Published in final edited form as:
Inflamm Bowel Dis. 2011 September ; 17(9): 1936–1942. doi:10.1002/ibd.21579.

Distinct and Overlapping Genetic Loci in Crohn's Disease and Ulcerative Colitis: Correlations with Pathogenesis

Matti Waterman, MD[*], **Wei Xu, PhD**[†], **Joanne M. Stempak, MSc**[*], **Raquel Milgrom, MD**[*], **Charles N. Bernstein, MD**[‡], **Anne M. Griffiths, MD**[§,‖], **Gordon R. Greenberg, MD**[*,§], **A. Hillary Steinhart, MD MSc**[*,§], and **Mark S. Silverberg, MD, PhD**[*,§]

[*]Zane Cohen Centre for Digestive Diseases, Mount Sinai Hospital IBD Group, Toronto, ON

[†]Dalla Lana School of Public Health, University of Toronto, ON

[‡]Division of Gastroenterology, University of Manitoba, Winnipeg, MB

[§]Faculty of Medicine, University of Toronto, Toronto ON

[‖]Division of Gastroenterology, Hospital for Sick Children, Toronto ON, Canada

Abstract

Background—A common genotypic basis for ulcerative colitis (UC) and Crohn's disease (CD) is implied by overlapping clinical characteristics, epidemiological studies, and association of genes with both UC and CD. We evaluated the overlap between CD and UC genetic loci stratified by pathogenetic pathways and by disease location.

Methods—The allele frequencies of six UC-associated and 34 CD-associated single nucleotide polymorphisms (SNPs) were determined in a Canadian IBD cohort (n = 2374). Differences between CD, UC, colon-only CD, ileal CD, and controls were analyzed controlling for ethnicity, age of diagnosis, and gender.

Results—In all, 21 of 34 CD-associated SNPs had similar allele frequencies in UC (n = 1230) and CD (n = 1144). Three of six UC-associated SNPs had significantly different frequencies in CD (n = 1144). Most of the divergence in allele frequency among CD and UC was noted in NOD2/autophagy pathway SNPs, while most SNPs with similar frequencies were in IL-22/23 Th17, adaptive immunity, and barrier pathways. Colon-only CD (n = 228) was compared with healthy controls: three of six UC SNPs (in *MST1*, *HLA-DRA*, and *IL-23R*) and 11 of 34 CD SNPs: in *IRGM*, *NOD2* (rs2066845), *CCNY*, *MST1*, *IL23R*, *PTPN22*, *C11orf30*, *ZNF365*, *PTPN2*, *PSMG1*, and rs1456893 were significantly associated. In all, 29 of 34 CD SNPs had similar allele frequencies in colonic CD compared with ileal CD (n = 366). All UC SNPs had similar frequencies in UC and colonic CD.

Conclusions—Our results suggest that CD and UC share common genetic associations related to impaired adaptive immunity and diverge in pathways of foreign antigen processing. Colon-only CD overlaps extensively with UC and considerably with ileal CD.

Keywords

Crohn's disease; ulcerative colitis; genotypic overlap; colon-only CD; IBD immunopathogenesis

Reprints: Dr. Mark S. Silverberg, Mount Sinai Hospital, Division of Gastroenterology, 600 University Ave., Rm. 441, Toronto ON M5G1X5, Canada (msilverberg@mtsinai.on.ca).

Additional Supporting Information may be found in the online version of this article.

The inflammatory bowel diseases (IBDs) are chronic inflammatory disorders of the intestine that are likely the result of a dysregulated immune response to the gut microflora in genetically predisposed individuals.[1] The two major forms of IBD, Crohn's disease (CD) and ulcerative colitis (UC), are often quite distinct. For example, CD complications due to penetration and stricturing occur in some 50% of patients[2] and only very rarely in UC.[3,4] These dissimilarities may reflect differences in the pathogenesis of the immune response between the two forms of IBD. Immunologically, CD is thought to be the result of an exaggerated T-helper cell 1 (Th1) response mounted against commensal bacteria[5] with elevation in the Th1 cytokine profile, including interferon (IFN)-gamma, interleukin (IL)-12 (composed of p35 and p40), and tumor necrosis factor alpha (TNF-α).[5] This bias towards a Th1 response is thought to be regulated through IL-17 production by Th17 regulatory T cells.[6] In contrast, the immunological response in UC is mediated by Th2 cells that have been activated by natural killer T (NK-T)-cells[5] with a distinct cytokine profile including IL-4 and IL-13 that have been found to be cytotoxic to epithelial cells.

