Published in final edited form as:

Science. 2006 December 1; 314(5804): 1461–1463. doi:10.1126/science.1135245.

A Genome-Wide Association Study Identifies *IL23R* as an Inflammatory Bowel Disease Gene

Richard H. Duerr^{1,2}, Kent D. Taylor^{3,4}, Steven R. Brant^{5,6}, John D. Rioux^{7,8}, Mark S. Silverberg⁹, Mark J. Daly^{8,10}, A. Hillary Steinhart⁹, Clara Abraham¹¹, Miguel Regueiro¹, Anne Griffiths¹², Themistocles Dassopoulos⁵, Alain Bitton¹³, Huiying Yang^{3,4}, Stephan Targan^{4,14}, Lisa Wu Datta⁵, Emily O. Kistner¹⁵, L. Philip Schumm¹⁵, Annette T. Lee¹⁶, Peter K. Gregersen¹⁶, M. Michael Barmada², Jerome I. Rotter^{3,4}, Dan L. Nicolae^{11,17}, and Judy H. Cho^{18,*}

¹Division of Gastroenterology, Hepatology and Nutrition, Department of Medicine, School of Medicine, University of Pittsburgh, University of Pittsburgh Medical Center Presbyterian, Mezzanine Level, C-Wing, 200 Lothrop Street, Pittsburgh, PA 15213, USA

²Department of Human Genetics, Graduate School of Public Health, University of Pittsburgh, Crabtree A300, 130 Desoto Street, Pittsburgh, PA 15261, USA

³Medical Genetics Institute, Cedars-Sinai Medical Center, 8700 Beverly Boulevard, Los Angeles, CA 90048, USA

⁴IBD Center, Division of Gastroenterology, Cedars-Sinai Medical Center, 8700 Beverly Boulevard, Los Angeles, CA 90048, USA

⁵Harvey M. and Lyn P. Meyerhoff Inflammatory Bowel Disease Center, Department of Medicine, Johns Hopkins University School of Medicine, B136, 1503 East Jefferson Street, Baltimore, MD 21231, USA

⁶Department of Epidemiology, Bloomberg School of Public Health, Johns Hopkins University, 615 North Wolfe Street, Baltimore, MD 21205, USA

⁷Université de Montréal and the Montreal Heart Institute, S-6400, 5000 Belanger Street, Montreal, Quebec H1T 1C8, Canada

⁸Medical and Population Genetics Program, Broad Institute of MIT and Harvard, 7 Cambridge Center, Cambridge, MA 02142, USA

⁹Mount Sinai Hospital IBD Centre, University of Toronto, 441–600 University Avenue, Toronto, Ontario M5G 1x5, Canada

¹⁰Massachusetts General Hospital, Harvard Medical School, 185 Cambridge Street, Boston, MA 02114, USA

*To whom correspondence should be addressed. judy.cho@yale.edu.

Supporting Online Material: www.sciencemag.org/cgi/content/full/1135245/DC1

Materials and Methods

Figs. S1

Tables S1 to S4

References and Notes

¹¹Department of Medicine, University of Chicago, 5841 South Maryland Avenue, Chicago, IL 60637, USA

¹²Department of Pediatrics, The Hospital for Sick Children, 555 University Avenue, Toronto, Ontario M5G 1x8, Canada

¹³Royal Victoria Hospital, McGill University Health Centre, 687 Pine Avenue West, Montreal, Quebec H3A 1A1, Canada

¹⁴Immunobiology Research Institute, Cedars-Sinai Medical Center, Davis 4063, 8700 Beverly Boulevard, Los Angeles, CA 90048, USA

¹⁵Department of Health Studies, University of Chicago, 5841 South Maryland Avenue, Chicago, IL 60637, USA

¹⁶The Feinstein Institute for Medical Research, 350 Community Drive, Manhasset, NY 11030, USA

¹⁷Department of Statistics, University of Chicago, 5734 South University Avenue, Chicago, IL 60637, USA

¹⁸IBD Center, Section of Digestive Diseases, Departments of Medicine and Genetics, Yale University, S155A, 300 Cedar Street, New Haven, CT 06519, USA

