Journal of Crohn's and Colitis, 2019, 916–930
doi:10.1093/ecco-jcc/jjz007
Advance Access publication January 18, 2019
Original Article

Original Article

Gene and Mirna Regulatory Networks During Different Stages of Crohn's Disease

Sare Verstockt[a], Gert De Hertogh[b,c], Jan Van der Goten[d], Bram Verstockt[d,e], Maaike Vancamelbeke[d], Kathleen Machiels[d], Leentje Van Lommel[f], Frans Schuit[f], Gert Van Assche[d,e], Paul Rutgeerts[d,e], Marc Ferrante[d,e], Séverine Vermeire[d,e], Ingrid Arijs[d,g,§], Isabelle Cleynen[a,§]

[a]Laboratory for Complex Genetics, Department of Human Genetics, KU Leuven, Leuven, Belgium [b]Department of Morphology and Molecular Pathology, University Hospitals, Leuven, Belgium [c]Department of Imaging & Pathology, Translational Cell & Tissue Research, KU Leuven, Leuven, Belgium [d]Translational Research Center for Gastrointestinal Disorders [TARGID], Department of Chronic Diseases, Metabolism & Aging [CHROMETA], KU Leuven, Leuven, Belgium [e]Department of Gastroenterology and Hepatology, University Hospitals, KU Leuven, Leuven, Belgium [f]Gene Expression Unit, Department of Cellular and Molecular Medicine, KU Leuven, Leuven, Belgium [g]Faculty of Medicine and Life Sciences, Hasselt University, Hasselt, Belgium; Jessa Hospital, Hasselt, Belgium

[§]These authors contributed equally.

Corresponding author: Isabelle Cleynen, Laboratory for Complex Genetics, Department of Human Genetics, KU Leuven, O&N4 Herestraat 49 Box 610, 3000 Leuven, Belgium. Tel: 32-16-37-74-80; fax: 32-16-34-60-98; email: isabelle.cleynen@kuleuven.be

Conference presentation:

1. Verstockt S., Van der Goten J., Vancamelbeke M., Verstockt B., et al. Molecular profiling of early Crohn's disease reveals a prominent role for WNT5A. Oral presentation at BIRD Meeting of the BWGE Conference [Antwerp, February 2017], Y-ECCO Basic Science Workshop at the ECCO Congress [Barcelona, February 2017], and Digestive Disease Week [Chicago, May 2017].
2. Cleynen I., Verstockt S., De Hertogh G., Van der Goten J., et al. Dysregulated gene and miRNA expression in different stages of Crohn's disease. Poster presentation at ASHG meeting [Orlando, October 2017]
3. Verstockt S., Van der Goten J., Vancamelbeke M., Verstockt B., et al. Molecular landscape of early Crohn's disease using an integrated approach of mRNA/miRNA profiling and genomics. Poster presentation at ECCO Congress [Amsterdam, March 2016] and BeSHG Meeting [Leuven, Feb. 2016].

Abstract

Background and Aims: Early treatment of Crohn's disease [CD] is required in order to optimize patient outcomes. To this end, we need to gain a better understanding of the molecular changes at the onset of CD.

Methods: As a model for the earliest mucosal CD lesions, we study post-operative recurrent CD [Rutgeerts score ≥ i2b]. We are the first to analyse gene and microRNA [miRNA] expression profiles in ileal biopsies from these patients, and compare them with those of newly diagnosed [≤18 months] and late-stage [>10 years after diagnosis] CD patients.

Results: Except for one gene [*WNT5A*], there are no differential genes in CD patients without post-operative recurrence [i0], showing that previous disease did not influence gene expression in the

neoterminal ileum, and that this model can be used to study early mucosal CD lesions. Gene expression and co-expression network dysregulation is more pronounced in newly diagnosed and late-stage CD than in post-operative recurrent CD, with most important modules associated with [a]granulocyte adhesion/diapedesis, and cholesterol biosynthesis. In contrast, we found a role for snoRNAs/miRNAs in recurrent CD, highlighting the potential importance of regulatory RNAs in early disease stages. Immunohistochemistry confirmed the expression of key dysregulated genes in damaged/regenerating epithelium and immune cells in recurrent CD.

Conclusions: Aside from regulatory RNAs, there are no clear gene signatures separating post-operative recurrent, newly diagnosed, and late-stage CD. The relative contribution of dysregulated genes and networks differs, and suggests that surgery may reset the disease at the mucosal site, and therefore post-operative recurrent CD might be a good model a good model to study to study early mucosal CD lesions.

Key Words: Early Crohn's disease; gene expression; gene regulation; molecular mechanisms

1. Background

Crohn's disease [CD] is a chronic, relapsing inflammatory bowel disease [IBD] with increasing prevalence worldwide.[1] CD is believed to be caused by an interaction of genetic, immune, and environmental factors, leading to an inappropriate chronic activation of the mucosal immune response,[2] often resulting in strictures or fistulas. Several studies suggest that treatment is more effective and can prevent complications when initiated early in the disease course,[3,4] although currently we do not know what the early triggers of disease are. The subclinical phase before diagnosis of CD often goes unnoticed, and at the time the diagnosis of CD is made, bowel damage has already occurred in a significant number of patients, and immune dysregulation, dysbiosis and tissue injury associated with full-blown disease is set, and in many cases, irreversible.[5]

Up to 50% of patients with CD need surgery within 10 years of diagnosis.[6] Rutgeerts et al. showed that after ileal resection with ileo-colonic anastomosis, in >70% of patients new CD lesions recur in the neoterminal ileum within weeks to months.[7] These lesions may progressively evolve to full blown CD, and develop only when the gut is exposed to luminal contents.[8] When luminal fluid is diverted through a proximal stoma and the ileo-colonic anastomosis is protected, the gut mucosa remains intact.[9] Infusion of intestinal contents through the diverted bowel triggers early changes consistent with recurrent inflammation.[10]

To get a better understanding of the underlying mechanisms of early CD, it is important to know which molecular changes are at the onset of CD. Large-scale gene expression studies have been performed in IBD and found dysregulated expression of genes mainly involved in immune response, cell adhesion, barrier integrity and tissue remodelling.[11–14] These however did not focus on the earliest mucosal lesions in CD. Gene expression is partly regulated by microRNAs [miRNAs], small noncoding RNAs that bind to the 3′ untranslated region of target mRNAs, hereby negatively regulating their stability or translation. A few studies have investigated the role of miRNAs in IBD, and found distinct miRNA expression profiles in IBD patients with active and inactive disease, and normal controls,[15,16] but again did not specifically include the earliest mucosal lesions in CD patients.

To understand what happens early in the disease course, we studied ileal mucosal mRNA and miRNA expression in patients with post-operative recurrence, and compared the expression with that in patients without post-operative recurrence, or with newly diagnosed CD, and with late-stage CD.

2. Materials and Methods

2.1. Ethics statement

This study was carried out at the University Hospitals Leuven [Leuven, Belgium]. The ethics committee of the University Hospitals Leuven approved the study [IRB approvals B322201213950/S53684 and B322201110724/S52544]. All individuals gave written informed consent.

2.2. Patients and biopsy specimens

Ileal mucosal biopsies were obtained during endoscopy from 24 post-operative recurrent CD patients [Rutgeerts score ≥ i2b,[17] taken ≤18 months after ileal resection with ileo-colonic anastomosis]; from 18 newly diagnosed, active CD patients [taken ≤18 months after diagnosis as defined by Peyrin-Biroulet et al.[18]; however, only fistulizing disease was excluded]; from 14 late-stage CD patients [>10 years after diagnosis, and with active disease]; and from non-IBD controls with normal endoscopy [12 for mRNA profiling; eight for miRNA profiling]. We also included three CD patients with uninflamed post-operative ileum since surgery [Rutgeerts score i0, and taken ≤18 months after ileal resection with ileo-colonic anastomosis]. Baseline characteristics of the different groups are summarized in Table 1, and Supplementary Table S1. An overview of the number of samples used for each analysis is listed in Supplementary Table S2.

Biopsies were taken at the most affected site of the [neoterminal] ileum, and always at the border of the ulcerative surface. Part of the biopsies were immediately snap-frozen in liquid nitrogen [mRNA profiling], or placed in RNAlater [Ambion, Austin, TX, USA] and snap-frozen in liquid nitrogen [miRNA profiling] and stored at −80°C. Remaining biopsies were used for routine histopathological examination. The features of chronic intestinal inflammation were scored in hematoxylin/eosin stained slides from the paraffin blocks of each patient using a previously reported scoring system for CD.[10]

2.3. mRNA isolation and oligonucleotide array hybridization

Total RNA was extracted from biopsies using the RNeasy Mini Kit [QIAGEN Valencia, California, USA]. The quantity and integrity was assessed with a Nanodrop ND-1000 spectrophotometer [Nanodrop Technologies, Wilmington, Delaware, USA] and 2100 Bioanalyzer [Agilent, Waldbronn, Germany] using the Agilent RNA 6000 Nano kit, respectively. All samples showed good RNA quality

Table 1. Baseline characteristics of subjects from gene expression study

Baseline characteristic	Post-operative recurrent CD	Newly diagnosed CD	Late-stage CD	Post-operative CD i0	Controls
	[*n* = 24]	[*n* = 18]	[*n* = 14]	[*n* = 3]	[*n* = 12]
Median [IQR] age [years]	39.0 [27.3–53.6]	24.7 [18.9–32.6]*	47.6 [40.0–59.4]	63.5 [22.4–66.1]	55.4 [29.4–74.1]
Male/Female [%]	12/12 [50.0/50.0]	12/6 [66.7/33.3]	6/8 [42.9/57.1]	0/3 [0/100]	7/5 [58.3–41.7]
Median [IQR] duration of disease [years]	10.1 [4.2–20.8]	0 [0.0–0.0][$]	23.7 [17.1–32.9]	19.4 [0.49–33.2]	NA
Median [IQR] time between surgery and biopsy [months]	5.6 [4.7–7.0]	NA	38.5 [13.8–47.6]	5.3 [5.0–5.5][$$]	NA
Disease location					
ileal L1 [%]	22 [91.7]	7 [38.9]	2 [14.3]	0 [0.0]	NA
ileo-colonic L3 [%]	2 [8.3]	11 [61.1]	12 [84.7]	3 [100.0]	NA
Disease behavior					
inflammatory B1 [%]	18 [75.0]	12 [66.7]	2 [14.3]	3 [100.0]	NA
stricturing B2 [%]	5 [20.8]	6 [33.3]	9 [64.3]	0 [0.0]	NA
fistulizing B3 [%]	1 [4.2]	0 [0.0]	3 [21.4]	0 [0.0]	NA
Prior use of anti-TNF [%]	13 [59.1]	0 [0.0]	3 [21.4]	2 [66.7]	NA
Concomitant medication:					
Mesalazine [%]	1 [4.2]	0 [0.0]	5 [35.7]	0 [0.0]	NA
Corticosteroids [%]	4 [16.7]	4 [21.1]	3 [21.4]	0 [0.0]	NA
Immunomodulators [%]	0 [0.0]	0 [0.0]	2 [14.3]	0 [0.0]	NA
Biologics [%]	0 [0.0]	0 [0.0]	0 [0.0]	0 [0.0]	NA
Antibiotics [%]	1 [4.2]	1 [5.6]	0 [0.0]	0 [0.0]	NA
Active smoking [%]	9 [37.5]	10 [55.6]	1 [7.0]	2 [67.0]	3 [25.0]

[$]Min–max duration of disease: 0.2–1.5 years; [$$]no recurrence since surgery; *$p < 0.05$ as compared with controls.
CD, Crohn's disease; IQR, interquartile range; *n*, number; NA, non-applicable.

