

Genome-wide transcriptional patterns of candidate genes linked to distinctive pathogenesis in Crohn's disease

Shengbo Chen

Southern Medical University Nanfang Hospital

Zhijun Li

Southern Medical University Nanfang Hospital

Chen Li

Southern Medical University Nanfang Hospital

Yi Liu

Southern Medical University Nanfang Hospital

Fachao Zhi

Southern Medical University Nanfang Hospital

Xinmei Zhao (✉ xmzhao914@163.com)

Southern Medical University Nanfang Hospital

Research article

Keywords: Crohn's disease, microarray, DASL

Posted Date: October 3rd, 2019

DOI: <https://doi.org/10.21203/rs.2.15540/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Background: Crohn's disease (CD) is an inflammatory bowel (IBD) disease with variable and complex pathophysiology. The objective of the present study was to identify genome-wide gene expression profiles underlying the progression of CD. Methods: Surgery biopsies (n=48) were analyzed by Illumina cDNA-mediated annealing, selection, extension, and a ligation microarray process specifically designed for formalin-fixed, paraffin-embedded clinical samples. Tissue samples were collected from consecutive actively-involved and uninvolved sites from the same CD patient. The CD-involved and CD-uninvolved samples were compared with non-inflammatory bowel disease healthy controls. Results: CD patients' uninvolved sites demonstrated an average gene expression between CD-involved patients and healthy controls suggesting the CD uninvolved site was unable to restore the full healthy control phenotype. In addition, peroxisome proliferator-activated receptors (PPAR) signaling-associated genes were involved in both CD involved and uninvolved sites, showing a stepwise decrease in the gene and protein expression of PPAR and downregulation of 3-hydroxy-3-methylglutaryl-CoA synthase 2, as confirmed by gene expression analysis and immunohistochemistry. Conclusions: The results of the present study provide evidence for clear differentiation of the two states, and in some cases repeated flares appeared chronically at the previous uninvolved locus.

Background

Crohn's disease (CD), an inflammatory bowel disease (IBD) subgroup characterized by transmural inflammation in the ileum and colon, causes various symptoms including abdominal pain, diarrhea, bloody stools, and weight loss^[1]. It is a relapsing and remitting disease commonly complicated by strictures, abscesses, and fistulas^[2].

The incidence of CD has increased worldwide, especially in areas that previously had a low incidence including Asia^[3,4]. Although the pathogenesis of CD is unclear, it is known that CD is mainly caused by a combination of factors involving susceptibility genes, environment, immune disorders, and dysbacteriosis^[5,6]. Early diagnosis and early treatment are important to reduce the risk of complications and the need for surgery^[7-9]. Since early treatment relies on early diagnosis, discovering a CD-specific molecular signature involved in disease initiation and progression is the key to better understanding the pathogenesis of CD and to allow an early diagnosis.

Recent studies on genome-wide microarray gene expression profiling of intestinal mucosal biopsies have showed that immune disorder, autophages, and the proinflammatory pathway play important roles in the pathogenesis of CD^[10,11]. Interestingly, several studies have demonstrated that the molecular-level CD-uninvolved site is unable to restore a full healthy control phenotype. Distinctive genes including interleukin 8 (IL-8), SSA1, regenerating family member (REGL), domain-containing protein 2(NOD2), and autophagy-related 16 like 1(ATG16L1) have been shown to be differentially regulated in inflamed or uninvolved CD biopsies compared to normal controls^[11]. Therefore, involved biopsies and uninvolved biopsies in the same patient with CD were collected to further study the early pathogenic process of CD.

The principal objective of the present study was to confirm if the CD-uninvolved phase did not have a gene expression profile consistent with that of healthy tissue. A biologically relevant gene expression slope extending from healthy checks through to CD-uninvolved and CD-involved phases as well as important pathways differentiating CD-involved from CD-uninvolved states are identified in this study thus providing a better understanding of the molecular processes leading directly to colon tissue dysplasia and colorectal cancer (CRC).

Results

Expression profiling differentiated the CD-involved phase from the CD-uninvolved phase and controls

Surgery biopsies collected from successively active-involved and uninvolved sites at the same time in individual CD patients were analyzed by genome-wide gene expression profiling. These CD-involved and CD-uninvolved samples were compared with each other as well as with non-IBD healthy controls. Principal component analysis (PCA) was conducted using the entire gene expression data from all specimens to yield a three-dimensional plot (Fig. 1). Both CD-involved (red) and CD-uninvolved (blue) sites were clearly isolated from healthy controls (green), and grouped into two groups. The results of a full dataset (carrying large parts of genes) strongly suggested differentiation between CD-involved and CD-uninvolved disease states that were both separate from healthy controls.

Gene expression profiles comparing healthy controls with diseased tissues (Fig. 2A) were conducted directly on a gene-by-gene basis. A total of 5,329 greatly perturbed genes (3,191 upregulated genes; 2,138 downregulated genes) relevant to controls were compared with CD cases involved with controls (Fig. 2B), and 3,528 genes were significantly perturbed (1,976 upregulated genes; 1,552 downregulated genes) relative to CD-uninvolved sites by directly comparing CD-involved with CD-uninvolved sites, respectively (Fig. 2B).

Among these genes, several genes were found to overlap between the two groups; 1339 genes performed at a higher level of expression (Fig. 2C) and 894 genes had a lower level of expression (Fig. 2D) in both CD-involved group linked to the controls and the CD-involved group relative to the CD-uninvolved group, respectively. By comparing CD-uninvolved samples directly to controls, several of the 2,948 genes were significantly perturbed (1,618 upregulated genes, and 1,330 downregulated genes) and those relevant to the controls were also identified (Fig. 2B).

Signature identification confirmed that the CD-uninvolved site totally disabled recovery to a healthy phenotype

A group of statistically significant regulated genes were selected for functional enrichment analysis (FEA) to study the biological functions distinguishing CD-involved sites from CD-uninvolved sites as well as from healthy controls (Fig. 2B). These selected genes were enriched into multiple biological pathways including inflammatory, immune responses, cell adhesion, apoptosis, EMT, and PPAR signaling (Table 2). An increase in biological functions associated with cell inflammation and immune reactions was found

in both CD-involved sites and CD uninvolved sites thus demonstrating strong inflammation in the remission stage and in active CD.