Differences in the genetic predisposition to UC and CD have also been recognized from earlier genetic studies. Mutations in the gene encoding the NOD2 protein were the first genetic associations reported specifically for CD.[7,8] Mutations in NOD2 result in defects in bacterial antigen processing.[9] Other proteins that are important in the recognition and processing of bacterial components in the gut, such as the autophagy genes *ATG16L110* and *IRGM11*, also have genetic variants associated with CD but not UC.[6,12] Genome-wide association (GWA) studies in UC have demonstrated associations of several loci on chromosomes 1p31 (*IL23R*), 1p36 and 12q15 and in a region on chromosome 6p21.[13] Additionally, while initial association of *MST1* gene variants with CD was reported, other variations have now been associated with UC[14] along with the *ECM1* gene, which encodes for the extracellular matrix protein 1 with a plausible role in epithelial-stromal interaction and NF-κB signaling.[12]

Despite the unique features of CD and UC, there are also a number of overlapping aspects that suggest a common pathogenesis, particularly in early onset disease. For example, in approximately 10% of cases colonic IBD cannot be classified as CD or UC and is referred to as IBD unclassified (IBDU).[15] Individuals with CD have a higher rate of UC-affected individuals within their families and vice versa as compared to population rates, which suggests shared susceptibility determinants.[16] Additionally, many medical therapies are equally efficacious for both diseases.[17,18] It may therefore be hypothesized that there are at least some shared susceptibility factors that may then be affected by environmental exposures that ultimately determine the clinical phenotype.[19]

The aim of this study was to evaluate the degree of genotypic overlap between CD and UC in a large IBD cohort by investigating the frequencies of CD single nucleotide polymorphisms (SNPs) in a UC population and UC SNPs in a CD population. Determining the pathways in which the overlapping and divergent SNPs cluster will permit the generation of hypotheses as to the role these may play in disease pathogenesis.

Materials and Methods

Patient and Control Recruitment

All patients were recruited from either Mount Sinai Hospital or the Hospital for Sick Children in Toronto, Ontario and from the University of Manitoba Inflammatory Bowel Disease Clinical and Research Centre, in Winnipeg, Manitoba, Canada. IBD phenotype was determined using the Montreal Classification[15] and based on clinical, pathology, endoscopic, and radiological data. Unrelated healthy controls (HCs) were recruited from Mount Sinai Hospital and provided by the National Institute for Diabetes and Digestive and

Kidney Diseases (NIDDK) IBD Genetics Consortium (Pittsburgh, Yale, and Chicago sites) as well as through the Ontario Population Genomics Platform (Toronto, Canada). All study participants provided written informed consent. Study protocols were approved by the Research Ethics Boards of the participating centers.

Genotyping

DNA samples obtained from peripheral whole-blood samples collected from the study subjects and were genotyped using the Illumina GoldenGate custom SNP assay on Illumina BeadStation500G (San Diego, CA). We selectively determined the allele frequencies of 31 recently reported CD-associated SNPs, three *NOD2* SNPs, and six UC-associated SNPs[12,13,20,21,22] in the entire cohort of IBD patients and controls. The selected SNPs as well as their original association results in CD and UC are described in Supporting Information Table 1.

Statistical Analysis

IBD patients were grouped into all CD and all UC (which included IBDU) categories to determine differences in the minor allele frequencies between the two groups. In a separate analysis, we analyzed the CD subgroup with colon-only involvement (Montreal Classification L2) and compared the allele frequency in this subgroup to healthy controls and to ileal-only (Montreal Classification L1) CD and UC/IBDU patients. Descriptive statistics of demographic variables were generated using SAS v. 9.2 (SAS Institute, Cary, NC). PLINK v. 1.06[23] was used to obtain descriptive statistics of the SNPs such as the allele frequency and genotype distribution; it also was used to test for Hardy–Weinberg equilibrium (HWE) for each marker based on Pearson's chi-square test. Association analyses were applied to detect the associations with the candidate SNPs and phenotypic endpoints. Logistic regression models were applied for association analysis. Although we used an additive genetic model for primary analysis[24] we also explored dominant and recessive genetic models for sensitivity analysis. Throughout this report the *P*-values are the additive genetic model unless otherwise stated. Odds ratios (OR) and 95% confidence intervals (CI) were estimated for the disease compared to the control group. For CD versus UC, and colon-only CD versus UC or ileal CD, unconditional logistic regression models were applied as well as conditional logistic regression models adjusting for age of diagnosis, gender, and ethnicity. Two-sided statistical tests were applied.

Results

A total of 2374 IBD patients were genotyped (CD = 1144, UC/IBDU = 1230 [1140 UC, 90 IBDU]). Patient demographics as well as disease phenotypes are shown in Table 1. The UC and CD patient groups showed significant differences compared to the healthy control group in ethnicity and gender distribution (both $P < 0.0001$). Of the 1144 individuals with CD, 228 (19.9%) had colon-only involvement (Montreal Classification L2).

Of 34 previously associated CD SNPs tested, 21 were not significantly different in their frequencies between UC and CD patients. Of six UC SNPs tested, three did not have significantly different allele frequencies in CD-affected subjects. A multivariate analysis adjusting for ethnicity, age at diagnosis, and gender distribution as well as a Caucasian only analysis did not significantly affect the results (Supporting Information Table 2).