Abstract

The inflammatory bowel diseases Crohn's disease and ulcerative colitis are common, chronic disorders that cause abdominal pain, diarrhea, and gastrointestinal bleeding. To identify genetic factors that might contribute to these disorders, we performed a genome-wide association study. We found a highly significant association between Crohn's disease and the *IL23R* gene on chromosome 1p31, which encodes a subunit of the receptor for the proinflammatory cytokine interleukin-23. An uncommon coding variant (rs11209026, c.1142G>A, p.Arg381Gln) confers strong protection against Crohn's disease, and additional noncoding *IL23R* variants are independently associated. Replication studies confirmed *IL23R* associations in independent cohorts of patients with Crohn's disease or ulcerative colitis. These results and previous studies on the proinflammatory role of IL-23 prioritize this signaling pathway as a therapeutic target in inflammatory bowel disease.

Crohn's disease (CD) and ulcerative colitis (UC), the two common forms of idiopathic inflammatory bowel disease (IBD), are chronic, relapsing inflammatory disorders of the gastrointestinal tract. Each has a peak age of onset in the second to fourth decades of life and prevalences in European ancestry populations that average about 100 to 150 per 100,000 (1, 2). Although the precise etiology of IBD remains to be elucidated, a widely accepted hypothesis is that ubiquitous, commensal intestinal bacteria trigger an inappropriate, overactive, and ongoing mucosal immune response that mediates intestinal tissue damage in genetically susceptible individuals (1). Genetic factors play an important role in IBD pathogenesis, as evidenced by the increased rates of IBD in Ashkenazi Jews, familial aggregation of IBD, and increased concordance for IBD in monozygotic compared to dizygotic twin pairs (3). Moreover, genetic analyses have linked IBD to specific genetic

variants, especially *CARD15* variants on chromosome 16q12 and the *IBD5* haplotype (spanning the organic cation transporters, *SLC22A4* and *SLC22A5*, and other genes) on chromosome 5q31 (3–7). CD and UC are thought to be related disorders that share some genetic susceptibility loci but differ at others.

The replicated associations between CD and variants in *CARD15* and the *IBD5* haplotype do not fully explain the genetic risk for CD, so we performed a genome-wide association study testing 308,332 autosomal single nucleotide polymorphisms (SNPs) on the Illumina HumanHap300 Genotyping BeadChip (8). Our study population consisted of 567 non-Jewish, European ancestry patients with ileal CD and 571 non-Jewish controls. We initially focused on ileal CD, the most common location of CD, to minimize pathogenic heterogeneity. After exclusion of study subjects with genotype completion rates less than 94%, we included 547 cases and 548 controls in subsequent analyses (8). Single-marker allelic tests were performed using χ^2 statistics for all autosomal markers. Three SNPs had nearly two orders of magnitude greater significance compared to the next most significant markers, and they are the only markers that remain significant at the 0.05 level after Bonferroni correction. Two of the three markers, rs2066843 ($P = 2.86 \times 10^{-9}$, corrected $P = 8.82 \times 10^{-4}$) and rs2076756 ($P = 5.12 \times 10^{-10}$, corrected $P = 1.58 \times 10^{-4}$), are in the known CD susceptibility gene, *CARD15* (4, 5). The third marker, rs11209026 ($P = 5.05 \times 10^{-9}$, corrected $P = 1.56 \times 10^{-3}$), is a nonsynonymous SNP (c.1142G>A, p.Arg381Gln) in the *IL23R* gene (GenBank accession: NM_144701, GeneID: 149233) on chromosome 1p31. This gene encodes a subunit of the receptor for the proinflammatory cytokine, interleukin-23 (IL-23), and is therefore an intriguing functional candidate. In addition to Arg381Gln, nine other markers in *IL23R* and in the intergenic region between *IL23R* and the adjacent IL-12 receptor, beta-2 gene (*IL12RB2*), had association P -values < 0.0001 in the non-Jewish, ileal CD case-control cohort (Table 1 and table S1a).

We next tested for association of *IL23R* markers in an independent ileal CD case-control cohort, consisting of 401 patients and 433 controls, all of Jewish ancestry (8). Significant associations were observed for several of the same markers that were associated in the non-Jewish cohort (Table 1 and table S1b). In a combined analysis of the data from the two ileal CD case-control cohorts (8), nine markers had highly significant association P -values ranging from 1.60×10^{-9} to 3.36×10^{-13} (Table 1 and table S1b).

