Published in: Nature Genetics (2009)
Status: Postprint (Author's version)

Common variants in the *NLRP3* region contribute to Crohn's disease susceptibility

Alexandra-Chloé Villain[1,2], Mathieu Lemire[3], Geneviève Fortin[1], Edouard Louis[4], Mark S Silverberg[5], Catherine Collette[1], Nobuyasu Baba[6], Cécile Libioulle[4], Jacques Belaiche[4], Alain Bitton[7], Daniel Gaudet[8], Albert Cohen[9], Diane Langelier[10], Paul R Fortin[11], Joan E Wither[11], Marika Sarfati[6], Paul Rutgeerts[12], John D Rioux[13], Séverine Vermeire[12], Thomas J Hudson[3] & Denis Franchimont[1,14]

[1] *Division of Gastroenterology, McGill University Health Centre, Montreal General Hospital, 1650 Cedar Avenue, Montréal, Québec H3G 1A4, Canada.*
[2] *McGill University and Génome Québec Innovation Centre, 740 Doctor Penfield Avenue, Montréal, Québec H3A 1A4, Canada.*
[3] *Ontario Institute for Cancer Research, MaRS Centre, South Tower, 101 College Street, Suite 800, Toronto, Ontario M5G 0A3, Canada.*
[4] *Division of Gastroenterology, CHU, Université de Liège, 13 avenue de l'Hôpital Bât. B35, 4000 Liège, Belgium.*
[5] *Mount Sinai Hospital IBD Centre, University of Toronto, 441-600 University Avenue, Toronto, Ontario M5G 1X5, Canada.*
[6] *lmmunoregulation Laboratory, CHUM Research Center, Université de Montréal, 1560 rue Sherbrooke Est, Montréal, Québec H2L 4M1, Canada.*
[7] *Division of Gastroenterology, McGill University Health Centre, Royal Victoria Hospital, 687 Pine Avenue West, Montréal, Québec H3A 1A1, Canada.*
[8] *Université de Montréal, Centre de médecine génique communautaire de l'Université de Montréal, Hôpital universitaire de Chicoutimi, 555 boulevard de l'Université, Chicoutimi, Québec G7H 2B1, Canada.*
[9] *Jewish General Hospital, 3755 chemin de la Côte-Sainte-Catherine, Montréal, Québec H3T 1E2, Canada.*
[10] *Centre Hospitalier Universitaire de Sherbrooke-Hôtel-Dieu, 580 rue Bowen Sud, Sherbrooke, Québec J1G 2E8, Canada.*
[11] *Arthritis Centre of Excellence, Toronto Western Hospital Research Institute, 399 Bathurst Street, University of Toronto, Toronto, Ontario M5T 2S8, Canada.*
[12] *Department of Gastroenterology, University Hospital Gasthuisberg, 49 Herestraat, 3000 Leuven, Belgium.*
[13] *Université de Montréal and Montréal Heart Institute, Research Center, 5000 rue Bélanger, Montréal, Québec HIT 1C8, Canada.*
[14] *Department of Gastroenterology, Erasme Hospital, ULB, 808 Lennik road, 1070 Brussels, Belgium.*

Abstract

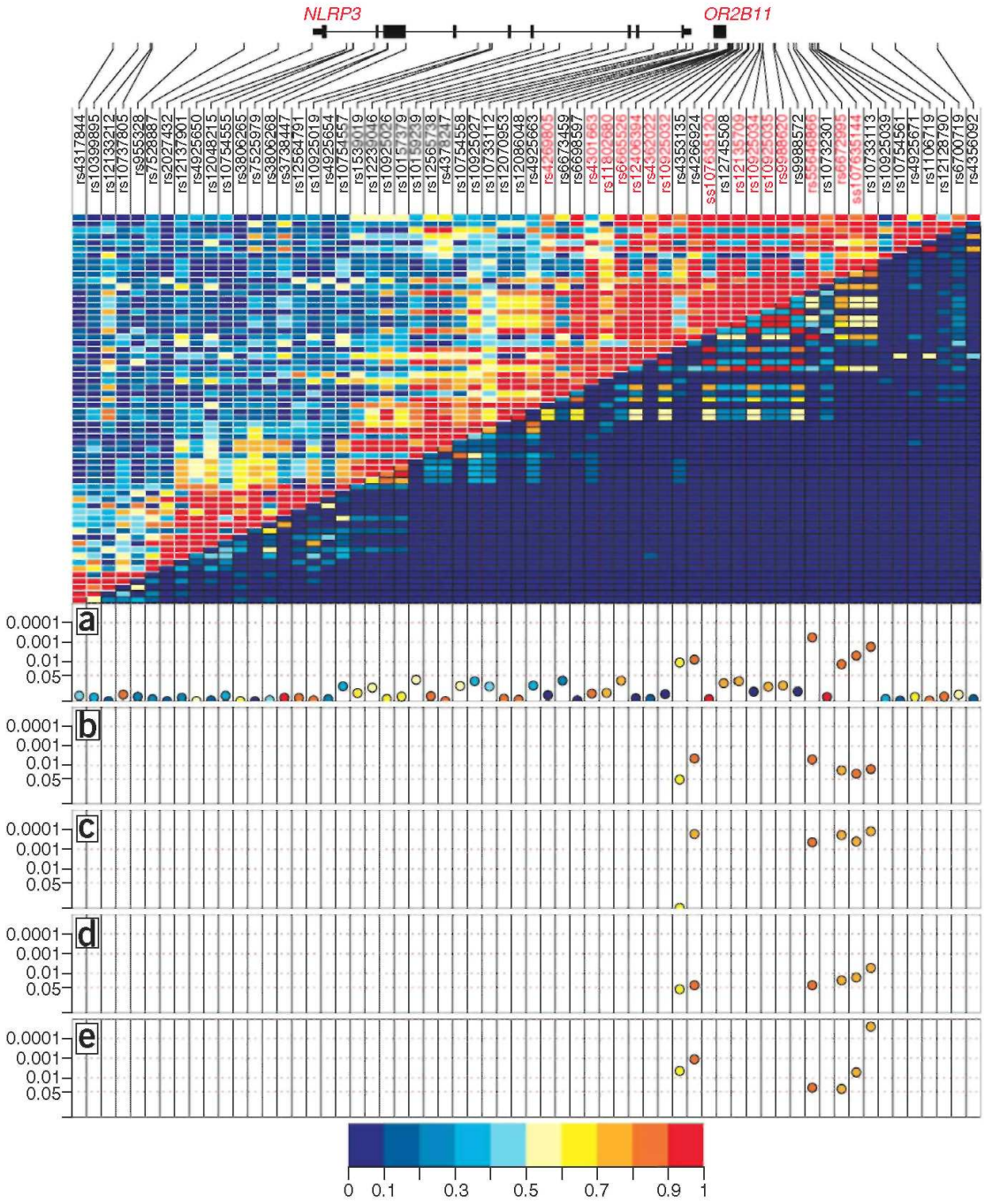
We used a candidate gene approach to identify a set of SNPs, located in a predicted regulatory region on chromosome 1q44 downstream of *NLRP3* (previously known as *CIAS1* and *NALP3)* that are associated with Crohn's disease. The associations were consistently replicated in four sample sets from individuals of European descent. In the combined analysis of all samples (710 father-mother-child trios, 239 cases and 107 controls), these SNPs were strongly associated with risk of Crohn's disease ($P_{combined} = 3.49 \times 10^{-9}$, odds ratio = 1.78, confidence interval = 1.47-2.16 for rs10733113), reaching a level consistent with the stringent significance thresholds imposed by whole-genome association studies. In addition, we observed significant associations between SNPs in the associated regions and *NLRP3* expression and IL-1β production. Mutations in *NLRP3* are known to be responsible for three rare autoinflammatory disorders[1,2]. These results suggest that the *NLRP3* region is also implicated in the susceptibility of more common inflammatory diseases such as Crohn's disease.

