Gut Online First, published on September 7, 2011 as 10.1136/gutjnl-2011-300078

ORIGINAL ARTICLE

NADPH oxidase complex and IBD candidate gene studies: identification of a rare variant in *NCF2* that results in reduced binding to RAC2

Aleixo M Muise,[1,2,3] Wei Xu,[4,5] Cong-Hui Guo,[2] Thomas D Walters,[1] Victorien M Wolters,[1] Ramzi Fattouh,[2] Grace Y Lam,[2] Pingzhao Hu,[6] Ryan Murchie,[2] Mary Sherlock,[1] Juan Cristóbal Gana,[7] NEOPICS,* Richard K Russell,[8] Michael Glogauer,[9] Richard H Duerr,[10 11] Judy H Cho,[12] Charlie W Lees,[13] Jack Satsangi,[13] David C Wilson,[14] Andrew D Paterson,[5,6] Anne M Griffiths,[1] Mark S Silverberg,[15] John H Brumell[2,3,16]

► Additional data are published online only. To view these files please visit the journal online (http://gut.bmj.com).

For numbered affiliations see end of article.

Correspondence to
Dr Aleixo Muise, Division of Gastroenterology, Program in Cell Biology, Hospital for Sick Children, 555 University Ave, Toronto, Ontario M5G 1X8, Canada;
aleixo.muise@utoronto.ca

MSS and JHB contributed equally.

*Canadian National Early Onset IBD Cohort Study (NEOPICS): see online supplement for details.

Revised 4 August 2011
Accepted 5 August 2011

ABSTRACT

Objective The NOX2 NADPH oxidase complex produces reactive oxygen species and plays a critical role in the killing of microbes by phagocytes. Genetic mutations in genes encoding components of the complex result in both X-linked and autosomal recessive forms of chronic granulomatous disease (CGD). Patients with CGD often develop intestinal inflammation that is histologically similar to Crohn's colitis, suggesting a common aetiology for both diseases. The aim of this study is to determine if polymorphisms in NOX2 NADPH oxidase complex genes that do not cause CGD are associated with the development of inflammatory bowel disease (IBD).

Methods Direct sequencing and candidate gene approaches were used to identify susceptibility loci in NADPH oxidase complex genes. Functional studies were carried out on identified variants. Novel findings were replicated in independent cohorts.

Results Sequence analysis identified a novel missense variant in the neutrophil cytosolic factor 2 (*NCF2*) gene that is associated with very early onset IBD (VEO-IBD) and subsequently found in 4% of patients with VEO-IBD compared with 0.2% of controls ($p=1.3\times10^{-5}$, OR 23.8 (95% CI 3.9 to 142.5); Fisher exact test). This variant reduced binding of the *NCF2* gene product $p67^{phox}$ to RAC2. This study found a novel genetic association of *RAC2* with Crohn's disease (CD) and replicated the previously reported association of *NCF4* with ileal CD.

Conclusion These studies suggest that the rare novel $p67^{phox}$ variant results in partial inhibition of oxidase function and are associated with CD in a subgroup of patients with VEO-IBD; and suggest that components of the NADPH oxidase complex are associated with CD.

Significance of this study

What is already known about this subject?

- Defects in the NADPH oxidase complex genes cause X-linked and autosomal recessive chronic granulomatous disease (CGD).
- Patients with CGD are more susceptible to developing Crohn's-like colitis and perianal disease.
- Polymorphisms in the NADPH oxidase gene *NCF4* were found to be associated with ileal Crohn's disease (CD).

What are the new findings?

- Identification of a novel variant in *NCF2* that is associated with very early onset inflammatory bowel disease (IBD) and results in reduced protein binding.
- Genetic studies suggest the NADPH oxidase complex gene *RAC2* as a CD susceptibility gene.
- Replication of the *NCF4* gene association with ileal CD.

How might it impact on clinical practice in the foreseeable future?

- These results implicate the NADPH oxidase complex in the pathogenesis of IBD.
- Identification of novel variants in these genes may lead to alternative therapies for a subgroup of patients with defective reactive oxygen species production.

INTRODUCTION

Inflammatory bowel diseases (IBD) are hypothesised to occur in genetically susceptible individuals as a result of dysregulated immune responses to gut flora after exposure to an as yet unidentified environmental stimulus.[1 2] Investigation of diseases that present with intestinal inflammation that are similar to IBD may provide an important insight into the pathogenesis of IBD. For example, chronic granulomatous disease (CGD) is a rare genetic disorder with a prevalence of 1/200 000 to 1/250 000 caused by X-linked and autosomal recessive mutations in genes encoding components of the nicotinamide adenine dinucleotide phosphate (NADPH) oxidase complex (also referred to as NOX2 NADPH oxidase or phagocyte oxidase).[3] Patients with CGD have severe and recurrent infections as a result of the inability of phagocytes to mount sufficient respiratory burst to kill invading pathogens.[4] Interestingly, up to 40% of patients with CGD develop a form of colitis that is

endoscopically and pathologically very similar to the colitis observed in Crohn's disease (CD).[5–7]

Defective neutrophil respiratory burst has been observed in patients with IBD,[8–10] and a genome-wide association study (GWAS) identified the *NCF4* (encoding the $p40^{phox}$ subunit of the NOX2 NADPH oxidase) region as an ileal CD-specific susceptibility gene. This association did not meet genome-wide significance in the original published GWAS[11 12] and was not replicated in a recent GWAS meta-analysis, perhaps due to the fact that the GWAS analysis focused on CD independent of disease location.[13] We have therefore undertaken studies to determine if components of the NADPH oxidase complex play a role in the development of IBD. Here we report a novel missense variant in *NCF2* in patients with very early onset IBD (VEO-IBD) that results in neutrophil dysfunction and susceptibility to CD. We also describe novel associations of the NADPH oxidase complex gene *RAC2* with CD and replicate the previously described association of *NCF4* with ileal CD. Together, these results demonstrate that the NADPH oxidase complex genes play a role in the pathogenesis of CD.

MATERIALS AND METHODS

RT-PCR

RNA was isolated from the whole blood by PAX gene blood RNA kit (Qiagen, USA) according to the manufacturer's instructions. cDNA was synthesised using SuperScript III Reverse Transcriptase (Life Technologies, Carlsbad, California, USA). Primers for full length *NCF2, NCF4* and *RAC2* were designed and synthesised at CDI (see table 1 in online supplement). PCR was performed according to a standard protocol and the purified PCR product was cloned into pJET cloning vector (Fermentas, Hanover, Maryland, USA) and sequenced by ABI 3730 DNA analyser (Applied Biosystems, Melbourne, Australia).

NCF2 genotyping

In order to determine if the NCF2 variant c.113 G→A R38Q was found in other patients with IBD, we designed a custom Taqman probe (tcagtgccgtccaggacccccactccc(g/a) gatttgcttcaacattggctgc). Four hundred and eighty patients and 480 controls and a second cohort of 119 patients with VEO-IBD (Toronto samples described below) were genotyped using Taqman at the Centre for Applied Genomics (TCAG), Hospital for Sick Children, Toronto (see online supplement for binding studies).

Single nucleotide polymorphism analysis and genotyping

International HapMap project[14] (http://www.hapmap.org) Caucasian (CEU) phase II data Release 23a were used to select tag single nucleotide polymorphisms (SNPs) (minor allelic frequency (MAF) >1%) that span the NADPH oxidase complex genes and flanking regions through the 'Tagger' software program (r^2 >0.8).[15] Twenty-one tag SNPs covering the *NCF4* region (chromosome 22, 35 581 544 to 35 598 557), 15 tag SNPs covering the *NCF2* region (chromosome 1, 180 256 354 to 180 291 372), 19 tag SNPs covering the *RAC2* region (chromosome 22, 35 945 811 to 35 964 804), five tag SNPs covering the *CYBA* region (chromosome 16, 87 237 198 to 87 244 957) and six tag SNPs covering the *NOX2/CYBB* (X-chromosome, 37 395 536 to 37 428 930) were used to capture all SNPs with r^2>0.8 to the tag SNPs (see table 1 for list of NADPH oxidase SNPs). As *NCF1* is located in tandem with two nearly identical pseudogenes, analysis of this region was not carried out. Genotyping of samples was performed using the Illumina Goldengate Custom Chip genotyping system (Toronto discovery) and Taqman (North America and Scotland Replication) at the Centre for Applied Genomics, Hospital for Sick Children, Toronto and the University of Edinburgh. Sixty-two SNPs in the NADPH oxidase complex (*RAC2*, *CYBA*, *CYBB*, *NCF2* and *NCF4*) passed quality control.

Subjects, quality control and population stratification

All subjects in this study were of European descent by self-reporting of ethnic heritage. All probands had a confirmed diagnosis of IBD and fulfilled standard diagnostic criteria. Phenotypic characterisation was based on the Montreal Classification.[16] Perianal disease (Montreal Classification 'p') included only those patients with perianal abscess and/or fistulae. Phenotypic information and DNA samples were obtained from study subjects with approval of the institutional review ethics board for IBD genetic studies at the Hospital for Sick Children and Mount Sinai Hospital in Toronto. Replication cohorts had review ethics board approval for genetic and phenotypic studies at the individual institutions. Written informed consent was obtained from all participants.

The discovery cohort included patients recruited from the Hospital for Sick Children (22%) and Mount Sinai Hospital (78%) in Toronto and local and National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK) control individuals. A total of 2049 subjects (656 with CD, 544 with ulcerative colitis (UC) and 849 controls; see table 2A in online supplement for demographic and phenotypic description) in the discovery analysis were used in the final analysis. The first replication cohort consisted of 1836 Caucasian individuals from North America including 443 patients with CD, 477 with UC and 916 controls (see table 2B in online supplement for demographic and phenotypic description); NIDDK patients recruited from Chicago and Pittsburgh with North America Caucasians of European origin control individuals obtained from the Centre for Applied Genomics (Ontario Population Genomics Platform (plates used: 1–4, 6, 9–12, 14); a complete description of this control population can be found at http://www.tcag.ca/cyto_population_control_DNA.html). The second cohort consisted of 2449 individuals exclusively recruited from Scotland including 691 patients with CD, 615 with UC and 1143 controls (see table 2C in online supplement for demographic and phenotypic description). All patients and control individuals were Caucasian and all related individuals were excluded by checking the estimated identity by descent for each pair of samples. Part of these cohorts has been used in previous GWAS including all the NIDDK patients in the North American replication cohort[11 17] and 374 individuals from Scotland in the Paediatric IBD GWAS.[18] None of the replication cohort control individuals were genotyped in IBD GWAS.

Preliminary analysis

HAPLOVIEW[19] was used to obtain LD patterns, obtaining descriptive statistics for the SNPs. PLINK version 1.06[20] was used to test for Hardy-Weinberg equilibrium (HWE) for each marker based on the Pearson χ^2 test. Descriptive statistics of demographic variables were generated using SAS V.9.2 (SAS Institute). The Wilcoxon rank sum test and χ^2 test were used to identify differences in demographic variables between subgroups.

Association analysis

The analysis was applied in stages. In stage 1, association analyses were applied to detect the associations with the 62 SNPs in five genes involved in the NADPH oxidase complex (*RAC2*, *CYBA*, *CYBB (NOX2)*, *NCF2* and *NCF4*) and IBD, CD and UC versus healthy controls. Logistic regression models were applied to an additive genetic model and Pearson χ^2 tests

Table 1 Association of NADPH oxidase genes *NCF4* and *RAC2* with CD in the discovery cohort

