Investigation of Crohn's Disease Risk Loci in Ulcerative Colitis Further Defines Their Molecular Relationship

Published in final edited form as:
Gastroenterology. 2009 February ; 136(2): 523–9.e3. doi:10.1053/j.gastro.2008.10.032.

Investigation of Crohn's Disease Risk Loci in Ulcerative Colitis Further Defines Their Molecular Relationship

CARL A. ANDERSON[*,1], **DUNECAN C. O. MASSEY**[‡,1], **JEFFREY C. BARRETT**[§], **NATALIE J. PRESCOTT**[||], **MARK TREMELLING**[‡], **SHEILA A. FISHER**[||], **RHIAN GWILLIAM**[§], **JEMIMA JACOB**[||], **ELAINE R. NIMMO**[¶], **HAZEL DRUMMOND**[¶], **CHARLIE W. LEES**[¶], **CLIVE M. ONNIE**[||], **CATHERINE HANSON**[#], **KATARZYNA BLASZCZYK**[||], **RADHI RAVINDRARAJAH**[§], **SARAH HUNT**[§], **DHIRAJ VARMA**[§], **NAOMI HAMMOND**[§], **GREGORY LEWIS**[‡], **HEATHER ATTLESEY**[‡], **NICK WATKINS**[**], **WILLEM OUWEHAND**[**], **DAVID STRACHAN**[‡‡], **WENDY MCARDLE**[§§], **CATHRYN M. LEWIS**[||], **The Wellcome Trust Case Control Consortium**, **ALAN LOBO**[||||], **JEREMY SANDERSON**[¶¶], **DEREK P. JEWELL**[##], **PANOS DELOUKAS**[§], **JOHN C. MANSFIELD**[#], **CHRISTOPHER G. MATHEW**[||], **JACK SATSANGI**[¶], and **MILES PARKES**[‡]

[*]Genetic and Genomic Epidemiology Unit, Wellcome Trust Centre for Human Genetics, University of Oxford, Roosevelt Drive, Oxford

[‡]Inflammatory Bowel Disease Research Group, Addenbrooke's Hospital, University of Cambridge, Cambridge

[§]The Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge

[||]Department of Medical and Molecular Genetics, King's College London School of Medicine, London

[¶]Gastrointestinal Unit, Division of Medical Sciences, School of Molecular and Clinical Medicine, University of Edinburgh, Western General Hospital, Edinburgh

[#]Department of Gastroenterology and Hepatology, University of Newcastle upon Tyne, Royal Victoria Infirmary, Newcastle upon Tyne

[**]Department of Haematology, University of Cambridge, National Blood Service Centre, Long Road, Cambridge

[‡‡]Division of Community Health Services, St. George's University of London, Cranmer Terrace, London

[§§]ALSPAC, Dept Social Medicine, University of Bristol, Oakfield House, Oakfield Grove, Bristol

[||||]Division of Molecular and Genetic Medicine, University of Sheffield Medical School, Royal Hallamshire Hospital, Sheffield

[¶¶]Department of Gastroenterology, Guy's and St. Thomas' NHS Foundation Trust, London

[##]Gastroenterology Unit, Radcliffe Infirmary, University of Oxford, Oxford, United Kingdom

Abstract

Address requests for reprints to: Miles Parkes, Inflammatory Bowel Disease Research Group, Addenbrooke's Hospital, University of Cambridge, Cambridge CB2 0QQ, United Kingdom. .

[1]C.A.A. and D.C.O.M. contributed equally to this work.

The authors disclose the following: Supported by the NIHR Cambridge Biomedical Research Centre, by the Wellcome Trust (to C.A.A.), and by the Medical Research Council (to D.C.O.M.).

Europe PMC Funders Author Manuscripts

Background & Aims—Identifying shared and disease-specific susceptibility loci for Crohn's disease (CD) and ulcerative colitis (UC) would help define the biologic relationship between the inflammatory bowel diseases. More than 30 CD susceptibility loci have been identified. These represent important candidate susceptibility loci for UC. Loci discovered by the index genome scans in CD have previously been tested for association with UC, but those identified in the recent meta-analysis await such investigation. Furthermore, the recently identified UC locus at *ECM1* requires formal testing for association with CD.

Methods—We analyzed 45 single nucleotide polymorphisms, tagging 29 of the loci recently associated with CD in 2527 UC cases and 4070 population controls. We also genotyped the UC-associated *ECM1* variant rs11205387 in 1560 CD patients and 3028 controls.

