Title: Crypt-top and crypt-bottom colonic epithelial cell microRNA profiling reveals cell type-specific response in active and quiescent ulcerative colitis

Authors: MSc Ruta Inciuraite1*, PhD Rima Ramonaite1*, PhD Juozas Kupcinskas1,2*, PhD Indre Dalgediene3, PhD Ugne Kulokiene1, MSc Vytautas Kiudelis1,2, PhD Greta Varkalaite1, PhD Aurelija Zvirbliene3, PhD Laimas Virginijus Jonaitis1,2, PhD Gediminas Kiudelis1,2, PhD Andre Franke4, PhD Simonas Juzenas1,3*, PhD Jurgita Skieceviciene1*

1 Institute for Digestive Research, Academy of Medicine, Lithuanian University of Health Sciences, A. Mickeviciaus st. 9, 44307, Kaunas, Lithuania
2 Department of Gastroenterology, Academy of Medicine, Lithuanian University of Health Sciences, A. Mickeviciaus st. 9, 44307, Kaunas, Lithuania
3 Institute of Biotechnology, Life Sciences Center, Vilnius University, Sauletekio av. 7, 10257, Vilnius, Lithuania
4 Institute of Clinical Molecular Biology, Christian-Albrechts-University of Kiel, Rosalind-Franklin St. 12, 24105, Kiel, Germany

* These authors contributed equally to this work

Correspondence: Jurgita Skieceviciene, PhD, Institute for Digestive Research, Academy of Medicine, Lithuanian University of Health Sciences, Mickeviciaus st. 9, 44307, Kaunas, Lithuania. Email: jurgita.skieceviciene@lsmuni.lt, phone: (370 37) 327236.

Funding: This work was supported by the Research Council of Lithuania and European Crohn’s and Colitis Organisation (grant numbers S-MIP-20-56 and ECCO Grant 2016, respectively).

Declaration of interests: The authors declare no conflict of interest.

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.
Summary: We present microRNA transcriptome analysis on different levels - colon tissue and epithelial cell population - in ulcerative colitis. Results indicate cell type- and disease stage-dependent microRNA deregulation, unveil associations with disease activity, and putative biological role of deregulated microRNAs.
Abstract and Key Words

**Background.** Colonic epithelial cells form a frontline intestinal barrier and maintain its function which deteriorates early in ulcerative colitis (UC). MicroRNAs (miRNAs) participate in regulation of intestinal epithelial integrity and barrier permeability. However, there is a lack of understanding about cell type-specific expression of miRNAs in UC.

**Methods.** Two independent cohorts composed of active and quiescent UC patients (n=74), and healthy controls (HC; n=50) were studied. Crypt-bottom (CD44+) and crypt-top (CD66a+) colonic epithelial cell populations were enriched using FACS. Small RNA-sequencing was performed on colon biopsy and colonic epithelial cell population samples. Data processing encompassed differential expression, gene-set enrichment analysis (GSEA) and clinical correlation analysis.

**Results.** We describe differentially expressed miRNAs among active and quiescent UC compared to HC colon tissue and propose their involvement in intestinal barrier integrity regulation. We further focus on crypt-bottom and crypt-top colonic epithelial cells and characterise common and cell population-specific miRNA expression in response to UC-caused inflammation. We suggest that differentially expressed miRNAs are commonly involved in inflammation- and intestinal barrier integrity-related processes (such as signalling of interleukin-4 and interleukin-13), while differences between cell populations might reflect their function, i.e., crypt-bottom cell miRNA target genes are enriched in regulation of cell differentiation. Moreover, we show cell population-specific miRNA expression correlations with endoscopic disease activity, i.e., let-7b-5p and let-7e-5p negatively correlate with activity score only in the crypt-bottom cells, while miR-24-3p and miR-27a-3p positively correlates only in the crypt-top cells.

**Conclusions.** Changes in miRNA expression during UC are epithelial cell type- and UC activity-specific (including correlations with endoscopic Mayo score). Further, irrespective of the UC stage and colonic cell population, deregulated miRNAs are potentially involved in signalling pathways responsible for regulation of intestinal barrier integrity and permeability.
Keywords. Colon crypt-bottom (CD44+) epithelial cells; colon crypt-top (CD66a+) epithelial cells; cell populations; microRNA; Ulcerative colitis.

Key Messages

What is already known? Colonic epithelium plays an important role in pathogenesis of ulcerative colitis (UC), while microRNAs (miRNA) have been implicated in modulation of intestinal homeostasis.

What is new here? Cell type-specific miRNA expression of colonic epithelial cells during UC was unknown. Here, we show cell population-specific miRNA expression in response to UC-caused inflammation in crypt-top and crypt-bottom colonic cells.

How can this study help patient care? The identified cell type-specific correlations of miRNA expression and endoscopic Mayo score might be further evaluated in UC monitoring and diagnostics as well as selected as targets for further therapeutics development.

Graphical abstract
Introduction

Ulcerative colitis (UC) is a chronic remitting/relapsing inflammatory disease affecting mucosa of the lower gastrointestinal tract, including colon and rectum, and can be classified as active or in remission (quiescent). Although development of UC is known to be a multifactorial process involving genetic susceptibility, deregulated immune responses, epithelial barrier defects, and environmental factors, the etiopathogenesis of UC remains still not fully understood. In terms of cell populations, besides diverse types of immune cells, intestinal epithelial cells also play a major role in the onset and course of UC. Colonic epithelial cells and their secreted products primarily form and maintain frontline intestinal barrier function which deteriorates early during disease. Colonic crypt epithelial cells such as colonocytes, Goblet and epithelial stem cells have been shown to play an important role in UC pathogenesis via reduced epithelial mucus secretion and/or increased barrier permeability.

Since one of the aims in UC management is to induce and then to maintain remission, disentangling molecular mechanisms regulating epithelial barrier might help to artificially modulate its permeability and retain long-lasting remission of UC.