Chromosome	SNP	Gene	Position	A1	A2	MAF	FA	FU	p Value	OR	L95	U95
22	rs6000448	NCF4	35 580 717	1	3	0.108045	0.1176	0.1007	0.1839	1.184	0.9228	1.52
22	rs10854693	NCF4	35 581 956	3	1	0.405648	0.378	0.427	0.003823	0.7925	0.6769	0.9278
22	rs1883113	NCF4	35 582 630	3	2	0.0677741	0.07546	0.06184	0.01325	1.472	1.084	2
22	rs4821542	NCF4	35 582 864	1	3	0.165449	0.1776	0.1561	0.1857	1.149	0.9354	1.411
22	rs7287350	NCF4	35 583 185	1	3	0.0548537	0.06031	0.05065	0.02581	1.47	1.048	2.063
22	rs1883112	NCF4	35 586 792	1	3	0.422821	0.3941	0.4451	0.001388	0.7732	0.6605	0.9053
22	rs4821544	NCF4	35 588 449	3	1	0.344518	0.3735	0.3221	0.005118	1.257	1.071	1.475
22	rs741997	NCF4	35 588 759	3	1	0.0843854	0.09299	0.07774	0.121	1.243	0.9442	1.636
22	rs746713	NCF4	35 589 305	3	1	0.306977	0.3011	0.3115	0.7173	0.9696	0.8202	1.146
22	rs909484	NCF4	35 590 547	1	3	0.171761	0.17	0.1731	0.8553	0.9808	0.7968	1.207
22	rs729749	NCF4	35 593 792	1	3	0.217158	0.2184	0.2162	0.7703	1.028	0.853	1.239
22	rs2075938	NCF4	35 596 268	1	3	0.279255	0.2679	0.288	0.3198	0.915	0.7682	1.09
22	rs8137456	NCF4	35 604 389	2	3	0.0687708	0.07774	0.06184	0.1693	1.24	0.9125	1.685
22	rs8137602	NCF4	35 604 595	3	2	0.0916944	0.08155	0.09953	0.07419	0.7759	0.5872	1.025
22	rs3959633	NCF4	35 605 501	1	3	0.0249169	0.03125	0.02002	0.07318	1.56	0.9591	2.537
22	rs5756379	NCF4	35 605 850	3	1	0.334551	0.3285	0.3392	0.5302	0.9483	0.8033	1.119
22	rs4821554	NCF4	35 606 030	1	3	0.257475	0.2424	0.2691	0.1157	0.8646	0.7213	1.036
22	rs5750326	NCF4	35 606 231	1	2	0.496678	0.5023	0.4923	0.9018	1.01	0.8663	1.177
22	rs5756564	RAC2	35 942 049	3	1	0.397674	0.3598	0.427	0.0006546	0.7562	0.6439	0.888
22	rs4820272	RAC2	35 945 560	1	3	0.162791	0.189	0.1425	0.001566	1.401	1.137	1.727
22	rs933222	RAC2	35 948 583	1	3	0.281395	0.295	0.2709	0.2225	1.117	0.9348	1.336
22	rs12166968	RAC2	35 951 304	2	3	0.0671096	0.0686	0.06596	0.9971	1.001	0.7332	1.365
22	rs6572	RAC2	35 951 391	3	2	0.453488	0.4947	0.4217	0.00038	1.337	1.139	1.57
22	rs739041	RAC2	35 954 945	3	1	0.434492	0.4331	0.4356	0.8488	1.015	0.8678	1.188
22	rs9607431	RAC2	35 959 884	2	1	0.135548	0.1677	0.1107	0.00008594	1.577	1.256	1.98
22	rs1476002	RAC2	35 960 734	1	3	0.12766	0.1321	0.1243	0.6058	1.065	0.8383	1.353
22	rs13058338	RAC2	35 962 716	4	1	0.256981	0.2294	0.2783	0.006391	0.7773	0.6486	0.9316
22	rs5756573	RAC2	35 963 721	1	3	0.145376	0.1437	0.1466	0.7453	1.037	0.8329	1.291
22	rs2284038	RAC2	35 965 001	3	1	0.398007	0.404	0.3934	0.6625	1.037	0.8821	1.218
22	rs2239774	RAC2	35 967 599	2	3	0.164452	0.1974	0.139	0.0001514	1.495	1.214	1.84
22	rs2239775	RAC2	35 967 787	1	2	0.134884	0.1319	0.1372	0.4163	0.9086	0.7212	1.145
22	rs2239773	RAC2	35 968 235	1	3	0.277002	0.247	0.2996	0.004515	0.7696	0.6424	0.9221
22	rs2213430	RAC2	35 968 906	1	3	0.434219	0.436	0.4329	0.6503	1.038	0.8845	1.217
22	rs6000632	RAC2	35 974 061	1	3	0.245183	0.237	0.2515	0.4481	0.9314	0.7751	1.119
22	rs4821615	RAC2	35 976 077	2	3	0.447841	0.4512	0.4452	0.8961	1.01	0.8644	1.181
22	rs12484031	RAC2	35 979 216	3	1	0.100664	0.09299	0.1066	0.2213	0.8521	0.6593	1.101
22	rs7288979	RAC2	35 979 440	3	1	0.176412	0.1608	0.1885	0.06473	0.8266	0.6754	1.012

The discovery cohort consisted of 2049 subjects (656 with Crohn's disease, 544 with ulcerative colitis and 849 controls).
p Values are presented as uncorrected. Bonferroni correction threshold for 62 SNPs and IBD/UC/CD, $\alpha = 2.0 \times 10^{-4}$.
A1/A2, allele 1/2; CD, Crohn's disease; MAF, minor allelic frequency; FA, frequency affected; FU, frequency unaffected; IBD, inflammatory bowel disease; L95 and U95, lower and upper 95th confidence interval; SNP, single nucleotide polymorphism.; UC, ulcerative colitis.

were applied for dominant and recessive genetic models. We used an additive genetic model for primary analysis.[21] For the *CYBB* analysis we used the chromosome-counting approach[22] employed by the Wellcome Trust Case-Control Consortium.[23] Throughout the report the p values are the additive model p value. ORs and 95% CIs were estimated for the disease group compared with the control group. The association, adjusting for selected principal component vectors from the Eigenstrat analysis, was tested using conditional logistic regression (SAS V.9.2).

In stage 2 the four *RAC2* SNPs identified from the discovery cohort were genotyped in a replication cohort (North America) and an independent validation cohort (Scotland). Independent analysis was applied to the replication and validation cohorts. Combined effect estimates from all three IBD cohorts were estimated using a logistic regression model including cohorts as adjusted covariates. All p values are two-sided.

Subphenotype analysis

In addition to comparing patients with IBD, CD and UC with healthy controls, we applied subphenotype analysis to evaluate the genetic effect on the disease risk of the IBD subpopulation according to the Montreal Classification system. The subpopulation comparisons were applied for each of the genetic markers on ileal only (L1), colonic only (L2), ileocolonic (L3), ileal any (L1/L3), colon only (L2 plus UC), colon any (L2/L3 plus UC), perianal (p) and early diagnosis patients (diagnosis age <16 years) versus healthy controls. The allelic model was used to test single marker associations between each of the subgroups. The analyses were applied to the discovery cohort, the North American replication cohort, the Scottish replication cohort and the pooled samples separately. Since the subphenotype analysis is exploratory and hypothesis-generating, $p<0.05$ was used to define nominal signals.

RESULTS

Novel *NCF2* variant associated with very early onset IBD

Many patients with VEO-IBD (defined as onset of disease before 10 years of age based on the Paris IBD Classification[24]) present with clinical intestinal features, including perianal disease and pancolitis, similar to those observed in CGD. We therefore sequenced coding regions of NADPH oxidase complex genes (including *NCF2*, *NCF4* and *RAC2*) in 10 patients with VEO-IBD with pancolitis and perianal disease without evidence of immunodeficiency, as determined by commercial genetic testing

for known CGD mutations, and a negative comprehensive immunological investigation including normal T, B and NK cell populations, response to immunisations, normal serum immunoglobulins and sequencing of the *IL10RA/B* genes. An outline of the experimental approach is shown in figure 1.

All exons and flanking intron sequences in *NCF2*, *NCF4* and *RAC2* were successfully sequenced in the 10 patients (data not shown). We found a novel *NCF2* (encoding $p67^{phox}$) heterozygote variant c.113 G→A ($p67^{phox}$ R38Q) in one patient. This male patient was diagnosed with IBD at 1 year of age with pancolitis (colonoscopy at age 1 and 3 years) and had developed perianal abscesses and small bowel disease by 2 years of age. He had low to normal reactive oxygen species (ROS) production as assessed by the nitroblue tetrazolium (NBT) test on two separate occasions (40 and 53; normal range 32–300). The patient had no history of chronic or significant infections. The identified *NCF2* missense variant (c.113 G→A) has not previously been reported and is not known to cause CGD (NCF2base; Mutation Registry for Autosomal Recessive Chronic Granulomatous Disease, http://bioinf.uta.fi/NCF2base/). The c.113 G-A variant in NCF2 has been subsequently assigned the rs#-rs147415774 (http://www.ncbi.nlm.nih.gov/SNP/snp_ref.cgi?rs=147415774).

To determine if this novel variant plays a role in VEO-IBD, we examined an additional 139 patients with VEO-IBD, 341 patients with IBD and 480 healthy controls. Four of the 139 patients with VEO-IBD were found to be heterozygous for the c.113 G→A variant compared with one of the 341 patients with IBD and one of the 480 controls. We next carried out genotyping of this novel variant in a second independent VEO-IBD cohort consisting of 119 patients. Four of these patients were found to be heterozygous and two homozygous for the c.113 G→A variant. Overall, 4% (11/268) of patients with VEO-IBD carried the c.113 G→A *NCF2* variant compared with 0.2% (1/480) of healthy controls studied ($p=1.3\times10^{-5}$, OR 23.8 (95% CI 3.9 to 142.5); Fisher exact test). All patients with VEO-IBD carrying the *NCF2* c.113 G→A variant had extensive colonic disease, five had perianal disease and one had significant arthritis (table 2).

To determine the functional significance of the c.113 G→A *NCF2* variant, we examined the protein product R38Q $p67^{phox}$. The arginine at position 38 of $p67^{phox}$ has been shown to be important in RAC binding in a yeast two-hybrid assay,[25] and binding of these proteins is known to be an essential step in the assembly and activation of the NOX2 NADPH oxidase.[3] Co-immunoprecipitation studies (figure 2A) showed that the $p67^{phox}$ R38Q variant has decreased binding to RAC2 in comparison with wild-type $p67^{phox}$. Computational analysis showed that the reduced binding of the $p67^{phox}$ R38Q variant to RAC occurs as a result of a shorter side chain and loss of charge at amino acid 38 in $p67^{phox}$; both are required to stabilise the interaction between $p67^{phox}$ and RAC (figure 2B).

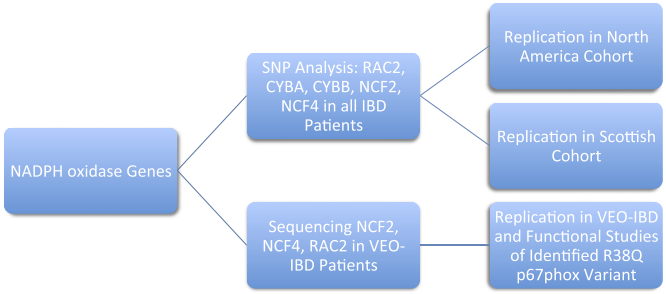


Figure 1 Flow diagram of NADPH oxidase genetic experiments. VEO-IBD, very early onset inflammatory bowel disease.

NADPH oxidase complex SNP analysis

We also examined the association of 62 genotyped SNPs in the NADPH oxidase complex, *RAC2*, *CYBA*, *CYBB*, *NCF2* and *NCF4* with IBD, CD and UC in the 2049 subjects in the discovery cohort (656 with CD, 544 with UC and 849 controls, table 1; see also table 3A–C in online supplement for full detailed analysis). The experimental approach is shown in figure 1. We found a novel association with *RAC2* and CD after Bonferroni correction threshold for 62 SNPs examined in IBD/UC/CD ($\alpha=2.0\times10^{-4}$; table 1). *RAC2* is located on chromosome 22 (35 951 258 to 35 970 251; figure 1 in the online supplement shows the LD plot of *RAC2* tag SNPs). To further examine this association, we genotyped four *RAC2* SNPs with modest association in the discovery cohort ($p<0.005$; rs5756564, rs6572, rs9607431 and rs2239774) in two independent cohorts (figure 1). The replication cohort comprised 1836 Caucasian subjects including 443 patients with CD, 477 with UC and 916 controls recruited from North America (table 4A in online supplement) and a second validation cohort from Scotland comprised 2449 Caucasian subjects including 691 patients with CD, 615 with UC and 1143 controls (table 4B in online supplement). No *RAC2* SNP association was replicated in both independent cohort (table 3). However, in a combined analysis of the discovery, replication and validation cohorts (1790 patients with CD, 1636 with UC and 2908 controls), the *RAC2* association signal for rs6572 was significantly associated with CD and not UC (table 3; see also table 5 in online supplement). It is important to note

Table 2 Phenotype of patients with inflammatory bowel disease with R38Q $p67^{phox}$

VEO-IBD R38Q $p67^{phox}$	Genotype of risk allele	Gender	Age at diagnosis (years)	Disease type at diagnosis	Disease location	Perianal disease	Other
1 (index patient)	Heterozygote	M	<1	CD	Ileocolonic (L3)	Yes	
2	Heterozygote	M	2	IBDU	Colonic (L2)	Yes	
3	Heterozygote	F	4.2	CD	Colonic (L2)	Yes	
4	Heterozygote	F	4.6	CD	Ileocolonic (L3)	No	Severe arthritis
5	Heterozygote	F	<1	CD	Ileocolonic (L3)	Yes	
6	Heterozygote	M	9	CD	Ileocolonic (L3)	No	Extensive small bowel disease
7	Heterozygote	F	9.7	IBDU	Colonic (L2)	No	
8	Heterozygote	M	8.8	IBDU	Ileocolonic (L3)	No	
9	Heterozygote	M	7.2	IBDU	Colonic (L2)	No	
10	Homozygote	F	5.5	IBDU	Colonic (L2)	No	
11	Homozygote	F	7.7	CD	Ileocolonic (L3)	Yes	

L3, ileocolonic CD; L2, colonic only CD.
CD, Crohn's disease; IBDU, inflammatory bowel disease unclassified; VEO-IBD, very early onset inflammatory bowel disease.

