**SMAD3** gene variant is a risk factor for recurrent surgery in patients with Crohn's disease

Sharyle A. Fowler\textsuperscript{a}, Ashwin N. Ananthakrishnan\textsuperscript{a}, Agnes Gardet\textsuperscript{a,b}, Christine R. Stevens\textsuperscript{d}, Joshua R. Korzenik\textsuperscript{a}, Bruce E. Sands\textsuperscript{c}, Mark J. Daly\textsuperscript{b}, Ramnik J. Xavier\textsuperscript{a,b}, Vijay Yajnik\textsuperscript{a,*}

\textsuperscript{a} Gastrointestinal Unit, Massachusetts General Hospital, Harvard Medical School, Boston, MA, USA
\textsuperscript{b} Center for Computational and Integrative Biology, Massachusetts General Hospital, Boston, MA, USA
\textsuperscript{c} Division of Gastroenterology, Mount Sinai School of Medicine, New York, NY, USA
\textsuperscript{d} The Broad Institute of MIT and Harvard, Cambridge, MA, USA

Received 20 August 2013; received in revised form 12 December 2013; accepted 3 January 2014

**KEYWORDS**
Crohn's disease; Surgery for IBD; Genotype/clinical phenotype; NOD2; SMAD3

Abstract

**Background and aims:** More than 80% of Crohn's disease (CD) patients will require surgery. Surgery is not curative and rates of re-operation are high. Identification of genetic variants associated with repeat surgery would allow risk stratification of patients who may benefit from early aggressive therapy and/or post-operative prophylactic treatment.

**Methods:** CD patients who had at least one CD-related bowel resection were identified from the Prospective Registry in IBD Study at Massachusetts General Hospital (PRISM). The primary outcome was surgical recurrence. Covariates and potential interactions were assessed using the Cox proportional hazard model. Kaplan–Meier curves for time to surgical recurrence were developed for each genetic variant and analyzed with the log-rank test.

**Results:** 194 patients were identified who had at least 1 resection. Of these, 69 had two or more resections. Clinical predictors for repeat surgery were strictureing (HR 4.18, \( p = 0.022 \)) and penetrating behavior (HR 3.97, \( p = 0.024 \)). Smoking cessation was protective for repeat surgery (HR 0.45, \( p = 0.018 \)). SMAD3 homozygosity for the risk allele was also independently associated with increased risk of repeat surgery (HR 4.04, \( p = 0.001 \)). NOD2 was not associated with increased risk of surgical recurrence.

**Conclusion:** Strictureing and penetrating behavior were associated with increased risk of surgical recurrence, while smoking cessation was associated with a decreased risk. A novel association between SMAD3 and increased risk of repeat operation and shorter time to repeat surgery is observed.
1. Introduction

Crohn’s disease (CD) and ulcerative colitis (UC) affect greater than 1.4 million Canadians and Americans.1 The behavior and clinical course of CD is heterogeneous. At the time of diagnosis, approximately 40–50% of patients present with ileocolonic disease, and 30% with isolated small bowel or colonic disease.2 The location of CD is relatively stable over time and only 10–15% of patients will have a change in disease location 10 years after diagnosis.3 However, disease behavior can change significantly. Initially, the majority of patients present with inflammatory disease, but over time, this phenotype progresses to fistulizing or penetrating disease in up to 70% of patients.4 Treatment of these complications frequently requires surgical intervention, with more than 80% of patients with CD requiring surgery during their disease course.5 Surgery does not always provide a prolonged remission, with endoscopic recurrence in up to 90% of patients at 1 year and clinical recurrence rates of 30% by 3 years and 60% by 10 years.5 The rate of re-operation ranges between 20 and 70%, depending on length of follow-up.6–9

Despite advances in our understanding of CD, the course of the disease in individual patients remains difficult to predict. Several clinical risk factors have been associated with a complicated disease course, as defined in most studies as progression to a non-inflammatory disease behavior or dependence on steroids. These include active smoking, young age at diagnosis, extensive bowel involvement, need for steroids at diagnosis, perianal disease, and extra-intestinal manifestations.10–12

Advances in the field of genetics have lead to the identification of numerous susceptibility genes for inflammatory bowel disease (IBD). The first susceptibility gene identified, and the most extensively studied to date, is NOD2/CARD15.13,14 NOD2 has been shown to be associated with an increased risk of CD with a 2–4 fold increase with one mutant allele and a 40-fold increase with two mutant alleles.11 NOD2 has also displayed phenotypic and prognostic implications, with an association with ileal location and fistulizing behavior.15–24 A large Dutch tertiary multicenter cohort study described a more severe disease phenotype, a higher need for surgery and a younger age of onset of CD with an increasing number of risk alleles (NOD2, IBD5, DLG5, ATG16L1, IL23R).25 Sehgal et al. recently reported that mutations in IRGM correlated with increased frequency of surgery in patients with ileocolonic CD.26 Despite the identification of numerous susceptibility loci for CD, the phenotypic and prognostic implications of these loci are largely unknown.27