[260/280 nm absorbance ratios > 1.9 and RNA Integrity Number [RIN] > 7].

Total RNA [150 ng] was used to analyze mRNA expression via Affymetrix GeneChip® Human Gene 1.0ST arrays [Affymetrix, Santa Clara, CA, USA], which are comprised of 33 252 probe sets (99% of all well-annotated human genes; and a limited fraction of non-coding RNAs, including small nucleolar RNAs [snoRNAs]). All steps were performed according to Affymetrix manufacturer's manual 4425209 Rev.B, and as described earlier.[16,19] The mRNA microarray data were deposited to the Gene Expression Omnibus database according to Minimum Information About a Microarray Experiment [MIAME] guidelines [series accession GSE102133, https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE102133].

2.4. miRNA isolation and oligonucleotide array hybridization

For miRNA analysis, total RNA, including small RNA, was isolated from adjacent biopsies with the mirVana miRNA Isolation kit [Ambion]. The quantity and integrity was assessed with a NanoDrop ND-1000 spectrophotometer and a 2100 Bioanalyzer using the Agilent Small RNA kit, respectively. All samples showed good RNA quality [260/280 nm absorbance ratios > 1.9 and RIN > 7].

Total RNA [250 ng] was analyzed with Affymetrix GeneChip® miRNA 2.0 arrays containing 4560 probe sets for human small RNAs, out of which 1082 human mature miRNAs were filtered. These probe sets guarantee 100% coverage of all human mature miRNAs in miRBase v.15 [April, 2010]. Probe sets that were deleted in a more recent version of miRBase were excluded for analysis. All steps were performed according to Affymetrix, as previously described.[16] Quality control was performed with the Affymetrix Expression Console Software, version 1.4. The miRNA microarray data were deposited in MIAME format to the Gene Expression Omnibus database [series accession GSE102127, https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE102127].

2.5. Differential expression analysis

Microarray data [mRNA and miRNA] were analyzed as previously described[19] using Bioconductor tools[20] in R [version 3.4.2, http://www.r-project.org]. Probe level analysis was performed on the Affymetrix raw data [.cel files] with the robust multichip average [RMA] method[21] to obtain log2 expression values. GeneChip® Human Gene 1.0 ST and GeneChip® miRNA 2.0 arrays were processed using the aroma.affymetrix R package [http://www.aroma-project.org/] or the Affymetrix Expression Console Software, respectively. A non-specific filtering was applied on the normalized log2 expression values to obtain only those probe sets hybridizing above background levels [intensity > log2[100] in 5% of samples]. In addition, only those probe sets with a high variability in intensities across the samples [interquartile range, IQR > 0.5] were selected. The probe sets that passed these filtering [*n* = 4110 on mRNA level; *n* = 1187 on miRNA level] processes were used for further analysis.

The linear models for microarray data [LIMMA] package[22] was used to identify differentially expressed probe sets. Obtained *p*-values were adjusted for multiple testing [Benjamini–Hochberg method].[23] An adjusted *p*-value < 0.05 and >2-fold change [FC] [mRNA] or >1.5-FC [miRNA] were considered biologically significant. Gene probe sets were annotated through the Affymetrix NetAffx website [NetAffx release 33.2, http://www.affymetrix.com/analysis/index.affx] or the NCBI website [http://www.ncbi.nlm.nih.gov]. Annotations of miRNA probe sets were derived from the Sanger miRBase database v.20 [June 2013, http://mirbase.org]. IPA [QIAGEN Inc., https://www.qiagenbioinformatics.com/products/

ingenuity-pathway-analysis] was used to identify biological [sub] functions and canonical pathways associated with the differentially expressed gene probe sets.

Differentially expressed genes and miRNAs were integrated using the IPA microRNA Target Filter. Tarbase, miRecords, and IPA Expert were used to annotate experimentally validated miRNA–mRNA pairs. TargetScan was used to determine the confidence levels for the remaining predicted pairs. The confidence level indicates whether the miRNA is predicted to repress its mRNA target to 40% [high] or to 65% [moderate] of the 'normal' expression level.

2.6. Quantitative real-time reverse-transcriptase polymerase chain reaction

To validate the array data, quantitative real-time reverse-transcriptase polymerase chain reaction [qRT-PCR] was performed for selected differentially expressed genes. β actin was used as the endogenous reference gene. An overview of the validated genes with the corresponding primer and probe sequences are given in Supplementary Table S3. All steps were performed as described earlier.[13]

miRNA microarray data were validated for selected differentially expressed miRNAs [hsa-let-7g-5p, hsa-miR-30d, and hsa-196a-5p] using qRT-PCR, as previously described.[16] Specific Locked Nucleic Acid miRNA primers were purchased from Exiqon [Vedbaek, Denmark]. The relative amount of each miRNA was calculated as a ratio to the amount of RNU6B, the endogenous reference, as before.[16]

2.7. Immunohistochemistry

To localize the corresponding proteins of validated genes [FOLH1, MUC1, MUC4, LCN2, S100A8 and WNT5A] in the ileal mucosa, immunohistochemistry was performed on 5-µm-thick sections from paraffin blocks of formalin-fixed endoscopic biopsies as previously described.[19] Specific protocols for each protein are summarized in Supplementary Table S4. All procedures were manually conducted by the same operator [SaV]. All stains were evaluated by an experienced IBD pathologist [GDH]. Microscopic images were acquired with Leica Application Suite V4.1.0 software using a Leica DFC290 HD camera [Leica Microsystems Ltd, Heerbrugg, Switzerland] mounted on a Leica DM2000 LED bright field microscope.

WNT5A staining was also performed on resected ileum from a CD patient undergoing surgery, and resected normal ileum from a non-IBD control with colorectal cancer. The resected tissue in the latter case was taken next to the tumor-free resection margin.

2.8. Co-expression network analysis

To identify group-specific network clusters [modules] based on highly correlated genes, Weighted Gene Correlation Network Analysis [WGCNA] was applied using R [Supplementary Figure S1].[24] Co-expression networks were generated by calculating a gene pair-wise Pearson correlation matrix. An additional filtering step was performed on gene probe sets analysed for differential expression [n = 4110], i.e. only annotated genes were included in this network analysis [n = 2979]. Gene probe sets were subsequently assigned to modules by average linkage hierarchical clustering. Minimal module size was set to 20 probe sets at a deepSplit = 3. Co-expression analysis was performed on the combined expression data of all samples [post-operative recurrent, newly diagnosed, and late-stage CD, and controls]. Each module was then tested for correlation with each of the traits [post-operative recurrent CD, newly diagnosed CD, late-stage CD]. Multiple testing correction [Benjamini–Hochberg method] was applied on the module-trait correlation testing. An adjusted p-value < 0.05 was considered significant. Module eigengene was defined as the first principal component, which summarizes the expression patterns of all probe sets into a single expression profile within a given module. Genes showing high correlation with the module eigengene were referred to as intramodular hub genes.

2.9. Statistical analysis

Continuous variables [except for microarray data] are expressed as median and IQR, and categorical variables as frequencies and percentages. Comparisons were done using the Mann–Whitney U-test [continuous variables], and Fisher's exact test [categorical variables]. A p-value < 0.05 was considered significant. Statistical analyses were done using IBM SPSS Statistics 24 [IBM SPSS, Costa Mesa, CA, USA] and R [version 3.4.2].

3. Results

3.1. Patient cohorts

Fifty-nine CD patients with post-operative recurrence [n = 24], without post-operative recurrence [n = 3], with a new diagnosis [n = 18], or with late-stage disease [n = 14], and controls [n = 12] were recruited and included in the gene expression study [Table 1]. There were no significant differences between the CD groups and non-IBD control group for any of the demographic variables, except for age, which was significantly different between newly diagnosed CD (median age 24.7 [IQR, 18.9–32.6] years) and non-IBD controls (median age 55.4 [IQR, 29.4–74.1] years).

Biopsies from newly diagnosed CD patients were collected at the time of index endoscopy, except for four patients (median time between diagnosis and biopsy collection 0.0 years [min–max 0.2–1.5 years). Moreover, newly diagnosed CD patients were treatment-naïve, except for four patients being exposed to corticosteroids. In addition, none of the 59 CD patients was being treated with a biologic at the time of biopsy collection. However, the proportion of patients exposed to anti-TNF treatment before was the highest in post-operative CD patients [without recurrence: 66.7%; with recurrence 59.1%], compared with late-stage CD patients [21.4%]. Besides, the proportions of L1/L3 differed among the several CD groups, with most of the post-operative recurrent CD patients showing ileal disease [91.7%]. In addition, as disease behaviour is partly associated with the disease stage, the proportion of stricturing and fistulizing disease was the highest in late-stage CD [64.3% and 21.4%, respectively].

3.2. Differential gene expression analysis in different CD stages

Compared with non-IBD controls, 353 [234 up, 119 down], 659 [382 up, 277 down] and 613 [388 up, 225 down] significantly differentially expressed gene probe sets were identified in post-operative recurrent, newly diagnosed and late-stage CD, respectively [Figure 1A–C; Supplementary Tables S5–7]. A total of 265 gene probe sets were overlapping for the three groups [Figure 2A], which encode for proteins involved in immune-related [sub]functions [Supplementary Figure S2A]: cell-to-cell signaling and interaction [e.g. adhesion, activation of immune cells], cellular movement [e.g. chemotaxis, cell migration], hematological system development and function, immune cell trafficking, and cellular growth and proliferation. Associated canonical pathways included [a]granulocyte adhesion and diapedesis, IL-17 signaling, and acute phase response [Supplementary Figure S2B].