Recent studies provide evidence for microRNAs (miRNAs) being modulators of intestinal homeostasis by calibrating immune cell responses and regulating autophagy. Furthermore, these small non-coding RNA molecules were shown to be involved in regulation of intestinal epithelial integrity and barrier permeability via interference with protein-coding genes responsible for tight and adherens junctions, also in UC. However, most of the previous studies were either based on immortalized cell cultures or bulk tissue experiments, thus the exact cellular context of miRNA deregulation during colonic inflammation remains elusive.

Therefore, the search for cell type specific deregulation of miRNAs in UC might uncover novel aspects of underlying regulatory pathways and molecular targets for therapeutic development.

Here, we report the very first study of cell type-specific miRNA expression profiles of colonic crypt-bottom (CD44+) and crypt-top (CD66a+) cell populations in active and quiescent UC.
Initially, for hypothesis generation and cell type population selection, we apply small RNA-seq on inflamed and non-inflamed colonic biopsies from UC patients and healthy controls. Further, we use fluorescence-activated cell sorting (FACS) to select crypt-bottom (CD44+) and crypt-top (CD66a+) cell populations from active and quiescent UC and healthy controls. Then, by employing small RNA-seq we determine UC activity-specific, as well as cell type-specific, signatures in miRNA expression during colonic inflammation. In addition, we describe putative biological pathways in which the deregulated miRNAs are involved and provide cell population-specific miRNA expression correlations with disease activity score.
Materials and Methods

Study samples

Study subject recruitment was conducted at the Department of Gastroenterology, Lithuanian University of Health Sciences (Kaunas, Lithuania) during the period of 2011-2014 (Study group I and III) and 2017-2019 (Study group II). For miRNA sequencing, colon biopsies were collected in two independent groups: the first study group contained 23 patients with active UC, 20 patients with quiescent UC and 33 healthy control individuals (HC), while the second study group was composed of 16 patients with active UC, 15 patients with quiescent UC and 17 HC individuals. Additionally, colon biopsies for qPCR-based gene expression analysis were collected in the third independent group which contained 75 patients with active UC, 50 patients with quiescent UC and 75 HC individuals. Demographic and clinical characteristics of subjects of Study groups I and II are presented in Table 1, Study group III – in Supplementary Table S1. Tissue samples of patients with active UC and quiescent UC were collected from sigmoid colon. Endoscopic activity was determined using the Mayo endoscopic subscore. Quiescent UC was defined as mucosal healing (endoscopic Mayo subscore ≤ 1) with no clinical symptoms. The variables collected for the individuals of both study groups were such as age, sex, body mass index (BMI), smoking status, disease activity, localization, rectal bleeding, stool frequency, endoscopic Mayo subscore, global assessment score, full Mayo score.

All subjects with active and quiescent UC had a routine colonoscopy performed as a part of their planned examination programme. The diagnosis of UC was based on standard clinical, endoscopic, radiological, and histological criteria. The control individuals included in both study groups consisted of healthy subjects who either underwent colonoscopy due to positive faecal occult blood test or were consulted on functional complaints, but had a normal colonoscopy, uninflamed mucosa during histopathological examination and no previous history of intestinal inflammation. All patients enrolled in the study were of European descent.
**Table 1.** Demographic and clinical characteristics of subjects

<table>
<thead>
<tr>
<th></th>
<th>Study group I, n=76</th>
<th>Study group II, n=48</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Active UC, n=23</td>
<td>Quiescent UC, n=20</td>
</tr>
<tr>
<td><strong>Age</strong></td>
<td></td>
<td>HC, n=33</td>
</tr>
<tr>
<td>Mean ± SD</td>
<td>43.9 ± 16.1</td>
<td>47.2 ± 13.5</td>
</tr>
<tr>
<td></td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>SD</td>
<td>58.5 ± 12.7</td>
<td>57.3 ± 12.1</td>
</tr>
<tr>
<td><strong>Sex, n (%)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>10 (43.5)</td>
<td>9 (45.0)</td>
</tr>
<tr>
<td></td>
<td>21 (73.6)</td>
<td></td>
</tr>
<tr>
<td><strong>Full Mayo score, n (%)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Remission</td>
<td>0 (0.0)</td>
<td>20 (100.0)</td>
</tr>
<tr>
<td><strong>Mild</strong></td>
<td>1 (4.4)</td>
<td>0 (0.0)</td>
</tr>
<tr>
<td><strong>Moderate</strong></td>
<td>21 (91.3)</td>
<td>0 (0.0)</td>
</tr>
<tr>
<td><strong>Severe</strong></td>
<td>1 (4.4)</td>
<td>0 (0.0)</td>
</tr>
<tr>
<td><strong>Endoscopic Mayo subscore, n (%)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Normal to mild</td>
<td>0 (0.0)</td>
<td>33 (100.0)</td>
</tr>
<tr>
<td>Moderate to severe</td>
<td>23 (100.0)</td>
<td>0 (0.0)</td>
</tr>
<tr>
<td><strong>Smoking status, n (%)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Never</td>
<td>8 (34.8)</td>
<td>11 (55.0)</td>
</tr>
<tr>
<td>Smoking</td>
<td>8 (34.8)</td>
<td>6 (30.0)</td>
</tr>
<tr>
<td>Ex-smoker</td>
<td>3 (13.1)</td>
<td>2 (10.0)</td>
</tr>
<tr>
<td>Unknown</td>
<td>4 (17.4)</td>
<td>1 (5.0)</td>
</tr>
<tr>
<td><strong>BMI, kg/m^2</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Median, (range)</td>
<td>26.8 (18.1–56.2)</td>
<td>26.2 (19.7–31.5)</td>
</tr>
<tr>
<td></td>
<td>27.7 (19.6–39.5)</td>
<td>25.5 (18.0–37.2)</td>
</tr>
<tr>
<td>Unknown, n (%)</td>
<td>1 (4.4)</td>
<td>1 (5.0)</td>
</tr>
</tbody>
</table>

