nature genetics

Sequence variants in *IL10*, *ARPC2* and multiple other loci contribute to ulcerative colitis susceptibility

Andre Franke[1,17], Tobias Balschun[1,17], Tom H Karlsen[2,17], Jurgita Sventoraityte[1], Susanna Nikolaus[3,4], Gabriele Mayr[5], Francisco S Domingues[5], Mario Albrecht[5], Michael Nothnagel[6], David Ellinghaus[1], Christian Sina[3,4], Clive M Onnie[7], Rinse K Weersma[8], Pieter C F Stokkers[9], Cisca Wijmenga[10], Maria Gazouli[11], David Strachan[12], Wendy L McArdle[13], Séverine Vermeire[14], Paul Rutgeerts[14], Philip Rosenstiel[1], Michael Krawczak[6], Morten H Vatn[2,15], the IBSEN study group[16], Christopher G Mathew[7] & Stefan Schreiber[1,4]

Inflammatory bowel disease (IBD) typically manifests as either ulcerative colitis (UC) or Crohn's disease (CD). Systematic identification of susceptibility genes for IBD has thus far focused mainly on CD, and little is known about the genetic architecture of UC. Here we report a genome-wide association study with 440,794 SNPs genotyped in 1,167 individuals with UC and 777 healthy controls. Twenty of the most significantly associated SNPs were tested for replication in three independent European case-control panels comprising a total of 1,855 individuals with UC and 3,091 controls. Among the four consistently replicated markers, SNP rs3024505 immediately flanking the *IL10* (interleukin 10) gene on chromosome 1q32.1 showed the most significant association in the combined verification samples ($P = 1.35 \times 10^{-12}$; OR = 1.46 (1.31–1.62)). The other markers were located in *ARPC2* and in the *HLA-BTNL2* region. Association between rs3024505 and CD (1,848 cases, 1,804 controls) was weak ($P = 0.013$; OR = 1.17 (1.01–1.34)). IL10 is an immunosuppressive cytokine that has long been proposed to influence IBD pathophysiology. Our findings strongly suggest that defective IL10 function is central to the pathogenesis of the UC subtype of IBD.

The two main subtypes of inflammatory bowel disease (IBD), ulcerative colitis (UC, MIM191390) and Crohn's disease (CD, MIM266600), are chronic relapsing-remitting inflammatory disorders affecting the intestinal mucosa. Both diseases represent major burdens of morbidity in Western countries, with prevalence rates in North America and Europe ranging from 21 to 246 per 100,000 for UC and 8 to 214 per 100,000 for CD[1]. Although some clinical and pathological features are shared by these two subphenotypes of IBD, there are important differences in disease localization, endoscopic appearance, histology and behavior, which suggest differences in the underlying pathophysiology. In both diseases, inappropriate control of chronic inflammation has a major role[2].

The genetic contribution to disease risk has been documented more clearly for CD than for UC (relative sibling risks: 15–35 for CD, 6–9 for UC), and the recent identification of several CD susceptibility genes has yielded valuable insights into the pathogenesis of this IBD subtype[3]. In the clinical picture, some overlap is seen and the systematic analysis of CD risk markers shows that several of them are also associated with UC, including *IL23R, IL12B, NKX2-3, CCNY* and the 3p21.31 (*MST1*) locus[4,5]. A genome-wide candidate gene experiment investigating 10,886 nonsynonymous SNPs in 1,470 British controls and 936 UC cases yielded *ECM1* on 1q21.2 as a new UC-specific susceptibility gene[6]. However, a systematic, genome-wide analysis of UC has not been reported so far.

We set out to identify UC susceptibility loci systematically in a large sample of 1,167 cases and 777 healthy controls (panel A, **Table 1**) by testing 440,794 SNPs with the Affymetrix SNP Array 5.0 (**Supplementary Methods** online). Screening panel A had 80% power to detect a variant with an odds ratio of 1.5 or higher at the 5%

