PLoS MEDICINE

Dissection of the Inflammatory Bowel Disease Transcriptome Using Genome-Wide cDNA Microarrays

Christine M. Costello[1¤a], Nancy Mah[1], Robert Häsler[1], Philip Rosenstiel[1], Georg H. Waetzig[2], Andreas Hahn[3], Tim Lu[1], Yesim Gurbuz[4], Susanna Nikolaus[5], Mario Albrecht[3], Jochen Hampe[1], Ralph Lucius[6], Günther Klöppel[4], Holger Eickhoff[7¤b], Hans Lehrach[7], Thomas Lengauer[3], Stefan Schreiber[1*]

1 Institute for Clinical Molecular Biology, University Hospital Schleswig Holstein, Christian-Albrechts University of Kiel, Kiel, Germany, 2 CONARIS Research Institute, Kiel, Germany, 3 Department of Computational Biology and Applied Algorithmics, Max-Planck-Institute for Informatics, Saarbrücken, Germany, 4 Institute of Pathology, University Hospital Schleswig Holstein, Kiel, Germany, 5 Department of General Internal Medicine, University Hospital Schleswig Holstein, Kiel, Germany, 6 Institute of Anatomy, University Hospital Schleswig Holstein, Kiel, Germany, 7 Max-Planck-Institute for Molecular Genetics, Berlin, Germany

ABSTRACT

Background

The differential pathophysiologic mechanisms that trigger and maintain the two forms of inflammatory bowel disease (IBD), Crohn disease (CD), and ulcerative colitis (UC) are only partially understood. cDNA microarrays can be used to decipher gene regulation events at a genome-wide level and to identify novel unknown genes that might be involved in perpetuating inflammatory disease progression.

Methods and Findings

High-density cDNA microarrays representing 33,792 UniGene clusters were prepared. Biopsies were taken from the sigmoid colon of normal controls ($n = 11$), CD patients ($n = 10$) and UC patients ($n = 10$). ^{33}P-radiolabeled cDNA from purified poly(A)$^+$ RNA extracted from biopsies (unpooled) was hybridized to the arrays. We identified 500 and 272 transcripts differentially regulated in CD and UC, respectively. Interesting hits were independently verified by real-time PCR in a second sample of 100 individuals, and immunohistochemistry was used for exemplary localization. The main findings point to novel molecules important in abnormal immune regulation and the highly disturbed cell biology of colonic epithelial cells in IBD pathogenesis, e.g., *CYLD (cylindromatosis, turban tumor syndrome)* and *CDH11 (cadherin 11, type 2)*. By the nature of the array setup, many of the genes identified were to our knowledge previously uncharacterized, and prediction of the putative function of a subsection of these genes indicate that some could be involved in early events in disease pathophysiology.

Conclusion

A comprehensive set of candidate genes not previously associated with IBD was revealed, which underlines the polygenic and complex nature of the disease. It points out substantial differences in pathophysiology between CD and UC. The multiple unknown genes identified may stimulate new research in the fields of barrier mechanisms and cell signalling in the context of IBD, and ultimately new therapeutic approaches.

Introduction

The two main forms of inflammatory bowel disease (IBD), Crohn disease (CD) and ulcerative colitis (UC), are both characterised by an aberrant immune response of the intestinal mucosa. The current understanding of disease pathogenesis suggests a complex interplay of multiple environmental and genetic factors [1]. Although clinical, endoscopic, histopathologic, and radiologic criteria exist to distinguish CD from UC, considerable overlap is found in clinical criteria, understanding of pathophysiology, and therapy [2]. The enormous complexity of pathophysiology mandates a systematic approach to identify the molecular events that cause and perpetuate these chronic, relapsing inflammatory disorders.

In the search for genes that cause IBD, several genetic linkage studies, which identify the approximate chromosomal locations of disease susceptibility genes, have been carried out [3]. This technique, in combination with the more classical candidate gene approach, led to the identification of the first disease-associated variants in the *NOD2/CARD15* gene on chromosome 16 [4–6]. More recently, we have identified *discs, large homolog 5 (DLG5; Drosophila)* on chromosome 10q23, which encodes a scaffolding protein potentially involved in the maintenance of epithelial integrity, as an IBD susceptibility gene [7]. Concomitantly, functional variants in the *solute carrier family 22 (organic cation transporter)* genes, *SLC22A4* and *5* on chromosome 5q31, were found to be associated with CD [8]. The high number of linkage regions in IBD and the multiplicity of association findings suggest enormous complexity behind the polygenic risk between patients. Genetic susceptibility factors are therefore unlikely to serve as molecular targets for direct therapeutic application. Key molecules in pathophysiology, downstream of points of convergence between the chains of regulatory events originating from different etiologic factors, are more likely targets for successful therapeutic interventions. This can be illustrated with the example of *tumour necrosis factor-alpha (TNF-α)*, a molecule that does not appear to be causative in CD, but allows effective interruption of the inflammatory cascades through administration of a neutralizing, recombinant antibody construct [9]. The example of *TNF-α* also illustrates that a successful therapeutic approach may not target disease-specific pathophysiology, but rather molecules that are of general importance for inflammation pathophysiology, and therefore involved in a host of inflammatory disorders.

The sequencing of the human genome and the concurrent establishment of the expressed sequence tag (EST) clone database [10] have greatly improved the possibility of finding new pathophysiology-relevant genes. One technique now adopted in this latter endeavour is microarray technology, where the transcripts of thousands of genes can be simultaneously investigated. Three exploratory microarray studies carried out on intestinal mucosa samples to date broadly concur on the known genes found to be associated with IBD [11–13].

In the present study, we have used mucosal biopsies obtained by endoscopy, and not isolated cell populations, because IBD represents the rare case of a nonmalignant human disorder in which relevant disease tissue can be obtained without surgery and without any technical variance introduced by a cell isolation process. We set up a system for expression profiling using PCR-amplified cDNA clone inserts from a large whole genome collection derived from clustered EST libraries spotted on nylon filters. This genome-wide gene set was optimized for clones representing clusters from unknown ESTs and genes. Due to the clustering algorithm used, some well-known genes are therefore not represented. Preliminary experiments have demonstrated a limited overlap with the genome-wide arrays from Affymetrix with reproducible but different findings generated on both systems [14]. The two-stage design of the present experiment used a sensitive radioactive detection method for the array signals (resulting in a lower dynamic range than Affymetrix arrays) followed by real-time PCR (TaqMan) for a quantitative assessment of the changes detected. In carrying out this study, we aimed to broaden our understanding of gene regulation events in CD and UC, and identify novel genes involved in perpetuating inflammatory disease progression.

Methods

Patients

Patient group 1. This group was composed of all samples (representing 31 individuals, recruited between 1999 and 2003) analysed by cDNA microarrays (Table 1). Eleven individuals (five females [F]; mean age 51.4 y; range 25–84 y) were included in the study as the normal population, with endoscopic and histological examination yielding no significant pathological findings. Indications for colonoscopy in this group included colonic cancer surveillance and previous nonspecific changes in stool habits. To investigate gene expression in IBD, ten patients with active CD (4F; mean age 33.6 y; range 19–66 y) and ten patients with active UC (2F; mean age 34.8 y; range 27–51 y) were sequentially recruited. All endoscopic biopsies were taken from a defined area of the sigmoid colon (at 20–30 cm measured during withdrawal), and immediately snap-frozen in liquid nitrogen. A second set of biopsies was taken from the same region for histological scoring. Clinical disease activity was documented using established clinical parameters of the Crohn disease activity index (CDAI) in CD [15] and the colitis activity index (CAI) [16,17] for UC patients. Inclusion criteria were clinically (CDAI > 150 or CAI ≥ 4) and endoscopically active disease in the sigmoid colon at the time of sampling. Patients also had to be free of all medication used for the treatment of IBD (other than low-dose 5-aminosalicylic acid [5-ASA]) for a minimum of 6 wk prior to endoscopy. More than 300 patients were screened to recruit the study population.

Patient group 2. To verify and investigate the molecular epidemiology of selected signals seen in genome-wide array experiments, patient group 1 was extended from 31 to 100 individuals and analysed by quantitative real-time PCR (Table 2). IBD patients used in group 2 consisted of 27 CD patients (17F; mean age 32.1 y; range 17–45 y) and 35 UC patients (20F; mean age 33.8 y; range 18–74 y). The patients were recruited between 1999 and 2004 and were selected using similar criteria as outlined above, i.e., all patients had active inflammation in the colon and were clinically active (CDAI > 150 or CAI ≥ 4). Patients were allowed 5-ASA or glucocorticoids, but not immunosuppressants or biologicals. The normal control population consisted of 21 individuals (11F; mean age 48.9 y; range 16–84 y), who had no significant

Table 1. Clinical Characteristics of the IBD Patient Group 1

Characteristic	Crohn Disease										Ulcerative Colitis									
	CD1	CD2	CD3	CD4	CD5	CD6	CD7	CD8	CD9	CD10	UC1	UC2	UC3	UC4	UC5	UC6	UC7	UC8	UC9	UC10
Age (y)	20	38	19	32	24	66	25	44	50	18	29	38	27	31	29	38	33	51	29	43
Sex	M	F	M	F	F	M	M	M	M	F	M	M	M	M	M	F	M	M	F	M
Disease duration (y)	2	19	2	14	6	13	4	3	2	2	9	5	5	8	2	2	15	5	4	7
CDAI/CAI[a]	418	272	284	329	386	280	372	404	286	261	9	9	8	8	8	7	7	8	8	9
Acute inflammatory activity score[b]	A0	A2	A3	A1	A0	A1	A1	A1	A2	A2	A3	A1	A1	A0	A1	A2	A2	A3	A1	A3
Chronicity score[b]	C2	C2	C3	C3	C1	C1	C1	C2	C2	C2	C3	C3	C2	C1	C2	C3	C2	C3	C3	C3
Other histopathologic findings[b]	E1, L	L	E1, E2	E1, E2	E1	L		E1, L	L	P	L	E1, E2, L	E2, P		P				E1, L	P
NOD2 genotype[c]	WT	R702W	WT	WT	WT	ND	R702W	L1007fsinsC	WT	G908R	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Ileal disease	Y	Y	N	N	Y	N	N	N	Y	Y	N	N	N	N	N	Y[d]	N	Y[d]	N	N
Previous surgery	N	Y	N	N	N	Y	N	N	N	Y	N	N	N	N	N	N	N	N	N	N
Disease behaviour[e]	C	A	C	C	C	A	A	A	C	A	A	A	A	A	C	A	C	C	C	C

[a] Inclusion criteria were clinically active disease defined as CDAI > 150 or CAI ≥ 4 and inflammation present in the sigmoid colon during endoscopy. All biopsies were taken from the sigmoid colon and patients were free of all medication (other than low-dose 5-ASA) for a minimum of 6 wk prior to endoscopy.
[b] Biopsies were obtained from the same site as expression analysis samples and subjected to routine histopathology. A blinded reviewer assessed the slides. As no validated scores exist for a grading of inflammatory activity in IBD, samples were graded (scale 0–3) along the presence of acute inflammatory cells (A0–A3) and presence of a chronic infiltrate (C0–C3). Additional findings were noted, such as eosinophil-rich infiltrate (E1); eosinophilic cryptitis (E2); paneth cell metaplasia (P); and lymphoid aggregates with germinal centres (L).
[c] Patients with CD were examined for the three disease-associated NOD2/CARD15 variants (R702W, G908R, and L1007fsinsC). Four patients were heterozygote carriers. The remaining five patients were wild-type (WT) for the predisposing alleles. One patient withdrew the genotype part of the consent, so the genotype was not determined (ND) in that case.
[d] UC patients with backwash ileitis.
[e] Disease behaviour is classified as acutely active (A) if the present episode started 4 wk or less than 4 wk ago and as chronically active (C) if the present episode started more than 4 wk ago. Only one CD patient had a history of fistulating disease, and the two CD patients with a history of surgery had fibrostenosing disease. NA, not applicable; Y/N, yes/no.
DOI: 10.1371/journal.pmed.0020199.t001

pathological findings following endoscopic examination for changes in stool habits, abdominal pain, or upper gastrointestinal bleeding or cancer surveillance. As disease specificity controls, 17 patients (9F; mean age 56 y; range 17–80 y) with colonic disease but not IBD, were also included (Table 2). These disease specificity controls (DCs) included nine non-inflamed and eight inflamed biopsies samples from patients with infectious diarrhoea, other forms of gastrointestinal inflammation, or irritable bowel syndrome. No restriction on medications was specified in the inclusion criteria for DCs and all concomitant medications are listed in Table 2. A minimum of 19 normal controls, 16 UC, 19 CD, and 14 DC samples were used for each real-time PCR assay, depending on the availability of the patient samples and the plate layout at the time the real-time cDNA plates were produced.

