Published in final edited form as:
Dis Colon Rectum. 2011 August ; 54(8): 1020–1025. doi:10.1007/DCR.0b013e31821b94b3.

Influence of Crohn's disease risk alleles and smoking on disease location

Hongyan Chen, M.S.[‡], **Alexander Lee, M.D.**[*], **Anne Bowcock, Ph.D.**[#], **Wei Zhu, Ph.D.**[‡], **Ellen Li, M.D.-Ph.D.**[†,*], **Matthew Ciorba, M.D.**[*], and **Steven Hunt, M.D.**[o]

[*]Department of Medicine, Washington University-St. Louis School of Medicine, Saint Louis, Missouri

[#]Department of Genetics, Washington University-St. Louis School of Medicine, Saint Louis, Missouri

[o]Department of Surgery, Washington University-St. Louis School of Medicine, Saint Louis, Missouri

[†]Department of Medicine, Stony Brook University, Stony Brook, New York

[‡]Department of Applied Mathematics and Statistics, Stony Brook University, Stony Brook, New York

Abstract

Objective—Our objective is to assess the effect of genetic and environmental factors on Crohn's disease location.

Design—We identified 628 Crohn's disease patients within the Washington University database (April 2005-February 2010) that had complete information on 31 Crohn's disease associated genotypes and clinical information on disease location (L1-L4), smoking, gender, race and age at diagnosis. For statistical reasons, the three major NOD2 alleles (rs2066844, rs2066845, rs2066847) were grouped together. Logistic regression incorporating all of the genotypes and clinical covariates, including smoking, was carried out with stepwise variable selection and by best subset selection.

Results—Stepwise variable selection selected three major covariates, composite NOD2 genotype, smoking, and TNFSF15 genotype, which are also the three covariates selected by the best subset method. While NOD2 genotype and smoking are positively associated with ileal (L1 + L3) disease, TNFSF15 genotype is positively associated with isolated colonic (L2) disease.

Limitations—The ability to detect disease site associations in this single center study may be limited by the population size, low allelic frequency and/or low odds ratio of certain CD risk alleles.

Address correspondence to: Steven R. Hunt, Department of Surgery, Campus Box 8109, 660 S. Euclid Avenue, TEL: 1-314-454-8593, FAX: 1-314-747-2977, HUNTS@wudosis.wustl.edu.

Hongyan Chen: Statistical analysis of data, drafting the article, final approval of version published.
Alexander Lee: Acquisition of data, revising the article, final approval of version published.
Anne Bowcock: Contribution to conception and design, acquisition of data, revising the article, final approval of version published.
Wei Zhu: Statistical analysis, revising the article, final approval of version published.
Ellen Li: Contribution to conception and design, revising the article, final approval of version published.
Matthew Ciorba: Contribution to conception and design, revising the article, final approval of version published.
Steven Hunt: Contribution to conception and design, revising the article, final approval of version published.

This work was presented in part at the American Gastroenterology Association meeting held in New Orleans, May 1-5, 2010.

Conclusion—These results indicate that NOD2 genotype, smoking status and TNFSF15 genotype should be included as co-variates in assessing the effect of genetic and environmental factors on Crohn's disease site location.

Classifications

Crohn's disease; genetics; smoking; colon

INTRODUCTION

Crohn's disease (CD) is a chronic relapsing inflammatory intestinal disorder that can affect any segment of the intestine often in a discontinuous manner.[1,2] Because CD is phenotypically heterogeneous, CD patients are further stratified on the basis of three parameters: 1.) age of diagnosis, 2.) disease location and 3.) disease behavior based on the Vienna and Montreal classification systems.[3,4] While the first two parameters are stable over time, disease behavior is dynamic and therefore difficult to correlate with genotype data.[5] Four major patterns of disease location have been defined: L1, ileal disease with or without cecal disease ("pure" ileal CD); L2, colonic disease only (Crohn's colitis); L3, ileal disease with colonic disease beyond the cecum (ileocolonic disease); L4, proximal intestinal disease. Approximately 70-85 % of CD patients have ileal disease involvement (L1 + L3) and 15-30% has colonic disease (L2) without ileal involvement.[6-8] Only a very small number of patients have disease restricted to the proximal gut (L4).