[1]Institute for Clinical Molecular Biology, Christian-Albrechts-University, Kiel 24105, Germany. [2]Medical Department, Rikshospitalet University Hospital, Oslo 0027, Norway. [3]popgen Biobank, Christian-Albrechts-University Kiel, Kiel 24105, Germany. [4]Department of General Internal Medicine, University Hospital Schleswig-Holstein, Christian-Albrechts-University, Kiel 24105, Germany. [5]Max-Planck Institute for Informatics, Saarbrücken 66123, Germany. [6]Institute of Medical Informatics and Statistics, Christian-Albrechts-University, Kiel 24105, Germany. [7]Department of Medical and Molecular Genetics, King's College London School of Medicine, London SE1 9RT, UK. [8]Department of Gastroenterology and Hepatology, University Medical Center Groningen and University of Groningen, Groningen 9700 RB, The Netherlands. [9]Department of Gastroenterology and Hepatology, Academic Medical Center, Amsterdam 1105 AZ, The Netherlands. [10]Department of Genetics, University Medical Center Groningen and University of Groningen, Groningen 9700 RB, The Netherlands. [11]Department of Biology, School of Medicine, University of Athens, Athens 11527, Greece. [12]St. Georges's University, Division of Community Health Sciences, London SW17 ORE, UK. [13]ALSPAC, Department of Social Medicine, University of Bristol, Bristol BS8 1TQ, UK. [14]Department of Gastroenterology, University Hospital Gasthuisberg, Leuven 3000, Belgium. [15]Faculty of Medicine, Epigen, Akershus University Hospital, 1474 Oslo, Norway. [16]A full list of members is provided in the **Supplementary Note** online. [17]These authors contributed equally to this work. Correspondence should be addressed to S.S. (s.schreiber@mucosa.de).

Received 17 April; accepted 25 July; published online 5 October 2008; doi:10.1038/ng.221

Table 1 Summary of association results of replication

			Panel A – GWAS 777 controls/1,167 UC cases			Panel B – Germany 985 controls, 523 UC cases			Panel C – UK 1,091 controls, 304 UC cases			Panel D – Belgium/Netherlands 1,015 controls, 1,028 UC cases			Combined analysis (panels B–D) 3,091 controls, 1,855 UC cases	
Position (bp)	dbSNP ID	A1 A2	AF_{A1} Con. Case	P_{CCA}	OR (95% CI)	AF_{A1} Con. Case	P_{CCA}	OR (95% CI)	AF_{A1} Con. Case	P_{CCA}	OR (95% CI)	AF_{A1} Con. Case	P_{CCA}	OR (95% CI)	P_{BD}	P_{CMH} OR (95% CI)
6p21.32	rs9268877	T	0.44	***5.23E–7***	1.51	0.44	***5.68E–8***	1.91	0.40	***9.94E–6***	1.87	0.45	***3.10E–7***	1.14	0.61	***6.48E–18***[a]
HLA-DRA		C	0.52		(1.23–1.86)	0.54		(1.48–2.46)	0.50		(1.39–2.49)	0.53		(0.94–1.38)		1.45 (1.33–1.58)
(32,539,125)																
6p21.32	rs9268858	G	0.30	***5.41E–7***	0.65	0.29	***1.27E–5***	0.67	0.37	***0.0067***	0.86	0.27	***1.20E–6***	0.63	0.61	***2.58E–12***[a]
HLA-DRA		A	0.22		(0.54–0.78)	0.22		(0.54–0.83)	0.31		(0.67–1.12)	0.20		(0.53–0.76)		0.71 (0.64–0.78)
(32,537,736)																
6p21.33	rs9268480	A	0.27	***2.21E–6***	0.62	0.27	***4.00E–4***	0.67	0.31	***0.0076***	0.72	0.25	***1.21E–4***	0.67	0.95	***3.15E–9***[a]
BTNL2		G	0.20		(0.52–0.75)	0.21		(0.54–0.84)	0.25		(0.56–0.93)	0.20		(0.56–0.80)		0.74 (0.67–0.82)
(32,471,822)																
1p31.2	rs11805303	T	0.26	***5.39E–6***	1.50	0.27	***0.0057***	1.39	0.31	***0.019***	1.20	0.28	***0.011***	1.26	0.86	***1.09E–5***[a]
IL23R		C	0.33		(1.25–1.80)	0.32		(1.12–1.72)	0.36		(0.92–1.55)	0.32		(1.05–1.50)		1.23 (1.12–1.35)
(67,387,537)																
2q35	rs12612347	T	0.45	***8.42E–6***	1.60	0.49	0.14	1.22	0.48	***0.0098***	1.37	0.46	***0.013***	1.23	0.59	***2.00E–4***[a]
ARPC2		C	0.52		(1.30–1.96)	0.52		(0.94–1.57)	0.54		(1.01–1.86)	0.50		(1.00–1.50)		1.18 (1.08–1.28)
(218,882,844)																
1q32.1	rs3024505	T	0.14	***1.43E–5***	1.52	0.16	***0.0012***	1.48	0.14	***3.51E–5***	1.64	0.15	***8.04E–6***	1.50	0.75[b]	***1.35E–12***[a,b]
IL10		C	0.20		(1.24–1.85)	0.21		(1.18–1.86)	0.21		(1.25–2.15)	0.21		(1.24–1.81)		1.46 (1.31–1.62)[b]
(203,328,299)																
3p12.1 *CADM2*	rs7611991	A	0.29	***6.15E–5***	0.71	0.25	***0.034***	0.75	0.26	0.76	0.98	0.27	0.14	0.91	0.49	***0.023***
(85,842,248)		G	0.23		(0.60–0.86)	0.22		(0.60–0.93)	0.25		(0.76–1.27)	0.25		(0.76–1.09)		0.89 (0.81–0.98)
9p24.1	rs10974944	G	0.25	***7.95E–5***	1.42	0.30	0.69	1.01	0.27	0.58	0.89	0.26	***6.98E–4***	1.30	***0.036***	***0.023***
JAK2		C	0.31		(1.18–1.70)	0.31		(0.82–1.26)	0.26		(0.69–1.15)	0.31		(1.09–1.55)		1.12 (1.02–1.23)
(5,060,831)																
5q13.3	rs7712957	G	0.07	***8.98E–5***	1.61	0.08	0.63	1.04	0.10	0.45	1.15	0.09	***0.042***	1.21	0.64	***0.039***
S100Z		A	0.11		(1.26–2.06)	0.09		(0.78–1.40)	0.11		(0.84–1.57)	0.11		(0.96–1.52)		1.16 (1.01–1.34)
(76,174,452)																

