THE UNIVERSITY *of* EDINBURGH

Edinburgh Research Explorer

Common variants at five new loci associated with early-onset inflammatory bowel disease

OPEN ACCESS

Common variants at five new loci associated with early-onset inflammatory bowel disease

Marcin Imielinski[1,28], **Robert N Baldassano**[2,3,28], **Anne Griffiths**[4,28], **Richard K Russell**[5,28], **Vito Annese**[6,28], **Marla Dubinsky**[7,28], **Subra Kugathasan**[8,28], **Jonathan P Bradfield**[1], **Thomas D Walters**[4], **Patrick Sleiman**[1], **Cecilia E Kim**[1], **Aleixo Muise**[4], **Kai Wang**[1], **Joseph T Glessner**[1], **Shehzad Saeed**[9], **Haitao Zhang**[1], **Edward C Frackelton**[1], **Cuiping Hou**[1], **James H Flory**[1], **George Otieno**[1], **Rosetta M Chiavacci**[1], **Robert Grundmeier**[2,10], **Massimo Castro**[11], **Anna Latiano**[11], **Bruno Dallapiccola**[12], **Joanne Stempak**[13], **Debra J Abrams**[3], **Kent Taylor**[7], **Dermot McGovern**[7], **Western Regional Research Alliance for Pediatric IBD**[30], **International IBD Genetics Consortium**[30], **Melvin B Heyman**[14], **George D Ferry**[15], **Barbara Kirschner**[16], **Jessica Lee**[17], **Jonah Essers**[17], **Richard Grand**[17], **Michael Stephens**[18], **Arie Levine**[19], **David Piccoli**[2,3], **Johan Van Limbergen**[20], **Salvatore Cucchiara**[21], **Dimitri S Monos**[22], **Stephen L Guthery**[23], **Lee Denson**[24], **David C Wilson**[25], **Struan F A Grant**[1,2,26], **Mark Daly**[27], **Mark S Silverberg**[13,29], **Jack Satsangi**[20,29], and **Hakon Hakonarson**[1,2,26,29]

[1]Center for Applied Genomics, The Children's Hospital of Philadelphia, Philadelphia, Pennsylvania, USA

[2]Department of Pediatrics, University of Pennsylvania School of Medicine, Philadelphia, Pennsylvania, USA

[3]Division of Gastroenterology, The Children's Hospital of Philadelphia, Philadelphia, Pennsylvania, USA

[4]The Hospital for Sick Children, University of Toronto, Toronto, Ontario, Canada

[5]Paediatric Gastroenterology and Nutrition, Royal Hospital for Sick Children, Glasgow, UK

[6]Gastroenterology Unit, Department of Medical Sciences, Istituto di Ricovero e Cura a Carattere Scientifico (IRCCS) 'Casa Sollievo della Sofferenza' (CSS), San Giovanni Rotondo, Italy

[7]Departments of Pediatrics and Common Disease Genetics, Cedars Sinai Medical Center, Los Angeles, California, USA

[8]Department of Pediatrics, Emory University School of Medicine and Children's Health Care of Atlanta, Atlanta, Georgia, USA

[9]Department of Pediatrics, University of Alabama at Birmingham, Alabama, USA

[10]Center for Biomedical Informatics, The Children's Hospital of Philadelphia, Philadelphia, Pennsylvania, USA

[11]Gastroenterology Unit, Bambino Gesù Children's Hospital, IRCCS, Rome and SIGENP, Italy

[12]Mendel Institute, La Sapienza University of Rome, IRCCS-CSS Hospital, S. Giovanni Rotondo, Italy

[13]Mount Sinai Hospital Inflammatory Bowel Disease Centre, University of Toronto, Toronto, Ontario, Canada

[14]Department of Pediatrics, University of California, San Francisco, California, USA

[15]Department of Pediatrics, The Baylor College of Medicine, Houston, Texas, USA

[16]University of Chicago Comer Children's Hospital, Chicago, Illinois, USA

[17]Division of Gastroenterology, Children's Hospital, Boston, Massachusetts, USA

[18]Department of Pediatrics, Medical College of Wisconsin, Milwaukee, Wisconsin, USA

[19]Pediatric Gastroenterology Unit, Wolfson Medical Center, Tel Aviv University, Tel Aviv, Israel