The differences in genetic variation between UC/IBDU and CD-affected subjects were evaluated by grouping SNPs according to their putative role in the pathogenesis of IBD. The SNP frequencies of genes that have a purported role in innate immunity including the bacterial recognition and autophagy pathways are shown in Table 2. For this analysis,

significantly different allele frequencies were found in UC/IBDU, compared to CD, for most (6/7) SNPs. Interestingly, we were not able to replicate the association between the SNP in the *LRRK2* gene and CD. Allele frequencies (in UC/IBDU, CD, and HC) of SNPs located in genes that involve Th-17 lymphocyte differentiation and the IL-23 receptor pathways are summarized in Table 3. Most (4/6) of the SNPs have similar frequencies in CD and UC/IBDU. There was a statistically significant difference in allele frequencies demonstrated for the SNPs in the genes *CCR6* and *STAT3* with UC/IBDU versus CD, but not versus HC. Other secondary immune response genetic variants previously associated with CD or UC were also investigated for difference in frequencies (Table 4). With the exception of the variants in the genes *HLA-DRA* and *PTPN22*, and all other SNPs (8/10) were similar in their frequencies between CD and UC/IBDU. The frequencies of genetic variants in genes with putative roles in the maintenance of the epithelial barrier integrity in our cohort are similar between UC/IBDU and CD, as shown in Supporting Information Table 3. Finally, we investigated genetic variants that have previously been highly associated with either UC or CD but with an unclear function or that are located within gene deserts. Their frequencies are summarized in Supporting Information Table 4. Of the 15 SNPs that were investigated, only six have shown difference in frequencies between CD and UC/IBDU. Four of the SNPs were within gene deserts and two were gene variants (*PSMG1* and *C11orf30*). The significance of these is yet to be determined.

Genetic Variants in Colonic CD Compared with HCs

The association of the above mentioned 40 SNPs (including the three *NOD2* variants) were evaluated in the cohort of patients with CD with colon-only involvement (Montreal L2, n = 228) compared with HCs (n = 1057). The results of the positively associated SNPs are outlined in Table 5. When compared with HCs (n = 1057), 3/6 UC-associated SNPs were associated with colon-only CD (SNPs within *MST1*, *IL-23R*, and *HLA-DRA* genes). Eleven of 34 CD SNPs were significantly associated with colon-only CD. Specifically, an association was found between SNPs in the autophagy and innate immunity genes *IRGM*, and *NOD2* (rs2066845). SNPs in the IL-23 receptor pathway and Th-17 regulatory lymphocyte differentiation pathway and secondary immune response pathways *CCNY*, *MST1*, *PTPN22*, *PTPN2*, and *IL23R* were also found to be associated with colonic CD. Three other SNPs in genes with unclear pathogenetic roles, *C11ORF30*, *ZNF365*, *PSMG1*, as well as a SNP without an identified gene, were associated with colonic CD.

Comparison of Colon-only CD to UC and Ileal CD

The frequencies of UC SNPs in colon-only CD patients were compared with UC patients as shown in Supporting Information Table 5. None of the six UC-associated SNPs demonstrated significantly different frequencies. Five of six SNPs that were previously associated with both UC and CD showed similar frequencies in L2 CD and UC. Further comparison of the CD-associated SNPs among ileal-only CD (Montreal L1, n = 366) and colon-only CD (Montreal L2, n = 228) showed significant though incomplete genotypic overlap between these subgroups (as shown in Supporting Information Table 6). Of the 34 CD-associated SNPs, 29 (85%) were not significantly different in their allele frequencies. The SNPs that did show significantly different frequencies are located within *NOD2* (two SNPs), *ATG16L1* (two SNPs), and *ICOSLG*.

Discussion

From among the 40 IBD-associated SNPs that were investigated, only 16 demonstrated significantly different allele frequencies between UC/IBDU and CD. Most of the difference is found in genes related to the innate immunity, pattern recognition, or autophagy pathways and the HLA Class II locus. This adds to the mounting evidence that bacterial sensing

mechanisms (NOD2) and the processing of bacterial components (autophagy) which leads to impaired bacterial clearance is likely more important in CD than in UC.[6,10] In contrast, in our sample of CD and UC/IBDU patients, gene variants that were previously associated with IBD and putatively play a role in the immune response amplification and perpetuation, including the IL-23/Th17 pathways and genes involved in barrier function, were similar in their prevalence between UC/IBDU and CD. Fourteen out of the 18 SNPs (78%) related to these pathogenetic pathways were similar in their frequencies. Our findings support previous reports that implicated genetic variations in several inflammatory pathways in the pathogenesis of both UC and CD. For example, GWA studies in CD and UC have identified variants in the IL-23 pathway, which plays an important role in the promotion and synthesis of Th1 and Th17 cytokines and further extension of the local immune response and tissue damage in IBD[25]. Specifically, the IL23 receptor gene, *IL12B* gene, which encodes and p40 subunit of IL-12, and *STAT3* and been associated with both UC and CD.[26–28] Even more recently, a GWAS in a very large European UC cohort identified an association between another IL17 pathway gene (*IL17REL*) and UC.[29] This is further corroborated by functional studies wherein most of the Th17/IL-23 axis related cytokines were increased in both UC and CD compared to healthy controls.[30] Other genes that have a role in the bridging between early innate immune response and late adaptive immune response such as *TNFSF15*[31,32] and genes that may have a role in the orchestration and amplification of the inflammatory response[33] have been associated in the past with both CD and UC—quite congruent with our findings in this IBD cohort.