Twenty SNPs were genotyped in three independent replication panels of different ancestry. Data are shown for the nine SNPs that were significant in the combined analysis (nominal P_{CMH} <0.05). Results of all 20 SNPs are shown in **Supplementary Table 7** online. SNPs are ranked according to the *P* value obtained in the GWAS. Nucleotide positions refer to NCBI build 35. A1 denotes the rare allele and A2 is the common allele in controls. The respective allele frequencies are shown for allele A1 (AF_{A1}). *P* values obtained in the case-control analysis using an allele-based χ^2 test (one degree of freedom) are listed (P_{CCA}). Significant *P* values ($P < 0.05$) are highlighted in bold italic. Odds ratios and 95% confidence intervals for carriership of the allele A1 are shown. Column P_{BD} lists the asymptotic *P* values of a Breslow-Day test for heterogeneity. A significant *P* value indicates a significant heterogeneity between replication panels in terms of the odds ratio of the disease association. Combined *P* values (P_{CMH}) and combined ORs of the Cochran-Mantel-Haenszel test statistic (one degree of freedom) are shown. Results for *IL10* lead SNP rs3024505 are highlighted by gray shading. Further details including genotype counts are listed in **Supplementary Table 2**.

[a]*P* value <0.05 after correction for multiple testing (corresponding to a Bonferroni threshold of 0.0025 for the uncorrected *P* value). [b]For the combined analysis of rs3024505, 204 Greek cases and 431 matched controls were included (panel E).