* Age differed between both active and quiescent UC groups, and HC, estimated \(P\) value < 0.05 (\(t\)-test)
**Age differed between both active and quiescent UC groups, and HC, estimated \( P \) value < 0.05 (\( t \)-test)

***Sex distribution differed between three groups, estimated \( P \) value < 0.05 (Pearson's chi-squared test)

SD - Standard deviation

Colon tissue disaggregation

Colon biopsies were mechanically and enzymatically separated into single-cell suspensions. Briefly, four to six biopsies (5–10 mg wet weight each) were washed with PBS solution containing 50 IU/mL penicillin, 50 µg/mL streptomycin, and 0.5 mg/mL gentamicin. Biopsies were minced into small pieces (approx. 1-2 mm\(^3\)) and incubated in 1X trypsin-EDTA solution for 40-45 min at room temperature on an agitator to dissociate single epithelial cells from the lamina propria. After incubation, tissue fragments were gently transferred into PBS and shaken. The isolated cell suspension was filtered and suspended in DMEM:Ham's F-12 medium (1:1) with 15 mM HEPES (Gibco, USA) buffer for flow cytometry procedures.

Flow cytometry and fluorescence activated cell sorting

In order to minimize the loss of cell viability, fresh cell suspensions prepared shortly before flow cytometry were used in all cell sorting experiments. Antibody staining was performed in PBS supplemented with 1% heat-inactivated fetal bovine serum. To minimize nonspecific binding of antibodies, cells were first incubated with Human TruStain FcX™ (Fc Receptor Blocking Solution) (BioLegend Inc., USA) for 10 min at a concentration of 3-10 \( \times \) 10\(^5\) cells/100 µl. Cells were subsequently stained without washing with antibodies at dilutions recommended by manufacturers. Antibodies used in this study were selected based on previous study\(^1\) and included: mouse anti-human CD326/EpCAM-FITC (clone VU-1D9, RRID: AB_2534535m Life Technologies, USA), mouse anti-human CD44-APC (clone G44-26, RRID: AB_395868, BD Biosciences, USA), mouse anti-human CD66a-PE (clone 283340, R&D Systems, USA). Cells positive for expression of non-epithelial lineage markers
were excluded by staining with mouse anti-human CD45-APC-Cy7 (clone 2D1, RRID: AB_2566375, Biolegend Inc., USA) (see gating strategy in Supplementary Figure S1a).

After 20 min, stained cells were washed of excess unbound antibodies and resuspended in PBS supplemented with 1% heat-inactivated fetal bovine serum. Flow-cytometry analysis was performed using a CyFlow® Space cell sorter (Sysmex Partec, Germany) (see representative plots of flow cytometry data in Supplementary Figure S1b-d). In cell sorting experiments, each population of interest was sorted individually, using a protocol already built-in within the CyFlow® Space flow cytometer software package (FloMax® 2.8), with appropriate adjustments. Data was analysed and visualised using FlowJo™ v10.7 (BD FlowJo, USA). Sorted cell samples were frozen and stored at -70 °C prior to total RNA extraction.

**Total RNA extraction**

Total RNA extraction from colon biopsies and sorted cells was performed using standard protocols of commercial miRNeasy Mini Kit (Qiagen, Germany) and Single Cell RNA Purification Kit (Norgen, Canada), respectively. Total RNA concentration was evaluated by NanoDrop2000 spectrophotometer (Thermo Scientific, USA) and Qubit 4 fluorometer (Invitrogen, USA). Total RNA quality was assessed using Agilent 2100 Bioanalyzer (Agilent Biotechnologies, USA).

**Quantitative Reverse Transcription PCR and data analysis**

To estimate the expression of *IL-4* and *IL-13* genes in colon tissue of UC patients, total RNA from colon tissue samples was reverse transcribed using a High-Capacity cDNA Reverse Transcription Kit (Applied Biosystems™, USA). Further, expression levels were measured using TaqMan™ Gene Expression Assays (Assay IDs: *IL-4* Hs00174122_m1; *IL-13* Hs00174379_m1) on the 7500 Fast Real-Time PCR System (Applied Biosystems™, USA). The cycle threshold (C_T) values of *IL-4* and *IL-13* were normalized to the value of *GAPDH* (Assay ID: Hs99999905_m1) reference gene. All the procedures were performed in
accordance with the manufacturer’s protocol. Statistical analysis was performed using R
studio software (version 4.0.3). Data distribution was determined using the Shapiro-Wilk test,
gene expression differences were analysed the two-sided Mann-Whitney U test. The
difference between the values was considered significant when $P < 0.05$.

**Preparation of small RNA libraries and next-generation sequencing**

Small RNA libraries from tissue samples were prepared using TruSeq Small RNA Sample
Preparation Kit (Illumina, USA) with 1 μg of total RNA input per sample. Small RNA libraries
from sorted cells were prepared using NEXTFLEX Small RNA-seq Kit v3 (Bioo Scientific,
USA) with up to 50 ng of total RNA input per sample. Procedures were conducted according
to the manufacturers’ protocols. The yield of sequencing libraries was assessed using the
Agilent 2200 TapeStation system (Agilent Biotechnologies, USA). Subsequently, the TruSeq
libraries were pooled with around 24 samples per lane, while the NEXTFLEX libraries were
pooled with around 16 samples per lane and then sequenced on HiSeq 4000 (Illumina, USA)
next-generation sequencing platform.