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

[21]Pediatric Gastroenterology and Liver Unit, La Sapienza University of Rome and SIGENP, Italy

[22]Department of Pathology, The Children's Hospital of Philadelphia, Philadelphia, Pennsylvania, USA

[23]Department of Pediatrics, University of Utah School of Medicine and Primary Children's Medical Center, Salt Lake City, Utah, USA

[24]The Center for Inflammatory Bowel Disease, Division of Gastroenterology, Cincinnati Children's Hospital Medical Center, Cincinnati, Ohio, USA

[25]Paediatric Gastroenterology and Nutrition, Royal Hospital for Sick Children, Edinburgh and Child Life and Health, University of Edinburgh, UK

[26]Division of Human Genetics, The Children's Hospital of Philadelphia, Philadelphia, Pennsylvania, USA

[27]Center for Human Genetic Research, Massachusetts General Hospital, Boston, Massachusetts, USA

Abstract

The inflammatory bowel diseases (IBD) Crohn's disease and ulcerative colitis are common causes of morbidity in children and young adults in the western world. Here we report the results of a genome-wide association study in early-onset IBD involving 3,426 affected individuals and 11,963 genetically matched controls recruited through international collaborations in Europe and North America, thereby extending the results from a previous study of 1,011 individuals with early-onset IBD[1]. We have identified five new regions associated with early-onset IBD susceptibility, including 16p11 near the cytokine gene *IL27* (rs8049439, $P = 2.41 \times 10^{-9}$), 22q12 (rs2412973, $P = 1.55 \times 10^{-9}$), 10q22 (rs1250550, $P = 5.63 \times 10^{-9}$), 2q37 (rs4676410, $P = 3.64 \times 10^{-8}$) and 19q13.11 (rs10500264, $P = 4.26 \times 10^{-10}$). Our scan also detected associations at 23 of 32 loci previously implicated in adult-onset Crohn's disease and at 8 of 17 loci implicated in

adult-onset ulcerative colitis, highlighting the close pathogenetic relationship between early- and adult-onset IBD.

Crohn's disease and ulcerative colitis are chronic inflammatory disorders of the gastrointestinal tract that most commonly arise during the second and third decades of life. Incidence, family, twin and phenotype concordance studies suggest that IBD is highly heritable, albeit complex, spurring an ongoing search for genetic factors that confer susceptibility to this disease[2,3]. Genome-wide association studies (GWASs) applying high-density SNP array technology have greatly expanded the number of genetic factors implicated in IBD pathogenesis to include 32 loci associated with Crohn's disease and 17 associated with ulcerative colitis, spanning pathways involved in adaptive (*IL23R, IL10, IL12B, STAT3*) and innate (*CARD15, ATG16L1, IRGM*) immunity[4–7].

Most genetic analyses in IBD have been performed in adult-onset disease[2,3]. Early-onset IBD, however, has unique characteristics of phenotype, severity and familiality[8,9], features that provide support for the search for loci that may be specific to early-onset disease. In addition, because early-onset IBD has a stronger familial component than the adult disease, studies targeting this subgroup potentially provide additional power to identify genes that contribute modest effects, as illustrated by the success of our previous scan in identifying 20q13 and 21q22 as IBD loci[1].