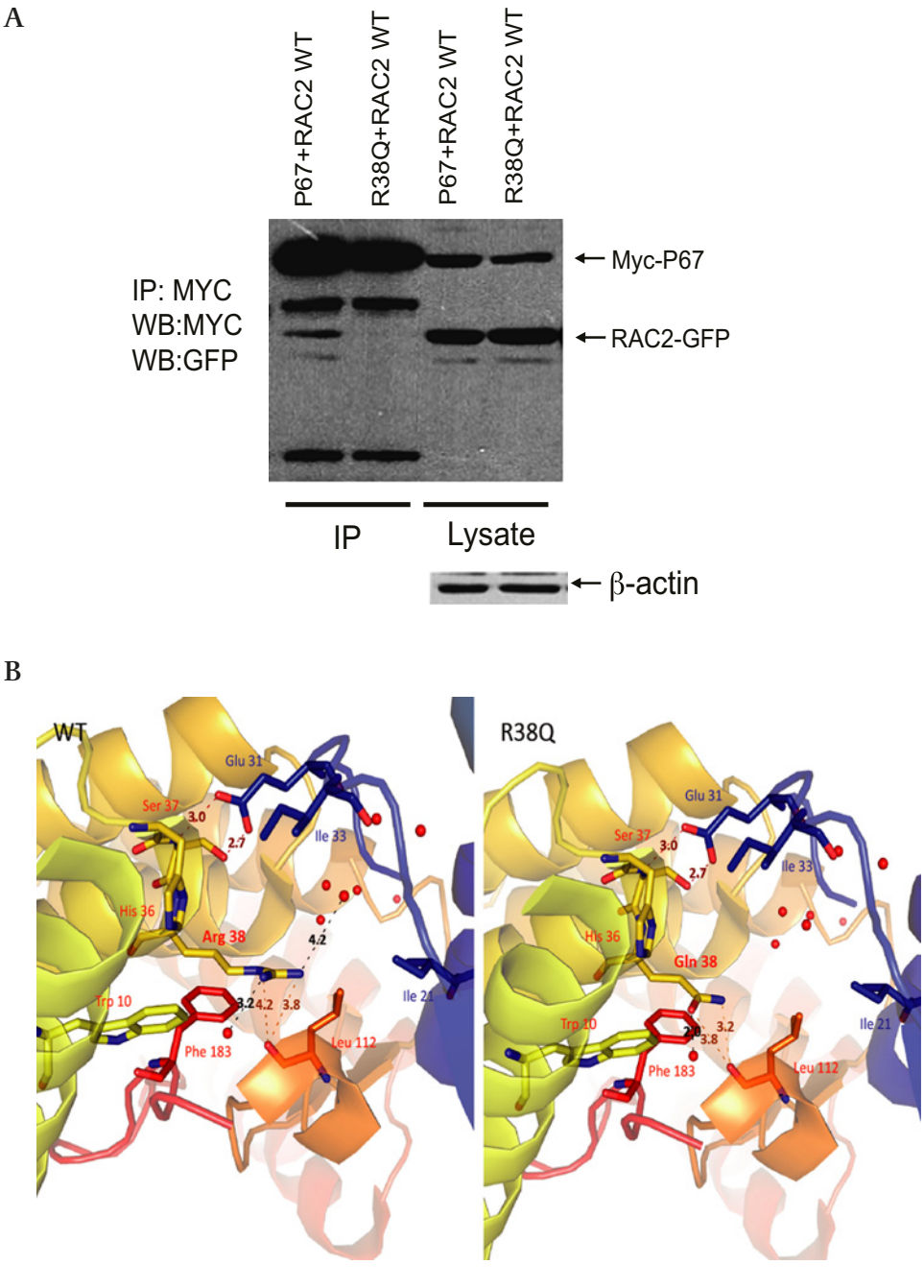


Figure 2 Functional studies of the p67^{phox} R38Q polymorphism. (A) p67^{phox} R38Q variant reduces binding to RAC. Interaction between RAC and wild-type p67^{phox} and p67^{phox} R38Q. GFP-tagged RAC2 and MYC-tagged p67^{phox}and p67^{phox} R38Q were co-transfected into T293 cells. After 20 h p67^{phox} was immunoprecipitated using anti-MYC antibody and blotted for MYC and GFP. RAC2 showed a 51% reduction of binding to p67^{phox} R38Q in comparison with RAC2 binding to wild-type p67^{phox}. Representative blot of three independent experiments. IP, immunoprecipitation; WB, western blot. (B) p67^{phox} R38Q variant results in a conformational change. The model shows a close-up view of the RAC binding cavity of the wild-type p67^{phox} and p67^{phox} R38Q as adapted from the RAC/p67^{phox} complex crystal structure (PDB 1E96). The proteins are shown in ribbon representation, coloured according to secondary structure, with RAC residues labelled in blue and p67^{phox} residues labelled in red. Residues within 7×10^{-8} cm of p67^{phox} R38 are represented by stick models with nitrogen and oxygen atoms indicated in blue and red, respectively. Water molecules found within the crystal structure are denoted by red spheres. Dashed lines denote putative hydrogen bond interactions between p67^{phox} wild-type and p67^{phox} L112 as well as water molecule 2013. In the p67^{phox} R38Q variant, bond distances between p67^{phox} R38Q and p67^{phox} L112 are reduced, the interaction with water molecule 2013 is lost, and a new putative interaction between p67^{phox} R38Q and water molecule 2025 forms. In addition, owing to the reduced carbon chain length between the arginine and glutamine side chains, a hydrophilic cavity is present perpendicular to the RAC/p67^{phox} R38Q interface that is visible through electron density mapping (not shown). All distances labelled are $\times10^{-8}$ cm. Images were generated with PyMOL (PyMOL Molecular Graphics System, Version 1.3, Schrödinger, LLC.)

that no *RAC2* SNP achieved genome-wide association significance in this study nor was significantly associated in all three populations studied. One SNP (rs6572; $p_{combined}=4.8\times10^{-5}$, OR 1.2 (95% CI 1.09 to 1.29)) remained significantly associated with CD after the Bonferroni correction threshold for the number of SNPs genotyped in this study and the phenotypes examined (62 SNPs examined in IBD/UC/CD, $\alpha=2.0\times10^{-4}$). Functional analysis showed that rs6572 did not affect *RAC2* splicing or gene expression based on genotype, and imputation analysis did not provide further information regarding the possible casual variants (data not shown).

In a secondary analysis we further examined the association of *RAC2* with CD by disease phenotype using the IBD Montreal Classification criteria.[16] In the combined analysis of the discovery, replication and validation cohorts (1790 patients with CD, 1636 with UC and 2908 controls) we found further associations with the *RAC2* SNPs, with the strongest signal in ileocolonic CD (Montreal Classification L3) with two SNPs: rs6572 ($p_{combined}=6.0\times10^{-6}$, OR 1.3 (95% CI 1.16 to 1.46)) and rs5756564 ($p_{combined}=3.0\times10^{-5}$, OR 0.77 (95% CI 0.68 to 0.87); see tables 5 and 6 in online supplement for summary). The association with *RAC2* SNPs and phenotypes was not replicated in all three cohorts examined.

Replication of *NCF4*

We also replicated the association between the *NCF4* rs4821544 and CD ($p=5.1\times10^{-3}$, OR 1.25 (95% CI 1.07 to 1.1.47); table 1). Figure 1 in the online supplement shows the LD plot of *NCF4* tag SNPs used in this study. As the association with *NCF4* rs4821544 was originally reported in an ileal CD GWAS and replication

Table 3 Combined analysis for discovery, NIDDK replication and Scottish validation cohorts showing association between *RAC2* SNPs and Crohn's disease

RAC2 gene SNP	Chr 22 Position	Discovery Toronto cohort		NIDDK cohort		Scottish cohort		Combined analysis	
		p Value	OR (95% CI)	p Value	OR (95% CI)	p Value	OR (95% CI)	p Value	OR (95% CI)
rs5756564	35 942 049	6.5×10^{-4}	0.75 (0.64 to 0.88)	5.0×10^{-2}	0.84 (0.74 to 0.1.0)	5.4×10^{-2}	0.95 (0.83 to 1.10)	4.0×10^{-4}	0.85 (0.78 to 0.93)
rs6572	35 951 391	3.8×10^{-4}	1.33 (1.13 to 1.57)	5.0×10^{-1}	1.04 (0.96 to 1.22)	1.0×10^{-2}	1.18 (1.03 to 1.36)	5.0×10^{-5}	1.19 (1.09 to 1.29)
rs9607431	35 959 884	8.5×10^{-5}	1.57 (1.25 to 1.98)	2.7×10^{-2}	0.75 (0.58 to 0.96)	5.4×10^{-1}	1.06 (0.87 to 1.30)	4.2×10^{-2}	1.36 (1.01 to 1.25)
rs2239774	35 967 599	1.5×10^{-4}	1.49 (1.21 to 1.84)	4.4×10^{-2}	0.78 (0.62 to 0.99)	6.9×10^{-2}	1.03 (0.86 to 1.25)	5.0×10^{-2}	1.12 (0.96 to 1.21)

p Values are presented as uncorrected.
Combined analysis of the discovery, replication and validation cohorts (1790 patients with Crohn's disease and 2908 controls).

study,[11 12] we examined ileal CD and also showed an association (Montreal Classification L1, $p=3.5\times10^{-3}$, OR 1.46 (95% CI 1.13 to 1.89); see tables 5 and 7 in online supplement for subphenotype analysis). The patients used in this study did not contribute to the original ileal CD NIDDK GWAS cohort that initially described the *NCF4* association signal.[11] The rs4821544 SNP did not alter gene expression or splicing of *NCF4* based on genotype, and imputation analysis did not provide further information regarding the possible casual variants (data not shown).

DISCUSSION

Phagocytic leucocytes (especially neutrophils) are critical for innate immunity and clear debris and prevent local damage in the bowel. The importance of phagocyte function in the development of non-infectious colitis is seen in a number of congenital disorders,[26] particularly the impaired respiratory burst demonstrated in glycogen storage disease Ib and CGD.[27–29] Patients with CGD often develop severe perianal disease, stricturing small bowel disease and discontinuous colitis that is endoscopically and histologically very similar to CD[5] and even misdiagnosed as CD.[30] A defective neutrophil respiratory burst has long been recognised as playing an important role in the pathogenesis of CD,[8–10] but only recently have NADPH oxidase complex genes been implicated. We observed an association between both *NCF2* and *RAC2* SNPs with clinical phenotypes that resemble that seen in CGD. We chose to examine the VEO-IBD age group closely as patients diagnosed before the age of 10 years have a distinct phenotype with a high familial aggregation and greater tendency to present with severe colonic disease.[31–33] This age group is most likely influenced by genetic alterations such as those seen in *IL10RA/B* and in other phenocopies of IBD such as CGD.[34] We hypothesised that polymorphisms in NADPH oxidase complex genes that reduce NADPH oxidase function but do not cause CGD are associated with an increased risk of developing IBD. We identified a novel variant in *NCF2* (encoding $p67^{phox}$). Rare homozygous mutations in *NCF2* are known to cause autosomal recessive CGD mostly through complete loss of $p67^{phox}$ expression and an inability to produce ROS, and these patients with CGD often present with a milder phenotype.[29 35 36] The identified variant arginine at position 38 of $p67^{phox}$ has been shown to be important in RAC2 binding[25]; variation at this position reduces RAC binding. As $p67^{phox}$ binding to RAC2 is an essential step in the assembly and activation of NOX2 NADPH oxidase,[3] we would expect that ROS production would be reduced but not absent as observed in our index patient. The patients with VEO-IBD with the $p67^{phox}$ R38Q variant all had extensive colonic involvement and five patients developed perianal disease. This phenotype resembles the colitis seen in CGD and in many patients with VEO-IBD. However, these patients had no evidence of immunodeficiency (as assessed by a comprehensive immunological investigation), indicating that the $p67^{phox}$ R38Q variant plays a role in the development of colitis only. Certainly, patients with CGD can present with CD-like colitis as the only manifestation of disease,[37 38] and patients with CGD are at much greater risk of developing CD-like colitis.[5–7] Interestingly, two patients were homozygotes for the $p67^{phox}$ R38Q variant. Neither patient had chronic infections or evidence to suggest CGD, and neither had worse disease at presentation. Long-term follow-up will be required to determine if there is a dosage effect, with more severe outcomes for individuals carrying two copies of this variant.

In this study we also identified the $p67^{phox}$ binding partner RAC2 as a CD susceptibility gene through a candidate gene approach. Although *RAC2* mutations are not known to cause CGD, a rare dominant negative mutation in *RAC2* has been shown to cause an immunodeficiency similar to CGD through inhibiting RAC2 from binding GTP.[39 40] The binding of RAC2 to $p67^{phox}$ is a critical step required for the activation of the NADPH oxidase complex and the production of ROS.[41] We also replicated the association signal for *NCF4* (encoding $p40^{phox}$) and CD. The $p40^{phox}$ protein serves to enhance delivery of $p47^{phox}$ and $p67^{phox}$ to the membrane.[3] Recently, a paediatric patient was found to have novel compound heterozygote *NCF4* mutations which reduced the gene product $p40^{phox}$ binding to PtdIns(3)P, essential for NADPH oxidase phagocytosis-induced oxidant production in human neutrophils.[38] Interestingly, this patient presented with granulomatous colitis and perianal disease without evidence of immune deficiency.[38] These studies demonstrate the importance of $p40^{phox}$ in the development of colitis.

Current estimates suggest that CD GWAS have explained 23% of the inherited contribution to the risk of CD.[13] Thus, there remains substantial 'missing heritability' and it is broadly anticipated that this will not be fully explained by simply expanding the GWAS approach with larger numbers of patients.[13] Our candidate gene approach suggests an association between *RAC2* and CD, although these results need to be interpreted with caution as no *RAC2* SNP replicated in all three cohorts examined and the association signal did not meet genome-wide significance. We observed both risk and protective signals for *RAC2* suggesting that, similar to the IL23R signal, these SNPs are involved in alternate splicing or altered expression of *RAC2* that may play opposing roles.[42] However, this association was not reported in the International IBD Genetics Consortium (IIBDGC) meta-analysis as the gene is poorly tagged in that genome-wide array.[13] The SNPs used in that study did not provide adequate coverage for our lead SNP (GWAS SNPs rs132515, rs2413552 and rs5757362 $r^2<0.25$ to rs6572); further large-scale analysis will therefore be required to validate these results. Our lead *NCF4* SNP (rs4821544) was associated with CD in the NIDDK GWAS ($p_{discovery}=2.89\times10^{-5}$; $p_{replication}=0.009$; OR 1.19) which only included patients with ileal disease and was independently replicated in ileal CD.[11 12] This association was not observed in a large European study.[43] In the first IIBDGC meta-analysis (including all disease locations of CD) the $p_{meta\text{-}analysis}$ was

0.0078 for rs48215442; hence, this SNP was not taken forward for independent replication. In the expanded second IIBDGC meta-analysis, $p_{meta\text{-}analysis}$ for rs4821544 was 1.80×10^{-5}.[13] It is therefore possible that *NCF4* SNP is associated with CD and that disease location may be important for further large-scale replication, although previous European studies suggest that geographical factors may also be important.