Some prior genetic studies have described variations associated with both CD and UC. Franke et al[34] investigated CD-associated loci in a German cohort of UC patients and found that gene variations in *NKX2-3* and *CCNY* that had been previously associated with CD[35] also are associated with UC. Similarly, Fisher et al[12] reported strong association of five previously CD-associated loci with UC (gene variations in the genes *IL23R*, *MST1*, two SNPs within the *IL12B* gene, and *NKX2-3*). Another GWA study has also demonstrated association of other IL23 receptor gene associated loci in UC.[13] Moreover, a recent meta-analysis demonstrated an association between UC and several previously associated CD SNPs in the autophagy genes *ATG16L1* (rs2241880) and *IRGM* (rs13361189, rs4958849).[36]

These recent advances in the field of IBD genetics as well as in mucosal immunology have contributed significantly to our current understanding of the immunopathogenesis of IBD. While autophagy and bacterial sensing and processing mechanisms have been implicated in the pathogenesis of CD, the IL-23 pathway as well as the T-lymphocyte differentiation regulation through IL-17 producing T-helper regulatory cells have been associated with both CD and UC.[31,33]

In our CD patients with colonic involvement (Montreal L2) we were able to replicate the association of 11 of the 34 CD-associated SNPs, including one of the *NOD2* variants. This is in contrast to previous reports that were not able to show an association between many of the CD-associated loci[37] and, in particular, *NOD2* variants[38,39] and colon-only CD. This observation may be explained in part by our young patient population and the well-recognized phenomenon seen in early-onset CD where a colon predominant location precedes the evolution to ileal involvement.[40] Indeed, *NOD2* variants have been associated not only with ileal CD but also with ileocolonic CD.[41] Moreover, we did not find an association of *NOD2* SNP 13 with colon-only CD, and this is the SNP that was repeatedly shown to be associated with ileal-only CD.

CD-affected individuals with colon-only involvement (Montreal L2) showed almost complete overlap with UC of the genotypes that were investigated, with 6/6 of UC SNPs and 5/6 shared SNPs showing similar frequencies in our colonic IBD patients. It may be that the

relatively smaller sample size of our colon-only CD contributed to the lack of ability to discern significant differences in allele frequencies between UC and L2 CD. However, common genetic loci for colon-only CD and UC have been reported in the past. Several studies showed that specific HLA loci were associated with both types of IBD.[42,43] Our results suggest that at least for the SNPs investigated, colonic CD and UC have common genetic basis-a finding that may explain the location predisposition in both UC and CD of the colon.

In summary, studying the difference in allele frequencies of CD- and UC-associated SNPs in a large cohort of patients with IBD from Canada showed that UC and CD overlap significantly in their genetic associations. It is likely that factors outside of inherited DNA polymorphisms account for the clinical variability seen in IBD, Moreover, it may be difficult to develop a panel of genetic markers that will reliably distinguish CD from UC. Instead, other dynamic biomarkers such as gene expression or microbiome signatures may be more useful for understanding the clinical variability in IBD. The results of this study extend the hypothesis that CD and UC share common genetic variants that are related to impaired adaptive immune response and diverge genetically in mechanisms that concern early events in foreign antigen recognition and cellular processing. It should be emphasized that studying SNP frequencies provide only indirect evidence to the different and common pathogenesis of CD and UC and further functional as well as replication studies should be conducted to substantiate and understand the pathogenesis of CD and UC.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank the patients and coordinators involved in this research as well as Lucy Zhang for technical help with the DNA samples.

Supported by funding from the Crohn's and Colitis Foundation of Canada (CCFC), National Institutes of Health (NIH)/NIDDK (National Institute of Diabetes and Digestive and Kidney Diseases), and Dr. Silverberg's Gale and Graham Wright Research Chair in Digestive Diseases at Mount Sinai Hospital.