We now report results from the largest GWAS conducted so far in early-onset IBD (Fig. 1). Our IBD discovery cohort (DC-IBD) comprised 2,413 individuals of European ancestry with IBD (cases), including 1,636 with Crohn's disease (DC-CD), 724 with ulcerative colitis (DC-UC) and 53 with IBD of unclassified type (IBD-U), and 6,158 genetically matched controls, and was genotyped on the Illumina HumanHap550 platform. Affected individuals were recruited from multiple centers from four geographically discrete countries and diagnosed before their nineteenth birthday according to standard IBD diagnostic criteria (Supplementary Table 1). Our study extends a previous IBD GWAS that was based on a subset of these cases (1,011 IBD cases, including 647 with Crohn's disease, 317 with ulcerative colitis and 47 with IBD-U; Supplementary Table 2)[1]. An independent replication cohort (RC1) of 482 early-onset IBD cases (289 with Crohn's disease, 120 with ulcerative colitis and 73 with IBD-U) and 1,696 genetically matched controls was gathered from the Children's Hospital of Philadelphia (CHOP) health system and collaborating centers. We refer to Crohn's disease and ulcerative colitis subanalyses of data set RC1 as RC1-CD and RC1-UC, respectively. A second replication cohort (RC2-CD) of 531 Crohn's disease cases diagnosed in childhood and 4,109 controls was assembled by the International IBD Genetics Consortium (IIBDGC). This cohort is based on subsets of data from genome-wide scans generated by the US National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK)[10], the Wellcome Trust Case Control Consortium (WTCCC)[11] and a Belgian-French[12] collaboration, which have been combined previously in a large-scale meta-analysis[5]. We computed *P* values in each cohort by comparing single-marker allele frequencies using χ^2 statistics on SNPs that passed quality-control criteria. We conducted meta-analysis across multiple studies using a *Z*-score transformation. See Online Methods for more detailed descriptions of cohorts and methods used in this study.

We first searched for previously unreported signals, including loci that met stringent genome-wide significant ($P < 5 \times 10^{-8}$) and suggestive ($P < 1 \times 10^{-6}$) thresholds in our three discovery scans (DC-CD, DC-UC and DC-IBD). For confirmation of these loci, we sought evidence for replication in two independent early-onset cohorts (RC1 and RC2-CD). Lastly, we combined our discovery and replication cohorts for Crohn's disease (Table 1), ulcerative colitis (Table 2) and IBD (Table 3) in a meta-analysis.

Analysis of DC-CD identified a region on 16p11 as the single new genome-wide significant locus; the most significant SNP in the block of linkage disequilibrium (LD) containing this locus, rs1968752, yielded a value of $P = 2.09 \times 10^{-8}$, with the minor allele (A) conferring risk (odds ratio (OR) = 1.25 (1.16–1.36)). We observed nominal replication of rs1968752 in the RC2-CD data set ($P = 0.036$, OR = 1.09 (0.94–1.27)) and a trend for association in the combined analysis of the replication cohorts (RC1-CD and RC2-CD), using two-sided P values ($P = 0.059$; Supplementary Table 3).

Analysis of our combined IBD discovery scan yielded suggestive association at the 16p11 locus (rs8049439, $P = 2.38 \times 10^{-7}$, OR = 1.20 (1.12–1.28)). rs8049439 is located ~200 kb upstream of rs1968752, and the two loci are in strong LD ($r^2 = 0.796$). rs8049439 showed a statistically significant association ($P = 0.00144$, OR = 1.14 (1.00–1.30)) in the RC2-CD data set, whereas meta-analysis of rs8049439 in the discovery data set and the two replication cohorts (RC1, RC2-CD) reached genome-wide significance in both IBD ($P = 2.41 \times 10^{-9}$) and Crohn's disease ($P = 2.87 \times 10^{-9}$; Tables 1 and 3). In addition, analysis of the available data in a meta-analysis data set for adult-onset Crohn's disease[5] demonstrated that an LD proxy for these two SNPs, rs4788084 ($r^2 = 0.83$ to rs1968752, and $r^2 = 0.86$ to rs8049439), was also associated with Crohn's disease ($P = 0.0035$; Supplementary Table 4). Of note, we found that the risk-conferring minor allele (G) at rs8049439 shares haplotypes with the risk-conferring ancestral allele (A) of the type 1 diabetes SNP rs4788084 (HapMap CEU, $r^2 = 0.864$)[13].

The LD block incorporating rs1968752, rs8049439 and rs4788084 contains multiple genes, including *IL27, CCDC101, CLN3, EIF3C, NUPR1, SULT1A1* and *SULT1A2*. Of these, we considered *IL27*, which encodes an immunomodulatory cytokine that regulates adaptive immunity responses, to be the most plausible candidate gene for susceptibility to IBD. Analysis of *IL27* gene expression in lymphoblastoid cell lines (LCLs) obtained from ten healthy individuals homozygous for rs1968752 (that is, either A/A or C/C) showed that those individuals with two copies of the risk allele (A/A) had a several-fold decrease in *IL27* expression relative to those with two copies of the nonrisk allele (C/C; $P = 0.0031$; Fig. 2a), suggesting that this SNP may exert a regulatory effect on *IL27* gene expression. In addition, colonic expression of *IL27* was significantly lower in 30 individuals with early-onset Crohn's disease than in 11 healthy controls ($P < 0.05$; Fig. 2b). These effects remained significant after correction for medication use and histological inflammation score.

