saravanarajan K, Douglas AR, Ismail MS, Omorogbe J, Semenov S, Muphy G, et al. Genomic profiling of intestinal T-cell receptor repertoires in inflammatory bowel disease. Genes Immun. 2020;21. doi:10.1038/s41435-020-0092-x

43. Turpin W, Goethel A, Bedrani L, Croitoru K. Determinants of IBD Heritability: Genes, Bugs, and More. Inflammatory Bowel Diseases. 2018. doi:10.1093/ibd/izy085

44. Cleynen I, Boucher G, Jostins L, Schumm LP, Zeissig S, Ahmad T, et al. Inherited determinants of Crohn's disease and ulcerative colitis phenotypes: A genetic association study. The Lancet. 2016;387. doi:10.1016/S0140-6736(15)00465-1

45. 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.

46. Dorofeyev AE, Dorofeyeva AA, Kiriyan EA, Rassokhina OA, Dynia YZ. Genetic polymorphism in patients with early and late onset of ulcerative colitis. Wiad Lek. 2020;73. doi:10.36740/wlek202001116

47. Xia Y, Li J, Gao H. Association Between TLR2, TLR4 and NOD2/CARD15 Gene Polymorphisms and Ulcerative Colitis: A Meta-analysis. Chinese Journal of Gastroenterology. 2021;26. doi:10.3969/j.issn.1008-7125.2021.02.004

48. Stokkers PCF, Reitsma PH, Tytgat GNJ, Van Deventer SJH. HLA-DR and -DQ phenotypes in inflammatory bowel disease: A meta-analysis. Gut. 1999;45. doi:10.1136/gut.45.3.395

49. McLean MH, Neurath MF, Durum SK. Targeting interleukins for the treatment of inflammatory bowel disease - What lies beyond anti-TNF therapy? Inflammatory Bowel Diseases. 2014. pp. 389–397. doi:10.1097/01.MIB.0000437616.37000.41

50. Stallhofer J, Friedrich M, Konrad-Zerna A, Wetzke M, Lohse P, Glas J, et al. Lipocalin-2 Is a Disease Activity Marker in Inflammatory Bowel Disease Regulated by IL-17A, IL-22, and TNF-alpha and Modulated by IL23R Genotype Status. Inflamm Bowel Dis. 2015/08/12. 2015;21: 2327–2340. doi:10.1097/MIB.0000000000000515

51. Barrett JC, Lee JC, Lees CW, Prescott NJ, Anderson CA, Phillips A, et al. Genome-wide association study of ulcerative colitis identifies three new susceptibility loci, including the HNF4A region. Nat Genet. 2009;41: 1330–1334. doi:10.1038/ng.483

52. Sarlos P, Kovesdi E, Magyari L, Banfai Z, Szabo A, Javorhazy A, et al. Genetic update on inflammatory factors in ulcerative colitis: Review of the current literature. World J Gastrointest Pathophysiol. 2014;5. doi:10.4291/wjgp.v5.i3.304

53. Carreras-Torres R, Ibáñez-Sanz G, Obón-Santacana M, Duell EJ, Moreno V. Identifying environmental risk factors for inflammatory bowel diseases: a Mendelian randomization study. Sci Rep. 2020;10. doi:10.1038/s41598-020-76361-2

54. Abegunde AT, Muhammad BH, Bhatti O, Ali T. Environmental risk factors for inflammatory bowel diseases: Evidence based literature review. World Journal of Gastroenterology. 2016. doi:10.3748/wjg.v22.i27.6296

55. MacGillivray DM, Kollmann TR. The role of environmental factors in modulating immune responses in early life. Frontiers in Immunology. 2014. doi:10.3389/fimmu.2014.00434

56. Radziszewska M, Smarkusz-zarzecka J, Ostrowska L, Pogodziński D. Nutrition and Supplementation in Ulcerative Colitis. Nutrients. 2022. doi:10.3390/nu14122469

57. Ponder A, Long MD. A clinical review of recent findings in the epidemiology of inflammatory bowel disease. Clinical Epidemiology. 2013. pp. 237–247. doi:10.2147/CLEP.S33961

58. Shen ZH, Zhu CX, Quan YS, Yang ZY, Wu S, Luo WW, et al. Relationship between intestinal microbiota and ulcerative colitis: Mechanisms and clinical application of

probiotics and fecal microbiota transplantation. World Journal of Gastroenterology. 2018. doi:10.3748/wjg.v24.i1.5

59. Gong D, Gong X, Wang L, Yu X, Dong Q. Involvement of Reduced Microbial Diversity in Inflammatory Bowel Disease. Gastroenterology Research and Practice. 2016. doi:10.1155/2016/6951091

60. Mills RH, Dulai PS, Vázquez-Baeza Y, Sauceda C, Daniel N, Gerner RR, et al. Multi-omics analyses of the ulcerative colitis gut microbiome link Bacteroides vulgatus proteases with disease severity. Nat Microbiol. 2022;7. doi:10.1038/s41564-021-01050-3

61. Angelberger S, Reinisch W, Makristathis A, Lichtenberger C, Dejaco C, Papay P, et al. Temporal bacterial community dynamics vary among ulcerative colitis patients after fecal microbiota transplantation. American Journal of Gastroenterology. 2013;108. doi:10.1038/ajg.2013.257

62. Tan P, Li X, Shen J, Feng Q. Fecal Microbiota Transplantation for the Treatment of Inflammatory Bowel Disease: An Update. Frontiers in Pharmacology. 2020. doi:10.3389/fphar.2020.574533

63. Lim SS, Vos T, Flaxman AD, Danaei G, Shibuya K, Adair-Rohani H, et al. A comparative risk assessment of burden of disease and injury attributable to 67 risk factors and risk factor clusters in 21 regions, 1990-2010: A systematic analysis for the Global Burden of Disease Study 2010. The Lancet. 2012;380: 2224–2260. doi:10.1016/S0140-6736(12)61766-8

64. Benchimol EI, Mack DR, Guttmann A, Nguyen GC, To T, Mojaverian N, et al. Inflammatory bowel disease in immigrants to Canada and their children: A population-based cohort study. American Journal of Gastroenterology. 2015;110. doi:10.1038/ajg.2015.52

65. Ho GT, Porter RJ, Kalla R. Ulcerative colitis: Recent advances in the understanding of disease pathogenesis. F1000Research. 2020. doi:10.12688/f1000research.20805.1

66. Coskun M. Intestinal epithelium in inflammatory bowel disease. Front Med (Lausanne). 2014;1. doi:10.3389/fmed.2014.00024

67. Wallaeys C, Garcia-Gonzalez N, Libert C. Paneth cells as the cornerstones of intestinal and organismal health: a primer. EMBO Mol Med. 2023;15. doi:10.15252/emmm.202216427

68. Seguella L, Gulbransen BD. Enteric glial biology, intercellular signalling and roles in gastrointestinal disease. Nature Reviews Gastroenterology and Hepatology. 2021. pp. 571–587. doi:10.1038/s41575-021-00423-7

69. Singh V, Johnson K, Yin J, Lee S, Lin R, Yu H, et al. Chronic Inflammation in Ulcerative Colitis Causes Long-Term Changes in Goblet Cell Function. CMGH. 2022;13. doi:10.1016/j.jcmgh.2021.08.010

70. Van Der Post S, Jabbar KS, Birchenough G, Arike L, Akhtar N, Sjovall H, et al. Structural weakening of the colonic mucus barrier is an early event in ulcerative colitis pathogenesis. Gut. 2019;68: 2142–2151. doi:10.1136/gutjnl-2018-317571

71. Kałużna A, Olczyk P, Komosińska-Vassev K. The Role of Innate and Adaptive Immune Cells in the Pathogenesis and Development of the Inflammatory Response in Ulcerative Colitis. Journal of Clinical Medicine. 2022. doi:10.3390/jcm11020400

72. Bennett JM, Reeves G, Billman GE, Sturmberg JP. Inflammation-nature's way to efficiently respond to all types of challenges: Implications for understanding and managing "the epidemic" of chronic diseases. Front Med (Lausanne). 2018;5. doi:10.3389/fmed.2018.00316

73. Bao L, Shi B, Shi YB. Intestinal homeostasis: A communication between life and death. Cell Biosci. 2020;10. doi:10.1186/s13578-020-00429-9

74. Boudry G, Yang PC, Perdue MH. Small intestine; Anatomy. Encyclopedia of Gastroenterology, Second Edition. 2019. pp. 482–486. doi:10.1016/B978-0-12-812460-4.00648-0

75. Newberry RD, Gustafsson JK. Anatomy and Function of the Gut Immune System. Encyclopedia of Immunobiology. 2016. pp. 427–433. doi:10.1016/B978-0-12-374279-7.07006-5

76. Okumura R, Takeda K. Maintenance of gut homeostasis by the mucosal immune system. Proceedings of the Japan Academy Series B: Physical and Biological Sciences. 2016. pp. 423–435. doi:10.2183/pjab.92.423

77. Geremia A, Biancheri P, Allan P, Corazza GR, Di Sabatino A. Innate and adaptive immunity in inflammatory bowel disease. Autoimmunity Reviews. 2014. pp. 3–10. doi:10.1016/j.autrev.2013.06.004

78. Mann ER, Li X. Intestinal antigen-presenting cells in mucosal immune homeostasis: Crosstalk between dendritic cells, macrophages and B-cells. World Journal of Gastroenterology. 2014. doi:10.3748/wjg.v20.i29.9653

79. Kmieć Z, Cyman M, Ślebioda TJ. Cells of the innate and adaptive immunity and their interactions in inflammatory bowel disease. Advances in Medical Sciences. 2017. doi:10.1016/j.advms.2016.09.001

80. Kordjazy N, Haj-Mirzaian A, Haj-Mirzaian A, Rohani MM, Gelfand EW, Rezaei N, et al. Role of toll-like receptors in inflammatory bowel disease. Pharmacological Research. 2018. pp. 204–215. doi:10.1016/j.phrs.2017.11.017

81. Marshall JS, Warrington R, Watson W, Kim HL. An introduction to immunology and immunopathology. Allergy, Asthma and Clinical Immunology. 2018. doi:10.1186/s13223-018-0278-1

82. Zhang N, Hartig H, Dzhagalov I, Draper D, He YW. The role of apoptosis in the development and function of T lymphocytes. Cell Research. 2005. pp. 749–769. doi:10.1038/sj.cr.7290345

83. Luckheeram RV, Zhou R, Verma AD, Xia B. CD4 +T cells: Differentiation and functions. Clinical and Developmental Immunology. 2012. doi:10.1155/2012/925135

84. Zhang Y, Shi W, Cao G, Li J, Wang H, Hao C. The significance of Th1,Th2,Th17and treg cells in the prediction and evaluation of ulcerative colitis. Eur J Inflamm. 2023;21. doi:10.1177/1721727X231167028

85. Zhao J, Lu Q, Liu Y, Shi Z, Hu L, Zeng Z, et al. Th17 Cells in Inflammatory Bowel Disease: Cytokines, Plasticity, and Therapies. Journal of Immunology Research. 2021. doi:10.1155/2021/8816041

86. Onuma EK, Amenta PS, Ramaswamy K, Lin JJC, Das KM. Autoimmunity in ulcerative colitis (UC): A predominant colonic mucosal B cell response against human tropomyosin isoform 5. Clin Exp Immunol. 2000;121: 466–471. doi:10.1046/j.1365-2249.2000.01330.x

87. Choy MC, Visvanathan K, De Cruz P. An overview of the innate and adaptive immune system in inflammatory bowel disease. Inflammatory Bowel Diseases. 2017. pp. 2–13. doi:10.1097/MIB.0000000000000955

88. Yeshi K, Ruscher R, Hunter L, Daly NL, Loukas A, Wangchuk P. Revisiting inflammatory bowel disease: Pathology, treatments, challenges and emerging therapeutics including drug leads from natural products. Journal of Clinical Medicine. 2020. doi:10.3390/jcm9051273

89. Gareb B, Otten AT, Frijlink HW, Dijkstra G, Kosterink JGW. Review: Local tumor necrosis factor-α inhibition in inflammatory bowel disease. Pharmaceutics. 2020. doi:10.3390/pharmaceutics12060539

90. Friedrich M, Pohin M, Powrie F. Cytokine Networks in the Pathophysiology of Inflammatory Bowel Disease. Immunity. 2019. doi:10.1016/j.immuni.2019.03.017

91. Gabay C. Interleukin-6 and chronic inflammation. Arthritis Res Ther. 2006;8. doi:10.1186/ar1917

92. Li Y, De Haar C, Chen M, Deuring J, Gerrits MM, Smits R, et al. Disease-related expression of the IL6/STAT3/SOCS3 signalling pathway in ulcerative colitis and

ulcerative colitis-related carcinogenesis. Gut. 2010;59: 227–235. doi:10.1136/gut.2009.184176

93. Kimura A, Kishimoto T. IL-6: Regulator of Treg/Th17 balance. European Journal of Immunology. 2010. pp. 1830–1835. doi:10.1002/eji.201040391

94. Atreya R, Neurath MF. Involvement of IL-6 in the pathogenesis of inflammatory bowel disease and colon cancer. Clinical Reviews in Allergy and Immunology. 2005. pp. 187–195. doi:10.1385/CRIAI:28:3:187

95. Imam T, Park S, Kaplan MH, Olson MR. Effector T helper cell subsets in inflammatory bowel diseases. Frontiers in Immunology. 2018. doi:10.3389/fimmu.2018.01212

96. Saadatdoust Z, Pandurangan AK, Ananda Sadagopan SK, Mohd. Esa N, Ismail A, Mustafa MR. Dietary cocoa inhibits colitis associated cancer: A crucial involvement of the IL-6/STAT3 pathway. Journal of Nutritional Biochemistry. 2015;26: 1547–1558. doi:10.1016/j.jnutbio.2015.07.024

97. Li MC, He SH. IL-10 and its related cytokines for treatment of inflammatory bowel disease. World Journal of Gastroenterology. 2004. pp. 620–625. doi:10.3748/wjg.v10.i5.620

98. Neurath MF. Targeting immune cell circuits and trafficking in inflammatory bowel disease. Nature Immunology. 2019. pp. 970–979. doi:10.1038/s41590-019-0415-0

99. Kucharzik T, Stoll R, Lugering N, Domschke W. Circulating antiinflammatory cytokine IL-10 in patients with inflammatory bowel disease (IBD). Clin Exp Immunol. 1995;100: 452–456. doi:10.1111/j.1365-2249.1995.tb03721.x

100. Bonasio R, Tu S, Reinberg D. Molecular signals of epigenetic states. Science. 2010. doi:10.1126/science.1191078

101. Henning SM, Wang P, Carpenter CL, Heber D. Epigenetic effects of green tea polyphenols in cancer. Epigenomics. 2013. pp. 729–741. doi:10.2217/epi.13.57

102. Filippini G, Farinotti M, Ferrarini M. Active and passive smoking during pregnancy and risk of central nervous system tumours in children. Paediatr Perinat Epidemiol. 2000;14: 78–84. doi:10.1046/j.1365-3016.2000.00230.x

103. Pierce R. Targeting Schistosome Histone Modifying Enzymes for Drug Development. Curr Pharm Des. 2012. doi:10.2174/138161212801327248

104. De Felice B, Manfellotto F, Palumbo A, Troisi J, Zullo F, Di Carlo C, et al. Genome-wide microRNA expression profiling in placentas from pregnant women exposed to BPA. BMC Med Genomics. 2015;8. doi:10.1186/s12920-015-0131-z

105. Luo D, Li X, Tang S, Song F, Li W, Xie G, et al. Epigenetic modifications in neuropathic pain. Molecular Pain. 2021. doi:10.1177/17448069211056767

106. Schuebel K, Gitik M, Domschke K, Goldman D. Making sense of epigenetics. International Journal of Neuropsychopharmacology. 2016. doi:10.1093/ijnp/pyw058

107. Toranõ EG, Garciá MG, Fernández-Morera JL, Ninõ-Garciá P, Fernández AF. The Impact of External Factors on the Epigenome: In Utero and over Lifetime. BioMed Research International. 2016. doi:10.1155/2016/2568635

108. Booth LN, Brunet A. The Aging Epigenome. Molecular Cell. 2016. doi:10.1016/j.molcel.2016.05.013

109. Vezzani B, Carinci M, Previati M, Giacovazzi S, Della Sala M, Gafà R, et al. Epigenetic Regulation: A Link between Inflammation and Carcinogenesis. Cancers. 2022. doi:10.3390/cancers14051221

110. Blackledge NP, Klose RJ. CpG island chromatin: A platform for gene regulation. Epigenetics. 2011;6: 147–152. doi:10.4161/epi.6.2.13640

111. Ventham NT, Kennedy NA, Adams AT, Kalla R, Heath S, O'Leary KR, et al. Integrative epigenome-wide analysis demonstrates that DNA methylation may mediate genetic risk in inflammatory bowel disease. Nat Commun. 2016;7. doi:10.1038/ncomms13507

112. Jin B, Li Y, Robertson KD. DNA methylation: Superior or subordinate in the epigenetic hierarchy? Genes and Cancer. 2011. pp. 607–617. doi:10.1177/1947601910393957

113. Mattei AL, Bailly N, Meissner A. DNA methylation: a historical perspective. Trends in Genetics. Elsevier Ltd; 2022. pp. 676–707. doi:10.1016/j.tig.2022.03.010

114. Jones PA, Takai D. The role of DNA methylation in mammalian epigenetics. Science. 2001. pp. 1068–1070. doi:10.1126/science.1063852

115. Lister R, Pelizzola M, Dowen RH, Hawkins RD, Hon G, Tonti-Filippini J, et al. Human DNA methylomes at base resolution show widespread epigenomic differences. Nature. 2009;462. doi:10.1038/nature08514

116. Ehrlich M. DNA hypermethylation in disease: mechanisms and clinical relevance. Epigenetics. 2019. pp. 1141–1163. doi:10.1080/15592294.2019.1638701

117. Xu X. DNA methylation and cognitive aging. Oncotarget. 2015;6: 13925–13934. doi:10.18632/oncotarget.4215

118. Li J, Li L, Wang Y, Huang G, Li X, Xie Z, et al. Insights Into the Role of DNA Methylation in Immune Cell Development and Autoimmune Disease. Frontiers in Cell and Developmental Biology. 2021. doi:10.3389/fcell.2021.757318

119. Ohkura N, Sakaguchi S. Transcriptional and epigenetic basis of Treg cell development and function: its genetic anomalies or variations in autoimmune diseases. Cell Research. 2020. pp. 465–474. doi:10.1038/s41422-020-0324-7

120. McLane LM, Abdel-Hakeem MS, Wherry EJ. CD8 T Cell Exhaustion During Chronic Viral Infection and Cancer. Annual Review of Immunology. 2019. pp. 457–495. doi:10.1146/annurev-immunol-041015-055318

121. Liu H, Li P, Wei Z, Zhang C, Xia M, Du Q, et al. Regulation of T cell differentiation and function by epigenetic modification enzymes. Seminars in Immunopathology. 2019. pp. 315–326. doi:10.1007/s00281-019-00731-w

122. Alt FW, Zhang Y, Meng FL, Guo C, Schwer B. Mechanisms of programmed DNA lesions and genomic instability in the immune system. Cell. 2013. pp. 417–429. doi:10.1016/j.cell.2013.01.007

123. Gasparetto M, Payne F, Nayak K, Kraiczy J, Glemas C, Philip-McKenzie Y, et al. Transcription and DNA Methylation Patterns of Blood-Derived CD8+ T Cells Are Associated With Age and Inflammatory Bowel Disease But Do Not Predict Prognosis. Gastroenterology. 2021;160: 232-244.e7. doi:10.1053/j.gastro.2020.08.017

124. Sullivan KE, Reddy ABM, Dietzmann K, Suriano AR, Kocieda VP, Stewart M, et al. Epigenetic Regulation of Tumor Necrosis Factor Alpha. Mol Cell Biol. 2007;27: 5147–5160. doi:10.1128/mcb.02429-06

125. Takahashi K, Sugi Y, Hosono A, Kaminogawa S. Epigenetic Regulation of TLR4 Gene Expression in Intestinal Epithelial Cells for the Maintenance of Intestinal Homeostasis. The Journal of Immunology. 2009;183: 6522–6529. doi:10.4049/jimmunol.0901271

126. Shuto T, Furuta T, Oba M, Xu H, Li J, Cheung J, et al. Promoter hypomethylation of Toll-like receptor-2 gene is associated with increased proinflammatory response towardbacterial peptidoglycan in cystic fibrosis bronchial epithelial cells. The FASEB Journal. 2006;20: 782–784. doi:10.1096/fj.05-4934fje

127. Taman H, Fenton CG, Hensel I V., Anderssen E, Florholmen J, Paulssen RH. Genome-wide DNA methylation intreatment-naïve ulcerative colitis. J Crohns Colitis. 2018;12: 1338–1347. doi:10.1093/ecco-jcc/jjy117

128. Glória L, Cravo M, Pinto A, De Sousa LS, Chaves P, Leitäo CN, et al. DNA hypomethylation and proliferative activity are increased in the rectal mucosa of patients with long-standing ulcerative colitis. Cancer. 1996;78: 2300–2306. doi:10.1002/(SICI)1097-0142(19961201)78:11<2300::AID-CNCR5>3.0.CO;2-Q

129. Taman H, Fenton CG, Anderssen E, Florholmen J, Paulssen RH. DNA hypo-methylation facilitates anti-inflammatory responses in severe ulcerative colitis. PLoS One. 2021;16. doi:10.1371/journal.pone.0248905

130. Tahara T, Shibata T, Nakamura M, Yamashita H, Yoshioka D, Okubo M, et al. Promoter methylation of protease-activated receptor (PAR2) is associated with severe clinical phenotypes of ulcerative colitis (UC). Clin Exp Med. 2009;9: 125–130. doi:10.1007/s10238-008-0025-x

131. Wielscher M, Mandaviya PR, Kuehnel B, Joehanes R, Mustafa R, Robinson O, et al. DNA methylation signature of chronic low-grade inflammation and its role in cardio-respiratory diseases. Nat Commun. 2022;13. doi:10.1038/s41467-022-29792-6

132. Scarpa M, Stylianou E. Epigenetics: Concepts and relevance to IBD pathogenesis. Inflammatory Bowel Diseases. 2012. pp. 1982–1996. doi:10.1002/ibd.22934

133. Chu X, Zhang B, Koeken VACM, Gupta MK, Li Y. Multi-Omics Approaches in Immunological Research. Frontiers in Immunology. 2021. doi:10.3389/fimmu.2021.668045

134. Saxena A, Carninci P. Whole transcriptome analysis: What are we still missing? Wiley Interdisciplinary Reviews: Systems Biology and Medicine. 2011. doi:10.1002/wsbm.135

135. Artika IM, Dewi YP, Nainggolan IM, Siregar JE, Antonjaya U. Real-Time Polymerase Chain Reaction: Current Techniques, Applications, and Role in COVID-19 Diagnosis. Genes. 2022. doi:10.3390/genes13122387

136. Kubista M, Andrade JM, Bengtsson M, Forootan A, Jonák J, Lind K, et al. The real-time polymerase chain reaction. Molecular Aspects of Medicine. 2006. doi:10.1016/j.mam.2005.12.007

137. VanGuilder HD, Vrana KE, Freeman WM. Twenty-five years of quantitative PCR for gene expression analysis. BioTechniques. 2008. doi:10.2144/000112776

138. Schena M, Shalon D, Davis RW, Brown PO. Quantitative Monitoring of Gene Expression Patterns with a Complementary DNA Microarray. Science (1979). 1995;270: 467–470. doi:10.1126/science.270.5235.467

139. Wang Z, Gerstein M, Snyder M. RNA-Seq: A revolutionary tool for transcriptomics. Nature Reviews Genetics. 2009. doi:10.1038/nrg2484

140. 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. doi:10.1371/journal.pone.0078644

141. Hong M, Tao S, Zhang L, Diao LT, Huang X, Huang S, et al. RNA sequencing: new technologies and applications in cancer research. Journal of Hematology and Oncology. 2020. doi:10.1186/s13045-020-01005-x

142. Mortazavi A, Williams BA, McCue K, Schaeffer L, Wold B. Mapping and quantifying mammalian transcriptomes by RNA-Seq. Nat Methods. 2008;5. doi:10.1038/nmeth.1226

143. Li Y, Tollefsbol TO. DNA Methylation Detection: Bisulfite Genomic Sequencing Analysis. Methods in Molecular Biology. 2011. doi:10.1007/978-1-61779-316-5_2

144. Laird PW. Principles and challenges of genome-wide DNA methylation analysis. Nature Reviews Genetics. 2010. doi:10.1038/nrg2732

145. Sarda S, Hannenhalli S. Next-Generation Sequencing and Epigenomics Research: A Hammer in Search of Nails. Genomics Inform. 2014;12: 2. doi:10.5808/gi.2014.12.1.2

146. Herman JG, Graff JR, Myöhänen S, Nelkin BD, Baylin SB. Methylation-specific PCR: A novel PCR assay for methylation status of CpG islands. Proc Natl Acad Sci U S A. 1996;93. doi:10.1073/pnas.93.18.9821

147. Esteller M. CpG island hypermethylation and tumor suppressor genes: A booming present, a brighter future. Oncogene. 2002. pp. 5427–5440. doi:10.1038/sj.onc.1205600

148. Non AL, Thayer ZM. Epigenetics for anthropologists: An introduction to methods. American Journal of Human Biology. 2015;27: 295–303. doi:10.1002/ajhb.22679

149. Wilhelm-Benartzi CS, Koestler DC, Karagas MR, Flanagan JM, Christensen BC, Kelsey KT, et al. Review of processing and analysis methods for DNA methylation array data. British Journal of Cancer. 2013. doi:10.1038/bjc.2013.496

150. Moran S, Arribas C, Esteller M. Validation of a DNA methylation microarray for 850,000 CpG sites of the human genome enriched in enhancer sequences. Epigenomics. 2016;8. doi:10.2217/epi.15.114

151. Heiss JA, Brennan KJ, Baccarelli AA, Téllez-Rojo MM, Estrada-Gutiérrez G, Wright RO, et al. Battle of epigenetic proportions: comparing Illumina's EPIC methylation microarrays and TruSeq targeted bisulfite sequencing. Epigenetics. 2020;15. doi:10.1080/15592294.2019.1656159

152. Paun O, Verhoeven KJF, Richards CL. Opportunities and limitations of reduced representation bisulfite sequencing in plant ecological epigenomics. New Phytologist. 2019. pp. 738–742. doi:10.1111/nph.15388

153. Veillard AC, Datlinger P, Laczik M, Squazzo S, Bock C. Diagenode® Premium RRBS technology: cost-effective DNA methylation mapping with superior coverage. Nat Methods. 2016;13. doi:10.1038/NMETH.F.391

154. Meissner A, Gnirke A, Bell GW, Ramsahoye B, Lander ES, Jaenisch R. Reduced representation bisulfite sequencing for comparative high-resolution DNA methylation analysis. Nucleic Acids Res. 2005;33. doi:10.1093/nar/gki901

155. Gu H, Smith ZD, Bock C, Boyle P, Gnirke A, Meissner A. Preparation of reduced representation bisulfite sequencing libraries for genome-scale DNA methylation profiling. Nat Protoc. 2011;6. doi:10.1038/nprot.2010.190