Biopsies from all patients in the microarray study were taken from the sigmoid colon. For the real-time PCR work, biopsies were also taken from other regions of the large intestine (Table 2). We conducted extensive preparatory studies in which we systematically compared the influence of anatomical region on the expression profiles obtained. Significant differences in expression patterns between the sigmoid colon, descending colon, and transverse colon/caecum were not observed (unpublished data). Additionally, animal data has shown that different regions of the GI tract (stomach, small intestine, and colon), but not regions within the colon, were distinguishable by expression profiles [18]. We therefore considered regions within the large bowel to be equivalent and that sampling from different areas of the large bowel would have negligible effect on differential gene expression.

Informed Consent and Approval by the Ethics Committee

All patients included in this study consented to additional research biopsies being taken 24 h prior to endoscopy. One patient withdrew the genotype part of his consent. The procedures in the study protocol were approved by the Ethic Committee of the Medical Faculty of the Christian-Albrechts-University prior to the start of the study.

Target Preparation and Hybridisation to cDNA Microarrays

cDNA clone inserts from 88 microtiter plates containing 384 wells (i.e., 33,792 inserts total) from the Human UniGene set RZPD 1 clone set (German Resource Center for Genome Research [RZPD], Berlin, Germany; http://www.rzpd.de/) were amplified by PCR and spotted onto nylon filters [14]. Since the redundancy of this clone set was previously estimated to be approximately 1.44-fold [19], the 33,792 clones on the cDNA microarray represent approximately 23,000 unique transcripts. Total RNA was isolated from snap-frozen biopsies, and mRNA isolation, radioactive labelling of mRNA, and subsequent steps (hybridisation of the radiolabelled target, washing, and scanning of filters) were carried out as previously described [14].

Data Analysis

Image gridding and spot quantitation were carried out using VisualGrid software (GPC Biotech, Martinsried, Germany; http://www.gpc-biotech.com), and results were imported into a custom-made database for further analysis. The ranked expression data distribution was found to follow Zipf's law, and the log-transformed datasets were normalized

Table 2. Characteristics of Patient Group 2, Including DCs

Characteristic	Hospitalized Normal Patients	CD Patients	UC Patients	DC Patients
Number of patients	21	27	35	17
Sex (F/M)	(11F/10M)	(17F/10M)	(20F/15M)	(9F/8M)
Age (y), mean (range)	48.9 (16–84)	32.1 (17–45)	33.8 (18–74)	56 (17–80)
Region of biopsy				
Sigmoid colon	13	26	35	16
Descending colon	1	1	0	0
Caecum	7	0	0	0
Transverse colon	0	0	0	1
Inflammation activity[a]				
Active	0	27	35	8
Inactive	21	0	0	9
Activity index, mean (range)[b]	NA	300 (250–450)	8 (6–9)	NA
Diagnosis (DC patients)				
Infectious diarrhoea	NA	NA	NA	8
Gastrointestinal inflammation	NA	NA	NA	7
Irritable bowel syndrome	NA	NA	NA	2
Disease location (IBD patients only)				
Ileocaecal	NA	10	0	NA
Pancolitis	NA	11	17	NA
Distal colon	NA	6	18	NA
Fistulae	NA	5	0	NA
Stenoses	NA	4	0	NA
Medication				
No treatment	14	9	11	9
Steroids	0	11	11	1
5-ASA	0	9	18	1
NSAID	2	0	1	3
Antibiotic	1	2	0	4
Other	4	2	4	4

[a]Inflammation activity of biopsied site, as observed by the endoscopist.
[b]CDAI and CAI reported only for CD and UC patients, respectively.
5-ASA, 5-aminosalicylic acid; DC, disease specificity control patients; NA, not applicable; NSAID, nonsteroidal anti-inflammatory drug.
DOI: 10.1371/journal.pmed.0020199.t002

following this principle [20,21]. Any hybridisation signals that had a value that fell below two standard deviations of the mean background were considered as indistinguishable from background and were set to zero. In the few cases where duplicate expression values for the same genes were not similar (outliers or false positives), all readings for this gene were eliminated. The entire microarray dataset was not normally distributed and, therefore, in order not to assume a specific distribution of data, significantly differentially regulated genes between conditions were identified by a nonparametric test (Mann-Whitney U test). Fold-changes were calculated based on the median of each experimental group.

In order to identify potential gene function, the SOURCE database (http://source.stanford.edu) was used to retrieve Gene Ontology terms, after genes expressed more than once were removed (Dataset S1). Systematic analysis of functional groups in microarray datasets is a complex issue, which is still under discussion. We used the published method of Tavazoie et al. [22] to support the validity of the three functional groups we selected for discussion. Briefly, this method determines whether the selection of a subset of genes results in over-representation/enrichment of functional groups, compared to the genes in these functional groups on the entire array. We compared our gene findings with those from three other smaller IBD studies [11–13] to determine if the proposed disease mechanisms were similar.

Quality Controls and Assessment of Technical Variance

In order to ascertain data reliability and reproducibility, the expression levels of each of the 33,792 clones were measured in duplicate in each experiment. The resulting correlation coefficient was very high ($r = 0.99$), indicating that hybridisation signals were highly reproducible (unpublished data). To assess microarray and hybridisation-based experimental variability, hybridisation was performed on five separate microarrays using the same RNA sample (extracted from an intestinal surgical specimen; unpublished data). We performed pairwise comparisons between all possible combinations of filters. We computed the fold-change observed for each of the 33,792 replicate signals, which should theoretically be 1 (i.e., no fold-change) in all cases, but varies due to experimental noise. We recorded the fold-change observed at the 95th percentile of the distribution for each pairwise comparison and took the average of these values. This provided us with a fold-change of 1.2 as the significance threshold to claim differential expression. This threshold of positive gene identification is lower than reported for many studies; i.e., array studies using Affymetrix arrays usually consider a 2-fold or greater change between two conditions as significant. It is worth noting that common fluorescently labelled oligonucleotide-based array systems (i.e., Affymetrix) amplify the signal with biotin-streptavidin complexes, which is not the case with radioactive detection, hence the fold-change cutoff criteria are not comparable between these two platforms. However, we substantiated our threshold findings

by carrying out additional statistical analysis with an additional well-established microarray analysis package. Using significance analysis of microarrays (SAM) software (set to a maximum false discovery rate of 4%) to retest our analysis based on fold-changes, 10% of our significantly regulated genes were rejected by SAM. This difference comes mainly from genes that have barely passed the significance level of our own analysis (unpublished data). Heat maps of differentially regulated genes were generated using Spotfire Decision site 8.1 software (Spotfire, Somerville, Massachusetts, United States; http://www.spotfire.com/).

Verification of Clone Sequences

Differentially expressed clones were bidirectionally sequenced from PCR products of clone inserts or bacterial plasmids as previously described [14]. Contigs were generated if the clone ends overlapped (Sequencher software; Gene Codes Corporation, Ann Arbor, Michigan, United States) and identified by BLAST searching (http://www.ncbi.nlm.nih.gov/BLAST/). All genes reported in this study have been sequence-verified in this manner. Incorrect clone annotations were found in 27% of cases. This annotation error rate is similar to that reported for IMAGE clone sets in other studies [23,24].

Independent Quantitation of Microarray Results by TaqMan PCR

Selected gene expression signals were independently quantitated with real-time PCR (TaqMan) on patient group 2. Probes and primers were either designed to non-redundant sequences using Primer Express V2.0 (Applied Biosystems, Foster City, California. United States), or ordered from Applied Biosystems as Assays-on-Demand Gene Expression Assays (Table S1). Total RNA (1 μg) was reverse-transcribed to cDNA according to the manufacturer's instructions (Multi-Scribe Reverse Transcriptase, Applied Biosystems). Reactions were carried out on the ABI PRISM Sequence 7700 Detection System (Applied Biosystems) and relative transcript levels were determined using β-actin as the endogenous control gene. Statistically significant differences between control and IBD samples were determined using a Mann-Whitney U test. A p-value of less than 0.05 was considered significant.

Histological Rating of Disease Activity

A single pathologist (Y. G.), blinded to disease status and the experimental results, performed the histological rating of inflammatory activity in biopsy samples taken from the same anatomical location as the samples used for microarray experiments. The normal group showed no pathologic signs of activity in the histological examination (unpublished data). As detailed in Table 1, acute inflammatory activity and chronic changes were rated, with additional endoscopic findings also noted.

Immunohistochemistry

In order to determine protein localization, paraffin-embedded biopsies from normal controls ($n = 5$) and from patients with CD ($n = 5$) and UC ($n = 6$), which were obtained in parallel from the same sites as the biopsies used for the expression analysis studies, were analysed. This procedure is described in more detail elsewhere [25]. Briefly, 7-μm sections were subjected to heat-induced antigen retrieval in 0.01 M EDTA solution (pH 8) for 10 min. After blocking in a solution of 0.75% BSA in PBS for 20 min, the sections were washed three times in PBS and incubated for 1 h with the respective primary antibodies (monoclonal carcinoembryonic antigen-related cell adhesion molecule 1 [CEACAM1] 4D1/C2, protein kinase C beta 1 [PRKCB1; Pharmingen, San Diego, California, United States], and casein kinase 1, delta [[CSNK1D; Santa Cruz Biotechnology, Santa Cruz, California, United States]. After washing in PBS (three times for 10 min), the sections were incubated with peroxidase-conjugated rabbit anti-mouse secondary antibody (Sigma, Deisenhofen, Germany; 1:100) for 30 min, washed, and stained with goat anti-rabbit IgG (Sigma; 1:200, 30 min), before processing with diamino-benzidine and embedding in Aquatek (Merck, Hawthorne, New York, United States). For specificity controls, (i) the primary antibodies were omitted and (ii) normal sera were used from the species in which the primary antibodies were raised. No specific staining could be detected using either of these combinations (unpublished data).

Prediction of Function in Unknown Genes

In order to identify novel genes, unknown differentially regulated clones were examined in more detail. cDNA clone sequences were searched against the GenBank nonredundant database by means of BLAST searching. The best BLAST hit (sequence identity close to 100%) was used to extend the sequences to longer mRNA transcripts. InterPro (http://www.ebi.ac.uk/InterProScan), SMART (http://smart.embl-heidelberg.de) and Pfam (http://www.sanger.ac.uk/Software/Pfam) search methods (with standard parameter sets) were used to detect known protein domains and functional sequence motifs in putative open reading frames (ORFs) longer than 100 amino acids. Close homologies of the ORFs were found through their association to UniGene clusters in humans and other species. In cases of ORFs without a detectable homology to known genes, these transcripts were mapped to genomic locations by BLAST searching and examined in terms of their genomic context. Additionally, we used a previously published method for prediction of function of mouse transcripts [26], and applied a similarity analysis in order to predict the function based on common regulation patterns between known functional groups and unknown transcripts.

Results

Expression Profiles of the IBD Colonic Mucosa

Samples from sigmoid colon that were taken from normal controls ($n = 11$), patients with CD ($n = 10$), and patients with UC ($n = 10$) were used to compare differences between normal and inflamed intestinal mucosa. After implementing

Figure 1. Heat Map of Differentially Expressed Genes in Normal Controls Compared to IBD

(A) The top 40 up-regulated and the top 40 down-regulated between normal controls ($n = 11$) and CD patients ($n = 10$).
(B) The top 40 up-regulated and the top 40 down-regulated genes between normal controls ($n = 11$) and UC patients ($n = 10$).
Selection criteria were $p \leq 0.0015$ (based on Mann-Whitney U test) and a fold-change of 1.2 or greater. All genes presented in this heat map were sequence verified.
DOI: 10.1371/journal.pmed.0020199.g001