The expressions of nine selected genes were analyzed using QRT-PCR in an independent cohort of patients (cohort 2, Table 1). IL8, colony-stimulating factor 3 (CSF3), and prostaglandin D2 synthase (PTGDS) genes related to inflammation and chemokine showed a pronounced upregulation in both CD-involved and CD-uninvolved samples, compared with controls. The cell adhesion genes matrix metalloproteinase 1 (MMP1), epidermal growth factor-like protein 6 (EGFL6), insulin-like growth factor 2 (IGF2), and LEP exhibited significant upregulation only in the CD-involved samples, compared with CD-uninvolved samples and controls. Gene expression of the colorectal cancer biomarker gene chordin-like 2 (CHRD2) was also confirmed by QRT-PCR in independent COHORT 2 by showing a noticeable upregulation in both CD-involved and CD-uninvolved samples. Conversely, WNK4, a gene related to cell junctions, was prominently downregulated in CD-involved and CD-uninvolved samples compared with controls (Fig. 5).

EMT marker genes distinguished CD-involved samples from CD-uninvolved samples and control tissues

The expressions of genes for biological functions related to the epithelial mesenchymal transition (EMT) pathway were found to be upregulated in CD involved samples, but not in CD uninvolved samples (Table 2). FEA and gene set enrichment analysis (GSEA) confirmed both up- and downregulated gene expressions for the EMT pathway as up- and downregulation was seen in both CD-involved samples relative to CD-uninvolved samples and controls, clearly supported by a unique pattern of enhancement as shown in Figure 3.

EMT is a process of cellular dedifferentiation allowing epithelial cells to undergo multiple biochemical changes for mesenchymal cells involved in tissue repair and pathological processes. However, EMT gene expression can also initiate metastasis in the progression of disseminated cancers^[12]. The distinctive difference of EMT genes in the molecular phenotype between CD-involved and CD-uninvolved samples is important in metastatic potential at active locations of ulceration in CD patients, which may increase the cancer risk in patients with active CD. The fundamental event in EMT conversion is the loss of E-cadherin from epithelial cells^[13]. Conversely, both initiation of cell movement and suppression of cytokeratin expression distinctive of the final stages of the epithelial to mesenchymal cell transition can be activated by vimentin^[14].

In the present study, RT-PCR in independent COHORT 2 confirmed that the expression of E-cadherin and vimentin genes was significantly deregulated in CD-involved sites, compared with CD-uninvolved sites and non-IBD controls (Fig. 3, $p < 0.05$).

CD involved and uninvolved sites showed gradually decreasing expression of PPAR signaling genes

The expression of genes related to biological functions of the peroxisome proliferator-activated receptor (PPAR) signaling pathway was found to be downregulated in CD-involved samples and CD-uninvolved

patients (Table 2). Figure 4 illustrates the gene expression of the PPAR pathway on a gene-by-gene and sample-by-sample basis. A clear gradient reduction of average gene expression in the heat map was detected from healthy controls to CD-uninvolved groups and finally to the CD-involved groups. The average gene expression for each sample (Fig. 4B) quantitatively demonstrated the relationship, and the statistical estimates (Fig. 4C) among the means of each group showed a clear distinction between CD-involved groups and controls, and between CD-uninvolved groups and controls as well in downregulation ($p < 3.8E-15$, and $p < 2.1E-05$, respectively). A significant downregulation was also found between CD-involved and CD-uninvolved groups ($p < 4.9E-09$). The CD-uninvolved group exhibited a clear intermediate phenotype between CD-involved and control groups (see Fig. 4C for gene set averages per patient).

The expression of the PPAR signaling marker gene HMGCS2 was confirmed by QRT-PCR, showing a marked downregulation in both CD involved and CD uninvolved groups, compared with controls (Fig. 4D). Immunohistochemistry (IHC) staining supported results of another independent COHORT 3 as shown in Table 1. IHC images are presented in Fig. 4F (scoring data is listed in Supplementary Table 2). Both HMGCS2 and PPAR- γ showed decreased expression among healthy tissues as well as in CD-uninvolved and CD-involved groups. The predicted levels of gene expression from both microarray and RT-PCR in CD-uninvolved samples are at an intermediate level between controls and CD-involved groups. Expression levels of HMGCS2 and PPAR- γ measured in tumors and adjacent inflamed tissue, and compared with healthy control mice, indicated significant lower expression in tumor tissues and adjacent inflamed tissues than those in controls (Fig. 4E).

Discussion

In this study the CD-uninvolved samples displayed a level of gene expression that was intermediate between controls and CD-involved samples, while gene and protein expression of PPAR markers HMGCS2 and PPAR- γ were reduced significantly in both CD-involved and uninvolved samples, as well as in tumors samples, thus indicating that the EMT pathway was specifically upregulated in the pathogenesis of CD-involved samples.

The results provide new evidence for better understanding increasing cancer occurrence risk linked to patients with severe and chronic CD [15, 16]. Chronic CD inflammation increases tissue dysplasia and metastatic potential underlying dysregulated colonic mucosal epithelial tissues.

The gene expression associated with the PPAR signaling pathway was downregulated in the samples for CD-involved and CD-uninvolved patients. The expression of many regulatory genes in lipid metabolism and insulin sensitization is governed by PPAR- γ , a nuclear receptor together with family members PPAR- α and PPAR- β/δ originally discovered in adipose tissue. Previous studies using animal models of colitis and IBD patients have also identified PPAR- γ 's role in regulating inflammation and the immune response in the colon through epithelial cells [17-19]. Furthermore, recent studies have revealed how PPAR- γ decreased expression in many types of tumors including cancers of the breast, lung, pancreas, and colon. Therefore, PPAR- γ is now known as a growth-limiting and differentiation leading factor suppressing

tumor development^[20, 21]. Moreover, naturally-occurring and synthetic PPAR- γ promotes agonist growth inhibition and apoptosis supported by our findings showing a significant decrease in gene and protein expression of PPAR markers HMGCS2 and PPAR- γ in CD-involved and uninvolved samples as well as in tumor samples.

The presence of a consecutive inflammatory state in CD-uninvolved samples is confirmed in the present study as well as in a previous study^[11] in which the gene expression of inflammation including IL8, CSF3, and PTGDS was not recovered to normal levels but remained much improved in CD-uninvolved samples. An intermediate inflammatory state in CD-uninvolved samples was also identified by other researchers^[10]. Furthermore, in addition to inflammatory pathways, a set of the important dysregulated genes in CD-involved samples were involved in multiple pathways similarly perturbed in both CD-involved and CD-uninvolved samples such as pathways responsible for the biological functions of cellular proliferation, angiogenesis, and cell junctions, and cancer-related pathways.