We then extended the replication study by performing family-based association testing of 27 *IL23R* region markers in an independent cohort of 883 nuclear families in which both parents and their IBD (CD, UC, or indeterminate IBD)-affected offspring were available for genotyping (Table 2 and table S2) (8). For Arg381Gln and other *IL23R* markers, we observed significant departure from random allele transmission to CD-affected offspring in both non-Jewish and Jewish families, providing further evidence for association between CD and *IL23R*. We also observed distortion of allele transmission to non-Jewish, UC-affected offspring, providing evidence for association of *IL23R* with non-Jewish UC. There was no evidence for association of Arg381Gln or other *IL23R* region markers in the Jewish UC families. In a combined analysis of the data from all 883 nuclear families and both case-control cohorts (8), all 10 *IL23R* markers in Table 2 showed highly significant association with IBD, with P -values ranging from 3.55×10^{-9} to 6.62×10^{-19} .

The *IL23R* gene is contained within two large blocks of linkage disequilibrium, and markers in the centromeric block containing exons 5 to 11 and part of the intergenic region between *IL23R* and *IL12RB2* have the strongest association signals (Fig. 1). There is no significant association within *IL12RB2* (Fig. 1), and we did not identify a *IL12RB2* SNP in the International HapMap CEU data that is correlated with an IBD-associated, *IL23R* region variant (8).

The IL23R protein contains an extracellular domain (composed of a signal sequence, an N-terminal immunoglobulin-like domain, and two cytokine receptor domains), a single transmembrane domain, and a 252-amino acid cytoplasmic domain (9). Arg-381, in the cytoplasmic domain, is the fifth amino acid internal to the transmembrane domain and is highly conserved between species (fig. S1). In contrast, two other non-synonymous *IL23R* SNPs, rs1884444 (His3Gln) and rs7530511 (Pro310Leu), which are located within the extracellular domain, show no evidence for disease association (table S1, a and b).

The glutamine allele of Arg381Gln is much less common than the arginine allele, with an allelic frequency of 1.9% in the non-Jewish patients with ileal CD and 7.0% in non-Jewish controls. The glutamine allele appears to protect against development of CD in both non-Jewish [odds ratio (OR) = 0.26, 95% confidence interval (CI) (0.15 to 0.43)] and Jewish [OR = 0.45, 95% CI (0.27 to 0.73)] case-control cohorts. The glutamine allele is also significantly undertransmitted from heterozygous parents to non-Jewish and Jewish CD-affected offspring, non-Jewish UC-affected offspring, and all IBD-affected offspring (transmitted:non-transmitted = 45:130, $P = 1.32 \times 10^{-10}$ for the IBD phenotype in all 883 families) (Table 2 and table S2). Our discovery of an uncommon protective allele, or conversely, a very common predisposing allele, reflects a major theme in complex genetics; namely, that functional genetic variation exerts a continuum of susceptibility, neutral, and protective effects. Furthermore, alleles conferring protection against one disease may result in increased risk for another (10).

In addition to Arg381Gln, we found several other variants within the *IL23R* gene that are also associated with IBD (Tables 1 and 2 and tables S1 and S2). Marker rs11465804 is an intronic variant in a nonconserved region and is in significant linkage disequilibrium with Arg381Gln (correlation coefficient $r^2 = 0.84$ in the case-control data) and therefore is unlikely to confer disease risk independent of the latter. However, other markers show evidence for association that appears to be independent of Arg381Gln. For example, rs7517847, which has the most significant association P -value (3.36×10^{-13}) in the combined analysis of both ileal CD case-control cohorts (Table 1), is not in significant linkage disequilibrium with Arg381Gln ($r^2 = 0.03$ in the case-control data). To identify variants that are independent of the Arg381Gln signal, we performed conditional association testing of the combined case-control data by stratifying on the Arg381Gln genotypes. The P -values for these conditional tests (table S3) demonstrate multiple residual association signals throughout *IL23R*, indicating that there are multiple risk variants in the region. The *IL23R* gene is expressed as at least six alternatively spliced mRNAs, which generate diverse isoforms of the receptor protein (11). The most common splice variants result in the deletion of exons 7 and/or 10. We therefore speculate that the multiple genetic association signals

detected in the centromeric portion of *IL23R* (Fig. 1) could exert their influence via differential splicing.