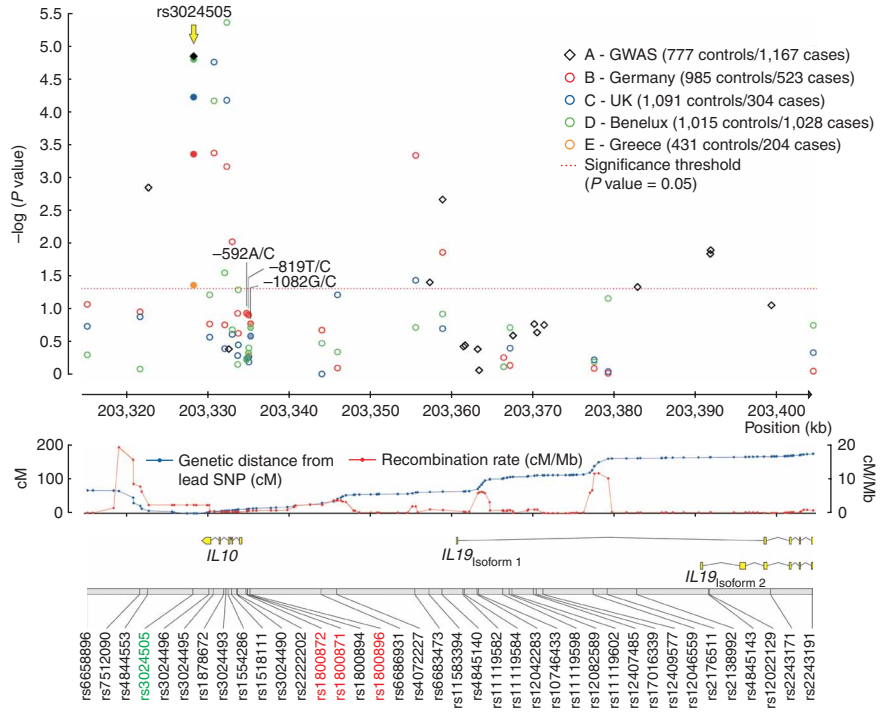

Figure 1 Regional plot of the confirmed UC association at *IL10*. Plot of the negative decadic logarithm of the *P* values obtained in the GWAS (panel A) and the fine mapping in replication panels B to E. Twenty-three tagging SNPs, including lead SNP rs3024505 (highlighted by filled symbols), were genotyped across the 89-kb region surrounding the *IL10* gene. The three *IL10* promoter SNPs rs1800872 (–C592A), rs1800871 (–C819T) and rs1800896 (–G1082A) are highlighted. Nominal *P* values for each UC case-control panel are shown and are based on a Pearson χ^2 test with one degree of freedom. The red dotted line corresponds to a threshold of 0.05 for the *P* value. The middle panel includes plots of the recombination intensity (cM/Mb) and the cumulative genetic distance in cM, and the bottom panel shows the position and intron-exon structure of *IL10* and part of the upstream *IL19* gene. Positions are given as NCBI build 35 coordinates. For details, see **Supplementary Table 2**.

significance level, assuming a frequency of the disease-associated allele of at least 20% in controls (**Supplementary Fig. 1** online). Genotyping was done blind to case-control status, and several HapMap samples with known genotype were included in each batch for quality control. After applying stringent quality control criteria (**Supplementary Methods**) to the genotype data from panel A, we included all 1,944 samples and a subset of 355,262 SNPs in the final association analysis (see **Supplementary Fig. 2** online). The total genotyping rate across these samples was 99.8%. We found genetic heterogeneity to be low, with an estimated genomic inflation factor[7] of $\lambda_{GC} = 1.11$ (**Supplementary Figs. 3** and **4** online). *P* values with and without correction for structure were similar and consequently, unadjusted *P* values are shown below. We genotyped the 20 most strongly associated SNPs that passed selection criteria for replication (**Supplementary Methods** and **Supplementary Fig. 5** online) in three additional panels of UC cases and healthy controls (panels B, C and D in **Table 1**, see also **Supplementary Methods**). The results of the association analysis are summarized in **Table 1**, and full details, including genotype counts, are given in **Supplementary Table 1** online.

New associations that withstood correction for multiple testing using Bonferroni correction ($\alpha = 0.0025$ (0.05/20)) across all three replication panels were obtained for rs3024505 near the 3′ UTR of the *IL10* gene at 1q32, for rs12612347 near the *ARP2C* locus at 2q35 and for rs9268480, rs9268858 and rs9268877 at the class II–class III junction in the HLA complex at 6p21 (**Table 1**). The findings at the latter three SNPs, located near the HLA class II genes on chromosome 6p21, are not surprising given the large body of evidence for an association between classical HLA loci and UC[8]. Because of the complex pattern of linkage disequilibrium (LD) in this region, comprehensive experiments beyond the scope of the present study will be required to clarify whether the observed associations are due to variation in the HLA class II genes themselves, at neighboring loci (for example, *BTNL2*), or both. Another notable finding in the present study is the consistent association between *ARPC2* and UC. The exact function of *ARPC2* is not known. The microbial equivalent of human Arp2/3 associates with the protein encoded by the *WAS* gene, which, when mutated in humans, causes Wiskott-Aldrich syndrome (WAS). Notably, the WAS protein is involved in the regulation of regulatory T cells[9], and manifestation of UC has been reported as the first sign of disease in an individual with WAS[10]. Genome-wide linkage studies have also provided evidence for UC susceptibility factors in the respective region at 2q[11]. That the previously reported *IL23R* association[4] replicated in all sample panels of the present study (SNP rs11805303: combined $P = 1.09 \times 10^{-5}$ and OR = 1.23 (1.12–1.35)), albeit at lower levels of significance than *IL10*, highlights a role for this locus in UC as well as CD. Population attributable risk fractions (%) in the combined panel (B–D) were 12.7, 13.3, 42.6, 47.8, 32.1 and 9.8 for SNPs rs3024505, rs12612347, rs9268480, rs9268858, rs9268877 and rs11805303, respectively (**Supplementary Methods**).