Recent studies have provided insight into the role of the NOX2 NADPH oxidase complex in the development of colitis. Animal models have shown that defects in NADPH oxidase complex genes can impair bacterial killing[44] and cause susceptibility to autoimmunity through defects in ROS production.[45] Furthermore, studies have shown that bacterial pathogens can cause excessive ROS production leading to disease through tissue destruction and inflammation and, on the other hand, other bacteria can modulate ROS-dependent neutrophil apoptosis thereby surviving and causing disease.[46] In human studies the non-infectious inflammation (especially colitis) that occurs in patients with CGD has been linked to the NADPH oxidase regulation of the inflammosome[47] and autophagy[48] pathways involved in the pathogenesis of CD. It is therefore plausible that polymorphisms that increase or decrease ROS production may be associated with the development of human colitis.

The NADPH oxidase complex has been implicated in the pathogenesis of human disease from arthritis to cancer.[46] Overall, our studies of the NADPH oxidase complex provide further evidence supporting defective neutrophil ROS production in the pathogenesis of CD. The $p67^{phox}$ R38Q variant suggests that reduced ROS is important in the development of early onset disease, while the candidate gene approach shows the importance of the NADPH oxidase complex in the development of CD. Our study demonstrates that SNPs in the core NADPH oxidase complex genes are not associated with IBD. This is an interesting finding as variations in genes that are responsible for localisation of the NADPH oxidase complex (including $p47^{phox}$ and $p67^{phox}$ and RAC2) appear to be associated with IBD, and this mislocalisation or delay in localisation of the NADPH oxidase complex to the membrane may be important in the pathogenesis of IBD.

We propose that partial impairment of NADPH oxidase function coupled with other genetic variations, especially those in innate immunity genes, contribute to the pathogenesis of CD without development of the severe innate immune deficiency observed in CGD.

Author affiliations
[1]Division of Gastroenterology, Hepatology, and Nutrition, Department of Paediatrics, Hospital for Sick Children, Toronto, Ontario, Canada
[2]Program in Cell Biology, Hospital for Sick Children, University of Toronto, Toronto, Ontario, Canada
[3]Institute of Medical Science, University of Toronto, Toronto, Ontario, Canada
[4]Princess Margaret Hospital, Toronto, Ontario, Canada
[5]Dalla Lana School of Public Health, University of Toronto, Toronto, Ontario, Canada
[6]The Centre for Applied Genomics and Program in Genetics and Genome Biology, The Hospital for Sick Children, Toronto, Ontario, Canada
[7]Division of Pediatrics; Gastroenterology, Hepatology and Nutrition Unit, Pontificia Universidad Católica de Chile, Chile
[8]Department of Paediatric Gastroenterology, Royal Hospital for Sick Children, Glasgow
[9]Faculty of Dentistry, University of Toronto, Toronto, Ontario, Canada
[10]Division of Gastroenterology, Hepatology and Nutrition, Department of Medicine, School of Medicine, University of Pittsburgh, Pittsburgh, USA
[11]Department of Human Genetics, Graduate School of Public Health, University of Pittsburgh, Pittsburgh, USA
[12]Inflammatory Bowel Disease Centre, Departments of Medicine and Genetics, Yale University, New Haven, Connecticut, USA
[13]Department of Gastroenterology, Western General Hospital, University of Edinburgh, Edinburgh, UK
[14]Child Life and Health, Department of Pediatrics, University of Edinburgh, Edinburgh, UK
[15]Mount Sinai Hospital Inflammatory Bowel Disease Group, Zane Cohen Centre for Digestive Diseases, University of Toronto, Toronto, Ontario, Canada
[16]Department of Molecular Genetics, University of Toronto, Toronto, Ontario, Canada

Acknowledgements The authors thank ET and his family, and all IBD patients and families for participating in this research. The authors also thank the NIDDK IBD Genetics Consortium for providing control samples and the International IBD Genetics Consortium (IIBDGC) for providing genotyping data for *RAC2* and *NCF4*. They also acknowledge the work of Karoline Fiedler at the Hospital for Sick Children, Joanne Stempak at Mount Sinai and Dr Elaine Nimmo, Dr Johan Van Limbergen and Hazel Drummond in Edinburgh.

Funding AMM is supported by a transition award from the Crohn's and Colitis Foundation of Canada (CCFC)/Canadian Association of Gastroenterology (CAG)/Canadian Institute for Health Research (CIHR), a Canadian Child Health Clinician Scientist Program (Strategic Training Initiatives in Health Research Program—CIHR) Award, an Early Researcher Award from the Ontario Ministry of Research and Innovation and a CDHNF/NASPGHAN George Ferry Young Investigator Development Award. DCW is the holder of a Medical Research Council Patient Cohorts Research Initiative award (G0800675). Financial assistance was also provided by the Wellcome Trust Programme Grant (072789/Z/03/Z), Action Medical Research, the Chief Scientist Office of the Scottish Government Health Department and the GI/Nutrition Research Fund, Child Life and Health, University of Edinburgh. RHD is supported by NIH/NIDDK grant (DK062420). ADP holds a Canada Research Chair in Genetics of Complex Diseases. MSS is supported by the Gale and Graham Wright Research Chair in Digestive Diseases at Mount Sinai Hospital and funding from CCFC and NIH/NIDDK (DK-06-2423). JHB holds an Investigator in the Pathogenesis of Infectious Disease Award from the Burroughs Wellcome Fund. Funding was provided by a CIHR operating grant (MOP97756) to AMM and JHB.

Competing interests None.

Ethics approval Hospital for Sick Children REB.

Contributors AMM, JHB and MSS conceived and designed all experiments. MSS, AMM, AMG, RHD, RR, JC, DCW and JS provided study samples. WX, TW and ADP analysed the data. VMW, GL, JCG, MS, RF, GL, JB, RM and CH performed functional analysis under supervision of AMM, JHB, MG. AMM wrote the manuscript with JHB and MSS and contributions from all authors.

Provenance and peer review Not commissioned; externally peer reviewed.

REFERENCES

1. **Abraham C,** Cho JH. Inflammatory bowel disease. *N Engl J Med* 2009;**361**:2066—78.
2. **Barrett JC,** Hansoul S, Nicolae DL, *et al*. Genome-wide association defines more than 30 distinct susceptibility loci for Crohn's disease. *Nat Genet* 2008;**8**:955—62.
3. **Nauseef WM.** Biological roles for the NOX family NADPH oxidases. *J Biol Chem* 2008;**283**:16961—5.
4. **Heyworth PG,** Cross AR, Curnutte JT. Chronic granulomatous disease. *Curr Opin Immunol* 2003;**15**:578—84.
5. **Marks DJ,** Miyagi K, Rahman FZ, *et al*. Inflammatory bowel disease in CGD reproduces the clinicopathological features of Crohn's disease. *Am J Gastroenterol* 2009;**104**:117—24.
6. **Schappi MG,** Smith VV, Goldblatt D, *et al*. Colitis in chronic granulomatous disease. *Arch Dis Child* 2001;**84**:147—51.
7. **Werlin SL,** Chusid MJ, Caya J, *et al*. Colitis in chronic granulomatous disease. *Gastroenterology* 1982;**82**:328—31.
8. **Ward M.** The pathogenesis of Crohn's disease. *Lancet* 1977;**2**:903—5.
9. **Verspaget HW,** Mieremet-Ooms MA, Weterman IT, *et al*. Partial defect of neutrophil oxidative metabolism in Crohn's disease. *Gut* 1984;**25**:849—53.
10. **Segal AW,** Loewi G. Neutrophil dysfunction in Crohn's disease. *Lancet* 1976;**2**:219—21.
11. **Rioux JD,** Xavier RJ, Taylor KD, *et al*. Genome-wide association study identifies new susceptibility loci for Crohn disease and implicates autophagy in disease pathogenesis. *Nat Genet* 2007;**39**:596—604.
12. **Roberts RL,** Hollis-Moffatt JE, Gearry RB, *et al*. Confirmation of association of IRGM and NCF4 with ileal Crohn's disease in a population-based cohort. *Genes Immun* 2008;**9**:561—5.
13. **Franke A,** McGovern DP, Barrett JC, *et al*. Genome-wide meta-analysis increases to 71 the number of confirmed Crohn's disease susceptibility loci. *Nat Genet* 2010;**42**:1118—25.
14. **Anon.** The International HapMap Project. *Nature* 2003;**426**:789—96.
15. **de Bakker PI,** Yelensky R, Pe'er I, *et al*. Efficiency and power in genetic association studies. *Nat Genet* 2005;**37**:1217—23.
16. **Silverberg MS,** Satsangi J, Ahmad T, *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* 2005;**19**(Suppl A):5—36.
17. **Silverberg MS,** Cho JH, Rioux JD, *et al*. Ulcerative colitis-risk loci on chromosomes 1p36 and 12q15 found by genome-wide association study. *Nat Genet* 2009;**41**:216—20.

18. **Imielinski M,** Baldassano RN, Griffiths A, *et al*. Common variants at five new loci associated with early-onset inflammatory bowel disease. *Nat Genet* 2009;**41**:1335—40.
19. **Barrett JC,** Fry B, Maller J, *et al*. Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics* 2005;**21**:263—5.
20. **Purcell S,** Neale B, Todd-Brown K, *et al*. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* 2007;**81**:559—75.
21. **Freidlin B,** Zheng G, Li Z, *et al*. Trend tests for case-control studies of genetic markers: power, sample size and robustness. *Hum Hered* 2002;**53**: 146—52.
22. **Clayton D.** Testing for association on the X chromosome. *Biostatistics* 2008;**9**:593—600.
23. **Wellcome Trust Case Control Consortium.** Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature* 2007;**447**:661—78.
24. **Levine A,** Griffiths A, Markowitz J, *et al*. Pediatric modification of the Montreal classification for inflammatory bowel disease: the Paris classification. *Inflamm Bowel Dis* 2011;**17**:1314—21.
25. **Koga H,** Terasawa H, Nunoi H, *et al*. Tetratricopeptide repeat (TPR) motifs of p67 (phox) participate in interaction with the small GTPase Rac and activation of the phagocyte NADPH oxidase. *J Biol Chem* 1999;**274**:25051—60.
26. **Rahman FZ,** Marks DJ, Hayee BH, *et al*. Phagocyte dysfunction and inflammatory bowel disease. *Inflamm Bowel Dis* 2008;**14**:1443—52.
27. **Couper R,** Kapelushnik J, Griffiths AM. Neutrophil dysfunction in glycogen storage disease Ib: association with Crohn's-like colitis. *Gastroenterology* 1991;**100**:549—54.
28. **Ament ME,** Ochs HD. Gastrointestinal manifestations of chronic granulomatous disease. *N Engl J Med* 1973;**288**:382—7.
29. **Stasia MJ,** Li XJ. Genetics and immunopathology of chronic granulomatous disease. *Semin Immunopathol* 2008;**30**:209—35.
30. **Freudenberg F,** Wintergerst U, Roesen-Wolff A, *et al*. Therapeutic strategy in p47-phox deficient chronic granulomatous disease presenting as inflammatory bowel disease. *J Allergy Clin Immunol* 2011;**125**:943—6.e1.
31. **Paul T,** Birnbaum A, Pal DK, *et al*. Distinct phenotype of early childhood inflammatory bowel disease. *J Clin Gastroenterol* 2006;**40**:583—6.
32. **Griffiths AM.** Specificities of inflammatory bowel disease in childhood. *Best Pract Res Clin Gastroenterol* 2004;**18**:509—23.
33. **Heyman MB,** Kirschner BS, Gold BD, *et al*. Children with early-onset inflammatory bowel disease (IBD): analysis of a pediatric IBD consortium registry. *J Pediatr* 2005;**146**:35—40.
34. **Glocker EO,** Kotlarz D, Boztug K, *et al*. Inflammatory bowel disease and mutations affecting the interleukin-10 receptor. *N Engl J Med* 2009;**361**:2033—45.
35. **Gentsch M,** Kaczmarczyk A, van Leeuwen K, *et al*. Alu-repeat-induced deletions within the NCF2 gene causing p67-phox-deficient chronic granulomatous disease (CGD). *Hum Mutat* 2011;**31**:151—8.
36. **Francke U,** Hsieh CL, Foellmer BE, *et al*. Genes for two autosomal recessive forms of chronic granulomatous disease assigned to 1q25 (NCF2) and 7q11.23 (NCF1). *Am J Hum Genet* 1990;**47**:483—92.
37. **Marciano BE,** Rosenzweig SD, Kleiner DE, *et al*. Gastrointestinal involvement in chronic granulomatous disease. *Pediatrics* 2004;**114**:462—8.
38. **Matute JD,** Arias AA, Wright NA, *et al*. A new genetic subgroup of chronic granulomatous disease with autosomal recessive mutations in p40 phox and selective defects in neutrophil NADPH oxidase activity. *Blood* 2009;**114**:3309—15.
39. **Williams DA,** Tao W, Yang F, *et al*. Dominant negative mutation of the hematopoietic-specific Rho GTPase, Rac2, is associated with a human phagocyte immunodeficiency. *Blood* 2000;**96**:1646—54.
40. **Ambruso DR,** Knall C, Abell AN, *et al*. Human neutrophil immunodeficiency syndrome is associated with an inhibitory Rac2 mutation. *Proc Natl Acad Sci U S A* 2000;**97**:4654—9.
41. **Nauseef WM.** Assembly of the phagocyte NADPH oxidase. *Histochem Cell Biol* 2004;**122**:277—91.
42. **Duerr RH,** Taylor KD, Brant SR, *et al*. A genome-wide association study identifies IL23R as an inflammatory bowel disease gene. *Science* 2006;**314**:1461—3.
43. **Glas J,** Seiderer J, Pasciuto G, *et al*. rs224136 on chromosome 10q21.1 and variants in PHOX2B, NCF4, and FAM92B are not major genetic risk factors for susceptibility to Crohn's disease in the German population. *Am J Gastroenterol* 2009;**104**:665—72.
44. **Ellson CD,** Davidson K, Ferguson GJ, *et al*. Neutrophils from p40phox-/- mice exhibit severe defects in NADPH oxidase regulation and oxidant-dependent bacterial killing. *J Exp Med* 2006;**203**:1927—37.
45. **Hultqvist M,** Olofsson P, Holmberg J, *et al*. Enhanced autoimmunity, arthritis, and encephalomyelitis in mice with a reduced oxidative burst due to a mutation in the Ncf1 gene. *Proc Natl Acad Sci U S A* 2004;**101**:12646—51.
46. **Quinn MT,** Ammons MC, Deleo FR. The expanding role of NADPH oxidases in health and disease: no longer just agents of death and destruction. *Clin Sci (Lond)* 2006;**111**:1—20.
47. **Meissner F,** Seger RA, Moshous D, *et al*. Inflammasome activation in NADPH oxidase defective mononuclear phagocytes from patients with chronic granulomatous disease. *Blood* 2010;**116**:1570—3.
48. **Huang J,** Canadien V, Lam GY, *et al*. Activation of antibacterial autophagy by NADPH oxidases. *Proc Natl Acad Sci U S A* 2009;**106**:6226—31.