**Bioinformatics and statistical analysis of small RNA-seq data**

The demultiplexed raw reads (.fastq) were processed with nf-core/smrnaseq v1.0.0 best-
practise analysis pipeline using default parameters (Nextflow version v20.01.0). The pipeline
was executed within a Docker container. Briefly, depending on the small RNA-seq library
preparation kits, “illumina” or “nextflex” protocol was selected for processing of the libraries
generated from tissue and sorted cell samples, respectively. First, Trim Galore (v0.6.3) was
used to remove 3’ adapter (5’ - TGAATTCTCGGGTGCCAAGG - 3’ for both “illumina” and
“nextflex” protocols) sequences from the reads and additional 4 nucleotides from both 5’ and
3’ ends of reads (only for “nextflex” protocol). Second, to reduce computational time, the
reads with identical sequences were collapsed using seqcluster while saving the information
about the read counts. Third, Bowtie1 v1.2.2 was used to perform the alignment of collapsed
reads against mature and hairpin miRNA sequences in miRbase database v22.1. Finally,
miRNA annotation was performed using the mirTOP v0.4.23. Further, sample and miRNA QC was performed: samples with initial read count < 1.5 IQR and number of detected miRNAs < 1.5 IQR on log₂ scale as well as non-expressed (mean raw < 1) and non-variable miRNAs were excluded from further analysis. Subsequently, to perform differential expression analyses of the size factor normalized counts of mature miRNAs between samples, negative binomial generalized linear models implemented in the R package DESeq2 were used including age (scaled and centred) and sex as covariates in the model. The P-values resulting from Wald tests were corrected for false discovery rate (FDR) according to Benjamini and Hochberg. The miRNAs with FDR < 0.05 and absolute value of log₂ FC > 1 were considered to be significantly differentially expressed. A multidimensional scaling (MDS) analysis using Euclidean distance was performed on variance stabilizing transformation (VST) normalized miRNA count data. Additionally, Spearman’s rank correlation analysis was performed between sex- and age-adjusted normalized miRNA read counts and endoscopic Mayo subscore. FDR < 0.05 was considered statistically significant. Removal of sex and age effects from normalized was performed using removeBatchEffect function from the limma R package. Statistical analyses and data processing were performed using R version 4.0.3. Visualization of graphs was performed using the ggplot2 package.

Gene set enrichment analysis

In order to obtain putative biological functions of differentially expressed miRNAs, gene set enrichment analysis (GSEA) was performed using Reactome pathways and GO categories. More specifically, luciferase assay-validated miRNA-target interactions (MTIs) were obtained from miRecords, miRTarBase and TarBase using the multiMiR package. The retrieved MTIs were then submitted to a hypergeometric test implemented in the enrichPathway (from the ReactomePA package) and enrichGO (from the clusterProfiler package) functions using genes that are expressed in colon crypt-bottom (CD44⁺) and crypt-top (CD66a⁺) cells as a background reference (defined as universe). The universe genes were obtained from single-
cell RNA-seq data of the human colon, available in GEO database with accession number GSE116222. The pathways with FDR < 0.05 were considered to be significantly deregulated.

Data Availability
The small RNA-seq data underlying this article will be available in Gene Expression Omnibus (GEO) Database and will be open upon publishing with accession numbers GSE185101 and GSE185102.
Results

Differentially expressed miRNAs in active and quiescent UC tissues are involved in regulation of inflammation-related pathways

To identify differentially expressed miRNAs and their putative regulatory processes during chronic colon inflammation, small RNA-seq was performed on inflamed (active) and non-inflamed (quiescent) colonic mucosal biopsies of UC patients and healthy controls (HC) (Figure 1a).

Figure 1: Small RNA-seq defines differentially expressed miRNAs involved in inflammation-associated pathways in active and quiescent UC tissues. (a) Design of the initial study phase. (b)
MDS plot showing the similarity structure of the miRNA transcriptomes in active (aUC) (n=23),
quiescent UC (qUC) (n=20) and HC (n=30) tissue samples based on normalized expression values.
The dots represent samples coloured by group. The centroid of ellipses corresponds to the group
mean, the shapes are defined by covariance within a given group. (c) Differentially expressed
miRNAs in active (aUC) (n=23) and quiescent UC (qUC) (n=20). The red colour represents
significantly (FDR < 0.05) differentially expressed miRNAs with an absolute value of log₂FC > 1, while
the blue colour represents non-differentially expressed miRNAs. (d) Top 10 overrepresented
pathways in active (aUC) (n=23) and quiescent UC (qUC) (n=20) tissues identified by miRNA set
enrichment analysis. Dot size represents the number of miRNA-target gene count in the significantly
enriched (FDR < 0.05) Reactome pathways.

After count data normalization and quality control, 573 unique miRNAs were found to be
expressed in colon tissue samples. The overall similarity structure (based on MDS analysis;
see Methods) of colon miRNA transcriptomes revealed two clearly resolved clusters
corresponding to active UC and HC tissues, while the third cluster corresponding to
quiescent UC overlapped with both active UC and HC clusters, suggesting a shift of miRNA
expression from healthy to inflammatory state (Figure 1b). To evaluate expression of
specific miRNAs in active UC and quiescent UC, differential gene expression analysis was
performed. As expected, the most profound miRNA deregulation was observed comparing
active UC to HC or to quiescent UC (93 and 59 differentially expressed miRNAs [FDR < 0.05
and |log₂FC| > 1], respectively). Interestingly, although substantially lower than in active UC,
a differential expression of miRNAs (n=32) was also observed in quiescent UC compared to
HC (Figure 1c; Supplementary Table S2). Among the differentially expressed miRNAs, a
considerable number (n=13) of molecules, including miR-106-5p, miR-125b-1-3p, miR-205-
5p and miR-3182 were deregulated in both active and quiescent UC compared to HC.
Contrarily, majority of miRNAs (n=80) including, miR-190a-5p, miR-3168, miR-378i, and
miR-223-3p were deregulated only in active UC, while miRNAs (n=19) such as miR-331-3p,
miR-409-5p, miR-4497 and miR-629-5p - only in the quiescent UC. In addition, the gradual
decrease in miR-1-3p expression was found in all pairwise comparisons (active UC vs quiescent UC, active UC vs HC and quiescent UC vs HC; Supplementary Table S2). Thus, besides shared miRNA deregulation patterns, the data also suggests unique and perhaps non-overlapping functions of some miRNAs in active and quiescent UC.