References

1. Podolsky DK. Inflammatory bowel disease. N Engl J Med. 2002; 347:417–429. [PubMed: 12167685]
2. Peyrin-Biroulet L, Loftus EV Jr, Colombel JF, et al. The natural history of adult Crohn's disease in population-based cohorts. Am J Gastroenterol. 2010; 105:289–297. [PubMed: 19861953]
3. Gumaste V, Sachar DB, Greenstein AJ. Benign and malignant colorectal strictures in ulcerative colitis. Gut. 1992; 33:938–941. [PubMed: 1644333]
4. Froines EJ, Palmer DL. Surgical therapy for rectovaginal fistulas in ulcerative colitis. Dis Colon Rectum. 1991; 34:925–930. [PubMed: 1914728]
5. Shih DQ, Targan SR. Immunopathogenesis of inflammatory bowel disease. World J Gastroenterol. 2008; 14:390–400. [PubMed: 18200661]
6. Cho JH. The genetics and immunopathogenesis of inflammatory bowel disease. Nat Rev Immunol. 2008; 8:458–466. [PubMed: 18500230]
7. Hugot JP, Chamaillard M, Zouali H, et al. Association of NOD2 leucine-rich repeat variants with susceptibility to Crohn's disease. Nature. 2001; 411:599–603. [PubMed: 11385576]
8. Ogura Y, Bonen DK, Inohara N, et al. A frameshift mutation in NOD2 associated with susceptibility to Crohn's disease. Nature. 2001; 411:603–606. [PubMed: 11385577]

9. Abraham C, Cho JH. Functional consequences of NOD2 (CARD15) mutations. Inflamm Bowel Dis. 2006; 12:641–650. [PubMed: 16804402]
10. Cummings JR, Cooney R, Pathan S, et al. Confirmation of the role of ATG16L1 as a Crohn's disease susceptibility gene. Inflamm Bowel Dis. 2007; 13:941–946. [PubMed: 17455206]
11. Parkes M, Barrett JC, Prescott NJ, et al. Sequence variants in the autophagy gene IRGM and multiple other replicating loci contribute to Crohn's disease susceptibility. Nat Genet. 2007; 39:830–832. [PubMed: 17554261]
12. Fisher SA, Tremelling M, Anderson CA, et al. Genetic determinants of ulcerative colitis include the ECM1 locus and five loci implicated in Crohn's disease. Nat Genet. 2008; 40:710–712. [PubMed: 18438406]
13. Silverberg MS, Cho JH, Rioux JD, et al. Ulcerative colitis-risk loci on chromosomes 1p36 and 12q15 found by genome-wide association study. Nat Genet. 2009; 41:216–220. [PubMed: 19122664]
14. Goyette P, Lefebvre C, Ng A, et al. Gene-centric association mapping of chromosome 3p implicates MST1 in IBD pathogenesis. Mucosal Immunol. 2008; 1:131–138. [PubMed: 19079170]
15. Silverberg MS, Satsangi J, Ahmad T, et al. Toward an integrated clinical, molecular and serological classification of inflammatory bowel disease: report of a Working Party of the 2005 Montreal World Congress of Gastroenterology. Can J Gastroenterol. 2005; 19(suppl A):5–36. [PubMed: 16151544]
16. Orholm M, Munkholm P, Langholz E, et al. Familial occurrence of inflammatory bowel disease. N Engl J Med. 1991; 324:84–88. [PubMed: 1984188]
17. Holtmann MH, Krummenauer F, Claas C, et al. Long-term effectiveness of azathioprine in IBD beyond 4 years: a European multicenter study in 1176 patients. Dig Dis Sci. 2006; 51:1516–1524. [PubMed: 16927148]
18. Behm BW, Bickston SJ. Efficacy of infliximab for luminal and fistulizing Crohn's disease and in ulcerative colitis. Curr Treat Options Gastroenterol. 2007; 10:171–177. [PubMed: 17547855]
19. Bridger S, Lee JC, Bjarnason I, et al. In siblings with similar genetic susceptibility for inflammatory bowel disease, smokers tend to develop Crohn's disease and non-smokers develop ulcerative colitis. Gut. 2002; 51:21–25. [PubMed: 12077086]
20. Imielinski M, Baldassano RN, Griffiths A, et al. Common variants at five new loci associated with early-onset inflammatory bowel disease. Nat Genet. 2009; 41:1335–1340. [PubMed: 19915574]
21. Kugathasan S, Baldassano RN, Bradfield JP, et al. Loci on 20q13 and 21q22 are associated with pediatric-onset inflammatory bowel disease. Nat Genet. 2008; 40:1211–1215. [PubMed: 18758464]
22. Barrett JC, Hansoul S, Nicolae DL, et al. Genome-wide association defines more than 30 distinct susceptibility loci for Crohn's disease. Nat Genet. 2008; 40:955–962. [PubMed: 18587394]
23. Purcell S, Neale B, Todd-Brown K, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet. 2007; 81:559–575. [PubMed: 17701901]
24. Freidlin B, Zheng G, Li Z, et al. Trend tests for case-control studies of genetic markers: power, sample size and robustness. Hum Hered. 2002; 53:146–152. [PubMed: 12145550]
25. Sarra M, Pallone F, Macdonald TT, et al. IL-23/IL-17 axis in IBD. Inflamm Bowel Dis. 2010; 16:1808–1813. [PubMed: 20222127]
26. Newman WG, Zhang Q, Liu X, et al. Genetic variants in IL-23R and ATG16L1 independently predispose to increased susceptibility to Crohn's disease in a Canadian population. J Clin Gastroenterol. 2009; 43:444–447. [PubMed: 19276991]
27. Márquez A, Mendoza JL, Taxonera C, et al. IL23R and IL12B polymorphisms in Spanish IBD patients: no evidence of interaction. Inflamm Bowel Dis. 2008; 14:1192–1196. [PubMed: 18383521]
28. Anderson CA, Massey DC, Barrett JC, et al. Investigation of Crohn's disease risk loci in ulcerative colitis further defines their molecular relationship. Gastroenterology. 2009; 136:523–529. [PubMed: 19068216]
29. Franke A, Balschun T, Sina C, et al. Genome-wide association study for ulcerative colitis identifies risk loci at 7q22 and 22q13 (IL17REL). Nat Genet. 2010; 42:292–294. [PubMed: 20228798]