Both genetic factors and environmental factors have been associated with ileal disease location (L1 + L3). Smoking has been linked to small bowel as opposed to colonic disease.[9-12] NOD2 and ATG16L1 risk alleles have been associated with ileal disease location based on comparisons of the ileal CD subset with control patients considering only a limited number of genetic loci.[13-19] Since then, more than 30 CD susceptibility loci have been identified through genome-wide association studies.[20] Taking these susceptibility loci and smoking into consideration, we have carried out logistic regression analysis on a single CD database in order to assess genetic and environmental factors that are differentially associated with the ileal CD subset L1 + L3) compared to the colonic CD subset without ileal disease (L2).

MATERIALS AND METHODS

Subjects

This study was approved by the Washington University-St. Louis and Stony Brook University Institutional Review Boards. We identified 628 CD subjects from the WU DDRCC TPF database (April 2005-February 2010) that had complete clinical information on Montreal disease location, smoking status, gender, race and age of diagnosis and complete genotype information on 31 established CD genetic susceptibility loci. Subjects were recruited to the Washington University Digestive Diseases Research Core Center Tissue Procurement Facility (WU DDRCC TPF) in a consecutive fashion during inpatient and outpatient visits to physicians within the Section of Colon Rectal Surgery and the Division of Gastroenterology at the Washington University Medical Center (WUMC) to donate de-identified clinical information and blood, saliva and/or tissues for genotyping. Patients who were unwilling or unable to give informed consent were excluded from the database.

Clinical variables

Gender and race were recorded as reported by the patients (see Table 1). Based on a review of the subjects' medical records they were classified into one of the following Montreal

disease location categories : L1, ileal disease with or without cecal disease (ileal CD); L2, colonic disease only (Crohn's colitis); L3, ileal disease with colonic disease beyond the cecum; L4, proximal intestinal disease.[3] The first three categories (L1, L2 and L3), include patients who also had proximal intestinal disease location as a modifier (L4+). In patients that had undergone subtotal colectomy, L2 CD was distinguished from ulcerative colitis (UC and indeterminate colitis (IC) based on the pathology of the resected specimen. The database includes patients initially diagnosed with UC but subsequently reclassified as CD after developing small bowel or perianal disease. It is therefore possible that some of the UC patients in our database who were excluded from the present study may be reclassified as CD in the future. Subjects were also classified into one of the Montreal disease location categories: A1, age < 17 years, A2, 17-40 years, A3, >40 years.[3] Because disease behavior is dynamic over time, we did not include this variable for this analysis. The patient's smoking history was determined at the time of initial surgery or most recent follow up prior to surgery by reviewing the medical records and interviewing the patient. A smoker was defined as smoking ≥ 7 cigarettes a week for at least a year as described previously.[5,9,11] An ex-smoker was defined as subjects who had stopped smoking for at least a year. A non-smoker was defined as subjects who had never smoked any cigarettes.

Genotyping

DNA was prepared from peripheral blood mononuclear cells, saliva and/or GI tissues using Qiagen Gentra Puregene (Qiagen Inc., Valencia, CA) reagents according to the manufacturer's recommendations. Genotyping of the NOD2 single nucleotide polymorphism (SNP), Leu1007fsInsC (rs2066847, SNP13) was conducted by Taqman® allelic discrimination as previously described.[21] Applied Biosystems Taqman® SNP Genotyping assays (catalogue nos. C_11717468 and C_11717466, Life Technologies Co., Carlsbad, CA) were used to genotype NOD2 R702W (rs2066844, SNP8 and G908R (rs2066845, SNP12) alleles respectively, according to the manufacturer's recommendations. The remaining SNPs (see Table 2) were genotyped using the Sequenom MassArray System (Sequenom Inc., San Diego, CA) in the Washington University Sequenom Technology Core.