In the combined analysis, the most significant association outside the HLA complex in the replication analysis was obtained for rs3024505 located 1 kb downstream of the 3 UTR of *IL10* ($P = 1.35 \times 10^{-12}$; OR = 1.46 (1.31–1.62)). In a consecutive analysis we did not find strong evidence for an association between rs3024505 and the CD phenotype ($P = 0.013$; OR = 1.17 (1.01–1.34)) in 1,804 healthy controls and 1,848 CD cases. We saturated the *IL10* locus using an additional 22 HapMap tagging SNPs in an attempt to narrow down the association signal and to support the disease association of the lead SNP. All fine mapping SNPs were genotyped in the UC replication panels B to D, revealing associations that included rs3024495 in intron 4 ($P = 2.69 \times 10^{-11}$ in combined analysis) and rs3024493 in intron 3 ($P = 6.16 \times 10^{-12}$ in combined analysis), together with lead SNP rs3024505 (**Fig. 1** and **Supplementary Table 2** online). The risk alleles of these three SNPs were in perfect LD with each other ($r^2 = 1.00$), but not with any of the other 20 SNPs used for fine mapping ($r^2 \leq 0.20$). In a logistic regression analysis of the combined panels B–D, applying forward selection to the 23 fine-mapping SNPs, we achieved the best model fit for SNPs rs3024496 (3′ UTR), rs6658896 (14 kb 3′ of *IL10*) and rs4845140 (25 kb 5′ of *IL10*), in addition to lead SNP rs3024505, suggesting that more than one causal variant might contribute to the association signal at the *IL10* locus. A haplotype analysis of the latter four SNPs supported this finding (see **Supplementary Table 3** online).

These data taken together clearly identified *IL10* as a susceptibility gene for UC, but the causative variant(s) within the gene remained to be found. We resequenced the entire *IL10* gene—that is, the promoter, introns and exons—in 94 individuals with UC, 94 individuals with CD and 94 healthy controls (**Supplementary Fig. 6** online). In total, we identified 25 known SNPs and 19 additional SNPs, all of which were private or rare variants (**Supplementary Tables 4** and **5** online). No indel polymorphism was identified. Two private nonsynonymous SNPs not previously described (encoding F129Y and R177Q) were detected alongside a rare synonymous SNP (K135K) and the previously identified G15R variant, which has failed to show significant association with CD in previous studies[12]. Our results corroborate this finding and extend the association to UC, as only one heterozygote sample was found among the 282 samples that were resequenced.

Three key observations can be made regarding the possible causality of the variants identified by the present fine mapping and association analysis. First, we did not find any evidence of a UC association for any of the previously investigated SNPs at *IL10* promoter positions –592 (rs1800872), –819 (rs1800871) and –1082 (rs1800896)[13–23]. Only two out of as many as ten previous studies have so far yielded evidence for an association of UC with the SNPs at these positions[15,21]. That eight of these previous studies have been unable to generate evidence for an *IL10* association in UC highlights general limitations of study design and statistical power in a large fraction of historical candidate gene studies[24]. The variable evidence for the association detected between the promoter SNPs and UC is probably due to different levels of LD between the promoter SNPs and the main SNP responsible for the association signal which, according to our data, is likely to reside elsewhere at the *IL10* locus. Second, the close proximity (79 bp) of lead SNP rs3024505 to a highly conserved stretch of DNA at the 3′ end of the gene is of interest. This region has a high regulatory potential score[25] and contains a putative AP-1 binding motif. AP-1 is activated upon stimulation of macrophages by the bacterial cell-wall component lipopolysaccharide[26], and IL10 production could provide an important anti-inflammatory feedback mechanism. Finally, we carried out an extensive analysis of the influence of the two newly identified exonic variants on the interaction sites between IL10 and the high-affinity IL10 receptor A (IL10RA; see **Supplementary Note** and **Supplementary Figs. 7** and **8** online). To what extent carriage of either of these variants can be functionally linked to UC susceptibility in the individuals in question can, however, only be speculated.