A set of EMT-mediated genes was dysregulated only in active CD, as supported by our findings of gene expressions in the EMT pathway that showed significant upregulation and downregulation in CD-involved samples relative to CD-uninvolved samples and controls. EMT is a process of transition activity of epithelial plasticity for all changes in cell morphology from epithelioid to a mesenchymal-like phenotype as myofibroblasts^[22]. Epithelial cells express E-cadherin, while mesenchymal cells display N-cadherin, fibronectin, and vimentin^[19, 23]. Regulation of the EMT procedure selectively losing E-cadherin and increasing vimentin in CD was validated and confirmed with QRT-PCR (Fig. 3) in the present study thus proving the EMT process occurs in CD-involved samples and distinguishes CD-involved samples from both CD uninvolved samples and controls. Since EMT plays an important role in the pathogenesis and invasion of colon cancer^[24, 25], the CD involved site which exhibits the active EMT process may increase the risk of colon cancer.

Conclusions

The present study provides rigorously defined whole genome expression profiles in the same patients with different phases of CD, in comparison with controls, and also confirms the occurrence of particular changes of gene expression determined from healthy controls as well as CD-uninvolved and CD-involved groups. The gene expression of multiple pathways is permanently dysregulated in CD patients thus allowing CD samples to be easily distinguished from healthy samples.

Methods

Patients and biopsy samples

All protocols were approved by the Nanfang Hospital Medical Ethics Committee, and all participating subjects provided written informed consent. All methods were performed in accordance with relevant guidelines and regulations. Experienced pathologists and physicians followed WHO diagnostic criteria as

well as the clinical disease activity index (CDAI) in diagnosing CD and measuring disease activities^[26, 27]. CD-involved and CD-uninvolved samples were collected from inflamed areas and uninfamed areas, respectively, in the same CD patients at the same time. Healthy control specimens were collected from the normal colon tissue of healthy individuals.

RNA extraction

Three 5- μ m specimens per FFPE block were applied for RNA extraction and purification using an AllPrep DNA/RNA FFPE Kit (Qiagen, Gaithersburg, MD, USA).

DASL cDNA microarray gene expression profiling and data statistical analysis

Whole genome DASL (Illumina, USA) was conducted as previously described in the study by Reddy et al (2016)^[28] and data was generated on Illumina array platforms using Genome Studio software. Raw and normalized data were accessed from the Gene Expression Omnibus (GEO) database (GSE95095). Detailed data analysis information is described in Supplementary Methods.

Quantitative reverse transcription PCR (QRT-PCR)

Using a cDNA Archive Kit (Applied Biosystems, Foster City, CA), cDNA from total RNA was used for QRT-PCR. Probes and primers were developed to adapt the shorter amplicons (average size is equal to 95 nucleotides) possibly from partially degraded FFPE RNA specimens, and were synthesized by Integrated DNA Technologies, Inc. (San Diego, CA, USA). QRT-PCR was conducted in triplicate by the SYBR Green system on a RT-PCR apparatus HT7300 (Applied Biosystems). Relative levels of gene expression were analyzed using the $2^{-\Delta\Delta C_t}$ method following the description of previous study²⁵. The ΔC_t value of each sample was estimated using GAPDH as endogenous control gene.

Azoxymethane (AOM)-dextran sodium sulfate (DSS) developing colorectal cancer (CRC) model

On day 1, balb/c mice weighed 18–20 g and aged 7-weeks were conducted i.p. injection with a single dose (10 mg/kg) of AOM. On day 7, mice were fed with 2.5% DDS for 1 week, followed by 2 weeks of regular (autoclaved) water to start a DSS cycle. Mice were continuously fed with the second and third DSS cycle. On day 70, mice were humanely killed and cut longitudinally, then cleaned with PBS. The distal half of each colon was washed with ethylene blue for counting and measuring tumors. Colons were fastened in 4% of paraformaldehyde and dyed with hematoxylin and eosin (H&E) for histology. Samples were cut longitudinally by a pathologist (blinded to the treatments received) into five sections to determine the numbers of tumors, adenomas and microadenomas.

List Of Abbreviations

Crohn's disease (CD)

Inflammatory bowel (IBD)

Regenerating family member (REGL)

Colony-stimulating factor 3 (CSF3)

Matrix metalloproteinase 1 (MMP1)

Epidermal growth factor-like protein 6 (EGFL6)

Declarations

Ethics approval and consent to participate: All protocols are approved by Nanfang Hospital Medical Ethics Committee, and all participating subjects provided written informed consent.

Consent for publication: Not applicable.

Availability of data and material: All data generated or analyzed during this study are included in this MS.

Competing interests: All authors declare there are no conflicts of interests involved.

Funding: This work was supported by National Nature Science Foundation of China (81402037), Guangdong Natural Science Foundation (2017A030313616). National Nature Science Foundation of China (81402037) played a role in designing the study and writing the manuscript. Guangdong Natural Science Foundation (2017A030313616) played a role in collection, analysis and interpretation of data.

Authors' contributions: XZ, and FZ conceived and designed the experiments; XZ, AL, ZL, and YL performed the experiments; XZ, AL, ZL, and YL analyzed the data; AL, SC, and CL contributed reagents/materials/analysis tools; and XZ and SC wrote the paper.

Acknowledgments: None.

References

- [1] Safroneeva E, Vavricka SR, Fournier N, et al. Impact of the early use of immunomodulators or TNF antagonists on bowel damage and surgery in Crohn's disease. *Aliment Pharmacol Ther.* 2015. 42(8): 977-89.
- [2] Thia KT, Loftus EV Jr, Sandborn WJ, Yang SK. An update on the epidemiology of inflammatory bowel disease in Asia. *Am J Gastroenterol.* 2008. 103(12): 3167-82.
- [3] 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(1): 46-54.e42; quiz e30.