Statistical analysis

The three major NOD2 (rs2066847, rs2066844, rs2066845) risk alleles were combined into a single "super" allele. Subjects with two risk alleles (either homozygote or compound heterozygote) or with a single risk allele were designated R (R = risk). Those with none of the three risk alleles were designated as NR (NR = nonrisk). In this analysis we present age of diagnosis as a categorical variable, based on the Montreal classification. Logistic regression was applied in order to model the effect on disease locations (L1+L3 vs. L2) of all the CD genotypes, gender, race, smoking and age of diagnosis. We did stepwise variable selection in SAS®9.0 to identify significantly effective single nucleotide polymorphisms (SNPs) or other covariates (α=0.05). All predictor variables were treated as categorical ones with three possible types. Likelihood-ratio test (approximately chi-squared test) was conducted during every round of variable selection. Afterwards, we also performed the best subset selection in R (see supplementary text for code) to confirm our findings. Akaike information criterion was set as the goodness of fit to assess subsets. We also conducted the analysis with only the SNPs.

RESULTS

The clinical characteristics of the CD patients with respect to disease location are summarized in Table 1. The relative distribution of CD patients with ileal disease (L1 + L3) and with only colonic disease (L2) in this database was approximately 80:20. Only 6 (~1%)

of the CD patients had isolated proximal bowel disease (L4) and these patients were not included in the subsequent analysis. The CD patients were predominantly Caucasian. The predominant categorical age of diagnosis based on Montreal classification was A2 (17-40 years). A lower proportion of L2 patients had undergone surgery than L1 + L3 patients as of the cutoff date (February 2010).

The distribution of CD genotypes for 31 established Crohn's disease genetic loci with respect to disease location are summarized in Table 2. We performed logistic regression analysis by stepwise variable selection to evaluate all 31 established Crohn's disease genetic loci along with the clinical co-variates, including smoking, on disease site. CD patients with two NOD2 risk alleles (R/R, O.R. 6.102, 95% C.I.1.855-20.074) and with one NOD2 risk allele (R/NR, O.R. 2.552, 95% C.I.1.479-4.402) have a 2 fold risk (O.R. 2.552, 95% C.I. 1.479-4.402) of developing ileal (L1 + L3) disease compared to pure colonic (L2) disease. Current smokers have an increased risk of developing ileal compared to colonic disease than both control (O.R. 1.689, 95% C.I. 1.071- 2.664), and ex-smokers (O.R. 2.611, 95% C.I. 1.279-5.319) of developing ileal vs. isolated colonic disease (see Table 3). In contrast, CD patients with one or two TNFSF15 risk alleles have ~4 fold risk of developing isolated colonic compared to ileal disease. When the logistic regression was carried out with only genotype data, NOD2 and TNFSF15 genotype were selected with similar O.R. and 95% C.I by both stepwise variable selection and best subset selection.

DISCUSSION

This is one of the first studies to evaluate the effect of 31 established CD genetic loci with clinical co-variates, including smoking, on CD site (L1 + L3 vs. L2) using logistic regression.[20] This study demonstrates that both NOD2 risk alleles and smoking were associated with an increased risk of developing ileal disease vs. isolated colonic disease. It thus confirms the results of a previous logistic regression study that examined NOD2 risk alleles alone with other clinical co-variates including smoking, and other studies that have examined these two variables separately.[9-15,22,23] Waterman et al. recently reported the allelic frequencies of 31 established genetic loci and 6 additional UC loci in 366 L1 compared to 228 L2 CD patients.[24] Two NOD2 loci (rs2066844, 0.1014 vs. 0.05, P = 0.002; rs2066847, 0.1113 vs. 0.0376, P = 0.0001), the ATG16L1 locus (rs2241880, 0.381 vs.0.455, P = 0.024) and the ICOSL1 locus (rs762421, 0.4345 vs. 0.3484, P = 0.0131) demonstrated significantly different allelic frequencies in univariate analyses, and did not include smoking as a covariate in the analysis. Our study demonstrates that current smoking has a tremendous impact on disease location in CD, and suggests that analysis of additional genetic and environmental factors should include smoking as a covariate. In this study smoking history was elicited by patient report alone. Measurement of serum cotinine levels may provide a more objective indicator of active smoking in future studies.[25]

In addition, logistic regression analysis identified the TNSF15 (rs4263839) genotype as significantly associated with a reduced risk of developing ileal disease (L1 + L3) vs. isolated colonic disease (L2). This particular SNP was chosen because it was the same SNP assayed previously by Barrett et al.[20] TNFSF15 SNPs have been previously associated with CD and in certain populations with UC.[26-30] The causative TNFSF15 SNP has not been identified thus far despite extensive sequencing of the gene. It is possible that this SNP is in linkage equilibrium, or nonrandomly associated, with the true disease allele, which may or may not affect TNFSF15 expression.