30. Liu ZJ, Yadav PK, Su JL, et al. Potential role of Th17 cells in the pathogenesis of inflammatory bowel disease. World J Gastroenterol. 2009; 15:5784–5788. [PubMed: 19998498]
31. Shih DQ, Targan SR. Insights into IBD pathogenesis. Curr Gastroenterol Rep. 2009; 11:473–480. [PubMed: 19903423]
32. Brand S. Crohn's disease: Th1, Th17 or both? The change of a paradigm: new immunological and genetic insights implicate Th17 cells in the pathogenesis of Crohn's disease. Gut. 2009; 58:1152–1167. [PubMed: 19592695]
33. Lees CW, Satsangi J. Genetics of inflammatory bowel disease: implications for disease pathogenesis and natural history. Expert Rev Gastroenterol Hepatol. 2009; 3:513–534. [PubMed: 19817673]
34. Franke A, Balschun T, Karlsen TH, et al. Replication of signals from recent studies of Crohn's disease identifies previously unknown disease loci for ulcerative colitis. Nat Genet. 2008; 40:713–715. [PubMed: 18438405]
35. Weersma RK, Stokkers PC, Cleynen I, et al. Confirmation of multiple Crohn's disease susceptibility loci in a large Dutch-Belgian cohort. Am J Gastroenterol. 2009; 104:630–638. [PubMed: 19174780]
36. Palomino-Morales RJ, Oliver J, Gómez-García M, et al. Association of ATG16L1 and IRGM genes polymorphisms with inflammatory bowel disease: a meta-analysis approach. Genes Immun. 2009; 10:356–364. [PubMed: 19491842]
37. Hancock L, Beckly J, Geremia A, et al. Clinical and molecular characteristics of isolated colonic Crohn's disease. Inflamm Bowel Dis. 2008; 14:1667–1677. [PubMed: 18521924]
38. Cuthbert AP, Fisher SA, Mirza MM, et al. The contribution of NOD2 gene mutations to the risk and site of disease in inflammatory bowel disease. Gastroenterology. 2002; 122:867–874. [PubMed: 11910337]
39. Lesage S, Zouali H, Cézard JP, et al. CARD15/NOD2 mutational analysis and genotype-phenotype correlation in 612 patients with inflammatory bowel disease. Am J Hum Genet. 2002; 70:845–857. [PubMed: 11875755]
40. Meinzer U, Idesström M, Alberti C, et al. Ileal involvement is age dependent in pediatric Crohn's disease. Inflamm Bowel Dis. 2005; 11:639–644. [PubMed: 15973117]
41. Radford-Smith G, Pandeya N. Associations between NOD2/CARD15 genotype and phenotype in Crohn's disease — are we there yet? World J Gastroenterol. 2006; 12:7097–7103. [PubMed: 17131470]
42. Silverberg MS, Mirea L, Bull SB, et al. A population- and family-based study of Canadian families reveals association of HLA DRB1*0103 with colonic involvement in inflammatory bowel disease. Inflamm Bowel Dis. 2003; 9:1–9. [PubMed: 12656131]
43. Ahmad T, Armuzzi A, Bunce M, et al. The molecular classification of the clinical manifestations of Crohn's disease. Gastroenterology. 2002; 122:854–866. [PubMed: 11910336]

Table 1

Demographics of Study Subjects

	UC/IBDU (*n* = 1230)	CD (*n* = 1144)	Colon-only CD (*n* = 228)	Healthy Controls (*n* = 1057)
Ethnicity*:				
Caucasians (%)	82.2	90	85.4	93.9
Non-Caucasians:				
Asians (%)	12.8	5.6	7.9	4.6
Blacks (%)	1.5	1.1	1.2	0.2
Mixed (%)	1.5	1.7	1.8	1
Others	2	1.6	2.7	0.3
Male %**:	49.2	52.7	52.1	35.8
Age at diagnosis (%)				
A1 (≤16 years)	31.8	52.2	53.4	
A2 (17–40 years)	53.7	44.1	40.4	
A3 (>40 years)	14.5	3.6	6.2	
Median (range)	23 (1–73)	16 (2–62)	15 (2–62)	
CD behavior (%):	NA			
B1		47.6	83.8	
B2		25.5	9.2	
B3		24.3	6.6	
UC extent		NA	NA	
E1	2.5			
E2	22			
E3	75.5			
CD location (%):	NA		NA	
L1		31.9 (366)		
L2		19.9 (228)		
L3		44.9 (514)		
L4		19.3 (221)		
Smoking:				
Current	11.8	20.5	13.8	
Ex-smoker	19.4	6.6	11.4	
Never	68.8	72.9	74.8	
Perianal disease (% positive)	NA	30.6	35.4	