When examining the colonic expression of other genes at this locus, we also detected significantly lower expression of *SULT1A1* and *SULT1A2* in both early-onset Crohn's disease ($P < 0.05$, $P < 0.001$) and ulcerative colitis ($P < 0.0001$, $P < 0.0001$) as compared with healthy controls (Fig. 2c,d). *SULT1A1* and *SULT1A2* encode sulfotransferases that catalyze sulfate conjugation of catecholamines, phenolic drugs and neurotransmitters. These biological functions make *SULT1A1* and *SULT1A2* less attractive as IBD candidate genes. We also observed a strongly expressed quantitative trait locus (eQTL) for *EIF3C* expression in LCLs at this locus (lod score = 5–8) on the basis of publicly available data[14]; however, we did not observe altered *EIF3C* expression in Crohn's disease or ulcerative colitis cases relative to healthy controls, nor did we detect an eQTL for *EIF3C* in colonic tissue at SNPs in this region (Supplementary Figs. 1 and 2). *EIF3C* encodes a eukaryotic translation initiation factor that forms part of the basic translational machinery, which also makes *EIF3C* less likely to be an IBD candidate gene. Additional allele-specific expression effects were not observed for the other genes at this locus either in our LCL and colonic expression data sets (Supplementary Fig. 1) or in the public database[14]. Taken together, these results point most strongly to *IL27* as a candidate gene associated with early-onset IBD; however, further functional and fine-mapping studies are warranted to confirm this and rule out the involvement of other genes at this locus.

The first of the two suggestive early-onset disease associations was identified in the DC-IBD cohort at 22q12 (rs2412973, $P = 9.14 \times 10^{-7}$, OR = 1.18 (1.10–1.26)). This SNP replicated in RC1 ($P = 0.0052$, OR = 1.23 (1.05–1.43)) and RC2-CD ($P = 0.016$, OR = 1.15 (1.01–1.31)), yielding $P = 1.55 \times 10^{-9}$ in a meta-analysis across all three early-onset IBD cohorts (Table 3). Meta-analysis of rs2412973 across the data sets DC-CD, RC1-CD and RC2-CD also reached genome-wide significance ($P = 3.77 \times 10^{-8}$). rs2412973 also showed association in the CD meta-analysis data set with an age of onset primarily in adulthood[5] ($P = 0.000953$; Supplementary Table 4). rs2412973 is located within *HORMAD2*, an open reading frame whose putative functions include mitotic checkpoints, chromosome synapsis and DNA repair. The LD block incorporating rs2412973 also contains *MTMR3*, encoding myotubularin-related protein-3, and *LIF*, encoding leukemia inhibitory factor, a cytokine that stimulates differentiation in leukocytes. We observed a significant difference in colonic *MTMR3* expression in biopsies from individuals with ulcerative colitis as compared with controls ($P < 0.001$), but not in those from individuals with Crohn's disease (Fig. 2e), and we did not detect colonic eQTL for *MTMR3* near rs2412973. Other genes in the LD block did not exhibit significant expression effects (Supplementary Fig. 3).

The second suggestive association was identified only in DC-UC at 2q37 (rs4676410, $P = 1.70 \times 10^{-7}$, OR 1.41 (1.24–1.61)). This SNP showed only a trend for replication in the small RC1-UC cohort ($P = 0.0611$, OR = 1.38 (1.08–1.77)), but yielded genome-wide significance in the combined analysis ($P = 3.64 \times 10^{-8}$; Table 2). rs4676410 lies within *GPR35*, which encodes an orphan receptor primarily expressed in the intestine of humans and rats. Other genes in the LD block of rs4676410 include *CAPN10, KIF1A* and *RNPEPL1*. *CAPN10* colonic gene expression was significantly lower in ulcerative colitis cases than in controls ($P < 0.05$; Fig. 2f). *CAPN10* encodes a Ca^{2+}-regulated thiol-protease involved in cytoskeletal remodeling and signal transduction. We did not observe significant expression effects in the remaining genes (Supplementary Figs. 1 and 2).