The ATG16L1T300A GG genotype has been associated with ileal CD compared to control patients.[16-19] We were unable to detect an effect of ATG16L1 genotype when comparing ileal CD with isolated colonic disease. Of note, a recent meta-analysis found excessive odds

ratio heterogeneity among previous studies comparing the distribution of ATG16l1 risk allele frequency in ileal compared with colonic disease.[18] It is interesting to note that the allelic frequency of the ATG16L1 risk allele was actually higher in the L2 CD patients than the L1 CD patients in the study reported by Waterman et al.[24]

Since smoking is self-reported, the classification of smoking status is not as robust as the genotyping data. The ability to detect disease site associations in this single center study may be limited by the population size, low allelic frequency and/or low O.R. of certain CD risk alleles. We anticipate that more SNPs will emerge in meta-analyses combining datasets from multiple centers. Nonetheless this study reinforces the concept that isolated Crohn's colitis exhibits molecular characteristics that are distinct from that of the two other major Crohn's disease subtypes, ileal (L1) and ileocolonic (L3) Crohn's disease.[22]

Acknowledgments

The authors thank the patients who have contributed their medical information to the Digestive Diseases Research Core Center (DDRCC) Clinical Database, Candace Miller and Drs. Elisa Birnbaum, James Fleshman, Ira Kodner, Anne Lin, Matthew Mutch and Bashar Safar in the Section of Colon and Rectal Surgery, Drs. Themistocles Dassopoulos, Nicholas Davidson, Dayna Early, Chandra Prakash, Deborah Rubin, Gregory Sayuk and William Stenson in the Division of Gastroenterology at Washington University, and the community gastroenterologists including Drs. Barry Abramson, Marc Bernstein, Paul Buse, David Cort, Steve Fern, John J. Kelly, Carl Lyss, Robert Shuman, Eric Thyssen, Janet Todorczuk, David Walden, Michele Woodley, Leonard Weinstock, Michael Zerega for their help and support in recruiting patients. We thank Elizabeth Cole, Melissa Reimers, Zegbeh Kpadeh, Katherine Stanley, Terry Li for assistance in genotyping the patients.

This work was partially supported by NIH grants UH2 DK grant (Li) and P30 DK 52574, a Senior Research Award from the Crohn's and Colitis Foundation of America (Li), a grant from the Simons Foundation (Li). We acknowledge use of the Washington University Digestive Diseases Research Core Center Tissue Procurement Facility (P30 DK52574) and the Human Genetics Genotyping Core. M.C. is a recipient of a Crohn's and Colitis Foundation Career Development Award and DK089016-01. A.L. participated in the Mentors in Medicine Program and the Clinical Science Training and Research Pathway in the Department of Medicine at Washington University.

References

1. Goyette P, Labbé C, Trinh TT, et al. Molecular pathogenesis of inflammatory bowel disease: genotypes, phenotypes and personalized medicine. Ann Med. 2007; 39:177–99. [PubMed: 17457716]
2. Abraham C, Cho J. Mechanisms of disease: inflammatory bowel disease. New Engl J Medicine. 2009; 361:2066–2078.
3. Satsangi J, Silverberg MS, Vermeire S, et al. The Montreal classification of inflammatory bowel disease: controversies, consensus, and implications. Gut. 2006; 55:749–753. [PubMed: 16698746]
4. Louis E, Collard A, Oger AF, et al. Behavior of Crohn's disease according to the Vienna classification: changing pattern over the course of the disease. Gut. 2001; 49:777–782. [PubMed: 11709511]
5. Unkart JT, Anderson L, Li E, et al. Risk factors for surgical recurrence after ileocolic resection of Crohn's disease. Dis Colon Rectum. 2008; 51:1211–1216. [PubMed: 18536967]
6. Farmer RG, Whelan G, Fazio VW. Long-term follow-up of patients with Crohn's disease. Relationship between the clinical pattern and prognosis. Gastroenterology. 1985; 88:1818–1825. [PubMed: 3922845]
7. Magro F, Portela F, Lago P, et al. Crohn's disease in a southern European country: Montreal classification and clinical activity. Inflamm Bowel Dis. 2009; 15:1343–1350. [PubMed: 19235885]
8. Freeman HJ. Application of the Montreal classification for Crohn's disease to a single clinician database of 1015 patients. Can J Gastroenterol. 2007; 21:363–366. [PubMed: 17571169]
9. Cosnes J, Carbonnel F, Beaugerie L, et al. Effects of cigarette smoking on the long-term course of Crohn's disease. Gastroenterology. 1996; 110:424–431. [PubMed: 8566589]