UC = ulcerative colitis, CD = Crohn's disease, IBDU = inflammatory bowel disease - unclassified, B = Montreal classification of CD behavior: B1 = inflammatory, B2 = stricturing, B3 = penetrating, E = Montreal classification of UC location: E1 = proctitis, E2 = left colitis, E3 = extensive, L = Montreal classification of CD anatomical location: L1 = ileal-only, L2 = colon-only, L3 = ileocolonic, L4 = proximal SB and upper GI, NA = not applicable.

* The *P*-value of the ethnicity comparison for CD vs. HC is 0.002, for UC/IBDU vs. HC is <0.0001, for CD vs. UC/IBDU is <0.0001, and L2 CD vs. HC is <0.0001.

** The *P*-value of the gender comparison for CD vs. HC is <0.0001, for UC/IBDU vs. HC is <0.0001, for CD vs. UC/IBDU is 0.24, for L2 CD vs. HC is <0.0001.

Table 2

Allele Frequencies of Pattern Recognition and Autophagy Genetic Variants in UC/IBDU and CD

Chr	Gene	Assoc IBD Subtype	SNP	Freq in CD *n* = 1144	Freq in UC/IBDU *n* = 1230	Freq in HC *n* = 1057	*P*-value UC/IBDU vs. CD	*P*-value CD vs. HC	*P*-value UC/IBDU vs. HC
2q27	***ATG16L1***	**CD**	**rs3828309**	**0.3976**	**0.4732**	**0.4876**	***P* = 3.79E-07**	***P* = 5.38E-09**	***P* = 0.4**
2q37.1	***ATG16L1***	**CD**	**rs2241880**	**0.3964**	**0.4725**	**0.4881**	***P* = 3.16E-07**	***P* = 2.77E-09**	***P* = 0.356**
5q33	***IRGM***	**CD**	**rs11747270**	**0.1549**	**0.1289**	**0.0978**	***P* = 0.0132**	***P* = 1.53E-08**	***P* = 0.001**
12q12	*LRRK2-MUC19*	CD	rs11175593	0.0403	0.0319	0.0277	*P* = 0.1347	*P* = 0.0208	*P* = 0.392
16q21	***NOD2***	**CD**	**rs2066844**	**0.0802**	**0.0377**	**0.0467**	***P* = 7.80E-9**	***P* = 2.27E-5**	***P* = 0.15**
16q21	***NOD2***	**CD**	**rs2066845**	**0.0530**	**0.0127**	**0.0172**	***P* = 2.27E-12**	***P* = 2.79E-9**	***P* = 0.22**
16q21	***NOD2***	**CD**	**rs2066847**	**0.0733**	**0.0138**	**0.0275**	***P* = 1.11E-16**	***P* = 1.47E-9**	***P* = 0.001**

Chr = chromosome band, CD = Crohn's disease, UC = ulcerative colitis, IBDU = inflammatory bowel disease - unclassified, HC = healthy controls, SNP = single nucleotide polymorphism, Freq = minor allele frequency.

SNPs in bold font, significantly different ($P < 0.05$) frequencies between CD and UC.

Table 3

Allele Frequencies of IL-23 Receptor Pathway and Th-17 Regulatory Lymphocyte Differentiation Pathway Genetic Variants in UC/IBDU and CD

Chr	Gene	Assoc IBD Subtype	SNP	Freq in CD *n* = 1144	Freq in UC/IBDU *n* = 1230	Freq in HC *n* = 1057	*P*-value UC/IBDU vs. CD	*P*-value CD vs. HC	*P*-value UC/IBDU vs. HC
1p31	*IL23R*	CD and UC	rs11465804	0.0287	0.0381	0.0567	*P* = 0.117	*P* = 1.98E-06	*P* = 0.001
1p31	*IL23R*	UC	rs10889677	0.3841	0.3842	0.3282	*P* = 0.754	*P* = 0.0003	*P* = 9.54E-05
5q33	*IL12B*	CD and UC	rs10045431	0.2542	0.2527	0.3094	*P* = 0.627	*P* = 5.39E-05	*P* = 6.05E-06
6q27	***CCR6***	**CD**	**rs2301436**	**0.4860**	**0.4559**	**0.4757**	***P* = 0.045**	***P* = 0.551**	*P* = **0.17**
17q21	*STAT3*	CD	rs744166	0.3947	0.3760	0.4016	*P* = 0.136	*P* = 0.085	*P* = 0.796
21q22	***ICOSLG***	**CD**	**rs762421**	**0.3786**	**0.4142**	**0.3766**	***P* = 0.015**	***P* = 0.011**	***P* = 0.865**

Chr = chromosome band, CD = Crohn's disease, UC = ulcerative colitis, IBDU = inflammatory bowel disease - unclassified, HC = healthy controls, SNP = single nucleotide polymorphism, Freq = minor allele frequency.