156. Paun O, Verhoeven KJF, Richards CL. Opportunities and limitations of reduced representation bisulfite sequencing in plant ecological epigenomics. New Phytologist. 2019. pp. 738–742. doi:10.1111/nph.15388

157. Zhang H, Liu L, Li M. Mini-review of DNA Methylation Detection Techniques and Their Potential Applications in Disease Diagnosis, Prognosis, and Treatment. ACS Sensors. 2024. doi:10.1021/acssensors.3c02328

158. Grunau C, Clark SJ, Rosenthal A. Bisulfite genomic sequencing: systematic investigation of critical experimental parameters. Nucleic Acids Res. 2001;29: 65–65. doi:10.1093/nar/29.13.e65

159. Liu Y, Rosikiewicz W, Pan Z, Jillette N, Wang P, Taghbalout A, et al. DNA methylation-calling tools for Oxford Nanopore sequencing: a survey and human epigenome-wide evaluation. Genome Biol. 2021;22. doi:10.1186/s13059-021-02510-z

160. Suzuki M, Liao W, Wos F, Johnston AD, DeGrazia J, Ishii J, et al. Whole-genome bisulfite sequencing with improved accuracy and cost. Genome Res. 2018;28. doi:10.1101/gr.232587.117

161. Stirzaker C, Taberlay PC, Statham AL, Clark SJ. Mining cancer methylomes: Prospects and challenges. Trends in Genetics. 2014. doi:10.1016/j.tig.2013.11.004

162. Huang Y, Pastor WA, Shen Y, Tahiliani M, Liu DR, Rao A. The behaviour of 5-hydroxymethylcytosine in bisulfite sequencing. PLoS One. 2010;5. doi:10.1371/journal.pone.0008888

163. Kint S, De Spiegelaere W, De Kesel J, Vandekerckhove L, Van Criekinge W. Evaluation of bisulfite kits for DNA methylation profiling in terms of DNA fragmentation and DNA recovery using digital PCR. PLoS One. 2018;13. doi:10.1371/journal.pone.0199091

164. Simpson JT, Workman RE, Zuzarte PC, David M, Dursi LJ, Timp W. Detecting DNA cytosine methylation using nanopore sequencing. Nat Methods. 2017;14. doi:10.1038/nmeth.4184

165. Sun YM, Chen YQ. Principles and innovative technologies for decrypting noncoding RNAs: From discovery and functional prediction to clinical application. Journal of Hematology and Oncology. 2020. doi:10.1186/s13045-020-00945-8

166. Huang Z hao, Du Y ping, Wen J tao, Lu B feng, Zhao Y. snoRNAs: functions and mechanisms in biological processes, and roles in tumor pathophysiology. Cell Death Discovery. 2022. doi:10.1038/s41420-022-01056-8

167. Ponting CP, Oliver PL, Reik W. Evolution and Functions of Long Noncoding RNAs. Cell. 2009. pp. 629–641. doi:10.1016/j.cell.2009.02.006

168. Li K, Tian Y, Yuan Y, Fan X, Yang M, He Z, et al. Insights into the functions of LncRNAs in drosophila. International Journal of Molecular Sciences. 2019. doi:10.3390/ijms20184646

169. Zhu J, Yu W, Wang Y, Xia K, Huang Y, Xu A, et al. LncRNAs: Function and mechanism in cartilage development, degeneration, and regeneration. Stem Cell Research and Therapy. 2019. doi:10.1186/s13287-019-1458-8

170. Huang W, Li X, Huang C, Tang Y, Zhou Q, Chen W. LncRNAs and Rheumatoid Arthritis: From Identifying Mechanisms to Clinical Investigation. Frontiers in Immunology. 2022. doi:10.3389/fimmu.2021.807738

171. Bocchetti M, Scrima M, Melisi F, Luce A, Sperlongano R, Caraglia M, et al. Lncrnas and immunity: Coding the immune system with noncoding oligonucleotides. International Journal of Molecular Sciences. 2021. pp. 1–30. doi:10.3390/ijms22041741

172. Yarani R, Mirza AH, Kaur S, Pociot F. The emerging role of lncrnas in inflammatory bowel disease. Experimental and Molecular Medicine. 2018. doi:10.1038/s12276-018-0188-9

173. Ray MK, Fenton CG, Paulssen RH. Novel long non-coding RNAs of relevance for ulcerative colitis pathogenesis. Noncoding RNA Res. 2022;7. doi:10.1016/j.ncrna.2022.02.001

174. Padua D, Mahurkar-Joshi S, Law IKM, Polytarchou C, Vu JP, Pisegna JR, 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. doi:10.1152/ajpgi.00212.2016

175. Zhu M, Xie J. LncRNA MALAT1 Promotes Ulcerative Colitis by Upregulating lncRNA ANRIL. Dig Dis Sci. 2020;65: 3191–3196. doi:10.1007/s10620-020-06093-w

176. 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;11: 649–670. doi:10.1093/ecco-jcc/jjx008

177. Nilsson Å, Danielsson Å, Löfberg R, Benno P, Bergman L, Fausa O, et al. Olsalazine Versus Sulphasalazine for Relapse Prevention in Ulcerative Colitis: A Multicenter Study. Am J Gastroenterol. 1995;90. doi:10.1111/j.1572-0241.1995.tb08231.x

178. 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. doi:10.1080/00365520701409035

179. Wendt J, Rosenbaum H, Richmond TA, Jeddeloh JA, Burgess DL. Targeted bisulfite sequencing using the SeqCap epi enrichment system. Methods in Molecular Biology. 2018. doi:10.1007/978-1-4939-7481-8_20

180. Ewels P, Magnusson M, Lundin S, Kaller M. MultiQC: summarize analysis results for multiple tools and samples in a single report. Bioinformatics. 2016/06/18. 2016;32: 3047–3048. doi:10.1093/bioinformatics/btw354

181. 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. doi:10.1093/bioinformatics/bts635

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

183. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. Nat Methods. 2012;9: 357–359. doi:10.1038/nmeth.1923

184. Krueger F, Andrews SR. Bismark: A flexible aligner and methylation caller for Bisulfite-Seq applications. Bioinformatics. 2011;27: 1571–1572. doi:10.1093/bioinformatics/btr167

185. Korthauer K, Chakraborty S, Benjamini Y, Irizarry RA. Detection and accurate false discovery rate control of differentially methylated regions from whole genome bisulfite sequencing. Biostatistics. 2019;20: 367–383. doi:10.1093/biostatistics/kxy007

186. Kendall MG. A New Measure of Rank Correlation. Biometrika. 1938;30: 81. doi:10.2307/2332226

187. Team BC, Maintainer BP. TxDb.Hsapiens.UCSC.hg38.knownGene: Annotation package for TxDb object(s). In: R package version 3.4.6 [Internet]. 2019. Available: 10.18129/B9.bioc.TxDb.Hsapiens.UCSC.hg38.knownGene

188. Wold S, Sjöström M, Eriksson L. PLS-regression: A basic tool of chemometrics. Chemometrics and Intelligent Laboratory Systems. 2001;58: 109–130. doi:10.1016/S0169-7439(01)00155-1

189. Gray RJ. Modeling Survival Data: Extending the Cox Model. Journal of the American Statistical Association. 2002. doi:10.1198/jasa.2002.s447

190. Kropko J, Harden JJ. Beyond the Hazard Ratio: Generating Expected Durations from the Cox Proportional Hazards Model. Br J Polit Sci. 2020;50: 303–320. doi:10.1017/S000712341700045X

191. Bastien P, Bertrand F, Meyer N, Maumy-Bertrand M. Deviance residuals-based sparse PLS and sparse kernel PLS regression for censored data. Bioinformatics. 2015;31: 397–404. doi:10.1093/bioinformatics/btu660

192. Yu G, Wang L-G, Han Y, He Q-Y. clusterProfiler: an R Package for Comparing Biological Themes Among Gene Clusters. OMICS. 2012;16: 284–287. doi:10.1089/omi.2011.0118

193. Mi H, Thomas P. PANTHER pathway: an ontology-based pathway database coupled with data analysis tools. Methods Mol Biol. 2009;563: 123–140. doi:10.1007/978-1-60761-175-2_7

194. Yu G, He QY. ReactomePA: An R/Bioconductor package for reactome pathway analysis and visualization. Mol Biosyst. 2016;12: 477–479. doi:10.1039/c5mb00663e

195. Yu G, Wang LG, Han Y, He QY. ClusterProfiler: An R package for comparing biological themes among gene clusters. OMICS. 2012/03/30. 2012;16: 284–287. doi:10.1089/omi.2011.0118

196. Zheng SC, Breeze CE, Beck S, Teschendorff AE. Identification of differentially methylated cell types in epigenome-wide association studies. Nat Methods. 2018;15. doi:10.1038/s41592-018-0213-x

197. Lury DA, Fisher RA. Statistical Methods for Research Workers. The Statistician. 1972;21: 229. doi:10.2307/2986695

198. Brown S, Melamed L. T Test. Experimental Design and Analysis. 2012. pp. 10–12. doi:10.4135/9781412984218.n3

199. Tukey JW. Comparing Individual Means in the Analysis of Variance. Biometrics. 1949;5: 99. doi:10.2307/3001913

200. Benjamini Y, Hochberg Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. Journal of the Royal Statistical Society: Series B (Methodological). 1995;57: 289–300. doi:10.1111/j.2517-6161.1995.tb02031.x

201. Al-Qahtani AA, Alhamlan FS, Al-Qahtani AA. Pro-Inflammatory and Anti-Inflammatory Interleukins in Infectious Diseases: A Comprehensive Review. Tropical Medicine and Infectious Disease. 2024. doi:10.3390/tropicalmed9010013

202. Lee JY, Tsolis RM, Bäumler AJ. The microbiome and gut homeostasis. Science. 2022. doi:10.1126/science.abp9960

203. Riley SA, Mani V, Goodman MJ, Lucas S. Why do patients with ulcerative colitis relapse? Gut. 1990;31: 179–183. doi:10.1136/gut.31.2.179

204. Bruserud Ø, Aarstad HH, Tvedt THA. Combined c-reactive protein and novel inflammatory parameters as a predictor in cancer—what can we learn from the hematological experience? Cancers (Basel). 2020;12: 1–23. doi:10.3390/cancers12071966

205. Mosli MH, Feagan BG, Sandborn WJ, D'Haens G, Behling C, Kaplan K, et al. Histologic evaluation of ulcerative colitis: A systematic review of disease activity indices. Inflammatory Bowel Diseases. 2014. doi:10.1097/01.MIB.0000437986.00190.71

206. Fenton CG, Taman H, Florholmen J, Sørbye SW, Paulssen RH. Transcriptional Signatures That Define Ulcerative Colitis in Remission. Inflamm Bowel Dis. 2021;27: 94–105. doi:10.1093/ibd/izaa075

207. Adamowicz M, Milkiewicz P, Kempinska-Podhorodecka A. 5-AMINOSALICYLIC ACID INHIBITS THE EXPRESSION OF ONCOmiRS AND PRO-INFLAMMATORY microRNAS: AN IN VITRO STUDY. Journal of Physiology and Pharmacology. 2021;72: 529–535. doi:10.26402/jpp.2021.4.04

208. Karatzas PS, Gazouli M, Safioleas M, Mantzaris GJ. DNA methylation changes in inflammatory bowel disease. Annals of Gastroenterology. 2014. pp. 125–132.

209. Planell N, Lozano JJ, Mora-Buch R, Masamunt MC, Jimeno M, Ordás I, et al. Transcriptional analysis of the intestinal mucosa of patients with ulcerative colitis in remission reveals lasting epithelial cell alterations. Gut. 2012/11/09. 2013;62: 967–976. doi:10.1136/gutjnl-2012-303333

210. Liu J, Huang B, Ding F, Li Y. Environment factors, DNA methylation, and cancer. Environmental Geochemistry and Health. 2023. doi:10.1007/s10653-023-01749-8

211. Howell KJ, Kraiczy J, Nayak KM, Gasparetto M, Ross A, Lee C, et al. DNA Methylation and Transcription Patterns in Intestinal Epithelial Cells From Pediatric Patients With

Inflammatory Bowel Diseases Differentiate Disease Subtypes and Associate With Outcome. Gastroenterology. 2018;154. doi:10.1053/j.gastro.2017.10.007

212. Wreczycka K, Gosdschan A, Yusuf D, Grüning B, Assenov Y, Akalin A. Strategies for analyzing bisulfite sequencing data. Journal of Biotechnology. 2017. pp. 105–115. doi:10.1016/j.jbiotec.2017.08.007

213. Gao ZJ, Li WP, Mao XT, Huang T, Wang HL, Li YN, et al. Single-nucleotide methylation specifically represses type I interferon in antiviral innate immunity. Journal of Experimental Medicine. 2021;218. doi:10.1084/JEM.20201798

214. Harris RA, Nagy-Szakal D, Mir SAV, Frank E, Szigeti R, Kaplan JL, et al. DNA methylation-associated colonic mucosal immune and defense responses in treatmentnaïve pediatric ulcerative colitis. Epigenetics. 2014/06/18. 2014;9: 1131–1137. doi:10.4161/epi.29446

215. Sanchez R, Yang X, Maher T, Mackenzie SA. Discrimination of DNA methylation signal from background variation for clinical diagnostics. Int J Mol Sci. 2019;20. doi:10.3390/ijms20215343

216. Park Y, Figueroa ME, Rozek LS, Sartor MA. MethylSig: A whole genome DNA methylation analysis pipeline. Bioinformatics. 2014;30: 2414–2422. doi:10.1093/bioinformatics/btu339

217. Akalin A, Kormaksson M, Li S, Garrett-Bakelman FE, Figueroa ME, Melnick A, et al. MethylKit: a comprehensive R package for the analysis of genome-wide DNA methylation profiles. Genome Biol. 2012;13. doi:10.1186/gb-2012-13-10-R87

218. Hansen KD, Langmead B, Irizarry RA. BSmooth: from whole genome bisulfite sequencing reads to differentially methylated regions. Genome Biol. 2012;13. doi:10.1186/gb-2012-13-10-R83

219. Crary-Dooley FK, Tam ME, Dunaway KW, Hertz-Picciotto I, Schmidt RJ, LaSalle JM. A comparison of existing global DNA methylation assays to low-coverage whole-genome bisulfite sequencing for epidemiological studies. Epigenetics. 2017;12: 206–214. doi:10.1080/15592294.2016.1276680

220. Libertini E, Heath SC, Hamoudi RA, Gut M, Ziller MJ, Czyz A, et al. Information recovery from low coverage whole-genome bisulfite sequencing. Nat Commun. 2016;7. doi:10.1038/ncomms11306

221. Qin Z, Li B, Conneely KN, Wu H, Hu M, Ayyala D, et al. Statistical Challenges in Analyzing Methylation and Long-Range Chromosomal Interaction Data. Stat Biosci. 2016;8: 284–309. doi:10.1007/s12561-016-9145-0

222. Han H, Cortez CC, Yang X, Nichols PW, Jones PA, Liang G. DNA methylation directly silences genes with non-CpG island promoters and establishes a nucleosome occupied promoter. Hum Mol Genet. 2011;20: 4299–4310. doi:10.1093/hmg/ddr356

223. Eckhardt F, Lewin J, Cortese R, Rakyan VK, Attwood J, Burger M, et al. DNA methylation profiling of human chromosomes 6, 20 and 22. Nat Genet. 2006;38: 1378–1385. doi:10.1038/ng1909

224. Richmond AL, Kabi A, Homer CR, Marina-García N, Nickerson KP, Nesvizhskii AI, et al. The nucleotide synthesis enzyme CAD inhibits NOD2 antibacterial function in human intestinal epithelial cells. Gastroenterology. 2012;142. doi:10.1053/j.gastro.2012.02.040

225. Huang J, Xiao L, Gong X, Shao W, Yin Y, Liao Q, et al. Cytokine-like molecule CCDC134 contributes to CD8+ T-cell effector functions in cancer immunotherapy. Cancer Res. 2014;74: 5734–5745. doi:10.1158/0008-5472.CAN-13-3132

226. Liu Y, Wang JX, Nie ZY, Wen Y, Jia XJ, Zhang LN, et al. Upregulation of ERp57 promotes clear cell renal cell carcinoma progression by initiating a STAT3/ILF3 feedback loop. Journal of Experimental and Clinical Cancer Research. 2019;38. doi:10.1186/s13046-019-1453-z

227. Umar S. Intestinal stem cells. Curr Gastroenterol Rep. 2010;12: 340–348. doi:10.1007/s11894-010-0130-3

228. Gustafsson JK, Hansson GC, Sjövall H. Ulcerative colitis patients in remission have an altered secretory capacity in the proximal colon despite macroscopically normal

mucosa. Neurogastroenterology and Motility. 2012;24. doi:10.1111/j.1365-2982.2012.01958.x

229. Dotti I, Mora-Buch R, Ferrer-Picón E, Planell N, Jung P, Masamunt MC, et al. Alterations in the epithelial stem cell compartment could contribute to permanent changes in the mucosa of patients with ulcerative colitis. Gut. 2017;66: 2069–2079. doi:10.1136/gutjnl-2016-312609

230. Novak EA, Mollen KP. Mitochondrial dysfunction in inflammatory bowel disease. Frontiers in Cell and Developmental Biology. 2015. doi:10.3389/fcell.2015.00062

231. Bachar AR, Scheffer L, Schroeder AS, Nakamura HK, Cobb LJ, Oh YK, et al. Humanin is expressed in human vascular walls and has a cytoprotective effect against oxidized LDL-induced oxidative stress. Cardiovasc Res. 2010;88: 360–366. doi:10.1093/cvr/cvq191

232. Lin W, Ma C, Su F, Jiang Y, Lai R, Zhang T, et al. Raf kinase inhibitor protein mediates intestinal epithelial cell apoptosis and promotes IBDs in humans and mice. Gut. 2017;66: 597–610. doi:10.1136/gutjnl-2015-310096

233. Di Cunto F, Calautti E, Hsiao J, Ong L, Topley G, Turco E, et al. Citron Rho-interacting kinase, a novel tissue-specific Ser/Thr kinase encompassing the Rho-Rac-binding protein citron. Journal of Biological Chemistry. 1998;273: 29706–29711. doi:10.1074/jbc.273.45.29706

234. Zhang J, Han X, Jiang L, Han Z, Wang Z. LncRNA CERNA2 is an independent predictor for clinical prognosis and is related to tumor development in gastric cancer. Int J Clin Exp Pathol. 2018;11: 5783–5791. Available: http://www.ncbi.nlm.nih.gov/pubmed/31949664%0Ahttp://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC6963070

235. Ma D, Zhou X, Wang Y, Dai L, Yuan J, Peng J, et al. Changes in the Small Noncoding RNAome During M1 and M2 Macrophage Polarization. Front Immunol. 2022;13. doi:10.3389/fimmu.2022.799733

236. Atri C, Guerfali FZ, Laouini D. Role of human macrophage polarization in inflammation during infectious diseases. International Journal of Molecular Sciences. 2018. doi:10.3390/ijms19061801

237. Mills CD. M1 and M2 macrophages: Oracles of health and disease. Crit Rev Immunol. 2012;32: 463–488. doi:10.1615/CritRevImmunol.v32.i6.10

238. Lopes MF da S, Felix J de S, Scaramele NF, Almeida MC, Furlan A de O, Troiano JA, et al. Co-expression analysis of lncRNA and mRNA identifies potential adipogenesis regulatory non-coding RNAs involved in the transgenerational effects of tributyltin. PLoS One. 2023;18: e0281240. doi:10.1371/journal.pone.0281240

239. Lu JW, Rouzigu A, Teng LH, Liu WL. The Construction and Comprehensive Analysis of Inflammation-Related ceRNA Networks and Tissue-Infiltrating Immune Cells in Ulcerative Progression. Biomed Res Int. 2021;2021. doi:10.1155/2021/6633442

240. Chen F, Hu Y, Fan YH, Lv B. Clinical Value of Fecal Calprotectin in Predicting Mucosal Healing in Patients With Ulcerative Colitis. Front Med (Lausanne). 2021;8. doi:10.3389/fmed.2021.679264

241. Malvão LDR, Madi K, Esberard BC, De Amorim RF, Silva KDS, Farias E Silva K, et al. Fecal calprotectin as a noninvasive test to predict deep remission in patients with ulcerative colitis. Medicine (United States). 2021;100. doi:10.1097/MD.0000000000024058

242. McGovern DPB, Gardet A, Törkvist L, Goyette P, Essers J, Taylor KD, et al. Genome-wide association identifies multiple ulcerative colitis susceptibility loci. Nat Genet. 2010;42: 332–337. doi:10.1038/ng.549

243. Haritunians T, Taylor KD, Targan SR, Dubinsky M, Ippoliti A, Kwon S, et al. Genetic predictors of medically refractory ulcerative colitis. Inflamm Bowel Dis. 2010;16: 1830–1840. doi:10.1002/ibd.21293

244. Hong D, Kim HK, Yang W, Yoon C, Kim M, Yang C-S, et al. Integrative analysis of single-cell RNA-seq and gut microbiome metabarcoding data elucidates macrophage

dysfunction in mice with DSS-induced ulcerative colitis. Commun Biol. 2024;7: 731. doi:10.1038/s42003-024-06409-w

245. Dai YC, Qiao D, Fang CY, Chen QQ, Que RY, Xiao TG, et al. Single-cell RNA-sequencing combined with bulk RNA-sequencing analysis of peripheral blood reveals the characteristics and key immune cell genes of ulcerative colitis. World J Clin Cases. 2022;10. doi:10.12998/wjcc.v10.i33.12116

246. Haberman Y, Karns R, Dexheimer PJ, Schirmer M, Somekh J, Jurickova I, et al. Ulcerative colitis mucosal transcriptomes reveal mitochondriopathy and personalized mechanisms underlying disease severity and treatment response. Nat Commun. 2019/01/04. 2019;10: 38. doi:10.1038/s41467-018-07841-3

247. Xiao H, Feng X, Liu M, Gong H, Zhou X. SnoRNA and lncSNHG: Advances of nucleolar small RNA host gene transcripts in anti-tumor immunity. Frontiers in Immunology. 2023. doi:10.3389/fimmu.2023.1143980

248. Dubey G, Singh M, Singh H, Agarwal M, Chandel SS, Mishra A, et al. Emerging roles of SnoRNAs in the pathogenesis and treatment of autoimmune disorders. Pathology Research and Practice. 2024. doi:10.1016/j.prp.2023.154952

249. Borren NZ, Plichta D, Joshi AD, Bonilla G, Sadreyev R, Vlamakis H, et al. Multi-omics" profiling in patients with quiescent inflammatory bowel disease identifies biomarkers predicting relapse. Inflamm Bowel Dis. 2020;26. doi:10.1093/ibd/izaa183

250. Toiyama Y, Okugawa Y, Tanaka K, Araki T, Uchida K, Hishida A, et al. A Panel of Methylated MicroRNA Biomarkers for Identifying High-Risk Patients With Ulcerative Colitis-Associated Colorectal Cancer. Gastroenterology. 2017;153. doi:10.1053/j.gastro.2017.08.037

251. Zhou Q, Shen ZF, Wu BS, Xu CB, He ZQ, Chen T, et al. Risk of Colorectal Cancer in Ulcerative Colitis Patients: A Systematic Review and Meta-Analysis. Gastroenterol Res Pract. 2019;2019. doi:10.1155/2019/5363261

252. Kim MS, Lee J, Sidransky D. DNA methylation markers in colorectal cancer. Cancer and Metastasis Reviews. 2010. doi:10.1007/s10555-010-9207-6

253. Pekow J, Hernandez K, Meckel K, Deng Z, Haider HI, Khalil A, et al. IBD-associated Colon Cancers Differ in DNA Methylation and Gene Expression Profiles Compared with Sporadic Colon Cancers. J Crohns Colitis. 2019;13. doi:10.1093/ecco-jcc/jjz014

254. Li Y, Zhu Z, Hou X, Sun Y. LncRNA AFAP1-AS1 Promotes the Progression of Colorectal Cancer through miR-195-5p and WISP1. J Oncol. 2021;2021. doi:10.1155/2021/6242798

255. Ramos Meyers G, Samouda H, Bohn T. Short Chain Fatty Acid Metabolism in Relation to Gut Microbiota and Genetic Variability. Nutrients. 2022. doi:10.3390/nu14245361

256. Guo W, Zhang Z, Li L, Liang X, Wu Y, Wang X, et al. Gut microbiota induces DNA methylation via SCFAs predisposing obesity-prone individuals to diabetes. Pharmacol Res. 2022;182. doi:10.1016/j.phrs.2022.106355

257. Herrera-Deguise C, Varela E, Sarrabayrouse G, Pozuelo Del Rió M, Alonso VR, Sainz NB, et al. Gut Microbiota Composition in Long-Remission Ulcerative Colitis is Close to a Healthy Gut Microbiota. Inflamm Bowel Dis. 2023;29. doi:10.1093/ibd/izad058

258. Mendes-Frias A, Moreira M, Vieira MC, Gaifem J, Costa P, Lopes L, et al. Akkermansia muciniphila and Parabacteroides distasonis as prognostic markers for relapse in ulcerative colitis patients. Front Cell Infect Microbiol. 2024;14: 1367998. doi:10.3389/fcimb.2024.1367998

259. Boland BS, He Z, Tsai MS, Olvera JG, Omilusik KD, Duong HG, et al. Heterogeneity and clonal relationships of adaptive immune cells in ulcerative colitis revealed by single-cell analyses. Sci Immunol. 2020;5. doi:10.1126/SCIIMMUNOL.ABB4432

260. Venkateswaran S, Somineni HK, Matthews JD, Kilaru V, Hyams JS, Denson LA, et al. Longitudinal DNA methylation profiling of the rectal mucosa identifies cell-specific signatures of disease status, severity and clinical outcomes in ulcerative colitis cell-

specific DNA methylation signatures of UC. Clin Epigenetics. 2023;15. doi:10.1186/s13148-023-01462-4

261. Mennillo E, Kim YJ, Lee G, Rusu I, Patel RK, Dorman LC, et al. Single-cell and spatial multi-omics highlight effects of anti-integrin therapy across cellular compartments in ulcerative colitis. Nat Commun. 2024;15. doi:10.1038/s41467-024-45665-6

Paper I

Functional & Integrative Genomics (2023) 23:165
https://doi.org/10.1007/s10142-023-01099-9

ORIGINAL ARTICLE

Anti-apoptotic genes and non-coding RNAs are potential outcome predictors for ulcerative colitis

Wei Meng[1] · Kay-Martin Johnsen[2,3] · Christopher G. Fenton[4] · Jon Florholmen[2] · Ruth H. Paulssen[1,4]

Received: 10 November 2022 / Revised: 10 May 2023 / Accepted: 11 May 2023 / Published online: 18 May 2023

Abstract

Due to the lack of clinical, immunologic, genetic, and laboratory markers to predict remission in ulcerative colitis (UC) without relapse, there is no clear recommendation regarding withdrawal of therapy. Therefore, this study was to investigate if transcriptional analysis together with Cox survival analysis might be able to reveal molecular markers that are specific for remission duration and outcome. Mucosal biopsies from patients in remission with active treatment-naïve UC and healthy control subjects underwent whole-transcriptome RNA-seq. Principal component analysis (PCA) and Cox proportional hazards regression analysis were applied to the remission data concerning duration and status of patients. A randomly chosen remission sample set was used for validation of the applied methods and results. The analyses distinguished two different UC remission patient groups with respect to remission duration and outcome (relapse). Both groups showed that altered states of UC with quiescent microscopic disease activity are still present. The patient group with the longest remission duration and no relapse revealed specific and increased expression of antiapoptotic factors belonging to the MTRNR2-like gene family and non-coding RNAs. In summary, the expression of anti-apoptotic factors and non-coding RNAs may contribute to personalized medicine approaches in UC by improving patient stratification for different treatment regimens.