To further determine the biological function of differentially expressed miRNAs in UC pathogenesis, GSEA was performed for each pairwise comparison (active UC vs HC; quiescent UC vs HC; active vs quiescent UC) using validated target genes of significantly deregulated miRNAs and Reactome pathways (see Methods). The analysis indicated similarities and differences in the over-represented pathways among the active UC, quiescent UC and HC comparisons (Supplementary Table S3). Intriguingly, both active UC and quiescent UC compared to HC had over-represented interleukin signalling-related pathways among the top significant ones, such as “Signaling by Interleukins” [R-HSA-449147] and “Interleukin-4 and Interleukin-13 signaling” [R-HSA-6785807] (Figure 1d). This similar enrichment in the interleukin signalling-related target genes suggested that these pathways remain dysregulated in the colonic mucosa of quiescent UC.

To assess if this deregulation can be observed not only in silico, but also in vivo, the expression of genes coding two main cytokines IL-4 and IL-13 was evaluated in colon tissue. The gradual increase in IL-13 expression was observed between UC patients and HC individuals (2.19-fold [P = 0.031] increase in quiescent UC vs. HC, 2.91-fold [P = 0.0007] increase in active UC vs. quiescent UC, 6.38-fold [P = 2 × 10^{-10}] increase in active UC vs. HC), whereas the expression of IL-4 did not differ between the groups (Supplementary Figure S2). These findings fall in concordance to GSEA results and confirm the changes in interleukin signalling during different stages of UC.

Taken together, differential miRNA expression and GSEA analyses in UC tissues suggest that dysregulated miRNAs are likely involved in regulation of inflammation-related pathways and that some of the pathways, i.e., interleukin-4 and interleukin-13 signalling, remarkably, remain dysregulated in the colonic mucosa of quiescent UC.
Sequencing of FACS-sorted colonic epithelial cells shows cell type-specific miRNA expression during colonic inflammation

The role of interleukin-4 and interleukin-13 in mediating permeability of epithelial barrier as well as its relation to the pathogenesis of UC has been previously described\textsuperscript{13}. Given that this pathway is deregulated in active UC tissues and remains deregulated in quiescent UC (based on miRNA and mRNA expression levels in our study) and the fact that it affects epithelial barrier function, we further decided to focus on miRNA expression in crypt-bottom (CD44\textsuperscript{+}) and crypt-top (CD66a\textsuperscript{+}) colonic epithelial cells.

FACS was applied to enrich and select for crypt-bottom and crypt-top colonic epithelial cells from active and quiescent UC patients and healthy controls using CD44\textsuperscript{+} and CD66a\textsuperscript{+} surface markers, respectively (Supplementary Figure S1). Analysis of FACS enriched cell populations over single-cell analysis was chosen due to limitations of single-cell transcriptome sequencing technology, which is not yet compatible with small RNA-seq. Interestingly, analysis of flow cytometry showed significant (FDR < 0.05) increase in crypt-bottom (CD44\textsuperscript{+}) cells in active UC compared to HC (Figure 2a), suggesting a potential inflammation-stimulated cell proliferation\textsuperscript{14,15}. 
Figure 2: Sequencing data reveals differentially expressed miRNAs in colonic epithelial cells of patients with active and quiescent UC. (a) FACS of colonic epithelial cells and distribution of crypt-top CD44+CD66+ and crypt-bottom CD44+CD66+ epithelial cell types in inflammatory (aUC) (n=16) and non-inflammatory (qUC and HC) (n=15 and 17) colon tissues. Each dot represents a sample from the patients of the second study group. Mean ±SD of each group is represented by vertical lines. To compare the groups a nonparametric Mann–Whitney U test was performed, p* < 0.05. (b) MDS plot showing the similarity structure of the miRNA transcriptomes in crypt-top (CD66a+) and crypt-bottom (CD44+) colonic epithelial cell populations in active (aUC) (n=16), quiescent UC (qUC) (n=15), and HC (n=17) based on normalized expression values. The dots represent samples shaped by cell population. Dot colours represent condition. The centroid of ellipses corresponds to the condition group mean, the shapes are defined by covariance within the group. (c, d) Volcano plots of differentially expressed miRNAs in crypt-top (CD66a+) and crypt-bottom (CD44+) colonic epithelial cell populations in active (aUC) (n=16), quiescent UC (qUC) (n=15), and HC (n=17). Colours indicate significantly (FDR < 0.05) differentially expressed miRNAs with an absolute value of log2 FC > 1 between compared groups. (e-g) Venn diagrams representing the numbers of commonly and uniquely differentially expressed miRNAs in (e) crypt-bottom (CD44+) and (f) crypt-top (CD66a+) epithelial cell
populations in different UC activity, and (g) between crypt-bottom and crypt-top cells in the same condition.