Crohn's disease and ulcerative colitis are multigenic and heterogeneous inflammatory bowel diseases of the gastrointestinal tract that seem to result from a dysregulated mucosal immune response to bacterial antigens in the gut lumen of a genetically susceptible host[3]. *NLRP3* is a member of the CATERPILLER[4] family of genes encoding proteins that comprise a nucleotide-binding domain and a leucine-rich repeat domain. Cryopyrin, the protein encoded by *NLRP3*, controls the inflammasome, a crucial molecular platform that regulates activation of caspase-1 and processing of interleukin (IL)-1β—two key mediators of inflammation[2,5,6]. The potential involvement of *NLRP3* in the pathogenesis of more common inflammatory disorders motivated us to conduct an in-depth genetic analysis of the *NLRP3* region.

We first assessed the association between 47 SNPs in the *NLRP3* region and Crohn's disease risk in 296 trios from Leuven University Hospital Gasthuisberg (see Methods, Supplementary Methods and Supplementary Table 1 online). The major alleles of three tagging SNPs were significantly associated with increased risk of Crohn's disease in the Leuven sample set ($P = 0.0107$ for rs4353135; $P = 7.63 \times 10^{-3}$ for rs4266924; $P = 1.68 \times 10^{-3}$ for rs10733113; Fig. 1 and Table 1). These SNPs span a 5.3-kb region and are located 4.7 kb downstream of *NLRP3*

and 1.85 kb upstream of the olfactory receptor gene *OR2B11* (Fig. 1). No associations were observed with SNPs within the *NLRP3* gene, and none of the associated SNPs were in linkage disequilibrium with tagging SNPs located in *NLRP3* or *OR2B11*. These preliminary associations pointed to the 5.3-kb region near *NLRP3*, bounded by rs4353135 and rs10733113, as a candidate region contributing to Crohn's disease susceptibility.

Figure 1 *Association results for the five Crohn's disease sample sets. Top panel shows SNPs, their positions in the genes and the linkage disequilibrium structure between them. SNP names in red were genotyped in the second phase of the study, subsequent to the sequencing experiment. Middle panel shows D' in the upper left and* r^2 *in the lower right, (a-e) Lower panels show results from association analysis of Leuven trios (a), Liège trios (b), Liège case-control cohort (c), Québec trios (d) and Toronto trios (e). Pvalues for individual alleles are reported in a logarithmic scale on the y axis. Color spectrum represents strength of linkage disequilibrium and frequency of the associated alleles.*



In the second phase of the study, we examined the association of the above three SNPs with risk of Crohn's disease in additional samples. Overall, we screened one case-control cohort (Liège) and three familial sample sets (Liège, Québec and Toronto; see Methods and Supplementary Methods). We replicated our initial significant Crohn's disease associations with the major alleles of tagging SNPs rs4266924 and rsl0733113 in all four sample sets ($P < 0.05$; Fig. 1 and Table 1). For the Québec and Toronto cohorts, the observed associations remained whether the analysis was done with or without samples from individuals of Jewish ancestry (data not shown). Association with rs4353135 was replicated only in the Toronto sample set ($P < 0.05$; Fig. 1 and Table 1). Combined analysis of all Crohn's disease samples revealed strong associations for rs4353135 ($P_{combined} = 8.36 \times 10^{-3}$, odds ratio = 1.21, confidence interval = 1.05-1.39, T allele frequency = 71% in cases and 65% in controls), rs4266924 ($P_{combined} = 6.01 \times 10^{-7}$, odds ratio = 1.69, confidence interval = 1.37-2.07, A allele frequency = 91% in cases and 84% in controls) and rs10733113 ($P_{combined} = 3.49 \times 10^{-9}$, odds ratio = 1.78, confidence interval = 1.47-2.16, G allele frequency = 90% in cases and 80% in controls).