We next combined our discovery DC-IBD and replication cohorts (RC1, RC2-CD) for a genome-wide meta-analyses of early-onset Crohn's disease, ulcerative colitis and IBD. These analyses yielded two new loci achieving genome-wide significance. The first new signal is associated with IBD (rs10500264, $P = 4.26 \times 10^{-10}$) and is located at 19q13 in a small block of LD devoid of known genes lying within 50 kb of *SLC7A10* and *CEBPA*. Notably, rs10500264 showed only nominal association in the adult-onset Crohn's disease meta-analysis ($P = 0.0217$), suggesting that this locus may be weighted more toward early-onset disease. The second new signal, rs1250550, lies on 10q22 inside the *ZMIZ1* gene and is associated with both Crohn's disease ($P = 4.41 \times 10^{-10}$) and IBD combined ($P = 5.63 \times 10^{-9}$; Tables 1 and 3). In addition to showing significance in the early-onset meta-analysis, rs1250550 robustly associates in the majority adult-onset Crohn's disease meta-analysis[5] ($P = 3.27 \times 10^{-5}$). *ZMIZ1* encodes a PIAS-like protein that interacts with Smad4 to regulate Smad3 transcription and modulate transforming growth factor-β signaling[15]. Despite achieving robust significance in our meta-analysis, these loci merit replication in an independent cohort.

We conducted a meta-analysis of our discovery and replication cohorts to determine association with the 49 previously reported IBD loci implicated in adult-onset disease, determining nominal ($P < 0.05$) and Bonferroni-corrected ($P < 0.001$, correcting for 49 hypotheses; Supplementary Table 5) significance. Of 32 previously confirmed loci associated with adult-onset Crohn's disease, 29 were nominally significant and 21 were significant after Bonferroni correction in meta-analysis of DC-CD and RC1-UC data sets. Of eight additional Crohn's disease loci that attained nominal significance ($P < 0.05$) in the previously reported majority adult-onset meta-analysis[5], two showed significant association with early-onset Crohn's disease, namely the *IL18R1-IL18RAP* locus on 2q12 (rs917997, P

= 6.84×10^{-5}, $Z = 3.98$) and the C-C motif chemokine (*CCL*) gene cluster on 17q12 (rs991804, $P = 2.31 \times 10^{-4}$, $Z = -3.68$; Table 4). We found that 13 of 17 previously identified adult-onset ulcerative colitis loci showed nominal significance, and 8 were significant after Bonferroni-corrected meta-analysis of DC-CD and RC1-UC data sets, including *IL23R* on 1p31 and *IL26* on 12q15 (Supplementary Table 5). Our data also supported the association of loci on 20q13 and 21q22 with early-onset IBD, as previously reported[1] on the basis of analyses of a subset of our discovery cohort (Supplementary Note).

We also evaluated previously reported loci associated with adult-onset Crohn's disease for association with early-onset ulcerative colitis, and vice versa. Examining 32 known Crohn's disease signals in our ulcerative colitis cohort implicated two loci that had not previously been associated with adult-onset ulcerative colitis susceptibility in early-onset ulcerative colitis: *ICOSLG* on 21q22 (rs762421, $P = 2.54 \times 10^{-5}$, $Z = 4.21$) and *ORMDL3* on 17q12 (rs2872507, $P = 7.62 \times 10^{-4}$, $Z = 3.37$; Supplementary Table 5). When examining the association of early-onset Crohn's disease with 17 previously reported adult-onset ulcerative colitis signals, we detected association only with the ulcerative colitis gene *IL10* on 1q32.1 (rs3024505, $P = 0.00048$, $Z = 3.49$), suggesting that this locus may also play a role in early-onset Crohn's disease susceptibility.

Our study adds insight into the pathogenic mechanisms mediating early-onset IBD and its close relationship with adult-onset disease. In particular, identification of *IL27* as a candidate gene for Crohn's disease susceptibility lends further support to the involvement of the T-helper 17 pathway[16,17] in pathogenesis of Crohn's disease, complementing gene discoveries in other genome scans (*IL23R, STAT3, JAK2, IL12B*)[2,3,5]. In addition, our discovery of five new IBD susceptibility loci through analysis of a genetically enriched early-onset disease cohort underscores the validity of this approach in the study of complex disease.