Recent controlled trials have highlighted newer approaches to the management of CD. The step-up top-down trial suggested that early exposure to biologic therapy (compared to introduction after failure of azathioprine therapy) allows for greater rates of mucosal healing.28 The SONIC trial further suggested that in newly diagnosed patients with CD, combination therapy with azathioprine and infliximab achieves superior clinical and endoscopic remission compared to monotherapy with either agent.29 In addition, the study by Regueiro et al. demonstrated that early initiation of biologics post-operatively can effectively prevent endoscopic and clinical recurrence.30 However, universal adoption of early aggressive therapy may not be appropriate as a significant proportion of patients with CD will not require such therapy and may have a good outcome with standard step-up care. In addition, the costs associated with such therapies are considerable and the side effects not insignificant. Thus, an unmet need in the field is the identification of markers that predict aggressive disease. Genetics offers an attractive strategy for such risk stratification. Genetic polymorphisms are fixed in an individual and are not dependent on duration of disease or lack of availability of phenotypic data. In addition, genetics is not modified by treatments or other interventions. However, systematic examination of the role of genetics in predicting the need for recurrent surgery in CD is lacking.

Previous studies assessing genetic risk factors for recurrent surgery have only evaluated a single, or limited number, of CD-associated risk alleles. With improvements in our understanding of the pathogenesis of CD, several different cellular/immune pathways have been identified as of particular importance. Thus, we performed this study to assess the influence of innate, autophagy, and IL-23/Th-17 pathway associated mutations on the clinical course of CD with respect to need for reoperation in CD. Identification of genetic risk factors associated with the need for reoperation in CD would allow risk stratification of patients who may benefit from early aggressive therapy and/or post-operative prophylactic treatment and may also provide insight into the pathophysiology of disease recurrence and identify novel therapeutic targets.

2. Materials and methods

2.1. Study subjects

The Prospective Registry in IBD Study at Massachusetts General Hospital (PRISM) is a large observational cohort of patients with IBD with detailed clinical and genetic information on approximately 1000 patients with IBD. This is an ongoing cohort based at a tertiary inflammatory bowel disease practice with recruitment initiated in 2005. At the time of enrollment to the study, information was collected in a retrospective fashion since the time of the diagnosis with CD. Patients were then followed prospectively. The diagnosis of CD was based on standard clinical, radiologic, endoscopic, and histologic criteria.31 Clinical information available included age at diagnosis, gender, disease duration, disease location and behavior as per the Montreal
classification, extra-intestinal manifestations (arthritis/arthralgia, ocular and cutaneous manifestations and primary sclerosing cholangitis), smoking status, family history, and date, type and indication of surgery.

Montreal classification of disease behavior and location was determined at the time of enrollment into PRISM and was assigned by a study gastroenterologist after review of the patient’s medical record. Upper GI location and perianal involvement were scored separately and were not mutually exclusive with ileal, colonic or ileocolonic location and inflammatory, penetrating and strictureing behavior respectively. Patients with indeterminate colitis were excluded. Patients initially diagnosed with UC who underwent colectomy with ileoanal pouch anastomosis were classified as CD only if mucosal inflammation was later demonstrated in the proximal ileal limb or upper GI tract. Patients who developed complications with fistulas or abscesses post-pouch formation without evidence of inflammation in the proximal bowel were excluded.

Smoking status was defined as non-smoker, former smoker or current smoker. Smoking status at time of surgical recurrence, or at time of enrollment into PRISM for patients with only a single surgery was included. Patients were classified as a current smoker if they smoked more than one cigarette per day. Patients were classified as former smokers if they had quit greater than one month prior to these time points.

Our primary outcome was undergoing a repeat CD-related bowel resection. Surgery for perianal disease including seton placement, fistulotomy and fecal diversion were excluded. Strictureplasty without resection and surgery for early complications of prior surgery were excluded. We further examined the indication for surgery defined as the primary reason for operation based on clinical presentation, preoperative diagnostic studies, and intraoperative findings. Surgeries prior to enrolment into the PRISM database were included in our analysis. As such, this data was collected in a retrospective manner.

All patients gave their written informed consent to participate in the study after approval of the project by the Partners Healthcare Institutional Review Board.