- [4] Ng SC, Tang W, Ching JY, et al. Incidence and phenotype of inflammatory bowel disease based on results from the Asia-pacific Crohn's and colitis epidemiology study. *Gastroenterology*. 2013. 145(1): 158-165.e2.
- [5] Baumgart DC, Sandborn WJ. Crohn's disease. *Lancet*. 2012. 380(9853): 1590-605.
- [6] Zhou Y, Chen H, He H, et al. Increased *Enterococcus faecalis* infection is associated with clinically active Crohn disease. *Medicine (Baltimore)*. 2016. 95(39): e5019.
- [7] Adami HO, Bretthauer M, Emilsson L, et al. The continuing uncertainty about cancer risk in inflammatory bowel disease. *Gut*. 2016. 65(6): 889-93.
- [8] Grivennikov SI, Cominelli F. Colitis-Associated and Sporadic Colon Cancers: Different Diseases, Different Mutations. *Gastroenterology*. 2016. 150(4): 808-10.
- [9] Walters TD, Kim MO, Denson LA, et al. Increased effectiveness of early therapy with anti-tumor necrosis factor- α vs an immunomodulator in children with Crohn's disease. *Gastroenterology*. 2014. 146(2): 383-91.
- [10] Adachi M, Kurotani R, Morimura K, et al. Peroxisome proliferator activated receptor gamma in colonic epithelial cells protects against experimental inflammatory bowel disease. *Gut*. 2006. 55(8): 1104-13.
- [11] Dubuquoy L, Rousseaux C, Thuru X, et al. PPARgamma as a new therapeutic target in inflammatory bowel diseases. *Gut*. 2006. 55(9): 1341-9.
- [12] Zhao X, Fan J, Zhi F, et al. Mobilization of epithelial mesenchymal transition genes distinguishes active from inactive lesional tissue in patients with ulcerative colitis. *Hum Mol Genet*. 2015. 24(16): 4615-24.
- [13] Noble CL, Abbas AR, Lees CW, et al. Characterization of intestinal gene expression profiles in Crohn's disease by genome-wide microarray analysis. *Inflamm Bowel Dis*. 2010. 16(10): 1717-28.
- [14] Wu F, Dassopoulos T, Cope L, et al. Genome-wide gene expression differences in Crohn's disease and ulcerative colitis from endoscopic pinch biopsies: insights into distinctive pathogenesis. *Inflamm Bowel Dis*. 2007. 13(7): 807-21.
- [15] Bernstein CN, Fried M, Krabshuis JH, et al. World Gastroenterology Organization Practice Guidelines for the diagnosis and management of IBD in 2010. *Inflamm Bowel Dis*. 2010. 16(1): 112-24.
- [16] Radisky DC, Kenny PA, Bissell MJ. Fibrosis and cancer: do myofibroblasts come also from epithelial cells via EMT. *J Cell Biochem*. 2007. 101(4): 830-9.
- [17] Harvey RF, Bradshaw JM. A simple index of Crohn's-disease activity. *Lancet*. 1980. 1(8167): 514.

- [18] Kalluri R, Neilson EG. Epithelial-mesenchymal transition and its implications for fibrosis. *J Clin Invest*. 2003. 112(12): 1776-84.
- [19] Kim B, Cheon JH, Moon HJ, et al. Crohn's disease prognosis and early immunomodulator therapy: Results from the CONNECT study. *J Gastroenterol Hepatol*. 2016. 31(1): 126-32.
- [20] Egan JM. Targeting Stem Cells in Chronic Myeloid Leukemia with a PPAR- γ Agonist. *N Engl J Med*. 2015. 373(20): 1973-5.
- [21] Brabletz T, Jung A, Spaderna S, Hlubek F, Kirchner T. Opinion: migrating cancer stem cells - an integrated concept of malignant tumour progression. *Nat Rev Cancer*. 2005. 5(9): 744-9.
- [22] Glass CK, Saijo K. Nuclear receptor transrepression pathways that regulate inflammation in macrophages and T cells. *Nat Rev Immunol*. 2010. 10(5): 365-76.
- [23] Thiery JP, Acloque H, Huang RY, Nieto MA. Epithelial-mesenchymal transitions in development and disease. *Cell*. 2009. 139(5): 871-90.
- [24] Zou J, Luo H, Zeng Q, Dong Z, Wu D, Liu L. Protein kinase CK2 α is overexpressed in colorectal cancer and modulates cell proliferation and invasion via regulating EMT-related genes. *J Transl Med*. 2011. 9: 97.
- [25] Gu Y, Wang Q, Guo K, et al. TUSC3 promotes colorectal cancer progression and epithelial-mesenchymal transition (EMT) through WNT/ β -catenin and MAPK signalling. *J Pathol*. 2016. 239(1): 60-71.
- [26] Mathew CG. New links to the pathogenesis of Crohn disease provided by genome-wide association scans. *Nat Rev Genet*. 2008. 9(1): 9-14.
- [27] Vuoriluoto K, Haugen H, Kiviluoto S, et al. Vimentin regulates EMT induction by Slug and oncogenic H-Ras and migration by governing Axl expression in breast cancer. *Oncogene*. 2011. 30(12): 1436-48.
- [28] Reddy AT, Lakshmi SP, Reddy RC. PPAR γ as a Novel Therapeutic Target in Lung Cancer. *PPAR Res*. 2016. 2016: 8972570.

Tables

Table 1. Clinical and demographic characteristics of patients involved in the present study

Baseline characteristics	CD	Control
COHORT 1 (microarray analysis)		
Number	24	12
Gender (M/F)	19/5	7/5
Age (years)	40.0±14.7	40.1±9.4
Duration of disease (years)	1.5±1.7	-
Extent of disease*	17/2/5	-
Smoking (Y/N)	7/17	1/11
Medication§	7/4/2/0	-
COHORT 2 (Real-time RT-PCR)		
Number	13	10
Gender (M/F)	9/4	5/5
Age (years)	37.0±7.9	40.9±14.6
Duration of disease (years)	2.6±2.6	-
Extent of disease*	3/2/8	-
Smoking (Y/N)	3/10	2/8
Medication§	6/2/3/2	-
COHORT 3 (immunohistochemistry analysis)		
Number	15	8
Gender (M/F)	10/5	5/3
Age (years)	37.0±9.5	36.0±14.3
Duration of disease (years)	2.8±3.2	-
Extent of disease*	4/1/9	-
Smoking (Y/N)	2/13	1/7
Medication§	9/1/3/4	-

*ileum-only/colon-only/ileum+colon.