This regional association was the first step in localizing the most likely causal variant(s). Because the observed association signal was not in linkage disequilibrium with any genotyped variant within *NLRP3,* we resequenced a 9-kb region extending from the *NLRP3* 3' UTR to the 5.3-kb region described above, inclusively, which also comprised *OR2B11* (Supplementary Table 2 online). Overall, we selected 16 Crohn's disease samples and 8 controls based on genotypes at markers rs4266924 and rs10733113 to fully define the linkage disequilibrium within the region and identify all polymorphisms in linkage disequilibrium with the associated SNPs. The resequencing effort identified 79 SNPs, 8 of which were previously genotyped. Among these SNPs, 14 were novel compared to dbSNP release 129, and 60 had a minor allele frequency > 0.05 (Supplementary Table 3 online).

We next conducted comprehensive genotyping to identify polymorphisms with stronger associations (Supplementary Methods). We genotyped a total of 24 SNPs in the Leuven exploratory Crohn's disease trios (Supplementary Table 3) and analyzed 15 of them after they passed quality control tests (see Methods). None of the SNPs within *OR2B11* (Supplementary Table 3) were associated with Crohn's disease. The major alleles of three SNPs, spanning a 1.8-kb region bounded by rs4353135 and rs10733113, were associated with Crohn's disease (Fig. 1 and Table 1). According to the computational method ESPERR[7], two of these SNPs (rs6672995 and rs55646866) are located in a predicted regulatory region. The third variant, SS107635144, was selected by the tagging algorithm[8]. These SNPs were in high linkage disequilibrium ($r^2 > 0.70$) with rs4266924 and rs10733113. None of these three SNPs were tags for the SNPs that were not genotyped in the region.

We subsequently screened these three variants in the other four sample sets, and their associations were consistently replicated (Fig. 1 and Table 1). Combined analysis of all Crohn's disease samples revealed strong associations for rs55646866 ($P_{combined} = 7.2 \times 10^{-7}$, odds ratio = 1.69, confidence interval = 1.38-2.08, C allele frequency = 91% in cases and 84% in controls), rs6672995 ($P_{combined} = 2.91 \times 10^{-6}$, odds ratio = 1.53, confidence interval = 1.28-1.82, G allele frequency = 87% in cases and 79% in controls) and SS107635144 ($P_{combined} = 8.50 \times 10^{-6}$, odds ratio = 1.53, confidence interval = 1.27-1.84, C allele frequency = 88% in cases and 80% in controls).

Conditioning on one of these associated SNPs to evaluate the residual significance of the others did not provide evidence of additive effects of the associated SNPs on the risk. Rather, the association they all showed with Crohn's disease can be explained solely by linkage disequilibrium. To account for multiple testing issues, we note that if all 62 SNPs had been genotyped in the Leuven exploratory sample in the same study phase, then rs55646866 would have reached a significance level ($P = 5.72 \times 10^{-4}$) that, after correction using a permutation procedure, would have still been significant ($P_{corrected} = 0.019$, estimated from 5,000 replicates; see Methods). This result was further strengthened by the consistent replication of rs55646866 in all studied samples.

We next evaluated associations between genotype and gene expression to determine whether the above SNP associations reflected exacting regulatory effects on *NLRP3*. We first assessed the influence of the six SNPs (Table 1) on *NLRP3* mRNA expression from freshly isolated peripheral blood cells (PBCs) and from monocytes isolated from the peripheral blood, as *NLRP3* is primarily expressed in granulocytes and monocytes[9] (Supplementary Methods). A significant association was observed between *NLRP3* expression and rs4353135 genotypes in PBCs ($P = 0.00246$; Fig. 2a) and monocytes ($P = 0.0124$; Fig. 2b), with homozygosity for the risk allele being associated with the lowest level of *NLRP3* expression. Results for the five other Crohn's disease-associated SNPs are shown in Supplementary Figure 1 and Table 4 online.

Published in: Nature Genetics (2009)
Status: Postprint (Author's version)

Table 1 *Replication studies and pooled analysis showing association between SNPs in the 5.3-kb region and risk of Crohn's disease*