Results—Nine regions showed association with UC at a threshold corrected for the 29 loci tested ($P < .0017$). The strongest association ($P = 4.13 \times 10^{-8}$; odds ratio = 1.27) was identified with a 170-kilobase region on chromosome 1q32 that contains 3 genes. We also found association with *JAK2* and replicated a recently reported association with *STAT3*, further implicating the role of this signaling pathway in inflammatory bowel disease. Additional novel UC susceptibility genes were *LYRM4* and *CDKAL1*. Twenty of the loci were not associated with UC, and several appear to be specific to CD. *ECM1* variation was not associated with CD.

Conclusions—Collectively, these data help define the genetic relationship between CD and UC and characterize common, as well as disease-specific mechanisms of pathogenesis.

Genetic epidemiology data provide compelling evidence that the chronic inflammatory bowel diseases Crohn's disease (CD) and ulcerative colitis (UC) are related polygenic diseases. This has catalyzed a series of molecular studies aimed at gene identification in both diseases. Following the initial success of hypothesis-free linkage analyses and positional cloning in identifying *NOD2*/*CARD15* as the first CD susceptibility gene in 2001,1,2 further progress was eagerly anticipated. The technology of genome-wide association scanning (GWAS) is now yielding multiple susceptibility loci across the range of complex disease, and nowhere has it been more successfully applied than in CD.

Others and we have recently reported the results of index GWAS studies in CD3-11 and followed this by undertaking an international meta-analysis of GWAS studies with replication in a large independent sample set.12 In total, 28 new CD susceptibility genes or loci have been identified and confirmed by these studies, with an additional 10 loci showing nominal evidence of association (that is, modest association, which in the sample sets studied does not attain sufficient statistical support to withstand correction for multiple testing). Two major pathogenic themes have emerged: first, the evident importance of interleukin 23 and the T helper cell (Th)17 pathway of T-cell differentiation, where CD associations have been confirmed in multiple components including genes *IL23R, IL12B, JAK2, STAT3*, and *CCR6*; second, the critical contribution of defects in innate immunity has been confirmed, with autophagy genes *ATG16L1* and *IRGM* being added to *NOD2* as established susceptibility genes for CD. An emerging hypothesis is that CD may be driven by commensal bacteria infecting cells of the intestinal mucosa and that it is the failure of innate immune mechanisms to appropriately clear intracellular bacteria that leads to activation of adaptive immunity and inflammation.

Many other genes and loci that are not known to contribute to either Th17 or *NOD2*/autophagy pathways have also been identified. The function and pathogenic role of many of these are less well characterized but will now be subjected to detailed investigation.

The recent meta-analysis involving data from the index CD GWAS studies carried out in Great Britain, North America, and Belgium/France has led to the identification of several new susceptibility loci in CD, which require interrogation in UC. Most are of modest effect

size in CD, hence only being revealed in the meta-analysis and follow-up rather than in the individual index GWAS studies. Indeed, the CD susceptibility loci identified to date, with the notable exceptions of *NOD2* and *IL-23R*, mostly confer odds ratios (OR) of less than 1.4 — effect sizes typical of those we might expect to see in UC. This point does emphasize the need to study these loci in large panels.

In a previous report, we carried out a nonsynonymous single nucleotide polymorphism (SNP) scan in UC, which confirmed the known contribution of the major histocompatibility complex, and identified a new UC locus at *ECM1*.13 In that study, we also followed-up the confirmed CD hits identified in our index genome-wide scan and identified variants in *IL23R, IL12B, NKX2-3*, and the *MST1* locus on chromosome 3p21 as showing significant association with UC. A parallel German case-control study, based on following up hits from our CD data set within the Wellcome Trust Case Control Consortium (WTCCC) study, identified association between UC and variants in *STAT3, BSN, NKX2-3, HERC2*, and *CCNY* and a borderline association at *PTPN2*.14 The study also detected nominal association to interleukin *(IL)12B*, thus replicating our finding of association between variants in this locus and UC risk.

Variants within nel-like 1 precursor (*NELL1*, chromosome 11p15) have previously been associated with UC risk.8 In addition, *MDR1* polymorphisms have been associated with UC risk,15,16 and, although not all studies have replicated this association,17 a recent meta-analysis firmly implicated this locus as a determinant both of disease susceptibility and of severity.18

GWAS studies have yet to be reported for UC, but are underway. While awaiting such studies, a potentially fruitful approach to the identification of new UC susceptibility genes is the interrogation of all GWAS-derived CD loci in UC sample sets. This has the additional benefit of identifying which loci are shared and which are specific to CD and UC, hence illuminating the genetic relationship between the different forms of inflammatory bowel disease (IBD).