- 5-aminosalicylic acid/steroids/immunomodulation/anti-TNF

Table 2. Pathway regulations in CD patients

Pathway	Group	Benjamini or FDR q-val (Count)		
		CD involved vs CD uninvolved	CD involved vs normal	CD uninvolved vs normal
Cell adhesion	UR	1.6E-11 (143)	1.2E-20 (211)	8.6E-03 (86)
Inflammatory response	UR	2.6E-09 (79)	2.0E-21 (195)	3.1E-08 (65)
Chemotaxis	UR	7.1E-06 (43)	5.0E-12 (66)	8.7E-04 (32)
Immune response	UR	1.1E-06 (58)	1.4E-09 (80)	3.7E-03 (27)
Angiogenesis	UR	1.9E-03 (34)	3.0E-07 (53)	2.3E-03 (29)
Cell migration	UR	1.1E-03 (54)	1.5E-08 (86)	1.9E-03 (45)
Lymphocyte activation	UR	1.2E-04 (46)	4.8E-12 (76)	2.9E-02(31)
Regulation of phosphorylation	UR	8.8E-05 (86)	9.0E-05 (103)	9.5E-03 (62)
MAPKKK cascade	UR	4.7E-02(34)	1.8E-03(50)	1.5E-01(26)
T cell differentiation	UR	2.2E-02 (15)	7.5E-05 (27)	2.6E-02(15)
Apoptosis	UR	3.6E-04 (64)	1.7E-07 (188)	8.9E-03 (96)
Leukocyte proliferation	UR	1.9E-03 (16)	2.8E-05 (22)	6.7E-02(11)
Regulation of cell growth	UR	1.1E-03 (42)	2.3E-05 (89)	6.2E-02 (29)
Cell cycle	UR	1.0E-01(28)	4.0E-02(38)	-

EGF	UR	1.1E-04 (46)	7.1E-07 (60)	3.2E-02(30)
TGF beta receptor signaling pathway	UR	0.028 (30)	0.003 (28)	0.13 (17)
NF-kB	UR	0.006 (33)	2.2E-04 (32)	6.0E-01 (14)
EMT_UP	UR	0.010 (17)	2.4E-04 (21)	0.062 (17)
E-cadherin signaling	DR	0.006(236)	0.005 (256)	0.482 (152)
Tight junction	DR	0.031(230)	0.030 (228)	0.365(128)
PPAR alpha	DR	0.478 (175)	0.012 (135)	0.0006 (78)
PPAR gamma	DR	0.473 (60)	0.007 (48)	0.003 (32)
ERBB2	DR	0.238 (203)	0.010 (213)	0.106 (119)
ERK1	DR	0.029 (312)	0.017 (330)	0.003 (198)
P53	DR	0.062 (499)	0.014 (531)	0.057 (350)
BRCA1	DR	0.043 (166)	0.014 (233)	0.030 (141)
ABC transporter	DR	0.048 (283)	0.042 (268)	0.039 (167)

Figures

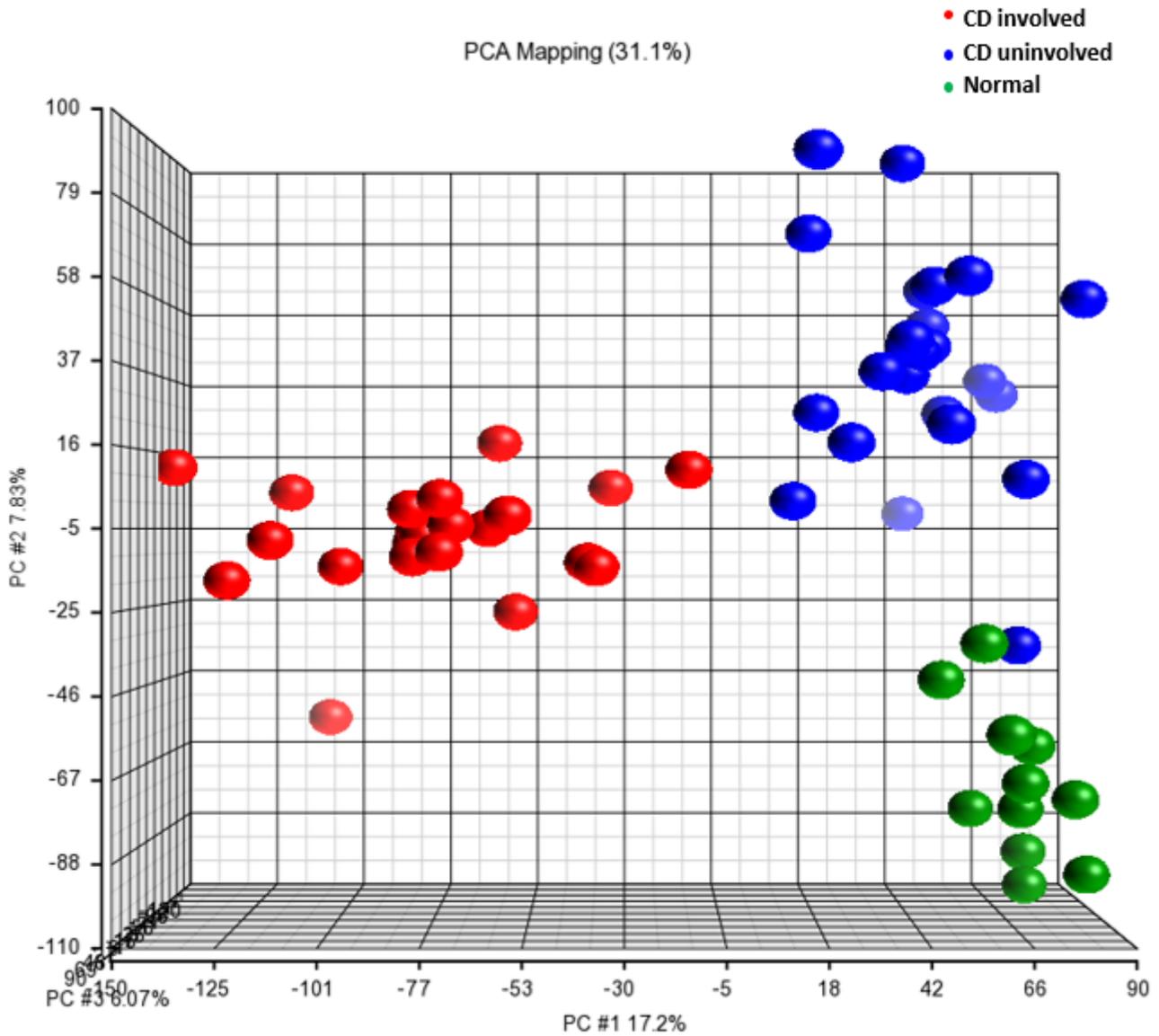


Figure 1

PCA of genome-wide microarray gene expression profiles from intestinal specimen. CD-involved (red, n=24) and CD-uninvolved (blue, n=24) samples were taken from actively-involved and uninvolved locations at the same time consecutively in CD patients, and controls (green, n=12) collected from healthy individuals. All microarray data captured from every sample were firstly normalized and then shown in a three-principal component plot where the first component (along the X-axis) shows separation between CD-involved and CD-uninvolved samples, and the second and third component (along the Y- and Z-axes, respectively) indicate a split between CD-uninvolved samples and controls.

