nature genetics

Genome-wide association identifies multiple ulcerative colitis susceptibility loci

Dermot P B McGovern[1*], Agnès Gardet[2], Leif Törkvist[3], Philippe Goyette[4], Jonah Essers[5], Kent D Taylor[6], Benjamin M Neale[5], Rick T H Ong[7], Caroline Lagacé[4], Chun Li[2], Todd Green[8], Christine R Stevens[8], Claudine Beauchamp[4], Phillip R Fleshner[1], Marie Carlson[9], Mauro D'Amato[10], Jonas Halfvarson[11], Martin L Hibberd[12], Mikael Lördal[13], Leonid Padyukov[14], Angelo Andriulli[15], Elisabetta Colombo[15], Anna Latiano[15], Orazio Palmieri[15], Edmond-Jean Bernard[16], Colette Deslandres[17], Daan W Hommes[18], Dirk J de Jong[19], Pieter C Stokkers[20], Rinse K Weersma[21], The NIDDK IBD Genetics Consortium[22], Yashoda Sharma[23], Mark S Silverberg[24], Judy H Cho[23,25], Jing Wu[26], Kathryn Roeder[27], Steven R Brant[27], L Phillip Schumm[28], Richard H Duerr[29], Marla C Dubinsky[1], Nicole L Glazer[30], Talin Haritunians[6], Andy Ippoliti[1], Gil Y Melmed[1], David S Siscovick[30], Eric A Vasiliauskas[1], Stephan R Targan[1], Vito Annese[15], Cisca Wijmenga[31], Sven Pettersson[32,33], Jerome I Rotter[6], Ramnik J Xavier[2,8], Mark J Daly[5,8], John D Rioux[4] & Mark Seielstad[7,34,35]

Ulcerative colitis is a chronic, relapsing inflammatory condition of the gastrointestinal tract with a complex genetic and environmental etiology. In an effort to identify genetic variation underlying ulcerative colitis risk, we present two distinct genome-wide association studies of ulcerative colitis and their joint analysis with a previously published scan[1], comprising, in aggregate, 2,693 individuals with ulcerative colitis and 6,791 control subjects. Fifty-nine SNPs from 14 independent loci attained an association significance of $P < 10^{-5}$. Seven of these loci exceeded genome-wide significance ($P < 5 \times 10^{-8}$). After testing an independent cohort of 2,009 cases of ulcerative colitis and 1,580 controls, we identified 13 loci that were significantly associated with ulcerative colitis ($P < 5 \times 10^{-8}$), including the immunoglobulin receptor gene *FCGR2A*, 5p15, 2p16 and *ORMDL3* (orosomucoid1-like 3). We confirmed association with 14 previously identified ulcerative colitis susceptibility loci, and an analysis of acknowledged Crohn's disease loci showed that roughly half of the known Crohn's disease associations are shared with ulcerative colitis. These data implicate approximately 30 loci in ulcerative colitis, thereby providing insight into disease pathogenesis.

Epidemiological studies suggest that ulcerative colitis and Crohn's disease share some, but not all, susceptibility genes, a hypothesis that is supported by genome-wide association (GWA) studies[1–5]. Meta-analysis of three GWA studies of Crohn's disease increased the number of known susceptibility loci for this disease to over 30 (ref. 2). Ulcerative colitis is a condition with lower heritability (λ_s (sibling recurrence risk) = 10–15) as compared with Crohn's disease (λ_s = 20–35) and, perhaps as a result, fewer loci have been identified for ulcerative colitis[1,3–6]. Previous experience with Crohn's disease implies that the as yet identified ulcerative colitis susceptibility loci explain only a fraction of the genetic contribution to disease susceptibility. Considerations of statistical power also suggest that additional loci might be found by enlarging the numbers of cases and controls used in genome-wide discovery. Here we have combined data from two new GWA studies of ulcerative colitis and performed a meta-analysis with a published study[1]. This analysis brings together a discovery set of 2,693 individuals with ulcerative colitis and 6,791 control subjects, all of European descent (**Supplementary Table 1**). We then independently replicated the top results from this meta-analysis with 2,009 cases and 1,580 controls, who were also of European origin.

All primary studies used similar Illumina BeadChips, enabling us to examine directly 266,047 (258,137 autosomal and 7,910 X-chromosomal) SNPs that passed quality control in each study. The three studies were analyzed and corrected for population structure separately. *P* values from each study were converted to *Z* scores summarizing the direction and magnitude of association evidence; these *Z* scores (weighted by the relative size of each study) were then combined by using standard methods[7]. A *Q*–*Q* plot (**Supplementary Fig. 1**) shows a significant excess of likely true positives in the tail of the distribution, against only modest overall inflation (λ_{GC} (genomic control inflation factor) = 1.036).

We sought to replicate 14 independent loci represented by one or more of the 59 SNPs with $P < 10^{-5}$ in the meta-analysis (**Table 1** and **Supplementary Table 2**). Several loci previously associated with ulcerative colitis were among those identified in the meta-analysis, although not all had previously attained genome-wide significance. Loci

*A full list of author affiliations appears at the end of the paper.

Received 25 September 2009; accepted 26 January 2010; published online 14 March 2010; doi:10.1038/ng.549

Table 1 Loci associated with ulcerative colitis in the GWA meta-analysis and replication studies