A

	Symbol	Name	Fold-change	p-Value
Normal / CD				
	BASP1	Brain abundant, membrane attached signal protein 1	1.98	0.00026
	B2M	Beta-2-microglobulin	1.84	0.00038
	CLDN3	Claudin 3	1.47	0.00079
	TAF4	TAF4 RNA polymerase II, TATA box binding protein (TBP)-associated factor, 135kDa	2.07	0.00011
	SAT2	Spermidine/spermine N1-acetyltransferase 2	1.67	0.00110
	FRZB	Frizzled-related protein	1.53	0.00055
	ALG6	Asparagine-linked glycosylation 6 homolog (yeast, alpha-1,3-glucosyltransferase)	2.47	0.00055
	ASS	Argininosuccinate synthetase	2.83	0.00017
		CDNA clone IMAGE:4807381, partial cds; Acc: BM972881	1.46	0.00026
		Transc. seq.; Acc: BX115981	2.15	0.00110
	FOXF2	Forkhead box F2	2.31	0.00110
	ENO1	Enolase 1, (alpha)	1.47	0.00004
	CLDN4	Claudin 4	1.70	0.00017
	SPARC	Secreted protein, acidic, cysteine-rich (osteonectin)	1.47	0.00110
	ZNF198	Zinc finger protein 198	1.52	0.00038
	INSIG2	Insulin induced gene 2	1.45	0.00038
	HLA-DRB3	Major histocompatibility complex, class II, DR beta 3	1.71	0.00017
		Transc. seq. with strong similarity to prot. pir:I38067; nitric-oxide synthase; Acc: AA393074	>10	0.00017
	RCOR3	REST corepressor 3	2.95	0.00036
	DHRS9	Dehydrogenase/reductase (SDR family) member 9	2.18	0.00001
	ITGB2	Integrin, beta 2 (antigen CD18 (p95), lymphocyte function-associated antigen 1 (...)	2.64	0.00038
	RRM2	Ribonucleotide reductase M2 polypeptide	1.90	0.00110
	LOC146174	Hypothetical protein LOC146174	2.60	0.00001
		CDNA clone IMAGE:5092935, partial cds; Acc: W45240	1.89	0.00004
	COL6A2	Collagen, type VI, alpha 2	2.50	0.00001
	DAF	Decay accelerating factor for complement (CD55, Cromer blood group system)	1.83	0.00038
	ANXA4	Annexin A4	1.76	0.00001
	TIMP1	Tissue inhibitor of metalloproteinase 1 (erythroid potentiating activity, collagenase inhibitor)	2.35	0.00007
	PFKP	Phosphofructokinase, platelet	1.45	0.00001
	ZYX	Zyxin	1.80	0.00079
	DCN	Decorin	1.48	0.00017
	COL6A1	Collagen, type VI, alpha 1	1.44	0.00055
	TNFRSF6B	Tumor necrosis factor receptor superfamily, member 6b, decoy	1.87	0.00002
	Prkcz	Protein kinase C, zeta	1.59	0.00026
	MMP2	Matrix metalloproteinase 2 (gelatinase A, 72kDa gelatinase, 72kDa type IV collagenase)	1.98	0.00026
	SLC39A9	Solute carrier family 39 (zinc transporter), member 9	1.62	0.00055
	MUC1	Mucin 1, transmembrane	1.68	0.00055
	PPP2R1B	Protein phosphatase 2 (formerly 2A), regulatory subunit A (PR 65), beta isoform	2.16	0.00017
	JUNB	Jun B proto-oncogene	1.52	0.00079
	IER3	Immediate early response 3	1.95	0.00001
	ZFP36	Zinc finger protein 36, C3H type, homolog (mouse)	1.50	0.00110
	IGHG1	Immunoglobulin heavy constant gamma 1 (G1m marker)	>10	0.00078
	PLA2G2A	Phospholipase A2, group IIA (platelets, synovial fluid)	1.72	0.00017
	DNAJB5	DnaJ (Hsp40) homolog, subfamily B, member 5	1.64	0.00055
	HLA-A	Major histocompatibility complex, class I, A	2.17	0.00017
	HLA-B	Major histocompatibility complex, class I, B	1.96	0.00079
	IGH@	Immunoglobulin heavy locus	5.50	0.00038
	ITM2C	Integral membrane protein 2C	1.63	0.00011
	CSNK1D	Casein kinase 1, delta	1.54	0.00110
	ZNF24	Zinc finger protein 24 (KOX 17)	1.86	0.00017
		Transc. seq.; Acc: AA401971	-1.83	0.00079
		EST; Acc: H66442	-1.65	0.00110
	HNMT	Histamine N-methyltransferase	-1.57	0.00017
	TNFRSF10B	Tumor necrosis factor receptor superfamily, member 10b	-1.87	0.00079
		H.s. transc. seq. with moderate similarity to protein sp:P39188 (...) Acc: BQ023720	-1.60	0.00011
	CLMN	Calmin (calponin-like, transmembrane)	-1.74	0.00017
		EST; Acc: R53815	-1.62	0.00002
	WDR10	WD repeat domain 10	-1.77	0.00055
		CDNA FLJ37384 fis, clone BRAMY2026347; Acc: R63134	-1.61	0.00079
		Transc. seq.; Acc: R62888	-1.61	0.00079
		EST; Acc: T52988	-1.60	0.00055
		Transc. seq.; Acc: W02851	-1.63	0.00079
		Full length insert cDNA YH93E08; Acc: BQ022853	-1.65	0.00007
		EST; Acc: R34024	-1.60	0.00004
		Similar to RIKEN cDNA 2210021J22 (LOC150383), mRNA; Acc: NM_032644	-1.62	0.00004
	FLJ21919	Hypothetical protein FLJ21919	-1.63	0.00026
	SLC6A2	Solute carrier family 6 (neurotransmitter transporter, noradrenalin), member 2	-1.63	0.00110
	CPD	Carboxypeptidase D	-1.62	0.00055
		Transc. seq. with weak similarity to prot. ref:NP_060190.1; FLJ20234; Acc: T40928	-1.62	0.00017
		Transc. seq.; Acc: T57946	-1.65	0.00026
	SLC9A6	Solute carrier family 9 (sodium/hydrogen exchanger), isoform 6	-1.64	0.00038
	POGZ	Pogo transposable element with ZNF domain	-1.58	0.00038
	VPS13D	Vacuolar protein sorting 13D (yeast)	-1.57	0.00002
		Transc. seq.; Acc: R34132	-1.60	0.00011
		Transc. seq.; Acc: R34570	-1.58	0.00004
		Transc. seq.; Acc: R65962	-1.58	0.00055
	COL8A2	Collagen, type VIII, alpha 2	-1.62	0.00038
	PKD1	Polycystic kidney disease 1 (autosomal dominant)	-1.67	0.00038
	CHST1	Carbohydrate (keratan sulfate Gal-6) sulfotransferase 1	-1.57	0.00004
		Transc. seq.; Acc: BM974959	-1.65	0.00017
	FLJ11000	Hypothetical protein FLJ11000	-1.61	0.00026
	KIF5C	Kinesin family member 5C	-1.62	0.00017
	NEDD9	Neural precursor cell expressed, developmentally down-regulated 9	-1.64	0.00011
	RQCD1	RCD1 required for cell differentiation1 homolog (S. pombe)	-1.62	0.00026
	LOC285540	Hypothetical protein LOC285540	-1.57	0.00002
	ALOX5	Arachidonate 5-lipoxygenase	-1.58	0.00079
	TAC3	Tachykinin 3 (neuromedin K, neurokinin beta)	-1.60	0.00007
	DKFZp547A023	Hypothetical protein DKFZp547A023	-1.65	0.00001
	ALB	Albumin	-1.59	0.00026
		Transc. seq.; Acc: BM724641	-1.80	0.00026
	DKFZp762E1312	Hypothetical protein DKFZp762E1312	-1.57	0.00055
	APRG1	AP20 region protein	-1.74	0.00007
	FLJ32440	Hypothetical protein FLJ32440	-1.78	0.00007
		Transc. seq.; Acc: R91316	-1.65	0.00079
	PVRL3	Poliovirus receptor-related 3	-1.58	0.00001
	AIM1L	Absent in melanoma 1-like	-1.60	0.00110
		Transc. seq.; Acc: R65798	-1.61	0.00079
	TLE4	Transducin-like enhancer of split 4 (E(sp1) homolog, Drosophila)	-1.65	0.00038
	LTB4R2	Leukotriene B4 receptor 2	-1.59	0.00007
	PRLR	Prolactin receptor	-1.59	0.00026

upregulated in CD

downregulated in CD

low — high expression value

B

	Symbol	Name	Fold-change	p-Value
	TIMM17B	Translocase of inner mitochondrial membrane 17 homolog B (yeast)	1.91	0.00038
		Clone 2-12 immunoglobulin light chain mRNA, partial cds ;Acc: AA405415	4.96	0.00017
		IG heavy chain variable region (VH4) mRNA, VH4-59 allele, partial cds ;Acc: AB027441	>10	0.00001
		Partial mRNA for IG light chain variable region (IGLV gene), clone RN86 ;Acc: Y14736	3.83	0.00007
		Immunoglobulin kappa light chain VKJ region mRNA, partial cds ;Acc: BC018761	3.88	0.00017
	IGKC	Immunoglobulin kappa constant	3.42	0.00038
	SYNPO2	Synaptopodin 2	3.40	0.00055
	SHB	SHB (Src homology 2 domain containing) adaptor protein B	1.91	0.00001
	HLA-DRB3	Major histocompatibility complex, class II, DR beta 3	2.09	0.00001
	GRP58	glucose regulated protein, 58kDa	1.84	0.00055
	ASS	Argininosuccinate synthetase	2.91	0.00007
	P4HB	Procollagen-proline, 2-oxoglutarate 4-dioxygenase (proline 4-hydroxylase) (...)	2.35	0.00002
	DAZAP1	DAZ associated protein 1	1.91	0.00001
	PSG2	Pregnancy specific beta-1-glycoprotein 2	1.95	0.00007
	S100A6	S100 calcium binding protein A6 (calcyclin)	1.84	0.00007
	MANEA	Mannosidase, endo-alpha	1.97	0.00038
	TFF1	Trefoil factor 1 (breast cancer, estrogen-inducible seq. expressed in)	3.06	0.00001
	CENPE	Centromere protein E, 312kDa	>10	0.00017
		Transc. seq. with strong similarity to prot pir:I38067; nitric-oxide synthase ;Acc: AA393074	>10	0.00036
	BF	B-factor, properdin	2.37	0.00038
	PRG1	Proteoglycan 1, secretory granule	4.23	0.00038
	DHX57	DEAH (Asp-Glu-Ala-Asp/His) box polypeptide 57	2.28	0.00055
	ITGB2	Integrin, beta 2 (antigen CD18 (p95), lymphocyte function-associated antigen 1 (...)	2.29	0.00017
	TRIP6	Thyroid hormone receptor interactor 6	2.02	0.00017
	AD158	Hypothetical protein AD158	3.45	0.00026
	PLAUR	Plasminogen activator, urokinase receptor	2.91	0.00055
	S100P	S100 calcium binding protein P	2.61	0.00011
	GLIS2	Kruppel-like zinc finger protein GLIS2	2.40	0.00079
		Transc. seq. ;Acc: N39241	3.10	0.00007
	IER3	Immediate early response 3	2.31	0.00110
	NOS2A	Nitric oxide synthase 2A (inducible, hepatocytes)	2.03	0.00055
	ZYX	Zyxin	2.30	0.00079
	TDG	Thymine-DNA glycosylase	2.42	0.00007
	DAF	Decay accelerating factor for complement (CD55, Cromer blood group system)	3.17	0.00004
	COL6A2	Collagen, type VI, alpha 2	2.11	0.00002
	TIMP1	Tissue inhibitor of metalloproteinase 1 (erythroid potentiating activity, collagenase inhibitor)	4.12	0.00001
	NLK	Nemo like kinase	>10	0.00025
	Prkcz	Protein kinase C, zeta	1.91	0.00002
	ZFP36	Zinc finger protein 36, C3H type, homolog (mouse)	2.02	0.00110
	TNFRSF6B	Tumor necrosis factor receptor superfamily, member 6b, decoy	2.35	0.00038
	PLA2G2A	Phospholipase A2, group IIA (platelets, synovial fluid)	2.37	0.00110
	IGHG1	Immunoglobulin heavy constant gamma 1 (G1m marker)	>10	0.00002
	CEACAM1	Carcinoembryonic antigen-related cell adhesion molecule 1 (biliary glycoprotein)	2.09	0.00079
	pp9099	PH domain-containing protein	2.54	0.00017
	CLTC	Clathrin, heavy polypeptide (Hc)	2.91	0.00038
	LOC91316	Similar to bK246H3.1 (immunoglobulin lambda-like polypeptide 1, pre-B-cell specific)	2.06	0.00026
	IGLL1	Immunoglobulin lambda-like polypeptide 1	1.94	0.00038
	HLA-B	Major histocompatibility complex, class I, B	2.19	0.00038
	ADAM12	A disintegrin and metalloproteinase domain 12 (meltrin alpha)	2.81	0.00038
	IGH@	Immunoglobulin heavy locus	>10	0.00002
	PSIP1	PC4 and SFRS1 interacting protein 1	-1.37	0.00110
	FLJ21125	Hypothetical protein FLJ21125	-1.42	0.00079
		CDNA clone MGC:40177 IMAGE:5167345, complete cds ;Acc: BC028123	-1.55	0.00017
		Transc. seq. with weak similarity to protein sp:P39195 (H.sapiens) (...) Acc: R85273	-1.37	0.00110
		Transc. seq. with weak similarity to protein sp:P39194 (H.sapiens) (...) Acc: N62647	-1.37	0.00110
		EST ;Acc: R53815	-1.36	0.00017
	BIRC4	Baculoviral IAP repeat-containing 4	-1.36	0.00026
	L3MBTL2	L(3)mbt-like 2 (Drosophila)	-1.38	0.00110
		Clone IMAGE:5273245, mRNA ;Acc: BI461842	-1.43	0.00026
	PVRL3	Poliovirus receptor-related 3	-1.38	0.00017
	FLJ32440	Hypothetical protein FLJ32440	-1.58	0.00002
		Transc. seq. ;Acc: BM724641	-1.56	0.00110
		Transc. seq. ;Acc: R69245	-1.39	0.00110
	ALB	Albumin	-1.50	0.00110
	ACSL1	Acyl-CoA synthetase long-chain family member 1	-1.41	0.00026
	DKFZp762E1312	Hypothetical protein DKFZp762E1312	-1.45	0.00079
	GNS	Glucosamine (N-acetyl)-6-sulfatase (Sanfilippo disease IIID)	-1.38	0.00011
	DKFZp547A023	Hypothetical protein DKFZp547A023	-1.47	0.00055
		Clone 24629 mRNA seq. ;Acc: AF052160	-1.41	0.00079
		Transc. seq. ;Acc: R15800	-1.49	0.00055
		Transc. seq. ;Acc: T57946	-1.55	0.00038
		H.s. Transc. seq. with weak similarity to protein ref:NP_060312.1; FLJ20489 ;Acc: T53274	-1.49	0.00079
	KIAA1128	KIAA1128 protein	-1.37	0.00110
	XRN1	5'-3' exoribonuclease 1	-1.36	0.00038
	KCNS2	Potassium voltage-gated channel, delayed-rectifier, subfamily S, member 2	-1.40	0.00026
		Transc. seq. ;Acc: AA704585	-1.54	0.00038
	AP1G2	Adaptor-related protein complex 1, gamma 2 subunit	-1.41	0.00038
		Transc. seq. ;Acc: AI821023	-1.41	0.00079
	ANGPTL3	Angiopoietin-like 3	-1.41	0.00017
		Similar to RIKEN cDNA 2210021J22 (LOC150383), mRNA ;Acc: NM_032644	-1.36	0.00079
		Full length insert cDNA clone YI46G04 ;Acc: AI580142	-1.37	0.00110
	LIFR	Leukemia inhibitory factor receptor	-1.38	0.00110
	PEF	PEF protein with a long N-terminal hydrophobic domain (peflin)	-1.37	0.00055
	ESRRG	Estrogen-related receptor gamma	-1.42	0.00110
	APPBP1	Amyloid beta precursor protein binding protein 1, 59kDa	-1.41	0.00110
	CASK	Calcium/calmodulin-dependent serine protein kinase (MAGUK family)	-1.37	0.00007
		H.s. Transc. seq. ;Acc: R17784	-1.39	0.00079
		EST ;Acc: T51800	-1.38	0.00110
	PDCD5	Programmed cell death 5	-1.50	0.00079
	GAS2	Growth arrest-specific 2	-1.45	0.00110
	FLJ37953	Hypothetical protein FLJ37953	-1.39	0.00079
	PRKAA1	Protein kinase, AMP-activated, alpha 1 catalytic subunit	-1.37	0.00026
		CDNA FLJ44273 fis, clone TOVAR2001281 ;Acc: BM311753	-1.40	0.00055
	PSCD1	Pleckstrin homology, Sec7 and coiled-coil domains 1(cytohesin 1)	-1.51	0.00026
		Transc. seq. ;Acc: BM974959	-1.52	0.00079
	FLJ11000	Hypothetical protein FLJ11000	-1.47	0.00055
	HAMP	Hepcidin antimicrobial peptide	-1.50	0.00110
	LOC285540	Hypothetical protein LOC285540	-1.37	0.00038
	SIPA1L3	Signal-induced proliferation-associated 1 like 3	-1.43	0.00011
	SLC26A2	Solute carrier family 26 (sulfate transporter), member 2	-1.37	0.00055