Our findings clearly suggest that IL10 may be a key cytokine in UC pathogenesis. This hypothesis is strongly supported by the spontaneous colitis phenotype developed by *Il10*$^{-/-}$ mice, which seems to result from a defective anti-inflammatory counter-regulation in response to the commensal flora. Notably, a reduced IL10 *in vitro* regulation in inflammatory immune cells obtained from the mucosa of individuals of UC has been described, and therapeutic administration of human recombinant IL10 to individuals with UC had a positive clinical effect[27–29]. Subcutaneous administration of IL10 was not further evaluated as a therapeutic in UC after failing in clinical studies of CD. In light of these results, systemic or topical delivery of IL10 should be worthy of consideration for clinical trials in UC. As a delivery mechanism, genetically engineered IL10-secreting *Lactococcus lactis* strains have been developed as a potent tool to influence colonic mucosal immunoregulation[30]. A role of IL10 has also been suggested in other forms of chronic inflammation, for example rheumatoid arthritis, lupus erythematosus and psoriasis, and it should be worthwhile to investigate the *IL10* locus by a haplotype-tagging approach in these conditions.

IL10 signals through STAT3- and MAPK-mediated pathways to trigger anti-inflammatory mechanisms dependent on suppressor of cytokine signaling[26]. Recently, a targeted assessment of CD susceptibility genes for their role in UC revealed an association between a SNP in *STAT3* and UC[5]. SNP rs744166, located in intron 2 of *STAT3* at 17q21.2, which had been associated with CD in the Wellcome Trust Case Control Consortium (WTCCC) genome-wide association study[7], proved to be associated with UC in that study (OR = 0.77 (0.66–0.90), $P = 5.00 \times 10^{-4}$), but not with CD. In this context, it is of great interest to note retrospectively that SNP rs7212299, located upstream of *STAT3*, was also found to be associated with UC in the present study (panel A). Because of the low level of significance ($P = 0.01$, see **Supplementary Table 6** online) the SNP was, however, not included in the replication phase of this study. We anticipate that further characterization of critical components of the IL10–STAT3 signaling pathway may point to important therapeutic targets and provide unique insights into the pathogenesis of UC.

METHODS

Participants. The diagnosis of UC or CD was based on typical clinical, radiological, histological and endoscopic (type and distribution of lesions) findings. The full recruitment details for all study panels are given in **Supplementary Methods**. All participants gave written informed consent, and the recruitment protocols were approved by the ethics committees at the respective recruiting institutions.

Genotyping and sequencing. The genotyping for the GWAS was performed as a service project by Affymetrix using the Genome-Wide Human SNP Array 5.0 (500K). Genotypes were assigned using the BRLMM-p algorithm. Samples with more than 5% missing genotypes, who showed excess genetic dissimilarity to the other subjects (see **Supplementary Fig. 2**), or who showed evidence for cryptic relatedness to other study participants (see **Supplementary Fig. 2**) were not included. These quality control measures left 1,167 UC samples and 777 control samples for inclusion in screening panel A. SNPs were excluded ($n = 85{,}532$; 19.4% of all SNPs) that had a low genotyping rate (<95% in cases or controls), were monomorphic or rare (minor allele frequency <2% in cases or controls), or deviated from Hardy-Weinberg equilibrium (HWE) in the control sample ($P_{HWE} \leq 0.01$).

All downstream genotyping was done with SNPlex and TaqMan technologies (Applied Biosystems) using an automated laboratory setup and all process data were written to and administered by a database-driven laboratory information management system. Sequencing of genomic DNA was done using BigDye Terminator v3.1 chemistry (Applied Biosystems) and an ABI3730 capillary sequencer (Applied Biosystems) according to manufacturer's protocols. Traces were inspected for SNPs and indels using novoSNP v2.03. See **Supplementary Methods** for further details.

Statistical analysis. Genome-wide association analyses were conducted with PLINK v1.01 in combination with gPLINK v2.049 and GENOMIZER v1.2.0. Single-marker analyses, permutation tests, estimation of pair-wise linkage disequilibrium (LD) and SNP selection were done using Haploview v4.0. Logistic regressions were done within LOGISTIC of the SAS software package (SAS Institute). For additional details, see **Supplementary Methods**.

Note: Supplementary information is available on the Nature Genetics website.