SNP	Chr.	Position	Genes	Allele (major/minor)	*P* value							
					GWA studies				Replication			GWA + replication
					Cedars	Swedish	NIDDK	Combined	Italian	Dutch	Combined	
rs3806308[a]	1	19612341	*RNF186, OTUD3, PLA2G2E*	G/A	0.034	0.047	4.7×10^{-8}	3.3×10^{-8}	–	–	NA	NA
rs1317209	1	20012623	*RNF186, OTUD3, PLA2G2E*	C/T	2.4×10^{-4}	6.3×10^{-3}	0.034	8.6×10^{-7}	0.046	1.6×10^{-4}	4.4×10^{-5}	1.6×10^{-10}
rs6426833	1	20044447	*RNF186, OTUD3, PLA2G2E*	G/A	1.0×10^{-4}	2.8×10^{-4}	6.8×10^{-10}	2.7×10^{-15}	4.9×10^{-4}	3.9×10^{-5}	7.7×10^{-8}	1.7×10^{-21}
rs2201841[b]	1	67466790	*IL23R*	T/C	0.023	4.9×10^{-3}	1.1×10^{-6}	8.9×10^{-9}	6.7×10^{-3}	9.9×10^{-5}	3.0×10^{-6}	1.3×10^{-13}
rs11209026	1	67478546	*IL23R*	G/A	7.3×10^{-4}	1.5×10^{-3}	3.1×10^{-5}	5.9×10^{-10}	0.012	1.3×10^{-3}	5.2×10^{-5}	1.9×10^{-13}
rs10800309	1	159738782	*FCGR2A, FCGR2C*	G/A	2.6×10^{-3}	0.083	7.3×10^{-4}	2.8×10^{-6}	2.2×10^{-4}	0.14	2.5×10^{-4}	2.8×10^{-9}
rs3024505	1	205006527	*IL10, IL19*	C/T	3.1×10^{-3}	0.012	9.3×10^{-3}	3.2×10^{-6}	2.1×10^{-3}	0.12	1.1×10^{-3}	1.4×10^{-8}
rs6706689	2	61024549	*REL, CCDC139, PUS10*	G/A	0.038	2.0×10^{-4}	0.031	4.4×10^{-6}	1.1×10^{-3}	0.15	8.9×10^{-4}	1.5×10^{-8}
rs13003464[b]	2	61040333	*REL, CCDC139, PUS10*	A/G	0.012	2.4×10^{-3}	9.5×10^{-5}	4.7×10^{-8}	0.21	0.027	0.014	7.4×10^{-9}
rs3197999	3	49696536	*MST1*	C/T	4.6×10^{-3}	2.0×10^{-4}	0.07	1.4×10^{-6}	1.9×10^{-4}	0.28	6.7×10^{-4}	3.8×10^{-9}
rs4957048	5	636180	*CEP72, TPPP*	C/T	0.19	5.3×10^{-3}	1.2×10^{-3}	2.3×10^{-5}	3.5×10^{-4}	7.1×10^{-3}	9.4×10^{-6}	1.2×10^{-9}
rs2395185[a]	6	32541145	*C6orf10, BTNL2*	G/T	1.4×10^{-7}	4.4×10^{-12}	1.4×10^{-6}	8.8×10^{-23}	–	–	NA	NA
rs4598195	7	107290677	*DLD, LAMB1*	A/C	0.098	4.9×10^{-4}	1.3×10^{-5}	4.2×10^{-8}	0.46	0.08	0.075	7.7×10^{-8}
rs4077515[b]	9	138386317	*CARD9*	A/G	2.0×10^{-3}	0.012	4.9×10^{-3}	1.2×10^{-6}	6.3×10^{-4}	0.75	8.3×10^{-3}	5.5×10^{-8}
rs11190140	10	101281583	*NKX2-3*	C/T	8.8×10^{-4}	0.028	0.058	1.9×10^{-5}	7.6×10^{-6}	0.37	1.5×10^{-4}	1.1×10^{-8}
rs1558744	12	66790859	*IFNG, IL26*	G/A	4.2×10^{-3}	0.028	5.5×10^{-10}	8.1×10^{-11}	2.4×10^{-4}	0.57	2.7×10^{-3}	4.2×10^{-12}
rs971545	12	66877952	*IFNG, IL26*	A/G	2.4×10^{-4}	0.097	3.1×10^{-4}	2.4×10^{-7}	0.012	0.055	1.7×10^{-3}	2.2×10^{-9}
rs2305480[b,c]	17	35315722	*ORMDL3* region	C/T	3.9×10^{-4}	8.5×10^{-4}	0.18	2.1×10^{-6}	0.19	4.3×10^{-3}	3.2×10^{-3}	3.0×10^{-8}
rs2836878[a]	21	39387404	Near *PSMG1*	G/A	1.9×10^{-3}	2.7×10^{-4}	2.1×10^{-3}	1.4×10^{-8}	–	–	NA	NA

Table lists SNPs with a significance of association at or near $P < 5 \times 10^{-8}$. Chr., chromosome; NA, not available.
[a]SNP assays failed to design (suitable genotyping primers could not be found) or failed genotyping in replication stage. [b]Highly correlated proxies ($r^2 > 0.9$) were typed in the replication stage. See **Supplementary Table 2** for further details. [c]Region contains *STAC2*, *FBXL20*, *MED1*, *CRKRS*, *NEUROD2*, *PPP1R1B*, *STARD3*, *TCAP*, *PNMT*, *PERLD1*, *ERBB2*, C17orf37, *GRB7*, *IKZF3*, *ZPBP2*, *GSDML* and *ORMDL3*.

previously associated with ulcerative colitis included *IL23R* (ref. 3), the HLA region (including the *BTNL2* association)[1,3,8], *MST1* (ref. 9), *CARD9* (ref. 6), 1q32 (near *IL10*)[5], 1p36 (*RNF186-OTUD3-PLA2G2E*)[1], *DLD-LAMB1* (ref. 1; also confirmed in a British ulcerative colitis GWA study[10]), 12q15 (neighboring *IFNG-IL26*)[1] and 21q22 (ref. 11). A second, back-up SNP was chosen for the six regions with $P < 10^{-6}$ in the discovery set. We chose additional SNPs ($n = 4$) from the 1p36, 2p16 and 12q15 regions because these SNPs seemed to be independently associated ($r^2 < 0.2$ in CEU HapMap). We also included ten SNPs with *P* values between 10^{-4} and 10^{-5}. Lastly, we included previously published associations with Crohn's disease and ulcerative colitis. These SNPs were typed in 2,009 individuals with ulcerative colitis and 1,580 control subjects of Dutch and Italian descent. (Complete data are given in **Supplementary Table 1**.)

Overall, 13 loci achieved association with genome-wide statistical significance ($P < 5 \times 10^{-8}$; **Table 1** and **Supplementary Table 2**). A fourteenth locus, containing the *CARD9* gene, did not reach genome-wide significance but was significant after Bonferroni correction ($P_{nominal} = 5.48 \times 10^{-8}$; $P_{corrected} = 0.014$). We achieved genome-wide significance for at least four loci new in populations of European origins, including 1q21 (*FCGR2A-FCGR2C*; this locus has been associated with ulcerative colitis in a Japanese population[12]), 2p16 (*REL-PUS10*), 17q12 (*ORMDL3*) and 5p15 (rs4957048; ~30 kb from *CEP72*; **Table 1**). Fourteen loci that have been either suggested to be or conclusively associated with ulcerative colitis in previous studies, namely *TNFSF15*, *NKX2-3*, *IL12B*, *MST1*, *IL18RAP*, the HLA region, the *IBD5* locus (5q31), *RNF186-OTUD3-PLA2G2E*, *DLD-LAMB1*, *IL10*, *CARD9*, 12q15 (*IFNG-IL26*), *JAK2* and *IL23R*, were replicated in our study. We also showed an association with ten more loci previously associated with Crohn's disease or inflammatory bowel disease (IBD), namely *IRGM*, *KIF21B*, *IKZF1*, *ICOSLG*, *CCL2-CCL7*, 5p13 (near *PTGER4*), 21q21, *CUL2-CREM*, *PSMG1* and *STAT3* (**Table 2**). Replication of these previously identified ulcerative colitis, Crohn's disease or IBD loci was defined as an association at a level of $P < 0.05$ with the same risk allele identified in the index studies.

Two studies performed in populations of European descent identified further loci associated with IBD[10,13]. We examined these loci within our discovery set (using the exact SNP or a perfect proxy, $r^2 = 1.0$) and found association between ulcerative colitis and 22q12 (an IBD-associated locus[13]; $P = 0.049$), 16q22/*CDH1* (an ulcerative colitis–associated locus[10]; $P = 0.0061$) and the *DLD-LAMB1* locus[10] discussed earlier. In addition, we found further evidence of association between the 1p36 locus (rs7524102, $P = 0.0015$), which showed strong (but not genome-wide significant) association in the WTCCC2

study[10]. We were not able to replicate the ulcerative colitis associations seen at 13q13 ($P = 0.15$), 20q13/*HNF4A* and 2q37/*GPR35* (SNPs not in our data set and no proxy SNPs available) or the IBD associations seen at 16p11 ($P = 0.17$), 10q22/*ZMIZ1* ($P = 0.22$) and 19q13 (SNP not in our data set and no proxy SNP available). Therefore, taking into account these additional associations, our analysis provides evidence for at least 30 distinct risk factors for ulcerative colitis.