NADPH oxidase complex and IBD candidate gene studies: identification of a rare variant in *NCF2* that results in reduced binding to RAC2

Aleixo M Muise, Wei Xu, Cong-Hui Guo, et al.

Gut published online September 7, 2011
doi: 10.1136/gutjnl-2011-300078

Notes

Canadian NEOPICS Consortium

Name	Institution
Dr. Jeff Critch MD, MSc, FRCPC Associate Professor of Pediatrics, Memorial University	Janeway Children's Health Centre, St John's Newfoundland
Anthony Otley MD, MSc FRCPC Associate Professor of Pediatrics, Dalhousie University	IWK Halifax, Nova Scotia
Sylviane Forget MD, MSc, FRCPC Assistant Professor of Pediatrics, McGill University Health Center	Montreal Children's Hospital Montreal, Quebec
Colette Deslandres MD Associate Professor of Pediatrics, University of Montreal	Sainte-Justine Hospital Montreal, Quebec
Eric Benchimol MD, PhD, FRCPC, Clinician-Scientist, Assistant Professor of Pediatrics, University of Ottawa	Children's Hospital of Eastern Ontario Ottawa, Ontario
David Mack MD, FRCPC Professor, Department of Pediatrics and Biochemistry, Microbiology and Immunology, Faculty of Medicine, University of Ottawa	Children's Hospital of Eastern Ontario Ottawa, Ontario
Anne Griffiths MD, MSc, FRCPC Professor of Pediatrics, University of Toronto, Associate Scientist, Genetics and Genome Biology Program, HSC Research Institute	Hospital for Sick Children Toronto, Ontario
John Brumell PhD Senior-Scientist, Cell Biology, HSC Research Institute, Associate Professor of Molecular Genetics, University of Toronto	Hospital for Sick Children Toronto, Ontario
Nicola Jones MD, PhD, FRCPC, Clinician-Scientist, Professor of Pediatrics, University of Toronto, Senior-Scientist, Cell Biology, HSC Research Institute	Hospital for Sick Children Toronto, Ontario
Thomas Walters MD, Assistant Professor of Pediatrics, University of Toronto	Hospital for Sick Children Toronto, Ontario
Aleixo Muise MD, PhD, FRCPC Clinician-Scientist, Assistant Professor of Pediatrics, University of Toronto, Scientist, Cell Biology, HSC Research Institute	Hospital for Sick Children Toronto, Ontario
Robert Issenman MD, FRCPC Professor of Pediatrics, McMaster University	McMaster Children's Hospital Hamilton, Ontario
Kevin Bax, MD, FAAP Assistant Professor of Pediatrics, University of Western Ontario	University of Western Ontario, Children's Hospital, London Health Sciences Centre
Iwona Wrobel MD, FRCPC Assistant Professor of Pediatrics, University of Calgary	Alberta Children's Hospital Calgary, Alberta
Christopher Waterhouse MD, PhD, FRCPC, Clinician-Scientist Assistant Professor of Pediatrics, University of Calgary	Alberta Children's Hospital Calgary, Alberta
Hien Huynh MD, FRCPC, FRACP Associate Professor, Department of Pediatrics, University of Alberta	Stollery Children's Hospital Edmonton, Alberta
Eytan Wine MD, PhD, Clinician-Scientist, Assistant Professor of Pediatrics, University of Alberta	Stollery Children's Hospital Edmonton, Alberta
Kevan Jacobson MBBCh, FRCPC Associate Professor of Pediatrics and Senior Clinician Scientist, University of British Columbia	British Columbia's Children's Hospital Vancouver, British Columbia
Dan Turner MD, PhD (Clinical Epidemiology) Shaare Zedek Medical Center, Jerusalem, Israel	Shaare Zedek Medical Center Jerusalem, Israel

Muise et al Supplemental Material: Methods, Tables, and Figures

Methods

Binding Studies

Human HEK293 cells maintained in Dulbecco's modified Eagle's medium (Invitrogen) containing 10% heat-inactivated fetal bovine serum (Invitrogen), antibiotics (100 units/ml penicillin and 100 μg/ml streptomycin) at 37 °C in 5% CO_2. Monoclonal anti-Myc antibody was purchased from Upstate. Monoclonal anti-GFP antibody was from Covance. ECL reagents were from Pierce, and Protein G- beads were from Sigma. Plasmid Construction. Myc-tagged human P67 cDNA was cloned into the pCDNA3 vector (Invitrogen). Mutation in R38Q (c:113G>A) in the RAC binding domain was generated by site-directed mutagenesis using PCR. All constructs were verified by sequencing. Empty GFP vector was used as negative control. Transient Transfections—HEK293 cells were plated at 1.5×10^6 cells/well in 10-cm dish and grown overnight in Dulbecco's modified Eagle's medium with 10% fetal bovine serum, 100 units/ml penicillin, and 100 μg/ml streptomycin to reach to 60-70% confluence. Cells were co-transfected with constructs harboring the inserted gene using calcium phosphate according to the manufacturer's instructions. Six hours after transfection the media was changed and cells were collected at the indicated time.

Cells grown in 10 cm dishes washed twice with cold 1× phosphate-buffered saline, and extracted in ice-cold lysis buffer (50 mM Hepes, pH 7.5, 150 mM NaCl, 1.5 mM $MgCl_2$, 1 mM EGTA, 10% glycerol (v/v), 1% Triton X-100 (v/v), supplemented with 1 mm phenylmethylsulfonyl fluoride, 1 μg/ml aprotinin, 1 μg/ml leupeptin, 1 μg/ml pepstatin, and 1 mM vanadate). Equal amounts (50 μg) of total protein were analyzed by western blotting. For protein immunoprecipitation experiments, supernatants (1mg) were incubated using both anti-myc and GFP antibody (1 μl, 1mg/ml).

Genetic Quality Control

To conduct systematic quality control on the raw genotyping data, we examined 303 SNPs genotyped for the discovery cohort only. Most of SNPs were not part of this candidate study (either known IBD SNPs or from other independent studies) but were genotyped using the same platform and therefore allowed detailed QC to be carried out. To

reduce population stratification, we limited the analysis to Caucasian subjects. We excluded 4 of 303 SNPs with a genotype call rate of less than 95% and 74 individuals where removed due to an overall genotype completion rate less than 95%. No individual was excluded due to sex discrepancies based on the heterozygosity rate from SNPs on chromosome. A total of 19 SNPs were removed as a result of a genotype distribution that deviated from expected by Hardy-Weinberg Equilibrium (HWE) among controls ($P < 0.001$). A total of 27 SNPs with a minor allele frequency < 1% in cases and controls were excluded from the analysis. After QC filtering, 253 SNPs were analyzed and the median sample call rate was 98.3%, and the median marker call rate was 99.7%.

Supplemental Table 1: Primers for *NCF2, RAC2* and *NCF4*.

Gene Name	Gene	primers	forward	reverse
NCF2-1	NM_001127651	244-1339	AGC TCT CTT GGC CTC CTA GTT TCT	TTG AGT GTG TAG GGC ATG GGA ACA
NCF2-2		812-1854	GGG CAA GCT GTT TCG ACC AAA TGA	AGA CTT CTC TCC GAG TGC TTT CCA
NCF4	NM_000631.4	170-1210	AAGCTTATGGCTGTGGCCCAGCAG	CAGCTCATGGCATCGTGTTGTAGAC
RAC2	NM_002872	118-696	ATGCAGGCCATCAAGTG	TAGAGGAGGCTGCAGGCGCGC

Supplemental Table 2: Demographic data for discovery, North American and Scottish replication cohorts

Supplemental Table 2A. The summary statistics for discovery cohort

Characteristics	CD (n=754)	UC (n=603)	HC (n=924)
Gender			
Male	395	279	332
Female	359	324	592
Age at diagnosis (median, range)	16 (2-62)	23 (1-73)	n/a
Early Onset (<16)	441	207	n/a
Jewish Heritage			
No	590	498	856
Yes	156	104	68
Perianal (p)			
No	527	n/a	n/a
Yes	227		
Location			
0	57	n/a	
1	181		n/a
2	191		
3	321		
Ileal Only CD (L1)	181	n/a	n/a
Any Ileal CD (L1/L3)	502	n/a	n/a
L2 CD + UC	191	603	n/a
L2/L3 CD + UC	512	603	n/a
Behavior			
1	366	n/a	n/a
2	186		
3	195		

The discovery cohort prior to QC includes 2281 Caucasian subjects (754 CD, 603 UC, and 924 controls). Location and Behavior based on Montreal Classification. ileal only (L1), colonic only (L2), ileo-colonic (L3), ileal any (L1/L3), colon only (L2 plus UC), colon any (L2/L3 plus UC), perianal (p), and early diagnosis patients (diagnosis age < 16) Nonstricturing/nonpenetration (B1), Stricturing (B2), and Penetration (B3). n/a – not applicable.

Supplemental Table 2B. The summary statistics for North America replication cohort

Characteristics	CD (n=443)	UC (n=477)	HC (n=916)
Gender			
Male	191	275	559
Female	252	202	357
Age at diagnosis (median, range)	24 (7-71)	30 (2-82)	n/a
Early Onset (<16)	70	46	n/a
Perianal (p)			
No	293	n/a	n/a
Yes	143		
Ileal Only CD (L1)	136	n/a	n/a
Any Ileal CD (L1/L3)	345	n/a	n/a
L2 CD + UC	84	477	n/a
L2/L3 CD + UC	293	477	n/a
Behavior			
1	187	n/a	n/a
2	110		
3	136		

The replicate cohort includes 1836 Caucasian subjects (443 CD, 477 UC, and 916 controls). Location and Behavior based on Montreal Classification. Ileal only (L1), colonic only (L2), ileo-colonic (L3), ileal any (L1/L3), colon only (L2 plus UC), colon any (L2/L3 plus UC), perianal (p), and early diagnosis patients (diagnosis age < 16) Nonstricturing/nonpenetration (B1), Stricturing (B2), and Penetration (B3). n/a – not applicable.

Supplemental Table 2C. The summary statistics for Scottish validation cohort

Characteristics	CD (n=691)	UC (n=615)	HC (n=1143)
Gender Male Female	 317 374	 302 313	 n/a
Age at diagnosis (median, range)	20 (2-83)	30 (2-83)	n/a
Early onset (<16)	322	46	n/a
Perianal (p) No Yes	 615 76	n/a	n/a
Ileal Only CD (L1)	159	n/a	n/a
Any Ileal CD (L1/L3)	413	n/a	n/a
L2 CD + UC	241	615	n/a
L2/L3 CD + UC	495	615	n/a
Behavior 1 2 3	 556 45 40	n/a	n/a

The Scotland cohort includes 2449 Caucasian subjects (691 CD, 615 UC, and 1143 controls). Location and Behavior based on Montreal Classification. ileal only (L1), colonic only (L2), ileo-colonic (L3), ileal any (L1/L3), colon only (L2 plus UC), colon any (L2/L3 plus UC), perianal (p), and early diagnosis patients (diagnosis age < 16) Nonstricturing/nonpenetration (B1), Stricturing (B2), and Penetration (B3). n/a – not applicable.