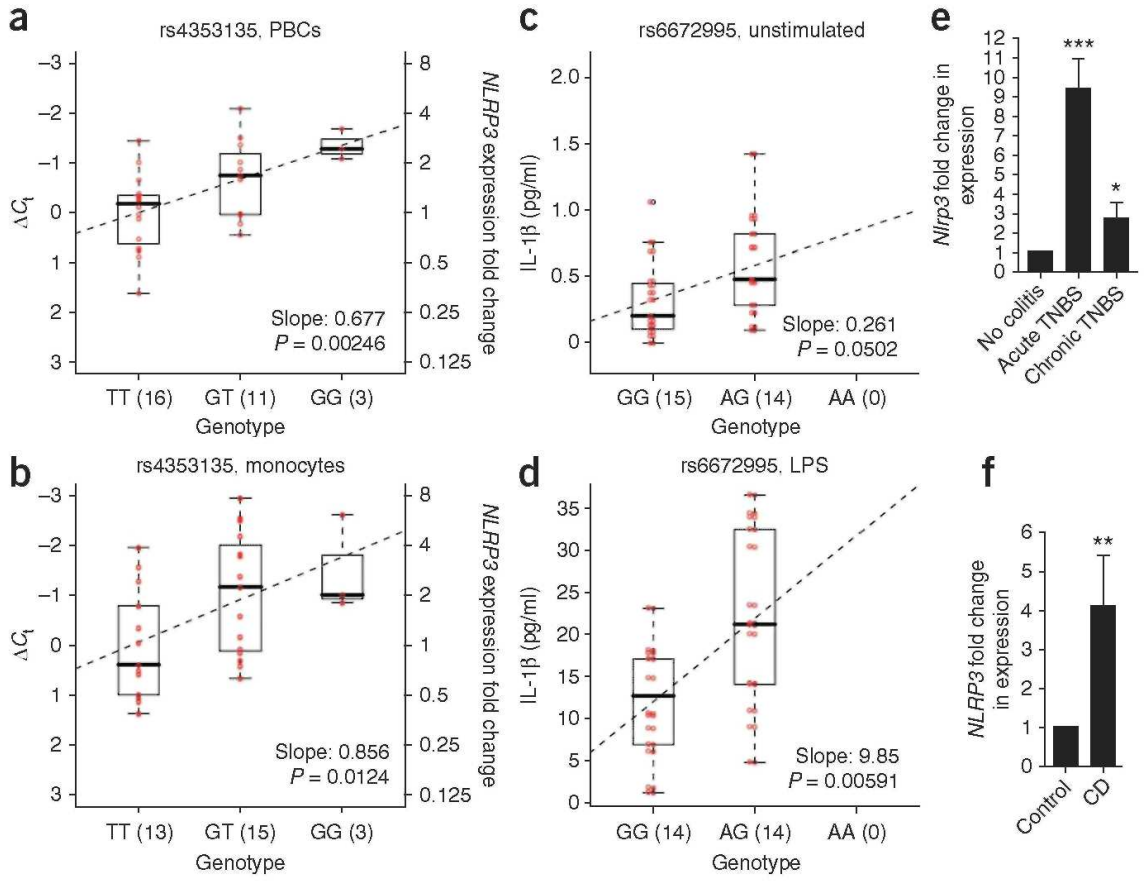
SNP	Associated allele	Control frequency[a]	Case frequency[b]	T:U ratio[c]	Control ratio[d]	Case ratio[d]	Odds ratio (95% Cl)[e]	*P* value[e]
Leuven (296 trios)								
rs4353135[f]	T	0.65	0.73	118:82			1.44 (1.08-1.91)	0.0107
rs4266924[f]	A	0.86	0.91	57:31			1.78 (1.14-2.79)	7.63×10^{-3}
rs55646866[g]	C	0.86	0.93	57:26			2.19 (1.37-3.52)	5.72×10^{-4}
rs6672995[g]	G	0.83	0.88	64:39			1.64 (1.09-2.48)	0.0133
ss107635144[g]	C	0.84	0.91	64:36			1.78 (1.16-2.73)	4.82×10^{-3}
rs10733113[f]	G	0.84	0.91	65:34			1.91 (1.24-2.95)	1.68×10^{-3}
Liège trios (155 trios)								
rs4353135[f]	T	0.62	0.73	33:20			1.65 (0.93-2.92)	0.0727
rs4266924[f]	A	0.84	0.94	22:7			3.14 (1.36-7.26)	4.31×10^{-3}
rs55646866[g]	C	0.85	0.95	20:6			3.33 (1.36-8.18)	4.80×10^{-3}
rs6672995[g]	G	0.79	0.89	31:15			2.07 (1.11-3.84)	0.0171
ss107635144[g]	C	0.80	0.90	27:13			2.08 (1.03-4.18)	0.0253
rs10733113[f]	G	0.83	0.92	24:10			2.40 (1.16-4.96)	0.0148
Liège CC (239 Crohn's disease and 107 controls)								
rs4353135[f]	T	0.65	0.65		39:44:11	92:95:27	1.01 (0.71-1.46)	0.944
rs4266924[f]	A	0.79	0.90		63:35:4	184:41:2	2.43 (1.53-3.85)	1.62×10^{-4}
rs55646866[g]	C	0.80	0.91		66:32:4	191:38:3	2.33 (1.47-3.69)	4.32×10^{-4}
rs6672995[g]	G	0.74	0.86		57:40:7	172:56:4	2.19 (1.46-3.30)	1.87×10^{-4}
ss107635144[g]	C	0.75	0.86		57:38:7	165:56:3	2.13 (1.41-3.23)	3.94×10^{-4}
rsl0733113[f]	G	0.76	0.88		58:38:5	177:47:3	2.36 (1.52-3.65)	1.20×10^{-4}
Québec (130 trios)[h]								
rs4353135[f]	T	0.67	0.73	68:48			1.42 (1.00-2.01)	0.0627
rs4266924[f]	A	0.85	0.91	43:26			1.65 (1.02-2.70)	0.0400
rs55646866[g]	C	0.84	0.90	43:26			1.65 (1.02-2.70)	0.0400
rs6672995[g]	G	0.78	0.85	53:32			1.66 (1.09-2.52)	0.0220
ss107635144[g]	C	0.80	0.88	49:28			1.75 (1.12-2.73)	0.0160
rs10733113[f]	G	0.79	0.89	53:28			1.89 (1.19-3.00)	5.10×10^{-3}
Toronto (129 trios)[i]								
rs4353135[f]	T	0.62	0.74	63:35			1.80 (1.17-2.77)	4.40×10^{-3}
rs4266924[f]	A	0.82	0.92	39:16			2.50 (1.44-4.35)	1.12×10^{-3}
rs55646866[g]	C	0.83	0.90	36:20			1.80 (1.07-3.04)	0.0313
rs6672995[g]	G	0.77	0.85	46:28			1.64 (1.05-2.56)	0.0355
ss107635144[g]	C	0.78	0.88	46:23			2.00 (1.22-3.27)	5.18×10^{-3}
rs10733113[f]	G	0.77	0.91	50:17			3.00 (1.76-5.10)	2.47×10^{-5}
Combined (710 trios, 239 Crohn's disease and 107 controls)[j]								
rs4353135[f]	T	0.65	0.71	282:185	39:44:11	92:95:27	1.21 (1.05-1.39)	8.36×10^{-3}
rs4266924[f]	A	0.84	0.91	161:80	63:35:4	184:41:2	1.69 (1.37-2.07)	6.01×10^{-7}
rs55646866[g]	C	0.84	0.91	156:78	66:32:4	191:38:3	1.69 (1.38-2.08)	7.20×10^{-7}
rs6672995[g]	G	0.79	0.87	194:114	57:40:7	172:56:4	1.53 (1.28-1.82)	2.91×10^{-6}
ss107635144[g]	C	0.80	0.88	186:100	57:38:7	165:56:3	1.53 (1.27-1.84)	8.50×10^{-6}
rs10733113[f]	G	0.80	0.90	192:89	58:38:5	177:47:3	1.78 (1.47-2.16)	3.49×10^{-9}