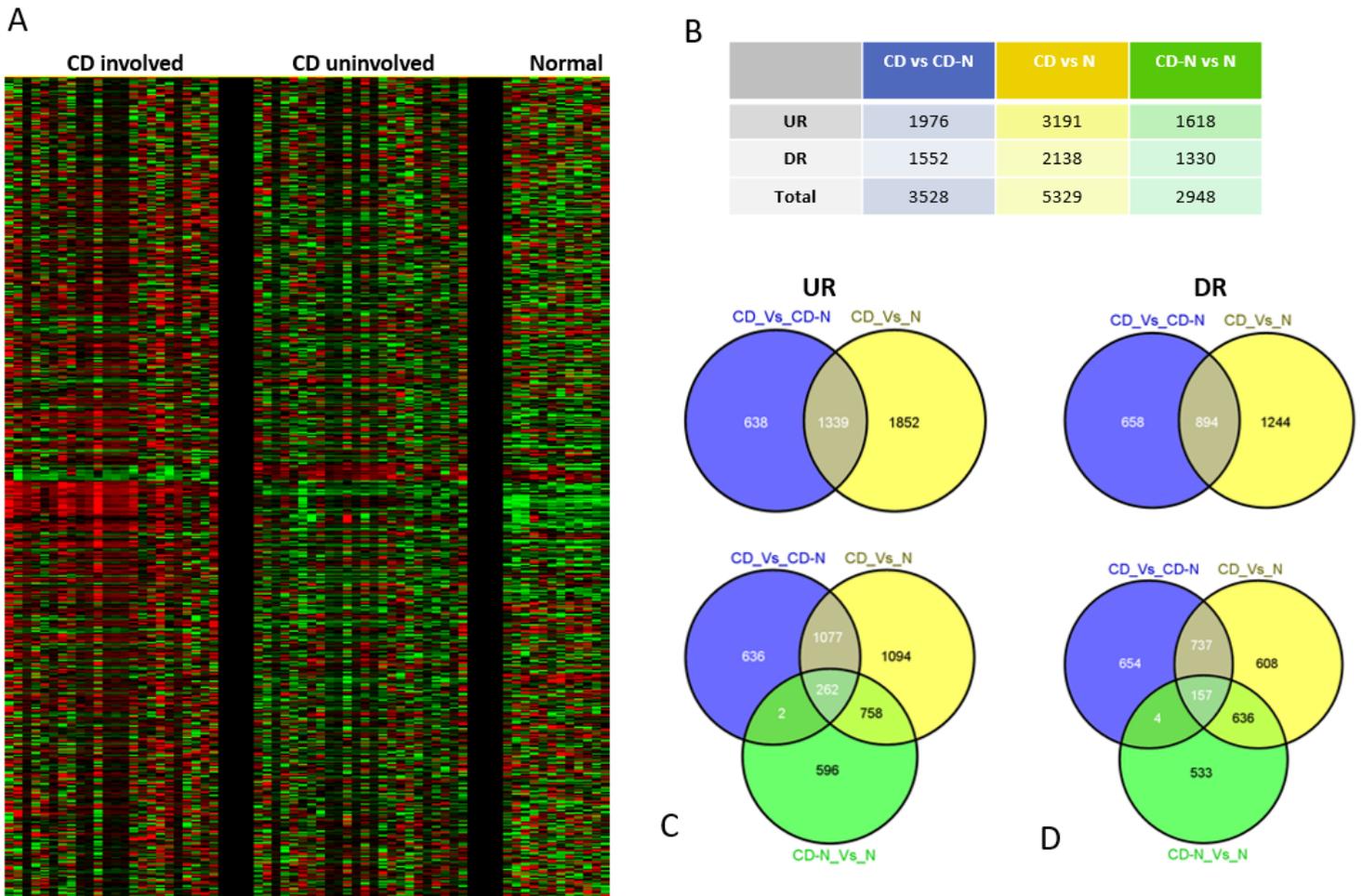


Figure 2

Transcriptional patterns of gene expression distinguishing between the CD-involved group, CD-uninvolved group, and healthy controls. (A) Heatmap of all normalized data from microarray analysis, every row is presented as an individual probe and every column shows a testing sample (red indicates high expression and green indicates low expression). The genes were ordered by unsupervised hierarchical clustering. Intestinal samples are from patients with CD-involved (n=24) and CD-uninvolved (n=24) samples as well as from controls (n=12). (B) Comparisons of number of significantly differential-regulated genes between up-(UR) and downregulated (DR) groups in CD-involved samples, CD-uninvolved samples, and controls. (C, D) Venn diagrams presenting the overlap of UR and DR genes.