Keywords Ulcerative colitis · Remission · Cox analysis · Biomarkers

✉ Ruth H. Paulssen
ruth.h.paulssen@uit.no

Wei Meng
wei.meng@uit.no

Kay-Martin Johnsen
kay-martin.Johnsen@unn.no

Christopher G. Fenton
christopher.fenton@uit.no

Jon Florholmen
jon.florholmen@uit.no

[1] Clinical Bioinformatics Research Group, Department of Clinical Medicine, Faculty of Health Sciences, UiT- The Arctic University of Norway, Tromsø, Norway

[2] Gastroenterology and Nutrition Research Group, Faculty of Health Sciences, UiT- The Arctic University of Norway, Tromsø, Norway

[3] Department of Medical Gastroenterology, University Hospital of North Norway, Tromsø, Norway

[4] Genomics Support Centre Tromsø, UiT- The Arctic University of Norway, Department of Clinical Medicine, Faculty of Health Sciences, UiT- The Arctic University of Norway, Sykehusveien 44, N-9037 Tromsø, Norway

Introduction

Ulcerative colitis (UC) is a chronic inflammatory disorder which requires long-term treatment in order to achieve remission (Ungaro et al. 2017). The inflammation status of UC patients is usually determined by endoscopic, histologic, and laboratory parameters (Peyrin-Biroulet et al. 2014; Rogler et al. 2013). Different guidelines for medical and surgical treatment of UC are available (Dassopoulos et al. 2015; Magro et al. 2017). In general, a step-up approach is recommended with the goal of obtaining clinical remission (Danese et al. 2014). Biological therapy is recommended for patients with moderate to severe disease refractory or patients dependent on steroid treatment. Side effects of both types of medication are common. The current management programs for UC aim for induction and maintenance of clinical remission to prevent treatment-induced and disease-related complications.

Today, different scoring systems for UC activity are in use to evaluate endoscopic disease activity and activity status, but none of the scoring systems have had all criteria fully determined (Travis et al. 2011; Rutter et al. 2004). There

is no validated current definition of remission, and therefore still no consensus on how to define clinical remission (Magro et al. 2017). The guidelines from the European Colitis and Crohn's organization (ECCO) for remission suggest the absence of visible mucosal lesions (Mayo endoscopic grade 0) in remission (Magro et al. 2013), whereas others allow Mayo ≤ 1 including endoscopic grade 1 in remission (Lamb et al. 2019; Rutgeerts et al. 2005; Schroeder et al. 1987). However, it is generally accepted that healed mucosa with the absence of mucosal lesions is a treatment goal. "Histological" healed mucosa is not included in clinical remission, but there is an increasing focus of including histological criteria in healed mucosa (Peyrin-Biroulet et al. 2014). It is well known that even in the absence of gastrointestinal symptoms as well as normal endoscopic and clinical findings, patients may have persisting microscopic inflammatory activity even in the absence of gastrointestinal symptoms (Korelitz 2010; Magro et al. 2018; DeRoche et al. 2014). This activity can result in progressive accumulation of bowel damage, such as fibrosis, dysmotility, and increased risk of colorectal neoplasm (Gupta et al. 2007).

It is self-evident that there is a need for standardization of both assessment and validation as well as prognostic values. There is still a need to characterize the complex pathogenic and healing mechanisms in UC. Due to the lack of clinical, immunologic, genetic, and laboratory markers to predict remission without relapse, there is no clear recommendation regarding withdrawal of therapy. Therefore, the current study aims to identify molecular signatures in a UC remission cohort obtained by whole-transcriptome RNA-Seq with the intent to provide a better understanding of the molecular mechanisms responsible for remission duration and disease outcome. Altogether, this knowledge might lead to novel personalized therapeutic approaches that will help patients to stay in remission.

Materials and methods

Patient material

A standardized sampling method was used to collect mucosal biopsies ($n = 56$) from patients in remission (RR; $n = 26$). For comparison purposes, normal patient biopsies (NN; $n = 16$) and biopsies from patients with active UC (UC; $n = 14$) were adapted from an earlier study (Fenton et al. 2021). The level of inflammation in UC patients was diagnosed based upon established clinical endoscopic and histological criteria as defined by the European Colitis and Crohn's Organization (ECCO) guidelines (Magro et al. 2017). A total Geboes score was determined for the remission samples (Geboes et al. 2000). TNF mRNA levels in biopsies were estimated by qPCR (Olsen et al. 2007). TNF-α values of <7000 copies/ug RNA were considered non-inflamed tissues. Faecal calprotectin was measured with the Calprest ELISA kit (Eurospital). All patient characteristics are depicted in Table 1. All methods were performed in accordance with the Declaration of Helsinki. The study participants signed informed and written consent forms. Approvals were granted by the Regional Committee of Medical Ethics of Northern Norway, Ref no: 14/2004, 1349/2012 and 29895/2020. The samples were taken from an established biobank approved by the Norwegian Board of Health (952/2006).

RNA isolation

Total RNA was isolated using the Allprep DNA/RNA Mini Kit from Qiagen (catalogue number 80204) and the QIAcube instrument (Qiagen), according to the manufacturer's protocol. Quantity and purity of the RNA were assessed by using the NanoDrop ND-1000

Table 1 Characteristics of patients

Characteristics	Control[§] ($n = 16$)	Remission ($n = 26$)	Treatment-naïve active UC[§] ($n = 14$)
Gender (male/female)	11/5	15/11	9/5
Age (years) mean ± SD	52.5 ± 16.9	48.4 ± 13.4	40.7 ± 13.9
Endo score mean ± SD	0	0	1.79 ±0.43
Geboes score (total) ± SD	n.d.	0.36 ± 1.38	6.35 ± 2.93
TNF-α copies/μg RNA ± SD	3663 ± 1973	5060 ± 3047*	15907 ± 9623
Calprotectin (mg/kg) mean ± SD	n.d.	23.8 ± 35.7[€]	587.5 ± 483.8[¥]
Extension of disease[£]	_	2/7/8/9	2/9/3
Duration of remission (years) ± SD	_	4.38 ± 4.28	_
Medication[#]	_	26/0/7/2	_

*TNF-α copies/μg RNA in 18 patients. [£]Proctitis/rectosigmoid/left-sided colitis/pancolitis. [#]5-ASA/steroids/immunosuppressives/biologics. [€]Average calprotectin levels in 16 patients. [¥]Average calprotectin levels in 11 patients. [§]Data adapted for comparison from Fenton et al. (Fenton et al. 2021)

spectrophotometer (Thermo Fisher Scientific, Wilmington, DE). The Experion Automated Electrophoresis System (Bio-Rad, Hercules, CA) and the RNA StdSens Analysis Kit (Bio-Rad, catalogue # 700–7103) were used to evaluate RNA integrity. The RNA samples were kept at −70 °C until further use. All RNA samples used for analyses showed an RNA integrity number (RIN) value of between 8.0 and 10.0.

Library preparation and next-generation sequencing

Whole transcriptome libraries of UC remission samples were prepared with the TruSeq Stranded Total RNA LT Sample Prep Kit from Illumina (Catalogue number RS-122–2203). The amount of input material was 1 µg of total RNA. The Bioanalyzer 2100 (Agilent Technologies, Santa Clara, CA) and the Agilent DNA 1000 kit (Catalogue number 5067-1504) were used to assess RNA library quality, according to the instruction manual. The libraries were normalized to 10 nM and subsequently paired- end sequenced with the NextSeq 550 instrument (Illumina) according to the manufacturer's instructions. The average number of uniquely mapped reads per sequencing run was 85 million reads per sample.

Data analysis

The entire design and workflow of the study is depicted in Fig. 1.

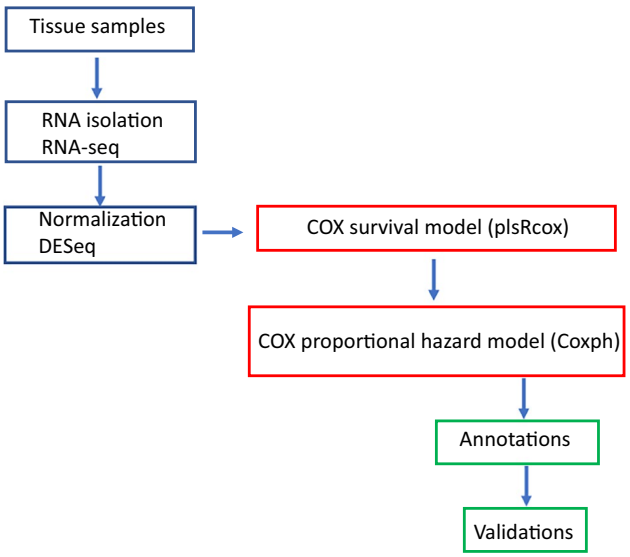


Fig. 1 Study design. The flow chart depicts the entire workflow of the study

Data quality assessment and initial principal component analysis (PCA)

Quality scoring and base calling were performed on the Illumina NextSeq 550 sequencing instrument. The output FastQ file was aligned with reference GENCODE Human Release 33 (Human Genome Assembly GRCh38.p13) (https://www.ncbi.nlm.nih.gov/grc/human/data) by STAR (Version 2.7.3a) with 2-pass mapping and gene counts parameters in STAR (Dobin et al. 2013). After alignment, the read quality was controlled by multiQC (Ewels et al. 2016). The gene counts were analysed and log-normalized by DESeq2 (Love et al. 2014); genes with an average log2 expression less than 4 were filtered out prior to normalization. Seven remission samples were randomly chosen for verification. Initial principal component analysis (PCA) was performed based on the top 15,000 variable genes after normalization.

Processed RNA-Seq data have been deposited in NCBI's Gene Expression Omnibus (GEO, https://www.ncbi.nlm.nih.gov/geo/) and are accessible through GEO series accession numbers GSE128682 and GSE169360.

Cox survival analysis of remission samples

After PCA, the remission patient group was investigated with Cox survival analysis in R using plsRcox (Bastien et al. 2015). Using remission patient information (Table 1) indicating state (relapse or not) and duration (time to relapse), a Cox model was created. The Cox model was applied on the normalized gene count matrix from remission patients. The initial Cox model was significant with a $p < 0.01$ in the likelihood ratio test and $p = 0.03$ in the Wald test for all normalized genes in the remission group, thus, suggesting that there is enough information in the gene matrix to explain patient risk. To further identify which genes influence risk for relapse, the R package survival (coxph) was then applied on each individual gene (Therneau 2021). The second analysis revealed 287 genes that significantly contribute to risk with a p value< 0.01. Those 287 genes were used for PCA analysis and visualization. Hazard beta-coefficients were calculated for the 23 selected genes.

Annotations

Genes were manually annotated using GeneCards (https://www.genecards.org/). EnrichGO of the clusterProfiler R package (Yu et al. 2012) was applied to the protein-coding genes. Only biological process GO terms for comparisons of patient groups with $p_{adj} < 0.05$ were kept.

Validations

Prior to analysis, seven remission samples were excluded from the remission patient cohort for validation. Gene counts from the four MTRNR2-like family genes (MTRNR2L6, MTRNR2L3, MTRNR2L12, and MTRNR2L8) from the validation samples were tested using the plsRcox package (Ginestet 2011). The statistics of all the Cox models are shown in Table S2.

Data visualization

Heatmaps were generated by ComplexHeatmap (Gu et al. 2016). Among 287 genes, protein coding genes and non-protein coding genes were ranked by means of each gene in the remission samples divided by the sum of means of each gene in each group, respectively. The rows were clustered for better visualization.

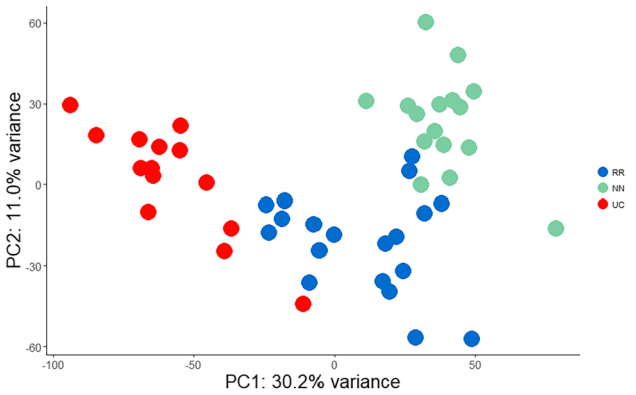


Fig. 2 Principal component analysis (PCA) of remission, ulcerative colitis, and normal control patient samples. PCA of remission (RR), ulcerative colitis (UC), and normal control patient samples (NN) of the 15,000 most variable genes after normalization. Principal component (PC1) explained 30.4% of the total variance, and principal component 2 (PC2) explained 11% of the total variance

Results

Transcriptomic analysis discriminates different states of UC

The whole transcriptome representing treatment-naïve active UC (UC; $n = 14$), UC in remission (RR; $n = 19$), and normal control samples (NN; $n = 16$) was established by RNA-seq. Pre-processing of the sequencing data revealed a total of 18,783 expressed genes. The normalization of the expression of gene matrices for all groups showed no batch effects (Table S1). The initial principal component analysis (PCA) with 15,000 most variable genes resulted in a clear distinction between normal (NN), ulcerative colitis (UC), and remission (RR) samples, along the first principal component (PC1) with a 30.2% explained variance and an 11% explained variance along with the second component (PC2) (Fig. 2). Note, that prior to the initial PCA analysis, seven remission patient samples were randomly removed from the remission data set for validation of a Cox survival model (see below).

Cox survival analysis discriminates genes related to remission duration and state

PCA alone did not result in a separation of remission samples, although they differed in terms of remission duration and time of relapse. Therefore, an attempt was made to distinguish remission samples by using different Cox models. Using the remission patient characteristics depicted in Table 1, a Cox survival model based on partial least squares was established using remission duration as survival time and relapse as an event. The model returned a likelihood test and Wald test $p < 0.05$ (Table S2), thus suggesting that there is enough information in the gene matrix to explain patient risk.

Remission significant genes obtained by Cox analysis

The second Cox analysis for each individual gene of the remission gene matrix returned 287 significant genes $p < 0.01$ related to risk (Tables S2 and S3). Of the 287 genes, 188 represented protein-coding genes, 28 small RNAs, 25 non-coding RNAs, 31 pseudogenes, and 15 miscellaneous RNAs, which are all listed in Table S3. Significant genes ($n = 287$) obtained by the remission Cox analysis were visualized by PCA with co-normalized normal control samples (NN) and UC samples (UC) included. The result of the PCA shows that the significant genes from of the Cox model can clearly separate the remission samples into two groups with 38.1% and 12.9% of explained variances for principal component 1 (PC1) and principal component 2 (PC2) (Fig. 3). Both remission groups showed clear differences with respect to endoscopic, histological, and laboratory parameters and were then denominated accordingly RM (remission without relapse) and RL (remission with relapse) (Fig. 3 and Table S4). The PCA biplot shows both PC scores of samples (dots) and loadings of variables (vectors). The further away these vectors are from a PC origin, the more influence they have on that PC. Notably, the RM samples grouped closer to the normal control samples, whereas RL samples clustered and in part overlapped with UC samples (Fig. 3).

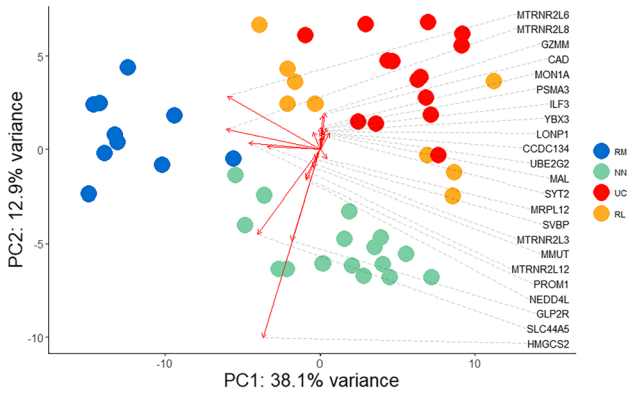


Fig. 3 Principal component analysis of genes revealed by Cox analysis. Genes revealed from the Cox regression analysis ($n = 287$) were used for principal component analysis (PCA) including remission (RR), ulcerative colitis (UC), and normal control patient samples (NN). Principal component 1 (PC1) explained 38.1% of the total variance, and principal component 2 (PC2) explained 12.9% of the total variance. The biplot depicts 23 protein-coding genes of 188 protein-coding genes obtained by Cox analysis. The arrows indicate the genes as loading projectiles that differ the group from the direction. The length of each arrow represents the effect of genes on the components. To improve the visibility, the loadings were multiplied by 25. Each arrow is labelled with a gene name as indicated. An entire list of genes can be found in Table S3. Figure S1 depicts a biplot including all protein-coding genes

Genes of the MTRNR2-like family separate remission duration

Two biplots were constructed on the PCA, one including all 188 protein-coding genes and one including 99 non-coding genes indicating the effect of each individual gene on the principal components (Figs. S1 and S2). For illustrative purposes, twenty-three relevant protein-coding genes with high influence were chosen for construction of a biplot (Fig. 3). In addition, the expressions of relevant 30 protein-coding genes and non-coding genes found for the different patient groups were visualized in a heatmap (Fig. 4).

Genes of the MTRNR2-like family influenced the remission samples separation. Solute carrier family 44 member 5 (SLC44A5), glucagon like peptide 2 receptor (GLP2R), prominin 1 (PROM1), NEDD4 like E3 ubiquitin protein ligase (NEDD4L), and methylmalonyl-CoA mutase (MMUT) are the main participants for components for separation towards the normal control group. Genes like interleukin enhancer binding factor 3 (ILF3), carbamoyl-phosphate synthetase 2, aspartate transcarbamylase, and dihydroorotase (CAD), Mal, T cell differentiation protein (MAL), and granzyme M (GZMM) are influencing the separation of UC samples and RL samples.

Hazard values confirm separation and specificity of expressed genes

To confirm this finding, a PCA using the 287 genes from the remission gene count matrix only was performed (Fig. 5). The result confirmed the separation and specificity of the expressed genes with 59.1% and 9.8% of explained variances for PC1 and PC2 in the remission matrix PCA. Beta-coefficients (hazard values) for selected 23 individual genes are shown in Fig. 6. Ten genes including ILF3, mitochondrial ATP-dependent protease Lon (LONP1), proteasome 20S subunit alpha 3 (PSMA3), and CAD were found to increase the chance of relapse which is reflected by negative coefficients. Thirteen genes including MTRNR2 like family, PROM1, and NEDD4L decrease the probability of relapse which is reflected by positive coefficients. A complete beta value list of all genes ($n = 287$) can be found in Table S5.

Annotation reveals involvement of apoptotic and RNA processing pathways

Among the 287 genes, 188 were protein-coding genes. GO enrichment of these 188 protein-coding genes is shown in Fig. 7. Significantly enriched gene sets revealed biological processes like negative regulation of the execution phase of apoptosis with genes of the MTRNR2-like family (MTRNR2L6, MTRNR2L8, MTRNR2L3, MTRNR2L12), ribosome biogenesis, rRNA processing, RNA splicing, signal transduction by p53 class mediator, and ribonucleoprotein complex biogenesis. The cellular component preribosome and molecular functions including single-stranded RNA binding and receptor antagonist activity were enriched. The full enrichment list is shown in Table S6.

MTRNR2-like genes are predictors for risk of relapse

Seven remission testing samples were used to validate the influence of the four MTRNR2-like genes (MTRNR2L6, MTRNR2L8, MTRNR2L3, and MTRNR2L12) using the Cox model. The correlation between predicted duration and actual duration $R = 0.641$ (Fig. 8). This indicates that the MTRNR2-like genes are good predictors for risk of relapse.

Discussion

Today, the recommendations regarding the withdrawal of therapy during UC are not clear. Therefore, the current study aimed to identify molecular signatures in a UC remission patient cohort with focus on remission duration after treatment and disease outcome. The analysis of transcriptional expression data of UC remission samples obtained by RNA-Seq, Cox survival analysis, and downstream PCA

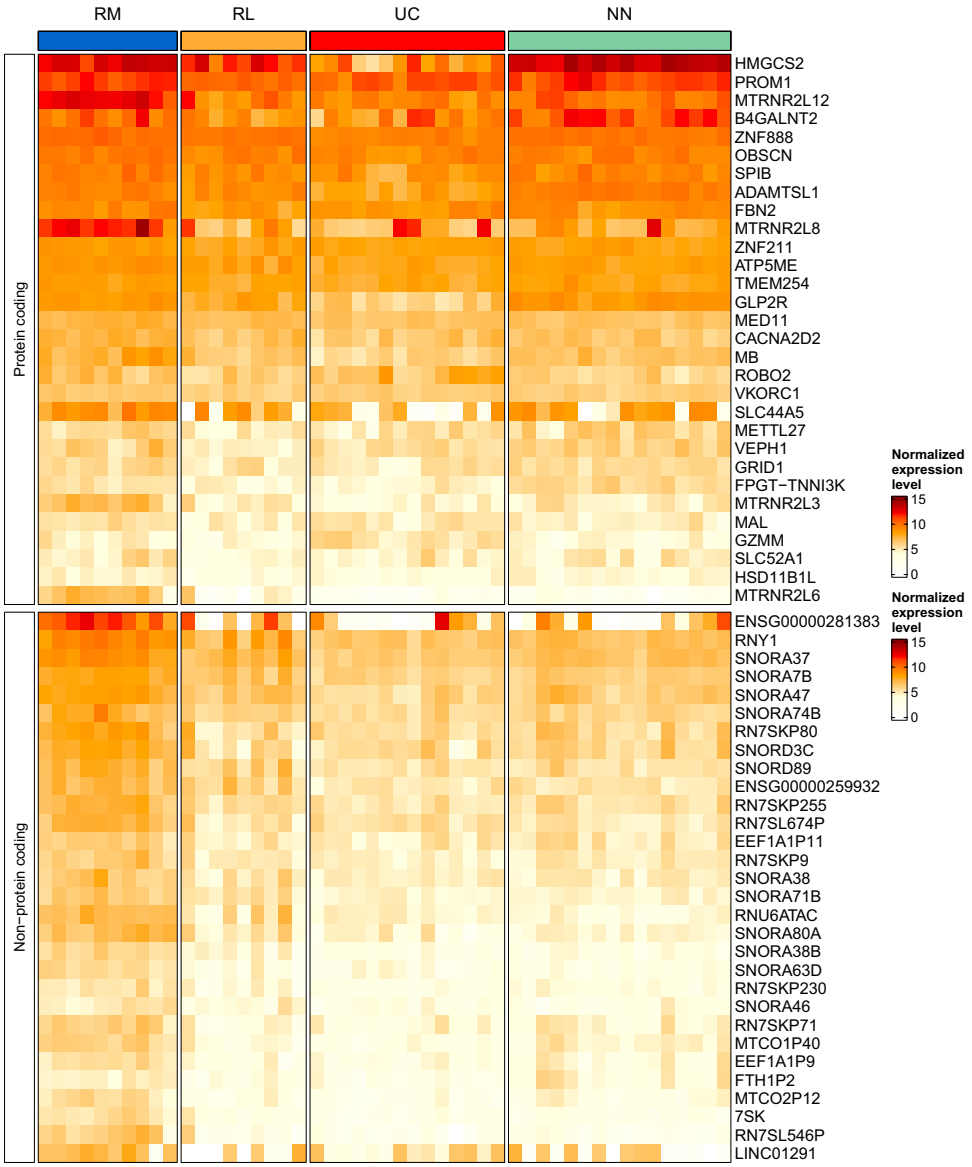


Fig. 4 Heatmap of selected genes of relevance for remission status. Heatmaps were generated by ComplexHeatmap as described in the "Materials and methods" section. Thirty protein coding genes and non-protein coding genes were ranked by means of each gene in the remission samples divided by the sum of means of each gene in each group, respectively. Long-term remission samples (RM), short-term remission samples (RL), treatment-naïve ulcerative colitis samples (UC), and normal control samples (NN) are depicted and normalized expression levels of a genes are indicated

analysis of genes obtained by Cox survival analysis clearly revealed a relationship between remission event (duration of remission) and remission states (relapse or no relapse). Initial PCA analysis (Fig. 2) on the normalized expression matrix of all patient samples confirmed similar distribution patterns for remission samples found for two independent UC remission cohort studies studying the differential expression of genes, showing a clear distinction between UC in remission, normal controls, and active UC samples (Fenton et al. 2021; Planell et al. 2013). PCA analysis of the genes obtained by Cox analysis could clearly separate remission samples into two groups representing UC remission, one with relapse (RL) and one without relapse (RM) (Figs. 3 and 5). The Cox analysis, using the remission gene matrix only, showed that the model was independent of the other sample groups, UC and NN (Fig. 5). Therefore, it is surprising that a clear relationship between selected genes and the UC and NN background samples could be observed (Fig. 3).

The obtained molecular signatures did show different inflammatory states in the remission groups (Fig. 3, Tables S3 and S4). A quiescent inflammation is still present in remission which is reflected by the expression of interleukin enhancer binding factor 3 (ILF3) which is involved in innate immune responses and myeloid dendritic cell maturation in IBD (Aitchison et al. 2021). The influence shown in the biplot (Figs. 3 and S1) on the first principal component and a high beta coefficient found for ILF3 confirms inflammatory signals in RL samples (Fig. 6 and Table S5). Likewise, increased expression of other inflammatory genes like CAD which is involved in the inhibition of NOD2 antibacterial function in intestinal epithelial cells (Richmond et al. 2012) and PSMA3

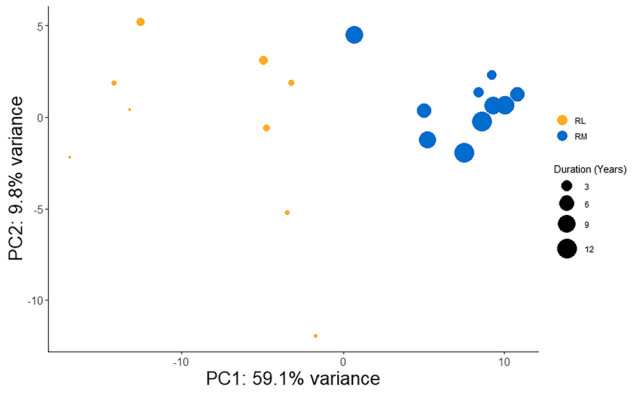


Fig. 5 Separation of UC remission samples by PCA. Separation of remission samples by principal component analysis (PCA) using 287 genes obtained by Cox analysis. The samples separate into two groups dependent on remission duration, remission without relapsing (RM, blue), and remission with relapsing (RL, yellow). The size of the circles represents the duration of remission. Principal component 1 (PC1) explained 59.1% of the total variance, and principal component 2 (PC2) explained 9.8% of the total variance

which is involved in the proteasome-mediated NF-dB activation in UC (Goetzke et al. 2021) was observed. Recently, a relationship between UC and atherosclerosis has been implicated (Weissman et al. 2020; Roifman et al. 2009). The reported higher risk of cardiovascular events in UC patients may be pertinent in inflammation-mediated atherosclerosis (Rungoe et al. 2013; Kristensen et al. 2013). The mitochondrial matrix protein LONP1 has been shown to be involved in atherosclerosis mitochondrial protein quality control (Hansen et al. 2008; Onat et al. 2019) and is a strong risk factor of relapse (Fig. 6). All the above-mentioned genes are shown to have an influence pointing towards inflammation and increased risk of relapse especially for patients in the RL group (Figs. 3, 6, and S1).