10. Russel MG, Volovics A, Schoon EJ, et al. Inflammatory bowel disease: is there any relation between smoking status and disease presentation? European Collaborative IBD Study Group. Inflamm Bowel Dis. 1998; 4:182–186. [PubMed: 9741019]
11. Brant SR, Picco MF, Achkar JP, et al. Defining complex contributions of NOD2/CARD15 gene mutations, age of onset, and tobacco use on Crohn's disease phenotypes. Inflamm Bowel Dis. 2003; 9:281–289. [PubMed: 14555911]
12. Aldhous MC, Drummond HE, Anderson N. Does cigarette smoking influence the phenotype of Crohn's disease? Analysis using the Montreal classification. Am J Gastroenterol. 2007; 102:577–588. [PubMed: 17338736]
13. Ahmad T, Armuzzi A, Bunce M, et al. The molecular classification of the clinical manifestations of Crohn's disease. Gastroenterology. 2002; 122:854–866. [PubMed: 11910336]
14. Cuthbert A, Fisher S, Croucher PJ, et al. The contribution of NOD2 gene mutations to the risk and site of disease in inflammatory bowel disease. Gastroenterology. 2002; 122:867–874. [PubMed: 11910337]
15. Lesage S, Zouali H, Cezard JP, et al. CARD15/NOD2 mutational analysis and genotype-phenotype correlation in 612 patients with inflammatory bowel disease. Am J Human Genet. 2002; 70:845–857. [PubMed: 11875755]
16. Prescott NJ, Fisher SA, Franke A, et al. A nonsynonymous SNP in ATG16L1 predisposes to ileal Crohn's disease and is independent of CARD15 and IBD5. Gastroenterology. 2007; 132:1665–1671. [PubMed: 17484864]
17. Márquez A, Núñez C, Martínez A. Role of ATG16L1 Thr300Ala polymorphism in inflammatory bowel disease: a study in the Spanish population and a meta-analysis. Inflamm Bowel Dis. 2009; 15:1697–1704. [PubMed: 19575361]
18. Fowler EV, Doecke J, Simms LA. ATG16L1 T300A shows strong associations with disease subgroups in a large Australian IBD population: further support for significant disease heterogeneity. Am J Gastroenterol. 2008; 103:2519–2526. [PubMed: 18671817]
19. Van Limbergen J, Russell RK, Nimmo ER. Autophagy gene ATG16L1 influences susceptibility and disease location but not childhood-onset in Crohn's disease in Northern Europe. Inflamm Bowel Dis. 2008; 14:338–346. [PubMed: 18088053]
20. 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; 40:955–962. [PubMed: 18587394]
21. Abreu MT, Taylor KD, Lin YC, et al. Mutations in NOD2 are associated with fibrostenosing disease in patients with Crohn's disease. Gastroenterology. 2002; 123:679–688. [PubMed: 12198692]
22. Hancock L, Beckly J, Geremia a, et al. Clinical and molecular characteristics of isolated colonic Crohn's disease. Inflamm Bowel Dis. 2008:1667–1677. [PubMed: 18521924]
23. Cosnes J. What is the link between the use of tobacco and IBD? Inflamm Bowel Dis. 2008; 14:S14–15. [PubMed: 18816683]
24. Waterman M, Xu W, Stempak JM, et al. Distinct and overlapping genetic loci in Crohn's disease and ulcerative colitis: correlations with pathogenesis. Inflamm Bowel Dis. 2010 epub ahead of print.
25. Tikkanen M, Surcel HM, Bloigu A, et al. Self-reported smoking habits and serum cotinine levels in women with placental abruption. Acta Obstet Gynecol Scand. 2010; 89:1538–1544. [PubMed: 21050148]
26. Yamazaki K, McGovern D, Ragoussis J, et al. Single nucleotide polymorphisms in TNFSF15 confer susceptibility to Crohn's disease. Hum Mol Genet. 2005 Nov 15.14:3499–3506. [PubMed: 16221758]
27. Tremelling M, Berzuini C, Massey D, et al. Contribution of TNFSF15 gene variants to Crohn's disease susceptibility confirmed in UK population. Inflamm Bowel Dis. 2008; 14:733–737. [PubMed: 18338776]
28. Thiébaut R, Kotti S, Jung C, et al. TNFSF15 polymorphisms are associated with susceptibility to inflammatory bowel disease in a new European cohort. Am J Gastroenterol. 2009; 104:384–391. [PubMed: 19174806]