Because *NLRP3* is involved in IL-1β processing, we therefore also evaluated whether these six SNPs influenced IL-1β production. We cultured monocytes in the presence or absence of crude lipopoly-saccharide (LPS; Supplementary Methods), as its derivatives have been shown to stimulate *NLRP3* expression[10]. We then assessed IL-1β levels in culture supernatants. We observed a borderline-significant association between IL-1β levels and rs6672995 genotype under the unstimulated condition ($P = 0.0502$; Fig. 2c) and a significant association under the LPS-stimulated condition ($P = 0.00591$; Fig. 2d). In both cases, homozygosity for the risk allele was associated with the lowest level of IL-1β. Results for the five other Crohn's disease-associated SNPs are shown in Supplementary Figure 2 and Table 5 online.

We also examined *Nlrp3* expression in colon tissues isolated from mice with trinitrobenzene sulfonic acid (TNBS)-induced colitis, a model that mimics Crohn's disease-like intestinal inflammation, and in biopsies from

individuals with Crohn's disease. *Nlrp3* expression was significantly higher in both acute (fold change = 9.38 ± 1.58; $P < 0.0009$) and chronic (fold change = 2.70 ± 0.88; $P < 0.0152$) TNBS-induced colitis models than in colon tissues from control mice (Fig. 2e; see Methods and Supplementary Methods). *NLRP3* expression was also significantly higher in the ulcerated intestinal mucosa from human Crohn's disease samples (fold change = 4.08 ± 1.33; $P < 0.0028$) than in healthy controls (Fig. 2f; see Methods and Supplementary Methods).

Figure 2 *NLRP3 functional study results. (a,b) Linear regression analysis of NLRP3 mRNA level versus rs4353135 genotype in DNA-RNA matched freshly isolated PBCs (a; n = 30) and monocytes (b; n = 31) obtained from healthy individuals. Genotypes of the six Crohn's disease (CD)-associated SNPs (Table 1) were obtained by sequencing. Mean threshold cycle (C_t) was calculated for each sample from three replicates and then used to calculate relative expression level (ΔC_t), which is the difference between NLRP3 C_t and endogenous control 18S RNA C_t. Fold change in NLRP3 expression was calculated using comparative C_t method (see Methods), using as a reference the average ΔC_t of homozygosity for the risk allele. (c,d) Linear regression analysis of IL-1β production (pg/ml) versus rs6672995 genotype for unstimulated (c) and LPS-stimulated (d; 1.0 μg/ml) conditions after 3 h of incubation. ΔC_t (a,b) and IL-1β level (c,d) for each individual are shown in red; regression lines are shown as dashed lines (a-d). (e) Quantitative real-time PCR analysis of Nlrp3 expression in colons of healthy mice (n = 6), mice with acute TNBS-induced colitis (n = 12) and mice with chronic TNBS-induced colitis (n = 6). (f) Quantitative real-time PCR analysis of NLRP3 expression in colon specimens from healthy individuals (n = 35) and individuals with Crohn's disease (n = 25). Expression was normalized to 18S RNA expression; each bar represents mean fold change in NLRP3 expression ± s.e.m. normalized to that of healthy colon specimens (e,f).*



NLRP3 (chromosome 1q44) encodes cryopyrin, which is involved in the inflammasome signaling platform by regulating caspase-1 activity and IL-1β processing. The importance of cryopyrin in inflammation is highlighted by gain-of-function mutations within its NOD domain that are associated with three hereditary periodic fever syndromes: Muckle-Wells syndrome, familial cold autoinflammatory syndrome and neonatal-onset multisystem inflammatory disease[1,2]. Hyperproduction of IL-1β is thought to be a central event leading to symptoms in these three syndromes[1]. Consistent with these observations is the successful use of IL-1β targeted therapy for treating Muckle-Wells syndrome[11] and familial cold autoinflammatory syndrome[12].

In our study, contrary to the gain-of-function mutations described above leading to hyperproduction of IL-1β, we uncovered a regulatory region downstream of *NLRP3* that contributes to Crohn's disease susceptibility and is associated with hypoproduction of IL-1β and decreased *NLRP3* expression. Indeed, the risk allele of rs6672995, located in a predicted regulatory region[7], was associated with a decrease in LPS-induced IL-1β production, and the risk allele of rs4353135 was associated with a decrease in baseline *NLRP3* expression in two independent sample sets of healthy donors. It is noteworthy that these two SNPs were in weak linkage disequilibrium ($r^2 <$ 0.105) in our combined sample set. Although the causal variant has not been conclusively shown and may still be unidentified, it is most likely to be in linkage disequilibrium with the tested variants. Nonetheless, we have shown that SNPs in the associated 5.3-kb region influence *NLRP3* at both the gene expression and functional levels, as indicated by altered *NLRP3* expression and IL-1β production. Notably, dysregulated IL-1β production has also been linked to Crohn's disease pathogenesis in which the three major *NOD2* (chromosome 16q12) mutations result in a loss-of-function phenotype, with decreased NF-κB activation in response to muramyl dipeptide stimulation and decreased IL-1β production in primary human mononuclear and dendritic cells from individuals with Crohn's disease[13-16]. Our results further support the recent Crohn's disease immunopathogenesis paradigm, which suggests that a defective innate immune response impairs clearance of luminal antigens and/or pathogens and leads to the development of chronic intestinal inflammation and Crohn's disease. *NLRP3* may thus have a role in the initiation phase of the disease, as indicated by our *in vitro* and expression experiments, as well as a role in perpetuating chronic inflammation through further activation of caspase-1 and processing of IL-1β, as indicated by the enhanced *NLRP3* expression in the Crohn's disease and chronic TNBS-induced colitis samples.