Methods

Methods and any associated references are available in the online version of the paper at http://www.nature.com/naturegenetics/.

Accession codes. Gene Expression Omnibus (GEO): colonic gene expression data set, GSE10616. Entrez Gene: *IL27*, 246779; *SULT1A1*, 171150; *SULT1A2*, 601292; *EIF3C*, 8663; *HORMAD2*, 150280; *MTMR3*, 8897; *LIF*, 159540; *ZMIZ1*, 57178; *CAPN10*, 11132; *GPR35*, 2859; *KIF1A*, 547; *RNPEPL1*, 57140.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

References

1. Kugathasan S, et al. Loci on 20q13 and 21q22 are associated with pediatric-onset inflammatory bowel disease. Nat. Genet. 2008; 40:1211–1215. [PubMed: 18758464]

2. Van Limbergen J, Wilson DC, Satsangi J. The genetics of Crohn's disease. Annu. Rev. Genomics Hum. Genet. 2009; 10:89–116. [PubMed: 19453248]

3. Cho JH. The genetics and immunopathogenesis of inflammatory bowel disease. Nat. Rev. Immunol. 2008; 8:458–466. [PubMed: 18500230]

4. Fisher SA, 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]

5. Barrett JC, et al. Genome-wide association defines more than 30 distinct susceptibility loci for Crohn's disease. Nat. Genet. 2008; 40:955–962. [PubMed: 18587394]

6. Franke A, et al. Sequence variants in *IL10, ARPC2* and multiple other loci contribute to ulcerative colitis susceptibility. Nat. Genet. 2008; 40:1319–1323. [PubMed: 18836448]

7. Silverberg MS, 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]

8. Vernier-Massouille G, et al. Natural history of pediatric Crohn's disease: a population-based cohort study. Gastroenterology. 2008; 135:1106–1113. [PubMed: 18692056]

9. Van Limbergen J, et al. Definition of phenotypic characteristics of childhood-onset inflammatory bowel disease. Gastroenterology. 2008; 135:1114–1122. [PubMed: 18725221]

10. Rioux JD, et al. Genome-wide association study identifies new susceptibility loci for Crohn disease and implicates autophagy in disease pathogenesis. Nat. Genet. 2007; 39:596–604. [PubMed: 17435756]

11. 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]

12. Libioulle C, et al. Novel Crohn 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]

13. Barrett JC, et al. Genome-wide association study and meta-analysis find that over 40 loci affect risk of type 1 diabetes. Nat. Genet. 2009; 41:703–707. [PubMed: 19430480]

14. Dixon AL, et al. A genome-wide association study of global gene expression. Nat. Genet. 2007; 39:1202–1207. [PubMed: 17873877]

15. Li X, Thyssen G, Beliakoff J, Sun Z. The novel PIAS-like protein hZimp10 enhances Smad transcriptional activity. J. Biol. Chem. 2006; 281:23748–23756. [PubMed: 16777850]

16. Steinman L. A brief history of T_H17, the first major revision in the T_H1/T_H2 hypothesis of T cell-mediated tissue damage. Nat. Med. 2007; 13:139–145. [PubMed: 17290272]