Normal | UC

upregulated in UC

downregulated in UC

low — high — expression value

Table 3. Functional Groups of Differentially Regulated Genes

Functional Category	Symbol	Name	NC Versus CD		NC Versus UC	
			Fold-Change	p-Value	Fold-Change	p-Value
Immune and Inflammatory Response	—	Transcribed sequence with strong similarity to protein pir:I38067 (H. sapiens) I38067 nitric-oxide synthase	>10	0.001066	>10	0.000363
	IGHG1	Immunoglobulin heavy constant gamma 1 (G1m marker)	>10	0.000783	>10	0.000023
	IGH@	Immunoglobulin heavy locus	5.5	0.000380	>10	0.000023
	ASS	Argininosuccinate synthetase	2.83	0.000170	2.91	0.000068
	ITGB2[a]	Integrin, beta 2 (antigen CD18 (p95), lymphocyte function-associated antigen 1; macrophage antigen 1 (mac-1) beta subunit	2.64	0.000380	2.29	0.000170
	DAF	Decay accelerating factor for complement (CD55, Cromer blood group system)	1.83	0.000380	3.17	0.000040
	IER3	Immediate early response 3	1.95	0.000006	2.31	0.001100
	HLA-B	Major histocompatibility complex, class I, B	1.96	0.000788	2.19	0.000380
	TNFRSF6B	Tumor necrosis factor receptor superfamily, member 6b, decoy	1.87	0.000023	2.35	0.000380
	PLA2G2A[a]	Phospholipase A2, group IIA (platelets, synovial fluid)	1.72	0.000170	2.37	0.001100
	HLA-DRB3	Major histocompatibility complex, class II, DR beta 3	1.71	0.000170	2.09	0.000006
	HAMP	Hepcidin antimicrobial peptide	−1.51	0.000788	−1.5	0.001100
	C5	Complement component 5	−1.42	0.000108	−1.28	0.000788
	BIRC4	Baculoviral IAP repeat-containing 4	−1.41	0.000006	−1.36	0.000255
	HLA-A	Major histocompatibility complex, class I, A	2.17	0.000170		
	IL1R1	Interleukin 1 receptor, type I	1.32	0.000380		
	LTB4R2	Leukotriene B4 receptor 2	−1.59	0.000068		
	ALOX5	Arachidonate 5-lipoxygenase	−1.58	0.000788		
	CYLD	Cylindromatosis (turban tumor syndrome)	−1.53	0.001111		
	ALOX5AP	Arachidonate 5-lipoxygenase-activating protein	−1.52	0.000550		
	ADORA3	Adenosine A3 receptor	−1.48	0.000068		
	HLA-DQA1	Major histocompatibility complex, class II, DQ alpha 1	−1.47	0.000108		
	CDKN1A	Cyclin-dependent kinase inhibitor 1A (p21, Cip1)	−1.46	0.000788		
	CCL28	Chemokine (C-C motif) ligand 28	−1.45	0.001100		
	MICA	MHC class I polypeptide-related sequence A	−1.4	0.001100		
	RNPEP	Arginyl aminopeptidase (aminopeptidase B)	−1.37	0.000170		
	MASP1	Mannan-binding lectin serine protease 1	−1.36	0.001100		
	IGKC[a]	Immunoglobulin kappa constant			3.42	0.000380
	P4HB	Proline 4-hydroxylase			2.37	0.000023
	NOS2A	Nitric oxide synthase 2A (inducible, hepatocytes)			2.03	0.000550
	CEACAM1[a]	Carcinoembryonic antigen-related cell adhesion molecule 1			2.01	0.000788
	IGLL1	Immunoglobulin lambda-like polypeptide 1			1.94	0.000380
	IFITM1	Interferon induced transmembrane protein 1 (9–27)			1.65	0.000068
	CHUK	Conserved helix-loop-helix ubiquitous kinase			1.51	0.001100
	EGLN3	Egl nine homolog 3 (Caenorhabditis elegans)			1.42	0.000170
	MYBL2	V-myb myeloblastosis viral oncogene homologue (avian)-like 2			1.35	0.000108
	ALDH2	Aldehyde dehydrogenase 2 family (mitochondrial)			1.28	0.00110
	BST1	Bone marrow stromal cell antigen 1			−1.34	0.000550
Oncogenesis, Cell Proliferation, and Growth	TIMP1	Tissue inhibitor of metalloproteinase 1	2.35	0.000068	4.12	0.000006
	PLA2G2A[a]	Phospholipase A2, group IIA (platelets, synovial fluid)	1.72	0.00017	2.37	0.0011
	CAPNS1	Calpain, small subunit 1	1.24	0.000255	1.36	0.000068
	GAS2	Growth arrest-specific 2	−1.54	0.00055	−1.45	0.0011
	STAG1	Stromal antigen 1	−1.46	0.000011	−1.31	0.00038
	DAB2	Disabled homolog 2, mitogen-responsive phosphoprotein (Drosophila)	−1.33	0.00004	−1.21	0.0011
	DNAJB5	Dnaj (Hsp40) homolog, subfamily B, member 5	1.64	0.00055		
	CSNK1D	Casein kinase 1, delta	1.54	0.0011		
	NME2	Non-metastatic cells 2, protein (NM23B) expressed in	1.28	0.0011		
	NEDD9	Neural precursor cell expressed, developmentally down-regulated 9	−1.64	0.000108		
	CYLD	Cylindromatosis (turban tumor syndrome)	−1.53	0.001111		
	TSSC1	Tumor suppressing subtransferable candidate 1	−1.51	0.000068		
	CISH	Cytokine inducible SH2-containing protein	−1.49	0.00055		
	Rock1	Rho-associated, coiled-coil containing protein kinase 1	−1.48	0.00017		
	PPP2CA	Protein phosphatase 2 (formerly 2A), catalytic subunit, alpha isoform	−1.48	0.00004		
	PICALM	Phosphatidylinositol binding clathrin assembly protein	−1.46	0.00004		
	WIT-1	Wilms tumor associated protein	−1.45	0.00055		
	FLJ13352	Hypothetical protein FLJ13352	−1.4	0.00038		
	RASA1	RAS p21 protein activator (GTPase-activating protein) 1	−1.39	0.0011		
	CDK5RAP1	CDK5 regulatory subunit associated protein 1	−1.37	0.000788		
	RAD1	RAD1 homolog (Saccharomyces pombe)	−1.33	0.00055		
	CENPE	Centromere protein E, 312 kDa			>10	0.00017
	TFF1[a]	Trefoil factor 1 (breast cancer)			3.06	0.000006
	S100P[a]	S100 calcium binding protein P			2.61	0.000108
	S100A6	S100 calcium binding protein A6 (calcyclin)			1.84	0.000068
	PRKCN	Protein kinase C, nu			1.68	0.0011

Table 3. Continued

Functional Category	Symbol	Name	NC Versus CD		NC Versus UC	
			Fold-Change	p-Value	Fold-Change	p-Value
	S100A11[a]	S100 calcium binding protein A11 (calgizzarin)			1.56	0.000011
	CAPN1	Calpain 1 (mu/l) large subunit			1.37	0.000108
	MYBL2	V-myb myeloblastosis viral oncogene homolog (avian)-like 2			1.35	0.000108
	NDRG3	NDRG family member 3			−1.24	0.0011
Structure and Permeability	TIMP1	Tissue inhibitor of metalloproteinase 1	2.35	0.000068	4.12	0.000006
	COL6A2	Collagen, type VI, alpha 2	2.50	0.000011	2.11	0.000023
	MMP2	Matrix metalloproteinase 2	1.98	0.000255	1.80	0.000108
	MUC1[a]	Mucin 1, transmembrane	1.68	0.00055	1.74	0.000011
	MUC13	Mucin 13, epithelial transmembrane	1.35	0.00017	1.64	0.000023
	CDH11	Cadherin 11, type 2, OB-cadherin (osteoblast)	1.59	0.00004	1.48	0.00017
	DKFZp547A023	Hypothetical protein dkfzp547a023	−1.65	0.000011	−1.47	0.00055
	FLJ36812	Hypothetical protein FLJ36812	−1.42	0.000023	−1.31	0.000023
	CLDN4	Claudin 4	1.70	0.00017		
	CLDN3	Claudin 3	1.47	0.000788		
	SPARC[a]	Secreted protein, acidic, cysteine-rich (osteonectin)	1.47	0.0011		
	COL8A2	Collagen, type VIII, alpha 2	−1.62	0.00038		
	COL15A1	Collagen, type XV, alpha 1	−1.55	0.00055		
	PRKCB1[a]	Protein kinase C, beta 1	−1.5	0.000788		
	ROCK1	Rho-associated kinase beta	−1.48	0.00017		
	LIM	LIM protein (similar to rat protein kinase C-binding enigma)	−1.48	0.000788		
	-	Transcribed sequences (GenBank N48794)	−1.46	0.000006		
	RDX	Radixin	−1.42	0.00055		
	ANK3	Ankyrin 3, node of Ranvier (ankyrin G)	−1.35	0.00038		
	OCLN	Occludin	−1.34	0.000788		
	FLJ25778	Hypothetical protein FLJ25778	−1.33	0.0011		
	CLTC	Clathrin, heavy polypeptide (Hc)			2.91	0.00038
	P4HB	Procollagen-proline, 2-oxoglutarate 4-dioxygenase			2.35	0.000023
	COL21A1	Collagen, type XXI, alpha 1			1.61	0.000788

Minus signs in the columns "Fold-Change" indicate down-regulation between normal controls (NC) and IBD individuals.
[a]Gene previously reported to be differentially expressed in IBD [11–13]. All data based on microarray results.
DOI: 10.1371/journal.pmed.0020199.t003

established analysis criteria, 650 genes were found to be differentially regulated between normal controls and at least one of the IBD subtypes. Of these 650 genes, 500 transcripts (81 up-regulated) were identified as differentially regulated between normal controls and CD patients. Two hundred and seventy-two differentially regulated transcripts (157 up-regulated) were identified between normal controls and UC patients. At our cutoff level, 122 genes were dysregulated in both diseases, and approximately 44% and 36% of all differentially regulated genes represented unknown genes in CD and UC, respectively. Heat maps shown in Figure 1 represent the top 40 up- and down-regulated genes for each experimental group. The complete list of all differentially regulated genes identified in this study (including reference to chromosomal areas previously implicated in IBD linkage studies) can be found in Table S2.