To identify additional UC susceptibility genes, we investigated the contribution to UC of 28 loci implicated in CD by the recent international meta-analysis.12 In addition, we formally tested the *ECM1* variant recently associated with UC for evidence of association with CD. Using the large, well-characterized panels of patients and controls available to the UK IBD Genetics Consortium, our data identify several new UC susceptibility loci.

Patients and Methods

Subjects/Genotyping

A total of 3026 UC and 1727 CD patients were recruited in participating centers across the United Kingdom. Diagnosis of UC and CD was based on standard clinical, endoscopic, radiologic, and histologic criteria. Patient demographics, family history, smoking history, and subphenotype data were ascertained by combination of questionnaire and case note review. Details of the cases analyzed after quality control are presented in Table 1. Ethics committee approval was granted in each of the four lead UK centers, and all participants gave their signed consent for inclusion.

UC cases—We genotyped 45 SNPs from 29 distinct candidate loci (primer sequences in Supplementary Table 1; see supplementary Table 1 online at www.gastrojournal.org). These comprised 28 loci from the meta-analysis (19 confirmed and 9 nominally replicated) plus one (*HERC2*) that was reported to be associated with UC by Franke et al14 in their follow-up of unconfirmed CD hits from the WTCCC. Genotyping was undertaken with iPLEX

chemistry on a matrix-assisted laser desorption/ionization time-of-flight MassARRAY platform (Sequenom, San Diego, CA), with the 45 SNPs split across 2 genomic pools. Genotyping was carried out in 2 batches, with 2120 individuals in the first batch and 906 in the second. For each batch of cases, and on a per-pool basis, individuals missing ≥20% of genotypes were removed. We also removed 66 UC patients identified as non-white ethnicity in our previous nonsynonymous SNP scan13 and 27 likely duplicates (identical by state [IBS] >0.98). Of UC cases, 2527 satisfied quality control criteria for at least 1 genomic pool. We removed 2 SNPs (rs2872507 and rs744166) with ≥20% missing genotype data from the largest batch of cases. Post-quality control, the genotype success rate was 0.932. Taking account of the small number of SNPs genotyped in this study, these measures represent stringent data quality control. All SNPs were in Hardy-Weinberg equilibrium.

CD cases—The *ECM1* variant rs11205384 previ ously associated with UC was genotyped in 1726 CD cases using the Taqman biallelic discrimination system (Applied Biosystems, Carlsbad, CA) with an ABI 7900HT analyzer. Following inspection of the genotype clusters, genotypes were available for 1560 individuals (9.6% missing data).

Control panel 1 (UC experiment)—One thousand one hundred thirty-two UK population controls were drawn from the National Blood Service (NBS) collection. The validity of blood donor DNA as a source of control genotypes for cases/control association studies has previously been established by the WTCCC.9 Genotyping was carried out using Sequenom iPLEX technology as part of the replication phase of the CD meta-analysis,12 with additional genotyping of a further 110 individuals. Forty-four of 45 SNPs genotyped in the UC cases were available for comparison in this cohort (rs916977 [*HERC2*] was not genotyped). Individuals of non-white ethnicity were identified and removed from further analysis.

Control panel 2 (UC experiment)—Of the 45 SNPs genotyped in our UC cases, 26 are present on the Affymetrix (Santa Clara, CA) SNP array 5.0 and 19 on the Illumina HumanHap550 platform (San Diego, CA). Use of these arrays in our earlier experiments allowed us to incorporate the previously generated control genotyping data for these markers in the current analysis.

Thus, 1480 individuals from the 1958 British Birth Cohort (58C) and 1458 from the NBS (total 2938) had been genotyped by the WTCCC and provided control data for the 26 of 45 SNPs present on the Affymetrix SNP array 5.0 platform. The 1480 58C subjects had also been genotyped using the Illumina HumanHap550 platform by the Sanger Institute and provided data for the remaining 19 SNPs. Again, individuals of non-white ethnicity were identified and removed prior to analysis. To allow a genotype concordance check between Sequenom iPLEX, Affymetrix SNP array 5.0, and Illumina HumanHap550 genotypes, we separately regenotyped 376 of the NBS DNAs and 294 of the 58C DNAs using the iPLEX platform, across all SNPs in the current study. Concordance for genotype calls between iPLEX and the Affymetrix array was 99.38% and with the Illumina array was 99.36%.

Control panel 3 (CD experiment)—We previously genotyped rs11205387 (*ECM1*) using Taqman technology (Applied Biosystems) in 1894 NBS individuals (1465 available after quality control).13

Control panel 4 (CD experiment)—As part of the same project, we genotyped 1728 NBS individuals for rs11205387 (*ECM1*) using iPLEX chemistry (1563 after quality control).13

Statistical Analysis

Cochran-Armitage trend tests, implemented through PLINK,19 were used to detect case-control association. For the UC experiment, *P* values less than .0017 are significant following Bonferroni correction for the number of independent loci (29) tested. Because only a single SNP was tested for association in our CD experiment, a *P* value less than .05 is statistically significant.