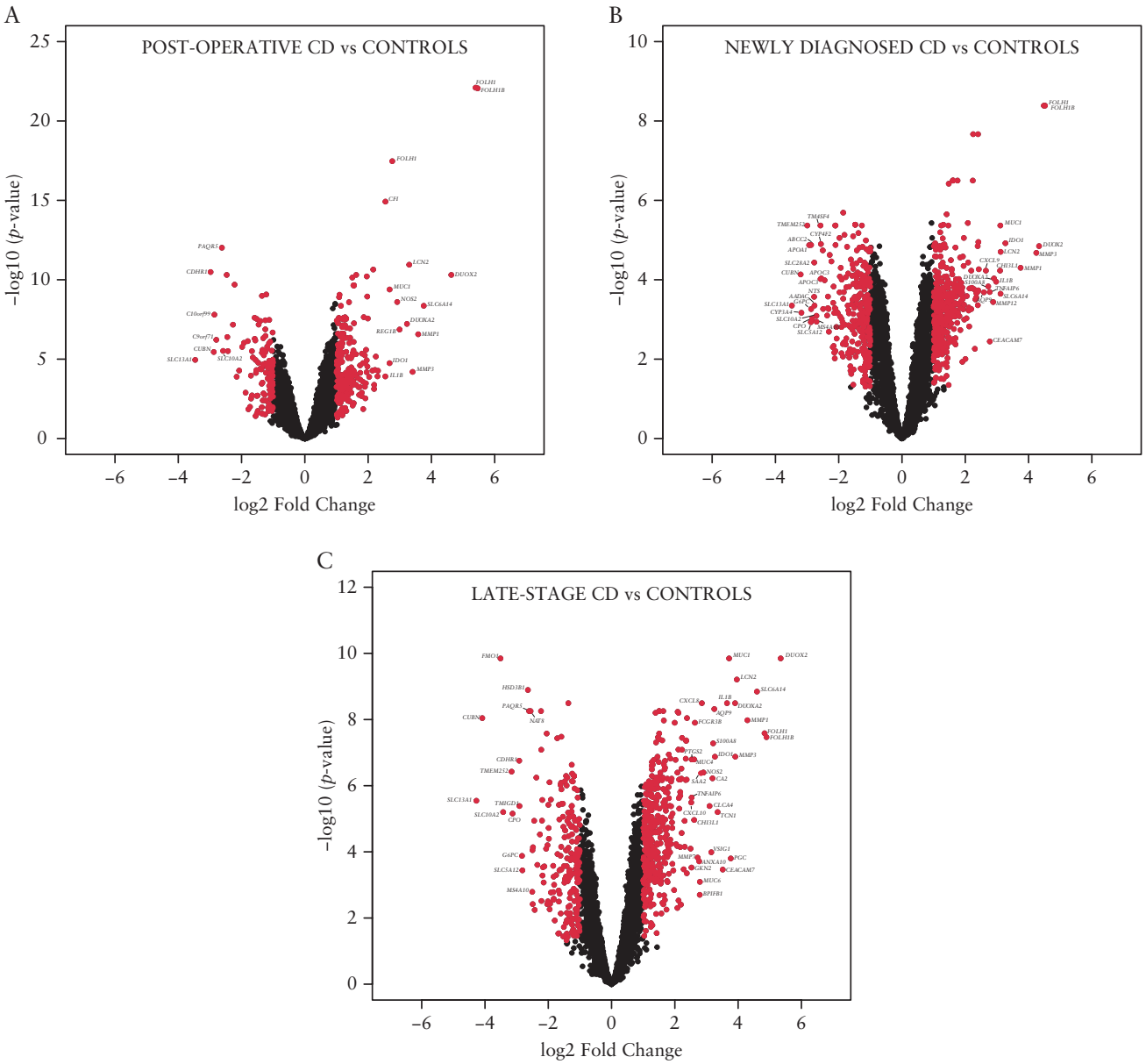

Figure 1. Volcano plots of differentially expressed gene probe sets in [A] post-operative recurrent CD vs controls, [B] newly diagnosed CD vs controls and [C] late-stage CD vs controls. The log2 ratio and -log10 adjusted *p*-values are plotted in the form of volcano plots. The probe sets in red represent biologically significantly different gene probe sets. The non-biologically significant probe sets are represented in black. CD, Crohn's disease; vs, versus.

The top dysregulated gene in post-operative recurrent CD was *FOLH1* [FC = 42.11, adjusted p = 8.01E-23]. This was also the most dysregulated gene in newly diagnosed CD [FC = 23.05, adjusted p = 1.34E-09], while it was ranked 27th in late-stage CD [FC = 28.66, adjusted p = 2.60E-08]. Several other genes related to CD pathogenesis (antimicrobial peptides [AMPs], barrier genes, cellular adhesion molecules [CAMs], and CD susceptibility genes) were also significantly differentially expressed in post-operative recurrent CD [listed in Table 2].

Five dysregulated genes in post-operative recurrent CD vs controls were chosen to confirm their altered expression with qRT-PCR in all patient groups. Selection was based on their biological significance in post-operative recurrent CD and/or clinical or functional relevance: [1] the most significantly differentially expressed gene [*FOLH1*]; [2] top intestinal barrier genes [*MUC1*, *MUC4*]; and [3] top AMPs [*S100A8*, *LCN2*]. Differential expression of all five mRNAs was confirmed [Supplementary Figure S3A–E].

3.3. Differential gene expression analysis in i0 CD

To be sure that expression levels in post-operative recurrence i2b/i3/i4 are not influenced by previous disease [before surgery], we also studied gene expression in uninflamed post-operative CD ileum [i0]. LIMMA analysis found only one biologically significant gene probe set [*WNT5A*, FC = 2.00, adjusted p = 3.45E-02] between i0 CD patients and non-IBD controls. No other gene probe sets were significantly different between i0 CD and controls.

Similarly to the microarray data, qRT-PCR validation [Supplementary Figure S3F] found significantly increased *WNT5A* expression levels in post-operative recurrent CD (median 0.45

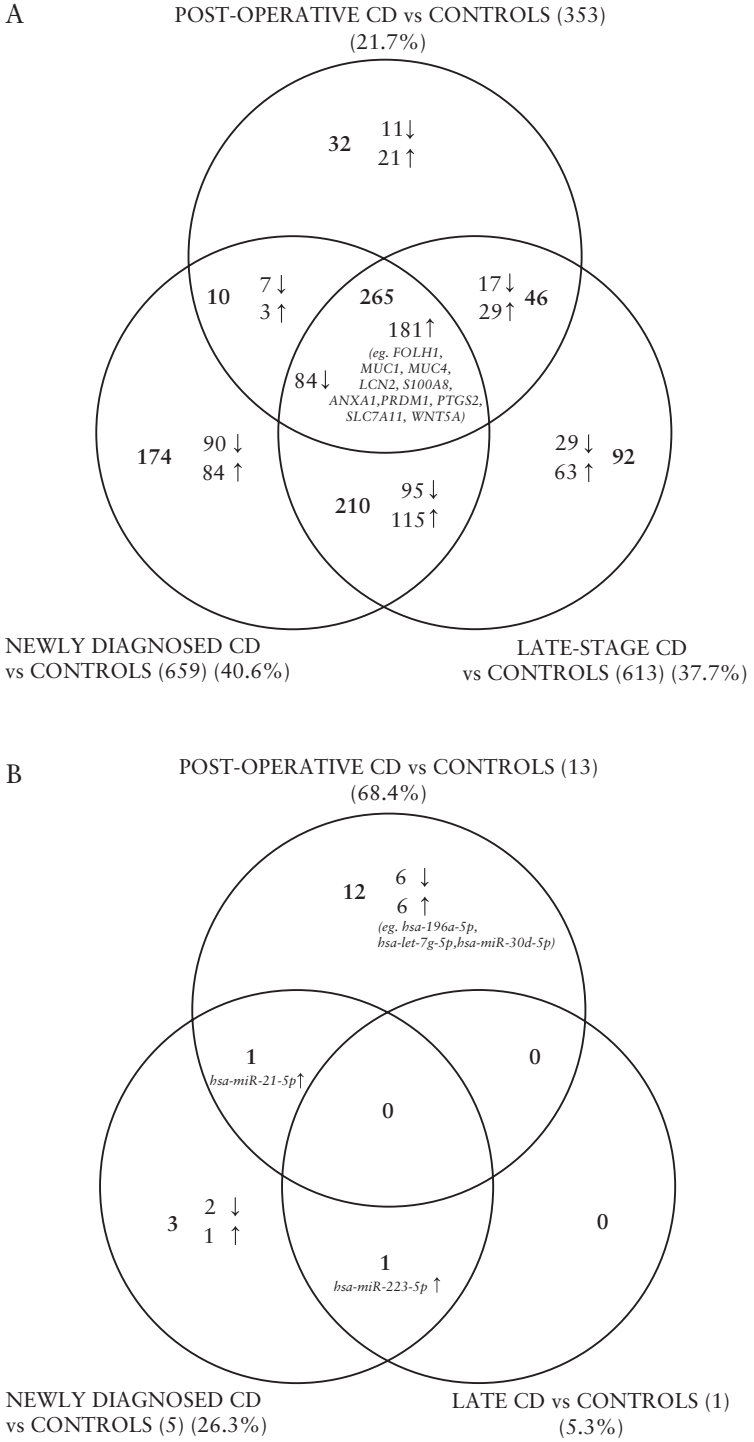

Figure 2. Venn diagram of [A] the significant gene probe sets in post-operative recurrent, newly diagnosed, and late-stage CD; and [B] significant miRNAs in post-operative recurrent, newly diagnosed, and late-stage CD compared with controls. The significantly differentially expressed gene or miRNA probe sets in the comparative analyses are depicted in overlapping circles. Percentages reflect the number of differentially expressed gene/miRNA probe sets for each comparison relative to the total number of differentially expressed gene/miRNA probe sets in the three comparisons. CD, Crohn's disease; miRNA, microRNA; vs, versus.

[IQR 0.34–0.70]) versus controls (median 0.21 [IQR 0.10–0.33], $p < 1E-04$), and even more in newly diagnosed CD (median 0.74 [IQR 0.61–1.12], $p < 1E-04$), while late-stage CD showed similar expression (median 0.43 [IQR 0.30–0.55, $p = 1E-03$]) to that of recurrent CD. We also observed a slightly increased *WNT5A* expression in the uninflamed post-operative ileum [i0], but unlike the microarray data this was not significantly different from controls [$p = 0.21$].