In order to delineate the molecular fingerprint of each of the two disease subtypes, the differentially regulated genes were assigned to functional groups based on classification by Gene Ontology (http://www.geneontology.org/) and references in the literature. The majority of genes fell into the following groups, which are listed as such in Table 3: immune and inflammatory responses; oncogenesis, cell proliferation, and growth; and structure and permeability. The significance of this observation was tested according to Tavazoie et al. [22]. The p-value for the overall significance for an over-representation of genes in the category "immune and inflammatory responses" was 0.00062. This enrichment was

mainly driven by genes that were up-regulated in CD ($p = 0.0011$) and UC ($p = 0.0026$). An enrichment of genes associated with cell growth and proliferation ($p = 0.0404$) was driven by genes that were up-regulated in UC, while genes associated with structure and permeability ($p = 0.0309$) are mostly found in the category of genes up-regulated in CD.

Functional Prediction of Unknown Genes

To decipher potential roles for unknown genes, we further analysed some of the top differentially expressed unknown genes between normal controls and IBD (Table 4). Some of the transcripts had no known sequence homologies (e.g., GenBank accession number BC006384), whereas others were in UniGene clusters with sequences homologous to structural/cytoskeletal proteins (e.g., AK022544). Additionally, some transcripts showed sequence homology to cell adhesion genes (e.g., N48794 and AF087994). The gene represented by GenBank accession number AL117511 may be involved in vacuolar sorting, and AK056932 is a putative transmembrane phosphate acyltransferase. Additional gene products contain DNA-binding domains and may function as transcription factors (e.g., N39296 and AW953679), or represent a possible kinase (e.g., AB067499) or may play a role in intracellular signalling (e.g., BC008744). Of the unknown genes analysed, 90% have homologues in other species, lending credibility to the hypothesis that these as yet uncharacterised genes must have an important function since they are conserved in the evolutionary process. Additionally, predictions based on

similarities of regulation patterns of known functional groups are listed in Table 4.

Independent Quantification by Fluorescent Real-Time PCR

Real-time TaqMan PCR on selected transcripts was carried out as a sensitive independent verification method of the microarray results (Figure 2). The population of 31 patients used for microarray studies (group 1) was extended for the real-time PCR experiments to 100 patients (group 2). Group 1 patients did not exhibit significant differences from group 2 patients on the level of gene expression, as verified by groupwise comparison. Genes for real-time PCR were chosen based upon their dysregulation in IBD, and representation within the three main functional groups discussed later. Differentially regulated transcripts of unknown function were also included. As an additional control for the suitability of the samples and the stability of the overall approach, *TNF-α* and *interleukin-8 (IL-8)* were used, as small soluble cytokine mediators not present on the array system but known as hallmarks of IBD pathophysiology. As expected, overexpression of *IL-8* in both CD and UC and *TNF-α* in CD was observed in the samples used. An additional 15 genes on the microarrays were then analysed in the extended real-time patient cohort (i.e., patient group 2). The 15 real-time PCR experiments that were carried out are summarised in Figures S1 and 2 and Table 5. Genes that were significantly up-regulated in both microarray and real-time PCR in the two IBD subtypes (compared to normal controls), include *cadherin-11 (CDH11), decay accelerating factor for complement (DAF), immunoglobulin heavy constant gamma 1 (IGHG1), mucin 1 (MUC1), phospholipase A2, group IIA (PLA2G2A),* and *tissue inhibitor of metalloproteinase 1 (TIMP1).* The unknown gene, *DKFZp547A023* (GenBank accession number AK022544; see Table 4), was confirmed by real-time PCR to be down-regulated in both disease groups. In the case of *cylindromatosis (CYLD), calcitonin gene-related peptide-receptor component protein (RCP9), LIM protein (LIM), occludin (OCLN), Rho-associated, coiled-coil containing protein kinase 1 (ROCK1),* and *zinc finger, CCHC domain containing 4 (ZCCHC4),* the microarray results showed that these genes were down-regulated in both IBD subtypes, but failed to reach significance in the UC cohort. However, results from the extended real-time analysis showed that these six genes were significantly down-regulated in both IBD subtypes. *PH domain-containing protein (PP9099,* BC008744; see Table 4) was up-regulated in both diseases on the microarray, but our microarray cutoff criteria were not reached in the CD population. Real-time analysis on the extended IBD population demonstrated this gene was significantly up-regulated in both IBD subtypes. One gene, *trefoil factor 1 (TFF1),* was found to be up-regulated in UC on the microarray but failed to reach significance for CD, a result that was confirmed by real-time PCR in the extended group 2 samples. These results strongly emphasize the importance of using a high-throughput quantitative technique to follow positive signals from microarray experiments into larger and diverse patient populations.

Expression Profiles in Non-IBD Disease Samples

Disease specificity is an important issue in the detection of differential gene expression between IBD and normal controls. Real-time PCR was carried out on patients that had colonic disease (DC, $n = 17$), but not IBD (Figures S1 and 3). These DC samples were again divided into noninflamed

and inflamed conditions. No significant difference was observed between the expression of any of the 15 genes tested between normal controls and noninflamed disease specificity controls. Eight genes (*CYLD, DAF, DKFZp547A023, MUC1, OCLN, PLA2G2A, TIMP1,* and *ZCCHC4;* Table 5) were significantly differentially regulated between normal controls and inflamed DC, and the direction of change was the same as that observed in IBD. This observation suggests that these eight gene findings probably reflect general inflammation pathophysiology rather than events specific for IBD. However, unlike in the IBD cohort, the expression of the remaining seven genes (*CDH11, IGHG1, PP9099, TFF1, LIM, ROCK,* and *RCP9;* Table 5) was not significantly different in between normal controls and the inflamed DC group.

Localisation by Immunohistochemistry

As expression analysis in complex tissues does not allow an identification of the cells responsible for the signal, we used immunohistochemistry to exemplify the approach to genes of interest. Paraffin-embedded sections of samples from the group 1 patients were used. Genes were chosen on the basis that they were formerly not associated with IBD etiopathogenesis and represented a gene in our three main functional groups. Consistent with increased transcript levels measured by microarrays, a marked up-regulation of CEACAM1 protein was seen in IBD when compared to normal controls (Figure 4A). Immunoreactivity was found in the apical epithelial lining in the normal mucosa, whereas it extended down into the epithelial cells of the crypts in the inflamed tissue of CD and UC patients. A staining of vascular structures and mononuclear cells, most likely with a lymphocytic phenotype, was only detected in the lamina propria of inflamed biopsies.

CSNK1D immunohistochemistry staining (Figure 4B) demonstrated expression in the colon, but the modest up-regulation seen in CD in the microarray was not detected by immunohistochemistry. In the biopsies from CD patients, staining of the apex of the crypts could be observed. Immunoreactivity revealed a strong granular staining pattern of intestinal epithelial cells in the normal and UC group, which was located basolaterally.

Staining of PRKCB1 (Figure 4C), that appeared to be down-regulated in the microarray results showed only a weak staining in the apical epithelial layer in the normal and CD mucosa, whereas a stronger staining was present in the UC group. Furthermore, scattered lamina propria mononuclear cells underlying the epithelial layer were also positive. Interestingly, in the CD group, immunoreactivity was found nearly exclusively in the marginal zone of small lymph follicles.

Discussion

The current understanding of IBD pathogenesis is that of a complex interplay of both genetic and environmental factors that results in an aberrant immune response of the intestinal mucosa. In the past few years, significant progress has been made in elucidating the root causes of CD and UC, with at least three genes now identified in which sequence variations confer disease susceptibility [4–8]. The advent of genomic technologies such as genome-wide microarrays can be used to dissect genes relevant to disease pathophysiology, and allows an unbiased view of both specific and nonspecific (such as

Table 4. In Silico (InterPro, SMART and Pfam Search Methods) Analysis of Top Differentially Expressed Unknown Genes between Normal Controls and CD or UC, Respectively

Name/Homologues in Other Species	Accession Number	Cytoband	Comment	Putative Function	NC Versus CD		NC Versus UC	
					Fold-Change	p-Value	Fold-Change	p-Value
Clone IMAGE:4110919, mRNA/no homologues	BC006384	17p11.2	ORF has no known protein domains; transcript maps to an intron of heparan sulfate (glucosamine) 3-O-sulfotransferase 3B1	Unknown	1.41	0.000255		n.s.
Hypothetical protein DKFZp547A023/Mm, Rn	AK022544	1p12	Homology to collagen, myosin tail domain, ATPase domain, high expression similarities to endocytosis genes	Endocytosis	−1.65[a]	0.000011	−1.47[a]	0.000550
Chromosome 14 open reading frame 125/Mm, Rn	AL117511	14q12	Homology to vacuolar sorting-associated protein VPS13 (SOI1), high expression similarities to genes regulating apoptosis	Regulation of apoptosis	−1.33	0.000040	−1.25	0.000170
Hypothetical protein FLJ32370/Mm, Rn, Dm, Ce	AK056932	8	Phosphate acyltransferase domain, high expression similarities to genes involved in lipid metabolism	Phospholipid metabolism	−1.54	0.000040		n.s.
Hs transcribed sequences/Mm, Rn	N48794	3p21	Cadherin repeats	Cell adhesion	−1.46	0.000006		n.s.
Hypothetical protein FLJ36812/Mm, Rn, Dm	AF087994	15q14	Leucine-rich repeats	Cell adhesion	−1.42	0.000023	−1.31	0.000023
KIAA1912 protein/Mm, Rn	AB067499	2p16.2	ORF shows similarities to nucleoside triphosphate hydrolases, expression similarities to genes with guanylate kinase activity	Guanylate kinase		n.s.	−1.25	0.000550
PH domain-containing protein/Mm, Rn	BC008744	15q22.1	Pleckstrin homology domain	Intracellular signaling		n.s.	2.54[a]	0.000170
zinc finger, CCHC domain containing 4/Mm, Rn, Dm, Ce	N39296	4p15.31	GRF zinc finger, shows expression similarities to transcription factors	Transcription factor	−1.50[a]	0.000380		n.s.
Chromosome 20 open reading frame 17	AW953679	20q13.2	Contains five zinc fingers and hox domain	Transcription factor	−1.22	0.000550		n.s.

Fold-changes are derived from microarray analysis; minus sign in the columns ''Fold-Change'' indicate down-regulation between normal controls (NC) and Crohn disease (CD). The fold-changes as reported by quantitative PCR for AK022544 (DKFZp547A023) were −1.96 for CD and −1.80 for UC, for BC008744 (PP9099) were 1.37 for CD and 1.60 for UC and for N39296 (ZCCHC4) were −1.61 and −1.46, respectively. The prediction of the function was based on two criteria: sequence homologies and expression pattern similarities. Expression pattern similarities were calculated based on cosine correlation, accepted similarity had to be greater than 97% and enrichment significance for the putative function based on GO was set to p ≤ 0.05.

[a] Direction of fold-change verified by quantitative real-time PCR in group 2 samples.

Ce, *Caenorhabditis elegans*; Dm, *Drosophila melanogaster*; Hs, *Homo sapiens*; Mm, *Mus musculus*; Rn, *Rattus norvegicus*; NC, normal control; n.s., not significant.