CHR	SNP	Gene	position	A1	A2	MAF	F_A	F_U	P	OR (ci=0.95)	L95	U95
1	rs1044879	NCF2	181789375	3	2	0.351173	0.348	0.3557	0.4905	0.9518	0.8272	1.095
1	rs789180	NCF2	181792243	4	1	0.105957	0.1075	0.1038	0.5588	1.068	0.8565	1.332
1	rs789181	NCF2	181792961	3	1	0.14202	0.1388	0.1466	0.6481	0.9579	0.7964	1.152
1	rs3845461	NCF2	181798823	1	3	0.0163494	0.0175	0.01472	0.5244	1.191	0.695	2.042
1	rs11811630	NCF2	181800533	1	3	0.449488	0.4492	0.4499	0.95	0.9957	0.8698	1.14
1	rs10797887	NCF2	181811317	1	3	0.0539287	0.055	0.05241	0.547	1.097	0.8116	1.483
1	rs12753665	NCF2	181812684	1	3	0.3204	0.3187	0.3227	0.9117	0.9918	0.8578	1.147
1	rs12094228	NCF2	181814495	2	1	0.0580771	0.05792	0.0583	0.8347	1.032	0.7707	1.381
1	rs3845466	NCF2	181818383	1	3	0.0807711	0.08292	0.07774	0.4482	1.101	0.8581	1.414
1	rs10911364	NCF2	181819149	3	1	0.0502684	0.04917	0.05183	0.6873	0.9378	0.6858	1.282
1	rs789185	NCF2	181820904	3	1	0.38409	0.3917	0.3734	0.1775	1.099	0.9581	1.261
1	rs11578964	NCF2	181822402	1	3	0.417277	0.4229	0.4093	0.2174	1.089	0.951	1.247
1	rs789191	NCF2	181822502	1	3	0.0248902	0.02542	0.02415	0.6334	1.109	0.7243	1.699
1	rs11588654	NCF2	181823018	3	1	0.341142	0.345	0.3357	0.4664	1.054	0.915	1.214
1	rs10911365	NCF2	181825728	1	3	0.079795	0.08375	0.0742	0.08033	1.247	0.9737	1.596
16	rs3180279	CYBA	87238334	2	3	0.4978	0.4996	0.4953	0.8726	1.011	0.8831	1.158
16	rs9940427	CYBA	87239320	1	2	0.0332031	0.03211	0.03475	0.3995	0.8536	0.5906	1.234
16	rs4673	CYBA	87240737	1	3	0.36164	0.3754	0.3422	0.01146	1.195	1.041	1.372
16	rs3794624	CYBA	87244575	1	3	0.322423	0.3109	0.3386	0.05254	0.8705	0.7566	1.002
22	rs6000448	NCF4	35580717	1	3	0.106445	0.1105	0.1007	0.2867	1.127	0.9044	1.405
22	rs10854693	NCF4	35581956	3	1	0.404588	0.3887	0.427	0.006411	0.8285	0.7237	0.9485
22	rs1883113	NCF4	35582630	3	2	0.0705222	0.07667	0.06184	0.006806	1.454	1.109	1.906
22	rs4821542	NCF4	35582864	1	3	0.157882	0.1592	0.1561	0.7042	1.036	0.8636	1.242
22	rs7287350	NCF4	35583185	1	3	0.0578613	0.06297	0.05065	0.007115	1.508	1.118	2.035
22	rs1883112	NCF4	35586792	1	3	0.424976	0.4108	0.4451	0.008813	0.8358	0.7309	0.9559
22	rs4821544	NCF4	35588449	3	1	0.336426	0.3465	0.3221	0.09234	1.129	0.9803	1.299
22	rs741997	NCF4	35588759	3	1	0.0810151	0.08333	0.07774	0.2644	1.15	0.8999	1.469
22	rs746713	NCF4	35589305	3	1	0.291402	0.2771	0.3115	0.04415	0.8607	0.7437	0.9961
22	rs909484	NCF4	35590547	1	3	0.177648	0.1808	0.1731	0.5071	1.062	0.8899	1.266
22	rs729749	NCF4	35593792	1	3	0.225283	0.2317	0.2162	0.1866	1.113	0.9494	1.305
22	rs2075938	NCF4	35596268	1	3	0.285645	0.284	0.288	0.9643	1.003	0.8651	1.164
22	rs8137456	NCF4	35604389	2	3	0.0697901	0.07542	0.06184	0.1701	1.203	0.9237	1.567
22	rs8137602	NCF4	35604595	3	2	0.0895559	0.0825	0.09953	0.1094	0.8276	0.6566	1.043
22	rs3959633	NCF4	35605501	1	3	0.0236701	0.02625	0.02002	0.2975	1.268	0.8111	1.983
22	rs5756379	NCF4	35605850	3	1	0.343262	0.3461	0.3392	0.5153	1.049	0.909	1.21
22	rs4821554	NCF4	35606030	1	3	0.263787	0.26	0.2691	0.689	0.9691	0.8312	1.13
22	rs5750326	NCF4	35606231	1	2	0.494143	0.4954	0.4923	0.7617	0.9795	0.8567	1.12
22	rs5756564	RAC2	35942049	3	1	0.38897	0.3621	0.427	0.0001631	0.7685	0.6701	0.8812
22	rs4820272	RAC2	35945560	1	3	0.16203	0.1758	0.1425	0.01319	1.266	1.051	1.525
22	rs933222	RAC2	35948583	1	3	0.287945	0.3	0.2709	0.04235	1.17	1.005	1.361
22	rs12166968	RAC2	35951304	2	3	0.0651537	0.06458	0.06596	0.727	0.9532	0.7284	1.247
22	rs6572	RAC2	35951391	3	2	0.45998	0.4871	0.4217	0.0001309	1.309	1.14	1.503
22	rs739041	RAC2	35954945	3	1	0.431127	0.428	0.4356	0.7734	0.9804	0.8567	1.122
22	rs9607431	RAC2	35959884	2	1	0.138116	0.1575	0.1107	0.0002877	1.456	1.188	1.783
22	rs1476002	RAC2	35960734	1	3	0.134521	0.1418	0.1243	0.1935	1.144	0.934	1.401
22	rs13058338	RAC2	35962716	4	1	0.253906	0.2367	0.2783	0.009238	0.8156	0.6995	0.9509
22	rs5756573	RAC2	35963721	1	3	0.149243	0.1511	0.1466	0.4227	1.08	0.8944	1.305
22	rs2284038	RAC2	35965001	3	1	0.396535	0.3987	0.3934	0.8148	1.017	0.8854	1.167

22	rs2239774	RAC2	35967599	2	3	0.164226	0.1821	0.139	0.00224	1.332	1.108	1.601
22	rs2239775	RAC2	35967787	1	2	0.138116	0.1388	0.1372	0.7624	0.9702	0.7971	1.181
22	rs2239773	RAC2	35968235	1	3	0.271598	0.2514	0.2996	0.00207	0.7858	0.6741	0.9161
22	rs2213430	RAC2	35968906	1	3	0.43509	0.4367	0.4329	0.5075	1.048	0.913	1.202
22	rs6000632	RAC2	35974061	1	3	0.243289	0.2375	0.2515	0.4456	0.9406	0.8035	1.101
22	rs4821615	RAC2	35976077	2	3	0.446559	0.4475	0.4452	0.8912	0.9906	0.8658	1.133
22	rs12484031	RAC2	35979216	3	1	0.101807	0.09842	0.1066	0.4141	0.9144	0.7378	1.133
22	rs7288979	RAC2	35979440	3	1	0.172035	0.1604	0.1885	0.02659	0.8219	0.6911	0.9775
X	rs7059081	CYBB	37526428	2	3	0.08321	0.0875	0.07715	0.3336	1.096	0.9097	1.322
X	rs6651773	CYBB	37528612	3	1	0.1652	0.165	0.1655	0.972	0.9975	0.8658	1.149
X	rs17312212	CYBB	37529550	3	1	0.05686	0.05542	0.05889	0.6978	0.9575	0.7691	1.192
X	rs4827298	CYBB	37536603	3	1	0.2469	0.2458	0.2485	0.8703	0.9901	0.8787	1.116
X	rs5917471	CYBB	37537458	1	3	0.4143	0.4175	0.4099	0.6871	1.022	0.9204	1.134
X	rs12848910	CYBB	37551189	3	1	0.1142	0.1121	0.1172	0.6664	0.9642	0.817	1.138

Supplemental Table 3A. Association of NADPH oxidase gene SNPs and IBD in the Discovery cohort. MAF: minor allelic frequency. F_A: frequency in affected patients. F_U: frequency in unaffected controls. L95 and U95: lower and upper confidence interval.

CHR	SNP	Gene	position	A1	A2	MAF	F_A	F_U	P	OR	L95	U95
1	rs1044879	NCF2	181789375	3	2	0.349169	0.3407	0.3557	0.2322	0.9048	0.7678	1.066
1	rs789180	NCF2	181792243	4	1	0.105053	0.1067	0.1038	0.6349	1.063	0.8261	1.368
1	rs789181	NCF2	181792961	3	1	0.143189	0.1387	0.1466	0.7249	0.9625	0.7778	1.191
1	rs3845461	NCF2	181798823	1	3	0.015614	0.01677	0.01472	0.5326	1.217	0.6573	2.252
1	rs11811630	NCF2	181800533	1	3	0.448173	0.4459	0.4499	0.7474	0.9747	0.8342	1.139
1	rs10797887	NCF2	181811317	1	3	0.055149	0.05869	0.05241	0.3318	1.186	0.8401	1.675
1	rs12753665	NCF2	181812684	1	3	0.323588	0.3247	0.3227	0.9595	1.004	0.8484	1.189
1	rs12094228	NCF2	181814495	2	1	0.058804	0.05945	0.0583	0.7232	1.063	0.7579	1.491
1	rs3845466	NCF2	181818383	1	3	0.079734	0.08232	0.07774	0.5068	1.104	0.8244	1.478
1	rs10911364	NCF2	181819149	3	1	0.052159	0.05259	0.05183	0.9873	0.9971	0.6962	1.428
1	rs789185	NCF2	181820904	3	1	0.382724	0.3948	0.3734	0.1795	1.115	0.9509	1.308
1	rs11578964	NCF2	181822402	1	3	0.418605	0.4306	0.4093	0.137	1.127	0.9628	1.318
1	rs789191	NCF2	181822502	1	3	0.023588	0.02287	0.02415	0.8909	1.036	0.6244	1.719
1	rs11588654	NCF2	181823018	3	1	0.344518	0.3559	0.3357	0.2062	1.111	0.9435	1.309
1	rs10911365	NCF2	181825728	1	3	0.077406	0.08155	0.0742	0.1909	1.208	0.9101	1.603
16	rs3180279	CYBA	87238334	2	3	0.491672	0.487	0.4953	0.4364	0.9394	0.8026	1.1
16	rs9940427	CYBA	87239320	1	2	0.033909	0.03282	0.03475	0.4885	0.8589	0.5586	1.321
16	rs4673	CYBA	87240737	1	3	0.351163	0.3628	0.3422	0.112	1.139	0.9701	1.337
16	rs3794624	CYBA	87244575	1	3	0.330452	0.3198	0.3386	0.1634	0.8906	0.7566	1.048
22	rs6000448	NCF4	35580717	1	3	0.108045	0.1176	0.1007	0.1839	1.184	0.9228	1.52
22	rs10854693	NCF4	35581956	3	1	0.405648	0.378	0.427	0.003823	0.7925	0.6769	0.9278
22	rs1883113	NCF4	35582630	3	2	0.067774	0.07546	0.06184	0.01325	1.472	1.084	2
22	rs4821542	NCF4	35582864	1	3	0.165449	0.1776	0.1561	0.1857	1.149	0.9354	1.411
22	rs7287350	NCF4	35583185	1	3	0.054853	0.06031	0.05065	0.02581	1.47	1.048	2.063
22	rs1883112	NCF4	35586792	1	3	0.422821	0.3941	0.4451	0.001388	0.7732	0.6605	0.9053
22	rs4821544	NCF4	35588449	3	1	0.344518	0.3735	0.3221	0.005118	1.257	1.071	1.475
22	rs741997	NCF4	35588759	3	1	0.084385	0.09299	0.07774	0.121	1.243	0.9442	1.636
22	rs746713	NCF4	35589305	3	1	0.306977	0.3011	0.3115	0.7173	0.9696	0.8202	1.146
22	rs909484	NCF4	35590547	1	3	0.171761	0.17	0.1731	0.8553	0.9808	0.7968	1.207
22	rs729749	NCF4	35593792	1	3	0.217158	0.2184	0.2162	0.7703	1.028	0.853	1.239
22	rs2075938	NCF4	35596268	1	3	0.279255	0.2679	0.288	0.3198	0.915	0.7682	1.09
22	rs8137456	NCF4	35604389	2	3	0.068770 8	0.07774	0.06184	0.1693	1.24	0.9125	1.685
22	rs8137602	NCF4	35604595	3	2	0.091694 4	0.08155	0.09953	0.07419	0.7759	0.5872	1.025
22	rs3959633	NCF4	35605501	1	3	0.024916 9	0.03125	0.02002	0.07318	1.56	0.9591	2.537
22	rs5756379	NCF4	35605850	3	1	0.334551	0.3285	0.3392	0.5302	0.9483	0.8033	1.119
22	rs4821554	NCF4	35606030	1	3	0.257475	0.2424	0.2691	0.1157	0.8646	0.7213	1.036
22	rs5750326	NCF4	35606231	1	2	0.496678	0.5023	0.4923	0.9018	1.01	0.8663	1.177
22	rs5756564	RAC2	35942049	3	1	0.397674	0.3598	0.427	0.0006546	0.7562	0.6439	0.888
22	rs4820272	RAC2	35945560	1	3	0.162791	0.189	0.1425	0.001566	1.401	1.137	1.727
22	rs933222	RAC2	35948583	1	3	0.281395	0.295	0.2709	0.2225	1.117	0.9348	1.336
22	rs12166968	RAC2	35951304	2	3	0.067109 6	0.0686	0.06596	0.9971	1.001	0.7332	1.365
22	rs6572	RAC2	35951391	3	2	0.453488	0.4947	0.4217	0.00038	1.337	1.139	1.57

22	rs739041	RAC2	35954945	3	1	0.434492	0.4331	0.4356	0.8488	1.015	0.8678	1.188
22	rs9607431	RAC2	35959884	2	1	0.135548	0.1677	0.1107	0.00008594	1.577	1.256	1.98
22	rs1476002	RAC2	35960734	1	3	0.12766	0.1321	0.1243	0.6058	1.065	0.8383	1.353
22	rs13058338	RAC2	35962716	4	1	0.256981	0.2294	0.2783	0.006391	0.7773	0.6486	0.9316
22	rs5756573	RAC2	35963721	1	3	0.145376	0.1437	0.1466	0.7453	1.037	0.8329	1.291
22	rs2284038	RAC2	35965001	3	1	0.398007	0.404	0.3934	0.6625	1.037	0.8821	1.218
22	rs2239774	RAC2	35967599	2	3	0.164452	0.1974	0.139	0.0001514	1.495	1.214	1.84
22	rs2239775	RAC2	35967787	1	2	0.134884	0.1319	0.1372	0.4163	0.9086	0.7212	1.145
22	rs2239773	RAC2	35968235	1	3	0.277002	0.247	0.2996	0.004515	0.7696	0.6424	0.9221
22	rs2213430	RAC2	35968906	1	3	0.434219	0.436	0.4329	0.6503	1.038	0.8845	1.217
22	rs6000632	RAC2	35974061	1	3	0.245183	0.237	0.2515	0.4481	0.9314	0.7751	1.119
22	rs4821615	RAC2	35976077	2	3	0.447841	0.4512	0.4452	0.8961	1.01	0.8644	1.181
22	rs12484031	RAC2	35979216	3	1	0.100664	0.09299	0.1066	0.2213	0.8521	0.6593	1.101
22	rs7288979	RAC2	35979440	3	1	0.176412	0.1608	0.1885	0.06473	0.8266	0.6754	1.012
X	rs7059081	CYBB	37526428	2	3	0.08272	0.08994	0.07715	0.3038	1.118	0.9042	1.381
X	rs6651773	CYBB	37528612	3	1	0.1621	0.1578	0.1655	0.6278	0.9595	0.812	1.134
X	rs17312212	CYBB	37529550	3	1	0.06047	0.0625	0.05889	0.7375	1.043	0.816	1.333
X	rs4827298	CYBB	37536603	3	1	0.2402	0.2294	0.2485	0.3079	0.9288	0.8057	1.071
X	rs5917471	CYBB	37537458	1	3	0.406	0.4009	0.4099	0.6786	0.9745	0.8623	1.101
X	rs12848910	CYBB	37551189	3	1	0.1169	0.1166	0.1172	0.9664	0.9959	0.8218	1.207

Supplemental Table 3B. Association of NADPH oxidase gene SNPs and CD in the Discovery cohort. MAF: minor allelic frequency. F_A: frequency in affected patients. F_U: frequency in unaffected controls. L95 and U95: lower and upper confidence interval.