Table 2. List of selected genes significantly differentially expressed in post-operative recurrent, newly diagnosed and late-stage CD as compared with controls

		Post-operative recurrent CD vs controls		Newly diagnosed CD vs controls		Late-stage CD vs controls	
	Symbol	FC	Adjusted *p*	FC	Adjusted *p*	FC	Adjusted *p*
AMPs	*LCN2*	9.84	1.13E–11	8.68	1.98E–05	15.56	6.18E–10
	NOS2	7.56	2.52E–09	4.56	5.48E–04	7.48	4.02E–07
	S100A8	4.97	5.36E–05	6.61	1.47E–04	9.26	5.25E–08
	C5	2.75	8.74E–06	–	–	–	–
	NTS	–2.51	4.00E–03	–6.85	2.69E–04	–5.08	2.48E–04
Barrier genes	*MUC1*	6.41	4.10E–10	8.62	4.30E–06	13.13	1.42E–10
	MUC4	3.57	1.78E–05	5.25	1.42E–05	6.11	1.61E–07
	CLDN1	2.35	2.13E–04	2.71	3.44E–04	3.78	1.90E–05
	TFF1	2.20	1.52E–04	–	–	4.51	2.01E–06
	CLDN18	2.08	5.93E–06	–	–	4.86	3.37E–04
	ABCB1	–2.26	3.11E–05	–3.24	5.87E–04	–2.87	1.75E–05
	CDHR1	–7.88	3.26E–11	–4.87	2.37E–05	–7.56	1.76E–07
CAMs	*CXCL10*	4.40	5.27E–05	4.99	3.13E–04	5.73	3.22E–06
	CXCL11	3.94	2.92E–05	3.67	3.43E–04	5.07	6.73E–07
	CCL28	3.92	1.14E–09	2.78	8.09E–04	4.50	6.54E–07
	CXCL9	3.81	3.33E–04	6.28	5.90E–05	4.95	1.16E–05
	IL8	3.73	5.49E–04	5.10	2.19E–04	7.23	3.20E–09
	CXCL1	2.73	8.94E–05	3.33	1.98E–04	4.33	6.28E–09
	SELE	2.57	4.12E–03	3.16	3.86E–03	3.39	5.87E–05
	CCL2	2.58	1.88E–03	4.56	5.90E–05	3.46	6.83E–05
	ICAM1	2.37	7.23E–04	2.95	3.25E–04	2.79	2.61E–05
	CXCR1	2.35	2.04E–03	2.92	6.79E–04	3.73	6.12E–07
	CXCL5	2.31	6.18E–04	3.48	9.57E–05	2.91	3.74E–05
	SELP	2.30	7.53E–04	2.76	1.08E–03	2.68	1.65E–04
	CXCL2	2.25	9.98E–05	–	–	2.76	1.31E–05
	CXCR2	2.09	2.50E–03	2.38	1.70E–03	3.12	3.22E–06
	CCL20	2.26	2.62E–02	–	–	2.69	2.86E–02
CD susceptibility genes[1]	*FCGR3A*	3.88	1.17E–05	5.38	5.44E–05	6.21	1.25E–08
	ITLN1	3.64	2.26E–08	2.65	1.23E–02	3.62	4.25E–05
	FCGR2A	3.32	5.67E–04	4.39	1.66E–04	5.12	4.43E–08
	S100A11	2.30	9.15E–08	2.31	6.39E–06	2.84	2.64E–08
	STAT1	2.28	7.46E–08	2.45	4.48E–05	2.42	1.93E–07
	SLAMF7	2.10	1.58E–03	2.64	1.14E–03	2.40	9.15E–05
	PRDM1	2.05	8.92E–05	2.47	1.87E–04	2.54	7.33E–07
	PLAU	2.03	3.92E–03	2.64	2.59E–04	2.31	4.73E–06
	FCRLA	–2.13	3.07E–03	–	–	–2.41	1.60E–03
	DHDH	–2.14	1.88E–04	–2.98	7.95E–05	–2.76	2.27E–05
	FAIM3	–2.19	6.71E–03	–	–	–2.37	1.16E–02
	SLC9A2	–2.34	7.22E–05	–2.12	7.10E–03	–	–
	MS4A10	–2.56	1.06E–02	–6.50	1.14E–03	–5.68	1.62E–03
	PLB1	–3.25	3.04E–03	–4.01	1.06E–03	–5.66	8.94E–05

When significant, the fold change and adjusted *p*-value in the other comparisons [newly diagnosed CD versus controls, or late-stage CD versus controls] are also given.

[1]Selection of CD susceptility genes was based on Jostins et al.[53]

AMPs, anti-microbial peptides; CAMS, cellular adhesion molecules; CD, Crohn's disease; FC, fold change; vs, versus.

3.4. Protein localization of selected genes

Immunohistochemistry was performed to localize FOLH1, MUC1, MUC4, LCN2, S100A8 and WNT5A in the ileal mucosa of post-operative recurrent CD patients and controls. FOLH1 was not detectable in normal mucosa of controls, whereas an increased protein expression was observed in surface epithelium and in regenerating and damaged epithelium of post-operative recurrent mucosa [Supplementary Figure S4A–C]. MUC1 was weakly detected along the apical membrane and in Paneth cells in normal mucosa of controls, whereas in post-operative recurrent mucosa, we observed a more intense and widespread expression in regenerating epithelium and in inflammatory cells [Supplementary Figure S5A–C]. MUC4 protein expression was accentuated in Goblet cells in normal mucosa of controls. A similar pattern was seen in post-operative recurrent mucosa with somewhat enhanced expression in Paneth cells [Supplementary Figure S5D and E]. LCN2 expression was detected in Paneth cells in the normal mucosa of controls [Supplementary Figure S6A]. In post-operative recurrent mucosa, an intensive staining was observed in epithelial cells, and also in inflammatory cells, although to a lesser extent Supplementary Figure S6B and C]. S100A8 expression was found in few inflammatory cells in normal mucosa, whereas in post-operative recurrent mucosa S100A8 was

detected in polymorphonuclear leukocytes and in endothelial cells [Supplementary Figures S6D and E].

Given that we found no prior reports on ileal WNT5A expression, we first studied WNT5A localization and expression in resected ileum. WNT5A expression was observed in Paneth cells in normal control ileum [Supplementary Figure S7A and D]. In active CD, there was extensive staining in damaged and regenerating epithelium [Supplementary Figure S7B, C and E]. Considering inflammatory cells, in normal ileum, WNT5A expression was only observed in lamina propria lymphocytes [Supplementary Figure S7A and F], whereas in active CD there was an increased expression in B and T cells, and in solitary giant cells or in epithelioid granulomas [Supplementary Figure S7G, H and J]. A clear staining in fibroblasts [scar tissue] and endothelial cells was also detected in active CD Supplementary Figure S7I, J and K]. WNT5A expression localization was confirmed in a series of ileal biopsies from post-operative recurrent CD patients, with a stronger staining when damaged and/or regenerating epithelium or an increased inflammatory infiltrate was present [Supplementary Figure S8A and B].

3.5. Co-expression gene network analysis

To identify clusters of genes with a similar ileal mucosal expression pattern, we performed WGCNA. Twenty-eight modules of correlated gene probe sets ranging in size from 31 to 453 transcripts were identified [Supplementary Figure S9]. Module–trait correlation analysis found that approximately half of these modules were significantly correlated with post-operative recurrent [n = 15], newly diagnosed [n = 19] and/or late-stage CD [n = 19] [Figure 3, Supplementary Table S8].

The strongest correlation with post-operative recurrent CD was found for a module of 64 gene probe sets [darkturquoise color code, r = 0.88, adjusted p = 7E-11] [Figure 3]. Strongest correlation with the darkturquoise eigengene included hub genes *CFB* and *CASP10*. Among the probe sets present in this module, 27 were significantly differentially expressed in post-operative recurrent CD as compared with controls [e.g. *FOLH1*]. The royalblue module [42 probe sets] was identified as the strongest correlated module in newly diagnosed CD [r = 0.79, adjusted p = 9E-06]. Hub genes within this module included *MYOF* and *SETD7*. Among the probe sets present in this module, 30 were dysregulated in newly diagnosed CD as compared with controls [e.g. *SLC39A8*]. The grey60 module [49 probe sets] was the strongest correlated module in late CD, including hub genes *APOM* and *TMEM252*; from this, 37 gene probe sets were dysregulated in late CD [e.g. *CDHR1*] [Supplementary Table S8].

Fourteen modules were correlated with all three traits, including the top modules mentioned above [Figure 3]. Other overlapping modules included magenta [e.g. *MUC1, MUC4, LCN2*], red [e.g. *S100A8, PTGS2, PRDM1, SLC7A11*], and greenyellow [*WNT5A, ANXA1*]. In addition, most of the overlapping modules [9 out of 14] followed a specific trend across the different CD groups, i.e. a high correlation in post-operative recurrence, even higher in newly diagnosed CD, and the highest in late-stage CD [Figure 3]. Pathways enriched for positively correlated [$r > 0$] modules, and following this trend [modules royalblue, greenyellow, red, white, pink, cyan, tan], were [a]granulocyte adhesion and diapedesis, and fibrosis; while associated biological [sub]functions included cellular movement [e.g. chemotaxis], cellular development [e.g. proliferation of mononuclear leukocytes], and cellular growth and proliferation [e.g. proliferation of fibroblasts]. Modules negatively correlating [$r < 0$] with each of the outcomes, and with increasing correlation across the different CD stages [grey60, blue], consisted of genes enriched for cholesterol biosynthesis; and biological functions included molecular transport and lipid metabolism.

Five modules were correlated with only one of the three traits [Figure 3]: the midnightblue module was uniquely correlated with post-operative recurrent CD [r = –0.43, adjusted p = 0.03] and it contained only snoRNAs. The modules salmon [r = 0.49, adjusted p = 2E-02] and lightgreen [r = –0.44, adjusted p = 4E-04] were specifically correlated with newly diagnosed CD, while orange [r = 0.59, adjusted p = 6E-03] and darkred [r = 0.54, adjusted p = 1E-02] were only correlated with late-stage CD. Modules uniquely correlated with newly diagnosed and late-stage CD [black, purple, turquoise] were mainly enriched for melatonin degradation.

3.6. Differential miRNA expression analysis

In post-operative recurrent CD, 13 miRNAs [7 up and 6 down] gave significantly different signals [Table 3] compared with non-IBD controls. In contrast, five and one miRNA[s] were differentially expressed in newly diagnosed and late-stage CD, respectively [Figure 2B]. One dysregulated miRNA was common to post-operative recurrent CD and newly diagnosed CD: hsa-miR-21-5p. The only significantly different miRNA in late CD, hsa-miR-223-5p, was also dysregulated in newly diagnosed CD.

3.7. Integrated analysis of gene and miRNA expression in post-operative recurrent CD

3.7.1. Inverse expression profiles

Next, we integrated differentially expressed genes and miRNAs in post-operative recurrent CD: 66 miRNA–mRNA pairs with inverse expression profiles were identified, comprising nine different miRNAs, and 57 different mRNAs [Figure 4; Supplementary Table S9]. The mRNA targets of these pairs encode for proteins mainly involved in cellular development, growth and proliferation, cell death and survival, and hematological system development and function. Of the 66 miRNA–mRNA pairs, five pairs were experimentally supported in literature [hsa-miR-196a-5p is known to target *ANXA1*; hsa-let-7g-5p targets *PRDM1* and *PTGS2*; and hsa-miR-30d-5p targets *SLC7A11* and *WNT5A*], seven pairs had a predicted repression to 40% of the 'normal' level, and 54 pairs had a predicted repression to 65% of the 'normal' level. The three miRNAs of all experimentally supported pairs were validated using qRT-PCR, together with their target mRNAs. We could confirm all five inverse expression profiles [Supplementary Figure S10].

3.7.2. Modules enriched for target genes of differentially expressed miRNAs

Of the 57 mRNAs targeted by dysregulated miRNAs, 41 were assigned to modules blue, darkgrey, darkturquoise, greenyellow, grey60, lightcyan, magenta, pink, red, tan or yellow. All these modules were correlated with post-operative recurrent CD [Figure 5; Supplementary Table S10]. The modules red and magenta were the most enriched for these dysregulated mRNAs [8. 8% and 8.3%, respectively] [Figure 5; Supplementary Table S10]. Genes in the module red [e.g. *PRDM1, PTGS2, SLC7A11, S100A8*] were all targeted by hsa-let-7g-5p and/or hsa-miR-30d-5p, except for one pair [hsa-miR-10b-5p/*CXCR2*]. In addition, genes belonging to the module magenta were all targeted by hsa-let-7g-5p and/or hsa-miR-629-5p [e.g. *MUC4*], except for one pair [hsa-miR-10b-5p/*RNF186*]. Pathways enriched for both modules included [a] granulocyte adhesion and diapedesis, and acute phase response signaling.