In three regions, 1p36 (*RNF186-OTUD3-PLA2G2E*), 2p16 (*PUS10-REL*) and 12q15 (*IFNG-IL26*), our scan identified several SNPs with $P < 10^{-5}$ that showed low linkage disequilibrium (LD) with each other

Table 2 Association of previously reported Crohn's disease risk loci with ulcerative colitis

Gene of interest	SNP from ref. 2	Chr.	CD meta-analysis *P* value (ref. 2)	*P* value							
				GWA studies				Replication			GWA + replication
				Cedars	Swedish	NIDDK	Combined	Italian	Dutch	Combined	
PTPN22	rs2476601	1p13	1.8×10^{-5}	0.11	0.32	0.096	0.19	0.25	0.15	0.067	0.88
IL23R	rs11465804[a–d]	1p31	6.4×10^{-34}	7.3×10^{-4}	1.5×10^{-3}	3.1×10^{-5}	5.9×10^{-10}	0.012	1.3×10^{-3}	5.2×10^{-5}	1.9×10^{-13}
ITLN1	rs2274910[b]	1q23	3.5×10^{-7}	0.11	0.15	0.24	0.45	0.51	0.52	0.36	0.24
–	rs9286879	1q24	4.0×10^{-7}	0.90	0.75	0.85	0.71	0.19	0.063	0.70	0.60
KIF21B	rs11584383[a]	1q32	1.9×10^{-6}	0.16	0.024	0.085	1.9×10^{-3}	0.020	1.1×10^{-4}	1.2×10^{-5}	2.2×10^{-7}
PUS10	rs13003464[c]	2p16	7.7×10^{-6}	0.012	2.4×10^{-3}	9.5×10^{-5}	4.7×10^{-8}	0.21	0.027	0.014	7.4×10^{-9}
GCKR	rs780094[b]	2p23	7.2×10^{-5}	0.70	0.95	0.18	0.30	0.29	0.47	0.81	0.81
IL18RAP	rs917997	2q11	2.2×10^{-5}	0.58	0.58	0.012	0.15	0.35	0.026	0.025	0.011
ATG16L1	rs3828309[a]	2q37	2.6×10^{-21}	0.24	0.18	0.25	0.57	0.28	0.63	0.67	0.48
MST1	rs3197999[c]	3p21	2.2×10^{-7}	4.6×10^{-3}	2.0×10^{-4}	0.07	1.4×10^{-6}	1.9×10^{-4}	0.28	6.7×10^{-4}	3.8×10^{-9}
PTGER4	rs4613763	5p13	5.0×10^{-22}	0.57	0.62	4.9×10^{-3}	0.11	0.048	1.6×10^{-3}	2.8×10^{-4}	4.2×10^{-4}
IBD5 locus	rs2188962[b]	5q31	4.6×10^{-9}	0.14	0.40	0.32	0.35	0.044	2.5×10^{-3}	3.7×10^{-4}	2.9×10^{-3}
IRGM	rs13361189	5q33	8.2×10^{-11}	–	–	–	NA	0.015	4.8×10^{-5}	4.3×10^{-6}	NA
IL12B	rs10045431	5q33	8.8×10^{-9}	–	–	–	NA	0.024	0.023	1.4×10^{-3}	NA
BTNL2, *SLC26A3*, *HLA-DRB1*, *HLA-DQA1*	rs3763313	6p21	1.5×10^{-8}	0.015	0.20	0.25	5.1×10^{-3}	0.020	0.71	0.17	0.19
CDKAL1	rs6908425	6p22	2.5×10^{-7}	0.49	0.61	0.17	0.14	0.29	0.97	0.44	0.51
LYRM4	rs12529198	6p25	7.1×10^{-7}	–	–	–	NA	0.23	0.78	0.51	NA
SLC22A23	rs17309827[e]	6p25	2.1×10^{-6}	–	–	–	NA	–	–	NA	NA
ATG5	rs7746082[a]	6q21	3.1×10^{-4}	0.069	5.7×10^{-3}	0.042	1.3×10^{-4}	0.23	0.88	0.46	6.1×10^{-4}
–	rs7758080	6q25	7.3×10^{-6}	0.67	0.18	0.54	0.38	0.50	0.079	0.44	0.84
CCR6	rs2301436	6q27	3.3×10^{-7}	0.044	0.16	0.094	3.3×10^{-3}	0.38	0.28	0.89	0.029
IKZF1	rs1456893[a]	7p12	3.2×10^{-5}	0.93	0.66	0.53	0.96	0.15	1.1×10^{-3}	9.1×10^{-4}	0.040
–	rs1551398[b]	8q24	4.9×10^{-6}	0.57	0.32	0.73	0.27	4.2×10^{-4}	0.057	0.25	0.90
JAK2	rs10758669	9p24	6.8×10^{-7}	0.023	0.028	0.49	2.9×10^{-3}	8.2×10^{-4}	0.023	7.1×10^{-5}	1.4×10^{-6}
TNFSF15	rs4263839[a,b]	9q32	2.6×10^{-7}	0.024	0.42	0.56	0.035	0.30	2.9×10^{-4}	9.7×10^{-4}	2.0×10^{-4}
CUL2, *CREM*	rs17582416[a,b]	10p11	2.2×10^{-5}	0.69	0.028	0.92	0.15	0.032	0.45	0.040	0.016
ZNF365	rs10995271[a]	10q21	1.6×10^{-7}	0.024	0.73	0.89	0.24	0.19	0.96	0.34	0.13
NKX2-3	rs11190140	10q24	1.7×10^{-10}	8.8×10^{-4}	0.028	0.058	1.9×10^{-5}	7.6×10^{-6}	0.37	1.5×10^{-4}	1.1×10^{-8}
C11orf30	rs7927894[a]	11q13	1.4×10^{-7}	0.021	0.79	0.062	0.010	0.20	0.69	0.23	0.22
LRRK2, *MUC19*	rs11175593	12q12	1.3×10^{-7}	–	–	–	NA	0.071	0.77	0.28	NA
–	rs3764147[a]	13q14	1.4×10^{-5}	0.27	0.010	0.041	0.34	0.59	0.76	0.55	0.27
CARD15 (R702W)	rs2066844	16q12	NA	–	–	–	NA	0.25	0.64	0.25	NA
CARD15 (G908R)	rs2066845	16q12	NA	–	–	–	NA	0.69	0.017	0.049	NA
CARD15 (insC)	rs2066847	16q12	NA	–	–	–	NA	0.90	0.077	0.18	NA
CARD15 (R459R)	rs2066843[e]	16q12	1.2×10^{-27}	0.36	0.99	0.62	0.41	–	–	NA	NA
CCL2, *CCL7*	rs991804[b]	17q12	4.0×10^{-6}	0.52	0.060	0.46	0.25	0.013	0.028	9.3×10^{-4}	2.9×10^{-3}
ORMDL3	rs2872507[a,c]	17q21	2.1×10^{-6}	3.9×10^{-4}	8.5×10^{-4}	0.18	2.1×10^{-6}	0.19	4.3×10^{-3}	3.2×10^{-3}	3.0×10^{-8}
STAT3	rs744166	17q21	5.9×10^{-6}	0.35	0.77	0.012	0.031	0.40	0.084	0.069	4.8×10^{-3}
PTPN2	rs2542151[a]	18p11	6.5×10^{-11}	0.14	0.57	0.75	0.17	0.98	0.17	0.32	0.091
–	rs8098673	18q11	3.2×10^{-5}	0.97	0.86	0.92	0.89	0.16	0.29	0.81	0.96
–	rs4807569[a,e]	19p13	1.3×10^{-8}	0.99	0.53	0.058	0.15	–	–	NA	
–	rs1736135[a]	21q21	3.3×10^{-5}	0.36	0.053	5.9×10^{-4}	2.9×10^{-4}	0.60	8.2×10^{-7}	1.2×10^{-4}	1.5×10^{-7}
ICOSLG	rs762421[b]	21q22	1.1×10^{-5}	–	–	–	NA	1.7×10^{-3}	0.59	9.3×10^{-3}	NA