ACKNOWLEDGMENTS

We wish to thank all individuals with IBD, families and physicians for their cooperation. We gratefully acknowledge the cooperation of the German Crohn and Colitis Foundation (Deutsche Morbus Crohn und Colitis Vereinigung e.V.), the BMBF competence network "IBD" and of the contributing gastroenterologists. Finally, we wish to thank T. Wesse, T. Henke, A. Dietsch, R. Vogler, C.v.d. Lancken and M. Friskovec for expert technical help. We thank T. Wienker and M. Steffens (IMBIE, University of Bonn) for performing the quality control of the GWAS datasets. We thank B.A. Lie and the Norwegian Bone Marrow Donor Registry at Rikshospitalet University Hospital, Oslo, for contributing the healthy Norwegian control population. This study was supported by the German Ministry of Education and Research (BMBF) through the

National Genome Research Network (NGFN), the popgen biobank and the SFB617 "Molecular Mechanisms of Epithelial Defense." The project received infrastructure support through the DFG excellence cluster "Inflammation at Interfaces." We acknowledge use of DNA from the 1958 British Birth Cohort collection (R. Jones, S. Ring, W. McArdle and M. Pembrey), funded by the UK MRC (grant G0000934) and The Wellcome Trust (grant 068545/Z/02). C.G.M. and C.M.O. were supported by The Wellcome Trust and CORE (UK).

AUTHOR CONTRIBUTIONS

A.F., T.B. and J.S. performed the SNP selection, genotyping, data analysis, resequencing, and prepared the figures and tables; T.H.K. helped with data analysis. S.N., C.S. and P. Rosenstiel coordinated the recruitment and collected the phenotype data. M.N. and D.E. helped with data analysis and quality control; G.M., F.S.D. and M.A. performed *in silico* protein analysis and contributed to writing the manuscript. Norwegian, Belgian, Dutch and Greek patient samples were provided by M.H.V., S.V., P. Rutgeerts, R.K.W., P.C.F.S. and M.G., respectively. W.L.M. and D.S. provided the UK and C.W. the Dutch control samples; M.K. supervised and performed the statistical analysis and edited the paper. C.M.O. and C.G.M. provided the UK patient sample and edited the manuscript. A.F., T.H.K. and S.S. designed and supervised the experiment and wrote the manuscript. All authors approved the final manuscript.

Published online at http://www.nature.com/naturegenetics/
Reprints and permissions information is available online at http://npg.nature.com/reprintsandpermissions/