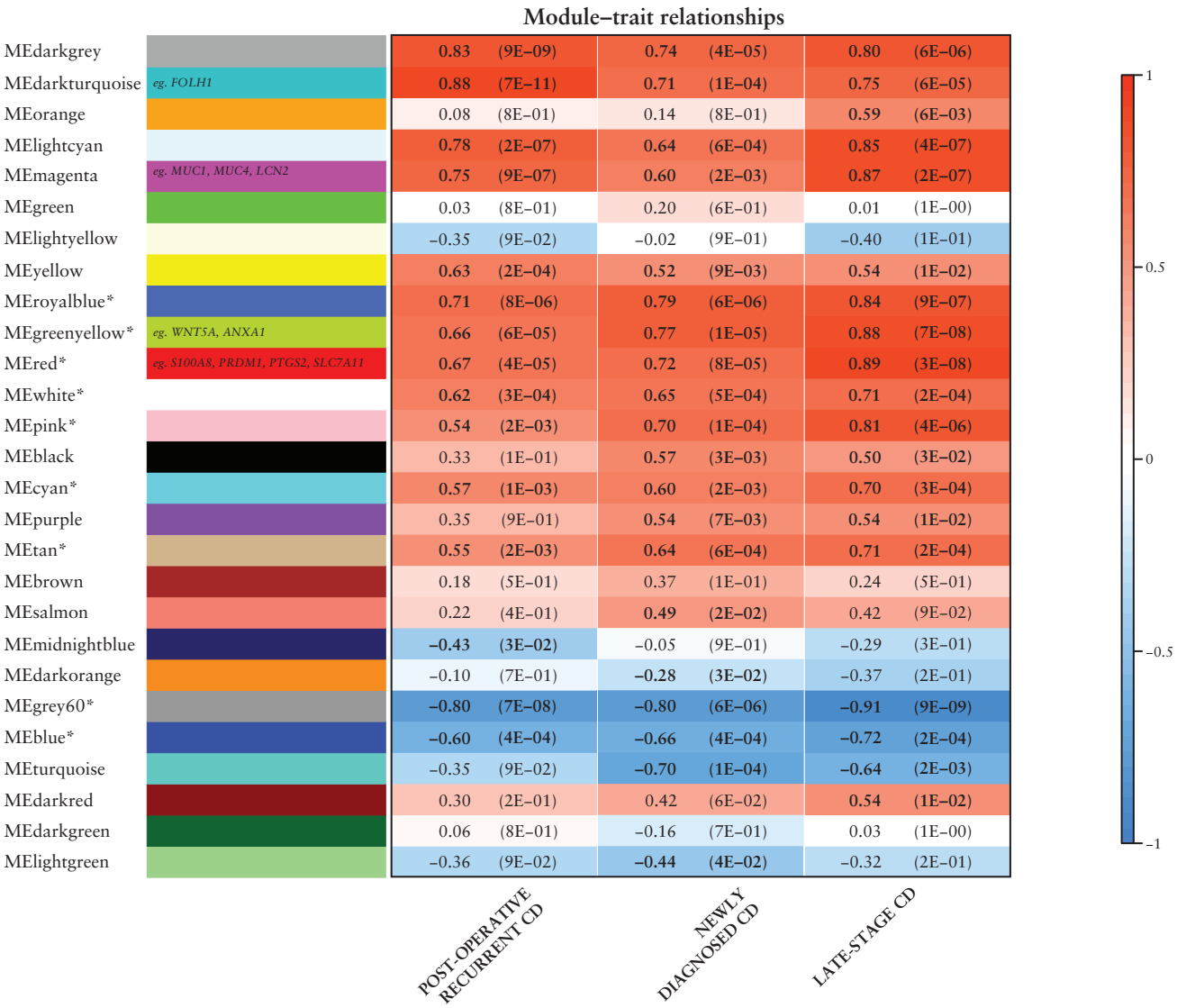

Figure 3. Co-expression network analysis: module–trait relationships. WGCNA on ileal mucosal biopsies of post-operative recurrent, newly diagnosed and late-stage CD, and controls. Correlation strengths *r* of each module were calculated for each CD subgroup with adjusted *p*-values in brackets. All significant correlations are indicated in bold. *Modules correlated with all three traits, which follow a spectrum across the different CD groups, i.e. a high correlation in post-operative recurrence, an even higher correlation in newly diagnosed CD, and the highest correlation in late-stage CD. CD, Crohn's disease; ME, module eigengene; *r*, correlation; WGCNA, Weighted Gene Correlation Network Analysis.

4. Discussion

Effective intervention before bowel damage occurs is required in order to optimize therapeutic outcomes in CD. To this end, defining the earliest lesions in CD is pivotal. Our study is the first to integrate whole-genome gene and miRNA expression profiles in ileal mucosal biopsies from CD patients with post-operative recurrence, the model that we used to study the earliest mucosal lesions in CD. Expression was compared with that in the ileal mucosa of newly diagnosed CD patients, and those with a disease duration of at least 10 years and having active disease at the time of sampling [late-stage CD]. This allowed us to see which genes and miRNAs remain dysregulated throughout different stages of disease.

In post-operative CD patients without recurrence [Rutgeerts score i0], only one differentially expressed gene was found. This confirms that previous disease does not influence baseline ileal gene expression in the neo-terminal ileum, and that the post-operative recurrence model can be used to study what happens in the early stages of CD pathogenesis. In addition, in newly diagnosed and late-stage CD, many more mRNAs were dysregulated than in post-operative recurrent CD. This was also seen in the co-expression analysis with [1] more modules significantly correlated with newly diagnosed and late CD; and [2] most of the modules correlated with all three CD stages showed a specific trend, i.e. the highest correlation in late-stage CD, followed by newly diagnosed, and then post-operative recurrent CD. These findings suggest that ileal resection with ileo-colonic anastomosis might reset ileal CD at the mucosal level, resulting in fewer genes being dysregulated in post-operative recurrent CD. Pathways enriched for in positively correlated modules following this trend were [a]granulocyte adhesion and diapedesis, and fibrosis. Interestingly, gene expression profiles of inflamed ileal biopsies of newly diagnosed pediatric CD patients [RISK cohort] were also enriched for [a]granulocyte adhesion and diapedesis.[25]

Table 3. Differentially expressed miRNAs in post-operative recurrent, newly diagnosed and late-stage CD as compared with controls

miRNA	Post-operative recurrent CD vs controls		Newly diagnosed CD vs controls		Late-stage CD vs controls	
	FC	Adjusted *p*	FC	Adjusted *p*	FC	Adjusted *p*
hsa-miR-345-5p	2.10	1.87E–03	–	–	–	–
hsa-miR-1307-5p	1.88	1.87E–03	–	–	–	–
hsa-miR-18a-5p	1.80	1.80E–03	–	–	–	–
hsa-miR-502-3p	1.73	1.94E–03	–	–	–	–
hsa-miR-182-5p	1.61	1.21E–02	–	–	–	–
hsa-miR-21-5p	1.60	1.54E–02	1.94	1.09E–03	–	–
hsa-miR-130b-5p	1.59	3.48E–02	–	–	–	–
hsa-miR-30d-5p	–1.53	1.21E–02	–	–	–	–
hsa-let-7g-5p	–1.73	1.21E–02	–	–	–	–
hsa-miR-10a-5p	–2.29	1.94E–03	–	–	–	–
hsa-miR-629-5p	–2.36	3.48E–02	–	–	–	–
hsa-miR-10b-5p	–4.15	1.87E–03	–	–	–	–
hsa-miR-196a-5p	–6.67	1.65E–02	–	–	–	–
hsa-miR-223-5p	–	–	4.03	2.25E–04	2.56	2.80E–04
hsa-miR-1246-5p	–	–	2.01	7.90E–04	–	–
hsa-miR-30c-5p	–	–	–1.54	1.76E–03	–	–
hsa-miR-378-3p	–	–	–1.94	1.75E–03	–	–

CD, Crohn's disease; FC, fold change; vs, versus.

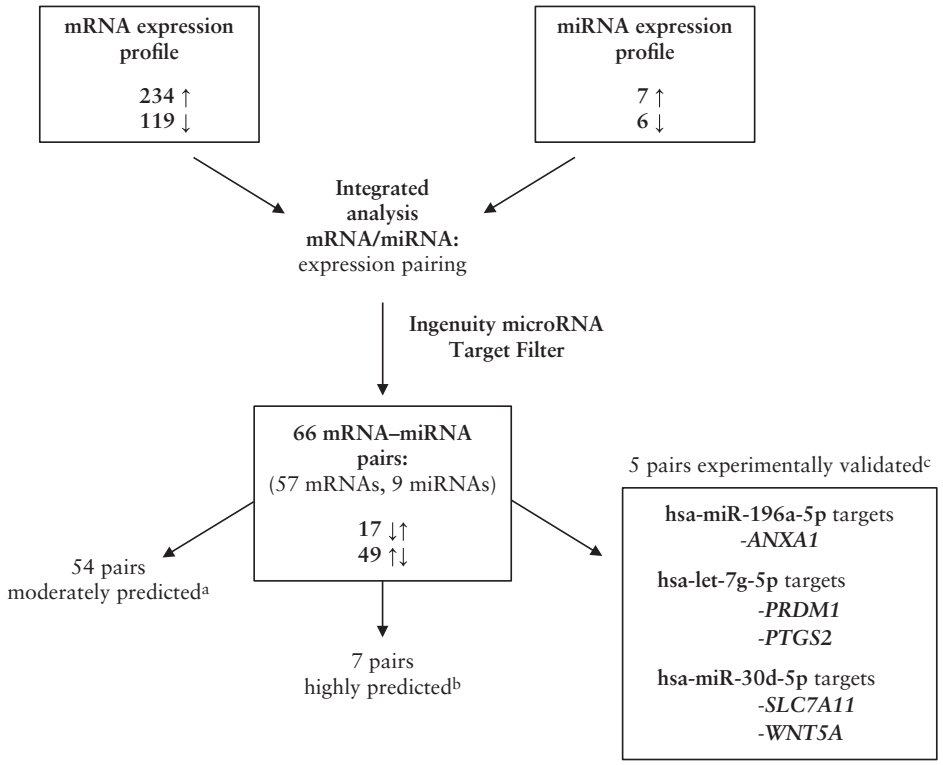

Figure 4. Integrated analysis of miRNA and mRNA expression profiles in ileal mucosal biopsies from post-operative recurrent CD patients compared with controls using the Ingenuity microRNA Target Filter. [a] The miRNA is predicted to repress the expression of its mRNA target to 65% of the 'normal' level with consultation of TargetScan; [b] the miRNA is predicted to repress the expression of its mRNA target to 40% of the 'normal' level with consultation of TargetScan; [c] indicates that the miRNA–mRNA pairs were experimentally supported in the literature using Tarbase, miRecords and Ingenuity Expert findings. CD, Crohn's disease.