DOI: 10.1371/journal.pmed.0020199.t004
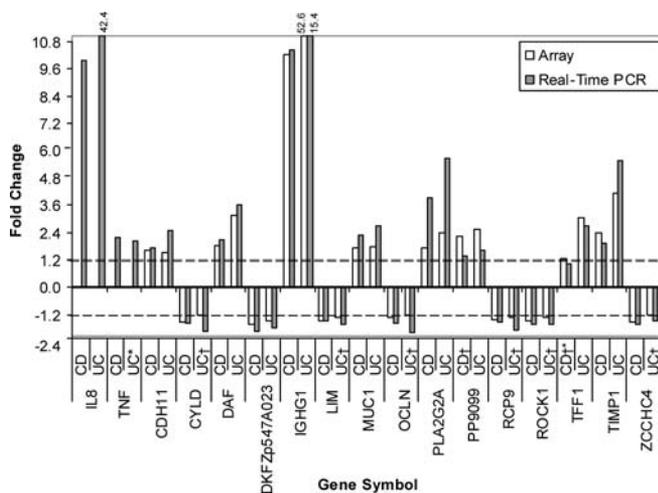
Figure 2. Microarray Results and Corresponding Quantitative Real-Time PCR for Differentially Regulated Genes in CD or UC Compared to Normal Controls

Genes were chosen on the basis of their dysregulation in IBD and represent both known genes from functional groups discussed and genes of unknown function. Quantitative real-time PCR was carried out on individual samples from group 2 patients (14–18 normal controls, 19–33 UC, and 17–22 CD samples, depending on the availability of the patient samples at the time the plates were produced), except for IL-8 and TNF-α (not on array), which were tested in group 1 patients (11 normal controls, ten UC, and ten CD patient samples) as a proof-of-principle measure. The extended cohort of group 2 patients includes those with active disease and using anti-inflammatory drugs (but not immunosuppressants or biologicals), whereas group 1 patients had active disease and were medication-free for 6 wk. Results are summarised by a ratio of medians (CD:normal or UC:normal). Complete results, including box-plots and number of samples analysed in each assay, are included in Figure S1. All results were significantly differentially regulated except for marked results; single dagger indicates that array result was not significant ($p > 0.0015$ or fold-change < 1.2); asterisk indicates that real-time PCR result was not significant ($p > 0.05$). A dashed line represents the fold-change level of 1.2.
DOI: 10.1371/journal.pmed.0020199.g002

Figure 3. Quantitative Real-Time PCR Results between Normal Controls and Both Non-Inflamed and Inflamed Disease Specificity Controls

Quantitative real-time PCR was carried out on individual samples from group 2 patients (including seven or eight inflamed and seven to nine non-inflamed DCs, depending on the availability of cDNA at the time the plates were produced). The DCs include patients with infectious diarrhoea, gastrointestinal inflammation, or irritable bowel syndrome. Patients in this group were not on immunosuppressants or biologicals, but the use of anti-inflammatory drugs was allowed. Results are summarised by a ratio of medians (inflamed DC:normal or non-inflamed DC:normal). Results marked with an asterisk indicate that the real-time PCR result was significant ($p < 0.05$). Complete results, including box-plots and sample numbers analysed in each assay, are included in Figure S1.
DOI: 10.1371/journal.pmed.0020199.g003

response to inflammation) events. In this study we aimed to use the power of genome-wide expression analysis by microarray to dissect novel important regulatory molecules downstream of the multiple primary genetic causes that may initiate the pathophysiologic cascade of inflammation in IBD. Therefore, both specific and nonspecific dysregulation may yield valuable information in identifying new targets for future disease therapy.

In the present study, a total of 650 genes were differentially regulated between normal controls and the two IBD subtypes. More specifically, 500 and 272 differentially regulated transcripts were identified between normal controls and CD and UC patients, respectively. Of note, we observed an imbalance between over- and underexpressed genes in the IBD subtypes. In CD, approximately 84% of differentially expressed genes were found to be down-regulated, compared with 42% of genes in UC. Although this is an interesting finding, it is not prudent at this stage to suggest that broad up- (in UC) or down-regulation (in CD) of genes is a distinctive disease feature, as this finding is highly influenced by types and numbers of genes that are present on any given microarray system. Without exception, none of the 122 differentially expressed genes that were found in both CD and UC, was found to be over expressed in one disease and under expressed in the other. This finding strongly supports the

robustness of the observation and the notion of a shared general inflammation profile underlying two clinically divergent forms of IBD, whereas more specific events in the pathophysiologic cascade could be disease-specific.

Disease divergence is also supported by a previous IBD microarray study. Interestingly, when we compared our findings with other CD and UC microarray studies [11–13], we observed few overlaps at the level of individual gene expression (see Table 3), but we did find a high degree of concordance at the level of proposed functional groups and potential mechanisms relevant to IBD (see discussion below). This study therefore not only complements these previous studies, but adds to our knowledge of disease, particularly as the Human UniGene RZPD Set 1 is enriched for as yet unannotated ESTs and, therefore, varies significantly from the commercially available arrays previously used [11,12]. As far as we are aware, the present study is the first to report a systematic assessment of novel unknown genes that might play a role in IBD pathophysiology (44% and 36% of differentially regulated genes identified in CD and UC respectively are, as of yet, not annotated).

The differentiation between CD and UC is a long-debated clinical problem. The large overlap that is seen in pathophysiology is also seen in clinical course. With the exception of gradual differences in activity, all established medical therapies used for colonic inflammation in IBD are applicable to both CD and UC [27]. A considerable percentage of patients appear to change clinical presentation during the course of the disease and are re-diagnosed with one of the other two subforms of IBD. Because of this scenario, the design of the present study did not intend to provide any diagnostic test to differentiate between CD and UC. As a matter of fact, at least one disease gene identified for IBD shows association with both CD and UC [7]. Therefore, the

Table 5. Summary of Array and Real-Time PCR Results

Symbol	Gene Name	Functional Group	NC Versus CD			NC Versus UC			NC Versus Inflamed DC
			Fold-Change on Array (Real-Time PCR)	p-Value for Array (Real-Time PCR)	Differences Confirmed by Real-Time PCR[a]	Fold-Change on Array (Real-Time PCR)	p-Value for Array (Real-Time PCR)	Differences Confirmed by Real-Time PCR[a]	Significant by Real-Time PCR (Fold-Change)
CDH11	Cadherin 11, type 2, OB-cadherin	SP	1.59 (1.71)	0.00004[b] (0.00562)	Yes	1.48 (2.46)	0.00017[b] (0.00001)	Yes	No (1.18)
CYLD	Cylindromatosis (turban tumor syndrome)	IIR; OCPG	−1.53 (−1.59)	0.00111[b] (0.00001)	Yes	−1.21 (−1.92)	0.01966 (0.000003)	Yes	Yes (−1.72)
DAF	Decay accelerating factor for complement	IIR	1.83 (2.05)	0.00038[b] (0.00121)	Yes	3.17 (3.63)	0.00004[b] (0.000002)	Yes	Yes (2.27)
DKFZp547A023	Hypothetical protein DKFZp547A023 / Mm, Rn	Unknown	−1.65 (−1.96)	0.00001[b] (0.00001)	Yes	−1.47 (−1.80)	0.0055[b] (0.000003)	Yes	Yes (−1.73)
IGHG1	Immunoglobulin heavy constant gamma 1 (G1m marker)	IIR	>10 (10.45)	0.00078[b] (0.04170)	Yes	>10 (15.41)	0.00002[b] (0.000001)	Yes	No (1.05)
LIM	LIM protein (similar to rat protein kinase C-binding enigma)	SP	−1.48 (−1.49)	0.00078[b] (0.00172)	Yes	−1.30 (−1.64)	0.03570 (0.00001)	Yes	No (−1.30)
MUC1	Mucin 1, transmembrane	SP	1.68 (2.27)	0.00055[b] (0.04217)	Yes	1.74 (2.68)	0.00001[b] (0.00052)	Yes	Yes (2.19)
OCLN	Occludin	SP	−1.34 (−1.57)	0.00078[b] (< 0.00001)	Yes	−1.22 (−1.98)	0.01587 (0.00001)	Yes	Yes (−1.39)
PLA2G2A	Phospholipase A2, group IIA (platelets, synovial fluid)	IIR; OCPG	1.72 (3.94)	0.00017[b] (0.04036)	Yes	2.37 (5.67)	0.00110[b] (0.00639)	Yes	Yes (11.40)
PP9099	PH domain-containing protein	Unknown	2.23 (1.37)	0.04296 (0.02075)	Yes	2.54 (1.60)	0.00017[b] (0.00315)	Yes	No (1.34)
RCP9	Calcitonin gene-related peptide-receptor component protein	Unknown	−1.44 (−1.54)	0.00025[b] (0.00003)	Yes	−1.31 (−1.88)	0.01009 (0.00001)	Yes	No (−1.16)
Rock1	Rho-associated kinase beta	SP; OCPG	−1.48 (−1.65)	0.00017[b] (0.00001)	Yes	−1.34 (−1.63)	0.00365 (0.00008)	Yes	No (−1.23)
TFF1	Trefoil factor 1 (breast cancer, estrogen-inducible sequence expressed in)	OCPG	1.22 (1.05)	0.00276 (0.77692)	No	3.06 (2.71)	0.000006[b] (0.01246)	Yes	No (3.56)
TIMP1	Tissue inhibitor of metalloproteinase 1	SP; OCPG	2.35 (1.90)	0.00006[b] (0.00037)	Yes	4.12 (5.57)	0.000006[b] (0.0000004)	Yes	Yes (2.04)
ZCCHC4	Zinc finger, CCHC domain containing 4	Unknown	−1.50[b] (−1.61)	0.00038[b] (< 0.00001)	Yes	−1.22 (−1.46)	0.05134 (0.00008)	Yes	Yes (−1.50)

Minus sign in the fold-change column indicates down-regulation between normal controls (NC) and IBD individuals. The direction of the dysregulation seen in the comparison between inflamed DC and NC was the same as that observed in the comparison between IBD and NC. No influence of sex, disease characteristics, age, and pretreatment could be detected by exploratory statistical groupwise comparison.

[a] $p < 0.05$ (uncorrected for multiple testing).

[b] Significant differential expression on microarrays ($p \leq 0.0015$, fold-change ≥ 1.2).

IIR, immune and inflammatory response; OCPG, oncogenesis, cell proliferation, and growth; SP, structure and permeability.

DOI: 10.1371/journal.pmed.0020199.t005
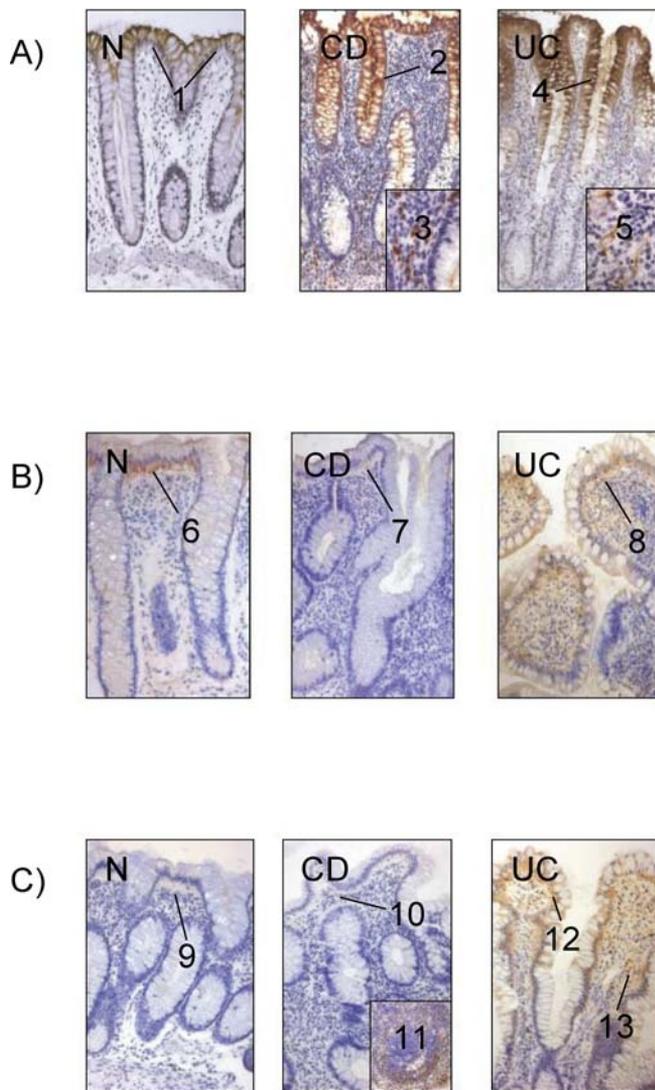
Figure 4. Immunohistochemical Localization of CEACAM1, CSNK1D, and PRKCB1 in Colonic Mucosa

Staining of a representative mucosal tissue samples from five normal controls (N), five Crohn disease (CD), and six ulcerative colitis (UC) patients using antibodies against (A) CEACAM1, (B) CSNK1D, and (C) PRKCB1.
(A) CEACAM1 immunoreactivity was found in the apical epithelial lining (1), crypts of inflamed tissue (2, 4). Additional staining was detected in immune cells (3) and blood vessels (5).
(B) CSNK1D immunoreactivity showed a strong granular staining pattern in normal (6) and the UC group (8), which was located basolaterally. Weaker staining of the apex of the crypts could be detected in CD (7).
(C) For PRKCB1, weak staining could be observed in the apical epithelial layer in the normal (9) and CD (10) mucosa and, interestingly, immunoreactivity was found nearly exclusively in the marginal zone of small lymph follicles (11), whereas the lamina propria was immuno-negative. In contrast, strong staining was detected in the apical epithelial layer of UC mucosa (12). Furthermore, lamina propria mononuclear cells underlying the epithelial layer were also positive in the UC mucosa (13).
DOI: 10.1371/journal.pmed.0020199.g004

overlap in etiologic factors and pathophysiology between both subforms of IBD could be larger than expected.

Interpreting the functional consequences of changes in gene expression observed in microarray expression screening is one of the major goals of exploratory microarray data analysis. One method used to attempt functional interpretation of microarray data is the use of annotation-based pathway databases

such as Gene Ontology (GO). The use of GO does have its limitations, in particular the amount and quality of annotation [28], but it does allow a broad overview of terms to decipher gene pathways. In the present study, we supplemented GO terms, with literature references, and could classify differentially expressed genes into three major groups: immune and inflammatory response; oncogenesis, cell proliferation, and growth; and structure and permeability (see Table 3).