It is well-known that mitochondrial function in the intestinal epithelium plays a critical role in maintaining intestinal health (Urbauer et al. 2021). A recent paediatric UC patient cohort study revealed suppressed expression of mitochondrial genes in active UC (Haberman et al. 2019). The here observed increased expression of MTRNR2-like genes might improve the remission state (Fig. 3, Table S3, Fig. S1). Mitochondrial dysfunction and dysbiosis of gut microbiota have been shown to be associated with IBD (Jackson and Theiss 2020). Therefore, a recovery of the gut-microbiota environment and restoring of rectal mitochondrial energy functions can be implied for remission without relapse (RM) where commensal bacterial-induced mitochondrial signalling potentiates epithelial homeostasis. The specific expression of MTRNR2-like genes in RM might represent these genes as potential molecular markers for disease outcome (Figs. 3 and 4, and Table S3). The GO annotations confirmed enrichment of genes involved in the regulation of execution phase of apoptosis (Fig. 7).

It is interesting to note that MTRNR2 treatment may exert beneficial effects in UC by decreasing inflammatory reactions and apoptosis (Gultekin et al. 2017). The mitochondrial metabolism in the intestinal stem cell niche plays also a pivotal role in regulating intestinal epithelial cell homeostasis, including self- renewal and differentiation (Urbauer et al. 2021). The observed expression of stem cell marker prominin 1 (PROM1) (Karim et al. 2014) and NEDD4 like E3 ubiquitin protein ligase (NEDD4L) points to a maintenance of proliferation and differentiation of the colonic epithelium in RM (Kimura et al. 2011) (Figs. 3 and 4). NEDDL4 strongly contributes to a lower risk of relapse (Fig. 6). In addition, increased expression of the vitamin B12 dependent, mitochondrial MMUT (Park et al. 2021) in RM points to a lower B12 deficiency reported for UC patients thereby lowering the risk of relapse (Fig. 6) (Battat et al. 2014; Mortimore and Florin 2010).

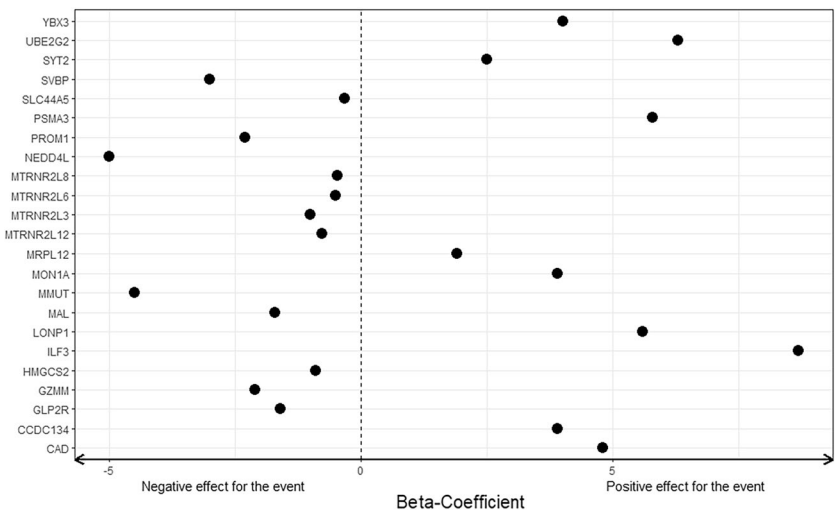


Fig. 6 Beta-coefficients obtained by Cox proportional hazards regression analysis. Beta-coefficients indicate the contribution of each gene to the relative risk of relapse in the Cox survival analysis for 23 UC-relevant genes shown in Fig. 2. The figure shows the beta-coefficient value on the *X*-axis for each gene. Zero is marked as a dashed line. A negative value indicates a protective effect of a gene with which it is associated, and vice versa

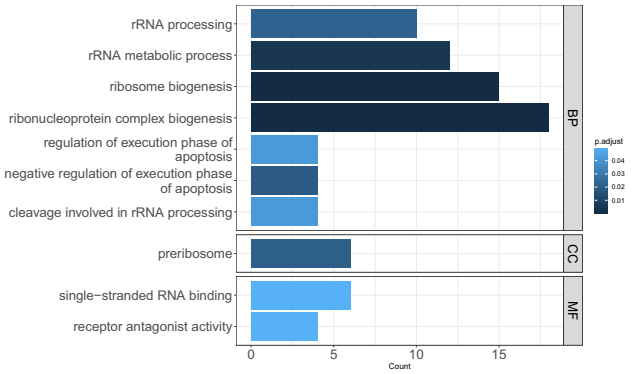


Fig. 7 Functional annotations of genes revealed by Cox analysis. The protein-coding genes were annotated with gene ontology (GO). Enriched pathways and genes involved are indicated. The length of bars indicates the number of genes involved in the GO terms for biological process (BP), cellular component (CC), and molecular function (MF). A complete list of enriched GO pathways annotations can be found in Table S6

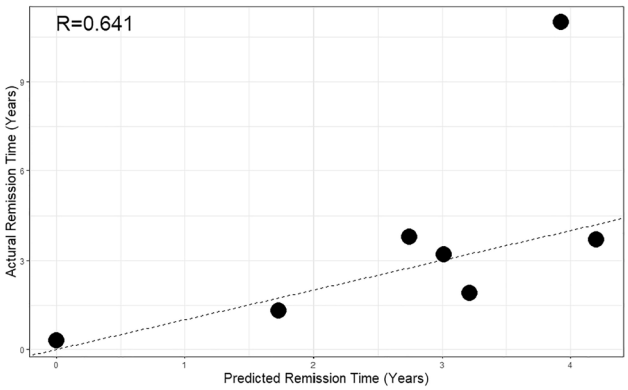


Fig. 8 Validation of the Cox survival model. The validation of the Cox model was tested with seven UC remission samples. The correlation plot depicts the predicted remission time (years) on the *X*-axis and the actual remission time (years) on the *Y*-axis for a group of 7 randomly picked patient samples using a gene set including 4 MTRNR-like genes (MTRNR2L6, MTRNR2L3, MTRNR2L12, MTRNR2L8). The correlation between the two parameters is estimated at $R = 0.641$

Top genes with great influences towards normal control samples in the PCA are HMGSCS2, BAGALNTT2, and GLP2R (Figs. 3 and S1). HMGCS2 encodes a mitochondrial protein that belongs to the HMG-CoA synthase family and catalyses the first reaction of ketogenesis. Elevated expression of HGMCS2 has been reported recently for long-duration ulcerative colitis (Low et al. 2019). Here, HMGCS2 showed increased expression in both remission groups when compared to UC and contributes to a lower risk of relapse (Table S3 and Fig. 6). However, a high expression of HMGCS2 has been associated with the development of colorectal cancer (CRC) which is contrary to these findings (Chen et al. 2017). Increased expression of glycosyltransferase B4GALNT2 in RM points to a maintenance of the intestinal mucus barrier function (Table S3) (Bergstrom et al. 2017). The increased expression of GLP2R involved in the stimulation of intestinal growth, increase of crypt cell proliferation and decrease of enterocyte apoptosis by glucagon-like peptides, prevents intestinal hypoplasia (Drucker 2003).

Interestingly, nearly all the non-coding genes shown in the biplot demonstrate an influence towards RM and normal controls in the PCA (Fig. S2). The expression of 20 small nucleolar RNAs (snoRNAs) (Fig. 4, Table S3, and Fig. S2) may be involved in the mediation of cell–cell communication and improvement of cell survival in the face of stress and/or infection (Rimer et al. 2018), and long non-coding RNAs (lncRNAs) have been shown to have relevance for ulcerative colitis pathogenesis (Ghafouri-Fard et al. 2020; Yarani et al. 2018; Ray et al. 2022). Functions of non-coding RNAs in ribosomal RNA (rRNA) regulation have been recently reported where especially snoRNAs and long non-coding RNAs play important roles in pre-rRNA transcription, processing, and maturation (Li et al. 2013). These pathways are shown to be enriched in RM (Fig. 7).

However, the relevance of specific expression of non-coding RNAs for UC remission duration and outcome needs further evaluation. In this context, it is interesting to note that synergistic gene regulation by pseudogenes and non-coding RNAs has been considered a novel regulatory mechanism which might have a role in UC pathogenesis (Li et al. 2013; Milligan et al. 2015).

This study is not without limitations and is limited by a restricted number of patient samples. Yet, a separation in the PCA after Cox analysis was clearly derived (Figs. 3 and 5). Although several studies present gene expression data of UC patients in remission, separate patient samples with indicated time of relapse were not available for validation of the Cox model. Knowing that the sample number was low, the Cox survival model was then validated with 7 randomly chosen remission patient samples and could confirm the model (Fig. 8) using four MTRNR2-like genes. In addition, a patient cohort with the possibility to investigate the remission state in the same patients consecutively was not available at the time of this study. Nevertheless, the different remission groups do not resemble a normal control phenotype. Patients in the RM group that have been previously treated with anti-TNF therapy (infliximab) until endoscopic remission and subsequently been treated with 5-aminosalicylic acid (5-ASA) only did not experience relapse (Johnsen et al. 2017). Patients in the RL remission group remained in remission for up to 8 months with additional immunosuppressive treatment but had a relapse at some point during the treatment period.

Conclusions

The data clearly demonstrate that remission is an altered state of UC with quiescent microscopic disease activity still present. This disease activity is independent of remission duration and outcome. Transcription expression analysis and Cox survival analysis revealed potential markers genes that could be useful to predict disease outcome. These markers include mitochondrial MTRNR2-like genes and non-coding RNAs. Especially, the expression of anti-apoptotic factors and snoRNAs may contribute to personalized medicine approaches in UC by improving patient stratification for different treatment regimens. The data presented might be of clinical utility in the future.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s10142-023-01099-9.

Acknowledgements The authors thank Ingrid Christiansen for the technical help performing the tumour necrosis factor-alpha level measurements, Lars Daniel Håland at the Genomics Support Centre Tromsø (GSCT; www.uit.no/helsefak/genomics) for sequencing, Sveinung W. Sørbye for histological evaluations, and Renate Meyer for administrating patient samples.

Authors' contributions W.M. performed statistical and functional analyses and wrote parts of the manuscript. K.M.J. was involved in evaluating and providing clinical samples from patients in long-term remission. C.G.F. was involved in data analysis. J.F. was involved in evaluating and providing clinical samples from patients and healthy controls. R.H.P was involved in project inception, project design, data analysis, supervision, and manuscript writing. All authors revised the manuscript and approved the final version of the manuscript.

Funding Open access funding provided by UiT The Arctic University of Norway (incl University Hospital of North Norway) This work was supported by institutional funding.

Data availability The dataset generated and analysed during the current study is available at GEO accession number GSE169360: Go to https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE169360. Enter token ovgloyiyptofdud into the box. All other data generated or analysed during this study are included in this published article and supplementary files.

Declarations

Ethics approval The study was conducted according to the guidelines of the Declaration of Helsinki. Approvals were granted by the Regional Committee of Medical Ethics of Northern Norway, Ref no: 14/2004, 1349/2012 and 29895/2020. The samples were taken from an established biobank approved by the Norwegian Board of Health (952/2006). Written informed consent has been obtained from the study participants to publish this paper.

Conflict of interest The authors declare no competing interests.

References

Aitchison D, Aderem A, Johnson JS, Rodolfo Nazitto AH, Amon LM, Mast FD (2021) Dendritic cell maturation of innate immune responses and myeloid ILF3 is a negative transcriptional regulator. https://doi.org/10.4049/jimmunol.2001235

Bastien P, Bertrand F, Meyer N, Maumy-Bertrand M (2015) Deviance residuals-based sparse PLS and sparse kernel PLS regression for censored data. Bioinformatics 31. https://doi.org/10.1093/bioinformatics/btu660

Battat R, Kopylov U, Szilagyi A, Saxena A, Rosenblatt DS, Warner M, Bessissow T, Seidman E, Bitton A (2014) Vitamin B 12 deficiency in inflammatory bowel disease: prevalence, risk factors, evaluation, and management. https://doi.org/10.1097/MIB.0000000000000024

Bergstrom K, Fu J, Johansson M, Liu X, Gao N, Wu Q, Song J, Mcdaniel JM, Mcgee S, Chen W et al (2017) Core 1- and 3-derived O-glycans collectively maintain the colonic mucus barrier and protect against spontaneous colitis in mice. https://doi.org/10.1038/mi.2016.45

Chen SW, Chou CT, Chang CC, Li YJ, Chen ST, Lin IC, Kok SH, Cheng SJ, Lee JJ, Wu TS et al (2017) HMGCS2 enhances invasion and metastasis via direct interaction with PPARα to activate Src signaling in colorectal cancer and oral cancer. Oncotarget 8:22460. https://doi.org/10.18632/ONCOTARGET.13006

Danese S, Siegel CA, Peyrin-Biroulet L (2014) Review article: integrating budesonide-MMX into treatment algorithms for mild-to-moderate ulcerative colitis. Aliment Pharmacol Ther 39:1095–1103

Dassopoulos T, Cohen RD, Scherl EJ, Schwartz RM, Kosinski L, Regueiro MD (2015) Ulcerative colitis care pathway. Gastroenterology 149:238–245. https://doi.org/10.1053/j.gastro.2015.05.036

DeRoche TC, Xiao SY, Liu X (2014) Histological evaluation in ulcerative colitis. Gastroenterol Rep (Oxf) 2:178–192. https://doi.org/10.1093/gastro/gou031

Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR (2013) STAR: ultrafast universal RNA-Seq aligner. Bioinformatics 29:15–21. https://doi.org/10.1093/bioinformatics/bts635

Drucker DJ (2003) Glucagon-like peptides: regulators of cell proliferation, differentiation, and apoptosis. https://doi.org/10.1210/me.2002-0306

Ewels P, Magnusson M, Lundin S, Kaller M (2016) MultiQC: summarize analysis results for multiple tools and samples in a single report. Bioinformatics 32:3047–3048. https://doi.org/10.1093/bioinformatics/btw354

Fenton CG, Taman H, Florholmen J, Sørbye SW, Paulssen RH (2021) Transcriptional signatures that define ulcerative colitis in remission. Inflamm Bowel Dis 27:94–105. https://doi.org/10.1093/ibd/izaa075

Geboes K, Riddell R, Ost A, Jensfelt B, Persson T, Lofberg R (2000) A reproducible grading scale for histological assessment of inflammation in ulcerative colitis. Gut 47:404–409

Ghafouri-Fard S, Eghtedarian R, Taheri M (2020) The crucial role of non-coding RNAs in the pathophysiology of inflammatory bowel disease. Biomed Pharmacother 129:110507. https://doi.org/10.1016/J.BIOPHA.2020.110507

Ginestet C (2011) Ggplot2: elegant graphics for data analysis. J R Stat Soc A Stat Soc 174. https://doi.org/10.1111/j.1467-985x.2010.00676_9.x

Goetzke CC, Ebstein F, Kallinich T, Feist E, Möller B (2021) Clinical medicine role of proteasomes in inflammation. Inflammation J Clin Med 10. https://doi.org/10.3390/jcm10081783

Gu Z, Eils R, Schlesner M (2016) Complex heatmaps reveal patterns and correlations in multidimensional genomic data. Bioinformatics (Oxford, England) 32:2847–2849. https://doi.org/10.1093/BIOINFORMATICS/BTW313

Gultekin FA, Emre AU, Celik SK, Barut F, Tali U, Sumer D, Turkcu UO (2017) Effects of humanin on experimental colitis induced by 2,4,6-trinitrobenzene sulphonic acid in rats. Saudi J Gastroenterol: Official J Saudi Gastroenterol Assoc 23(2):105. https://doi.org/10.4103/sjg.SJG_318_16

Gupta RB, Harpaz N, Itzkowitz S, Hossain S, Matula S, Kornbluth A, Bodian C, Ullman T (2007) Histologic inflammation is a risk factor for progression to colorectal neoplasia in ulcerative colitis: a cohort study. Gastroenterology 133:1091–1099. https://doi.org/10.1053/j.gastro.2007.08.001

Haberman Y, Karns R, Dexheimer PJ, Schirmer M, Somekh J, Jurickova I, Braun T, Novak E, Bauman L, Collins MH et al (2019) Ulcerative colitis mucosal transcriptomes reveal mitochondriopathy and personalized mechanisms underlying disease severity and treatment response. Nat Commun 10:38. https://doi.org/10.1038/s41467-018-07841-3

Hansen J, Corydon TJ, Palmfeldt J, Dürr A, Fontaine B, Nielsen MN, Christensen JH, Gregersen N, Bross P (2008) Decreased expression of the mitochondrial matrix proteases Lon and ClpP in cells from a patient with hereditary spastic paraplegia (SPG13). Neuroscience 153:474–482. https://doi.org/10.1016/j.neuroscience.2008.01.070

Jackson DN, Theiss AL (2020) Gut bacteria signaling to mitochondria in intestinal inflammation and cancer. Gut Microbes 11:285–304

Johnsen K-M, Goll R, Hansen V, Olsen T, Rismo R, Heitmann R, Gundersen MD, Kvamme JM, Paulssen EJ, Kileng H et al (2017) Repeated intensified infliximab induction - results from an 11-year prospective study of ulcerative colitis using a novel treatment algorithm. Eur J Gastroenterol Hepatol 29:98–104. https://doi.org/10.1097/MEG.0000000000000753

Karim BO, Rhee KJ, Liu G, Yun K, Brant SR (2014) Prom1 function in development, intestinal inflammation, and intestinal tumorigenesis. Frontiers. Oncology 4. https://doi.org/10.3389/fonc.2014.00323

Kimura T, Kawabe H, Jiang C, Zhang W, Xiang YY, Lu C, Salter MW, Brose N, Lu WY, Rotin D (2011) Deletion of the ubiquitin ligase Nedd4L in lung epithelia causes cystic fibrosis-like disease. Proc Natl Acad Sci U S A 108:3216–3221. https://doi.org/10.1073/PNAS.1010334108/SUPPL_FILE/PNAS.1010334108_SI.PDF

Korelitz BI (2010) Mucosal healing as an index of colitis activity: back to histological healing for future indices. Inflamm Bowel Dis 16:1628–1630. https://doi.org/10.1002/ibd.21268

Kristensen SL, Ahlehoff O, Lindhardsen J, Erichsen R, Jensen GV, Torp-Pedersen C, Nielsen OH, Gislason GH, Hansen PR (2013) Disease activity in inflammatory bowel disease is associated with increased risk of myocardial infarction, stroke and cardiovascular death--a Danish nationwide cohort study. PLoS One 8. https://doi.org/10.1371/JOURNAL.PONE.0056944

Lamb CA, Kennedy NA, Raine T, Hendy PA, Smith PJ, Limdi JK, Hayee B, Lomer MCE, Parkes GC, Selinger C et al (2019) British Society of Gastroenterology consensus guidelines on the management of inflammatory bowel disease in adults. Gut 68:s1–s106

Li W, Yang W, Wang XJ (2013) Pseudogenes: pseudo or real functional elements? J Genet Genomics = Yi Chuan Xue Bao 40:171–177. https://doi.org/10.1016/J.JGG.2013.03.003

Love MI, Huber W, Anders S (2014) Moderated estimation of fold change and dispersion for RNA-Seq data with DESeq2. Genome Biol 15. https://doi.org/10.1186/s13059-014-0550-8

Low END, Mokhtar NM, Wong Z, Raja Ali RA (2019) Colonic mucosal transcriptomic changes in patients with long-duration ulcerative colitis revealed colitis-associated cancer pathways. J Crohns Colitis 13:755–763. https://doi.org/10.1093/ecco-jcc/jjz002

Magro F, Gionchetti P, Eliakim R, Ardizzone S, Armuzzi A, Barreiro-de Acosta M, Burisch J, Gecse KB, Hart AL, Hindryckx P et al (2017) 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 11:649–670. https://doi.org/10.1093/ecco-jcc/jjx008

Magro F, Langner C, Driessen A, Ensari A, Geboes K, Mantzaris GJ, Villanacci V, Becheanu G, Borralho Nunes P, Cathomas G et al (2013) European consensus on the histopathology of inflammatory bowel disease. J Crohns Colitis 7:827–851. https://doi.org/10.1016/j.crohns.2013.06.001

Magro F, Lopes J, Borralho P, Lopes S, Coelho R, Cotter J, Castro FD, Sousa HT, Salgado M, Andrade P et al (2018) 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. https://doi.org/10.1136/gutjnl-2017-315545

Milligan MJ, Lipovich L, Rossbach M, Alvarez-Dominguez JR (2015) Pseudogene-derived LncRNAs: emerging regulators of gene expression. https://doi.org/10.3389/fgene.2014.00476

Mortimore M, Florin THJ (2010) A role for B 12 in inflammatory bowel disease patients with suppurative dermatoses? An experience with high dose vitamin B 12 therapy. https://doi.org/10.1016/j.crohns.2010.02.007

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

Onat UI, Yildirim AD, Tufanli Ö, Çimen I, Kocatürk B, Veli Z, Hamid SM, Shimada K, Chen S, Sin J et al (2019) Intercepting the lipid-induced integrated stress response reduces atherosclerosis. J Am Coll Cardiol 73:1149–1169. https://doi.org/10.1016/J.JACC.2018.12.055

Park YE, Park SJ, Park JJ, Cheon JH, Kim T, Il; Kim, W.H. (2021) Incidence and risk factors of micronutrient deficiency in patients with IBD and intestinal Behçet's disease: folate, vitamin B12, 25-OH-vitamin D, and ferritin. BMC Gastroenterol 21:1–9. https://doi.org/10.1186/S12876-021-01609-8/FIGURES/3

Peyrin-Biroulet L, Bressenot A, Kampman W (2014) Histologic remission: the ultimate therapeutic goal in ulcerative colitis? Clin Gastroenterol Hepatol 12:929–34 e2. https://doi.org/10.1016/j.cgh.2013.07.022

Planell N, Lozano JJ, Mora-Buch R, Masamunt MC, Jimeno M, Ordas I, Esteller M, Ricart E, Pique JM, Panes J et al (2013) Transcriptional analysis of the intestinal mucosa of patients with ulcerative colitis in remission reveals lasting epithelial cell alterations. Gut 62:967–976. https://doi.org/10.1136/gutjnl-2012-303333

Ray MK, Fenton CG, Paulssen RH (2022) Novel long non-coding RNAs of relevance for ulcerative colitis pathogenesis. Non-coding

RNA Res 7:40–47. https://doi.org/10.1016/J.NCRNA.2022.02.001

Richmond AL, Kabi A, Homer CR, Marina-García N, Nickerson KP, Nesvizhskii AI, Sreekumar A, Chinnaiyan AM, Nuñez G, McDonald C (2012) The nucleotide synthesis enzyme CAD inhibits NOD2 antibacterial function in human intestinal epithelial cells. Gastroenterology p 142. https://doi.org/10.1053/J.GASTRO.2012.02.040

Rimer JM, Lee J, Holley CL, Crowder RJ, Chen DL, Hanson PI, Ory DS, Schaffer JE (2018) Long-range function of secreted small nucleolar RNAs that direct 2-O-methylation. J Biol Chem 293:13284–13296. https://doi.org/10.1074/jbc.RA118.003410

Rogler G, Vavricka S, Schoepfer A, Lakatos PL (2013) Mucosal healing and deep remission: what does it mean? World J Gastroenterol 19:7552–7560. https://doi.org/10.3748/wjg.v19.i43.7552

Roifman I, Sun YC, Fedwick JP, Panaccione R, Buret AG, Liu H, Rostom A, Anderson TJ, Beck PL (2009) Evidence of endothelial dysfunction in patients with inflammatory bowel disease. Clin Gastroenterol Hepatol 7:175–182. https://doi.org/10.1016/J.CGH.2008.10.021

Rungoe C, Basit S, Ranthe MF, Wohlfahrt J, Langholz E, Jess T (2013) Risk of ischaemic heart disease in patients with inflammatory bowel disease: a nationwide Danish cohort study. Gut 62:689–694. https://doi.org/10.1136/GUTJNL-2012-303285

Rutgeerts P, Sandborn WJ, Feagan BG, Reinisch W, Olson A, Johanns J, Travers S, Rachmilewitz D, Hanauer SB, Lichtenstein GR et al (2005) Infliximab for induction and maintenance therapy for ulcerative colitis. N Engl J Med 353:2462–2476. https://doi.org/10.1056/NEJMoa050516

Rutter M, Bernstein C, Matsumoto T, Kiesslich R, Neurath M (2004) Endoscopic appearance of dysplasia in ulcerative colitis and the role of staining. Endoscopy 36:1109–1114. https://doi.org/10.1055/s-2004-826049

Schroeder KW, Tremaine WJ, Ilstrup DM (1987) Coated oral 5-aminosalicylic acid therapy for mildly to moderately active ulcerative colitis. A randomized study. N Engl J Med 317:1625–1629. https://doi.org/10.1056/NEJM198712243172603

Therneau TM (2021) Survival: a package for survival analysis in R. R Package Version 2:38

Travis SP, Higgins PD, Orchard T, Van Der Woude CJ, Panaccione R, Bitton A, O'Morain C, Panes J, Sturm A, Reinisch W et al (2011) Review article: defining remission in ulcerative colitis. Aliment Pharmacol Ther 34:113–124. https://doi.org/10.1111/j.1365-2036.2011.04701.x

Ungaro R, Mehandru S, Allen PB, Peyrin-Biroulet L, Colombel JF (2017) Ulcerative colitis. Lancet 389:1756–1770

Urbauer E, Rath E, Haller D (2021) Mitochondrial metabolism in the intestinal stem cell niche—sensing and signaling in health and disease. Front Cell Dev Biol 8. https://doi.org/10.3389/fcell.2020.602814

Weissman S, Sinh P, Mehta TI, Thaker RK, Derman A, Heiberger C, Qureshi N, Amrutiya V, Atoot A, Dave M et al (2020) Atherosclerotic cardiovascular disease in inflammatory bowel disease: the role of chronic inflammation. World J Gastrointest Pathophysiol 11:104–113. https://doi.org/10.4291/WJGP.V11.I5.104

Yarani R, Mirza AH, Kaur S, Pociot F (2018) The emerging role of Lncrnas in inflammatory bowel disease. Exp Mol Med 50. https://doi.org/10.1038/S12276-018-0188-9

Yu G, Wang LG, Han Y, He QY (2012) ClusterProfiler: an R package for comparing biological themes among gene clusters. OMICS J Integr Biol 16:284–287. https://doi.org/10.1089/omi.2011.0118

Paper II

scientific reports

Check for updates

OPEN

DNA methylation fine-tunes pro-and anti-inflammatory signalling pathways in inactive ulcerative colitis tissue biopsies

Wei Meng[1], Christopher G. Fenton[1,4], Kay-Martin Johnsen[2,3], Hagar Taman[1,4], Jon Florholmen[2,3] & Ruth H. Paulssen[1,4]✉

DNA methylation has been implied to play a role in the immune dysfunction associated with inflammatory bowel disease (IBD) and the disease development of ulcerative colitis (UC). Changes of the DNA methylation and correlated gene expression in patient samples with inactive UC might reveal possible regulatory features important for further treatment options for UC. Targeted bisulfite sequencing and whole transcriptome sequencing were performed on mucosal biopsies from patients with active UC (UC, n = 14), inactive UC (RM, n = 20), and non-IBD patients which served as controls (NN, n = 11). The differentially methylated regions (DMRs) were identified by DMRseq. Correlation analysis was performed between DMRs and their nearest differentially expressed genes (DEGs). Principal component analysis (PCA) was performed based on correlated DMR regulated genes. DMR regulated genes then were functional annotated. Cell-type deconvolutions were performed based on methylation levels. The comparisons revealed a total of 38 methylation-regulated genes in inactive UC that are potentially regulated by DMRs (correlation p value < 0.1). Several methylation-regulated genes could be identified in inactive UC participating in IL-10 and cytokine signalling pathways such as IL1B and STAT3. DNA methylation events in inactive UC seem to be fine-tuned by the balancing pro- and anti- inflammatory pathways to maintain a prevailed healing process to restore dynamic epithelium homeostasis.