The *NLRP3* locus can be added to the list of several newly uncovered Crohn's disease loci at which the common allele has been reported to be the risk allele[17-22]. Although it is difficult to strictly distinguish between one allele being a susceptibility risk factor and the other being a protective one, estimating attributable fractions and prevented fractions in addition to odds ratios offers insight into how to interpret these associations with very common risk factors in the context of complex diseases. If the SNPs with the strongest associations from Table 1 are interpreted as risk factors, then the attributable fractions (that is, the reduction in prevalence if the risk factor were removed from the population) of the alleles fall in the range of 45-55%. These proportions of 'cases explained' (all other factors being ignored) are large, but not surprisingly so, as most members of the population are carriers of the risk factors. Conversely, if the minor alleles are interpreted as protective factors, then their prevented fractions (that is, the proportion by which the prevalence would increase if the protective factor were removed from the population) fall in the range of 10-14%, a range easier to interpret in the context of complex genetic diseases.

Several recent genome-wide association studies have identified new Crohn's disease susceptibility genes using the Illumina HumanHap300 Genotyping BeadChip[17-19] and the Affymetrix GeneChip Human Mapping 500K Array Set[21]. Although rs4353135 is present on the Human Mapping 500K Array Set, this SNP showed the weakest significance of all six SNPs from Table 1, was the only one not consistently replicated across all samples and was in weak linkage disequilibrium with the other SNPs (max $r^2 = 0.28$). Supplementary Table 6 online shows the linkage disequilibrium between SNPs on the Illumina HumanHap300 and the Affymetrix GeneChip Human Mapping 500K arrays that were not genotyped in the present study and the SNPs from Table **1** found in HapMap. With a maximum r^2 of only 0.16, these observations may explain why this region escaped detection in these genome-wide association studies.

A recent meta-analysis of three large genome-wide association studies of Crohn's disease reported that well-established associations with Crohn's disease account for ~20% of the genetic variance observed in Crohn's disease, suggesting that additional genetic contributions have yet to be discovered[22]. With the exception of variations within *NOD2* and *IL23R*, established susceptibility alleles have been reported to have relatively modest effects, with odds ratios ranging from 0.7 to 1.7 (ref. 22). Despite the modest contribution of *NLRP3* to the risk of Crohn's disease, our results strongly implicate *NLRP3*, a gene with an essential role in regulating the inflammasome, as a risk factor for Crohn's disease. Our results also suggest that a gene such as *NLRP3* that is associated with rare, severe autoinflammatory disorders can also be implicated in the susceptibility of more

common inflammatory diseases such as Crohn's disease.

METHODS

Subjects. Five sample sets from four different centers, totaling 710 Crohn's disease trios, 239 Crohn's disease cases and 107 controls, were assembled for this project (Table 1). All participants gave informed consent, and studies were approved by the Institutional Review Board of each institution that sent samples. A clinical subtype of Crohn's disease was assigned using standard clinical criteria[23,24], except for a few 'indeterminate colitis' cases that were excluded from the study (Supplementary Methods). Belgian subjects from Leuven University Hospital Gasthuisberg (Leuven) were used for the exploratory experiments (n = 296 trios). The replication cohorts consisted of Belgian subjects from Université de Liège and of two Canadian cohorts (Québec and Toronto). Samples from Liège were subdivided into a family-based cohort (155 Liège trios) and a case-control cohort (Liège case-control, 239 Crohn's disease and 107 controls). The Québec cohort (n = 130 trios) comprised subjects from multiple sites in the province of Québec and included 22 probands of Ashkenazi and 5 of Sephardic Jewish ancestry. The Toronto samples (n = 129 trios) were collected from multiple sites in Toronto and included 26 probands of Ashkenazi Jewish ancestry. All study participants were of European descent except for 15 probands from Toronto that were excluded from the analyses.

Genotyping. We first investigated a 67.8-kb region spanning 1q44 (243890897-243958709; NCBI build 35, hg17), including *NLRP3* (32.9 kb). SNP selection details are given in Supplementary Methods. Samples were genotyped using the SNPstream ultra-high-throughput genotyping system (Orchid Biosciences)[25] and Sequenom homogenous MassExtend (hME) assays (matrix-assisted laser desorption/ionization time-of-flight mass spectrometry)[26]. Primers and probes are available in Supplementary Table 1. Analysis was restricted to SNPs passing quality filters, excluding SNPs with success rate < 95%, minor allele frequency < 5% or deviation from Hardy-Weinberg equilibrium (P < 0.01). Measures of linkage disequilibrium between SNPs and departures from Hardy-Weinberg equilibrium were computed using Haploview v4.0 (ref. 27). Families showing mendelian errors in 5% or more of the markers were excluded from the analysis (25 Crohn's disease families).

Sequencing. Primers were designed to have a T_m of 60 °C using the Primer3 program (Supplementary Table 2). PCR reactions were done using Hot Start Taq polymerase (Qiagen) in an 8-µl final volume comprising 9 ng of DNA (one cycle of 96 °C for 10 min; 40 cycles of 95 °C for 30 s, 58 °C for 1 min and 72 °C for 35 s; and one cycle of 72 °C for 7 min). Sequencing was done on an ABI 3730 DNA sequencer (Applied Biosystems) according to standard protocols. Sequence traces were assembled and analyzed using the PolyPhred software package and were compared to annotated sequences from NCBI build 35, hg17.