17. Bettelli E, Oukka M, Kuchroo VKT. T_H-17 cells in the circle of immunity and autoimmunity. Nat. Immunol. 2007; 8:345–350. [PubMed: 17375096]
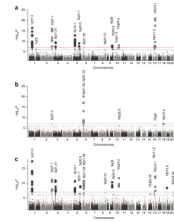
Figure 1.
Overview of genome-wide association results. (**a–c**) Chromosomal plots of –log*P* against genomic location for discovery scans in the early-onset DC-CD (**a**), DC-UC (**b**) and DC-IBD (**c**) cohorts. Red and blue dotted lines represent thresholds for genome-wide significant ($P = 1 \times 10^{-7}$) and suggestive ($P = 1 \times 10^{-6}$) signals, respectively. Loci on 16p11, 22q12 and 2q37 emerged from these analyses as suggestive signals and were validated in the follow-up cohorts RC1 and RC2-CD. The following signals did not replicate or show significance in the meta-analysis: 1q22 (CD), 10q25 (UC), 16q21 (UC), 18q12 (UC), 8q24 (IBD) and 15q22 (IBD). The remaining loci have been previously identified in majority adult-onset genome scans of Crohn's disease and ulcerative colitis. Loci on 22q12 and 20q13 were identified in a previous early-onset IBD scan involving a subset of the cohort analyzed in this study[1].
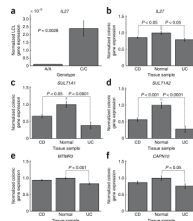
Figure 2.
Colonic and LCL mRNA expression of genes in significantly associated loci. The A allele of rs1968752 confers risk in our early-onset Crohn's disease and IBD cohorts (OR = 1.23 (1.12–1.40)). rs1968752 lies in a block of LD containing the *IL27* gene. (**a**) Cell lines from individuals (*n* = 5) with the A/A genotype at rs1968752 have over 90% decrease in *IL27* gene expression as compared with those with the C/C genotype. (**b**) *IL27* colonic expression was significantly lower in samples obtained from individuals with early-onset Crohn's disease and ulcerative colitis cases than in normal tissue. (**c,d**) Expression of both *SULT1A1* (**c**) and *SULT1A2* (**d**) genes (also in the *IL27* locus) was also significantly lower in Crohn's disease and ulcerative colitis tissue than in normal tissue. (**e,f**) Expression of putative IBD gene *MTMR3* at 22q12 (**e**) and the putative ulcerative colitis gene *CAPN10* at 2q37 (**f**) was significantly lower in ulcerative colitis tissue than in normal tissue. All significant expression effects were found to be independent of treatment and histological inflammation score (data not shown). Error bars represent s.e.m.

Table 1

Newly discovered loci significant in the GWAS of the DC-CD, RC1-CD and RC2-CD cohorts

Band (Mb)	Genes	SNP	Allele	CD discovery (DC-CD) (n = 1,636 / 6,158)				CHOP CD replication (RC1-CD) (n = 289 / 1,696)				IIBDGC replication (RC2-CD) (n = 531 / 4,109)				Combined (n = 2,456 / 11,963)	
				P	MAF_A	MAF_U	OR [95% CI]	P	MAF_A	MAF_U	OR [95% CI]	P	MAF_A	MAF_U	OR [95% CI]	P	Z
10q22.3 (80.71–80.73)	ZMIZ1	rs1250550	A/T	1.59E-05	0.28	0.32	0.83 [0.76–0.90]	0.001	0.28	0.35	0.73 [0.60–0.88]	0.005	0.20	0.24	0.82 [0.70–0.96]	4.41E-10	-6.24
16p11.2 (28.45–28.81)	IL27, SULT1A1, SULT1A2, EIF3C	rs8049439	C/G	2.72E-07	0.42	0.37	1.23 [1.13–1.32]	0.28	0.40	0.38	1.10 [0.92–1.32]	0.0014	0.42	0.39	1.14 [1.00–1.30]	2.87E-09	5.94
22q12.2 (28.75–28.86)	HORMAD2, MTMR3, LIF	rs2412973	A/T	6.16E-06	0.50	0.46	1.19 [1.10–1.29]	0.0095	0.52	0.46	1.26 [1.06–1.51]	0.0160	0.50	0.46	1.15 [1.01–1.31]	3.77E-08	5.50

MAF, minor allele frequency.

Table 2

Newly discovered loci significant in the GWAS of the DC-UC and RC1-UC cohorts

Band (Mb)	Genes	SNP	Allele	UC discovery (DC-UC) ($n = 724 / 6{,}158$)				CHOP UC replication (RC1-UC) ($n = 120 / 1{,}696$)				Combined ($n = 844 / 7{,}854$)	
				P	MAF$_A$	MAF$_U$	OR [95% CI]	P	MAF$_A$	MAF$_U$	OR [95% CI]	P	Z
2q37.3 (241.21–241.42)	*CAPN10, GPR35, KIF1A. RNPEPL1*	rs4676410	A/T	1.70×10^{-7}	0.24	0.18	1.41 [1.24–1.61]	0.0611	0.25	0.20	1.34 [0.99–1.82]	3.64×10^{-8}	5.51