We inspected 3 subphenotypes of UC (disease extent defined macroscopically as extensive, left-sided, or proctitis only), 3 disease modifiers (age at diagnosis, smoking status at diagnosis, and family history of disease), and 1 disease outcome measure (surgery). A linear regression model was used to test for association between the genotyped SNPs and age at diagnosis (in whole years). For the remaining within-case analyses, individuals were partitioned into 2 groups according to affection status. Individuals smoking at diagnosis were contrasted to nonsmokers/exsmokers at diagnosis. Cochran-Armitage trend tests were used to detect association between each binary phenotype and the genotyped SNPs.

Power calculations were performed using the online Genetic Power Calculator (http://pngu.mgh.harvard.edu/~purcell/gpc/)20 to allow a better interpretation of negative association results. Calculations were undertaken assuming a population prevalence of 0.0024 for UC,21 a multiplicative disease model, a false-positive rate (α) of .0017, and an r^2 of 1 between the marker and disease locus. For each SNP, power was estimated using the given case/control ratio and the estimated allelic risk for CD from the combined case/control and transmission disequilibrium test analysis from Barrett et al 200812 (ie, assuming equivalent effect size for UC as seen for CD).

Results

UC Experiment

Of the 45 SNPs included in the study, 13 showed significant association with UC (see supplementary Table 2 online at www.gastrojournal.org), and these comprise 9 independent loci (Table 2). The candidate region showing the strongest association was 1q32. Both SNPs genotyped within this locus showed robust evidence of association ($P_{rs2297909} = 4.13 \times 10^{-8}$, $P_{rs11584383} = 5.71 \times 10^{-7}$). A further 3 loci are novel associations with UC (*JAK2, LYRM4*, and *CDKAL1*). Both SNPs genotyped within the *JAK2* locus show evidence of association ($P_{rs10758669} = 1.02 \times 10^{-5}$, $P_{rs7849191} = .0015$). Only a single SNP was genotyped within *LYRM4* ($P_{rs12529198} = 1.10 \times 10^{-4}$) and *CDKAL1* ($P_{rs6908425} = 1.89 \times 10^{-4}$).

Our study also provides the first independent replication of association between UC and variants within loci encoding *IL18RAP*,22 *CCNY*,14 and *STAT3*14 (Table 2). Reassuringly from a technical perspective, given the known sample overlap with our earlier nonsynonymous SNP study of 905 UC cases and 1465 controls,13 we observed strong association with 2 positive controls, *IL12B* and the *MHC* locus. Importantly, for each locus showing association, we observed the same risk-increasing allele as reported previously for CD.

We observed nominal evidence of association ($P < .05$ but $> .0017$) at 8 of the remaining 20 loci. Our power to detect association (assuming the same effect size as documented in CD) at each of the 19 remaining CD risk loci is given in Table 3. Estimated power varied widely between SNPs, and, for those loci for which we had little power, analyses in larger cohorts will be required to elucidate their role in UC. The previously reported association between *HERC2* and UC risk was not replicated in our data ($P = .085$).

We saw no evidence of association between the genotyped SNPs and any of the UC subphenotypes, disease modifiers, or outcomes that we investigated (see supplementary Table 3 online at www.gastrojournal.org). This is true even assuming a nonconservative significance threshold of $P < .0017$, ie, uncorrected for the number of subphenotypes tested.

CD Experiment

There was no evidence of association between the *ECM1* variant rs11205387 and CD ($P = .269$). In our CD cases, the minor allele frequency was 26.8%, and control panels 3 and 4 had a minor allele frequency of 25.9% and 25.5%, respectively.

Discussion

Through follow-up of discoveries from the recent CD meta-analysis, this study identifies 4 new susceptibility loci for UC. In addition, we provide replicated association and thereby confirmation of 3 recently reported UC loci. These results therefore add significantly to the body of evidence for shared pathways between the 2 diseases, an observation of considerable importance both in understanding diseases pathogenesis and in discovering new therapeutic targets.

Choice of appropriate statistical thresholds for significance in association studies is important in understanding the validity of the findings. The field now recognizes the need for stringent thresholds to distinguish true signals from false-positive findings. In the current study, we only tested loci showing confirmed or nominal association with CD, for which the prior probability of association with UC is significantly elevated. Accounting for the 29 independent loci tested, Bonferroni correction suggests that $P < .0017$ is appropriate. Of note, all the new UC loci that we identify meet this criterion by at least an order of magnitude.