CHR	SNP	Gene	position	A1	A2	MAF	F_A	F_U	P	OR (ci=0.95)	L95	U95
1	rs1044879	NCF2	181789375	3	2	0.356115	0.3567	0.3557	0.9082	1.01	0.8518	1.198
1	rs789180	NCF2	181792243	4	1	0.105603	0.1085	0.1038	0.6136	1.072	0.8192	1.402
1	rs789181	NCF2	181792961	3	1	0.143575	0.1388	0.1466	0.6748	0.9517	0.7554	1.199
1	rs3845461	NCF2	181798823	1	3	0.0161522	0.01838	0.01472	0.6547	1.161	0.6034	2.234
1	rs11811630	NCF2	181800533	1	3	0.451184	0.4531	0.4499	0.7968	1.022	0.8655	1.207
1	rs10797887	NCF2	181811317	1	3	0.051687	0.05055	0.05241	0.9811	0.9955	0.6825	1.452
1	rs12753665	NCF2	181812684	1	3	0.318378	0.3116	0.3227	0.8003	0.9776	0.8199	1.166
1	rs12094228	NCF2	181814495	2	1	0.05743	0.05607	0.0583	0.979	0.9952	0.6952	1.425
1	rs3845466	NCF2	181818383	1	3	0.0800431	0.08364	0.07774	0.5311	1.102	0.8128	1.495
1	rs10911364	NCF2	181819149	3	1	0.0491744	0.04504	0.05183	0.4769	0.8662	0.5831	1.287
1	rs789185	NCF2	181820904	3	1	0.379038	0.3879	0.3734	0.3703	1.079	0.9132	1.276
1	rs11578964	NCF2	181822402	1	3	0.410983	0.4136	0.4093	0.601	1.046	0.8848	1.236
1	rs789191	NCF2	181822502	1	3	0.0258435	0.02849	0.02415	0.4865	1.196	0.7223	1.981
1	rs11588654	NCF2	181823018	3	1	0.334171	0.3318	0.3357	0.895	0.9883	0.8296	1.177
1	rs10911365	NCF2	181825728	1	3	0.0789663	0.0864	0.0742	0.0904	1.291	0.9606	1.736
16	rs3180279	CYBA	87238334	3	2	0.497128	0.4853	0.5047	0.2459	0.907	0.7691	1.07
16	rs9940427	CYBA	87239320	1	2	0.0333812	0.03125	0.03475	0.4866	0.8506	0.5393	1.342
16	rs4673	CYBA	87240737	1	3	0.361091	0.3906	0.3422	0.006009	1.265	1.07	1.497
16	rs3794624	CYBA	87244575	1	3	0.323635	0.3002	0.3386	0.06181	0.8484	0.7139	1.008
22	rs6000448	NCF4	35580717	1	3	0.10122	0.102	0.1007	0.6934	1.056	0.8043	1.387
22	rs10854693	NCF4	35581956	3	1	0.417085	0.4017	0.427	0.1105	0.8723	0.7376	1.032
22	rs1883113	NCF4	35582630	3	2	0.0681981	0.07812	0.06184	0.03028	1.429	1.035	1.973
22	rs4821542	NCF4	35582864	1	3	0.1486	0.1369	0.1561	0.3859	0.904	0.7197	1.136
22	rs7287350	NCF4	35583185	1	3	0.0567121	0.06618	0.05065	0.01461	1.552	1.091	2.208
22	rs1883112	NCF4	35586792	1	3	0.439568	0.4309	0.4451	0.2815	0.9129	0.7735	1.078
22	rs4821544	NCF4	35588449	3	1	0.318966	0.314	0.3221	0.816	0.9793	0.8213	1.168
22	rs741997	NCF4	35588759	3	1	0.0753769	0.07169	0.07774	0.8289	1.034	0.7623	1.403
22	rs746713	NCF4	35589305	3	1	0.286844	0.2482	0.3115	0.001418	0.7404	0.6156	0.8905
22	rs909484	NCF4	35590547	1	3	0.181263	0.1939	0.1731	0.1684	1.16	0.939	1.434
22	rs729749	NCF4	35593792	1	3	0.22852	0.2477	0.2162	0.04018	1.225	1.009	1.487
22	rs2075938	NCF4	35596268	1	3	0.29397	0.3033	0.288	0.2362	1.115	0.9315	1.334
22	rs8137456	NCF4	35604389	2	3	0.0660445	0.07261	0.06184	0.3465	1.165	0.8473	1.603
22	rs8137602	NCF4	35604595	3	2	0.0933238	0.08364	0.09953	0.4091	0.8884	0.6708	1.177
22	rs3959633	NCF4	35605501	1	3	0.0201005	0.02022	0.02002	0.7815	0.9196	0.5084	1.663
22	rs5756379	NCF4	35605850	3	1	0.350216	0.3674	0.3392	0.05921	1.184	0.9935	1.411
22	rs4821554	NCF4	35606030	1	3	0.273869	0.2812	0.2691	0.2966	1.104	0.9169	1.329
22	rs5750326	NCF4	35606231	1	2	0.490309	0.4871	0.4923	0.4902	0.9434	0.7997	1.113
22	rs5756564	RAC2	35942049	3	1	0.402728	0.3649	0.427	0.004023	0.7811	0.6601	0.9243
22	rs4820272	RAC2	35945560	1	3	0.149318	0.1599	0.1425	0.4011	1.106	0.8745	1.398
22	rs933222	RAC2	35948583	1	3	0.284637	0.3061	0.2709	0.02311	1.236	1.029	1.483
22	rs12166968	RAC2	35951304	2	3	0.0635319	0.05974	0.06596	0.5271	0.8973	0.6413	1.255
22	rs6572	RAC2	35951391	3	2	0.443647	0.4779	0.4217	0.004258	1.281	1.081	1.517
22	rs739041	RAC2	35954945	3	1	0.430216	0.4219	0.4356	0.4664	0.9406	0.7977	1.109
22	rs9607431	RAC2	35959884	2	1	0.124192	0.1452	0.1107	0.03367	1.307	1.021	1.673
22	rs1476002	RAC2	35960734	1	3	0.135678	0.1535	0.1243	0.07964	1.242	0.9747	1.584
22	rs13058338	RAC2	35962716	4	1	0.265445	0.2454	0.2783	0.1251	0.8641	0.7169	1.041
22	rs5756573	RAC2	35963721	1	3	0.151831	0.1599	0.1466	0.2873	1.131	0.9013	1.42
22	rs2284038	RAC2	35965001	3	1	0.393037	0.3925	0.3934	0.9423	0.9938	0.8401	1.176

22	rs2239774	RAC2	35967599	2	3	0.1486	0.1636	0.139	0.2295	1.149	0.9163	1.44
22	rs2239775	RAC2	35967787	1	2	0.141062	0.1471	0.1372	0.7057	1.048	0.8231	1.333
22	rs2239773	RAC2	35968235	1	3	0.282952	0.2566	0.2996	0.02475	0.8069	0.6691	0.9731
22	rs2213430	RAC2	35968906	1	3	0.434673	0.4375	0.4329	0.5006	1.06	0.8954	1.254
22	rs6000632	RAC2	35974061	1	3	0.246231	0.2381	0.2515	0.6145	0.9515	0.784	1.155
22	rs4821615	RAC2	35976077	2	3	0.444365	0.443	0.4452	0.6948	0.9676	0.8206	1.141
22	rs12484031	RAC2	35979216	3	1	0.105963	0.105	0.1066	0.9289	0.9882	0.7623	1.281
22	rs7288979	RAC2	35979440	3	1	0.177315	0.1599	0.1885	0.06777	0.8183	0.6598	1.015
X	rs7059081	CYBB	37526428	2	3	0.08004	0.08456	0.07715	0.5594	1.071	0.8503	1.349
X	rs6651773	CYBB	37528612	3	1	0.1687	0.1737	0.1655	0.6291	1.044	0.8779	1.24
X	rs17312212	CYBB	37529550	3	1	0.0542	0.04688	0.05889	0.2549	0.8446	0.6315	1.13
X	rs4827298	CYBB	37536603	3	1	0.2552	0.2656	0.2485	0.3988	1.065	0.9206	1.231
X	rs5917471	CYBB	37537458	1	3	0.4207	0.4375	0.4099	0.2275	1.083	0.9516	1.232
X	rs12848910	CYBB	37551189	3	1	0.1131	0.1066	0.1172	0.4655	0.926	0.7533	1.138

Table 3C. Association of NADPH oxidase gene SNPs with UC in Discovery Cohort: MAF: minor allelic frequency. F_A: frequency in affected patients. F_U: frequency in unaffected controls. L95 and U95: lower and upper confidence interval.

Supplemental Table 4 Association of NADPH oxidase gene SNPs and IBD, CD, and UC in North American and Scottish cohorts and combined analysis

4A. North America

IBD								ADD MODEL			
CHR	SNP	position	A1	A2	MAF	F_A	F_U	P	OR (ci=0.95)	L95	U95
22	rs5756564	35,942,049	C	T	0.3881	0.3755	0.4007	0.1121	0.8961	0.7826	1.026
22	rs6572	35,951,391	C	G	0.4581	0.4549	0.4612	0.701	0.975	0.857	1.109
22	rs9607431	35,959,884	C	A	0.1255	0.1163	0.1348	0.09077	0.8445	0.6943	1.027
22	rs2239774	35,967,599	C	G	0.1503	0.1462	0.1545	0.4832	0.9373	0.7822	1.123

CD								ADD MODEL			
CHR	SNP	position	A1	A2	MAF	F_A	F_U	P	OR (ci=0.95)	L95	U95
22	rs5756564	35,942,049	C	T	0.3882	0.3623	0.4007	0.0522	0.8468	0.7159	1.002
22	rs6572	35,951,391	C	G	0.465	0.4729	0.4612	0.5694	1.048	0.8926	1.229
22	rs9607431	35,959,884	C	A	0.1251	0.105	0.1348	0.02724	0.7504	0.5816	0.968
22	rs2239774	35,967,599	C	G	0.145	0.1253	0.1545	0.04474	0.7863	0.6218	0.994

UC								ADD MODEL			
CHR	SNP	position	A1	A2	MAF	F_A	F_U	P	OR (ci=0.95)	L95	U95
22	rs5756564	35,942,049	C	T	0.3963	0.3878	0.4007	0.5055	0.9462	0.8041	1.113
22	rs6572	35,951,391	C	G	0.4533	0.4382	0.4612	0.2514	0.913	0.7814	1.067
22	rs9607431	35,959,884	C	A	0.1321	0.1268	0.1348	0.5554	0.9324	0.7389	1.177
22	rs2239774	35,967,599	C	G	0.1583	0.1656	0.1545	0.4411	1.088	0.878	1.348

MAF: minor allelic frequency. F_A: frequency in affected patients. F_U: frequency in unaffected controls. L95 and U95: lower and upper confidence interval.

4B. Scotland

IBD

								ADD MODEL			
CHR	SNP	position	A1	A2	MAF	F_A	F_U	P	OR (ci=0.95)	L95	U95
22	rs5756564	35,942,049	2	4	0.4251	0.4248	0.4255	0.9596	0.9969	0.8858	1.122
22	rs6572	35,951,391	2	3	0.4292	0.4412	0.4158	0.07943	1.11	0.9878	1.248
22	rs9607431	35,959,884	2	1	0.1246	0.1251	0.1241	0.9183	1.009	0.8501	1.198
22	rs2239774	35,967,599	2	3	0.1459	0.1468	0.145	0.8615	1.014	0.8652	1.189

CD

								ADD MODEL			
CHR	SNP	position	A1	A2	MAF	F_A	F_U	P	OR (ci=0.95)	L95	U95
22	rs5756564	35,942,049	2	4	0.4217	0.4151	0.4255	0.5428	0.9572	0.8315	1.102
22	rs6572	35,951,391	2	3	0.4312	0.4577	0.4158	0.01663	1.184	1.031	1.36
22	rs9607431	35,959,884	2	1	0.1268	0.1313	0.1241	0.5404	1.065	0.8709	1.302
22	rs2239774	35,967,599	2	3	0.1468	0.1499	0.145	0.6052	1.038	0.861	1.252

UC

								ADD MODEL			
CHR	SNP	position	A1	A2	MAF	F_A	F_U	P	OR (ci=0.95)	L95	U95
22	rs5756564	35,942,049	2	4	0.429	0.4359	0.4255	0.5638	1.044	0.9023	1.208
22	rs6572	35,951,391	2	3	0.4184	0.4232	0.4158	0.6789	1.031	0.8929	1.19
22	rs9607431	35,959,884	2	1	0.1221	0.1182	0.1241	0.6224	0.9476	0.765	1.174
22	rs2239774	35,967,599	2	3	0.1444	0.1433	0.145	0.8998	0.9875	0.8121	1.201

MAF: minor allelic frequency. F_A: frequency in affected patients. F_U: frequency in unaffected controls. L95 and U95: lower and upper confidence interval.