Immune and Inflammatory Response

Compared to normal mucosa, many genes associated with an aberrant immune response were identified (enrichment significance for this functional group: $p = 0.00062$; category "Immune and Inflammatory Response," Table 3). Probably more than any other organ in the body, the intestine is a hostile environment, and it is not surprising that a general up-regulation of immune response (enrichment significance $p = 0.000973$) and antigen presentation (enrichment significance $p = 0.00298$) are a common feature of both IBD subtypes, considering the mucosal injury associated with these diseases. In contrast to these shared profiles, leukotriene biosynthesis appears to be more strongly associated with CD (enrichment significance $p = 0.00483$). With the exception of *phospholipase A2, group IIA (PLA2G2A)*, which has an inferred role in prostaglandin and leukotriene metabolism and was up-regulated in both subtypes, genes associated with leukotriene metabolism, such as *arachidonate 5-lipoxygenase (ALOX5), arachidonate 5-lipoxygenase-activating protein (ALOX5-AP), leukotriene B4 receptor 2 (LTB4R2),* and *arginyl amino-peptidase (aminopeptidase B; RNPEP)* were significantly down-regulated in the CD profile (Table 3). In the same vein, the marked down-regulation of CYLD is of interest as this gene has been identified as a key negative regulator of nuclear factor-kappa B (NF-κB) regulation [29–31]. Recent in vitro studies on CYLD regulation have uncovered a novel autoregulatory pathway in which up-regulation of CYLD, due to activation of NF-κB by TNF-α and bacteria, leads in turn to down-regulation of NF-κB signalling [32]. Although real-time analysis showed that this gene was also down-regulated in UC, decreased CYLD expression may represent an inflammation control response that is lost or impaired in some way in IBD. Similarly, inhibition of *NF-κB* signalling can be mediated by *nitrogen oxide (NO)* through the up-regulation of nitrogen oxide synthases (encoded by *NOS2A*, Table 3), which produces NO from arginine, along with other genes associated with proline and arginine metabolism, such as proline 4-hydroxylase (*P4HB*, Table 3) and aldehyde dehydrogenase 2 (*ALDH2*, Table 3), suggests over stimulation of this metabolic pathways in the UC profile. It is worth noting that it is not clear whether NO is deleterious or beneficial in gastrointestinal disease [33]. However, considering the well-documented potent inflammatory role NF-κB activation has in IBD [29–31], further examination of this complex issue is well worthwhile.

Finally, within this gene category, expression of *CEACAM1* (a known neutrophil activation marker and regulator of T cell function [34,35]) was examined by immunohistochemistry. We have shown that this protein is present in blood or lymphatic vessels of the inflamed mucosa (Figure 4A, insets), consistent with its role in angiogenesis [36]. Infiltrating immune cells in the lamina propria were stained by anti-CEACAM1 antibody in CD, but not in UC, suggesting that

CEACAM1 may be a better therapeutic target in CD. Expression of this adhesion molecule was distributed along the surface of the normal and IBD mucosa (Figure 4A), but distribution tended to more deeply penetrate the crypts in the IBD mucosa. It is tempting to speculate that this finding may have a role to play in increasing bacterial load in IBD, since CEACAM1 can act as a receptor for certain bacteria [37]. However, association of immunohistochemical findings with IBD pathogenesis are strictly speculative, and further studies, outside the scope of the present study, would need to be carried out to establish their exact role in disease.

Oncogenesis, Cell Proliferation, and Growth

Within this category, an enrichment of genes associated with cell growth and proliferation was found for genes up-regulated in UC (enrichment $p = 0.0404$; category "Onco-genesis, Cell Proliferation, and Growth," Table 3). This finding is similar to results from previous microarray studies, which reported involvement of cancer-related genes in IBD [11,12], although we report mostly different genes. During chronic inflammation, the constant repair mechanisms require precise controlling of molecular remodelling, cell proliferation, and growth [38], and dysregulation of any of these processes is followed by abnormalities that can lead to development of cancer [39], which is one of the main long-term complications of UC. Interestingly, one of the uncharacterised transcripts (Chromosome 14 ORF 125/Mm, Rn, GenBank accession number AL117511; see Table 4) was assigned as potentially involved in regulation of apoptosis, which concurs with the above-mentioned studies. Significant up-regulation of cancer-related genes in the UC profile is potentially important, considering reports of increased risks of developing colorectal carcinoma in this disease [40,41]. Up-regulated genes in the UC profile included *v-myb myeloblastosis viral oncogene homolog 2 (MYBL2,* Table 3) and several members of the S100 protein family, whose expression has been reported to promote cell proliferation and are linked with increased malignancy or tumour progression in colon cancer [42–45]. As confirmed by real-time PCR, TFF1 expression was also markedly higher in the UC group. This finding may be very interesting considering earlier studies describing increased TFF1 protein in colorectal cancer [46] and cancer development in TFF1 knock-out mice [47]. The potentially specific role this gene may have in UC is emphasized by the fact that we find no evidence of dysregulation of this gene in either CD or disease specificity controls. Similarly, the Wnt pathway is of intense interest, as aberrant activation is a common signalling abnormality in human cancers [36,37]. As members of the *casein kinase I (CKI)* gene family have been implicated in the regulation of Wnt-targeted gene expression [48,49], we used immunohistochemical analysis to localise the expression of this gene in IBD. Although contrary to the array results (where this gene was up-regulated in CD but not UC), the novel immunohistochemical finding of basolateral staining of CSNK1D in normal and UC epithelial cells, but only faint staining in the crypts of CD, may imply a role for Wnt signalling in IBD pathogenesis. This possibility would need to be investigated by more in-depth functional studies. The divergent expression patterns of genes involved in this category might yet prove useful in understanding the different risks of developing colorectal carcinoma between the two IBD subtypes [50].

Structure and Permeability

Constant tissue damage and injury of the intestinal surface are part of the pathophysiologic mechanisms in chronic disorders such as IBD that require continuous repair of the epithelium [38]. Enhanced permeability for inert macro-molecules is a well-described clinical feature of CD [51]. Most interestingly, *DLG5,* a recently discovered disease gene in IBD, also appears to be involved in mechanical integrity of epithelial barriers [7,52,53]. The organic cation transporter genes *SLC22A4* and *SLC22A5,* which are expressed in epithelial cells, have also been identified recently as disease genes for CD [8]. There is an enrichment of genes associated with structure and permeability in this study (enrichment $p = 0.0309$; category "Structure and Permeability," Table 3) as well as in previous microarray studies [11–13]. Several genes in this category were ubiquitously regulated in both IBD and non-IBD samples, reflecting known gene dysregulation in disease where inflammation and wound healing are recurrent events. These include paracellular permeability (down-regulation of *OCLN*) [54], degradation of extracellular matrix (up-regulation of *TIMP* and *MMP2*) [55,56] and barrier protection against bacterial invasion of the epithelial surface (up-regulation of *MUC1*) [57]. Conversely, several genes *(LIM* and *ROCK)* appear specifically relevant to disease processes in IBD, as no difference in regulation was observed in inflamed non-IBD patients by real-time PCR. Members of the cadherin superfamily, integral membrane proteins that mediate calcium-dependent cell-cell adhesion, have been shown to be involved in epithelial cell migration and resealing the site of tissue damage in the intestinal mucosa [58]. In this study, we show shared up-regulation of *cadherin 11 (CDH11)* in both IBD subtypes, too, but not in inflamed non-IBD tissue. It is tempting to speculate that this member of the cadherin family could also be involved in restructuring processes in the intestinal mucosa. In the same context, we identified two uncharacterised transcripts (GenBank accession numbers N48794 and AF087994) as differentially regulated, which were assigned a role in cell-cell adhesion (one of them containing cadherin repeats).

Cell migration in response to wound healing is regulated in part by cell adhesion processes, such as Rho-ROCK-mediated cytoskeletal reorganization [59,60]. Interestingly, members of this pathway, *ROCK1* and *LIM,* were down-regulated in both IBD subtypes in our PCR experiments (but not in inflamed DCs), which might indicate a potential decline in cell migration and an impaired ability to maintain epithelial integrity in IBD. Furthermore, immunohistochemistry staining of *PRKCB1*, which has been shown to interact with *LIM* [59], showed that lamina propria cells were preferentially stained in UC, but not CD or normal controls, possibly indicating a unique role for this protein in UC. Interestingly, a genetic variant in *DLG5,* which was recently associated with IBD [7], is a PDZ-containing scaffolding protein that associates with β-catenin and is a binding partner of vinexin at cell-cell contacts [52]. *LIM*, in addition to its signalling function by interacting with *PRKCB1* through its *LIM* domain [59], also contains PDZ domains, which may be used as sites for interaction of scaffolding proteins. These findings might represent interesting targets for further functional characterization in the context of wound healing and regeneration.

This study represents the largest sample number used in IBD microarray studies to date and focuses on three groups, namely normal control, CD, and UC. It would be of interest to substratify the groups according to other factors, which may then reveal factor-specific effects. For example, heterogeneity in the age distribution of the patient cohort could have the potential to introduce a bias. To address age-related effects, an independent analysis omitting age outliers from the analysis was performed. This did not result in significant changes in the presented list of differentially expressed genes, with only one out of 650 genes (Zyxin) failing to meet the cutoff criteria. Similarly, some patients, although meeting both clinical and endoscopic study inclusion criteria, had no acute infiltrate in their histology assessment (an "A0" score on histology). The lack of a reliable histological score reflecting disease activity is a well-known phenomenon in clinical research in IBD [27]. Elimination of individuals with an "A0" histology score did not change the results in a de novo analysis. We therefore conclude that the presented cohort composition was appropriately selected to identify differentially expressed genes in CD and UC, while noting that higher numbers of microarray samples would be necessary to sufficiently delineate the effects of additional clinical parameters. A future agenda would be the assessment of differential gene expression according to multiple sub-phenotypes and natural disease course. A prospective cohort that is adequately powered in size has been assembled and is currently under investigation.

In conclusion, this study reports our gene findings using a large, whole-genome filter system (Human UniGene set RZPD 1). Our results indicate that there are differences in the gene expression patterns between normal colonic mucosa and both CD and UC. The main findings point to novel, important molecules in abnormal immune regulation and the highly disturbed cell biology of the intestine in IBD pathogenesis. We suggest a host of novel genes that could be implicated in IBD pathophysiology and pathogenesis. The large number of involved mechanisms underlines the complexity of the IBD phenotype, but offers a large number of potential starting points for the development of new therapeutic strategies in the management of IBD.

Supporting Information

Dataset S1. Expression values and GO annotations for 650 sequence-verified genes discussed in the manuscript.

Found at DOI: 10.1371/journal.pmed.0020199.sd001 (409 KB TXT).

Figure S1. Boxplot Results of Real-Time PCR Experiments on 15 Genes

Real-time PCR was carried out for 15 genes as described in main text and shown in Figures 2 and 3. The number of samples per analysis is given in parentheses for each group analysed: normal controls; CD; UC; DC, actively inflamed (Inflamed__DC); and DC, not actively inflamed (Non-inflamed__DC). Graphs are provided for: *CDH11* (1a), *CYLD* (1b), *DAF* (1c), *DKFZp547A023* (1d), *IGHG1* (1e), *LIM* (1f), *MUC1* (1g), *OCLN* (1h), *PLA2G2A* (1i), *PP9099* (1j), *RCP9* (1k), *ROCK1* (1l), *TFF1* (1m), *TIMP1* (1n), and *ZCCHC4* (1o).

Found at DOI: 10.1371/journal.pmed.0020199.sg001 (78 KB PPT).

Table S1. Real-Time TaqMan Probe and Primer Sequences

Found at DOI: 10.1371/journal.pmed.0020199.st001 (131 KB PDF).

Table S2. Differentially Regulated Genes in IBD Linkage Regions Identified on Microarrays

Found at DOI: 10.1371/journal.pmed.0020199.st002 (131 KB PDF).