The strongest association in this study was with 2 SNPs in tight linkage disequilibrium on chromosome 1q32, establishing this association in UC unequivocally for the first time. Following up WTCCC hits in CD, Franke et al14 had previously found some evidence for association between this locus and UC, but at $P = .0017$ this was of borderline significance given the 50 loci tested. Three genes map to the 170-kilobase region (as defined by HapMap recombination hot spots) highlighted by the association signal. These are *C1orf106, KIF21B*, and *CACNA1S*. Fine mapping is required to identify which gene is relevant to UC (and indeed CD) pathogenesis.

One of the strongest new signals was observed at a locus encoding Janus kinase 2 (*JAK2*) on chromosome 9p24. In addition, we confirmed association between UC and signal transducer and activator of transcription 3 (*STAT3*) as recently reported by Franke et al.14 These findings thus confirm that the *JAK-STAT* pathway is a common feature of both UC and CD pathogenesis. This is a focal point in the downstream transmission of signals from cell surface receptors to the nucleus to modify transcription of various genes. Several cytokines and other immunoactive molecules utilize this signal transduction pathway. These include interferons, epidermal growth factor, IL-5, IL-6, and notably also the IL12/23 axis within which variants are also associated in common with both UC and CD.10,13 *JAK2* and *STAT3* play a key role in *IL23R* signaling, and *STAT3* is critical for maturation of naïve CD4+ve T cells to the proinflammatory Th17 cells increasingly recognized as central to driving inflammation in IBD.

The genes highlighted by the 2 other novel UC association signals are currently poorly characterized. *LYRM4* (*C6orf149*) on chromosome 6p25 codes for a mitochondrial

ribosomal protein with sequence homology to *NADH*,23 and *CDKAL1* on chromosome 6p22 is noteworthy for being recently confirmed as a type 2 diabetes susceptibility gene.24

Our study provides the first independent replication of association between UC and variants in *IL18RAP* and *CCNY*. The IL-18 receptor accessory protein (*IL18RAP*) is involved in IL-18 signaling and has sequence homology to the IL-1 receptor accessory protein (*IL1RACP*). IL-18 is released by macrophages and, with IL-12, induces cell-mediated immunity following microbial infection.25-27 As well as association with CD, variants in *IL18RAP* have also been reported to be associated with celiac disease.22,28 *CCNY* (Cyclin fold protein-1) has also previously shown association with both CD and UC. The protein product belongs to the cyclin protein superfamily and contains a protein-binding domain that plays a role in cell-cycle and transcription control by regulating cyclin-dependent kinases.

None of the UC loci that we have identified showed significant subphenotype association. With regard to disease extent, this is perhaps unsurprising given that these are all also CD susceptibility loci. As generic IBD loci, it would be surprising if any were associated with one subclass of UC as defined by extent of colonic involvement but not others. Such subphenotype specific loci are more likely to be associated with UC alone and to derive from forthcoming GWAS studies in UC, although a surprising feature has been the paucity of such effects in CD.10

We do not see significant evidence of association at the previously reported UC susceptibility locus, *HERC2* ($P = .085$), even though our power to detect association at rs916977 given the reported UC effect size (OR, 1.46) is estimated to be 100% (assuming a multiplicative disease model, a population prevalence of 0.0024,21 and a falsepositive rate [α] of .0017). The minor allele frequency in the Franke et al 2008 population controls is 0.106, yet, in our combined controls, the frequency is 0.140 (similar to the case frequencies of both studies). The minor allele frequency of this SNP in the CEU HapMap B36 data is 0.133. It appears that the unusually low minor allele frequency at this SNP in the control samples from the Franke et al 2008 study underlies the unreplicated association at this locus. Additional, high-powered, case-control, and family-based association studies are therefore needed to elucidate fully the role of *HERC2* in UC.

A number of loci previously associated with CD did not show evidence of association with UC. For some loci, this may reflect a lack of power because extremely large sample sets would be required to reliably detect small effects with ORs <1.15. For others, our UC sample set had good or high power to detect an effect comparable with that seen in CD (Table 3). Of note, such loci include those encoding genes *ICOSLG* and *CCR6*. Both encode proteins that appear to play a key role in T-cell activation and differentiation. Therefore, it appears that CD-specific risk loci are not limited to disruption of innate immune pathways such as autophagy and *NOD2*.

We compared effect sizes for the shared UC/CD loci but did not observe any evidence of significant heterogeneity (Cochrane Q test: $P < .05$). Although there is some uncertainty in these effect size estimates, we can rule out scenarios in which any of these loci has a substantially different effect in UC vs CD. Overall, both phenotypes conform to the expectation of many genes of modest effect.