Notably, we found no evidence for association in our non-Jewish, ileal CD case-control cohort (table S4) with the *IL12RB1* gene, which encodes the second subunit of the IL-23 receptor (9), or the *IL23A* and *IL12B* genes, which encode the p19 and p40 subunits, respectively, of the heterodimeric IL-23 cytokine (12).

Previous work with mouse models has documented a requirement for IL-23 in murine colitis (13), experimental autoimmune encephalitis (14), and collagen-induced arthritis (15). IL-23 activity is present in the terminal ileum (16) and colon (17), and the present study demonstrates that *IL23R* variants are associated with both small intestinal (ileal CD) and large intestinal (UC) inflammation. Furthermore, transgenic expression of IL-23 subunit p19 results in severe systemic inflammation, including in the small and large intestine (18), highlighting this pathway's particular role in promoting strong activation of effector T cells and perpetuation of organ-specific inflammatory responses. At least part of this effect is likely mediated via inflammatory, IL-17–producing T cells (19–23), and elevated IL-17 levels have been observed in the colonic mucosa of both CD and UC patients (24).

Taken together, these findings suggest that blockade of the IL-23 signaling pathway would be a rational therapeutic strategy for IBD. In support of this, a monoclonal antibody directed against the p40 subunit of the receptor, which blocks both IL-23 and IL-12 proinflammatory activities, has produced promising results in a clinical trial of Crohn's disease (25). It has been postulated that specific targeting of the IL23p19/IL23R pathway may be particularly effective in blocking organ-specific inflammation, with less compromise of protective responses (26). However, at least one model of murine colitis is worsened in the absence of IL-23, implicating a role for IL-23 in the down-regulation of IL-12 (27). In addition, IL-23 function may be important for proper responses to mycobacterial (28, 29) and intestinal infections (22). In assessing therapeutic approaches, the strong protective effect of the Arg381Gln allele could potentially be exploited to define desired functional outcomes (10). The contribution of the *IL23R* pathway to IBD will likely involve more than simple gain- or loss-of-function *IL23R* variants, and therapeutic interventions will be improved by a better understanding of the context and tissue-specific events associated with functional *IL23R* polymorphisms.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank the patients and their families for participating in the studies. The National Institute of Diabetes and Digestive and Kidney Diseases IBD Genetics Consortium is funded by the following grants: DK62431 (S.R.B.), DK62422 (J.H.C.), DK62420 (R.H.D.), DK62432 (J.D.R.), DK62423 (M.S.S.), DK62413 (K.D.T.), and DK62429 (J.H.C.). A.H.S. is on the Scientific Advisory Boards of Shire Pharmaceuticals, Schering (Canada), and Procter & Gamble Pharmaceuticals.