29. Nakagome S, Takeyama Y, Mano S, et al. Population-specific susceptibility to Crohn's disease and ulcerative colitis; dominant and recessive relative risks in the Japanese population. Ann Hum Genet. 2010; 74:126–136. [PubMed: 20367632]

30. Kakuta Y, Ueki N, Kinouchi Y, et al. TNFSF15 transcripts from risk haplotype for Crohn's disease are overexpressed in stimulated T cells. Hum Mol Genet. 2009 Mar 15.18:1089–1098. [PubMed: 19124533]

Table 1
Distribution of clinical characteristics between CD disease locations

The clinical characteristics are shown with respect to the four major disease locations. L1, L2, L3 include subjects with L4 as a modifier. L4 refers to patients with only L4 disease location.

	Ileal CD (L1) *n = 288*	Colonic CD (L2) *n = 131*	Ileocolonic CD (L3) *n = 203*	Upper GI CD (L4) *n = 6*
Gender (male)	44%	50%	45%	92%
Race				
White	91%	84%	90%	92%
Black	8%	14%	9%	0%
Other	1%	2%	1%	8%
Smoking Habit				
Smoker	39%	25%	36%	33%
Ex-smoker	6%	13%	7%	17%
Non-smoker	55%	62%	57%	50%
Age at Diagnosis				
A1 (<17 y)	11%	13%	21%	0%
A2 (17-40y)	70%	60%	61%	92%
A3 (>40y)	19%	27%	18%	8%
Surgery	77%	41%	65%	67%

Table 2
Distribution of CD genotypes between CD disease locations

The distribution of genotypes is shown with respect to the four major disease locations. L1, L2, L3 include subjects with L4 as a modifier. Because there were only 6 L4 patients with only L4 disease, these patients were excluded from the analysis.

Gene or Chromosome	SNP	L1 *n = 288*	L2 *n = 131*	L3 *n = 203*
NOD2	(composite)			
R/R	rs2066847	12%	2%	7%
R/NR	rs2066844	30%	14%	23%
NR/NR	rs2066845	58%	84%	70%
ATG16L1	rs2241880			
R/R		35%	30%	36%
R/NR		45%	43%	45%
NR/NR		20%	27%	19%
IL23R	rs11209026			
R/R		96%	95%	92%
R/NR		4%	5%	8%
NR/NR		0%	0%	0%
IRGM	rs13361189			
R/R		4%	5%	2%
R/NR		27%	18%	20%
NR/NR		69%	77%	78%
STAT3	rs744166			
R/R		33%	37%	32%
R/NR		50%	45%	53%
NR/NR		17%	18%	15%
ICOSLG	rs762421			
R/R		17%	14%	12%
R/NR		48%	48%	54%
NR/NR		35%	38%	34%
21q21	rs1736135			
R/R		38%	45%	49%
R/NR		50%	37%	39%
NR/NR		12%	18%	12%
7p12	rs1456893			
R/R		52%	50%	50%
R/NR		41%	42%	42%
NR/NR		7%	8%	8%