Table 3

Newly discovered loci significant in the GWAS of the DC-IBD, RC1 and RC2-CD cohorts

Band (Mb)	Genes	SNP	Allele	IBD discovery (DC-IBD) (n = 2,413 / 6,158)				CHOP IBD replication (RC1-IBD) (n = 482 / 1,696)				IIBDGC replication (RC2-CD) (n = 531 / 4,109)				Combined (n = 3,426 / 11,963)	
				P	MAF$_A$	MAF$_U$	OR [95% CI]	P	MAF$_A$	MAF$_U$	OR [95% CI]	P	MAF$_A$	MAF$_U$	OR [95% CI]	P	Z
19q13.11 (38.44–38.45)		rs10500264	A/T	1.18E-05	0.21	0.18	1.21 [1.11–1.31]	0.0034	0.22	0.18	1.30 [1.07–1.58]	0.00064	0.12	0.12	1.08 [0.89–1.31]	4.26E-10	6.24
22q12.2 (28.75–28.86)	HORMAD2, MTMR3, LIF	rs2412973	A/T	9.14E-07	0.50	0.46	1.18 [1.11–1.26]	0.0052	0.51	0.46	1.23 [1.05–1.43]	0.016	0.50	0.46	1.15 [1.01–1.31]	1.55E-09	6.04
16p11.2 (28.45–28.81)	IL27, SULT1A1, SULT1A2, EIF3C	rs8049439	C/G	2.38E-07	0.41	0.37	1.20 [1.12–1.28]	0.34	0.40	0.38	1.08 [0.92–1.26]	0.0014	0.42	0.39	1.14 [1.00–1.30]	2.41E-09	5.97
10q22.3 (80.71–80.73)	ZMIZ1	rs1250550	A/T	3.55E-05	0.29	0.32	0.86 [0.80–0.92]	0.0016	0.29	0.35	0.78 [0.66–0.92]	0.00461	0.20	0.24	0.82 [0.70–0.96]	5.63E-09	−5.83

Table 4

Newly discovered effects of previously known or suggested adult IBD risk loci in early-onset IBD[a]

CDk[b]	CDp[b]	UCk[b]	UCp[b]	Band	Mb	Genes	SNP	CD (DC-CD, RC1-CD) (n = 1,925 / 7,854)		UC (DC-UC, RC1-UC) (n = 844 / 7,854)		IBD (DC-IBD, RC1) (n = 2,895 / 7,854)	
								P	Z	P	Z	P	Z
		•		1q32.1	205.01	*IL10, IL19, IL20*	rs3024505	4.84×10^{-4}	3.49	6.20×10^{-4}	3.42	**6.84×10^{-6}**	**4.50**
	•			2q12.1	102.44	*IL18R1, IL18RAP*	rs917997	**6.84×10^{-5}**	**3.98**	0.17	1.38	**5.98×10^{-5}**	**4.01**
	•			17q12	29.61	*CCL11, CCL2, CCL7*	rs991804	**2.31×10^{-4}**	**−3.68**	2.78×10^{-2}	−2.20	**6.64×10^{-5}**	**−3.99**
	•			17q12	35.29	*ORMDL3*	rs2872507	3.65×10^{-4}	3.56	**7.62×10^{-4}**	**3.37**	**3.13×10^{-6}**	**4.66**
	•			21q22.3	44.44	*ICOSLG1*	rs762421	**3.32×10^{-7}**	**5.10**	**2.54×10^{-5}**	**4.21**	**3.19×10^{-10}**	**6.29**

[a] Listed are 5 of 49 previously identified adult-onset IBD loci that were either (1) previously nominal adult IBD signals that are verified by our data or (2) validated adult signals previously shown to have an effect on early-onset UC (CD) and found in our study to have an effect on early-onset CD (UC).

[b] Filled circles in the first four columns of the table specify whether the given row represents a known CD locus (CDk), putative/nominal CD locus (CDp), known UC locus (UCk) and/or putative/nominal UC locus (UCp). Loci replicating at a Bonferroni-corrected $P < 0.05$ (corrected for 49 hypotheses) are denoted in bold; newly discovered significant effects are denoted in bold italics.