References and Notes

1. Podolsky DK. *N Engl J Med.* 2002; 347:417. [PubMed: 12167685]
2. Loftus EV Jr. *Gastroenterology.* 2004; 126:1504. [PubMed: 15168363]
3. Vermeire S, Rutgeerts P. *Genes Immun.* 2005; 6:637. [PubMed: 16107869]
4. Hugot JP, et al. *Nature.* 2001; 411:599. [PubMed: 11385576]
5. Ogura Y, et al. *Nature.* 2001; 411:603. [PubMed: 11385577]
6. Rioux JD, et al. *Nat Genet.* 2001; 29:223. [PubMed: 11586304]
7. Peltekova VD, et al. *Nat Genet.* 2004; 36:471. [PubMed: 15107849]
8. Materials and methods are available as supporting material on Science Online.
9. Parham C, et al. *J Immunol.* 2002; 168:5699. [PubMed: 12023369]
10. Nadeau JH, Topol EJ. *Nat Genet.* 2006; 38:1095. [PubMed: 17006459]
11. Zhang XY, et al. *Immunogenetics.* 2006; 57:934. [PubMed: 16372191]
12. Oppmann B, et al. *Immunity.* 2000; 13:715. [PubMed: 11114383]
13. Yen D, et al. *J Clin Invest.* 2006; 116:1310. [PubMed: 16670770]
14. Cua DJ, et al. *Nature.* 2003; 421:744. [PubMed: 12610626]
15. Murphy CA, et al. *J Exp Med.* 2003; 198:1951. [PubMed: 14662908]
16. Becker C, et al. *J Clin Invest.* 2003; 112:693. [PubMed: 12952918]
17. Fuss IJ, et al. *Inflamm Bowel Dis.* 2006; 12:9. [PubMed: 16374252]
18. Wiekowski MT, et al. *J Immunol.* 2001; 166:7563. [PubMed: 11390512]
19. Aggarwal S, Ghilardi N, Xie MH, de Sauvage FJ, Gurney AL. *J Biol Chem.* 2003; 278:1910. [PubMed: 12417590]
20. Langrish CL, et al. *J Exp Med.* 2005; 201:233. [PubMed: 15657292]
21. Bettelli E, et al. *Nature.* 2006; 441:235. [PubMed: 16648838]
22. Mangan PR, et al. *Nature.* 2006; 441:231. [PubMed: 16648837]
23. Veldhoen M, Hocking RJ, Atkins CJ, Locksley RM, Stockinger B. *Immunity.* 2006; 24:179. [PubMed: 16473830]
24. Fujino S, et al. *Gut.* 2003; 52:65. [PubMed: 12477762]
25. Mannon PJ, et al. *N Engl J Med.* 2004; 351:2069. [PubMed: 15537905]
26. McKenzie BS, Kastelein RA, Cua DJ. *Trends Immunol.* 2006; 27:17. [PubMed: 16290228]
27. Becker C, et al. *J Immunol.* 2006; 177:2760. [PubMed: 16920909]
28. Verreck FA, et al. *Proc Natl Acad Sci U S A.* 2004; 101:4560. [PubMed: 15070757]
29. Cooper AM, et al. *J Immunol.* 2002; 168:1322. [PubMed: 11801672]

Fig. 1.

Association signals in the *IL23R* gene region on chromosome 1p31. (A) Genomic locations of genes on chromosome 1p31 between 67,260,000 and 67,580,000 base pairs (Build 35). (B) The negative log₁₀ association *P*-values (Cochran-Mantel-Haenszel chi-square test) from the combined Jewish and non-Jewish case-control cohorts are plotted for genotyped markers in the region. (C) Pairwise r^2 plot for International HapMap CEU data. The intensity of the shading is proportional to r^2 . The *IL23R* gene is contained within two blocks of linkage disequilibrium, and the association signals are strongest in the centromeric block, which contains exons 5 to 11 and extends into the intergenic region between *IL23R* and *IL12RB2*. Note that markers in the block encompassing the *IL12RB2* gene do not demonstrate significant association.

Table 1

Non-Jewish and Jewish ileal Crohn's disease (CD) case-control association study results for IL23R region markers with P -values < 0.0001 in the non-Jewish cohort. Minor allele frequencies (MAF), allelic test P -values, and odds ratios (OR) with 95% confidence intervals (CI) are shown for each case-control cohort (8). The ORs shown are for the minor allele. Combined Cochran-Mantel-Haenszel P -values are also shown (8). UTR, untranslated region.