Chr., chromosome; NA, not available; CD, Crohn's disease.
[a]Highly correlated proxies were evaluated in the ulcerative colitis GWA studies (see **Supplementary Table 2** for further details). [b]Highly correlated proxies were typed in the replication stage (see **Supplementary Table 2** for further details). [c]Also reported in **Table 1**. [d]rs11465804 was captured by the highly correlated proxy rs11209026 in both scan and replication (**Table 1** and **Supplementary Table 2**). [e]SNP assays failed to design (suitable genotyping primers could not be identified) or failed genotyping in replication stage.

Table 3 Conditional analyses for three regions with multiple associated SNPs in linkage equilibrium

SNP	Location	Conditional on	*P* value					Combined *P*
			Cedars	NIDDK	Swedish	Italian	Dutch	
rs1317209	1p36	rs3806308	8.0×10^{-5}	0.043	0.0045			2.8×10^{-6}
		rs6426833	2.7×10^{-4}	0.044	0.01	0.0498	1.9×10^{-4}	3.1×10^{-8}
rs3806308	1p36	rs1317209	0.022	3.4×10^{-8}	0.06			1.4×10^{-8}
		rs6426833	0.035	2.5×10^{-7}	0.13			2.6×10^{-7}
rs6426833	1p36	rs3806308	6.4×10^{-5}	3.1×10^{-9}	3.5×10^{-4}			5.1×10^{-14}
		rs1317209	1.1×10^{-4}	5.7×10^{-10}	2.2×10^{-4}	5.5×10^{-4}	7.5×10^{-5}	$<1 \times 10^{-16}$
rs6706689	2p16	rs13003464	0.31	0.83	0.017	0.0026	0.71	0.0091
rs13003464	2p16	rs6706689	0.086	7.0×10^{-4}	0.12	0.88	0.068	0.0011
rs1558744	12q15	rs971545	0.044	5.0×10^{-8}	0.20	0.0018	0.95	2.2×10^{-8}
rs971545	12q15	rs1558744	0.0021	0.046	0.20	0.092	0.061	3.9×10^{-4}

($r^2 < 0.2$) in HapMap CEU samples. To determine whether several independent alleles were contributing at these loci, we performed conditional analyses in all scan and replication samples (**Table 3**). Although the signals for each SNP conditional on the other were greatly diminished in the 2p16 region (suggesting that these SNPs may both be in LD with a single, as yet unidentified, causal variant), we extended previous evidence for several independent associations at 1p36 and 12q15 (ref. 1). Particularly noteworthy are three SNPs at 1p36, each of which is genome-wide significant even while conditional on the other two.

We also analyzed interaction between all pairs of SNPs listed in **Table 1**. Among the 496 pairs of SNPs examined (**Supplementary Table 3**), one pair (an interaction between SNPs at *CARD9* and *REL-PUS10*) was significant after correction for the number of tests performed, and a second interaction at this locus approached significance after replication. Few such interactions have been documented in complex disease[14], and further replication is warranted before this finding can be considered confirmed, but the known functional interaction between the CARD9 and REL proteins should not be overlooked.

It is clear from these (**Table 2**) and other data that ulcerative colitis and Crohn's disease share some mechanistic pathways and susceptibility genes but that some pathways and genes are particular to each condition. We sought to estimate the proportion of alleles with an influence on both Crohn's disease and ulcerative colitis by calculating the likelihood of the observed ulcerative colitis genotype data at each Crohn's disease locus, under the alternative hypotheses that the ulcerative colitis sample has the same allele frequency as a Crohn's disease sample or that it has the same allele frequency as a control sample. The maximum data set likelihood was achieved when 15 or 16 of the Crohn's disease loci (essentially half of the 31) were presumed to affect risk for ulcerative colitis. This observation is consistent with the summarized results in **Table 2**, where 14 of the 31 confirmed previous findings are significant at $P < 0.01$, three have $0.01 < P < 0.05$ and the remainder fit the null distribution. The lack of complete overlap is unlikely to result entirely from limited power because our likelihood analysis limits the possibility that more than 20 of the Crohn's loci are shared, and some of the strongest Crohn's disease variants (for example, *CARD15* and *ATG16L1*) are among the loci not associated with ulcerative colitis, despite the higher power to detect these variants of larger effect size.

All of the genes implicated by our work could plausibly have a functional role in ulcerative colitis (see **Supplementary Table 4** for a summary of genes located near replicated ulcerative colitis risk SNPs not shared with Crohn's disease). A key next step in translating genetic loci to function requires that we understand gene function in the context of cell and tissue types relevant to human disease. Expression studies can help to identify relevant cell types for functional studies, the nature of which are likely to differ for genes expressed in epithelial cells, macrophages or lymphocytes. Here we focused on five ulcerative colitis susceptibility loci containing genes whose expression and function are not well established. We performed quantitative real-time PCR in human intestinal and immune cDNA samples (**Supplementary Fig. 2a**).

The *RNF186*, *PLA2G2E* and *OTUD3* genes are located at 1p36, which harbors two distinct ulcerative colitis risk SNPs (rs1317209 and rs6426833). Although RNF186 and OTUD3 are proteins with unknown function, both contain protein domains that have been associated with protein ubiquitination. This type of modification is known to regulate immune responses, as the OTU (ovarian tumor doma) protein TNFAIP3 has been identified as an important negative regulator of the transcription factor NF-κB[15]. Notably, the *TNFAIP3* locus, which has been associated with several autoimmune conditions[16,17], is in the vicinity of one of the ulcerative colitis risk regions (rs2327832), although this locus achieved association at only a suggestive level of genome-wide significance ($P = 3.92 \times 10^{-5}$). Expression of *RNF186* was higher in intestinal tissues than in immune tissues. Immunostaining indicated that the RNF186 protein was expressed at the basal pole of epithelial cells and lamina propria within colonic tissues (**Supplementary Fig. 2a**,**b**). By contrast, *OTUD3* transcripts had higher levels in immune tissues (spleen, lymph nodes and peripheral blood mononuclear cells) and lymphocytes than in CD14-positive cells (data not shown). Phospholipase A2 group IIE (PLA2G2E) is a secretory PLA2 involved in the production of various types of proinflammatory lipid mediators[18]. *PLA2G2E* was undetectable in most tissues but showed weak expression in bone marrow, lymph node and thymus, suggesting that it has a role in immunity. Unexpectedly, we detected *PLA2G2E* very specifically in the small intestine but not in the colon. *RNF186*, *OTUD3* and *PLA2G2E* showed very different expression patterns, suggesting that investigations of the biological functions of these candidate genes will require disparate strategies.