Statistical analysis. Tests of association were done using the likelihood methods implemented in UNPHASED v3.0.10 (ref. 28), which can analyze samples of nuclear families, unrelated subjects or a combination of both. For nuclear families (cases and their parents), the likelihood is equivalent to the conditional likelihood models on which the transmission disequilibrium test is based[29]. For samples of unrelated cases and controls, the likelihood is equivalent to a logistic model, which allows estimations of risk effects in terms of odds ratios along with their confidence intervals. It also allows conditioning on the observed association at one marker, to test whether or not the observed significance at others can solely be explained by linkage disequilibrium. A permutation procedure is implemented to allow the calculation of significance levels that are corrected for the number of tests. Attributable and prevented fractions were calculated under the assumption that alleles had additive effects on the penetrance scale, that odds ratios from Table 1 are good approximations for the relative risks and that allele frequencies in the controls are reasonable estimates of population frequencies.

RNA extraction and quantitative real-time PCR. Details on tissue collection and monocyte isolation are given in Supplementary Methods. Biopsies preserved in RNAlater (Qiagen) were homogenized, and total RNA was extracted using TRIzol (Invitrogen). Total RNA was extracted from monocytes using an RNeasy kit (Qiagen). Total RNA was extracted from PBCs using a PAXgene blood RNA kit (Qiagen) with RNase inhibitor, using off-column DNase I digestion and ethanol precipitation to improve RNA yield and quality. First-strand cDNA was synthesized from 1 µg of RNA template with a cDNA archive kit (Applied Biosystems), using MultiScribe reverse transcriptase and random primers. Quantitative real-time PCR assays (mouse *Nlrp3*, Mm00840904_ml; human *NLRP3*, Hs00918082_ml; 18S RNA, 4319413E; Applied Biosystems) were conducted using an ABI PRISM 7900 sequence detection system based on the 5' nuclease assay[30] and quantified using Applied Biosystems' comparative threshold cycle (C_t) method. The Wilcoxon signed-rank test was used to evaluate tissue (mouse and human) expression differences. Associations between *NLRP3* expression or IL-1β level and the genotypes of the Crohn's disease-associated SNPs were assessed using linear regression. Analyses were done using GraphPad Software.

DNA extraction. Monocyte DNA was isolated from whole blood using a FlexiGene DNA kit (Qiagen). PBC DNA was isolated using a Gentra Autopure automated system (Qiagen) according to the manufacturer's protocol.

Enzyme-linked immunosorbent assay. IL-1β was quantified in monocyte culture supernatants using a human IL-1β DuoSet ELISA kit (R&D Systems) according to the manufacturer's protocol.

URLs. ESPERR software, http://www.bx.psu.edu/projects/esperr.

Accession codes. GenBank: *NLRP3,* AF054176; *OR2B11*, NM_001004492. OMIM: *NLRP3*, 606416; *OR2B11*, 605956. SNP data have been submitted to NCBI dbSNP under the numbers ss107635120, ss107635122, ss107635124, SS107635126, ss107635128, ss107635130, ss107635132, ss107635133, SS107635136, ss107635138, ss107635140, ss107635142, ss107635144, SS107635146.

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

ACKNOWLEDGMENTS

We thank the study participants and their families for their collaboration; A. Montpetit and the genotyping and sequencing platforms of the McGill University and Genome Québec Innovation Centre for their technical assistance; T. McKenzie, G. Bonventi and C. Landolt-Maticorena for their help with acquisition and preparation of the PBCs; F. Martinon and J. Tschopp for their technical assistance in designing the functional studies; and T. Kwan for helpful discussion. This research was supported by grants to D.F. from the Canada Foundation for Innovation, the Canada Research Chair program, the Research Institute of the McGill University Health Centre and the Crohn's & Colitis Foundation of Canada (CCFC). A.-C.V. is supported by a PhD training award from the Fonds de la recherche en santé du Québec. M.S.S. is funded by the National Institute of Diabetes and Digestive and Kidney Diseases (DK624230), the CCFC and the Mount Sinai Hospital Gale and Graham Wright Research Chair in Digestive Diseases. J.D.R. is funded by grants from the National Institutes of Allergy and Infectious Diseases (AI065687 and AI067152), the National Institute of Diabetes and Digestive and Kidney Diseases (DK064869 and DK062432) and the Crohn's and Colitis Foundation of America (SRA512). E.L. is a senior research associate at the Belgium Foundation for Scientific Research (FNRS). S.V. is a clinical investigator for the Funds for Scientific Research of Flanders. T.J.H. is the recipient of an Investigator Award from the Ontario Institute for Cancer Research through funding from the Government of Ontario's Ministry of Research and Innovation. D.F. is a senior clinical scientist at the Belgium FNRS. None of the funding agencies had any role in the design and conduct of the study; in the collection, analysis and interpretation of the data; or in the preparation, review or approval of the manuscript.

AUTHOR CONTRIBUTIONS

A.-C.V., G.F. and D.F. conceived and designed the experiments. A.-C.V., G.F., C.C. and N.B. did the experiments. A.-C.V. and M.L. analyzed the data. E.L., M.S.S., C.L., J.B., A.B., D.G., A.C., D.L., P.R.F., J.E.W., M.S., P.R., J.D.R., S.V, T.J.H. and D.F. provided study samples. A.-C.V., M.L. and D.F. wrote the paper, with contributions from all authors.

Published online at http://www.nature.com/naturegenetics/

REFERENCES

1. Agostini, L. *et al.* NALP3 forms an IL-1beta-processing inflammasome with increased activity in Muckle-Wells autoinflammatory disorder. *Immunity* **20,** 319-325 (2004).