Modules negatively correlating across the different CD stages were associated with cholesterol biosynthesis. Indeed, hypocholesterolemia is observed in CD patients, which is influenced by inflammatory status, impaired intestinal adsorption, and bowel resection.[26]

In addition, one module, containing only snoRNAs, was uniquely correlated with recurrent CD, whereas modules only correlated with newly diagnosed and/or late-stage CD were enriched for melatonin degradation. These modules were negatively correlated with both

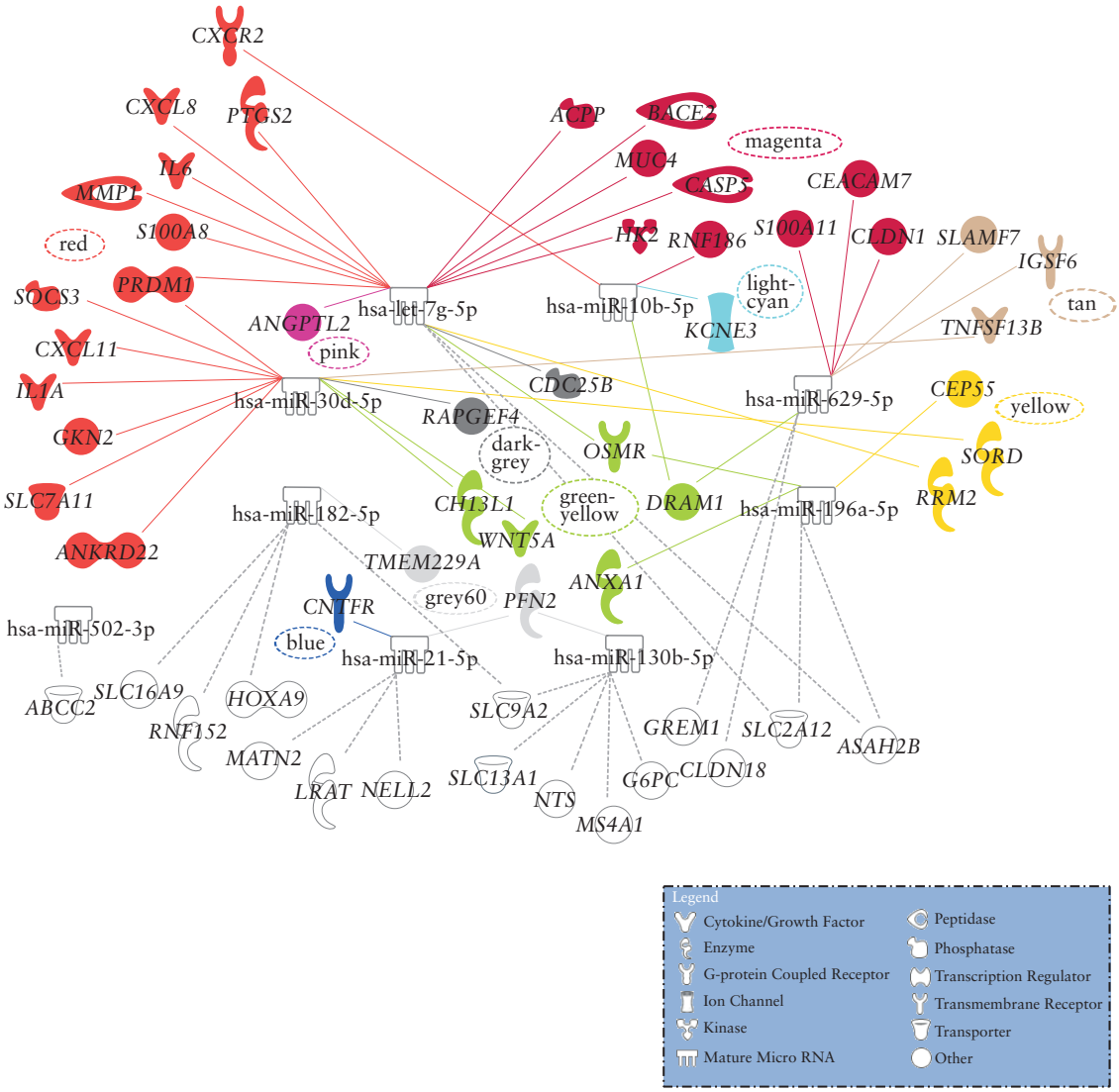

Figure 5. Regulatory network of inversely correlated miRNAs and mRNAs in post-operative recurrent CD. Differentially expressed mRNAs targeted by differentially expressed miRNAs are represented, and coloured according to the corresponding co-expression module [red, magenta, greenyellow, tan, yellow, tan, grey60, darkgrey, blue, lightcyan] if applicable. Uncoloured mRNAs do not belong to a module or belong to the module grey [i.e. non-correlated genes that cannot be clustered into one of the modules].

traits, which might suggest a decrease in melatonin degradation, leading to anti-apoptotic, anti-oxidative and anti-inflammatory activities.[27] This might be associated with more extensive bowel damage observed in newly diagnosed and late-stage disease patients, compared with post-operative recurrent lesions. Again, we could confirm previous findings of the RISK cohort, as melatonin degradation also belonged to the top enriched pathways.[25]

We identified interesting dysregulations in early post-operative recurrent CD [see Summary Figure 6 for an overview]. *FOLH1* was the most significantly differentially expressed gene in both post-operative recurrence and newly diagnosed CD, and ranked 27th in late-stage CD. It encodes a glutamate carboxypeptidase that acts on substrates such as folate.[28] This gene was also identified as the most dysregulated gene in ileal mucosal biopsies from CD patients with a mean duration of disease of 6 years.[29] Immunohistochemical staining localized FOLH1 expression in ileal post-operative recurrent CD mostly to the surface epithelium. Zhang et al. reported a similar FOLH1 expression pattern in the villous epithelium of unaffected ileal CD.[30] We also found several mucosal barrier genes to be dysregulated in post-operative recurrent CD. Impaired barrier function plays an important role in IBD, but there is debate whether this is a primary or secondary [to gut inflammation and damage] event.[31] A dysregulated *MUC1* and *MUC4* expression in early recurrent CD confirms our earlier findings of a persistently increased *MUC1* and *MUC4* expression in uninflamed CD ileum.[32] Both our previous and current findings imply an early, primary mucosal barrier defect in CD and support the development of barrier-enhancing therapeutic agents. Likewise, several AMPs were found to be dysregulated IBD.[11,33] LCN2, a bacteriostatic protein, has been found as a serological disease activity marker,[34] and as a fecal biomarker in IBD.[35] We detected a clear LCN2 expression in the epithelium and in inflammatory cells of post-operative recurrent mucosa, which is in accordance with the study of Thorsvik et al.[36] S100A8, a calcium-binding protein, forms a heterodimer [calprotectin] with S100A9, and fecal calprotectin has been used to detect intestinal inflammation.[37] We detected S100A8 in neutrophils and in endothelium of

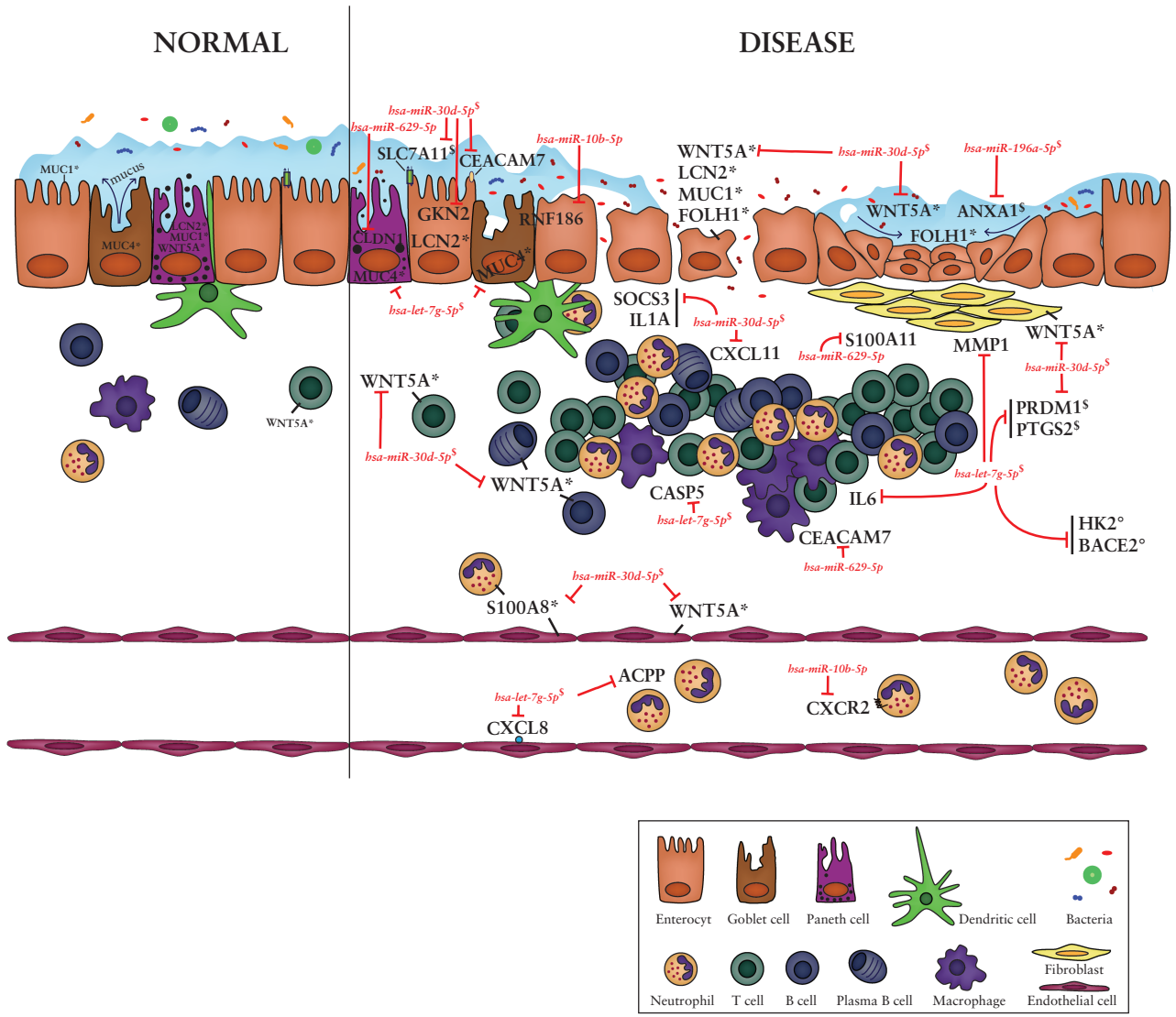

Figure 6. Schematic overview of some of the major players identified in post-operative recurrent CD. Several of the dysregulated genes observed in post-operative recurrent CD [e.g. *WNT5A, FOLH1, LCN2, MUC1*] show increased expression in damaged epithelium, and/or are important in mucosal regeneration [e.g. *ANXA1, FOLH1* and *WNT5A*]. The crossroads between the innate and adaptive immune system are also involved in the early CD stages, with among others dysregulated expression of *S100A8, PTGS2, PRDM1* and *WNT5A*, of which in addition the latter three are all targets of dysregulated miRNAs in post-operative recurrent CD. Major players represented are [1] genes validated using qRT-PCR and of which the corresponding protein was localized using immunohistochemistry [*]; [2] identified gene/miRNA pairs, which are experimentally supported in the literature and therefore validated using qRT-PCR [$]; and/or [3] genes belonging to modules enriched for target genes of differentially expressed miRNAs [modules red and magenta]. For genes/miRNAs not satisfying [1], the position relative to the different cell types was based on their biological function reported in the literature; except for two genes [°]. Gene expression level is indicated by font size. CD, Crohn's disease; miRNA, microRNA.