CEP72 and *TPPP* are located close to the ulcerative colitis risk SNP rs4957048. The CEP72 and TPPP proteins are both involved in microtubule organization and are expressed ubiquitously (**Supplementary Fig. 2a**,**b**). *LAMB1* and *DLD* are located near the ulcerative colitis risk SNPs rs4598195 and rs2237686 at 7q31 and are likewise expressed ubiquitously (**Supplementary Fig. 2a**). LAMB1 is an extracellular matrix glycoprotein constituent of basement membranes. These data, together with reported associations between ulcerative colitis and *CHD1* and *CHD3* (ref. 10), further implicate defects in barrier integrity in the development of colonic inflammation.

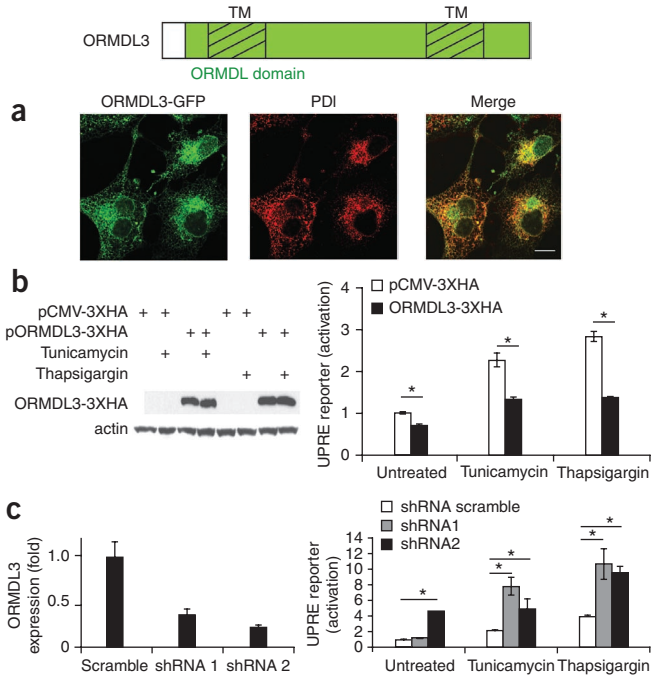

Figure 1 Expression and function of ORMDL3 (shown at top in green). (**a**) Expression of ORMDL3 protein. Cos-7 cells were transfected with pCMV-ORMDL3-GFP. After 24 h, cells were fixed and subjected to immunostaining using antibody to PDI as an ER marker. Fluorescence signals were acquired with a confocal microscope. Each image represents a single confocal section. (**b**) Effect of ORMDL3 overexpression on unfolded protein response. HEK293T cells were co-transfected with a firefly luciferase reporter p5xUPRE-GL3, *Renilla* luciferase as a transfection control and pCMV-3XHA or pCMV-ORMDL3-3XHA. After 24 h, cells were left untreated or treated using 5 μg/ml of tunicamycin or 10 μM thapsigargin for 6 h. ORMDL3-3XHA was detected with an antibody to HA. Luciferase activities were measured to detect activation of UPRE transcription. (**c**) Effect of ORMDL3 knockdown on unfolded protein response. HEK293T cells were co-transfected with plasmids encoding short hairpin RNA (shRNA) scramble or shRNA directed against ORMDL3, firefly luciferase reporter p5xUPRE-GL3 and *Renilla* luciferase as a transfection control. After 48 h, cells were left untreated or treated using 5 μg/ml of tunicamycin or 10 μM thapsigargin. *ORMDL3* expression was determined by quantitative RT-PCR (left). Luciferase activities were measured to detect activation of UPRE transcription (right). shRNA1, shRNA2, see Online Methods.

Future studies will need to integrate disease associations and the consequences of allele-specific expression to uncover the functional roles of genes in diseases such as ulcerative colitis.

GSDMB and *ORMDL3* are located near the ulcerative colitis risk SNPs rs2305480 and rs8067378. Both *ORMDL3* and *GSDMB* showed higher expression in immune tissues as compared with intestinal tissues (**Supplementary Fig. 2a**). We selected the *ORMDL3* region for functional analysis not only because of the association with ulcerative colitis presented herein but also because *ORMDL3* has been implicated in many diseases involving dysregulated immune responses, although the underlying mechanisms of this association remain unclear[2,19–21]. The ORMDL3 protein is thought to be involved in protein folding, and growing evidence indicates that there are interactions between the unfolded protein response (UPR) and immune responses[22,23]. We found that an ORMDL3-GFP fusion protein localized to the endoplasmic reticulum (ER; **Fig. 1a**), confirming previous data[24]. We next investigated whether ORMDL3 expression is involved in the UPR in epithelial cells. Overexpression of ORMDL3 decreased both the basal and ER-stress-induced UPR (**Fig. 1b**). Knockdown of ORMDL3 expression induced a higher UPR after tunicamycin or thapsigargin stimulation (**Fig. 1c**), indicating that ORMDL3 expression levels can regulate UPR and that ORMDL3 might be an important factor in ensuring ER homeostasis.

This analysis increases our understanding of the pathogenesis of ulcerative colitis and its relationship to Crohn's disease, in addition to providing expression and functional data for some of the genes implicated. The genetic associations described here, together with other data[10,12,13], highlight the importance of alterations in barrier function, cell-specific innate responses (ER stress, microbe-elicited responses, production of reactive oxygen species and activation of NF-κB), gene sets that coordinately regulate key functional programs in adaptive immunity and resolution of inflammation in the pathogenesis of ulcerative colitis. Taken together, our findings explain less than 10% of the variance of ulcerative colitis, and the challenge now is both to identify additional genetic factors and to translate these advances into real benefits for individuals with ulcerative colitis.

METHODS

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

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

ACKNOWLEDGMENTS

This study was supported in part by US National Center for Research Resources (NCRR) grant M01-RR00425 to the Cedars-Sinai General Research Center Genotyping core; US National Institutes of Health/NIDDK grant P01-DK046763; Diabetes Endocrinology Research Center grant DK063491; Cedars-Sinai Medical Center Inflammatory Bowel Disease Research Funds. Additional funding was provided by grants DK76984 (M.C.D.) and DK084554 (M.C.D. and D.P.B.M.). Cardiovascular Health Study research reported in this article was supported by contract numbers N01-HC-85079 through N01-HC-85086, N01-HC-35129, N01 HC-15103, N01 HC-55222, N01-HC-75150 and N01-HC-45133; grant numbers U01 HL080295 and R01 HL087652 from the US National Heart, Lung, and Blood Institute; and additional contribution from the US National Institute of Neurological Disorders and Stroke. A full list of principal Cardiovascular Health Study investigators and institutions can be found at http://www.chs-nhlbi.org/pi.htm. A.G. is supported by the Crohn's and Colitis Foundation of America. R.J.X. and M.J.D. are supported by grants DK83756, DK086502 and DK043351 (NIDDK).