1. Loftus, E.V. Jr. Clinical epidemiology of inflammatory bowel disease: incidence, prevalence, and environmental influences. *Gastroenterology* **126**, 1504–1517 (2004).
2. Bouma, G. & Strober, W. The immunological and genetic basis of inflammatory bowel disease. *Nat. Rev. Immunol.* **3**, 521–533 (2003).
3. Mathew, C.G. New links to the pathogenesis of Crohn disease provided by genome-wide association scans. *Nat. Rev. Genet.* **9**, 9–14 (2008).
4. Duerr, R.H. *et al.* A genome-wide association study identifies IL23R as an inflammatory bowel disease gene. *Science* **314**, 1461–1463 (2006).
5. Franke, A. *et al.* Replication of signals from recent studies of Crohn's disease identifies previously unknown disease loci for ulcerative colitis. *Nat. Genet.* **40**, 713–715 (2008).
6. Fisher, S.A. *et al.* Genetic determinants of ulcerative colitis include the ECM1 locus and five loci implicated in Crohn's disease. *Nat. Genet.* **40**, 710–712 (2008).
7. Devlin, B. & Roeder, K. Genomic control for association studies. *Biometrics* **55**, 997–1004 (1999).
8. Stokkers, P.C., Reitsma, P.H., Tytgat, G.N. & van Deventer, S.J. HLA-DR and -DQ phenotypes in inflammatory bowel disease: a meta-analysis. *Gut* **45**, 395–401 (1999).
9. Marangoni, F. *et al.* WASP regulates suppressor activity of human and murine CD4(+)CD25(+)FOXP3(+) natural regulatory T cells. *J. Exp. Med.* **204**, 369–380 (2007).
10. Webb, M.C., Andrews, P.A., Koffman, C.G. & Cameron, J.S. Renal transplantation in Wiskott-Aldrich syndrome. *Transplantation* **56**, 1585 (1993).
11. van Heel, D.A. *et al.* Inflammatory bowel disease susceptibility loci defined by genome scan meta-analysis of 1952 affected relative pairs. *Hum. Mol. Genet.* **13**, 763–770 (2004).
12. van der Linde, K. *et al.* A functional interleukin-10 mutation in Dutch patients with Crohn's disease. *Dig. Liver Dis.* **37**, 330–335 (2005).
13. Balding, J. *et al.* Inflammatory bowel disease: the role of inflammatory cytokine gene polymorphisms. *Mediators Inflamm.* **13**, 181–187 (2004).
14. Parkes, M., Satsangi, J. & Jewell, D. Contribution of the IL-2 and IL-10 genes to inflammatory bowel disease (IBD) susceptibility. *Clin. Exp. Immunol.* **113**, 28–32 (1998).
15. Tagore, A. *et al.* Interleukin-10 (IL-10) genotypes in inflammatory bowel disease. *Tissue Antigens* **54**, 386–390 (1999).
16. Koss, K., Satsangi, J., Fanning, G.C., Welsh, K.I. & Jewell, D.P. Cytokine (TNF alpha, LT alpha and IL-10) polymorphisms in inflammatory bowel diseases and normal controls: differential effects on production and allele frequencies. *Genes Immun.* **1**, 185–190 (2000).
17. Klein, W. *et al.* The IL-10 gene is not involved in the predisposition to inflammatory bowel disease. *Electrophoresis* **21**, 3578–3582 (2000).
18. Aithal, G.P. *et al.* Role of polymorphisms in the interleukin-10 gene in determining disease susceptibility and phenotype in inflamatory bowel disease. *Dig. Dis. Sci.* **46**, 1520–1525 (2001).
19. Kim, T.H. *et al.* Tumor necrosis factor-alpha and interleukin-10 gene polymorphisms in Korean patients with inflammatory bowel disease. *Korean J. Gastroenterol.* **42**, 377–386 (2003).
20. Celik, Y. *et al.* Cytokine gene polymorphisms in Turkish patients with inflammatory bowel disease. *Scand. J. Gastroenterol.* **41**, 559–565 (2006).
21. Fernandez, L. *et al.* Interleukin-10 polymorphisms in Spanish patients with IBD. *Inflamm. Bowel Dis.* **11**, 739–743 (2005).
22. Cantor, M.J., Nickerson, P. & Bernstein, C.N. The role of cytokine gene polymorphisms in determining disease susceptibility and phenotype in inflammatory bowel disease. *Am. J. Gastroenterol.* **100**, 1134–1142 (2005).
23. Castro-Santos, P., Suarez, A., Lopez-Rivas, L., Mozo, L. & Gutierrez, C. TNFalpha and IL-10 gene polymorphisms in inflammatory bowel disease. Association of -1082 AA low producer IL-10 genotype with steroid dependency. *Am. J. Gastroenterol.* **101**, 1039–1047 (2006).
24. Trikalinos, T.A., Ntzani, E.E., Contopoulos-Ioannidis, D.G. & Ioannidis, J.P. Establishment of genetic associations for complex diseases is independent of early study findings. *Eur. J. Hum. Genet.* **12**, 762–769 (2004).
25. Kolbe, D. *et al.* Regulatory potential scores from genome-wide three-way alignments of human, mouse, and rat. *Genome Res.* **14**, 700–707 (2004).
26. Hu, X., Chen, J., Wang, L. & Ivashkiv, L.B. Crosstalk among Jak-STAT, Toll-like receptor, and ITAM-dependent pathways in macrophage activation. *J. Leukoc. Biol.* **82**, 237–243 (2007).
27. Mitsuyama, K. *et al.* Interleukin-10 in the pathophysiology of inflammatory bowel disease: increased serum concentrations during the recovery phase. *Mediators Inflamm.* **2006**, 26875 (2006).
28. Schreiber, S., Heinig, T., Thiele, H.G. & Raedler, A. Immunoregulatory role of interleukin 10 in patients with inflammatory bowel disease. *Gastroenterology* **108**, 1434–1444 (1995).
29. Schreiber, S. *et al.* Ulcerative Colitis IL-10 Cooperative Study Grp TI Safety and tolerance of rHuIL-10 treatment in patients with mild/moderate active ulcerative colitis. *Gastroenterology* **114**, A1080–A1081 (1998).
30. Steidler, L. *et al.* Treatment of murine colitis by Lactococcus lactis secreting interleukin-10. *Science* **289**, 1352–1355 (2000).