post-operative recurrent mucosa. Indeed, in inflamed tissues, the S100A8/A9 complex is deposited onto the endothelium associated with extravasating leukocytes.[38] Based on the microarray data, we found an aberrant *WNT5A* expression in uninflamed post-operative ileum [i0]. Although this finding could not be confirmed with qRT-PCR [probably because of the small sample size or because different biopsies from i0 patients were used for microarray and qRT-PCR validation], its potentially important role in early CD was also demonstrated by the integrated analysis, showing that *WNT5A* was targeted by one of the dysregulated miRNAs. Wnt signaling pathway plays a pivotal role in the gut mucosal homeostasis and therefore in the intestinal epithelium.[39] WNT5A, a non-canonical ligand, plays a critical role from embryonic development to the maintenance of post-natal homeostasis.[40] Its role in the context of IBD is unclear: WNT5A might be involved in intestinal crypt regeneration,[41] but also promotes Th1 differentiation in a colitis mouse model.[42] In resected ileal mucosa from active CD, we detected WNT5A protein expression in damaged and regenerating epithelium; as well as in B and T cells, fibroblasts, giant cells, and endothelial cells. In a series of ileal biopsies from post-operative CD patients, similar findings were observed: when damaged/regenerating epithelium or increased inflammatory infiltrate was present, WNT5A levels were strongly increased. These findings are in accordance with previously reported studies,[41–43] and complement WNT5A observations in other inflammatory diseases [sepsis, atherosclerosis, rheumatoid arthritis, psoriasis].[44]

At the miRNA level, the opposite was observed: the most dysregulated miRNAs were found in post-operative recurrent CD [13], then newly diagnosed [5], and last late-stage CD [1];]; suggesting that miRNA dysregulation plays an important role at the mucosal level after resetting the disease by surgery, and therefore potentially in the early stage of CD. Gene co-expression analysis complemented these findings, i.e. one module containing only snoRNAs was only correlated with post-operative recurrence. These small non-coding snoRNAs influence mRNA splicing and serve as miRNA-precursors,[45] giving further evidence for an important regulating function of small non-coding RNA in early CD.

The small number of dysregulated miRNAs is in accordance with a previous report concerning miRNA profiling of the terminal ileum of chronically active CD patients. They found one miRNA with decreased expression, and six miRNAs, including hsa-miR-21 and hsa-miR-223, with increased expression in Crohn's ileitis as compared with control tissues.[46] In contrast, Guo et al. found 55 differentially expressed miRNAs in inflamed CD terminal ileal mucosa. One miRNA [hsa-let-7g-5p] showed a similar decreased expression as in our post-operative recurrent CD cohort, and two [hsa-miR-21-5p and hsa-30c-5p] had comparable dysregulated expression with our newly diagnosed CD group.[47] It is unclear why the number of miRNAs differ across the reported studies and our study. However, it is unlikely that the differences result from different sample sizes because both reported studies included six subjects in each study group. In addition, different patient characteristics [e.g. therapy use, disease duration, IBD-related surgery, ethnicity] and also the type of microarray and data processing might partly explain the expression differences. Furthermore, some patient characteristics were not [e.g. smoking status] or not clearly [e.g. definitions disease activity] reported, which hampers the comparison of miRNA expression studies in ileal CD.

Integrated analysis of the dysregulated genes and miRNAs in post-operative recurrent CD found five miRNA–mRNA pairs with inverse expression profiles and with experimental support in literature, which we also validated using qRT-PCR [Summary Figure 6].[48–51] All five mRNA targets were also dysregulated in newly diagnosed and late-stage CD, and four [*PRDM1, PTGS2, SLC7A11* and *WNT5A*] were also found to be increased by Ben-Shachar et al. in ileal mucosal biopsies from late-stage CD patients.[29]

In order to identify functional networks of genes and miRNAs contributing to early CD, we used multi-level integration of gene co-expression network analysis with mRNA–miRNA pair correlation in post-operative recurrence. The modules red and magenta contained the highest proportion of dysregulated mRNAs targeted by a dysregulated miRNA, in which a central role for hsa-let-7g-5p and hsa-miR-30d-5p was observed [Summary Figure 6]. Pathways enriched for both modules included [a]granulocyte adhesion/diapedesis, and acute-phase response signaling, which are linked to CD pathogenesis.[52]

Although this study provides a better insight into the earliest mucosal CD lesions, it has some limitations. We studied different CD stages in different patient sets, not in patients that were being followed over time. Besides, studies with high-throughput methodologies applied on clinical samples frequently face confounding factors. First, mucosal biopsies contain varying proportions of different cell types, meaning that we studied averaged expressions of all cells within the biopsy. Using cell deconvolution methods may provide insight regarding the proportions of cell types within each biopsy and the role of each cell type in disease pathogenesis. In addition, other potential confounders are related to the clinical heterogeneity of both the control and different CD groups [i.e. age, disease duration, failed prior therapy, medication at time of biopsy, disease activity and inflammation status]. However, correction for age resulted in similar results to the ones presented [i.e. similar main pathways and top dysregulated genes, data not shown]; and based on hierarchical clustering analysis, we could exclude potential confounding effects of medication [data not shown]. Moreover, most of the newly diagnosed patients were treatment-naïve. Taken together, we believe the different patient groups studied were fairly comparable. Furthermore, by definition, transcriptomic analyses are descriptive in nature, and therefore further studies are warranted to determine the functional impact of these expression changes.

In conclusion, we demonstrated that early post-operative recurrent CD is a good model for studying early mucosal CD as [1] previous disease [before surgery] did not influence gene expression levels; and [2] our findings suggest that surgery may reset the disease at the mucosal site. We suggest a potentially important role for snoRNAs and miRNAs in post-operative recurrent CD, and found that dysregulated gene expression and gene co-expression networks were more pronounced in newly diagnosed and late-stage CD. Functional studies of interactions among and between the identified mRNAs and miRNAs will provide further insights into the role of these key molecules during early-stage CD, which might allow pointers for new therapies, and help to identify early diagnostic markers.

Funding

This work was supported by the Research Foundation Flanders [FWO] [no. G.0440.06, no. G.0582.08 and no. G.0479.10]. BV is a doctoral fellow, KM is a postdoctoral fellow and GVA, MF and SV are senior clinical investigators of the FWO.

Conflict of Interest

GDH received consultancy fees from Genentech, Galapagos, Centocor, Takeda, and Fast Forward Pharmaceuticals. BV received speakers fees from Ferring Pharmaceuticals, and Takeda. MV is supported by a European Research Council [ERC] Advanced Grant. GVA received financial support for research from Abbott and Ferring Pharmaceuticals; consultancy fees from PDL BioPharma, UCB Pharma, Sanofi-Aventis, Abbott, Abbvie, Ferring, Novartis, Biogen Idec, Janssen Biologics, NovoNordisk, Zealand Pharma A/S, Millenium/Takeda, Shire, Novartis and Bristol Mayer Squibb; and speakers fees from Janssen, MSD and Abbott. PR received grant support from Abbvie, Centocor, and Merck; consulting fees from Centocor, Merck, UCB, Abbvie, Millenium/Takeda, Genentech/Hoffman LaRoche, Merck/Serono, Bristol Myers Squibb, Robarts, Tillotts, Pfizer, and Falk Pharma; and speakers fees from Centocor, Merck, and Abbvi.e. MF received grant support from Janssen and Takeda; consulting fees from Abbvie, Boehringer-Ingelheim, Ferring, Janssen, and MSD; and speakers fees from Abbvie, Boehringer-Ingelheim, Chiesi, Falk, Ferring, Janssen, Mitsubishi Tanabe, MSD, Takeda, Tillotts, and Zeria. SV reports research support from Merck, Abbvie, Pfizer, Takeda, and Janssen; consultancy fees from Abbvie, MSD, Prometheus Laboratories, Tillotts Pharma, EliLilly, Pfizer, Takeda, Janssen, Celgene, Ferring Pharmaceuticals, Galagapos, Gilead, Hospira, MundiPharma, Second Genome, and Biogen; and payments for lectures/speakers bureau from Abbvie, Pfizer, Takeda, Janssen, Ferring Pharmaceuticals, Galagapos, and Hospira. For the remaining authors, no conflicts of interest were declared.

Acknowledgments

We would like to thank the patients who contributed to this study. We would also like to thank Vera Ballet for the large amount of work she did

in organizing and maintaining the Biobank database, and Isabelle Therasson for her help in collecting the samples. We acknowledge the great assistance of Nooshin Ardeshir Davani, Karolien Claes, Sophie Organe and Willem-Jan Wollants in processing all the samples.

Author Contributions

SaV contributed to the study concept and design, acquisition of data, analysis and interpretation of data, drafting of the manuscript, statistical analysis and technical support. GDH contributed to the study concept and design, analysis and interpretation of data and drafting of the manuscript, and critical revision of the manuscript for important intellectual content. JVDG and BV contributed technical support, analysis and interpretation of data, and critical revision of the manuscript for important intellectual content. MV and KM contributed technical support and critical revision of the manuscript for important intellectual content; LVL contributed technical support. FS contributed material support and critical revision of the manuscript for important intellectual content. MF and GVA contributed material support, and critical revision of the manuscript for important intellectual content. PR contributed to the study concept and design, acquisition of data, material support, critical revision of the manuscript for important intellectual content, and obtained funding and study supervision. SV contributed to the study concept and design, acquisition of data, analysis and interpretation of data, material support, and obtained funding and study supervision. IA and IC contributed equally to the article. IA contributed to the study concept and design, acquisition of data, analysis and interpretation of data, drafting of the manuscript, and obtained funding. IC contributed to the study concept and design, acquisition of data, analysis and interpretation of data, material support, drafting of the manuscript, and study supervision.