The NIDDK IBD Genetics Consortium is funded by the following grants: DK062431 (S.R.B.), DK062422 (J.H.C.), DK062420 (R.H.D.), DK062432 (J.D.R.), DK062423 (M.S.S.), DK062413 (D.P.B.M.) and DK062429 (J.H.C.). J.H.C. is also funded by Bohmfalk Funds for Medical Research, Burroughs Wellcome Medical Foundation and the Crohn's and Colitis Foundation of America. J.D.R. is also funded by grants from the US National Institute of Allergy and Infectious Diseases (AI065687; AI067152) and from the US National Institute of Diabetes and Digestive and Kidney Diseases (DK064869).

Activities in Sweden were supported by the Swedish Society of Medicine, the Bengt Ihre Foundation, the Karolinska Institutet, the Swedish National Program for IBD Genetics, the Swedish Organization for IBD, the Swedish Medical Research Council, the Soderbergh Foundation and the Swedish Cancer Foundation. Support for genotyping and genetic data analysis was provided by the Singapore National Cancer Centre, Singapore General Hospital and the Singapore Millennium Foundation (to S.P.) and the Agency for Science Technology and Research (A*STAR), Singapore (to M.L.H. and M.S.). Genotyping and DNA handling at the Genome Institute of Singapore were performed by W.Y. Meah, K.K. Heng, H.B. Toh, X. Lin, S. Rajaram, D. Tan and C.H. Wong. We are grateful to the funders and investigators of the Epidemiological Investigation of Rheumatoid Arthritis for providing genotype data from healthy Swedish individuals.

AUTHOR CONTRIBUTIONS

D.P.B.M., M.J.D., R.J.X., J.D.R., J.H.C., P.G., R.H.D. and M.S. participated in the study design and conception. D.P.B.M., A.G., M.J.D., R.J.X., J.D.R. and M.S. wrote the manuscript with contributions from R.H.D. and J.I.R. D.P.B.M., L.T., K.D.T., C.Li, C.B., P.R.F., M.C., M.D., J.H., M.L.H., M.L., L.P., A.A., E.C., A.L., O.P.,

E.-J.B., C.D., D.W.H., D.J.d.J., P.C.S., R.K.W., Y.S., M.S.S., J.H.C., S.R.B., L.P.S., R.H.D., M.C.D., N.L.G., T.H., A.I., G.Y.M., D.S.S., E.A.V., S.R.T., V.A., C.W. and S.P. performed patient diagnosis, patient enrollment and collection of clinical data. Replication genotyping was performed by C. Lagacé, C.R.S. and C.B. in the laboratory of J.D.R. Expression analysis, immunohistochemistry and shRNA studies were designed by A.G. and R.J.X. and performed by C. Li and A.G. M.J.D., J.E., B.M.N., K.R., J.W., J.D.R., P.G., T.G. and R.T.H.O. provided statistical analyses. All authors contributed to the final paper.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

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

1. Silverberg, M.S. *et al.* Ulcerative colitis-risk loci on chromosomes 1p36 and 12q15 found by genome-wide association study. *Nat. Genet.* **41**, 216–220 (2009).
2. Barrett, J.C. *et al.* Genome-wide association defines more than 30 distinct susceptibility loci for Crohn's disease. *Nat. Genet.* **40**, 955–962 (2008).
3. 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).
4. 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).
5. Franke, A. *et al.* Sequence variants in *IL10, ARPC2* and multiple other loci contribute to ulcerative colitis susceptibility. *Nat. Genet.* **40**, 1319–1323 (2008).
6. Zhernakova, A. *et al.* Genetic analysis of innate immunity in Crohn's disease and ulcerative colitis identifies two susceptibility loci harboring *CARD9* and *IL18RAP*. *Am. J. Hum. Genet.* **82**, 1202–1210 (2008).
7. de Bakker, P.I. *et al.* Practical aspects of imputation-driven meta-analysis of genome-wide association studies. *Hum. Mol. Genet.* **17**, R122–R128 (2008).
8. Franke, A. *et al.* Systematic association mapping identifies *NELL1* as a novel IBD disease gene. *PLoS One* **2**, e691 (2007).
9. Goyette, P. *et al.* Gene-centric association mapping of chromosome 3p implicates *MST1* in IBD pathogenesis. *Mucosal Immunol.* **1**, 131–138 (2008).
10. Barrett, J.C. *et al.* Genome-wide association study of ulcerative colitis identifies three new susceptibility loci, including the *HNF4A* region. *Nat. Genet.* **41**, 1330–1334 (2009).
11. Kugathasan, S. *et al.* Loci on 20q13 and 21q22 are associated with pediatric-onset inflammatory bowel disease. *Nat. Genet.* **40**, 1211–1215 (2008).
12. Asano, K. *et al.* A genome-wide association study identifies three new susceptibility loci for ulcerative colitis in the Japanese population. *Nat. Genet.* **41**, 1325–1329 (2009).
13. Imielinski, M. *et al.* Common variants at five new loci associated with early-onset inflammatory bowel disease. *Nat. Genet.* **41**, 1335–1340 (2009).
14. McGovern, D.P. *et al.* Genetic epistasis of IL23/IL17 pathway genes in Crohn's disease. *Inflamm. Bowel Dis.* **15**, 883–889 (2009).
15. Sun, S.C. Deubiquitylation and regulation of the immune response. *Nat. Rev. Immunol.* **8**, 501–511 (2008).
16. Graham, R.R. *et al.* Genetic variants near *TNFAIP3* on 6q23 are associated with systemic lupus erythematosus. *Nat. Genet.* **40**, 1059–1061 (2008).
17. Trynka, G. *et al.* Coeliac disease-associated risk variants in *TNFAIP3* and *REL* implicate altered NF-κB signalling. *Gut* **58**, 1078–1083 (2009).
18. Burke, J.E. & Dennis, E.A. Phospholipase A2 structure/function, mechanism, and signaling. *J. Lipid Res.* **50** Suppl, S237–S242 (2009).
19. Hirschfield, G.M. *et al.* Primary biliary cirrhosis associated with *HLA, IL12A*, and *IL12RB2* variants. *N. Engl. J. Med.* **360**, 2544–2555 (2009).
20. Verlaan, D.J. *et al.* Allele-specific chromatin remodeling in the *ZPBP2/GSDMB/ORMDL3* locus associated with the risk of asthma and autoimmune disease. *Am. J. Hum. Genet.* **85**, 377–393 (2009).
21. Moffatt, M.F. *et al.* Genetic variants regulating *ORMDL3* expression contribute to the risk of childhood asthma. *Nature* **448**, 470–473 (2007).
22. Kitamura, M. Biphasic, bidirectional regulation of NF-κB by endoplasmic reticulum stress. *Antioxid. Redox Signal.* **11**, 2353–2364 (2009).
23. Todd, D.J., Lee, A.H. & Glimcher, L.H. The endoplasmic reticulum stress response in immunity and autoimmunity. *Nat. Rev. Immunol.* **8**, 663–674 (2008).
24. Hjelmqvist, L. *et al.* ORMDL proteins are a conserved new family of endoplasmic reticulum membrane proteins. *Genome Biol.* **3**, RESEARCH0027 (2002).