Keywords Ulcerative colitis, Remission, DNA methylation, Epigenetics

Ulcerative colitis (UC) is a chronic inflammatory disorder of the colon with a relapsing course[1]. To achieve remission, long-term treatment is often required[2]. Multiple factors can cause the disease, and research has focused on investigating genetic susceptibility, microbiome communities, environmental factors, and immune responses in UC patients[3–5] Current research denotes the importance of interactions between inflammation and environmental factors[6–8]. Therefore, dynamic processes such as DNA methylation, have been implied to play a role in the immune dysfunction associated with inflammatory bowel disease (IBD)[9–11] and disease development of UC[12–15]. DNA methylation in active UC has been reported recently[9,10,12,13]. Variations in DNA methylation patterns have been associated with homeostasis and defence, immune responses, and progression and development of colorectal cancer (CRC)[12,13].

However, DNA methylation during UC in remission has not been explored in detail. Long-term treatment of UC patients is often necessary to achieve induction and maintenance of clinical remission[16–18]. Previous work on transcriptional signatures in UC has revealed that UC in remission is a permanently altered state of UC with a still ongoing quiescent inflammation[19,20]. Therefore, induced epigenetic changes due to long-term treatment can be expected. In this study, the genome-wide DNA methylome in a UC remission patient cohort was investigated to see if methylation contributes to the expression regulation of specific molecular signatures recently found in a patient cohort with different remission duration[21]. It is believed that changes in the DNA methylation status

[1]Clinical Bioinformatics Research Group, Department of Clinical Medicine, Faculty of Health Sciences, UiT- The Arctic University of Norway, Tromsø, Norway. [2]Gastroenterology and Nutrition Research Group, Faculty of Health Sciences, UiT- The Arctic University of Norway, Tromsø, Norway. [3]Department of Medical Gastroenterology, University Hospital of North Norway, Tromsø, Norway. [4]Genomics Support Centre Tromsø, Department of Clinical Medicine, Faculty of Health Sciences, UiT- The Arctic University of Norway, Sykehusveien 44, 9037 Tromsø, Norway. ✉email: ruth.h.paulssen@uit.no

in remission patients, when correlated to differentially expressed genes (DEGs), may reveal possible regulatory features important for further treatment options for UC patients.

Materials and methods

Patient material

Mucosal biopsies (n = 43) were collected with a standardised sampling method of 12 newly diagnosed, treatment-naïve UC patients with mild to moderate disease activity, 20 patients with inactive UC (remission) and 11 control subjects from a former study[21]. Controls were derived from subjects performing a cancer screening, with normal colonoscopy and normal colonic histological examination. UC was diagnosed based upon established clinical endoscopic and histological criteria as defined by the ECCO guidelines[22]. 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[23]. TNF-α mRNA expression levels were measured by real-time PCR to determine disease activity. All UC patients were initially treated with 5-aminosalicylic acid (5-ASA), in some cases also supplemented with immunosuppressive drugs (Imurel, methotrexate (MTX), and Infliximab) until disease remission. Patients with achieved clinical and endoscopic remission and normalised TNF-α levels were included in this study, with a defined UCDAI score ≤ 2, an endoscopic sub-score of 0 or 1, and TNFα-levels < 7500 copies/µg protein. All patient characteristics are depicted in Table 1. The samples were taken from an established Biobank approved by the Norwegian Board of Health (952/2006). The study participants signed informed and written consent forms. Approvals were granted by the Regional Committee of Medical Ethics of Northern Norway, Ref no: 14/2004, 1349/2012 and 29895/2020.

DNA and RNA isolation

Genomic DNA and total RNA was isolated using the Allprep DNA/ RNA Mini Kit from Qiagen (catalogue number 80204) and the QIAcube instrument (Qiagen), according to the manufacturer's protocol. The quantity and purity of both DNA and RNA were assessed by using the NanoDrop ND-1000 spectrophotometer (Thermo Fisher Scientific, Wilmington, DE). The Experion Automated Electrophoresis System (Bio-Rad, Hercules, CA.) and the RNA StdSens Analysis Kit (Bio-Rad, catalogue number 700–7103) were used to evaluate RNA integrity. All RNA samples used for analyses showed an RNA integrity number (RIN) value between 8.0 and 10.0. DNA and RNA samples were kept at – 70°C until further use.

Institutional review board statement

The study was conducted according to the guidelines of the Declaration of Helsinki. Approvals were granted by the Regional Committee of Medical Ethics of Northern Norway, Ref no: 14/2004, 1349/2012 and 29,895/2020.

Informed consent statement

Written informed consent has been obtained from the study participants to publish this paper.

Library preparation and next-generation sequencing

The libraries were prepared using the SeqCap Epi CpGiant Enrichment Kit (Roche, Switzerland) which enables the targeting of selected genomic regions from bisulfite-treated genomic DNA to identify specific regions in the genome for methylation variation assessment and as previously described[13]. DNA was bisulfite converted using the EZ DNA Methylation-lightning Kit (Zymo Research, USA, cat no: D5030) prior to the hybridisation step and according to the manufacturer's instructions. The amount of input material was 1060 ng of genomic DNA per sample. DNA library quality was 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 4nM before sequencing.

Characteristics	Controls (n = 11)	UC remission (n = 20)	Active UC§# (n = 14)
Gender (male/female)	8/3	10/10	9/5
Age (years) mean ± SD	52.2 ± 19.3	50.0 ± 13.5	40.7 ± 13.9
Clinical score ± SD	0	0	7.78 ± 1.52
Endo Score mean ± SD	0	0.25 ± 0.50¶	1.79 ± 0.43
Geboes score (total) ± SD	n.d	0.13 ± 0.70¤	6.35 ± 2.93
TNF-α copies/µg RNA ± SD	4366 ± 1998	4500 ± 1509*	15,907 ± 9623
Calprotectin (mg/kg) mean ± SD	n.d	17.0 ± 53.78€	587.5 ± 483.8¥
Extension of disease£	_	2/12/6	2/9/3
Medication#	_	20/0/6	_

Table 1. Patient characteristics. *TNF-α copies/µg RNA in 16 patients. £ proctitis/left-sided colitis/pancolitis. #5-ASA/steroids/immunosuppressives. ¤Average score of 13 patients. €Average calprotectin levels in 15 patients. ¥Average calprotectin levels in 11 patients. ¶Average score of 16 patients. §Data adapted for comparison from Fenton et al.[20]. ± SD, Standard Deviation.

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 normalised to 10 nM and diluted to 4 nM prior to sequencing. 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

A flow chart illustrating the downstream analysis process is shown in Fig. 1.

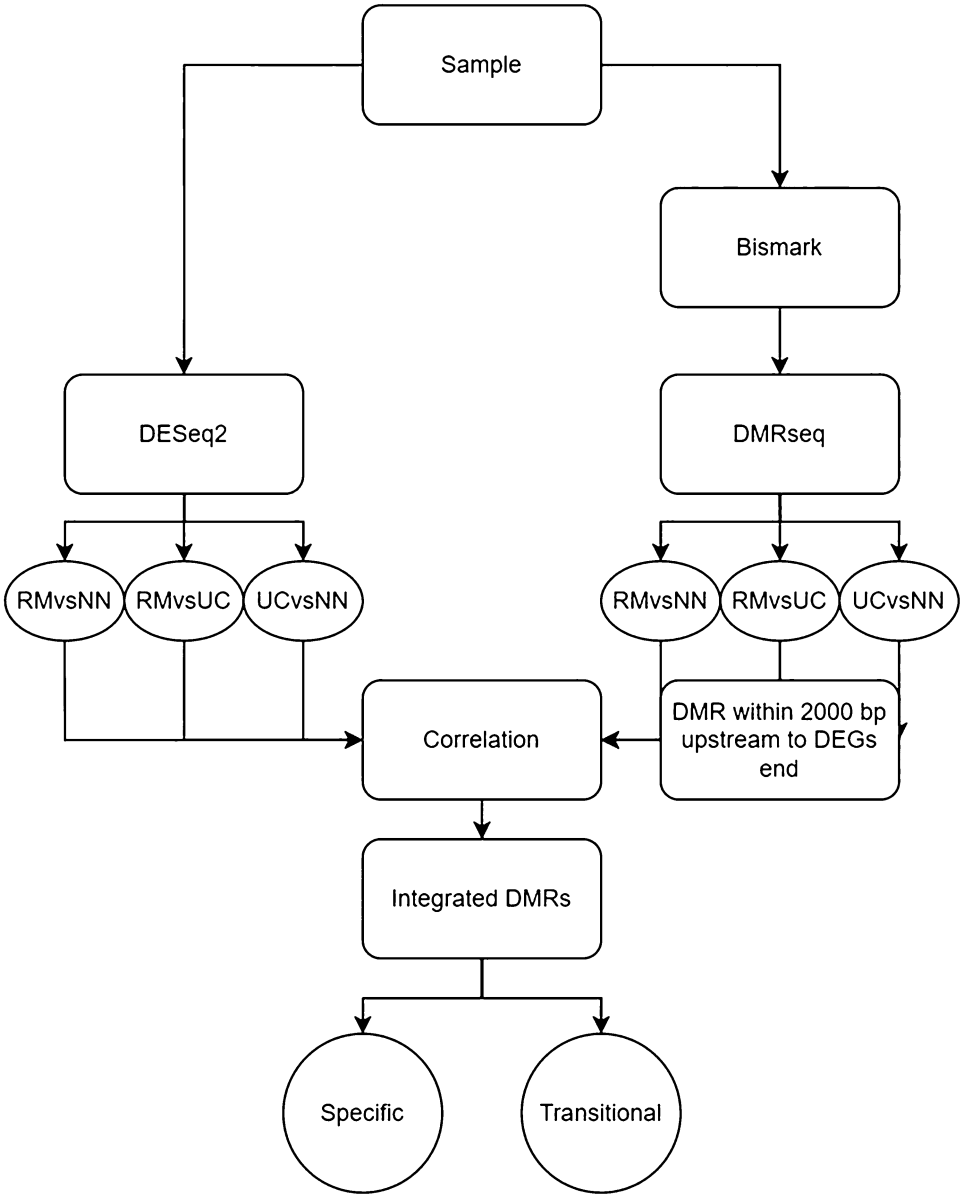


Figure 1. Flow chart for generating differentially methylated region (DMR) correlated gene patterns for groups of inactive UC (RM), active UC (UC), and controls (NN). Processed sequencing data of 45 samples underwent Bismark and DMRseq, incorporating three comparisons to identify DMRs (q value < 0.05) and underwent DEseq2 for DEGs (p.adj < 0.05). DMRs located within 2000 bp upstream of DEGs were correlated with the DEGs. The correlated DMRs were then integrated. Integrated DMRs were grouped based on expression level and t-test on methylation level ($p < 0.05$) to ensure a fit with the pattern conditions.

RNA sequencing

Quality scoring, base calling, and adapter removal were performed on the Illumina NextSeq 550 sequencing instrument. The output FastQ file was aligned with reference GENCODE Human Release 33 (Human Genome Assembly GRCh38.p13) ((https://www.ncbi.nlm.nih.gov/grc/human/data) by Kallisto[24]. DEseq2 was used to generate a list of differential expressed gene transcripts (DEGs)[25]. DEGs with only p.adj < 0.05 were kept.

Bisulfite converted DNA and DMRSeq

Quality scoring, base calling, and adapter removal were performed on the Illumina NextSeq 550 sequencing instrument. The output FastQ file was aligned with reference GENCODE Human Release 38 by Bismark with Bowtie2[26,27]. The output BAM files were then used to generate coverage data and methylation data by using Bismark methylation extractor[27]. The results of Bismark were further processed with DMRseq[28] to find differentially methylated regions. Only differentially methylated regions (DMRs) with a q-value < 0.05 were kept. In this context, the q-value served as an adjustment for multiple comparison to control the false discovery rate.

Identification of possible specific DMRs of inactive UC (RM)

DMRseq was performed on three individual comparisons: inactive UC (RM) vs active UC (UC), inactive UC (RM) vs controls (NN), and active UC (UC) vs controls (NN) resulting in three sets of DMRs with q < 0.05. The resulting DMRs were merged by overlapping genomic locations, and non-overlapping DMRs were discarded. To identify inactive UC specific DMRs additional t-tests were performed between the average DMR methylation levels of inactive UC samples against both control and active UC samples. DMRs whose p-value was less than 0.05 in both t-test comparisons were considered as inactive UC specific. The additional t-test ensured that the inactive UC group was in fact different from both the active UC and control group.

Identifying DEGs correlated to specific DMRs of inactive UC (RM)

For each DMR located within the 2000 bp upstream region of the transcription start site (TSS), sample DMR and DEG were correlated using average DMR methylation and DEG expression levels by Kendall correlation[29]. UCSC known gene (GRCh38) with R package TxDb.Hsapiens.UCSC.hg38.knownGene were used for TSS sites and functional region reference[30]. Correlated genes with $p < 0.1$ were kept based on the correlation coefficient $\tau < 0$.

Annotations and pathway enrichment

Genes associated with transitional methylation patterns, where inactive UC (RM) methylation levels are between active UC (UC) and controls (NN) went through enrichment analysis for pathways enrichment using the Panther/Reactome overrepresentation analysis (Reactome version 77 released 2021–10-01) using the Fisher's exact test[31]. For GO annotations of all 38 genes, clusterProfiler was used[32].

Cell deconvolutions

Cell deconvolutions were performed on DMRs that overlapped Illumina Epic array coordinates using the EpiDISH R package[33]. Average sample relative methylation values for these DMRs were used as input to EpiDISH in Robust Partial Correlation (RPC) mode. Differences between groups were calculated using ANOVA and Tukey's range test[34]. Cell deconvolutions for the expression data were performed with CIBERSORTx (https://cibersortx.stanford.edu/). The LM22 (22 immune cell types) was selected as a signature matrix. The normalized expression matrix was chosen as the input matrix file. The remaining parameters were left at default values[35].

Results

Characterisation of DNA methylation in inactive UC

By combining genome-wide methylation data and whole transcriptome data, insight into the molecular mechanisms of inactive UC was established. Bisulfite sequencing provided DNA methylation levels in patient biopsy samples from inactive UC (n = 20), treatment-naïve active UC (n = 12) and non-IBD controls (n = 11). (Table 1). The different methylation patterns between inactive UC, active UC and controls were identified by DMRseq[28]. DMRs were detected in the following comparisons: 313 DMRs were detected in inactive UC vs. controls, 5,316 DMRs were detected in inactive UC vs. active UC, and 8,262 DMRs were detected in active UC vs. controls. By considering all three comparisons, the methylation levels of a total of 3098 DMRs were negatively correlated with neighbouring transcript expression. Analysis of the combined DMRs revealed 52 DMRs (38 genes) with specific and transitional patterns as depicted in Supplementary Data 1, Fig. 2. Principal component analysis (PCA) with the correlated DMRs could discriminate samples of inactive UC, active UC, and controls at the transcriptomic level (Fig. 3), with 51.0% and 13.3% variance. An example for detailed comparisons between expression and methylation levels are visualised for annexin A11 (ANXA11) (Fig. 4). Here, the presence of DMRs provides valuable information regarding the potential regulation positions and overlapping on the promoter region (shown as "Prom" in Supplementary data 2). ANXA11 is specifically hypo-methylated in DMRs on chr10.470 compared to controls and active UC, thus the expression of ANXA11 in inactive UC is uniquely over-expressed compared to controls and UC (upper left panel, Fig. 4). It is hereby noted that the DMRs labels are specific to this study and do not represent universal IDs.

Methylation-regulated gene profiling in inactive UC.

Among the 38 genes, two methylation patterns were found, one which is specific for inactive UC compared to controls and active UC, and one where methylation patterns of inactive UC are in transitional state between controls and active UC (Fig. 2). Two specific profiles were identified where inactive UC can be further subdivided

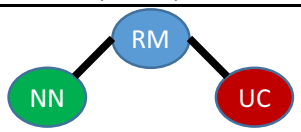
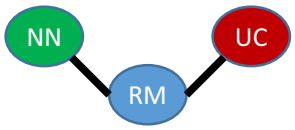
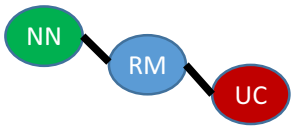
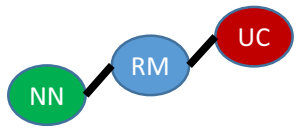
Methylation profiles of 38 genes found in inactive UC (RM)			
Type	Methylation patterns	Profiles	Gene symbol
Specific		NN<RM>UC	CERNA2
			CIT
			ITGA4
			ZNF626
		NN>RM<UC	AFAP1-AS1
			ANXA11
			HYAL1
			MST1R
			S100P
Transitional		NN>RM>UC	ADAM8
			CCDC88B
			CD80
			DENND2D
			DMBT1
			DUSP10
			FAM167A
			GALNT2
			HIF1A
			HIVEP2
			IL1B
			ISG20
			LIMK1
			PARP9
			PEA15
			PFKFB3
			PLA2G2D
			PPP1R18
			PSMB8
			RNF166
			RUBCN
			SBNO2
			SGMS1
			STAT3
			TG
			TNFRSF10C
			TRIM22
			ZC3H12A
		NN<RM<UC	PRR26

Figure 2. Methylation profiles of 38 DEGs found for inactive UC (RM). DEGs are assigned to specific or transitional profiles. All genes have at least one DMR which correlated to the differential expression (p.adj < 0.05) with a negative correlation p < 0.1. Patient groups representing inactive UC (RM), active UC (UC) and controls (NN) are indicated. Methylation patterns, methylation profiles, and gene symbols are indicated. A complete overview is listed in Supplementary Data 1.

into hyper- and hypomethylated compared to UC and controls. Two different transitional profiles were identified where inactive UC shows an intermediate state of methylation in relation to controls and active UC (Supplementary Data 1).

Genes with specific methylation patterns that are hyper-methylated in inactive UC include competing endogenous lncRNA 1 for mir-4707-5p and mir-4767 (CERNA2), citron rho-interacting serine/threonine kinase (CIT), integrin subunit alpha 4 (ITGA4), and zinc finger protein 626 (ZNF626). Genes uniquely hypo-methylated in inactive UC compared to active UC includes actin filament associated protein 1 antisense RNA 1

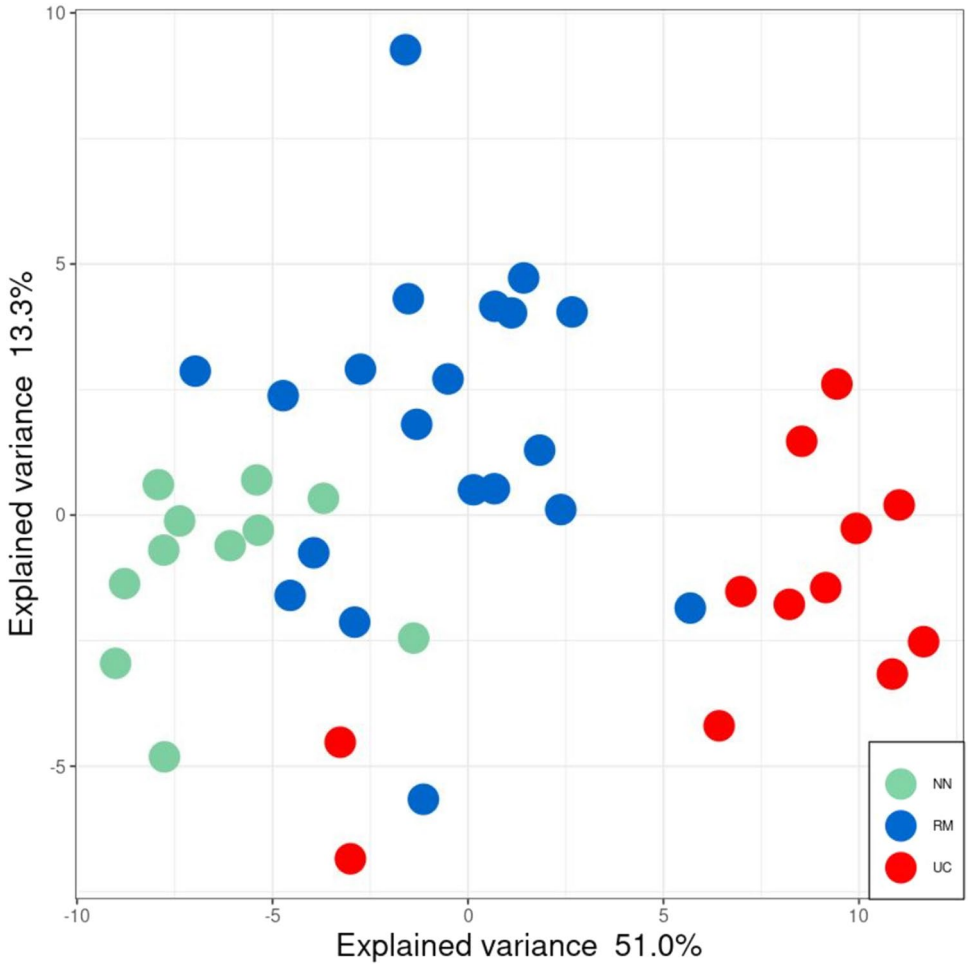


Figure 3. Principal component analysis (PCA). Expression levels of 38 methylation-regulated transcripts with specific and transitional patterns. Patient and control samples are indicated as followed: inactive UC (RM; blue), treatment-naïve active UC (UC; red), and controls (NN; green). Differential expression showed a 51.0% explained variance in PC1 and a 13.3% explained variance in PC2.

(AFAP1-AS1), annexin A11 (ANXA11), hyaluronidase 1 (HYAL1), macrophage stimulating 1 receptor (MST1R), and S100 calcium binding protein P (S100P).

The transitional genes patterns represent genes that have an intermediate methylation status between UC and controls (Fig. 2). 28 genes were found to be hyper-methylated in controls compared to active UC, including ADAM metallopeptidase domain 8 (ADAM8), coiled-coil domain containing 88B (CCDC88B), CD80, DENN domain containing 2D (DENND2D), dual specificity phosphatase 10 (DUSP10), signal transducer and activator of transcription 3 (STAT3), and interleukin 1B (IL1B), interferon stimulated exonuclease gene 20 (ISG20) and LIM domain kinase 1 (LIMK1). DIP2C antisense RNA 1 (PRR26) is the only gene found to be hyper-methylated in active UC compared to controls. A comprehensive list of figures of all methylation-regulated genes can be found in Supplementary Data 1.

Methylation-regulated genes are related to inflammation

The 28 methylation regulated genes in the intermediate state were functionally annotated with Gene Ontology (GO) terms with the Panther/Reactome overrepresentation test (Reactome v.77, released 2021-10-01) (Fig. 5). Genes like TRIM22, PSMB8, CD80, IL1B, ISG20, HIF1A, STAT3 were all hyper-methylated and downregulated to a lesser extent in inactive UC compared to UC. These genes were annotated to the IL-10 pathway and cytokine signalling in immune system (Fig. 5).

Functional enrichment revealed 101 immunological and inflammation related signalling pathways ($p.adj < 0.05$), which include interleukin-6 production, CD4-positive, alpha–beta T cell activation and lymphocyte differentiation represented by genes like ADAM8, IL1B, ISG20, CD80, STAT3 and ZC3H12A (Fig. 5; Supplementary Data 3). Genes in specific patterns including CIT, MST1R, HYAL1, ITGA4 were annotated as hyaluronan metabolic process, epithelium migration and phagocytosis. The total of 103 functional GO terms are listed in Supplementary Data 3.