Using sequencing, a total number of 436 unique miRNAs were found to be expressed in crypt-bottom (CD44+ and crypt-top (CD66a+ epithelial cells. Although miRNA transcriptomes of these colonic epithelial cell populations were strongly overlapping (Figure 2b), significant changes in expression profiles were observed within and between cell populations in different stages of the disease (Figure 2c and 2d; Supplementary Table S4). Initially, pairwise comparisons were performed in the same epithelial cell population to identify UC inflammation-associated miRNAs. As in the tissue data, the number of differentially expressed miRNAs (FDR < 0.05 and |log2FC| > 1) in both colonic epithelial cell populations were gradually increased depending on the disease activity stage, i.e., the highest number of differentially expressed miRNAs were observed in active UC compared to HC and to quiescent UC (Figure 2c). Among deregulated molecules, none of miRNAs were commonly differentially expressed across all three comparisons (active UC vs HC; quiescent UC vs HC; active UC vs quiescent UC) in both crypt-bottom (CD44+) and crypt-top (CD66a+) colonic epithelial cells. However, six miRNAs (miR-15b-5p, miR-222-3p, miR-223-3p, miR-194-3p, miR-3195 and miR-574-3p) were identified as commonly differentially expressed in crypt-bottom (CD44+) cells and eight miRNAs (let-7c-5p, miR-106b-3p, miR-125b-5p, miR-1-3p, miR-1290, miR-194-3p, miR-335-5p, miR-552-3p) in crypt-top (CD66a+) cells in the active and quiescent UC when compared to HC (Figure 2e and 2f). Further, the response of distinct epithelial cell populations to inflammation was determined by performing pairwise comparisons with the separate populations of colonic epithelial cells. Twenty-four miRNAs (such as miR-501-3p, miR-1-3p, miR-296-5p, and miR-122-5p) were identified to be differentially expressed in active UC, nine miRNAs (such as miR-1290, miR-3168, and miR-660-5p) in quiescent UC, and twenty-two miRNAs (such as miR-598-3p, miR-340-3p, and miR-223-3p) - HC when compared crypt-bottom (CD44+) and crypt-top (CD66a+) colonic epithelial cells (Figure 2d, Supplementary Table S4). Importantly, function of these miRNAs...
miRNAs might represent a different cellular response to active inflammation in UC, e.g., inflammation-induced cell proliferation or apoptosis. For example, in CD44+ cells over-expressed miR-501-3p has been previously shown to increase cell proliferation in colorectal cancer\textsuperscript{16}. Notably, the vast majority of identified differentially expressed miRNAs between distinct epithelial cell types in different stages of inflammation were found to be uniquely dysregulated, only two commonly differentially expressed miRNAs were observed between two comparison groups (miR-106b-3p and miR-1290 in active UC CD44+ vs CD66a+ and quiescent UC CD44+ vs CD66a+; miR-296-5p and miR-432-5p active UC CD44+ vs CD66a+ and HC CD44+ vs CD66a+) (Figure 2g), which suggests that even at different disease activity stages some cell population-specific responses, in terms of miRNA expression, are different.

Collectively, the results show some colonic cell population-specific miRNA expression patterns in UC and at different stages of disease activity.

Aberrantly expressed miRNAs in crypt-top (CD66a+) and crypt-bottom (CD44+) colonic epithelial cells are involved in regulation of intestinal epithelial barrier function-related processes

To further evaluate the potential functional role of differentially expressed miRNAs in crypt-bottom (CD44+) and crypt-top (CD66a+) colonic epithelial cells in UC pathogenesis, GSEA was performed on validated target genes of differentially expressed miRNAs (Supplementary Table S5). Similarly to the colonic tissue data (Figure 1b), majority of the most overrepresented pathways in both colonic epithelial cell populations in both stages of disease activity overlapped and included signalling pathways such as “Signaling by Interleukins” (R-HSA-449147), “Interleukin-4 and Interleukin-13 signaling” [R-HSA-6785807], “Signaling by Receptor Tyrosine Kinases” [R-HSA-9006934] (Figure 3a). This supports that observations in the colon biopsy samples were mainly driven by colonic epithelial cells. Notably, not all comparison groups in crypt-bottom (CD44+) cells showed uniform results. The most overrepresented Reactome pathways in crypt-bottom (CD44+) cells of quiescent
UC patient group differed from those in active UC groups and exclusively included “Signaling by Nuclear Receptors” (R-HSA-9006931), “Extra-nuclear estrogen signaling” [R-HSA-9009391] pathways which were shown to regulate the permeability of intestinal epithelium and thus be associated with pathogenesis of inflammatory bowel disease\(^1\) (Figure 3a). Finally, to assess which biological processes (from Gene Ontology terms) are overrepresented in crypt-bottom (CD44\(^+\)) compared to crypt-top (CD66a\(^+\)) cells, GSEA of deregulated miRNAs in active UC, quiescent UC, and HC groups was performed. The results revealed the overrepresented processes between the cell populations were mainly related to cell differentiation and motility in both active UC and HC (Figure 3b), suggesting that in inflamed and healthy colon mucosa these pathways are differentially regulated between the cell types. In addition, significant enrichment between crypt-bottom (CD44\(^+\)) and crypt-top (CD66a\(^+\)) cells in “epithelium migration” [GO:0090132] and “epithelial cell migration” [GO:0010631] were uniquely identified only in active UC among the most overrepresented biological processes (Figure 3b). Target genes of differentially expressed miRNAs between cell populations in the UC remission group were mainly related to cell migration and were least different between those populations (Figure 3b). Noteworthy, the GSEA results should be treated with caution, since the selection of miRNA targets significantly affects the results\(^1\). However, currently, there are no methods to solve this issue, since miRNA target prediction as well as its dosage to affect target expression are still rather unsolved problems in the field.
Figure 3: Aberrantly expressed miRNAs of crypt-top (CD66a+) and crypt-bottom (CD44+) colonic epithelial cells are involved in UC-related processes and reflect disease activity. Top 10 overrepresented pathways within (a) and between (b) crypt-top (CD66a+) and crypt-bottom (CD44+) colonic epithelial cell populations during active (aUC) (n=16), quiescent UC (qUC) (n=15) and in controls (HC) (n=17) identified by miRNA-target gene set enrichment analysis. Dot size represents the...
number of miRNA gene-target count in the significantly enriched (FDR < 0.05) Reactome pathways
(a) and GO biological process (BP) categories (b); (c) a heatmap showing correlations between
miRNA expression and endoscopic Mayo subscore in crypt-bottom (CD44⁺) (n=48) and crypt-top
(CD66a⁺) (n=48) colonic epithelial cell populations. Colour of the box represent the value of
Spearman's correlation coefficient (rho). Dots mark significant correlations (FDR < 0.05) in each cell
population.

To summarize, despite the significant overlap of aberrantly expressed miRNAs of both
colonic epithelial cell populations in regulatory signalling pathways, GSEA results reveal
unique involvement of particular miRNAs in UC-, inflammation- and intestinal barrier
integrity-related processes in different stages of disease activity and/or cell population.