ECM1 was not associated with CD in the current study. We had 100% power to detect association at this SNP (assuming a multiplicative allelic OR of 1.23, a population prevalence of 0.00145,21 and $\alpha = .05$). Even allowing for a much reduced OR of 1.10, our power to detect association to this SNP is 82%. It therefore appears that variants within *ECM1* specifically confer susceptibility to UC risk.

What are the implications of this study with regard to the overall genetic architecture of the IBDs and the molecular relationship between CD and UC? Clearly, a complete understanding must await full, well-powered GWAS experiments in UC to compare with the GWAS studies in CD. In the meantime, interrogation of CD GWAS hits in our UC panel has identified a number of shared loci but has also found some key differences. Of the shared pathways, the Th17 inflammatory axis is particularly noteworthy and highlighted by the fact that a number of its molecular components show association with susceptibility to CD and UC. This is in the context of earlier reports highlighting innate immunity pathways of *NOD2* and autophagy as specific to CD and the current study suggesting that *ECM1* is UC specific. The precise causal variants of the 9 loci shared between UC and CD have not yet been defined, and their functional significance remains to be elucidated. These are immediate priorities in the fast moving field of IBD genetics, and their resolution should further illuminate both the underlying mechanisms of chronic intestinal inflammation and the cellular processes that lead to the distinct phenotypes of CD and UC.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The authors thank the Medical Research Council and Wellcome Trust for funding the 1958 British Birth Cohort, the National Association for Colitis and Crohn's Disease and the Wellcome Trust, which supported case collections, and all subjects who contributed samples plus consultants and nursing staff across the United Kingdom who helped with recruitment of study subjects.

Abbreviations used in this paper

GWAS	genome-wide association scanning
SNP	single nucleotide polymorphism

References

1. Hugot J-P, 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]
2. Ogura Y, Bonen DK, Inohara N, et al. A frame shift mutation in NOD2 associated with susceptibility to Crohn's disease. Nature. 2001; 411:603–606. [PubMed: 11385577]
3. Yamazaki K, McGovern D, Ragoussis J, et al. Single nucleotide polymorphisms in TNFSF15 confer susceptibility to Crohn's disase. Hum Mol Genet. 2005; 14:3499–3506. [PubMed: 16221758]
4. Duerr RH, Taylor KD, Brant SR, et al. A genome-wide association study identifies IL23R as an inflammatory bowel disease gene. Science. 2006; 314:1461–1463. [PubMed: 17068223]
5. Rioux JD, Xavier RJ, Taylor KD, et al. Genome-wide association study identifies new susceptibility loci for Crohn's disease and implicates autophagy in disease pathogenesis. Nat Genet. 2007; 39:596–604. [PubMed: 17435756]
6. Hampe J, Franke A, Rosenstiel P, et al. A genome-wide association scan of nonsynonymous SNPs identifies a susceptibility variant for Crohn's disease in ATG16L1. Nat Genet. 2007; 39:207–211. [PubMed: 17200669]
7. Libioulle C, Louis E, Hansoul S, et al. Novel Crohn's disease locus identified by genome-wide association maps to a gene desert on 5p13.1 and modulates expression of PTGER4. PLoS Genet. 2007; 3:e58. [PubMed: 17447842]
8. Franke A, Hampe J, Rosenstiel P, et al. Systematic association mapping identifies NELL1 as a novel IBD disease gene. PLoS ONE. 2007; 2:e691. [PubMed: 17684544]