2. Mariathasan, S. & Monack, D.M. Inflammasome adaptors and sensors: intracellular regulators of infection and inflammation. *Nat. Rev. Immunol.* **7,** 31-40 (2007).

3. Podolsky, D.K. Inflammatory bowel disease. *N. Engl. J. Med.* **347,** 417-429 (2002).

4. Ting, J.P., Kastner, D.L. & Hoffman, H.M. CATERPILLERs, pyrin and hereditary immunological disorders. *Nat. Rev. Immunol.* **6,** 183-195 (2006).

5. Pétrilli, V., Dostert, C., Muruve, D.A. & Tschopp, J. The inflammasome: a danger sensing complex triggering innate immunity. *Curr. Opin. Immunol.* **19,** 615-622 (2007).

6. Martinon, F. & Tschopp, J. Inflammatory caspases and inflammasomes: master switches of inflammation. *Cell Death Differ.* **14,** 10-22 (2007).

7. Taylor, J. *et al.* ESPERR: learning strong and weak signals in genomic sequence alignments to identify functional elements. *Genome Res.* **16,** 1596-1604 (2006).

8. de Bakker, P.I. *et al.* Efficiency and power in genetic association studies. *Nat. Genet.* **37,** 1217-1223 (2005).

9. Kummer, J.A. *et al.* Inflammasome components NALP 1 and 3 show distinct but separate expression profiles in human tissues suggesting a site-specific role in the inflammatory response. *J. Histochem. Cytochem.* **55,** 443-452 (2007).

10. Martinon, F., Agostini, L., Meylan, E. & Tschopp, J. Identification of bacterial muramyl dipeptide as activator of the NALP3/cryopyrin inflammasome. *Curr. Biol.* **14,** 1929-1934 (2004).

11. Hawkins, P.N., Lachmann, H.J. & McDermott, M.F. Interleukin-l-receptor antagonist in the Muckle-Wells syndrome. *N. Engl. J. Med.* **348,** 2583-2584 (2003).

12. Hoffman, H.M. *et al.* Prevention of cold-associated acute inflammation in familial cold autoinflammatory syndrome by interleukin-1 receptor antagonist. *Lancet* **364,** 1779-1785 (2004).

13. Li, J. *et al.* Regulation of IL-8 and IL-1beta expression in Crohn's disease associated NOD2/CARD15 mutations. *Hum. Mol. Genet.* **13,** 1715-1725 (2004).

14. Van Heel, D.A. *et al.* Muramyl dipeptide and toll-like receptor sensitivity in N0D2-associated Crohn's disease. *Lancet* **365,** 1794-1796 (2005).

15. Kramer, M., Netea, M.G., de Jong, D.J., Kullberg, B.J. & Adema, G.J. Impaired dendritic cell function in Crohn's disease patients with N0D2 3020insC mutation. *J. Leukoc. Biol.* **79,** 860-866 (2006).

16. van Beelen, A.J. *et al.* Stimulation of the intracellular bacterial sensor N0D2 programs dendritic cells to promote interleukin-17 production in human memory T cells. *Immunity* **27,** 660-669 (2007).

17. Duerr, R.H. *et al.* A genome-wide association study identifies IL23R as an inflammatory bowel disease gene. *Science* **314,** 1461-1463 (2006).

18. Rioux, J.D. *et al.* Genome-wide association study identifies new susceptibility loci for Crohn disease and implicates autophagy in disease pathogenesis. *Nat. Genet.* 39, 596-604 (2007).

19. Libioulle, C. *et al.* Novel Crohn disease locus identified by genome-wide association maps to a gene desert on 5p13.1 and modulates expression of PTGER4. *PLoS Genet.* **3,** e58 (2007).

20. Parkes, M. *et al.* Sequence variants in the autophagy gene IRGM and multiple other replicating loci contribute to Crohn's disease susceptibility. *Nat. Genet.* **39,** 830-832 (2007).

21. Wellcome Trust Case Control Consortium. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature* **447,** 661-678 (2007).

22. 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).

23. Lennard-Jones, J.E. Classification of inflammatory bowel disease. *Scand. J. Gastroenterol. Suppl.* **170,** 2-6 (1989).

24. Silverberg, M.S. *et al.* Toward an integrated clinical, molecular and serological classification of inflammatory bowel disease: Report of a Working Party of the 2005 Montreal World Congress of Gastroenterology. *Can. J. Gastroenterol.* (Suppl A), **19,** 5-36 (2005).

25. Bell, P.A. *et al.* SNPstream UHT: ultra-high throughput SNP genotyping for pharmaco-genomics and drug discovery. *Biotechniques* Suppl 70-72, 74, 76-77 (2002).

26. van den Boom, D. & Ehrich, M. Discovery and identification of sequence polymorphisms and mutations with MALDI-TOF MS. *Methods Mol. Biol.* **366,** 287-306 (2007).

27. Barrett, J.C, Fry, B., Mailer, J. & Daly, M.J. Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics* **21,** 263-265 (2005).

28. Dudbridge, F. Likelihood-based association analysis for nuclear families and unrelated subjects with missing genotype data. *Hum. Hered.* **66,** 87-98 (2008).

29. Spielman, R.S., McGinnis, R.E. & Ewens, W.J. Transmission test for linkage disequilibrium: the insulin gene region and insulin-dependent diabetes mellitus (IDDM). *Am. J. Hum. Genet.* **52,** 506-516 (1993).

30. Holland, P.M., Abramson, R.D., Watson, R. & Gelfand, D.H. Detection of specific polymerase chain reaction product by utilizing the 5'→3' exonuclease activity of *Thermus aquaticus* DNA polymerase. *Proc. Natl. Acad. Sci. USA* **88,** 7276-7280 (1991).