**Expression levels of miRNAs in crypt-top (CD66a⁺) and crypt-bottom (CD44⁺) colonic
epithelial cells reflect UC activity**

To unveil the relation between miRNA expression levels and endoscopic Mayo subscore in
crypt-top (CD66a⁺) and crypt-bottom (CD44⁺) colonic epithelial cells, Spearman correlation
analysis was applied. In crypt-bottom (CD44⁺) and crypt-top (CD66a⁺) colonic epithelial cells,
a number (n=34 and n=23, respectively) moderate positive (0.4 < rho < 0.7; FDR < 0.05)
and a few (n=6 and n=7, respectively) moderate negative (-0.7 < rho < -0.4; FDR < 0.05)
correlations were observed among the normalized miRNA expression levels and endoscopic
Mayo subscore (**Figure 3c; Supplementary Table S6; Supplementary Table S7**).

Although there was substantial overlapping in disease activity-associated miRNAs between
both colonic epithelial cell populations (29 common moderately correlating miRNAs), cell
population-unique correlations were also observed. The expression of 21 and 15 miRNAs
moderately correlated with endoscopic Mayo subscore exclusively in either crypt-bottom
(CD44⁺) or crypt-top (CD66a⁺) colonic epithelial cell population, respectively (**Figure 3c**). For
example, the expression of several miRNAs such as let-7b-5p, let-7e-5p and miR-141-5p
correlated with disease activity only in crypt-bottom (CD44⁺) cells, while expression of
miRNAs such as miR-127-3p, miR-193b-5p and miR-30c-5p only in crypt-top (CD66a⁺) cells.

In addition, out of all correlating cell population-specific miRNAs, eight (miR-10b-5p, let-7b-5p, let-7e-5p, miR-6869-5p, miR-223-3p, miR-31-5p, miR-15b-5p, miR-182-5p) were also identified as differentially expressed in crypt-bottom (CD44⁺) colonic epithelial cells, and seven - in crypt-top (CD66a⁺) cells (miR-194-3p, miR-196b-5p, miR-10b-5p, miR-224-5p, miR-222-3p, miR-135b-5p, miR-21-5p, miR-24-3p, miR-146a-5p, miR-27a) when comparing different stages of UC activity. Some of these miRNAs, such as miR-223-3p, miR-146a-5p, miR-21-5p and miR-31-5p are known to be anti-inflammatory¹⁹, which explains their correlation with disease activity.

Generally, the results demonstrate not only common, but also colonic epithelial cell population-specific correlations between miRNA expression and endoscopic disease activity score.
Discussion

Although UC is a well-studied complex disease and huge efforts have been made to explore its molecular mechanisms, the disease pathogenesis still remains largely unclear\(^2\).

Especially, there is a substantial knowledge gap about expression patterns of regulatory non-coding miRNA in UC in a cell type-specific context. Thus, here we present detailed colonic epithelial cell population-specific miRNA expression profiles from active and quiescent UC patients as well as control individuals and describe differences in miRNA expression patterns between two distinct - crypt-bottom (CD44\(^+\)) and crypt-top (CD66a\(^+\)) - cell populations. Furthermore, we also describe putative biological pathways in which the UC deregulated miRNAs might be involved and determine associations between cell-population specific miRNA expression and disease activity score.