Supplementary Data

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

References

1. Molodecky NA, Soon IS, Rabi DM, *et al.* Increasing incidence and prevalence of the inflammatory bowel diseases with time, based on systematic review. *Gastroenterology* 2012;**142**:46–54.e42; quiz e30.
2. de Souza HS, Fiocchi C. Immunopathogenesis of IBD: current state of the art. *Nat Rev Gastroenterol Hepatol* 2016;**13**:13–27.
3. Khanna R, Bressler B, Levesque BG, *et al.*; REACT Study Investigators. Early combined immunosuppression for the management of Crohn's disease [REACT]: a cluster randomised controlled trial. *Lancet* 2015;**386**:1825–34.
4. D'Haens G, Baert F, van Assche G, *et al.*; Belgian Inflammatory Bowel Disease Research Group; North-Holland Gut Club. Early combined immunosuppression or conventional management in patients with newly diagnosed Crohn's disease: an open randomised trial. *Lancet* 2008;**371**:660–7.
5. Torres J, Burisch J, Riddle M, Dubinsky M, Colombel JF. Preclinical disease and preventive strategies in IBD: perspectives, challenges and opportunities. *Gut* 2016;**65**:1061–9.
6. Frolkis AD, Dykeman J, Negrón ME, *et al.* Risk of surgery for inflammatory bowel diseases has decreased over time: a systematic review and meta-analysis of population-based studies. *Gastroenterology* 2013;**145**:996–1006.
7. Rutgeerts P, Geboes K, Vantrappen G, Kerremans R, Coenegrachts JL, Coremans G. Natural history of recurrent Crohn's disease at the ileocolonic anastomosis after curative surgery. *Gut* 1984;**25**:665–72.
8. Rutgeerts PJ. From aphthous ulcer to full-blown Crohn's disease. *Dig Dis* 2011;**29**:211–4.
9. Rutgeerts P, Goboes K, Peeters M, *et al.* Effect of faecal stream diversion on recurrence of Crohn's disease in the neoterminal ileum. *Lancet* 1991;**338**:771–4.
10. D'Haens GR, Geboes K, Peeters M, Baert F, Penninckx F, Rutgeerts P. Early lesions of recurrent Crohn's disease caused by infusion of intestinal contents in excluded ileum. *Gastroenterology* 1998;**114**:262–7.
11. Arijs I, De Hertogh G, Lemaire K, *et al.* Mucosal gene expression of antimicrobial peptides in inflammatory bowel disease before and after first infliximab treatment. *PLoS One* 2009;**4**:e7984.
12. Arijs I, De Hertogh G, Machiels K, *et al.* Mucosal gene expression of cell adhesion molecules, chemokines, and chemokine receptors in patients with inflammatory bowel disease before and after infliximab treatment. *Am J Gastroenterol* 2011;**106**:748–61.
13. Arijs I, Li K, Toedter G, *et al.* Mucosal gene signatures to predict response to infliximab in patients with ulcerative colitis. *Gut* 2009;**58**:1612–9.
14. de Bruyn M, Machiels K, Vandooren J, *et al.* Infliximab restores the dysfunctional matrix remodeling protein and growth factor gene expression in patients with inflammatory bowel disease. *Inflamm Bowel Dis* 2014;**20**:339–52.
15. Pekow JR, Kwon JH. MicroRNAs in inflammatory bowel disease. *Inflamm Bowel Dis* 2012;**18**:187–93.
16. Van der Goten J, Vanhove W, Lemaire K, *et al.* Integrated miRNA and mRNA expression profiling in inflamed colon of patients with ulcerative colitis. *PLoS One* 2014;**9**:e116117.
17. Gecse K, Löwenberg M, Bossuyt P, *et al.* Sa1198 Agreement among experts in the endoscopic evaluation of postoperative recurrence in Crohn's disease using the Rutgeerts score. *Gastroenterology* 2014;**146**:S227.
18. Peyrin-Biroulet L, Billioud V, D'Haens G, *et al.* Development of the Paris definition of early Crohn's disease for disease-modification trials: results of an international expert opinion process. *Am J Gastroenterol* 2012;**107**:1770–6.
19. Arijs I, De Hertogh G, Lemmens B, *et al.* Effect of vedolizumab [anti-α4β7-integrin] therapy on histological healing and mucosal gene expression in patients with UC. *Gut* 2018;**67**:43–52.
20. Gentleman RC, Carey VJ, Bates DM, *et al.* Bioconductor: open software development for computational biology and bioinformatics. *Genome Biol* 2004;**5**:R80.
21. Irizarry RA, Hobbs B, Collin F, *et al.* Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostatistics* 2003;**4**:249–64.
22. Smyth GK. Linear models and empirical Bayes methods for assessing differential expression in microarray experiments. *Stat Appl Genet Mol Biol* 2004;**3**:Article3.
23. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc Series B Stat Methodol* 1995;**57**:289–300.
24. Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics* 2008;**9**:559.
25. Haberman Y, BenShoshan M, Di Segni A, *et al.* Long ncRNA landscape in the ileum of treatment-naive early-onset Crohn disease. *Inflamm Bowel Dis* 2018;**24**:346–60.
26. Agouridis AP, Elisaf M, Milionis HJ. An overview of lipid abnormalities in patients with inflammatory bowel disease. *Ann Gastroenterol* 2011;**24**:181–7.
27. Lin GJ, Huang SH, Chen SJ, Wang CH, Chang DM, Sytwu HK. Modulation by melatonin of the pathogenesis of inflammatory autoimmune diseases. *Int J Mol Sci* 2013;**14**:11742–66.
28. Silver DA, Pellicer I, Fair WR, Heston WD, Cordon-Cardo C. Prostate-specific membrane antigen expression in normal and malignant human tissues. *Clin Cancer Res* 1997;**3**:81–5.
29. Ben-Shachar S, Yanai H, Baram L, *et al.* Gene expression profiles of ileal inflammatory bowel disease correlate with disease phenotype and advance understanding of its immunopathogenesis. *Inflamm Bowel Dis* 2013;**19**:2509–21.
30. Zhang T, Song B, Zhu W, *et al.* An ileal Crohn's disease gene signature based on whole human genome expression profiles of disease unaffected ileal mucosal biopsies. *PLoS One* 2012;**7**:e37139.
31. König J, Wells J, Cani PD, *et al.* Human intestinal barrier function in health and disease. *Clin Transl Gastroenterol* 2016;**7**:e196.

32. Vancamelbeke M, Vanuytsel T, Farré R, *et al.* Genetic and transcriptomic bases of intestinal epithelial barrier dysfunction in inflammatory bowel disease. *Inflamm Bowel Dis* 2017;**23**:1718–29.
33. Kim JM. Antimicrobial proteins in intestine and inflammatory bowel diseases. *Intest Res* 2014;**12**:20–33.
34. Stallhofer J, Friedrich M, Konrad-Zerna A, *et al.* Lipocalin-2 is a disease activity marker in inflammatory bowel disease regulated by IL-17A, IL-22, and TNF-α and modulated by IL23R genotype status. *Inflamm Bowel Dis* 2015;**21**:2327–40.
35. Thorsvik S, Damås JK, Granlund AV,Flo TH. Fecal neutrophil gelatinase-associated lipocalin as a biomarker for inflammatory bowel disease. *J Gastroenterol Hepatol* 2017;**32**:128–35.
36. Thorsvik S, Damås JK, Van Beelen Granlund A,Flo TH, Østvik A, Sandvik A. P-110 YI Fecal neutrophil gelatinase-associated lipocalin [NGAL] is a promising biomarker for inflammatory bowel disease and NGAL is expressed in paneth cells. *Inflamm Bowel Dis* 2016;**22**(suppl_1):S44–S44.
37. Gisbert JP, McNicholl AG, Gomollon F. Questions and answers on the role of fecal lactoferrin as a biological marker in inflammatory bowel disease. *Inflamm Bowel Dis* 2009;**15**:1746–54.
38. Robinson MJ, Tessier P, Poulsom R, Hogg N. The S100 family heterodimer, MRP-8/14, binds with high affinity to heparin and heparan sulfate glycosaminoglycans on endothelial cells. *J Biol Chem* 2002;**277**:3658–65.
39. Armbruster NS, Stange EF, Wehkamp J. In the Wnt of paneth cells: immune–epithelial crosstalk in small intestinal Crohn's disease. *Front Immunol* 2017;**8**:1204.
40. Kumawat K, Gosens R. WNT-5A: signaling and functions in health and disease. *Cell Mol Life Sci* 2016;**73**:567–87.
41. Miyoshi H, Ajima R, Luo CT, Yamaguchi TP, Stappenbeck TS. Wnt5a potentiates TGF-β signaling to promote colonic crypt regeneration after tissue injury. *Science* 2012;**338**:108–13.
42. Sato A, Kayama H, Shojima K, *et al.* The Wnt5a-Ror2 axis promotes the signaling circuit between interleukin-12 and interferon-γ in colitis. *Sci Rep* 2015;**5**:10536.
43. Kim J, Kim J, Kim DW, *et al.* Wnt5a induces endothelial inflammation via beta-catenin-independent signaling. *J Immunol* 2010;**185**:1274–82.
44. Pashirzad M, Shafiee M, Rahmani F, *et al.* Role of Wnt5a in the pathogenesis of inflammatory diseases. *J Cell Physiol* 2017;**232**:1611–6.
45. Holley CL, Topkara VK. An introduction to small non-coding RNAs: miRNA and snoRNA. *Cardiovasc Drugs Ther* 2011;**25**:151–9.
46. Wu F, Zhang S, Dassopoulos T, *et al.* Identification of microRNAs associated with ileal and colonic Crohn's disease. *Inflamm Bowel Dis* 2010;**16**:1729–38.
47. Guo Z, Wu R, Gong J, *et al.* Altered microRNA expression in inflamed and non-inflamed terminal ileal mucosa of adult patients with active Crohn's disease. *J Gastroenterol Hepatol* 2015;**30**:109–16.
48. Selbach M, Schwanhäusser B, Thierfelder N, Fang Z, Khanin R, Rajewsky N. Widespread changes in protein synthesis induced by microRNAs. *Nature* 2008;**455**:58–63.
49. Nie K, Gomez M, Landgraf P, *et al.* MicroRNA-mediated down-regulation of PRDM1/Blimp-1 in Hodgkin/Reed-Sternberg cells: a potential pathogenetic lesion in Hodgkin lymphomas. *Am J Pathol* 2008;**173**:242–52.
50. Lin J, Lwin T, Zhao JJ, *et al.* Follicular dendritic cell–induced microRNA-mediated upregulation of PRDM1 and downregulation of BCL-6 in non-Hodgkin's B-cell lymphomas. *Leukemia* 2011;**25**:145–52.
51. Chen C, Zhang Y, Zhang L, Weakley SM, Yao Q. MicroRNA-196: critical roles and clinical applications in development and cancer. *J Cell Mol Med* 2011;**15**:14–23.
52. Rogler G, Biedermann L, Scharl M. New insights into the pathophysiology of inflammatory bowel disease: microbiota, epigenetics and common signalling pathways. *Swiss Med Wkly* 2018;**148**:w14599.
53. Jostins L, Ripke S, Weersma RK, *et al.*; International IBD Genetics Consortium [IIBDGC]. Host–microbe interactions have shaped the genetic architecture of inflammatory bowel disease. *Nature* 2012;**491**:119–24.