Marker	Location	Non-Jewish case-control cohort				Jewish case-control cohort				Combined P -value
		CD ($n = 547$) MAF	Control ($n = 548$) MAF	P -value	OR [95% CI]	CD ($n = 401$) MAF	Control ($n = 433$) MAF	P -value	OR [95% CI]	
rs1004819	Intron	0.374	0.280	3.79×10^{-6}	1.53 [1.27,1.84]	0.426	0.334	1.00×10^{-4}	1.48 [1.21,1.82]	1.54×10^{-9}
rs7517847	Intron	0.331	0.443	1.09×10^{-7}	0.62 [0.52,0.74]	0.240	0.352	5.84×10^{-7}	0.58 [0.47,0.72]	3.36×10^{-13}
rs10489629	Intron	0.378	0.475	4.27×10^{-6}	0.67 [0.56,0.80]	0.355	0.465	5.79×10^{-6}	0.63 [0.52,0.77]	1.14×10^{-10}
rs2201841	Intron	0.385	0.291	4.57×10^{-6}	1.52 [1.27,1.83]	0.414	0.315	2.92×10^{-5}	1.53 [1.25,1.89]	5.46×10^{-10}
rs11465804	Intron	0.020	0.063	7.52×10^{-7}	0.30 [0.18,0.51]	0.048	0.096	1.39×10^{-4}	0.47 [0.31,0.71]	5.97×10^{-10}
rs11209026	Arg381Gln	0.019	0.070	5.05×10^{-9}	0.26 [0.15,0.43]	0.033	0.070	7.95×10^{-4}	0.45 [0.27,0.73]	3.55×10^{-11}
rs1343151	Intron	0.275	0.370	2.26×10^{-6}	0.65 [0.54,0.78]	0.229	0.336	1.69×10^{-6}	0.59 [0.47,0.73]	1.64×10^{-11}
rs10889677	Exon-3'UTR	0.385	0.288	1.82×10^{-6}	1.55 [1.29,1.86]	0.419	0.316	1.51×10^{-5}	1.56 [1.27,1.91]	9.58×10^{-11}
rs11209032	Intergenic	0.393	0.293	1.03×10^{-6}	1.56 [1.30,1.87]	0.382	0.298	3.49×10^{-4}	1.45 [1.18,1.79]	1.60×10^{-9}
rs1495965	Intergenic	0.498	0.412	2.93×10^{-5}	1.44 [1.21,1.71]	0.469	0.412	2.04×10^{-2}	1.26 [1.03,1.53]	2.55×10^{-6}

Table 2

Family-based and combined (case-control and family-based) association results. Family-based association *P*-values were computed using the empirical variance estimator implemented in the FBAT software package (8). Combined Fisher *P*-values for all case-control (Table 1) and nuclear family cohorts are also shown (8). UTR, untranslated region.

Marker	Location	Non-Jewish CD (518 families, 651 affected offspring)	Non-Jewish UC (215 families, 251 affected offspring)	Jewish CD (77 families, 99 affected offspring)	Jewish UC (80 families, 91 affected offspring)	All IBD (883 families, 1,119 affected offspring)	Combined (family-based and case-control) <i>P</i> -value
		<i>P</i> -value	<i>P</i> -value	<i>P</i> -value	<i>P</i> -value	<i>P</i> -value	
rs1004819	Intron	3.60×10^{-5}	1.20×10^{-3}	1.24×10^{-2}	5.47×10^{-1}	6.06×10^{-8}	1.78×10^{-14}
rs7517847	Intron	2.30×10^{-5}	2.71×10^{-1}	3.50×10^{-2}	5.00×10^{-1}	1.80×10^{-5}	9.99×10^{-16}
rs10489629	Intron	1.87×10^{-3}	2.70×10^{-1}	4.33×10^{-1}	8.21×10^{-1}	1.27×10^{-3}	1.62×10^{-11}
rs2201841	Intron	5.80×10^{-4}	3.21×10^{-4}	3.50×10^{-2}	5.69×10^{-1}	1.04×10^{-7}	1.10×10^{-14}
rs11465804	Intron	1.32×10^{-4}	2.70×10^{-3}	8.90×10^{-5}	3.71×10^{-1}	3.46×10^{-9}	3.33×10^{-16}
rs11209026	Arg381Gln	8.00×10^{-6}	2.97×10^{-4}	9.41×10^{-4}	4.91×10^{-1}	1.32×10^{-10}	6.62×10^{-19}
rs1343151	Intron	9.63×10^{-2}	8.51×10^{-2}	3.30×10^{-2}	1.89×10^{-1}	1.24×10^{-3}	2.74×10^{-12}
rs10889677	Exon-3'UTR	2.60×10^{-3}	3.35×10^{-4}	5.88×10^{-2}	7.32×10^{-1}	1.65×10^{-6}	3.40×10^{-14}
rs11209032	Intergenic	2.68×10^{-3}	3.57×10^{-4}	3.48×10^{-2}	7.50×10^{-1}	2.41×10^{-6}	5.50×10^{-13}
rs1495965	Intergenic	4.07×10^{-4}	1.74×10^{-2}	3.93×10^{-2}	9.21×10^{-1}	1.72×10^{-5}	3.55×10^{-9}