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

Richard H. Duerr1,2, Kent D. Taylor3,4, Steven R. Brant5,6, John D. Rioux7,8, Mark S. Silverberg9, Mark J. Daly8,10, A. Hillary Steinhart9, Clara Abraham11, Miguel Regueiro1, Anne Griffiths12, Themistocles Dassopoulos5, Alain Bitton13, Huiying Yang3,4, Stephan Targan4,14, Lisa Wu Datta5, Emily O. Kistner15, L. Philip Schumm15, Annette T. Lee16, Peter K. Gregersen16, M. Michael Barmada2, Jerome I. Rotter3,4, Dan L. Nicolae11,17, and Judy H. Cho18, *

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 1×5, 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
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.
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 $\chi^2$ 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 $1.60 \times 10^{-9}$ to $3.36 \times 10^{-13}$ (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 $3.55 \times 10^{-9}$ to $6.62 \times 10^{-19}$. 

*Science. Author manuscript; available in PMC 2015 April 27.*
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 × 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

8. Materials and methods are available as supporting material on Science Online.
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.

<table>
<thead>
<tr>
<th>Marker</th>
<th>Location</th>
<th>Non-Jewish case-control cohort</th>
<th>Jewish case-control cohort</th>
<th>Combined ( P )-value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>CD (n = 547) MAF</td>
<td>Control (n = 548) MAF</td>
<td>( P )-value</td>
</tr>
<tr>
<td>rs1004819</td>
<td>Intron</td>
<td>0.374</td>
<td>0.280</td>
<td>3.79 \times 10^{-6}</td>
</tr>
<tr>
<td>rs7517847</td>
<td>Intron</td>
<td>0.331</td>
<td>0.443</td>
<td>1.09 \times 10^{-7}</td>
</tr>
<tr>
<td>rs10489629</td>
<td>Intron</td>
<td>0.378</td>
<td>0.475</td>
<td>4.27 \times 10^{-6}</td>
</tr>
<tr>
<td>rs2201841</td>
<td>Intron</td>
<td>0.385</td>
<td>0.291</td>
<td>4.57 \times 10^{-6}</td>
</tr>
<tr>
<td>rs11465804</td>
<td>Intron</td>
<td>0.020</td>
<td>0.063</td>
<td>7.52 \times 10^{-7}</td>
</tr>
<tr>
<td>rs11209026</td>
<td>Arg381Gln</td>
<td>0.019</td>
<td>0.070</td>
<td>5.05 \times 10^{-9}</td>
</tr>
<tr>
<td>rs1343151</td>
<td>Intron</td>
<td>0.275</td>
<td>0.370</td>
<td>2.26 \times 10^{-6}</td>
</tr>
<tr>
<td>rs10889677</td>
<td>Exon-3'UTR</td>
<td>0.385</td>
<td>0.288</td>
<td>1.82 \times 10^{-6}</td>
</tr>
<tr>
<td>rs1209032</td>
<td>Intergenic</td>
<td>0.393</td>
<td>0.293</td>
<td>1.03 \times 10^{-6}</td>
</tr>
<tr>
<td>rs1495965</td>
<td>Intergenic</td>
<td>0.498</td>
<td>0.412</td>
<td>2.93 \times 10^{-5}</td>
</tr>
</tbody>
</table>
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.

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