9. The Wellcome Trust Case Control Consortium. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. Nature. 2007; 447:661–678. [PubMed: 17554300]
10. 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]
11. Raelson JV, Little RD, Ruether A, et al. Genome-wide association study for Crohn's disease in the Quebec Founder population identifies multiple validated disease loci. Proc Natl Acad Sci U S A. 2007; 104:14747–14752. [PubMed: 17804789]
12. 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]
13. 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]
14. 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]
15. Schwab M, Schaeffeler E, Marx C, et al. Association between the C3435T MDR1 gene polymorphism and susceptibility for ulcerative colitis. Gastroenterology. 2003; 124:26–33. [PubMed: 12512026]
16. Brant SR, Panhuysen CI, Nicolae D, et al. MDR1 Ala893 polymorphism is associated with inflammatory bowel disease. Am J Hum Genet. 2003; 73:1282–1292. [PubMed: 14610718]
17. Croucher PJ, Mascheretti S, Foelsch UR, et al. Lack of association between the C3435T MDR1 gene polymorphism and inflammatory bowel disease in two independent Northern European populations. Gastroenterology. 2003; 125:1919–1921. [PubMed: 14727637]
18. Annese V, Valvano MR, Palmieri O, et al. Multidrug resistance 1 gene in inflammatory bowel disease: a meta-analysis. World J Gastroenterol. 2006; 12:3636–3644. [PubMed: 16773678]
19. 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]
20. Purcell S, Cherny SS, Sham PC. Genetic power calculator: design of linkage and association genetic mapping studies of complex traits. Bioinformatics. 2003; 19:149–150. [PubMed: 12499305]
21. Rubin GP, Hungin AP, Kelly PJ, et al. Inflammatory bowel disease: epidemiology and management in an English general practice population. Aliment Pharmacol Ther. 2000; 14:1553–1559. [PubMed: 11121902]
22. Zhernakova A, Festen EM, Franke L, et al. Genetic analysis of innate immunity in Crohn's disease and ulcerative colitis identifies two susceptibility loci harboring CARD9 and IL18RAP. Am J Hum Genet. 2008; 82:1202–1210. [PubMed: 18439550]
23. Mungall AJ, Palmer SA, Sims SK, et al. The DNA sequence and analysis of human chromosome 6. Nature. 2003; 425:805–811. [PubMed: 14574404]
24. Zeggini E, Weedon MN, Lindgren CM, et al. Replication of genome-wide association signals in UK samples reveals risk loci for type 2 diabetes. Science. 2007; 316:1336–1341. [PubMed: 17463249]
25. Kohno K, Kurimoto M. Interleukin 18, a cytokine which resembles IL-1 structurally and IL-12 functionally but exerts its effect independently of both. Clin Immunol Immunopathol. 1998; 86:11–15. [PubMed: 9434792]
26. Munder M, Mallo M, Eichmann K, et al. Murine macrophages secrete interferon γ upon combined stimulation with interleukin (IL)-12 and IL-18: a novel pathway of autocrine macrophage activation. J Exp Med. 1998; 187:2103–2108. [PubMed: 9625771]
27. Miettinen M, Matikainen S, Vuopio-Varkila J, et al. Lactobacilli and streptococci induce interleukin-12 (IL-12), IL-18, and γ interferon production in human peripheral blood mononuclear cells. Infect Immun. 1998; 66:6058–6062. [PubMed: 9826398]
28. Hunt KA, Zhernakova A, Turner G, et al. Newly identified genetic risk variants for celiac disease related to the immune response. Nat Genet. 2008; 40:395–402. [PubMed: 18311140]

Europe PMC Funders Author Manuscripts

Table 1

Demographic Details of 4077 IBD Patients

	CD (n = 1560)	UC (n = 2527)
Median age at diagnosis (*y*)	29.7	36.8
Sex (*female/male*)	933/617	1263/1264
Smoking at diagnosis (%)[a]		
Never	44.4	52.5
Ex-smoker	20.3	33.9
Current	35.3	13.6
Surgery (%)	60.6	15.0
Location/extent (%)[b]	Ileal: 36.9 Colonic: 26.4 Ileocolonic: 36.7 Perianal: 28.1	Proctitis only: 18.6 Left sided: 34.5 Extensive: 46.9
Behavior (%)	Stenosing: 41.6 Penetrating: 14.0	

[a]Smoking data available for 88% CD and 57% UC cases.

[b]Phenotypic data available for 88% CD and 68% UC cases. Left-sided location is defined as distal to the splenic flexure, and extensive disease is defined as proximal to the splenic flexure.

Europe PMC Funders Author Manuscripts

Table 2

Loci Showing Significant Association With Ulcerative Colitis

				UC cases		Control panel 1		Control panel 2		Combined					CD
SNP	**Chr**	**Gene**	**Risk allele**	**RAF**	**N**	**RAF**	**N**	**RAF**	**N**	**RAF controls**	***P* value**	**OR**	**OR L95**	**OR U95**	**OR**
rs11584383	1q32	*KIF21B*	T	0.724	2479	0.679	1117	0.684	2932	0.682	5.71×10^{-7}	1.220	1.129	1.319	1.19
rs917997	2q11	*IL18RAP*	T	0.242	2451	0.216	1110	0.214	1376	0.215	.0013	1.168	1.063	1.284	1.07
rs10045431	5q33	*IL12B*	C	0.736	2434	0.711	1100	0.699	1339	0.704	5.21×10^{-4}	1.169	1.070	1.278	1.19
rs12529198	6p25	*LYRM4*	G	0.070	2489	0.055	1115	0.053	2929	0.053	1.10×10^{-4}	1.330	1.149	1.539	1.15
rs6908425	6p22	*CDKAL1*	C	0.800	2453	0.778	1102	0.770	2932	0.772	1.89×10^{-4}	1.182	1.083	1.289	1.16
rs3763313	6p21	*BTNL2* *SLC26A3* *HLA-DRB1* *HLA-DQA*	C	0.218	2460	0.193	1111	0.182	1375	0.187	1.41×10^{-4}	1.211	1.097	1.336	1.12
rs10758669	9p24	*JAK2*	C	0.380	2445	0.350	1118	0.325	1346	0.336	1.02×10^{-5}	1.207	1.112	1.311	1.14
rs17582416	10p11	*CCNY*	G	0.366	2464	0.355	1108	0.328	2882	0.335	4.27×10^{-4}	1.143	1.061	1.232	1.18
rs12948909	17q21	*STAT3*	A	0.760	2474	0.727	1111	0.731	2927	0.730	1.33×10^{-4}	1.172	1.080	1.272	1.10