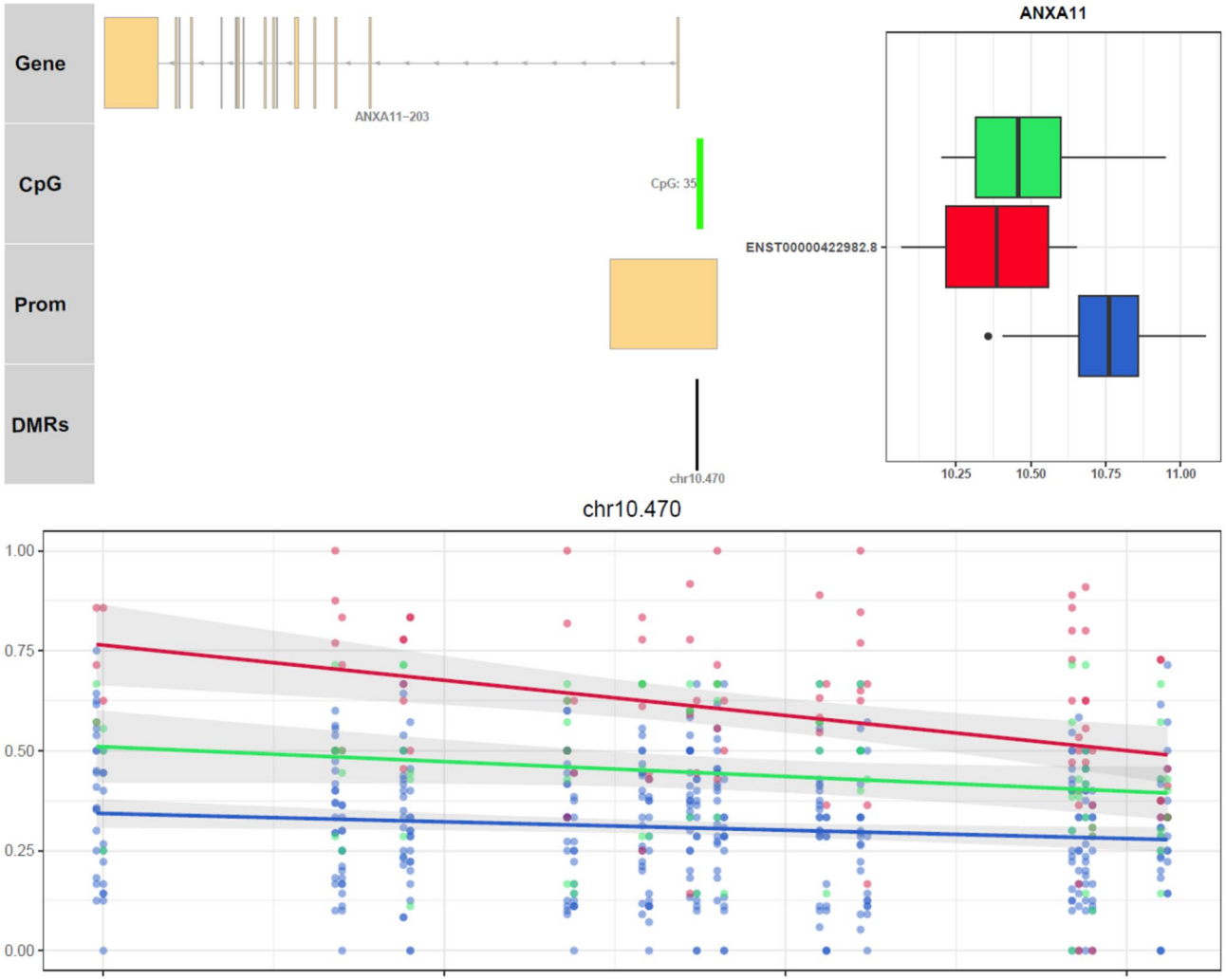


Figure 4. Correlation between DMRs status and transcriptional levels of ANXA11. The transcript and regulatory position of the transcript is aligned with identified DMRs on the top left, showing the regional transcript information. The data type is listed on the left: Prom stands for the promoter region in the transcript. DMRs are the differential methylated region from DMRseq, named with a chromosome and a tag. The differential expression level is shown in the top right as a box plot. The X-axis shows the log2 normalised expression levels. The differential methylation level of each region is shown at the bottom. The X-axis is the position of methylated sites in the region, and Y-axis is the methylation percentage (from 0 to 1). Each dot is one percentage of methylation position in each sample. The linear regression is shown as a line with inactive UC (blue), UC (red) and controls (green). The grey area of the line stands for a 95% confidence level. Raw difference of methylation level, differential expression level, and correlation value can be found in Supplementary Data 1. Figures of all 38 genes can be found in Supplementary Data 2, different transcripts of one gene are indicated if multiple transcripts are involved.

Cell deconvolutions based on DNA methylations discriminate cell fractions

Cell deconvolution was performed by mapping DMR genomic coordinates to Illumina EPIC array identifiers[33]. The deconvolution results revealed differences in cell fractions for inactive UC (RM), active UC (UC) and controls (NN). Epithelial cell fractions were higher in normal and inactive UC than active UC (padj < 0.01). The proportion of immune cells is significantly higher in active UC compared to inactive UC and controls (padj < 0.01). Similarly, the fibroblast cell fractions in inactive UC and controls are slightly higher than in active UC (Fig. 6, Supplementary Data 4). Using CIBERSORTx on the normalized gene expression matrix showed an increase in most immuno-derived cells, especially of neutrophils in active UC as compared to inactive UC and controls (Supplementary Data 5)[35].

Discussion

In a previous work, DNA methylation patterns have been identified for active, treatment-naïve UC[12,13,36]. In this study, the methylation status in inactive UC (RM) has been explored by determining gene expression regulated by global DNA methylation overlapping the promoter region of genes. By correlating DNA methylation data to

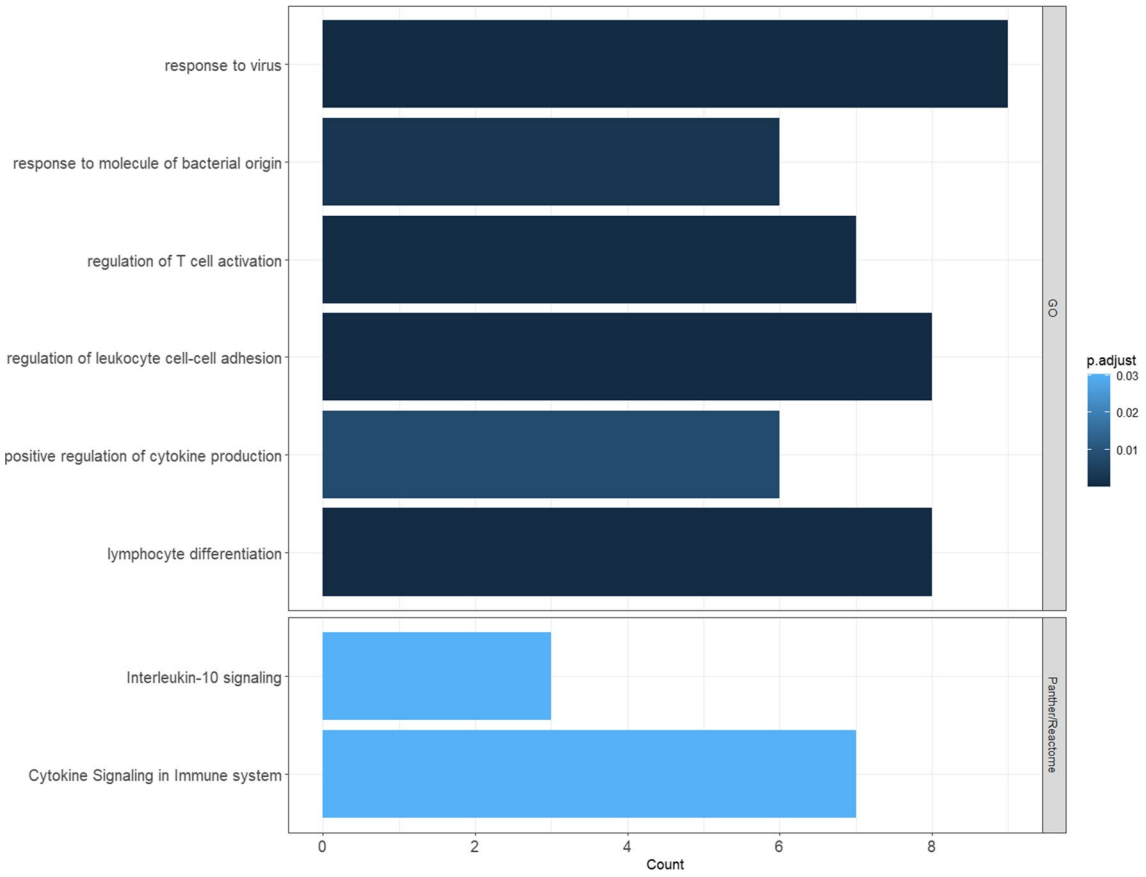


Figure 5. Gene ontology enrichment of the methylation correlated genes. The bar length shows the number of genes enriched in each term (x-axis). The colour stands for the padj value from dark blue to light blue (0 to 0.05, respectively). The detailed genes included in each term are listed in Supplementary Data 3.

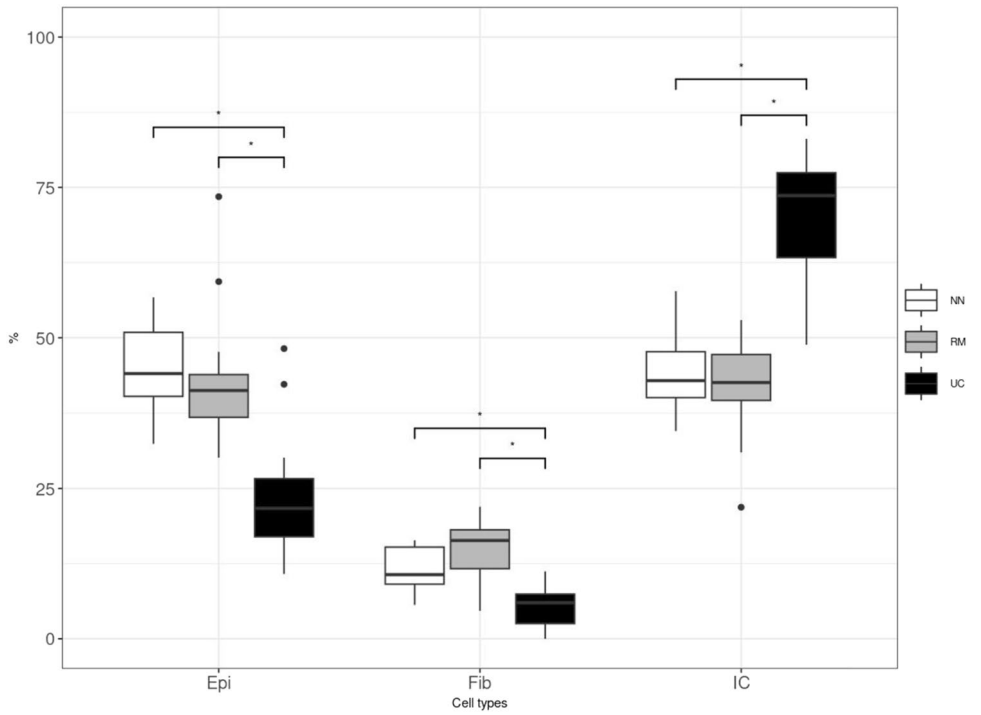


Figure 6. Cell deconvolutions. The cell types are shown in X-axis as Epi (epithelium), Fib (fibroblasts) and IC (immune cells). The percentages of the cell types are shown on the Y-axis.

expression levels of genes several regulatory DNA methylation features of relevance for inactive UC could be identified and are discussed below.

Long-term treatment of UC patients is often necessary to achieve induction and maintenance of clinical remission[16–18]. UC medications such as immunosuppressive drugs have been shown to have side effects on immune response and can change the DNA methylations status[37,38]. Induced epigenetic changes due to long-term treatment can be expected for 5-ASA which is commonly used as a first-line treatment for UC and might therefore have the potential to change the methylation status. This notion is supported by in vitro studies that have shown that 5-ASA treatment increases the expression of DNA methyltransferase 1 (DMNT1) which is responsible for most of the methylation events occurring on the human genome[39] Therefore, it can be anticipated that the observed methylation changes in inactive UC patients might be a result of 5-ASA treatment.

Four genes have been found to be specifically hyper-methylated in inactive UC compared to active UC and controls (NN) (Fig. 2), CERNA2, CIT, IGTA4 and ZNF 626 (Supplementary Data 1). The observed specific hyper-methylation of CERNA2 in inactive UC compared to UC indicates anti-inflammatory characteristics. CERNA2 has been recently reported to play a role in inflammation[40] and has been identified as an independent predictor for clinical prognosis of gastric cancer[41]. It is notable that the expression of CERNA2 has been shown to correlate with poor clinical parameters and an unfavourable prognosis of different cancer patient groups while silencing of CERNA2 expression inhibits cancer cell growth and promotes cell apoptosis[42,43]. In this context, CIT was specifically hyper-methylated in inactive UC. Silencing of CIT has been shown to reduce tumour growth in multiple myeloma and breast cancer cells by promoting apoptosis[44,45]. Therefore, specific hypermethylation of CIT and CERNA2 might play a role in the fine-tuning of the regulation of apoptotic events during inactive UC[21]. ITGA4 is another specifically hyper-methylated gene in inactive UC compared to UC and is a well-known therapeutic target for the treatment of IBD (Fig. 2 & Supplementary Data 2). The observed downregulation of ITGA4 expression implies reduced leukocyte infiltration into the GI tract through the interaction with MAdCAM-1 which is expressed on high endothelial venules (HEV) within vessels of mucosal tissue[46–50].

Quiescent inflammation present in inactive UC has previously been reported[20,51,52] and many of the observed methylation-regulated expressed genes identified in inactive UC are involved in inflammation. Surprisingly, increased levels of a well-known marker for inflammation S100P were observed to a greater extent in inactive UC compared to active UC and controls. The observed specific hypo-methylation and upregulation of S100P in inactive UC might, in addition to S100P's inflammatory responses, contribute to the regulation of tissue development and regeneration or repair as previously reported[53] (Fig. 2 & Supplementary Data 2).

Chronic inflammation like UC results in dramatic deposition of hyaluronic acid (HA) within affected tissues which both precedes and promotes immune cell infiltration, tissue destruction, and coagulation[54]. The observed increased expression of HYAL1 in inactive UC might lead to decreased levels of HA in active UC, thereby modulating the promotion and resolution of the disease by controlling recruitment of immune cells, by release of inflammatory cytokines, and by balancing haemostasis[55] (Supplementary Table 2,Fig. 2). Hypo-methylation and increased expression of MSTR-1 in inactive UC compared to UC might directly be involved in the wound healing process by promoting epithelial cell migration and proliferation, as this innate immune response regulates the migration of macrophages increasing the phagocytic activity[56]. Increased fractions of epithelial cells and decreased fractions of immune cells in inactive UC are in concordance with the results obtained by cell deconvolutions (Fig. 6).

Two other specifically hypo-methylated genes in inactive UC of relevance for IBD include ANXA11 and lncRNA AFAP1-AS1. ANXA11 has been shown to be a sarcoidosis susceptibility gene[57]. An association between sarcoidosis and ulcerative colitis has been reported[58]. The expression of lncRNA AFAP1-AS1 has been shown to promote the progression of CRC[59] and has been acknowledged as a biomarker for diagnosis and prognosis estimation of CRC patients[60]. However, the potential role of the here observed hypo-methylation of AFAP1-AS in inactive UC is still unclear.

Annotations of genes with intermediate methylation patterns for inactive UC revealed two pathways, IL10 signalling and cytokine signalling in immune system pathways (Fig. 5, Supplementary Data 3). These genes were all hyper-methylated and downregulated in inactive UC compared to active UC but to a lesser extent then in controls, meaning that normal levels of expression are not completely achieved by hyper-methylation. The downregulation of the proinflammatory cytokine IL1B, might reduce T cell immune response by downregulation of co-stimulatory molecules such as CD80[61–63]. In addition, IL1B production is diminished by the observed downregulation of STAT3 expression and implies reduced phosphorylation of IL1B. Hyper-methylation of STAT3 in inactive UC compared to active UC might be involved in the regulation of adaptive immune responses by reducing survival of pathogenic T cells and TNF-α[64]. The transcriptional coregulator SBNO2 (strawberry notch homolog 2) is hyper-methylated in inactive UC compared to UC and contributes to the downstream anti-inflammatory effects of IL-10 which is dependent on STAT3 activation[65].

Compared to controls, ZC3H12A is hypo-methylated in inactive UC which might indicate a modulation of the inflammatory response by promoting the degradation of a set of translationally active cytokine-induced inflammation-related mRNAs, such as IL6 and IL12B[66]. ZC3H12A induces the deubiquitylation of the transcription factor HIF1A which is also hyper-methylated in inactive UC compared to active UC, thereby positively regulating the expression of proangiogenic HIF1A-targeted genes[67]. The decrease of HIF1A expression in inactive UC may function as a transcriptional regulator of the adaptive response to hypoxia maintaining biological homeostasis[68]. In this context, cell deconvolutions revealed epithelial cell fractions in inactive UC were comparable to epithelial fractions in control samples (Fig. 6). It is notable that somatic mutations of ZC3H12A have been found in UC patients' epithelium which might have an unknown influence on the DNA methylation regulated expression[69]. The hypo-methylation of LIMK1 in inactive UC compared to controls could lead to reduced T-cell regulation in inactive UC through Rho/Rac pathways. A single nucleotide polymorphism (SNP)

rs6460071 in LIMK1 has been reported to be most significantly associated with proximal endoscopic extension in CRC and is a predictor of outcome in UC[70].

DNA methylation has been found to influence the regulation of interferon's antiviral processes mediated by TRIM22, IGS20, and DENND2D. The hypermethylation of the interferon-induced antiviral protein TRIM22 compared to UC might contribute to a decrease in disease development through the NF-κB signalling pathway[71]. The increased expression of antiviral ISG20 in UC still needs to be confirmed. However, the hypo-methylation and increased expression of ISG20 in inactive UC compared to controls may be a potential susceptibility biomarker or pharmacological target as has been shown for other inflammatory conditions[72]. It is hereby noted that a prognostic impact of expression and methylation status of DENN/MADD domain-containing protein 2D in gastric cancer has been proposed[73–75]. GALNT2 catalyses the initial reaction in O-linked oligosaccharide biosynthesis and has a broad spectrum of substrates for peptides such as MUC5AC, MUC1A, MUC1B. An increase of GALNT2 expression has been recently reported for UC patients in the active stage compared to patients in the remission[76]. This result might be in part also be explained and confirmed by the observed hyper-methylation of GALNT2 in inactive UC compared to UC (Supplementary Data 1, Fig. 2).

A limitation of this work is the small sample size and the heterogeneity of tissue samples as it was not possible to discriminate inactive UC in terms of remission duration and DNA methylation as has been recently seen for gene expression[21]. It is important to note that epigenome-wide association studies do not always precede changes in transcription as has been recently reported[77].

Conclusions

Several differentially expressed genes involved in IL-10/cytokine signalling pathways may be under the control of DNA methylation events which might indicate fine-tuned processes regulating the balance between quiescent inflammation and mucosal healing in inactive UC.

Data availability

Processed RNA-seq data are deposited in NCBI's Gene Expression Omnibus (GEO, https://www.ncbi.nlm.nih.gov/geo/) and are accessible through GEO series accession numbers GSE128682 and GSE169360. Regarding the availability of DNA data, it is hereby noted that, according to the Norwegian Health Research Act §34, the processing of health information can only take place in accordance with the consent given. In this case, the availability of unprocessed DNA information would not be in accordance with the participants' consent. All data generated or analysed during this study are included in this published article and supplementary data files.

Received: 9 October 2023; Accepted: 18 March 2024
Published online: 21 March 2024

References

1. Khor, B., Gardet, A. & Xavier, R. J. Genetics and pathogenesis of inflammatory bowel disease. *Nature* **474**, 307–317 (2011).
2. Peyrin-Biroulet, L., Bressenot, A. & Kampman, W. Histologic remission: the ultimate therapeutic goal in ulcerative colitis?. *Clin Gastroenterol Hepatol* **12**, 929–34 e2 (2014).
3. Jostins, L. *et al.* Host–microbe interactions have shaped the genetic architecture of inflammatory bowel disease. *Nature* **491**, 119–124 (2012).
4. Lee, J. C. *et al.* Genome-wide association study identifies distinct genetic contributions to prognosis and susceptibility in Crohn's disease. *Nat. Genetics* **49**, 262–268 (2017).
5. Moustafa, A. *et al.* Genetic risk, dysbiosis, and treatment stratification using host genome and gut microbiome in inflammatory bowel disease. *Clin. Trans. Gastroenterol.* **9**, e132 (2018).
6. Fakhoury, M., Al-Salami, H., Negrulj, R. & Mooranian, A. Inflammatory bowel disease: clinical aspects and treatments. *J. Inflamm. Res.* **7**, 113 (2014).
7. Aleksandrova, K., Romero-Mosquera, B. & Hernandez, V. Diet, gut microbiome and epigenetics: emerging links with inflammatory bowel diseases and prospects for management and prevention. *Nutrients* **9**, 962 (2017).
8. Zhang, H., Zeng, Z., Mukherjee, A. & Shen, B. Molecular diagnosis and classification of inflammatory bowel disease. *Expert Rev. Mol. Diagn.* **18**, 867–886 (2018).
9. Karatzas, P. S., Mantzaris, G. J., Safioleas, M. & Gazouli, M. DNA methylation profile of genes involved in inflammation and autoimmunity in inflammatory bowel disease. *Medicine* **93**, e309 (2014).
10. McDermott, E. *et al.* DNA methylation profiling in inflammatory bowel disease provides new insights into disease pathogenesis. *J Crohns Colitis* **10**, 77–86 (2016).
11. Harris, R. A. *et al.* DNA methylation-associated colonic mucosal immune and defense responses in treatment-naive pediatric ulcerative colitis. *Epigenetics* **9**, 1131–1137 (2014).
12. 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**, e0248905 (2021).
13. Taman, H. *et al.* Genome-wide DNA methylation in treatment-naïve ulcerative colitis. *J. Crohn's Colitis* **12**, 1338–1347 (2018).
14. Hasler, R. *et al.* A functional methylome map of ulcerative colitis. *Genome. Res.* **22**, 2130–2137 (2012).
15. Dunham, I. *et al.* An integrated encyclopedia of DNA elements in the human genome. *Nature* **489**, 57–74 (2012).
16. Bryant, R., Sandborn, W. & Travis, S. Introducing vedolizumab to clinical practice: who, when, and how?. *J. Crohn's Colitis* **9**, 356–366 (2015).
17. D'Haens, G. Risks and benefits of biologic therapy for inflammatory bowel diseases. *Gut* **56**, 725–732 (2007).
18. Neurath, M. F. & Travis, S. P. Mucosal healing in inflammatory bowel diseases: a systematic review. *Gut* **61**, 1619–1635 (2012).
19. 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).
20. Fenton, C. G., Taman, H., Florholmen, J., Sørbye, S. W. & Paulssen, R. H. Transcriptional signatures that define ulcerative colitis in remission. *Inflamm. Bowel Dis.* **27**, 94–105 (2021).
21. Meng, W., Johnsen, K. M., Fenton, C. G., Florholmen, J. & Paulssen, R. H. Anti-apoptotic genes and non-coding RNAs are potential outcome predictors for ulcerative colitis. *Funct. Integrat. Genom.* **23**, 165 (2023).

22. 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. *J. Crohns Colitis* **11**, 649–670 (2017).
23. 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).
24. Bray, N. L., Pimentel, H., Melsted, P. & Pachter, L. Near-optimal probabilistic RNA-seq quantification. *Nat. Biotechnol.* **34**, 525–527 (2016).
25. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).
26. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–359 (2012).
27. Krueger, F. & Andrews, S. R. Bismark: a flexible aligner and methylation caller for Bisulfite-Seq applications. *Bioinformatics* **27**, 1571–1572 (2011).
28. Korthauer, K., Chakraborty, S., Benjamini, Y. & Irizarry, R. A. Detection and accurate false discovery rate control of differentially methylated regions from whole genome bisulfite sequencing. *Biostatistics* **20**, 367–383 (2019).
29. Kendall, M. G. A new measure of rank correlation. *Biometrika* **30**, 81 (1938).
30. Team BC, M. B. TxDb.Hsapiens.UCSC.hg38.knownGene: annotation package for TxDb object(s). *R package version 3.4.6* (2019).
31. Mi, H. & Thomas, P. PANTHER Pathway: an ontology-based pathway database coupled with dataanalysis tools. *Methods Mol. Biol.* **563**, 123 (2009).
32. Yu, G., Wang, L.-G., Han, Y. & He, Q.-Y. Cluster profiler: an R Package for comparing biological themes among gene clusters. *OMICS A J. Integrat. Biol.* **16**, 284–287 (2012).
33. Zheng, S. C., Breeze, C. E., Beck, S. & Teschendorff, A. E. Identification of differentially methylated cell types in epigenome-wide association studies. *Nat. Methods* **15**, 1059–1066 (2018).
34. Tukey, J. W. Comparing individual means in the analysis of variance. *Biometrics* **5**, 99 (1949).
35. Newman, A. M. *et al.* Robust enumeration of cell subsets from tissue expression profiles. *Nat. Methods* **2015**(12), 453–457 (2015).
36. Paulssen, R. *et al.* Novel regulatory DNA methylation features of relevance for ulcerative colitis. *Am. J. Gastroenterol.* **113**, S12 (2018).
37. Selvestrel, D. *et al.* DNA methylation of the TPMT gene and azathioprine pharmacokinetics in children with very early onset inflammatory bowel disease. *Biomed. Pharmacother.* **157**, 113901 (2023).
38. Guderud, K. *et al.* Methotrexate treatment of newly diagnosed RA patients is associated with DNA methylation differences at genes relevant for disease pathogenesis and pharmacological action. *Front. Immunol.* **12**, 713611 (2021).
39. Adamowicz, M., Milkiewicz, P. & Kempinska-Podhorodecka, A. 5-aminosalicylic acid inhibits the expression of oncomiRs and pro-inflammatory microRNAs: an in vitro study. *J. Physiol. Pharmacol.* **72**, 529–535 (2021).
40. Lu, J. W., Rouzigu, A., Teng, L. H. & Liu, W. L. The construction and comprehensive analysis of inflammation-related ceRNA networks and tissue-infiltrating immune cells in ulcerative progression. *BioMed Res. Int.* **2021**, 6633442 (2021).
41. Zhang, J., Han, X., Jiang, L., Han, Z. & Wang, Z. LncRNA CERNA2 is an independent predictor for clinical prognosis and is related to tumor development in gastric cancer. *Int. J. Clin. Exp. Pathol.* **11**, 5783–5791 (2018).
42. Lin, J., Li, X., Zhang, F., Zhu, L. & Chen, Y. Transcriptome wide analysis of long non-coding RNA-associated ceRNA regulatory circuits in psoriasis. *J. Cell. Mol. Med.* **25**, 6925–6935 (2021).
43. Wang, M., Ouyang, J. & Li, H. CERNA2: A predictor for clinical progression and poor prognosis in cervical carcinoma. *J. Cell. Biochem.* **120**, 11216–11221 (2019).
44. Sahin, I. *et al.* Citron Rho-interacting kinase silencing causes cytokinesis failure and reduces tumor growth in multiple myeloma. *Blood Adv.* **3**, 995–1002 (2019).
45. Meng, D. *et al.* Citron kinase (CIT-K) promotes aggressiveness and tumorigenesis of breast cancer cells in vitro and in vivo: preliminary study of the underlying mechanism. *Clin. Transl. Oncol.* **21**, 910–923 (2019).
46. Berlin, C. *et al.* Alpha 4 beta 7 integrin mediates lymphocyte binding to the mucosal vascular addressin MAdCAM-1. *Cell* **74**, 185–195 (1993).
47. Habtezion, A., Nguyen, L. P., Hadeiba, H. & Butcher, E. C. Leukocyte trafficking to the small intestine and colon. *Gastroenterology* **150**, 340–354 (2016).
48. Butcher, E. C., Williams, M., Youngman, K., Rott, L. & Briskin, M. Lymphocyte trafficking and regional immunity. *Adv. Immunol.* **72**, 209–253 (1999).
49. Dotan, I. *et al.* The role of integrins in the pathogenesis of inflammatory bowel disease: Approved and investigational anti-integrin therapies. *Med. Res. Rev.* **40**, 245 (2020).
50. Nakache, M., Lakey-Berg, E., Streeter, P. R. & Butcher, E. C. The mucosal vascular addressin is a tissue-specific endothelial cell adhesion molecule for circulating lymphocytes. *Nature* **337**, 179–181 (1989).
51. Riley, S. A., Mani, V., Goodman, M. J., Dutt, S. & Herd, M. E. Microscopic activity in ulcerative colitis: what does it mean?. *Gut* **32**, 174–178 (1991).
52. Korelitz, B. I. Mucosal healing as an index of colitis activity: back to histological healing for future indices. *Inflamm. Bowel Dis.* **16**, 1628–1630 (2010).
53. Xia, C., Braunstein, Z., Toomey, A. C., Zhong, J. & Rao, X. S100 proteins as an important regulator of macrophage inflammation. *Front. Immunol.* **8**, 1908 (2018).
54. Petrey, A. C. & de la Motte, C. A. Hyaluronan in inflammatory bowel disease: cross-linking inflammation and coagulation. *Mat. Biol.* **78–79**, 314 (2019).
55. Dokoshi, T. *et al.* Hyaluronidase inhibits reactive adipogenesis and inflammation of colon and skin. *JCI Insight* **3**, e123072 (2018).
56. Han, X., Ding, S., Jiang, H. & Liu, G. Roles of macrophages in the development and treatment of gut inflammation. *Front.Cell Dev. Biol.* **9**, 625423 (2021).
57. Karakaya, B. *et al.* ANXA11 rs1049550 associates with Löfgren's syndrome and chronic sarcoidosis patients. *Cells* **11**, 1557 (2022).
58. Fries, W. *et al.* Association between inflammatory bowel disease and sarcoidosis. Report of two cases and review of the literature. *Scand. J. Gastroenterol.* **30**, 1221–1223 (1995).
59. Li, Y., Zhu, Z., Hou, X. & Sun, Y. LncRNA AFAP1-AS1 promotes the progression of colorectal cancer through miR-195–5p and WISP1. *J. Oncol.* **2021**, (2021).
60. Elabd, N. S. *et al.* Long non-coding RNAs ASB16-AS1 and AFAP1-AS1: diagnostic, prognostic impact and survival analysis in colorectal cancer. *Appl. Clin. Genetics* **15**, 97 (2022).
61. MacDermott, R. P. Chemokines in the inflammatory bowel diseases. *J. Clin. Immunol.* **19**, 266–272 (1999).
62. Scarpa, M. *et al.* CD80-CD28 signaling controls the progression of inflammatory colorectal carcinogenesis. *Oncotarget* **6**, 20058–20069 (2015).
63. Nakazawa, A. *et al.* Functional expression of costimulatory molecule CD86 on epithelial cells in the inflamed colonic mucosa. *Gastroenterology* **117**, 536–545 (1999).
64. Sugimoto, K. Role of STAT3 in inflammatory bowel disease. *World J. Gastroenterol.* **14**, 5110 (2008).
65. El Kasmi, K. C. *et al.* Cutting edge: A transcriptional repressor and corepressor induced by the STAT3-regulated anti-inflammatory signaling pathway. *J. Immunol.* **179**, 7215–7219 (2007).