References

1. Schreiber S, Rosenstiel P, Albrecht M, Hampe J, Krawczak M (2005) Genetics of Crohn disease, an archetypal inflammatory barrier disease. Nat Rev Genet 6: 376–388.
2. Podolsky DK (1991) Inflammatory bowel disease (1). N Engl J Med 325: 928–937.
3. Bonen DK, Cho JH (2003) The genetics of inflammatory bowel disease. Gastroenterology 124: 521–536.
4. Hugot JP, Chamaillard M, Zouali H, Lesage S, Cezard JP, et al. (2001) Association of NOD2 leucine-rich repeat variants with susceptibility to Crohn's disease. Nature 411: 599–603.
5. Ogura Y, Bonen DK, Inohara N, Nicolae DL, Chen FF, et al. (2001) A frameshift mutation in NOD2 associated with susceptibility to Crohn's disease. Nature 411: 603–606.
6. Hampe J, Cuthbert A, Croucher PJ, Mirza MM, Mascheretti S, et al. (2001) Association between insertion mutation in NOD2 gene and Crohn's disease in German and British populations. Lancet 357: 1925–1928.
7. Stoll M, Corneliussen B, Costello CM, Waetzig GH, Mellgard B, et al. (2004) Genetic variation in DLG5 is associated with inflammatory bowel disease. Nat Genet 36: 476–480.
8. Peltekova VD, Wintle RF, Rubin LA, Amos CI, Huang Q, et al. (2004) Functional variants of OCTN cation transporter genes are associated with Crohn disease. Nat Genet 36: 471–475.
9. van Dullemen HM, de Jong E, Slors F, Tytgat GN, van Deventer SJ (1998) Treatment of therapy-resistant perineal metastatic Crohn's disease after proctectomy using anti-tumor necrosis factor chimeric monoclonal antibody, cA2: Report of two cases. Dis Colon Rectum 41: 98–102.
10. Hillier LD, Lennon G, Becker M, Bonaldo MF, Chiapelli B, et al. (1996) Generation and analysis of 280,000 human expressed sequence tags. Genome Res 6: 807–828.
11. Lawrance IC, Fiocchi C, Chakravarti S (2001) Ulcerative colitis and Crohn's disease: Distinctive gene expression profiles and novel susceptibility candidate genes. Hum Mol Genet 10: 445–456.
12. Dieckgraefe BK, Stenson WF, Korzenik JR, Swanson PE, Harrington CA (2000) Analysis of mucosal gene expression in inflammatory bowel disease by parallel oligonucleotide arrays. Physiol Genomics 4: 1–11.
13. Dooley TP, Curto EV, Reddy SP, Davis RL, Lambert GW, et al. (2004) Regulation of gene expression in inflammatory bowel disease and correlation with IBD drugs: Screening by DNA microarrays. Inflamm Bowel Dis 10: 1–14.
14. Mah N, Thelin A, Lu T, Nikolaus S, Kuhbacher T, et al. (2004) A comparison of oligonucleotide and cDNA-based microarray systems. Physiol Genomics 16: 361–370.
15. Best WR, Becktel JM, Singleton JW, Kern F Jr (1976) Development of a Crohn's disease activity index. National Cooperative Crohn's Disease Study. Gastroenterology 70: 439–444.
16. Rachmilewitz D (1989) Coated mesalazine (5-aminosalicylic acid) versus sulphasalazine in the treatment of active ulcerative colitis: A randomised trial. BMJ 298: 82–86.
17. Truelove SC, Witts LJ (1955) Cortisone in ulcerative colitis; final report on a therapeutic trial. Br Med J 4947: 1041–1048.
18. Bates MD, Erwin CR, Sanford LP, Wiginton D, Bezerra JA, et al. (2002) Novel genes and functional relationships in the adult mouse gastrointestinal tract identified by microarray analysis. Gastroenterology 122: 1467–1482.
19. Boer JM, Huber WK, Sultmann H, Wilmer F, von Heydebreck A, et al. (2001) Identification and classification of differentially expressed genes in renal cell carcinoma by expression profiling on a global human 31,500-element cDNA array. Genome Res 11: 1861–1870.

20. Hoyle DC, Rattray M, Jupp R, Brass A (2002) Making sense of microarray data distributions. Bioinformatics 18: 576–584.
21. Lu T, Costello CM, Croucher PJ, Hasler R, Deuschl G, et al. (2005) Can Zipf's law be adapted to normalize microarrays? BMC Bioinformatics 6: 37.
22. Tavazoie S, Hughes JD, Campbell MJ, Cho RJ, Church GM (1999) Systematic determination of genetic network architecture. Nat Genet 22: 281–285.
23. Boguski MS, Lowe TM, Tolstoshev CM (1993) dbEST—Database for "expressed sequence tags." Nat Genet 4: 332–333.
24. Halgren RG, Fielden MR, Fong CJ, Zacharewski TR (2001) Assessment of clone identity and sequence fidelity for 1,189 IMAGE cDNA clones. Nucleic Acids Res 29: 582–588.
25. Waetzig GH, Seegert D, Rosenstiel P, Nikolaus S, Schreiber S (2002) p38 mitogen-activated protein kinase is activated and linked to TNF-alpha signaling in inflammatory bowel disease. J Immunol 168: 5342–5351.
26. Zhang W, Morris QD, Chang R, Shai O, Bakowski MA, et al. (2004) The functional landscape of mouse gene expression. J Biol 3: 21.
27. Geboes K (2004) Is histology useful for the assessment of the efficacy of immunosuppressive agents in IBD and if so, how should it be applied? Acta Gastroenterol Belg 67: 285–289.
28. Fraser AG, Marcotte EM (2004) A probabilistic view of gene function. Nat Genet 36: 559–564.
29. Neurath MF, Pettersson S, Meyer zum Buschenfelde KH, Strober W (1996) Local administration of antisense phosphorothioate oligonucleotides to the p65 subunit of NF-kappa B abrogates established experimental colitis in mice. Nat Med 2: 998–1004.
30. Rogler G, Brand K, Vogl D, Page S, Hofmeister R, et al. (1998) Nuclear factor kappaB is activated in macrophages and epithelial cells of inflamed intestinal mucosa. Gastroenterology 115: 357–369.
31. Schreiber S, Nikolaus S, Hampe J (1998) Activation of nuclear factor kappa B inflammatory bowel disease. Gut 42: 477–484.
32. Jono H, Lim JH, Chen LF, Xu H, Trompouki E, et al. (2004) NF-kB is essential for induction of CYLD, the negative regulator of NF-kB: Evidence for a novel inducible auto-regulatory feedback pathway. J Biol Chem 28: 28.
33. Dijkstra G, van Goor H, Jansen PL, Moshage H (2004) Targeting nitric oxide in the gastrointestinal tract. Curr Opin Investig Drugs 5: 529–536.
34. Iijima H, Neurath MF, Nagaishi T, Glickman JN, Nieuwenhuis EE, et al. (2004) Specific regulation of T helper cell 1-mediated murine colitis by CEACAM1. J Exp Med 199: 471–482.
35. Rubel C, Fernandez GC, Dran G, Bompadre MB, Isturiz MA, et al. (2001) Fibrinogen promotes neutrophil activation and delays apoptosis. J Immunol 166: 2002–2010.
36. Ergun S, Kilik N, Ziegeler G, Hansen A, Nollau P, et al. (2000) CEA-related cell adhesion molecule 1: A potent angiogenic factor and a major effector of vascular endothelial growth factor. Mol Cell 5: 311–320.
37. Hauck CR (2002) Cell adhesion receptors—Signaling capacity and exploitation by bacterial pathogens. Med Microbiol Immunol (Berl) 191: 55–62.
38. Dignass AU (2001) Mechanisms and modulation of intestinal epithelial repair. Inflamm Bowel Dis 7: 68–77.
39. Kuper H, Adami HO, Trichopoulos D (2000) Infections as a major preventable cause of human cancer. J Intern Med 248: 171–183.
40. Edwards FC, Truelove SC (1963) The course and prognosis of ulcerative colitis. Gut 41: 299–315.
41. Brostrom O, Lofberg R, Nordenvall B, Ost A, Hellers G (1987) The risk of colorectal cancer in ulcerative colitis. An epidemiological study. Scand J Gastroenterol 22: 1193–1199.
42. Joaquin M, Watson RJ (2003) The cell cycle-regulated B-Myb transcription factor overcomes cyclin-dependent kinase inhibitory activity of p57(KIP2) by interacting with its cyclin-binding domain. J Biol Chem 278: 44255–44264.
43. Komatsu K, Murata K, Kameyama M, Ayaki M, Mukai M, et al. (2002) Expression of S100A6 and S100A4 in matched samples of human colorectal mucosa, primary colorectal adenocarcinomas and liver metastases. Oncology 63: 192–200.
44. Arumugam T, Simeone DM, Schmidt AM, Logsdon CD (2004) S100P stimulates cell proliferation and survival via receptor for activated glycation end products (RAGE). J Biol Chem 279: 5059–5065.
45. Bronckart Y, Decaestecker C, Nagy N, Harper L, Schafer BW, et al. (2001) Development and progression of malignancy in human colon tissues are correlated with expression of specific Ca^{2+}-binding S100 proteins. Histol Histopathol 16: 707–712.
46. Vizoso FJ, Fagilde MC, Corte MD, Corte MG, Gava R, et al. (2003) Cytosolic levels of an estrogen-induced breast cancer-associated peptide (TFF1/pS2) in colorectal cancer: Clinical significance and relationship with steroid receptors. Int J Biol Markers 18: 301–310.
47. Lefebvre O, Chenard MP, Masson R, Linares J, Dierich A, et al. (1996) Gastric mucosa abnormalities and tumorigenesis in mice lacking the pS2 trefoil protein. Science 274: 259–262.
48. Amit S, Hatzubai A, Birman Y, Andersen JS, Ben-Shushan E, et al. (2002) Axin-mediated CKI phosphorylation of beta-catenin at Ser 45: A molecular switch for the Wnt pathway. Genes Dev 16: 1066–1076.
49. Liu C, Li Y, Semenov M, Han C, Baeg GH, et al. (2002) Control of beta-catenin phosphorylation/degradation by a dual-kinase mechanism. Cell 108: 837–847.
50. van Hogezand RA, Eichhorn RF, Choudry A, Veenendaal RA, Lamers CB (2002) . van Hogezand RA, Eichhorn RF, Choudry A, Veenendaal RA, Lamers CB (2002) Malignancies in inflammatory bowel disease: Fact or fiction? Scand J Gastroenterol (Suppl): 48–53.
51. Wyatt J, Oberhuber G, Pongratz S, Puspok A, Moser G, et al. (1997) Increased gastric and intestinal permeability in patients with Crohn's disease. Am J Gastroenterol 92: 1891–1896.
52. Wakabayashi M, Ito T, Mitsushima M, Aizawa S, Ueda K, et al. (2003) Interaction of lp-dlg/KIAA0583, a membrane-associated guanylate kinase family protein, with vinexin and beta-catenin at sites of cell-cell contact. J Biol Chem 278: 21709–21714.
53. Humbert P, Russell S, Richardson H (2003) Dlg, Scribble and Lgl in cell polarity, cell proliferation and cancer. Bioessays 25: 542–553.
54. Kucharzik T, Walsh SV, Chen J, Parkos CA, Nusrat A (2001) Neutrophil transmigration in inflammatory bowel disease is associated with differential expression of epithelial intercellular junction proteins. Am J Pathol 159: 2001–2009.
55. Massova I, Kotra LP, Fridman R, Mobashery S (1998) Matrix metalloproteinases: Structures, evolution, and diversification. FASEB J 12: 1075–1095.
56. von Lampe B, Barthel B, Coupland SE, Riecken EO, Rosewicz S (2000) Differential expression of matrix metalloproteinases and their tissue inhibitors in colon mucosa of patients with inflammatory bowel disease. Gut 47: 63–73.
57. Carraway KL, Ramsauer VP, Haq B, Carothers Carraway CA (2003) Cell signaling through membrane mucins. Bioessays 25: 66–71.
58. Wilson AJ, Gibson PR (1997) Epithelial migration in the colon: Filling in the gaps. Clin Sci (Lond) 93: 97–108.
59. Kuroda Si, Tokunaga C, Kiyohara Y, Higuchi O, Konishi H, et al. (1996) Protein-protein interaction of zinc finger LIM domains with protein kinase C. J Biol Chem 271: 31029–31032.
60. Nakagawa N, Hoshijima M, Oyasu M, Saito N, Tanizawa K, et al. (2000) ENH, containing PDZ and LIM domains, heart/skeletal muscle-specific protein, associates with cytoskeletal proteins through the PDZ domain. Biochem Biophys Res Commun 272: 505–512.

Patient Summary

Background Crohn's disease and ulcerative colitis are both inflammatory diseases of the gut, and there is a great deal of overlap between them in the symptoms that patients have. However, both diseases also have distinctly different clinical features that allow, in about 80% of patients, a differentiation between the two diseases. Although some genes have been shown to be associated with one or the other type of the disease, the cascade of events that leads to the symptoms we call disease is not fully understood.

Why Was This Study Done? New technology allows researchers to compare the function of many thousands of genes side by side. These researchers wanted to look at how the genes that are expressed (active) differ between samples taken from the guts of people with these two conditions.

What Did the Researchers Do and Find? They compared gut samples from people with inflammatory bowel disease with others taken from people with normal colon, and others with unrelated colonic diseases. They found differences in the pattern of expression of hundreds of genes between Crohn's disease and ulcerative colitis and samples from other diseases and normal colonic mucosa. The types of genes mostly altered were those involved in immunity (how the body protects itself against outside substances); those involved in mechanisms regulating how cells multiply (which can go wrong in cancer—one of the complications of ulcerative colitis) and those that control how the cells of the gut remain intact.

What Do These Findings Mean? They point to some new areas of research for these diseases, and possible targets for new drugs. They also indicate how complex these two diseases are, and suggest that there is not just one trigger for the diseases.

Where Can I Get More Information Online? National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health has an alphabetical list of digestive diseases and links to clinical trials in this area: http://digestive.niddk.nih.gov/ddiseases/
American and the German Crohn's and Colitis foundations:
http://www.ccfa.org; http://www.dccv.de
Kompetenznetz—Darmerkrankungen (in German):
http://www.kompetenznetz-ced.de
The researchers' home page:
http://www.ikmb.uni-kiel.de