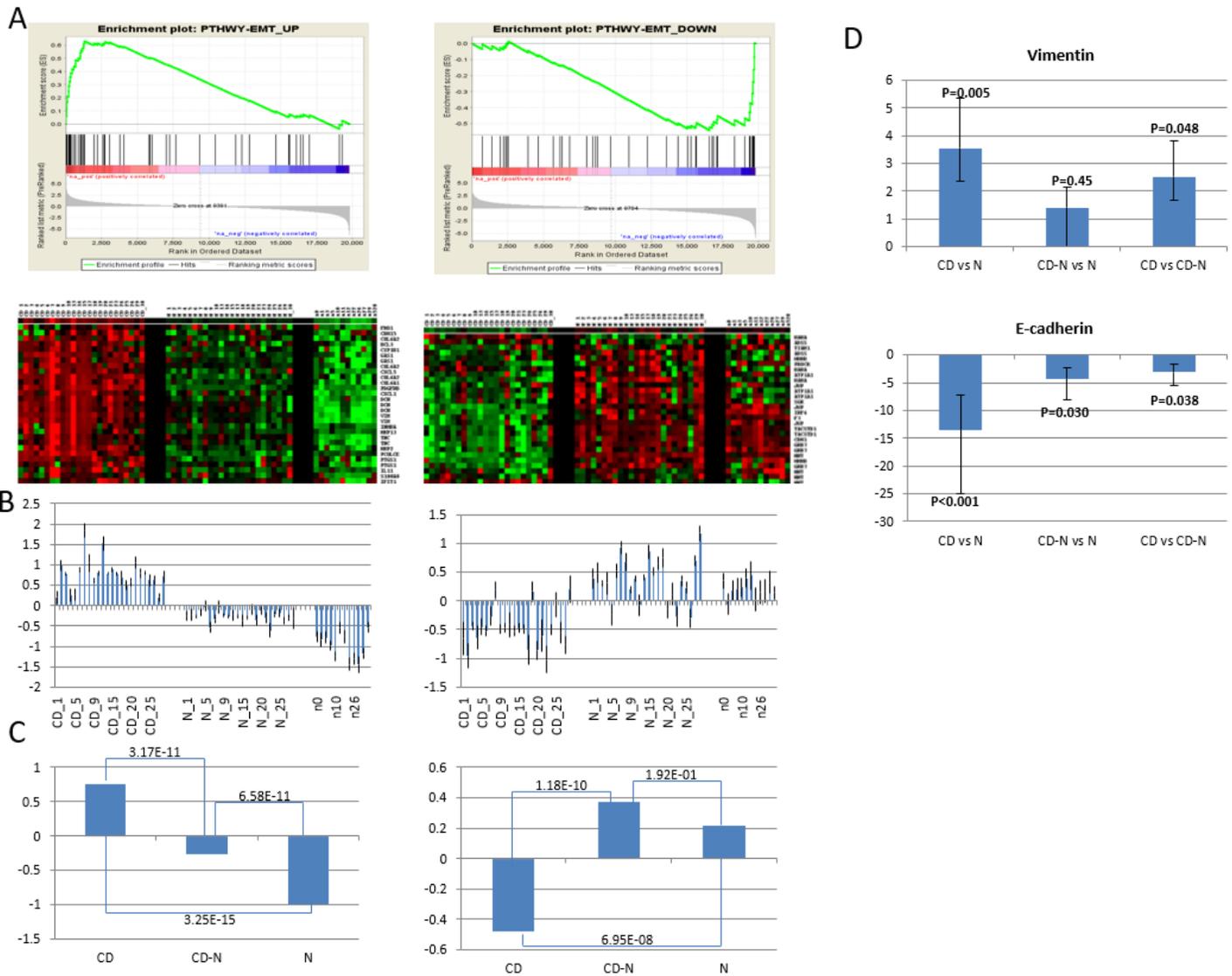


Figure 3

EMT marker genes distinguishing CD-involved samples from CD uninvolved and controls. (A) Up- and downregulated EMT relative to healthy controls are regulated in CD-involved samples, and intermediately regulated in CD-uninvolved samples. (B) Averaged EMT gene expression of each sample in every group in both UR and DR genes. (C) Averaged EMT gene expression for every phenotype in both UR and DR genes with p-values against averaged pathway gene expressions from Fig 3B. (D) QRT-PCR results of two selected genes from EMT-vimentin and E-cadherin in an independent cohort. The ratio between CD-involved samples and checks (CD vs N), CD uninvolved samples and checks (CD-N vs N), and CD-involved and CD-uninvolved samples (CD vs CD-N) are marked along with p-values for each comparison.