66. Garg, A. V. *et al.* MCPIP1 endoribonuclease activity negatively regulates interleukin-17-mediated signaling and inflammation. *Immunity* **43**, 475–487 (2015).
67. Roy, A., Zhang, M., Saad, Y. & Kolattukudy, P. E. Antidicer RNAse activity of monocyte chemotactic protein-induced protein-1 is critical for inducing angiogenesis. *Am. J. Physiol. Cell Physiol.* **305**, C1021–C1032 (2013).
68. Kerber, E. L. *et al.* The importance of hypoxia-inducible factors (HIF-1 and HIF-2) for the pathophysiology of inflammatory bowel Disease. *Int. J. Mol. Sci.* **21**, 1–27 (2020).
69. Nanki, K. *et al.* Somatic inflammatory gene mutations in human ulcerative colitis epithelium. *Nature* **577**, 254–259 (2020).
70. Waterman, M. *et al.* Predictors of outcome in ulcerative colitis. *Inflamm. Bowel Dis.* **21**, 2097 (2015).
71. Ye, B. & Lu, Z. Role of TRIM22 in ulcerative colitis and its underlying mechanisms. *Mol. Med. Rep.* **26**, 1–10 (2022).
72. Zheng, Z., Wang, L. & Pan, J. Interferon-stimulated gene 20-kDa protein (ISG20) in infection and disease: review and outlook. *Intractable Rare Dis. Res.* **6**, 35–40 (2017).
73. Kanda, M. *et al.* Prognostic impact of expression and methylation status of DENN/MADD domain-containing protein 2D in gastric cancer. *Gastric Cancer* **18**, 288–296 (2015).
74. Hibino, S. *et al.* Reduced expression of DENND2D through promoter hypermethylation is an adverse prognostic factor in squamous cell carcinoma of the esophagus. *Oncol. Rep.* **31**, 693–700 (2014).
75. Kanda, M. *et al.* Downregulation of DENND2D by promoter hypermethylation is associated with early recurrence of hepatocellular carcinoma. *Int. J. Oncol.* **44**, 44–52 (2014).
76. Tian, R. *et al.* A pharmacoinformatics analysis of artemisinin targets and de novo design of hits for treating ulcerative colitis. *Front. Pharmacol.* **13**, 843043 (2022).
77. Pacis, A. *et al.* Gene activation precedes DNA demethylation in response to infection in human dendritic cells. *Proc. Natl. Acad. Sci. USA* **116**, 6938–6943 (2019).

Acknowledgements

The authors thank Lars Daniel Håland at the Genomics Support Centre Tromsø (GSCT; (www.uit.no/helsefak/genomics) for technical assistance.

Author contributions

W.M. performed statistical and functional analyses, visualization, and wrote parts of the manuscript. K.M.J. was involved in evaluating and providing clinical samples from patients in long-term remission. C.G.F. was involved in data analysis and visualization. H.T. was involved in library preparation and sequencing, J.F. was involved in evaluating and providing clinical samples from patients and healthy controls. R.H.P was involved in project inception, project design, data analysis, supervision, manuscript writing. All authors revised the manuscript and approved the last version of the manuscript.

Funding

Open access funding provided by UiT The Arctic University of Norway (incl University Hospital of North Norway). This work was supported by institutional funding.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-024-57440-0.

Correspondence and requests for materials should be addressed to R.H.P.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Paper III

International Journal of
Molecular Sciences

MDPI

Article

Methylation-Regulated Long Non-Coding RNA Expression in Ulcerative Colitis

Christopher G. Fenton [1,2,†], Mithlesh Kumar Ray [1,†], Wei Meng [1] and Ruth H. Paulssen [1,2,*]

[1] Clinical Bioinformatics Research Group, Department of Clinical Medicine, UiT-The Arctic University of Norway, N-9037 Tromsø, Norway; christopher.fenton@uit.no (C.G.F.); wei.meng@uit.no (W.M.)
[2] Genomic Support Centre Tromsø (GSCT), Department of Clinical Medicine, UiT-The Arctic University of Norway, N-9037 Tromsø, Norway
* Correspondence: ruth.h.paulssen@uit.no; Tel.: +47-77645480
† These authors contributed equally to this work.

Abstract: Long non-coding RNAs (lncRNAs) have been shown to play a role in the pathogenesis of ulcerative colitis (UC). Although epigenetic processes such as DNA methylation and lncRNA expression are well studied in UC, the importance of the interplay between the two processes has not yet been fully explored. It is, therefore, believed that interactions between environmental factors and epigenetics contribute to disease development. Mucosal biopsies from 11 treatment-naïve UC patients and 13 normal controls were used in this study. From each individual sample, both whole-genome bisulfite sequencing data (WGBS) and lncRNA expression data were analyzed. Correlation analysis between lncRNA expression and upstream differentially methylated regions (DMRs) was used to identify lncRNAs that might be regulated by DMRs. Furthermore, proximal protein-coding genes associated with DMR-regulated lncRNAs were identified by correlating their expression. The study identified UC-associated lncRNAs such as MIR4435-2HG, ZFAS1, IL6-AS1, and Pvt1, which may be regulated by DMRs. Several genes that are involved in inflammatory immune responses were found downstream of DMR-regulated lncRNAs, including SERPINB1, CCL18, and SLC15A4. The interplay between lncRNA expression regulated by DNA methylation in UC might improve our understanding of UC pathogenesis.

Keywords: long non-coding RNAs; DNA methylation; ulcerative colitis; epigenetics

Citation: Fenton, C.G.; Ray, M.K.; Meng, W.; Paulssen, R.H. Methylation-Regulated Long Non-Coding RNA Expression in Ulcerative Colitis. *Int. J. Mol. Sci.* **2023**, *24*, 10500. https://doi.org/10.3390/ijms241310500

Academic Editors: Shinichi Fukushige and Caterina Cinti

Received: 13 April 2023
Revised: 26 May 2023
Accepted: 20 June 2023
Published: 22 June 2023

1. Introduction

Ulcerative colitis (UC) is a relapsing chronic inflammatory disease of the colon and one of the most common conditions of inflammatory bowel disease (IBD) [1]. The development of UC is influenced by a complex interplay between the host immune system, genetic variation, intestinal microbiota, and environmental factors [2,3]. The link between environmental factors and the genome is thought to be via epigenetic mechanisms, including DNA methylation [4], histone modifications [5], and interactions with non-coding RNAs [6]. Methylation can alter the expression of genes associated with UC pathogenesis [7–9].

Long non-coding RNAs (lncRNAs) are transcripts that are longer than 200 nt and have no protein-coding capacity. LncRNAs have multiple mechanisms to regulate gene expression including the modulation of transcription, mRNA stability, translation, and protein subcellular location by interacting with DNA, RNA, or protein to form large complexes [10]. LncRNAs have been shown to play a significant role in various biological processes including the regulation of gene expression, epigenetic regulation, and disease development [10]. Several studies have identified lncRNAs playing a role in the disease development and pathogenesis of UC [11–17]. DNA methylation is a key regulator of gene expression and contributes to lncRNA expression [18].

The interplay between DNA methylation and lncRNA expression has been implicated in various biological processes, including embryonic development, cancer, and neurolog-

ical disorders [19–21]. The interplay between lncRNAs and methylation is not limited to promoter methylation but represents part of a complex regulatory network [21]. Like protein-coding genes, the transcription of lncRNAs can be affected by promoter methylation [22]. LncRNAs may in turn regulate the epigenome by interacting with different epigenetic factors including DNMTs or other genes involved in chromatin organization [23]. The crosstalk between DNA methylation and lncRNAs has been confirmed by findings regarding lncRNA promoter methylation and dysregulation in response to methylation inhibitor treatments [21]. Changes in the promoter methylation state cause the significant dysregulation of many lncRNAs, including Pvt1, NEAT1, and LINC00261, and play a role in disease pathogenesis [21,24]. This study focuses on lncRNAs that may be regulated by differentially methylated regions (DMRs).

This study aims to provide valuable knowledge for future functional studies of lncRNAs associated with UC pathogenesis.

2. Results

A schematic overview of the methods and software used to generate the results used in this study is presented in Figure 1. The study workflow comprised several steps: WGBS (whole-genome bisulfite sequencing) data were aligned to the human reference genome using Bismark, and RNAseq fastq files were aligned to the human reference transcriptome using Kallisto. Differentially methylated regions (DMR) and differentially expressed (DE) transcripts were identified using DMRSeq and DESeq2, respectively. Using correlation analysis between lncRNA expression and adjacent DMR methylation levels, potentially methylation-regulated lncRNAs were selected. Methylation-regulated lncRNA expression was correlated with adjacent protein-coding transcript expression to predict target protein-coding genes for selected lncRNAs. The obtained results were visualized using Gviz and verified with ten other GEO UC datasets.

Material	UC cohort DNA (11 UC, 13 control)	
Type of analysis	Expression	Methylation
Laboratory method	RNAseq	WGBS
Alignment softare	Kallisto	Bismark
Analysis software	DEseq2	DMRseq
Selection	Adjusted p.value < 0.5 abs(log2Foldchange) > 0.5	q.value < 0.05
Result	Differentially expressed transcripts	Differentially methylated regions (DMR)
Integration	DMR within 20 kb of lncRNA DMR negatively correlated with lncRNA expression Protein coding genes within 500 kb of lncRNA Protein coding gene expression negatively correlated with lncRNA	
Visualization	Genome region plots	
Verification	Collect and compare results with 11 GEO datasets	

Figure 1. Schematic overview of material, methods, and software used in the study.

2.1. Identification of Differentially Expressed Transcripts with DESeq2

DEseq2 was run on the transcript count matrix generated by the Kallisto aligner on raw Illumina fastq reads, generated from 11 treatment-naïve mucosal biopsy UC samples and 13 control samples. A total of 1292 lncRNAs had an adjusted *p*-value less than 0.05 and an absolute fold change value greater than 0.5.

2.2. Identification of Differentially Methylated Regions (DMRs) with DMRseq

A total of 5796 DMRs were obtained with a q-value < 0.05 in the UC samples (n = 11) compared with the normal control group (n = 13). The DMRs included 1380 hypermethylated and 4416 hypomethylated regions (Table S1). The average size of the DMRs was 288 bp, and the average number of CpGs in the DMRs was 15.

2.3. LncRNAs That May Be Regulated by DMRs

LncRNAs that were within 20 kb upstream or downstream of a DMR and whose expression negatively correlated with DMR methylation levels were considered lncRNAs that are potentially regulated by a proximal DMR. A total of 254 lncRNAs met the above criteria. A total of 188 lncRNA were upregulated in UC, and 66 were downregulated in UC (Table S2).

2.4. Proteins That May Be Influenced by DMR-Regulated LncRNAs

Differentially expressed protein-coding genes that were within 500 kb upstream or downstream of a DMR-regulated lncRNA were considered for correlational expression analysis. A total of 244 protein-coding genes were found whose expressions were significantly and negatively correlated with lncRNA expression. This discussion focuses on those genes that may play a role in UC pathogenesis. Of the above proteins, 110 were upregulated in UC, and 134 were downregulated in UC versus the control. The results are summarized in Tables S3 and S4. Figure 2 shows an example of a genomic region containing a DMR, DE lncRNA transcripts, and DE protein-coding transcripts. An example of the correlation between the DMRs, DE lncRNA transcripts, and adjacent DE protein-coding transcripts is shown in Figure 3. All genomic regions of interest can be seen in Figure S1.

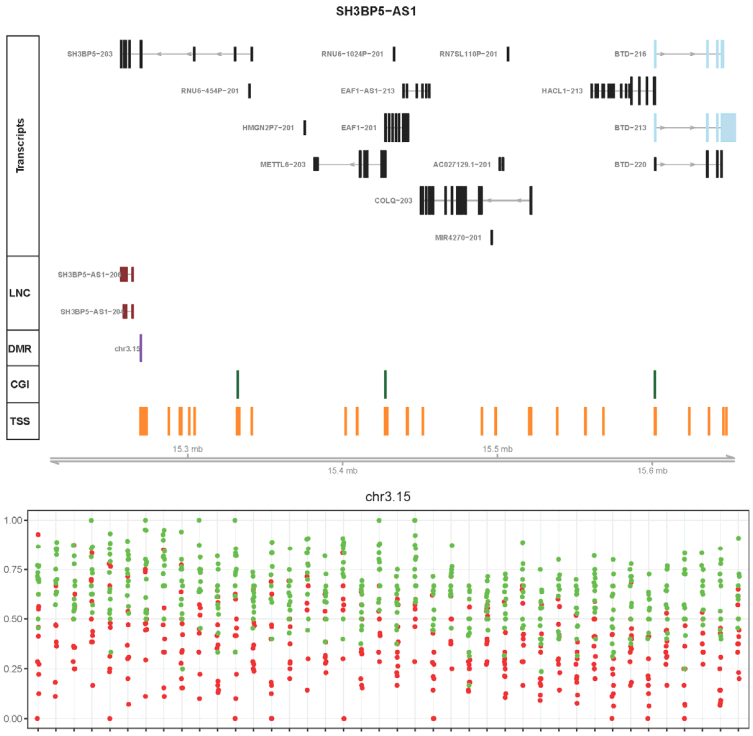


Figure 2. Example of a genomic region containing a differentially methylated region (DMR) chr3.15 and differentially expressed (DE) lncRNA SH3BP5-AS1 transcripts. The top transcript track represents the regions found between the DMR, lncRNA transcripts, and DE protein-coding transcripts of interest. Transcripts indicated in light blue denote DE protein-coding transcripts that may be influenced by DMR-regulated lncRNA transcripts, which are shown in brown. Transcripts indicated in black are the largest transcripts for each gene found within the region. The LNC track denotes the position of the DE lncRNA transcripts; the DMR track denotes the position of the DMR, which is shown in purple. The CGI track denotes the position of known CpG islands, which are shown in green. The TSS (transcription starting site) track denotes the position of known TSSs, which are shown in orange. The bottom track of the top panel shows the approximate distance in Mb. The bottom panel shows the relative methylation levels for the chr3.15 DMR. Red dots indicate the relative methylation values of the UC samples. The relative methylation values from the control samples are indicated in green.

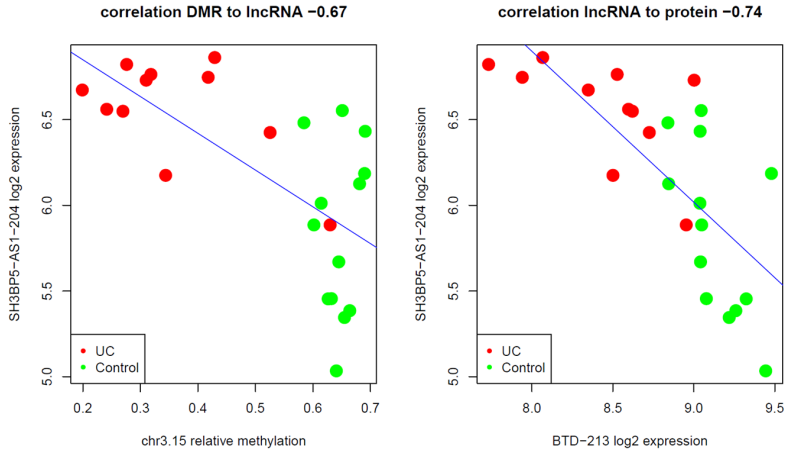


Figure 3. An example of correlations between sample DMR methylation levels, lncRNA, and adjacent protein-coding transcript expressions. On the left, the correlation between differentially expressed (DE) lncRNA transcript SH3BP5-AS1-204 and the mean-sample relative methylation levels of DMR chr3.15. On the right is the correlation between DE lncRNA transcript SH3BP5-AS1-204 expression and proximal protein-coding DE BDT transcripts.

2.5. Cell Deconvolution

To estimate types of cell fractions in UC and the normal controls' mucosal tissues, the EpiDISH cell deconvolution algorithm was adapted for use with methylation data. The deconvolution estimated relative fractions of epithelial, fibroblast, and immune cells present in the tissue samples. A cell-type fraction estimate revealed increased fractions of immune cells in tissues from UC patients, whereas fractions of epithelial cells and fibroblasts were increased in the control samples (Figure 4).

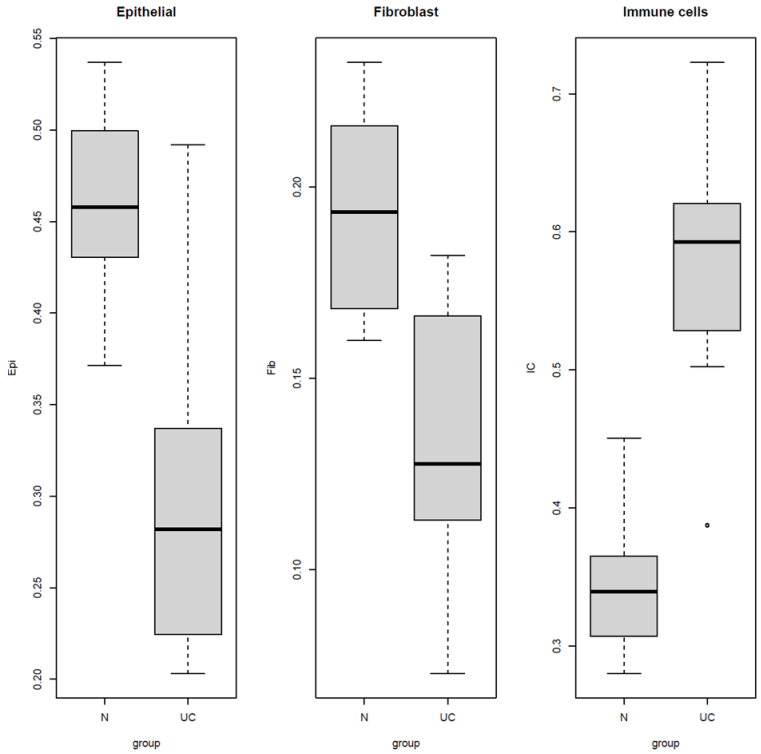


Figure 4. Box plots of fractions of cell types present in normal and UC tissue samples. Each plot indicates a significant difference in cell distribution between UC and normal samples. The Y-axis depicts cell fractions of tissue samples ranging from 0 to 1. The X-axis indicate the range of cell fractions in control (N) and UC samples.

2.6. Verification of DMR-Regulated lncRNAs and Proximal Proteins

To help verify the correlation between lncRNAs and adjacent protein expression, normalized matrices from 11 datasets were collected: GSE109142, GSE128682, GSE206285, GSE36807, GSE38713, GSE47908, GSE13367, GSE16879, GSE48958, GSE59071, and GSE73661. A total of 35 lncRNAs showed a significant correlation with adjacent protein expression in at least one dataset (Table S5). An overview of the number of samples in each GEO dataset, as well as sample locations, is shown in Table S6.

3. Discussion

Environmental factors have been implicated in both the incidence of UC and the likelihood of relapse in UC patients [25] and are thought to have a direct effect on the epigenome, including the expression of lncRNAs and methylation status [26]. Both lncRNA and DNA methylation have been shown to regulate the transcription of protein-coding genes [18]. However, the interplay between DNA methylation, the expression of lncRNAs, and the expression of protein-coding genes has not been explored in detail in UC.

The focus of this study was to identify lncRNAs that were negatively correlated with adjacent DMR methylation levels. The implication is that elevated levels of DMR methylation (hypermethylation) in UC samples should result in lower adjacent lncRNA expression and vice versa (hypomethylation). To explore the possible cis effects of these lncRNAs, neighboring DE protein-coding genes whose expression negatively correlated with lncRNA expression were identified. This ensures that lncRNAs and adjacent protein-coding genes are unlikely to be regulated by the same DMR. Defining the lncRNA cis-regulation of gene expression is difficult, as lncRNAs have been shown to regulate the expression of both proximal and distal genes [27]. Recent reports suggest that the 3D conformation of the genome guides lncRNAs to distal binding sites [28]. Therefore, several studies have considered the possible effects of lncRNA expression on genes within 500 kb of lncRNAs [29,30].

Recent publications have shown that methylation events outside 1–2 kb of the promoter can have effects on gene expression. It has been shown that increasing the range queried from 5 kb to 20 kb can add an additional ~0.5% of DEGs that associate with the identified DMRs [31]. Therefore, the influence of methylation on lncRNA expression in DMRs within 20 kb was considered.

The results identified protein-coding genes and lncRNAs that were previously associated with UC. Protein-coding genes adjacent to possible DMR-regulated lncRNAs include chemokine C-C motif ligand 18 (CCL18), potassium voltage-gated channel subfamily B member 1 (KCNB1), and serpin family B member 1 (SERPINB1). The increased expression of CCL18, which has been linked to inflammation and the migration of T cells, is correlated with the expression of lncRNA AC244100.3 [32]. KCNB1 is correlated with DE lncRNA ZFAS1 and is downregulated in active UC. KCNB1 regulates the cellular K^+-efflux necessary for enterocyte apoptosis and has been proposed as a therapeutic target for IBD [33]. In addition, KCNB1 has been identified in several cancers, including gastric and colorectal cancers (CRC). KCNB1 is downregulated in both CRC and gastric cancers [34,35]. The expression of lncRNA GMDS-DT is correlated with the expression of neutrophil elastase (NE) inhibitor protein-coding gene SERPINB1. In UC, activated neutrophils secrete NE, which plays a key role in colonic epithelial cell destruction. The increased expression levels of SERPINB1 might protect colonic epithelial cells by reducing NE activity [36].

Potentially DMR-regulated lncRNAs have been implicated in immunity, inflammation, and IBD, including AC007750.1 (lnc-SLC4A10-7), SH3BP5 antisense RNA 1 (SH3BP5-AS1), FOXD2-adjacent opposite strand RNA 1 (FOXD2-AS1), mir4435-2 host gene (MIR4435-2HG), and cytoskeleton regulator RNA (CYTOR). The expression of AC007750.1 is correlated with DPP-4 (dipeptidyl peptidase-4) expression, which is a potential biomarker for IBD. DPP-4 stimulates the production and release of cytokines, chemokines, and neuropeptides, thereby playing a role in the inflammatory response [37,38]. LncRNA SH3BP5-AS1 is correlated with biotinidase (BTD). The association between DMR, SH3BP5-AS1, and BTD

is shown in Figures 2 and 3. Biotin deficiency plays a role in the induction of Th1- and TH17-mediated proinflammatory responses [39]. The observed downregulation of BTD in UC may result in the dysfunction of cellular immune responses [40].

A reduction in FOXD2-AS1 expression correlates with an upregulation of PDZK1-interacting protein 1 (PDZK1IP1) in UC, which may contribute to the inflammatory responses associated with UC [41].

The dysregulation of MIR4435-2HG in UC might play a key role in the inflammatory process and has been shown to be associated with CRC [37,42,43]. MIR4435-2HG is correlated with the expression of B cell lymphoma 2 (Bcl-2)-interacting protein (BCL2L11), which is associated with an increase in apoptosis resistance, resulting in impaired epithelial cell turnover [44]. In addition, BCL2L11 also plays a major role in immune tolerance in UC [45]. CYTOR plays a role in promoting inflammation and epithelial–mesenchymal transition, ultimately promoting cellular invasion and CRC progression [46]. The expression of lncRNA CYTOR is correlated with the expression of FABP1, which is involved in the intestinal absorption of dietary long-chain fatty acids [47]. The dysregulation of CYTOR may disrupt FABP1-mediated fatty acid metabolism, which has been implied to contribute to the pathophysiology of UC [48,49].

Tissue samples are heterogeneous, and DNA methylation is a highly cell-type-specific event [50]. Therefore, EpiDISH cell deconvolution was adapted for use with methylation data and used to estimate cell-type fractions in both UC and control samples (Figure 2). EpiDISH was chosen simply because over 70% of the DMR sites overlapped known Illumina EPIC array sites. EPIC arrays are widely used to study methylation. The distribution of cell fractions was consistent with previous deconvolution results obtained from transcriptomic analysis of active UC [42]. The reduced epithelial fraction may be indicative of cell degradation, which is a major characteristic of UC [51].

Our results show several potentially DMR-regulated lncRNAs associated with epithelial cell proliferation and migration, including HOXA-AS2 and HOXA-AS3 [52,53]. Interestingly, these lncRNAs are under DMR regulation and are downregulated in UC. The downregulation of HOXA-AS2 and HOXA-AS3 may reduce epithelial cell differentiation and migration during UC. The increased proportion of immune cells in the colon of patients with UC is due to the recruitment and activation of these cells in response to ongoing inflammation in the gut [54]. The epigenetically upregulated lncRNAs ADORA2A-AS1 [55] and IL6-AS1 [56] may be associated with immune cell infiltration, which is a characteristic of inflammation. These potentially DMR-regulated lncRNAs may help explain the higher abundance of immune cells in UC patients. Several of the DMR-regulated lncRNA expressions in this study were found to be differentially expressed in UC in our previous study (114 of 254) [17].