[1]Inflammatory Bowel and Immunobiology Research Institute, Cedars-Sinai Medical Center, Los Angeles, California, USA. [2]Center for Computational and Integrative Biology and Gastrointestinal Unit, Massachusetts General Hospital, Harvard Medical School, Boston, Massachusetts, USA. [3]Department for Clinical Science Intervention and Technology, Karolinska Institutet and IBD Clinical Research Group at Karolinska University Hospital, Stockholm, Sweden. [4]Université de Montréal and the Montreal Heart Institute, Research Center, Montréal, Québec, Canada. [5]Center for Human Genetic Research, Massachusetts General Hospital, Harvard Medical School, Boston, Massachusetts, USA. [6]Medical Genetics Institute, Cedars-Sinai Medical Center, Los Angeles, California, USA. [7]Human Genetics, Genome Institute of Singapore, Singapore. [8]The Broad Institute of MIT and Harvard, Cambridge, Massachusetts, USA. [9]Department of Medical Sciences, Gastroenterology Research Group, Uppsala University Hospital, Uppsala, Sweden. [10]Department of Biosciences and Nutrition, Karolinska Institutet, Stockholm, Sweden. [11]Division of Gastroenterology, Department of Internal Medicine, Örebro University Hospital, Örebro, Sweden. [12]Infectious Diseases, Genome Institute of Singapore, Singapore. [13]Department of Medicine, and IBD Clinical Research Group at Karolinska University Hospital, Stockholm, Sweden. [14]Rheumatology Unit, Department of Medicine, Karolinska Institutet at Karolinska University Hospital Solna, Stockholm, Sweden. [15]Gastroenterologia ed Endoscopia Digestiva, Ospedale 'Casa Sollievo della Sofferenza', Istituto di Ricovero e Cura a Carattere Scientifico, San Giovanni Rotondo, Italy. [16]Université de Montréal and Centre Hospitalier Universitaire de l'Université de Montréal, Montréal, Québec, Canada. [17]Department of Gastroenterology, Hôpital Sainte-Justine, Montréal, Québec, Canada. [18]Department of Gastroenterology and Hepatology, Leiden University Medical Center, Leiden, The Netherlands. [19]Department of Gastroenterology and Hepatology, Radboud University Nijmegen Medical Centre, Nijmegen, The Netherlands. [20]Department of Gastroenterology and Hepatology, Academic Medical Center, Amsterdam, The Netherlands. [21]Department of Gastroenterology and Hepatology, University Medical Center Groningen, University of Groningen, Groningen, The Netherlands. [22]National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK). [23]Section of Digestive Diseases, Department of Medicine, Yale University, New Haven, Connecticut, USA. [24]Mount Sinai Hospital Inflammatory Bowel Disease Group, University of Toronto, Toronto, Ontario, Canada. [25]Department of Genetics, Yale University, New Haven, Connecticut, USA. [26]Department of Statistics, Carnegie Mellon University, Pittsburgh, Pennsylvania, USA. [27]Johns Hopkins University School of Medicine, Department of Medicine, and Johns Hopkins University Bloomberg School of Public Health, Department of Epidemiology, Baltimore, Maryland, USA. [28]Department of Health Studies, University of Chicago, Chicago, Illinois, USA. [29]Division of Gastroenterology, Hepatology and Nutrition, Department of Medicine, University of Pittsburgh School of Medicine, and Department of Human Genetics, Graduate School of Public Health, University of Pittsburgh, Pittsburgh, Pennsylvania, USA. [30]Cardiovascular Health Research Unit, Departments of Epidemiology and General Medicine, University of Washington, Seattle, Washington, USA. [31]Department of Genetics, University Medical Center Groningen and Groningen University, Groningen, The Netherlands. [32]Department of Microbiology Tumor and Cell Biology, Karolinska Institutet, Stockholm, Sweden. [33]Laboratory of Inflammation Biology, Singapore General Hospital, Singapore. [34]Program in Molecular and Genetic Epidemiology, Harvard School of Public Health, Boston, Massachusetts, USA. [35]Institute for Human Genetics, University of California San Fransisco, San Francisco, California, USA. Correspondence should be addressed to D.P.B.M. (mcgovernd@cshs.org), J.D.R. (john.david.rioux@umontreal.ca) or M.S. (mark.seielstad@ucsf.edu).

ONLINE METHODS

Study design. Two unpublished ulcerative colitis GWA studies were combined with the published NIDDK ulcerative colitis GWA study[1] (**Supplementary Table 1**).

Cedars-Sinai ulcerative colitis GWA study. The Cedars-Sinai ulcerative colitis scan consisted of 852 cases and 3,271 controls. Samples with >1% missing genotype data (1 case, 278 controls) were removed. Standard searches for relatedness (using PLINK –genome) and population structure (using PLINK –mds)[25] were performed. Identity-by-descent of >20% (detecting half-sibs and above) was identified in 81 pairs (primarily controls), and one member from each pair was removed. The first ten principal components were detected. The first principal component represented the expected Northern Europe–Southern Europe–Ashkenazi axis. The second principal component identified outliers that we removed because they were not readily matched in a case-control sense. For the final analysis of 723 cases and 2,880 controls, we performed a logistic regression analysis correcting for multidimensional scaling covariates 1 and 3–10. This quality-control process reduced the value of λ_{GC} from ~2.5 to 1.074. This residual inflation was genomic-control corrected before inclusion in the meta-analysis.

Swedish ulcerative colitis GWA study. Individuals with ulcerative colitis were enrolled at three sites in Sweden. All genotyping was performed on an Illumina Hap550 array and Illumina Quad-610 BeadChips at the Genome Institute of Singapore. Genotype data from 640 control individuals, previously checked for quality and free of inflammatory disease, were combined with data from common controls from the Epidemiological Investigation of Rheumatoid Arthritis[26] genotyped on Illumina Hap550 (n = 460) and Quad-610 (n = 378) BeadChips (a total of 1,478 controls). Alleles were called in the Illumina BeadStudio software by reclustering, and cases and controls were included on each chip type. Sample call rates exceeded 97.29%, and no samples were removed for this reason. We removed from further analysis four samples with low heterozygosity. Fifteen samples (10 cases and 5 controls) were excluded on the basis of sex discrepancies between database records and X chromosome zygosity. We identified 112 (106 case and 6 control) sample duplicates and first-degree relatives by excess allele sharing, as calculated in PLINK, and subsequently excluded them. We identified 138 population outliers (82 cases and 56 controls) by a principal components analysis implemented in EIGENSTRAT.

SNPs with a call rate of <95%, a minor allele frequency <0.005 and a Hardy-Weinberg equilibrium of $P < 10^{-7}$ in controls, and nonautosomal SNPs were removed. Lastly, a common set of directly genotyped SNPs across all three chip types was used; this procedure resulted in a data set of 948 cases and 1,408 controls for 297,031 SNPs after quality control. Trend tests of association were calculated in PLINK. The λ_{GC} value for the whole data set was 1.04, uncorrected. EIGENSTRAT correction based on the top ten PCs (bringing λ_{GC} to 1.03), followed by genomic control correction, was used to remove the small residual stratification.

Meta-analysis. Because all studies used compatible Illumina platforms, we combined the 266,047 (258,137 autosomal and 7,910 X-chromosomal) SNPs passing quality control in all three studies. P values from the population structure–corrected analyses were converted to Z scores consistently oriented to the combined minor allele. These consistently oriented Z scores were then combined, and the square of these scores was evaluated by comparison with a χ^2 distribution (see **Supplementary Fig. 1** for the Q–Q plot). Although overall there was only very modest inflation (λ_{GC} = 1.036), there was a significant excess of true positive results in the tail of the distribution. There were 59, 126 and 511 SNPs from the 266,047 SNPs that exceeded P = 0.00001, 0.0001 and 0.001, respectively (~20-, 5- and 2-fold in excess of null expectation).