SNPs in bold font, significantly different ($P < 0.05$) frequencies between CD and UC.

Table 4

Allele Frequencies of Variants in Genes Involved in the Orchestration of the Secondary Immune Response in UC/IBDU and CD

Chr	Gene	Assoc IBD Subtype	SNP	Freq in CD *n* = 1144	Freq in UC/IBDU *n* = 1230	Freq in HC *n* = 1057	*P*-value UC/IBDU vs. CD	*P*-value CD vs. HC	*P*-value UC/IBDU vs. HC
1p13	***PTPN22***	**CD**	**rs2476601**	**0.0628**	**0.0799**	**0.0864**	***P* = 0.0372**	***P* = 0.0051**	***P* = 0.4513**
3p21	*MST1*	CD and UC	rs3197999	0.3348	0.3353	0.2876	*P* = 0.6164	*P* = 0.0001	*P* = 0.0007
3p21	*MST1*	UC	rs9858542	0.341	0.3426	0.2939	*P* = 0.8049	*P* = 0.0004	*P* = 0.0009
6p21.32	***HLA-DRA***	**UC**	**rs2395185**	**0.3530**	**0.2384**	**0.3255**	***P* = 3.27E-15**	***P* = 0.1617**	***P* = 1.85E-10**
9q32	*TNFSF15*	CD	rs6478108	0.2980	0.2781	0.3161	*P* = 0.1428	*P* = 0.2020	*P* = 0.0065
9q32	*TNFSF15*	CD	rs4263839	0.2883	0.2650	0.3016	*P* = 0.0628	*P* = 0.3346	*P* = 0.0051
10p11	*CCNY*	CD	rs17582416	0.3936	0.3746	0.3416	*P* = 0.193	*P* = 0.0004	*P* = 0.0255
10q24	*NKX23*	CD and UC	rs11190140	0.4590	0.4641	0.5057	*P* = 0.7952	*P* = 0.0023	*P* = 0.0051
18p11	*PTPN2*	CD	rs2542151	0.1777	0.1682	0.1553	*P* = 0.4016	*P* = 0.0541	*P* = 0.2685
20q13.3	*TNFRSF6B*	CD and UC	rs2315008	0.2774	0.2777	0.3012	*P* = 0.6385	*P* = 0.2190	*P* = 0.0902

Chr = chromosome band, CD = Crohn's disease, UC = ulcerative colitis, IBDU = inflammatory bowel disease - unclassified, HC = healthy controls, SNP = single nucleotide polymorphism, Freq = minor allele frequency.

SNPs in bold font, significantly different ($P < 0.05$) frequencies between CD and UC.

Table 5

Allele Frequency of Highly Significant IBD Associated SNPs in Crohn's Disease with Colonic Involvement and Healthy Controls

Chr	SNP	Gene	Assoc IBD Subtype	Freq in L2 CD *n* = 228	Freq in Healthy Controls = 1057	OR	*P*-value
1	rs2476601	*PTPN22*	CD	0.0540	0.0863	0.606	0.025
1	rs11465804	*IL23R*	CD and UC	0.0181	0.0567	0.309	0.001
1	rs10889677	*IL23R*	UC	0.4077	0.3282	1.387	0.002
3	rs3197999	*MST1*	CD and UC	0.3439	0.2870	1.230	0.018
3	rs9858542	*MST1*	UC	0.3484	0.2939	1.283	0.024
6	rs2395185	*HLA-DRA*	UC	0.2658	0.3255	0.744	0.013
5	rs11747270	*IRGM*	CD	0.1674	0.0978	1.837	<0.001
7	rs1456893	?	CD	0.2455	0.3100	0.721	0.007
10	rs17582416	*CCNY*	CD	0.4099	0.3416	1.330	0.007
10	rs10995271	*ZNF365*	CD	0.4638	0.3976	1.313	0.010
11	rs7927894	*C11orf30*	CD	0.4568	0.3942	1.287	0.016
16	rs2066845	*NOD2*	CD	0.0383	0.0172	2.334	0.005
18	rs2542151	*PTPN2*	CD	0.2014	0.1553	1.388	0.016
21	rs2836878	*PSMG1*	CD and UC	0.2185	0.2715	0.753	0.022

Chr = chromosome, SNP = single nucleotide polymorphism, CD = Crohn's disease, UC = ulcerative colitis, IBDU = inflammatory bowel disease-unclassified, L2 = colon-only CD location, Freq = minor allele frequency, OR = odds ratio.