Verifying results in GEO (Gene Expression Omnibus) is difficult. No independent datasets with both methylation levels and gene expression levels for UC could be found. Therefore, an attempt was made to see if significant negative correlations between the lncRNAs and adjacent expression of protein-coding genes could be found in 11 published UC GEO datasets. Comparing annotations between GEO datasets is difficult, as recently annotated lncRNAs such as AL359962 simply do not appear in previously deposited microarray datasets, leaving approximately 58 lncRNAs that could be found in at least 1 of the 11 UC–control GEO datasets. Another challenge is that several of the 11 GEO datasets selected to verify the correlation between lncRNAs and adjacent protein-coding genes were samples collected from locations other than mucosal biopsies, including the ileum, the rectum, etc. (Table S6). For 35 lncRNAs, at least 1 GEO set confirmed a significant correlation between the lncRNA and adjacent protein expression (Table S5). For the 35 lncRNAs, a significant correlation was found, on average, in 25% of the datasets. Given the diversity of the GEO datasets, this represents a positive result. The normalized count matrix for this experiment can be found in Supplementary Table S7.

As a limitation of this work, it is hereby noted that the results presented are derived from in silico analysis and need experimental validation in the future.

4. Materials and Methods

4.1. Study Cohort

The study cohort comprised mucosal biopsies from patients with newly diagnosed, treatment-naïve UC with mild-to-moderate disease (n = 11) and control subjects (n = 13). Tissue samples from subjects which underwent cancer screening and showed normal colonoscopy and normal colonic histological examinations, served as controls. UC was diagnosed based on established clinical endoscopic and histological criteria, as defined by ECCO guidelines [57]. The grade of inflammation was assessed during colonoscopy using the UC disease activity index (UCDAI) endoscopic sub-score, with 3 to 10 indicating mild-to-moderate disease [58]. The biopsies from UC samples showed clinical scores of 8.2 ± SD 1.3 and endoscopic scores of 1.9 ± SD 0.5. The biopsies from the control subjects showed normal colonoscopies, colon histology, and immunochemistry, with clinical and endoscopic scores of 0. All biopsies were taken from the sigmoid part of the colon. The age distribution within the groups was 39 ± SD 12 years in the UC group and 53 ± SD 18 in the control subjects. The gender distribution was 7 males and 4 females in the UC group and 11 males and 2 females in the control group. 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 Norwegian Social Science Data Services (REK Nord 2012/1349). The raw fastq files of the transcriptomes were generated previously (GSE 128682), and raw WGBS fastq files from a previously published work were used [7]. However, to obtain optimal results, only the highest-coverage WGBS samples were included in the cohort of this study. Both transcriptomic data and data obtained by WGBS were reanalyzed for this manuscript, with a newer human genome build (GENCODE V38).

4.2. DNA and RNA Isolation

Both DNA and RNA were isolated using the Allprep DNA/RNA Mini Kit from Qiagen (Cat no: 80204) and the QIAcube instrument (Qiagen, Venlo, The Netherlands) according to the manufacturer's protocol. RNA and DNA quantity and purity were assessed as previously described [7,42]. All RNA samples used for analyses had a RIN value between 8.0 and 10.0. DNA and RNA samples were kept at −70 °C until further use.

4.3. Library Preparation and Next-Generation Sequencing

Library preparations and sequencing were conducted as described previously [7,42].

4.4. Preprocessing of Data

The human reference genome hg38 was downloaded from GENCODE and indexed using Bismark version 0.22.3. The data from each sample were then aligned to the indexed reference genome using the Bowtie2 aligner within Bismark. The methylation level in each cytosine was then determined using Bismark with the following parameters: –gzip –bedGraph—cytosine_report –no_overlap—buffer_size 10 G –paired –ignore 3 –ignore_r2 3 —-ignore_3prime_r2 2. Methylation data output contained read coverage and the percentage of methylated cytosine at each cytosine position of the genome.

4.5. Identification of DMRs

The R DMRseq package (version 1.4.9) was used to find differentially methylated regions (DMRs) between UC samples and normal samples from the Bismark output files. CpG sites with less than 6× coverage were set to 0 prior to DMRseq analysis, and only CpG sites with a minimum of 6× coverage in 50% of both groups were kept, as recommended by the software. DMRs with DMRseq q-values of less than 0.05 were considered significantly differentially regulated regions (Table S1).

4.6. Cell Deconvolution

To compare methylation with transcriptional cell deconvolution, the EpiDISH package in R (https://bioconductor.org/packages/release/bioc/html/EpiDISH.html, accessed on 21 January 2023) was adapted to estimate the relative proportions of different cell types present in a tissue sample. EpiDISH requires Illumina EPIC array identifiers and a matrix of beta values. DMRs were given EPIC array identifiers by overlapping DMR genomic positions with EPIC array positions. Approximately 70% of DMR locations overlapped within EPIC-array-annotated genomic positions. A matrix of the average relative methylation value per sample per DMR was used as the beta matrix. The Robust Partial Correlation (RPC) mode in EpiDISH was utilized to estimate the relative numbers of epithelial, fibroblast, and immune cells in each sample (UC and control).

4.7. RNAseq

Illumina-generated fastq sequences were aligned with a reference human transcriptome using the Kallisto RNA-seq aligner. The transcript read count table from the Kallisto output was imported into the DESseq2 R package for identifying differentially expressed transcripts. The lncRNA catalog was retrieved from GENCODE V38 using the transcript type "lncRNA". Only transcripts with a DESeq2-adjusted p-value of < 0.05 and an absolute foldchange greater than 0.5. were considered differentially expressed DE transcripts. The vst function of the DESeq2 package was used to create a normalized count matrix in the correlational analyses.

4.8. Identifying lncRNAs That May Be under DMR Regulation

DMRs located within 20 kb of a DE lncRNA were considered for correlation analysis. The R cor.test package was used to calculate the correlation and correlational p-value between the mean-sample relative methylation and DE lncRNA-normalized transcript counts. Only DE lncRNAs whose transcript expressions were negatively correlated with DMR methylation levels (correlation p-value of < 0.05) were considered possible DMR-regulated lncRNAs (Table S2).

4.9. Identifying Proteins That May Be under DMR-Regulated lncRNA Regulation

Only differentially expressed protein-coding transcripts within 500 kb of the DMR-regulated lncRNAs were considered. The lncRNA expression was then correlated with the neighboring proteins using the R cor.test package. Only protein-coding transcripts that significantly negatively correlated (correlation p-value of < 0.05) with DMR-regulated lncRNA transcripts were considered (Table S3). The R Gviz package was used to help visualize the relationship between the DMR methylation level, lncRNA transcript expression, lncRNA-DMR correlation, CpG islands, and TSS (Figure S1). TSS annotation was downloaded from the refTSS database (http://reftss.clst.riken.jp/reftss/Main_Page, accessed on 17 December 20222). The CpG island positions of the human genome (hg38) were downloaded from the UCSC table browser (https://genome.ucsc.edu/cgi-bin/hgTables, accessed on 17 December 2022).

4.10. Verification of DMR-Regulated lncRNAs and Proximal Partners in Other GEO Datasets

To help verify the DMR-regulated lncRNA and proximal protein results, the normalized matrices of the UC and control samples from 11 UC datasets (GSE109142, GSE128682, GSE206285, GSE36807, GSE38713, GSE47908, GSE13367, GSE16879, GSE48958, GSE59071, and GSE73661) were used. Table S5 compares the expression of lncRNAs, and adjacent proteins found in this study with the above datasets. Specifically, other datasets where a significant negative correlation between lncRNAs and adjacent protein-coding regions could be found. Additional information about the mean difference in expression (UC vs. control) for lncRNAs and adjacent proteins is provided in Table S5. Background information about the GEO datasets can be found in Table S6, including the number of UC and control samples, and their origin.

5. Conclusions

This study suggests a fine-tuned and complex regulatory mechanism between methylation, lncRNAs, and protein expression in UC. The results might open new avenues for diagnostic or therapeutic strategies.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/ijms241310500/s1.

Author Contributions: C.G.F.: Data curation, conceptualization, methodology, investigation, visualization, validation, software, writing, and review and editing. M.K.R.: Formal analysis, validation, writing, and reviewing the final draft. W.M.: validation and visualization. R.H.P.: Conceptualization, investigation, validation, project administration, resources, methodology, supervision, writing, and review and editing. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by institutional funding.

Institutional Review Board Statement: This study was conducted according to the guidelines of the Declaration of Helsinki. Approvals were granted by the Regional Committee of Medical Ethics of Northern Norway, REK Nord 2012/1349. The samples were taken from an established biobank approved by the Norwegian Board of Health (952/2006).

Informed Consent Statement: Written informed consent has been obtained from the study participants to publish this paper.

Data Availability Statement: The DESeq2 VST-normalized RNA-seq Kallisto transcript count matrix for the samples (Table S4) and all other data generated or analyzed during this study are included in the published article and Supplementary Materials. Regarding the availability of the DNA data, it is hereby noted that, according to Norwegian Health Research Act § 34, the processing of health information can only take place in accordance with the consent given. In this case, the availability of unprocessed DNA information would not be in accordance with the participants' consent.

Acknowledgments: The authors thank Hagar Taman at the Genomics Support Centre Tromsø (GSCT; (www.uit.no/helsefak/genomics) for technical assistance.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Du, L.; Ha, C. Epidemiology and Pathogenesis of Ulcerative Colitis. *Gastroenterol. Clin. N. Am.* **2020**, *49*, 643–654. [CrossRef] [PubMed]
2. De Souza, H.S.P.; Fiocchi, C.; Iliopoulos, D. The IBD interactome: An integrated view of aetiology, pathogenesis and therapy. *Nat. Rev. Gastroenterol. Hepatol.* **2017**, *14*, 739–749. [CrossRef] [PubMed]
3. Ramos, G.P.; Papadakis, K.A. Mechanisms of Disease: Inflammatory Bowel Diseases. *Mayo Clin. Proc.* **2019**, *94*, 155–165. [CrossRef] [PubMed]
4. Aleksandrova, K.; Romero-Mosquera, B.; Hernandez, V. Diet, Gut Microbiome and Epigenetics: Emerging Links with Inflammatory Bowel Diseases and Prospects for Management and Prevention. *Nutrients* **2017**, *9*, 962. [CrossRef] [PubMed]
5. Fellows, R.; Varga-Weisz, P. Chromatin dynamics and histone modifications in intestinal microbiota-host crosstalk. *Mol. Metab.* **2020**, *38*, 100925. [CrossRef]
6. Malmuthuge, N.; Guan, L.L. Noncoding RNAs: Regulatory Molecules of Host–Microbiome Crosstalk. *Trends Microbiol.* **2021**, *29*, 713–724. [CrossRef]
7. Taman, H.; Fenton, C.G.; Hensel, I.V.; Anderssen, E.; Florholmen, J.; Paulssen, R.H. Genome-wide DNA Methylation in Treatment-naïve Ulcerative Colitis. *J. Crohn's Colitis* **2018**, *12*, 1338–1347. [CrossRef]
8. Lin, Z.; Hegarty, J.P.; Yu, W.; Cappel, J.A.; Chen, X.; Faber, P.W.; Wang, Y.; Poritz, L.S.; Fan, J.-B.; Koltun, W.A. Identification of Disease-Associated DNA Methylation in B Cells from Crohn's Disease and Ulcerative Colitis Patients. *Dig. Dis. Sci.* **2012**, *57*, 3145–3153. [CrossRef]
9. Karatzas, P.S.; Mantzaris, G.J.; Safioleas, M.; Gazouli, M. DNA methylation profile of genes involved in inflammation and autoimmunity in inflammatory bowel disease. *Medicine* **2014**, *93*, e309. [CrossRef]
10. Statello, L.; Guo, C.J.; Chen, L.L.; Huarte, M. Gene regulation by long non-coding RNAs and its biological functions. *Nat. Rev. Mol. Cell Biol.* **2020**, *22*, 96–118. [CrossRef]
11. Chen, S.W.; Wang, P.Y.; Liu, Y.C.; Sun, L.; Zhu, J.; Zuo, S.; Ma, J.; Li, T.Y.; Zhang, J.L.; Chen, G.W.; et al. Effect of Long Noncoding RNA H19 Overexpression on Intestinal Barrier Function and Its Potential Role in the Pathogenesis of Ulcerative Colitis. *Inflamm. Bowel Dis.* **2016**, *22*, 2582–2592. [CrossRef] [PubMed]

12. Padua, D.; Mahurkar-Joshi, S.; Law, I.K.M.; Polytarchou, C.; Vu, J.P.; Pisegna, J.R.; Shih, D.; Iliopoulos, D.; Pothoulakis, C. 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–G457. [CrossRef] [PubMed]
13. Wu, F.; Huang, Y.; Dong, F.; Kwon, J.H. Ulcerative colitis-associated long noncoding RNA, BC012900, regulates intestinal epithelial cell apoptosis. *Inflamm. Bowel Dis.* **2016**, *22*, 782–795. [CrossRef]
14. Yarani, R.; Mirza, A.H.; Kaur, S.; Pociot, F. The emerging role of lncrnas in inflammatory bowel disease. *Exp. Mol. Med.* **2018**, *50*, 1–14. [CrossRef] [PubMed]
15. Qiao, C.; Yang, L.; Wan, J.; Liu, X.; Pang, C.; You, W.; Zhao, G. Long noncoding RNA ANRIL contributes to the development of ulcerative colitis by miR-323b-5p/TLR4/MyD88/NF-κB pathway. *Biochem. Biophys. Res. Commun.* **2019**, *508*, 217–224. [CrossRef] [PubMed]
16. Rankin, C.R.; Lokhandwala, Z.A.; Huang, R.; Pekow, J.; Pothoulakis, C.; Padua, D. Linear and circular CDKN2B-AS1 expression is associated with Inflammatory Bowel Disease and participates in intestinal barrier formation. *Life Sci.* **2019**, *231*, 116571. [CrossRef] [PubMed]
17. Ray, M.K.; Fenton, C.G.; Paulssen, R.H. Novel long non-coding RNAs of relevance for ulcerative colitis pathogenesis. *Non-Coding RNA Res.* **2022**, *7*, 40–47. [CrossRef]
18. Zhao, Y.; Sun, H.; Wang, H. Long noncoding RNAs in DNA methylation: New players stepping into the old game. *Cell Biosci.* **2016**, *6*, 45. [CrossRef]
19. Bhattacharyya, N.; Pandey, V.; Bhattacharyya, M.; Dey, A. Regulatory role of long non coding RNAs (lncRNAs) in neurological disorders: From novel biomarkers to promising therapeutic strategies. *Asian J. Pharm. Sci.* **2021**, *16*, 533. [CrossRef]
20. Li, J.; Han, W.; Shen, X.; Han, S.; Ye, H.; Huang, G. DNA methylation signature of long noncoding RNA genes during human pre-implantation embryonic development. *Oncotarget* **2017**, *8*, 56829. [CrossRef]
21. Yang, Z.; Xu, F.; Teschendorff, A.E.; Zhao, Y.; Yao, L.; Li, J.; He, Y. Insights into the role of long non-coding RNAs in DNA methylation mediated transcriptional regulation. *Front. Mol. Biosci.* **2022**, *9*, 1067406. [CrossRef] [PubMed]
22. Xu, J.; Wang, Z.S.; Li, S.; Chen, J.; Zhang, J.; Jiang, C.; Zhao, Z.; Li, J.; Li, Y.; Li, X. Combinatorial epigenetic regulation of non-coding RNAs has profound effects on oncogenic pathways in breast cancer subtypes. *Brief. Bioinform.* **2018**, *19*, 52–64. [CrossRef] [PubMed]
23. Mercer, T.R.; Mattick, J.S. Structure and function of long noncoding RNAs in epigenetic regulation. *Nat. Struct. Mol. Biol.* **2013**, *20*, 300–307. [CrossRef]
24. Ranjbar, M.; Heydarzadeh, S.; Shekari Khaniani, M.; Foruzandeh, Z.; Seif, F.; Pornour, M.; Rahmanpour, D.; Tarhriz, V.; Alivand, M.R. Mutual interaction of lncRNAs and epigenetics: Focusing on cancer. *Egypt. J. Med. Hum. Genet.* **2023**, *24*, 21. [CrossRef]
25. Martin, T.D.; Chan, S.S.M.; Hart, A.R. Environmental factors in the relapse and recurrence of inflammatory bowel disease: A review of the literature. *Dig. Dis. Sci.* **2015**, *60*, 1396–1405. [CrossRef] [PubMed]
26. Karlsson, O.; Baccarelli, A.A. Environmental Health and Long Non-coding RNAs. *Curr. Environ. Heal. Rep.* **2016**, *3*, 178. [CrossRef]
27. Vance, K.W.; Ponting, C.P. Transcriptional regulatory functions of nuclear long noncoding RNAs. *Trends Genet.* **2014**, *30*, 348–355. [CrossRef]
28. Engreitz, J.M.; Pandya-Jones, A.; McDonel, P.; Shishkin, A.; Sirokman, K.; Surka, C.; Kadri, S.; Xing, J.; Goren, A.; Lander, E.S.; et al. The Xist lncRNA exploits three-dimensional genome architecture to spread across the X chromosome. *Science* **2013**, *341*, 1237973. [CrossRef]
29. Nath, A.; Lau, E.Y.T.; Lee, A.M.; Geeleher, P.; Cho, W.C.S.; Huang, R.S. Discovering long noncoding RNA predictors of anticancer drug sensitivity beyond protein-coding genes. *Proc. Natl. Acad. Sci. USA* **2019**, *116*, 22020–22029. [CrossRef]
30. Mirza, A.H.; Berthelsen, C.H.; Seemann, S.E.; Pan, X.; Frederiksen, K.S.; Vilien, M.; Gorodkin, J.; Pociot, F. Transcriptomic landscape of lncRNAs in inflammatory bowel disease. *Genome Med.* **2015**, *7*, 39. [CrossRef]
31. Chen, Z.; Hagen, D.; Ji, T.; Elsik, C.G.; Rivera, R.M. Global misregulation of genes largely uncoupled to DNA methylome epimutations characterizes a congenital overgrowth syndrome. *Sci. Rep.* **2017**, *7*, 1–14. [CrossRef]
32. Malhotra, P.; Haslett, P.; Sherry, B.; Shepp, D.H.; Barber, P.; Abshier, J.; Roy, U.; Schmidtmayerova, H. Increased Plasma Levels of the TH2 chemokine CCL18 associated with low CD4+ T cell counts in HIV-1-infected Patients with a Suppressed Viral Load. *Sci. Rep.* **2019**, *9*, 5963. [CrossRef] [PubMed]
33. Grishin, A.; Ford, H.; Wang, J.; Li, H.; Salvador-Recatala, V.; Levitan, E.S.; Zaks-Makhina, E. Attenuation of apoptosis in enterocytes by blockade of potassium channels. *Am. J. Physiol.—Gastrointest. Liver Physiol.* **2005**, *289*, G815–G821. [CrossRef] [PubMed]
34. Lan, M.; Shi, Y.; Han, Z.; Hao, Z.; Pan, Y.; Liu, N.; Guo, C.; Hong, L.; Wang, J.; Qiao, T.; et al. Expression of delayed rectifier potassium channels and their possible roles in proliferation of human gastric cancer cells. *Cancer Biol. Ther.* **2005**, *4*, 1342–1347. [CrossRef] [PubMed]
35. Farah, A.; Kabbage, M.; Atafi, S.; Gabteni, A.J.; Barbirou, M.; Madhioub, M.; Hamzaoui, L.; Mohamed, M.A.; Touinsi, H.; Kchaou, A.O.; et al. Selective expression of KCNA5 and KCNB1 genes in gastric and colorectal carcinoma. *BMC Cancer* **2020**, *20*, 1179. [CrossRef]

36. Uchiyama, K.; Naito, Y.; Takagi, T.; Mizushima, K.; Hirai, Y.; Hayashi, N.; Harusato, A.; Inoue, K.; Fukumoto, K.; Yamada, S.; et al. Serpin B1 protects colonic epithelial cell via blockage of neutrophil elastase activity and its expression is enhanced in patients with ulcerative colitis. *Am. J. Physiol. Gastrointest. Liver Physiol.* **2012**, *302*, G1163–G1170. [CrossRef]
37. Lu, J.W.; Rouzigu, A.; Teng, L.H.; Liu, W.L. The Construction and Comprehensive Analysis of Inflammation-Related ceRNA Networks and Tissue-Infiltrating Immune Cells in Ulcerative Progression. *Biomed Res. Int.* **2021**, *2021*, 1–20. [CrossRef]
38. Perry, C.; Kapur, N.; Barrett, T.A. DPP-4 as a Novel Biomarker for Inflammatory Bowel Disease: Is It Ready for Clinical Use? *Inflamm. Bowel Dis.* **2020**, *26*, 1720–1721. [CrossRef]
39. Elahi, A.; Sabui, S.; Narasappa, N.N.; Agrawal, S.; Lambrecht, N.W.; Agrawal, A.; Said, H.M. Biotin Deficiency Induces Th1- and Th17-Mediated Proinflammatory Responses in Human CD4+ T Lymphocytes via Activation of the mTOR Signaling Pathway. *J. Immunol.* **2018**, *200*, 2563–2570. [CrossRef]
40. Pindolia, K.; Li, H.; Cardwell, C.; Wolf, B. Characterization and functional analysis of cellular immunity in mice with biotinidase deficiency. *Mol. Genet. Metab.* **2014**, *112*, 49–56. [CrossRef]
41. García-Heredia, J.M.; Carnero, A. The cargo protein MAP17 (PDZK1IP1) regulates the immune microenvironment. *Oncotarget* **2017**, *8*, 98580. [CrossRef] [PubMed]
42. Taman, H.; Fenton, C.G.; Hensel, I.V.; Anderssen, E.; Florholmen, J.; Paulssen, R.H. Transcriptomic landscape of treatment-naïve ulcerative colitis. *J. Crohn's Colitis* **2018**, *12*, 327–336. [CrossRef] [PubMed]
43. Gu, Y.; Zhao, H.; Zheng, L.; Zhou, C.; Han, Y.; Wu, A.; Jia, Z.; Xia, T.; Zhi, Q. Non-coding RNAs and colitis-associated cancer: Mechanisms and clinical applications. *Clin. Transl. Med.* **2023**, *13*, e1253. [CrossRef] [PubMed]
44. Soroosh, A.; Fang, K.; Hoffman, J.M.; Law, I.K.M.; Videlock, E.; Lokhandwala, Z.A.; Zhao, J.J.; Hamidi, S.; Padua, D.M.; Frey, M.R.; et al. Loss of miR-24-3p promotes epithelial cell apoptosis and impairs the recovery from intestinal inflammation. *Cell Death Dis.* **2022**, *13*, 8. [CrossRef] [PubMed]
45. Chapman, R. Genome-wide association studies in primary sclerosing cholangitis: Still more questions than answers? *Hepatology* **2011**, *53*, 2133–2135. [CrossRef] [PubMed]
46. Emam, O.; Wasfey, E.F.; Hamdy, N.M. Notch-associated lncRNAs profiling circuiting epigenetic modification in colorectal cancer. *Cancer Cell Int.* **2022**, *22*, 316. [CrossRef]
47. Abbasi, N.; Long, T.; Li, Y.; Yee, B.A.; Cho, B.S.; Hernandez, J.E.; Ma, E.; Patel, P.R.; Sahoo, D.; Sayed, I.M.; et al. DDX5 promotes oncogene C3 and FABP1 expressions and drives intestinal inflammation and tumorigenesis. *Life Sci. Alliance* **2020**, *3*, e202000772. [CrossRef]
48. Hotamisligil, G.S.; Bernlohr, D.A. Metabolic functions of FABPs—Mechanisms and therapeutic implications. *Nat. Rev. Endocrinol.* **2015**, *11*, 592–605. [CrossRef]
49. Heimerl, S.; Moehle, C.; Zahn, A.; Boettcher, A.; Stremmel, W.; Langmann, T.; Schmitz, G. Alterations in intestinal fatty acid metabolism in inflammatory bowel disease. *Biochim. Biophys. Acta* **2006**, *1762*, 341–350. [CrossRef]
50. Rahmani, E.; Schweiger, R.; Rhead, B.; Criswell, L.A.; Barcellos, L.F.; Eskin, E.; Rosset, S.; Sankararaman, S.; Halperin, E. Cell-type-specific resolution epigenetics without the need for cell sorting or single-cell biology. *Nat. Commun.* **2019**, *10*, 3417. [CrossRef]
51. Blander, J.M. Death in the intestinal epithelium—Basic biology and implications for inflammatory bowel disease. *FEBS J.* **2016**, *283*, 2720–2730. [CrossRef] [PubMed]
52. Ding, J.; Xie, M.; Lian, Y.; Zhu, Y.; Peng, P.; Wang, J.; Wang, L.; Wang, K. Long noncoding RNA HOXA-AS2 represses P21 and KLF2 expression transcription by binding with EZH2, LSD1 in colorectal cancer. *Oncogenesis* **2017**, *6*, e288. [CrossRef] [PubMed]
53. Yao, Q.; Wang, C.; Wang, Y.; Zhang, X.; Jiang, H.; Chen, D. The integrated comprehension of lncRNA HOXA-AS3 implication on human diseases. *Clin. Transl. Oncol.* **2022**, *24*, 2342. [CrossRef]
54. Lee, S.H.; Kwon, J.E.; Cho, M. La Immunological pathogenesis of inflammatory bowel disease. *Intest. Res.* **2018**, *16*, 26. [CrossRef] [PubMed]
55. Chini, A.; Guha, P.; Malladi, V.S.; Guo, Z.; Mandal, S.S. Novel long non-coding RNAs associated with inflammation and macrophage activation in human. *Sci. Rep.* **2023**, *13*, 4036. [CrossRef] [PubMed]
56. Yi, E.; Zhang, J.; Zheng, M.; Zhang, Y.; Liang, C.; Hao, B.; Hong, W.; Lin, B.; Pu, J.; Lin, Z.; et al. Long noncoding RNA IL6-AS1 is highly expressed in chronic obstructive pulmonary disease and is associated with interleukin 6 by targeting miR-149-5p and early B-cell factor 1. *Clin. Transl. Med.* **2021**, *11*, 11. [CrossRef]
57. Magro, F.; Gionchetti, P.; Eliakim, R.; Ardizzone, S.; Armuzzi, A.; Barreiro-de Acosta, M.; Burisch, J.; Gecse, K.B.; Hart, A.L.; Hindryckx, P.; 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*, 649–670. [CrossRef]
58. Lichtenstein, G.R.; Travis, S.; Danese, S.; D'Haens, G.; Moro, L.; Jones, R.; Huang, M.; Ballard, E.D.; Bagin, R.; Hardiman, Y.; et al. Budesonide MMX for the Induction of Remission of Mild to Moderate Ulcerative Colitis: A Pooled Safety Analysis. *J. Crohn's Colitis* **2015**, *9*, 738. [CrossRef]