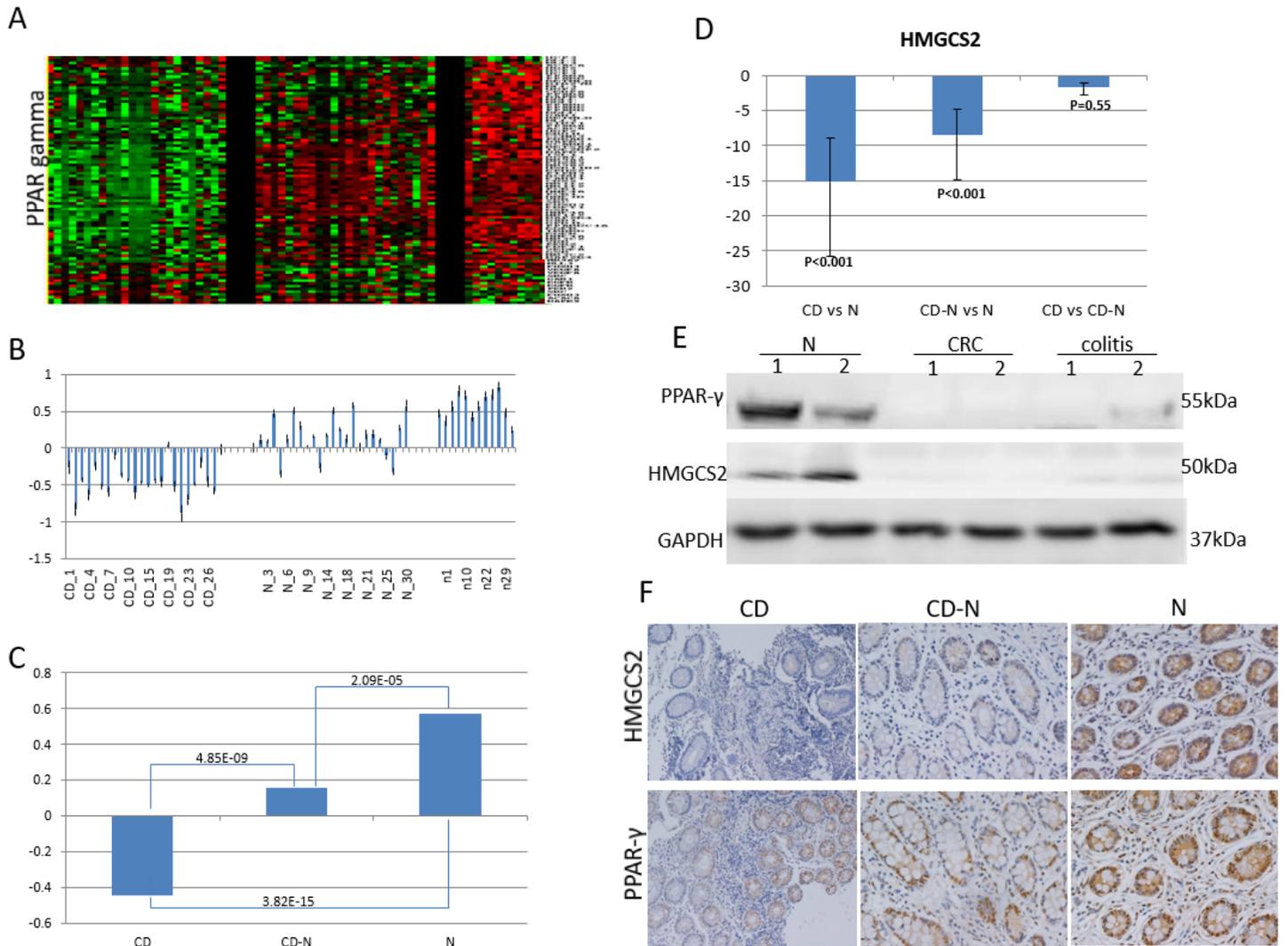


Figure 4

Expression of PPAR- γ pathway genes is downregulated in CD patients. (A) Heat map of microarray gene expressions of the selected genes linked to cell adhesion presenting a decrease from healthy samples to CD-uninvolved samples to CD-involved samples. (B) Averaged PPAR- γ gene expression of every sample in each of all groups. (C) Averaged PPAR- γ gene expression of every phenotype with p-values against averaged pathway gene expression from Fig 4B. (D) QRT-PCR results of the PPAR- γ signaling marker gene, HMGCS2, in an independent cohort. (E) In AOM/DSS colitis-related cancer mice, the levels of HMGCS2 and PPAR- γ protein were significantly lower in tumor tissue and adjacent inflamed tissue compared to controls. (F) Representative IHC staining of CD-involved samples, CD-uninvolved samples, and controls. Samples were dyed with HMGCS2 and PPAR- γ principal antibodies. The fractions were counterstained with hematoxylin for viewing with a 40 \times objective lens. Sites showing positive staining with three antibodies were located in the cytoplasm of the intestinal epithelium. The aggregated scores (the portion of specimens scored positive in every phenotyping group for every individual antibody) are shown in Supplementary Table 2.

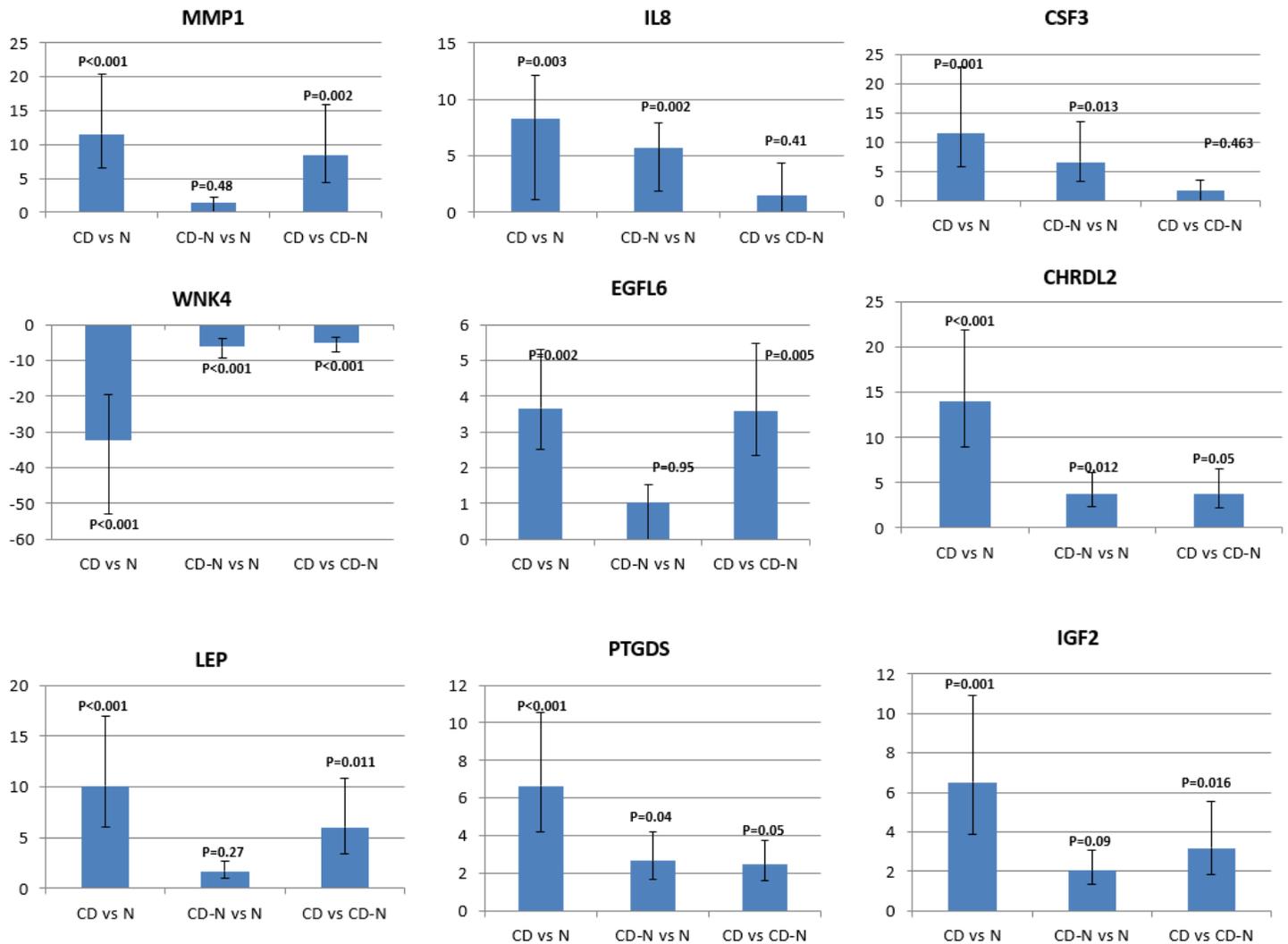


Figure 5

Expressions of genes linked to selected pathways are determined by QRT-PCR in an independent cohort 2. The ratio between CD-involved samples and checks (CD vs N), CD-uninvolved samples, and controls (CD-N vs N), and CD involved samples and CD uninvolved (CD vs CD-N) are marked along with p-values for each comparison.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Supplementaryfile.pdf](#)