Most importantly, we determined distinct responses in miRNA expression of different colonic epithelial cell populations during UC. Our findings showed that inflammation promoted/suppressed the expression of several miRNAs in colon crypt-bottom (CD44\(^+\)) cells (compared to crypt-top [CD66a\(^-\)] cells) which might be involved in cell proliferation, differentiation and/or permeability of intestinal barrier. For example, let-7c-5p showed considerable \((\log_2 FC < -4)\) down-regulation and miR-501-3p up-regulation \((\log_2 FC > 3.9)\) in crypt-bottom (CD44\(^+\)) cells during active UC. It has previously been shown that overexpression of let-7c-5p as well as inhibition of miR-501-3p can reduce the proliferation of colorectal cancer cells\(^{16,20}\). Thus, deregulation of these miRNAs might be related to the relative increase of the crypt-bottom (CD44\(^+\)) cells in active UC when compared to controls, as it has been shown in our flow cytometry experiment. However, this should be confirmed in an independent functional study employing colonic organoids. On the other hand, we observed increased expression of miR-1-3p and decreased expression of miR-125b-5p in crypt-bottom compared to crypt-top cells only during active UC. Both miRNAs were shown to be involved in barrier function dysregulation, where decrease of miR-125b-5p\(^{21}\) and increase of miR-1-3p\(^{22}\) was shown to disrupt epithelial barrier in colon tissue. This would suggest that
epithelial barrier is already impaired in crypt-bottom epithelial cells during active UC; however, it remains unclear if this is a UC specific event or rather a normal cell response to inflammation in the gut. Furthermore, we observed that some of the differentially expressed miRNAs in crypt-bottom (CD44⁺) and/or crypt-top (CD66a⁺) cells also correlated with endoscopic Mayo subscore. Among these disease activity-associated miRNAs, there were well known anti-inflammatory molecules, including miR-223-3p, miR-146a-5p, miR-21-5p and miR-31-5p. For example, the expression of miR-223-3p and miR-146a-5p are known to be induced via TLR-NF-κB pathway, which explains their correlation with disease activity, since the NF-κB is the central regulator of inflammation. The up-regulation of these two miRNAs has been previously observed in biopsies of colonic mucosa in active UC. We also determined the positive correlation between miR-141-5p expression in crypt-bottom (CD44⁺) cells and UC activity. It has already been shown that miR-141 is predominantly expressed in human colonic epithelial cells and takes part in pathogenesis of UC via targeting CXCL5. In addition, we identified two miRNAs belonging to let-7 family (let-7b-5p and let-7e-5p), the expression of which negatively correlated with the endoscopic Mayo subscore exclusively in crypt-bottom (CD44⁺) colonic epithelial cells, whereas no such association was observed in crypt-top (CD66a⁺) cells. In the intestinal epithelium let-7b appears to be among the highest-expressed miRNAs in let-7 group/family (Supplementary Figure S3). It has previously been shown that expression of let-7 miRNAs (especially let-7e) is increased and affects maintenance of cell differentiation. Altogether, these observations suggest the relevance of let-7 miRNAs during intestinal inflammation via maintenance of stemness of crypt-bottom (CD44⁺) cells and thereby explain their expression correlation with disease activity, exclusively in undifferentiated colonic epithelial cells. Finally, we described potential involvement of differentially expressed miRNAs in regulatory biological pathways. Our initial small RNA-seq of colonic mucosa biopsies from active and quiescent UC compared to healthy controls, at first, revealed multiple deregulated miRNAs, which were significantly enriched in inflammation- and intestinal epithelial barrier function-
related biological pathways. Interestingly, miRNAs enriched in interleukin biological pathways, such as interleukin-4 and interleukin-13 signalling, were deregulated not only in active, but also in quiescent UC, suggesting lasting derangement of this pathway in mucosa of UC. The interleukin-4 and interleukin-13 pathway is known to differentially regulate epithelial chloride secretion and cause epithelial barrier dysfunction. It has been shown that large amounts of interleukin-13 are produced in colon mucosa of UC patients and thereby impair epithelial barrier function by affecting epithelial apoptosis, tight junctions, and restitution velocity. We also confirmed the increased expression of interleukin-13 gene during the course of UC when analysing active and quiescent UC patient colon tissue samples. Contrarily, some studies report decreased mucosal amounts of interleukin-13 in active UC. Nevertheless, the attempts are still being made to adapt the inhibition of interleukin-13-based treatment to induce UC remission (e.g., clinical trials of anti-interleukin-13 monoclonal antibodies (tralokinumab and anruxinumab) and preclinical studies of anti-interleukin-Rα). Our findings on differentially expressed miRNA involvement in interleukin-4 and interleukin-13 regulation even in quiescent UC as well as controversial data in the literature led us to further focus on miRNA expression analysis in undifferentiated crypt-bottom (CD44+) and differentiated crypt-top (CD66a+) colonic epithelial cells, which are responsible for intestinal barrier integrity and permeability. Similarly to the results in colonic biopsies, both crypt-bottom (CD44+) and crypt-top (CD66a+) epithelial cell populations showed deregulation in miRNAs during UC, the targets of which were significantly enriched in the interleukin-4 and interleukin-13 signalling pathway. Since interleukin-4 and interleukin-13 cytokines are predominantly produced by immune cells, the expected regulatory action of deregulated miRNAs in the colonic epithelial cells would be downstream targets of the pathway, such as STAT3, FOXO3, SOCS1, etc. During active UC in both crypt-bottom (CD44+) and crypt-top (CD66a+) cells, we found miR-221-3p, miR-182-6p, miR-222-3p and miR-31-5p to be up-regulated, which are known to target FOXO3 gene. The up-regulation of the aforementioned miRNAs, theoretically, would lead to decreased expression of the FOXO3 gene, which was already observed in colonic mucosa of UC patients. This in turn,
may lead to more severe colonic inflammation during UC\textsuperscript{37}. Additionally, both crypt-bottom (CD44\textsuperscript{+}) and crypt-top (CD66a\textsuperscript{+}) cells of patients with active UC had increased expression of hsa-miR-221-3p and hsa-miR-21-5p, that target SOCS1 gene\textsuperscript{35}. SOCS1 is an important regulator of interleukin-4 signalling, and its forced expression was shown to inhibit interleukin-13 signalling in epithelial cells\textsuperscript{38}. In addition to involvement in interleukin signalling pathways, we also detected unique intestinal epithelial barrier function-related processes regulated by disease stage and/or cell population-specific miRNAs. For example, differentially expressed miRNAs, such as miR-222-3p, miR-223-3p, miR-15b-5p, in crypt-bottom (CD44\textsuperscript{+}) colonic epithelial cells of quiescent UC could potentially regulate nuclear receptors and extra-nuclear estrogen signalling pathways. These findings fall in accordance with other studies demonstrating the importance of impairment of signal transduction via nuclear receptors in inflammatory bowel disease, as nuclear receptors regulate essential aspects of intestinal barrier functions such as mucus secretion, expression of tight junction proteins and others\textsuperscript{17}. Moreover, we observed a few aberrantly expressed miRNAs between crypt-bottom (CD44\textsuperscript{+}) and crypt-top (CD66a\textsuperscript{+}) cells in active UC, that possibly exert their biological function through regulation of epithelial cell migration, which is known to happen along the crypt-villus axis\textsuperscript{39} and is increased during IBD\textsuperscript{40}.

In summary, we present a comprehensive study, which determined crypt-bottom (CD44\textsuperscript{+}) and crypt-top (CD66a\textsuperscript{+}) colonic epithelial cell-specific miRNA deregulation in UC in cell type- and disease stage-dependent manner. We also revealed and emphasized differences of cell population-specific miRNA expression associations with disease activity. Furthermore, we unveiled the potential functional role of differentially expressed miRNAs in/between two unique crypt-bottom (CD44\textsuperscript{+}) and crypt-top (CD66a\textsuperscript{+}) colonic epithelial cell populations. We observed potential involvement of deregulated miRNAs in biological pathways associated with maintenance of intestinal barrier function in active as well as quiescent UC, in both epithelial cell populations. Together, these observations not only highlight regulatory importance of miRNAs in distinct colonic epithelial cell populations during pathogenesis of
UC, but also could be applied for the development of new treatment strategies to maintain the remission of mucosal inflammation.
Acknowledgements

This work was supported by the Research Council of Lithuania and European Crohn’s and Colitis Organisation (grant numbers S-MIP-20-56 and ECCO Grant 2016, respectively).

Ethical Considerations

The approval to perform the study was received from Kaunas Regional Biomedical Research Ethics Committee (No. BE-2-31, 22-03-2018). All subjects have signed a written informed consent form to participate in the study.
References