Replication cohorts. Two independent case-control cohorts were examined for the replication phase: an Italian study population composed of 1,094 individuals with ulcerative colitis and 908 control subjects enrolled at the S. Giovanni Rotondo 'CSS' (SGRC) Hospital in Italy; and a Dutch study population composed of 1,090 individuals with ulcerative colitis of Caucasian ethnicity recruited through the Inflammatory Bowel Disease unit of the University Medical Center Groningen, Groningen; the Academic Medical Center, Amsterdam; the Leiden University Medical Center, Leiden; and the Radboud University Medical Center, Nijmegen. Healthy controls (n = 804) of self-declared European ancestry were drawn from volunteers at the University Medical Center, Utrecht. Ulcerative colitis was diagnosed according to accepted clinical, endoscopic, radiological and histological findings. All patients and control subjects gave informed consent, and the study was approved by the ethics review committee of each participating hospital. All DNA samples and data in this study were handled anonymously.

Replication genotyping and quality control. Selected SNPs were designed into multiplex assays and genotyped by using primer extension chemistry and mass spectrometric analysis (iPlex assay, Sequenom) on the Sequenom MassArray at the Laboratory for Genetics and Genomic Medicine of Inflammation of the Université de Montréal. When SNPs could not be designed into multiplex assays or failed the quality-control thresholds, a proxy was selected. Samples showing >10% missing data, in addition to SNPs that had >10% missing data or were significantly out of Hardy-Weinberg equilibrium ($P < 0.001$), were excluded from the analyses. The overall genotyping call rate in the replication data set following quality control was >99% and consisted of 993 cases and 826 controls (Italian) and 1,016 cases and 754 controls (Dutch).

Association analysis of the replication phase. Association testing of single SNPs in the replication cohorts was done by a standard χ^2 test carried out on a 2 × 2 contingency table (PLINK). Combination of the results from the two replication cohorts, and combination of the results from the screen and replication, was achieved by calculating a combined weighted Z score. The threshold for significant independent replication was set at $P < 0.05$ in the combined Italian and Dutch data sets.

Antibodies. We used the following antibodies: anti-RNF186 (Abnova), anti-CEP72 (Novus Biological NB100-60661), anti–Troma-1 (Developmental Studies Hybridoma Bank) and anti-PDI (Stressgen).

Reverse transcription and real-time PCR. For expression maps, immune and intestinal cDNA samples were obtained from Clontech (two independent panels of cDNA). For ORMDL3 functional studies, RNA extraction was performed by using an RNeasy kit (Qiagen) in accordance with the manufacturer's instructions. One microgram of total RNA was reverse-transcribed using an iScript cDNA synthesis kit (Bio-Rad). The gene expression reported is representative of three independent experiments. Real-time quantitative PCR was performed in duplicate in a Bio-Rad iCycler thermal cycler equipped with an iQ5 optical module using the iQ SYBR Green Super Mix (Bio-Rad). In brief, 100 ng of reverse-transcribed cDNA was used for each PCR with forward and reverse primers at 250 nM. The thermal cycling conditions were 4 min at 95 °C, followed by 40 cycles at 94 °C for 15 s and 59 °C for 1 min. Values were normalized to that of *GAPDH*, and the condition containing the lowest mRNA content was defined as 1 arbitrary unit. All PCR products were analyzed on a 2% agarose gel to verify the correct size of the amplicons.

Intestinal biopsies. After obtaining informed consent and approval from the Human Research Committee of the Massachusetts General Hospital, colonic biopsies from uninflamed regions were obtained from individuals with IBD undergoing colonoscopy. Tissues were embedded and frozen in OCT compound (Fisher).

Immunostaining and fluorescence microscopy. Cells or frozen tissue sections were fixed with 4% paraformaldehyde and permeabilized with 0.1% PBS-Triton X-100. After being washed with PBS, the sections were incubated for 30 min in PBS containing 3 M glycine to block the reactive groups of paraformaldehyde. The sections were incubated (1 h) with a blocking solution containing 10% donkey serum (Rockland Immunochemicals) and 10% human Fc block reagent (Miltenyi Biotec). The preparations were then incubated with primary antibody for 1 h, washed with PBS, incubated with fluorescent secondary antibody (Jackson Immunoresearch) for 1 h, washed with PBS and incubated with PBS containing 100 µg/ml of DABCO (Sigma) as an antifading reagent before being mounted in Glycergel medium (Dako). Fluorescence

signals were captured with a Radiance 2000 laser confocal microscope (Bio-Rad). Image acquisition was performed with LaserSharpScanning software (Bio-Rad).

Plasmids. To engineer *ORMDL3* constructs for mammalian expression, a clone of *ORMDL3* in pCMV-SPORT6 was obtained from Open Biosystems. The *ORMDL3* coding sequence was amplified by PCR using forward and reverse primers (**Supplementary Table 5**) containing the enzyme restriction sites EcoRI and NotI, respectively. After EcoRI and NotI digestion, the *ORMDL3* coding sequence was subcloned into a C-terminally tagged pCMV-3xHA vector derived from the pCMV-Myc vector (Clontech, catalog no. 631604), and into a pcDNA4/TO-GFP-C vector, which was derived by inserting an EGFP sequence from the XhoI and ApaI sites into a pcDNATM4/TO vector (Invitrogen).

For shRNA studies, shRNA hairpins directed against the human transcript of *ORMDL3* were designed by using the tools of the RNAi consortium from the Broad Institute. The pairs of oligonucleotides purchased to construct the hairpins are listed in **Supplementary Table 5**. After annealing, and according to RNAi consortium instructions, we inserted the paired oligonucleotides by using AgeI and EcoRI enzyme restriction sites in the TRC22 vector, derived from pLKO1.

ER-stress reporter. HEK293T cells were plated on 24-well plates at a density of 2×10^5 cells per well. After 24 h, the cells of each well were transfected with 2 ng of the ER-stress firefly luciferase reporters p5xUPRE-GL3 (ref. 27) and 0.025 ng of *Renilla* luciferase (Promega) by using TransFectin Lipid Reagent (Bio-Rad). Luciferase activities were measured by the Dual-Luciferase Reporter Assay System (Promega) and were normalized to the internal transfection control of *Renilla* luciferase activity.

URLs. Laboratory for Genetics and Genomic Medicine of Inflammation, http://www.inflammgen.org/; RNAi consortium from the Broad Institute, http://www.broadinstitute.org/rnai/trc/lib/.

25. Purcell, S. *et al.* PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* **81**, 559–575 (2007).
26. Plenge, R.M. *et al.* TRAF1-C5 is a risk locus for rheumatiod arthritis: a genomewide study. *N. Engl. J. Med.* **357**, 1199–1209 (2007).
27. Wang, Y. *et al.* Activation of ATF6 and an ATF6 DNA binding site by the endoplasmic reticulum stress response. *J. Biol. Chem.* **275**, 27013–27020 (2000).

doi:10.1038/ng.549