Supplemental Table 5. Association of NADPH oxidase gene SNPs and IBD, CD, and UC in North American and Scottish cohorts and combined subphenotype analysis

Colon Any L2/L3/CD						ADD MODEL			
CHR	SNP	position	MAF	F_A	F_U	P	OR (ci=0.95)	L95	U95
22	rs5756564	35,942,049	0.4029	0.3873	0.418	0.0008126	0.8783	0.8141	0.9476
22	rs6572	35,951,391	0.4473	0.4629	0.4321	0.0009037	1.135	1.053	1.222
22	rs9607431	35,959,884	0.1292	0.1349	0.1236	0.07272	1.105	0.9908	1.233
22	rs2239774	35,967,599	0.1536	0.1612	0.1462	0.02892	1.119	1.012	1.237

Colon Only CD/L2						ADD MODEL			
CHR	SNP	position	MAF	F_A	F_U	P	OR (ci=0.95)	L95	U95
22	rs5756564	35,942,049	0.409	0.3966	0.418	0.03143	0.9137	0.8415	0.992
22	rs6572	35,951,391	0.4401	0.4512	0.4321	0.05689	1.082	0.9977	1.173
22	rs9607431	35,959,884	0.1263	0.13	0.1236	0.3399	1.06	0.9406	1.194
22	rs2239774	35,967,599	0.1497	0.1546	0.1462	0.2546	1.066	0.955	1.19

Ileal Only (L1)						ADD MODEL			
CHR	SNP	position	MAF	F_A	F_U	P	OR (ci=0.95)	L95	U95
22	rs5756564	35,942,049	0.4132	0.3814	0.418	0.04175	0.8577	0.7399	0.9943
22	rs6572	35,951,391	0.4363	0.4647	0.4321	0.07483	1.14	0.987	1.316
22	rs9607431	35,959,884	0.1239	0.1259	0.1236	0.8495	1.021	0.8238	1.265
22	rs2239774	35,967,599	0.1476	0.1569	0.1462	0.4206	1.083	0.8917	1.316

Perianal (p)						ADD MODEL			
CHR	SNP	position	MAF	F_A	F_U	P	OR (ci=0.95)	L95	U95
22	rs5756564	35,942,049	0.4134	0.4044	0.418	0.2207	0.9445	0.8621	1.035
22	rs6572	35,951,391	0.4358	0.4431	0.4321	0.3243	1.046	0.9566	1.144
22	rs9607431	35,959,884	0.1265	0.1321	0.1236	0.263	1.077	0.9455	1.228
22	rs2239774	35,967,599	0.1518	0.1625	0.1462	0.04741	1.129	1.001	1.273

Early Onset < 16 years of age						ADD MODEL			
CHR	SNP	position	MAF	F_A	F_U	P	OR (ci=0.95)	L95	U95
22	rs5756564	35,942,049	0.4069	0.3794	0.418	0.001456	0.8512	0.7709	0.94
22	rs6572	35,951,391	0.4419	0.4664	0.4321	0.004964	1.149	1.043	1.266
22	rs9607431	35,959,884	0.1301	0.1464	0.1236	0.006036	1.215	1.057	1.396
22	rs2239774	35,967,599	0.1511	0.163	0.1462	0.05994	1.133	0.9948	1.29

Ileal Any (L1/L3)						ADD MODEL			
CHR	SNP	position	MAF	F_A	F_U	P	OR (ci=0.95)	L95	U95
22	rs5756564	35,942,049	0.4053	0.374	0.418	0.000247	0.8301	0.7515	0.917
22	rs6572	35,951,391	0.4459	0.4798	0.4321	0.00009542	1.212	1.101	1.335
22	rs9607431	35,959,884	0.1289	0.1421	0.1236	0.02565	1.171	1.019	1.346
22	rs2239774	35,967,599	0.1524	0.1677	0.1462	0.01694	1.169	1.028	1.33

Ileo-colonic (L3)						ADD MODEL			
CHR	SNP	position	MAF	F_A	F_U	P	OR (ci=0.95)	L95	U95
22	rs5756564	35,942,049	0.3889	0.3366	0.4007	0.00003088	0.7729	0.6847	0.8724
22	rs6572	35,951,391	0.4692	0.5049	0.4612	6.768E-06	1.307	1.163	1.469
22	rs9607431	35,959,884	0.1347	0.1341	0.1348	0.009905	1.24	1.053	1.46
22	rs2239774	35967599	0.1548	0.1561	0.1545	0.001021	1.286	1.107	1.494

Combined analysis of the discovery, replication and validation cohorts (1790 CD and 1636 UC patients, and 2908 controls). MAF: minor allelic frequency. F_A: frequency in affected patients. F_U: frequency in unaffected controls. L95 and U95: lower and upper confidence interval.

Supplemental Table 6. Sub-phenotype analysis *RAC2* SNPs in combined analysis

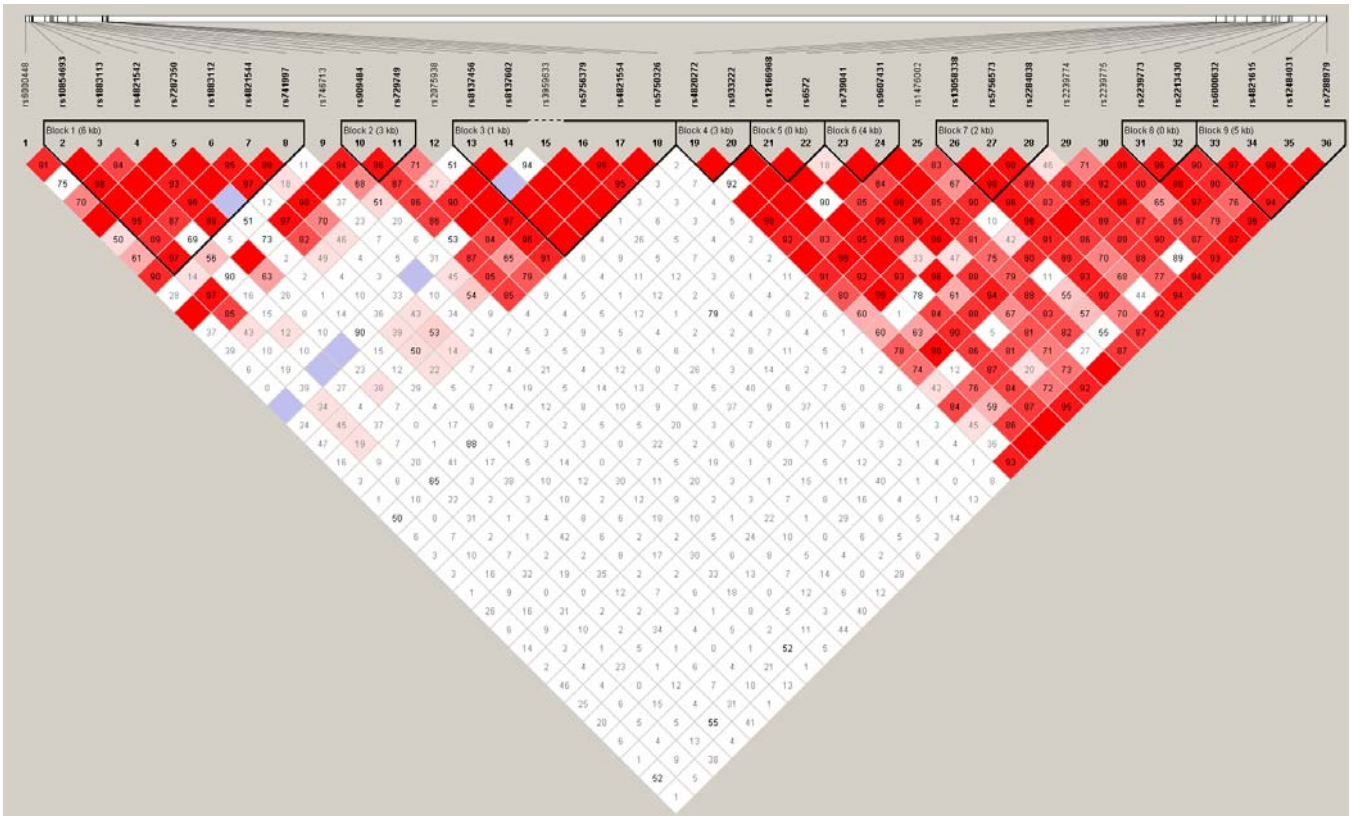
	Ileal "only" CD (L1)	Ileal "any" (L1/L3)	Colonic "only" CD (L2)	Ileo-colonic CD (L3)	Colon "Only" IBD (UC/L2 CD)	Colon "Any" IBD (UC/L2/L3 CD)	EO- IBD	CD with perianal disease (p)
rs5756564	P = 4.0 x 10-2, OR = 0.85 (0.73-0.99)	P = 2.5 x 10-4, OR = 0.83 (0.70-0.94)	ns	P = 3.0 x 10-5, OR = 0.77 (0.68-0.87)	P = 3.1 x 10-2, OR = 0.91 (0.84-0.99)	P = 8.1 x 10-4, OR = 0.87 (0.81-0.94)	P = 2.9 x 10-3, OR = 0.85 (0.77-0.94)	ns
rs6572	ns	P = 4.9 x 10-5, OR = 1.14 (1.04-1.26)	P = 3.0 x 10-2, OR = 1.15 (1.00-1.32)	P = 6.0 x 10-6, OR = 1.30 (1.16-1.46)	ns	P = 9.0 x 10-4, OR = 1.13 (1.05-1.22)	P = 4.9 x 10-3, OR = 1.14 (1.04-1.26)	ns
rs9607431	ns	P = 2.5 x 10-2, OR = 1.17 (1.01-1.35)	ns	P = 9.9 x 10-3, OR = 1.24 (1.05-1.46)	ns	ns	P = 6.0 x 10-3, OR = 1.21 (1.05-1.39)	ns
rs2239774	ns	P = 1.6 x 10-2, OR = 1.16 (1.02-1.33)	ns	P = 1.0 x 10-3, OR = 1.28 (1.10-1.49)	ns	P = 2.8 x 10-2, OR = 1.19 (1.02-1.23)	P = 3.4 x 10-2, OR = 1.17 (1.01-1.35)	P = 4.0 x 10-2, OR = 1.12 (1.02-1.27)

Combined analysis of the discovery, replication and validation cohorts (1790 CD and 1636 UC patients, and 2908 controls).. ns – not significant – $p > 0.05$. EO-IBD – early onset IBD (diagnosis prior to age 16). Bonferroni correction threshold: for 62 SNP and 8 IBD sub-phenotypes examined; $a = 1.0 \times 10^{-4}$

Supplemental Table 7. Sub-phenotype analysis *NCF4* SNPs in discovery analysis

	IBD	CD	Colonic "only" CD (L2)	Colon "any" CD (L2/L3 CD)	IBD Dx < 19 years of age	CD with perianal disease (p)
rs10854693	$P = 6.4 \times 10^{-3}$, OR 0.82 (0.72-0.94)	$P = 3.8 \times 10^{-3}$, OR 0.79 (0.67-0.92)	ns	$P = 2.9 \times 10^{-2}$, OR 0.86 (0.75-0.98)	$P = 1.0 \times 10^{-2}$, OR 0.82 (0.71-0.96)	$P = 1.5 \times 10^{-3}$, OR 0.69 (0.55-0.86)
rs1883113	$P = 6.4 \times 10^{-3}$, OR 1.45 (1.10-1.90)	$P = 1.3 \times 10^{-2}$, OR 1.47 (1.08-2.00)	ns	ns	$P = 4.8 \times 10^{-3}$, OR 1.49 (1.13-1.97)	$P = 2.0 \times 10^{-2}$, OR 1.57 (1.07-2.32)
rs7287350	$P = 7.1 \times 10^{-3}$, OR 1.50 (1.11-2.03)	ns	ns	ns	$P = 2.7 \times 10^{-2}$, OR 1.42 (1.04-1.93)	ns
rs1883112	$P = 8.8 \times 10^{-3}$, OR 0.83 (0.73-0.95)	$P = 1.3 \times 10^{-2}$, OR 0.77 (0.66-0.90)	ns	$P = 3.9 \times 10^{-2}$, OR 0.87 (0.76-0.99)	$P = 4.8 \times 10^{-3}$, OR 0.85 (0.73-0.95	$P = 1.1 \times 10^{-3}$, OR 0.68 (0.54-0.86))
rs746713	$P = 4.4 \times 10^{-2}$, OR 0.86 (0.74-0.99)	ns	$P = 1.6 \times 10^{-3}$, OR 0.77 (0.66-0.90)	$P = 6.1 \times 10^{-3}$, OR 0.81 (0.70-0.94)	ns	ns
rs4821544	ns	$P = 5.1 \times 10^{-3}$, OR 1.25 (1.07-1.47)	ns	ns	ns	ns

ns – not significant – $p > 0.05$. EO-IBD – early onset IBD (diagnosis prior to age 16). Bonferroni correction threshold: for 62 SNP and 8 IBD sub-phenotypes examined; $a = 1.0 \times 10^{-4}$



Supplemental Figure 1: LD plot (LD statistic in D') of *NCF4* (35,586,991 to 35,604,004) and *RAC2* (35,951,258 to 35,970,251)