Gene or Chromosome	SNP	L1 *n = 288*	L2 *n = 131*	L3 *n = 203*
IBD-5 locus	rs2188962			
R/R		20%	13%	24%
R/NR		45%	47%	47%
NR/NR		35%	40%	29%
ITLN1	rs2274910			
R/R		42%	46%	47%
R/NR		47%	44%	43%
NR/NR		11%	10%	10%
CCR6	rs2301436			
R/R		22%	22%	25%
R/NR		52%	54%	50%
NR/NR		26%	24%	25%
PTPN2	rs2542151			
R/R		7%	2%	4%
R/NR		27%	34%	34%
NR/NR		66%	64%	62%
PTPN22	rs2476601			
R/R		87%	85%	86%
R/NR		13%	13%	13%
NR/NR		0%	2%	0%
TNFSF15	rs4263839			
R/R		55%	55%	46%
R/NR		38%	43%	42%
NR/NR		7%	2%	11%
ORMDL3	rs2872507			
R/R		20%	27%	17%
R/NR		46%	47%	51%
NR/NR		34%	26%	32%
MST1	rs3197999			
R/R		11%	12%	9%
R/NR		42%	43%	41%
NR/NR		47%	45%	50%
C11orf30	rs7927894			
R/R		19%	17%	20%
R/NR		45%	51%	45%
NR/NR		36%	32%	34%
C13orf31	rs3764147			

Gene or Chromosome	SNP	L1 *n = 288*	L2 *n = 131*	L3 *n = 203*
R/R		7%	11%	7%
R/NR		36%	36%	44%
NR/NR		57%	53%	49%
PTGER4	rs4613763			
R/R		6%	2%	5%
R/NR		30%	27%	30%
NR/NR		64%	71%	65%
CDKAL1	rs6908425			
R/R		63%	61%	60%
R/NR		33%	34%	37%
NR/NR		4%	5%	3%
6q21	rs7746082			
R/R		8%	6%	11%
R/NR		36%	43%	39%
NR/NR		56%	51%	50%
1q24	rs9286879			
R/R		9%	5%	10%
R/NR		39%	50%	37%
NR/NR		52%	45%	53%
IL12B	rs10045431			
R/R		56%	58%	47%
R/NR		39%	33%	47%
NR/NR		5%	9%	6%
JAK2	rs10758669			
R/R		44%	45%	37%
R/NR		43%	44%	51%
NR/NR		13%	11%	12%
10p11	rs17582416			
R/R		14%	9%	14%
R/NR		44%	48%	46%
NR/NR		42%	43%	40%
NKX2-3	rs11190140			
R/R		26%	24%	24%
R/NR		49%	54%	49%
NR/NR		25%	21%	27%
ZNF365	rs10995271			
R/R		21%	14%	17%

Gene or Chromosome	SNP	L1 *n = 288*	L2 *n = 131*	L3 *n = 203*
R/NR		44%	53%	50%
NR/NR		35%	33%	33%
LOC651731	rs11584383			
R/R		60%	54%	54%
R/NR		31%	40%	41%
NR/NR		9%	6%	5%
MUC19	rs11175593			
R/R		0%	0%	0%
R/NR		7%	5%	5%
NR/NR		93%	95%	95%

Table 3

L1 +L3 (ileal CD and ileocolonic) vs. L2 (colonic only without ileal disease) logistic regression analysis with stepwise variable selection including genetic and clinical co-variates

The odds ratio (O.R.) and 95% confidence interval (C.I.) are shown.

	P-value	R/R vs. NR/NR Smoker vs. nonsmoker O.R. (95% C.I.)	R/NR vs. NR/NR Ex-smoker vs. nonsmoker O.R. (95% C.I.)	R/R vs R/NR Smoker vs. ex-smoker O.R. (95% C.I.)
NOD2	<0.0001	6.102 (1.855-20.074)	2.552 (1.479-4.402)	2.391 (0.671-8.525)
Smoking	0.0082	1.689 (1.071- 2.664)	0.647 (0.335-1.249)	2.611 (1.279-5.319)
TNSF15	0.0479	0.227 (0.068-0.763)	0.218 (0.064-0.737)	1.045 (0.694-1.573)