NOTE. UC cases include all individuals with UC whose genotyping data passed QC. Control panel 1: UK population controls genotyped using Sequenom iPLEX technology as part of the recent Crohn's disease meta-analysis; control panel 2: WTCCC controls genotyped on the Affymetrix SNP-array 5.0 platform and the 1958 British Birth Cohort subset also genotyped on the Illumina HumanHap-550 platform. *P* values are those from a Cochran-Armitage trend test.

Chr, chromosome; RAF, risk increasing allele frequency.

Europe PMC Funders Author Manuscripts

Gastroenterology. Author manuscript; available in PMC 2009 August 01.

Table 3

Loci Showing No Significant Association With Ulcerative Colitis

				UC cases		Combined controls					
SNP	**CHR**	**Gene**	**CD risk allete**	**RAF**	**N**	**RAF controls**	**N**	***P* values**	**CD OR**	**Power (%)**	**N cases for 80% power**
rs7927894	11q13	*C11orf30*	T	0.420	2447	0.392	4026	.0019	1.13	57	3525
rs10188217[a]	2p16	*PUS10*	C	0.538	2406	0.507	2497	.0024	1.14	54	3634
rs1456893	7p12		A	0.698	2456	0.678	4005	.0130	1.18	85	2223
rs762421	21q22	*ICOSLG*	G	0.415	2434	0.393	4045	.0165	1.16	81	2373
rs1551398	8q24		A	0.630	2465	0.609	2475	.0334	1.14	51	3929
rs7758080	6q25		G	0.276	2342	0.257	2491	.0352	1.07	5	17,290
rs4807569	19p13		C	0.218	2425	0.203	4047	.0363	1.12	28	5851
rs991804	17q12	*CCL2, CCL7*	C	0.739	2466	0.721	2492	.0428	1.09	11	10,760
rs11175593	12q12	*LRRK2, MUC19*	T	0.019	2455	0.014	4060	.0541	1.49	41	4578
rs780094	2p23	*GCKR*	T	0.399	2464	0.382	4002	.0590	1.10	29	5891
rs2274910	1q23	*ITLN2*	C	0.691	2417	0.675	2472	.0790	1.17	67	2985
rs3764147	13q14		G	0.222	2424	0.210	4017	.1075	1.23	95	1664
rs1736135	21q21		T	0.588	2432	0.575	4032	.1624	1.15	74	2702
rs7749278	6q27	*CCR6*	T	0.519	2466	0.525	4046	.4382	1.15	76	2635
rs17309827	6p25	*SLC22A23*	T	0.631	2468	0.628	4045	.7070	1.07	9	12,060
rs2872507[b]	17q12	*ORMDL3*	A	0.490	643	0.484	4050	.7175	1.15	21	1880
rs2476601	1p13	*PTPN22*	G	0.905	2471	0.903	2483	.7426	1.26	54	3720
rs9286879	1q24		G	0.243	2334	0.245	2490	.8002	1.15	45	4100
rs8098673	18q11		C	0.324	2467	0.325	2493	.9300	1.06	4	20,950

NOTE. UC cases include all individuals with UC whose genotyping data passed QC. *P* values are those from a Cochran-Armitage trend test. Power calculated assuming a multiplicative disease model, a population prevalence of 0.0024,21 and a false-positive rate of 0.0017. Crohn's disease odds ratios are taken from Barrett et al (2008)12 and represent the estimated allelic risk from the combined case/control and TDT analyses. RAF, risk increasing allele frequency.

[a] SNP selection for the current experiment was carried out using preliminary results from Barrett et al 2008,12 and an alternative SNP to that reported in the final analysis (rs13003464) was selected at this locus (for the purpose of calculating power, we use the OR reported for rs13003464).

[b] Genotyping did not pass QC procedures for 1 of 2 case batches; hence, only 643 cases were available at this SNP.