2.2. Genotyping

Genomic DNA was isolated from blood samples collected from patients from the PRISM registry. Oligonucleotides were synthesized and quality control using mass spectrometry was carried out at Integrated DNA Technologies. Genotyping was performed using a Sequenom genotyping platform (Sequenom, Inc., San Diego, CA). CD-associated SNPs available in the PRISM database that related to innate immunity, autophagy, and IL-23/Th-17 pathways were assessed. A total of 26 SNPs were analyzed (Table 1). As a Quality control step, the dataset was filtered to exclude SNPs with a Hardy–Weinberg p-value < 0.001 and a call rate > 95%. Individuals with < 80% genotyping were excluded.

2.3. Data analysis and statistics

Our primary endpoint was repeat surgery. Continuous variables were summarized using means and standard deviations. Categorical variables were expressed as proportions. The t test was used to compare continuous variables while the chi-square test (or Fisher’s exact test) was used for categorical variables. As there is limited knowledge about the function of several of the risk-alleles for CD, we modeled individual genes based on both dominant (homo/heterozygote vs. wild type) and recessive (homozygote vs. heterozygote/wild type) models. For the NOD2 allele, which can involve three different NOD2 allele, we introduced an additional stratum of compound heterozygote (two distinct SNPs at the NOD2 allele) in addition to homozygotes, heterozygotes, and wild type variants. Cox proportional hazards models were constructed to examine the effect of the various genetic polymorphisms adjusting for potential confounders. Time from first surgery to surgical recurrence (or to end of follow up if single surgery) was used as the time variable in the Cox proportional hazards model. Variables reaching a p value of \( p \leq 0.10 \) in the univariate analysis

### Table 1 Summary of all the CD risk alleles examined in the present study.

<table>
<thead>
<tr>
<th>Gene</th>
<th>SNP</th>
<th>Risk allele</th>
<th>MAF Single surgery n=125</th>
<th>MAF Repeat surgery n=69</th>
</tr>
</thead>
<tbody>
<tr>
<td>NOD2</td>
<td>R702W</td>
<td>T</td>
<td>0.09</td>
<td>0.16</td>
</tr>
<tr>
<td></td>
<td>G908R</td>
<td>C</td>
<td>0.06</td>
<td>0.07</td>
</tr>
<tr>
<td></td>
<td>Leu1007fsinC</td>
<td>C</td>
<td>0.06</td>
<td>0.07</td>
</tr>
<tr>
<td>CARD9</td>
<td>ITLN1</td>
<td>A</td>
<td>0.45</td>
<td>0.49</td>
</tr>
<tr>
<td></td>
<td>rs3766356</td>
<td>T</td>
<td>0.34</td>
<td>0.31</td>
</tr>
<tr>
<td></td>
<td>chr1_199134195</td>
<td>A</td>
<td>0.00</td>
<td>0.01</td>
</tr>
<tr>
<td>ATG16L1</td>
<td>rs2241880</td>
<td>A</td>
<td>0.40</td>
<td>0.43</td>
</tr>
<tr>
<td></td>
<td>rs1456896</td>
<td>C</td>
<td>0.30</td>
<td>0.30</td>
</tr>
<tr>
<td></td>
<td>rs921720</td>
<td>A</td>
<td>0.35</td>
<td>0.37</td>
</tr>
<tr>
<td>IRGM</td>
<td>rs11584383</td>
<td>C</td>
<td>0.26</td>
<td>0.25</td>
</tr>
<tr>
<td></td>
<td>rs7714584</td>
<td>G</td>
<td>0.15</td>
<td>0.13</td>
</tr>
<tr>
<td></td>
<td>rs13361189</td>
<td>C</td>
<td>0.16</td>
<td>0.14</td>
</tr>
<tr>
<td>LRRK2-MUC19</td>
<td>rs11175593</td>
<td>T</td>
<td>0.03</td>
<td>0.03</td>
</tr>
<tr>
<td></td>
<td>chr12_39026953</td>
<td>G</td>
<td>0.03</td>
<td>0.02</td>
</tr>
<tr>
<td></td>
<td>chr12_39108062</td>
<td>A</td>
<td>0.04</td>
<td>0.03</td>
</tr>
<tr>
<td>IL23R</td>
<td>rs7517847</td>
<td>G</td>
<td>0.31</td>
<td>0.32</td>
</tr>
<tr>
<td></td>
<td>rs11209026</td>
<td>A</td>
<td>0.04</td>
<td>0.01</td>
</tr>
<tr>
<td>JAK2</td>
<td>rs10758669</td>
<td>C</td>
<td>0.40</td>
<td>0.36</td>
</tr>
<tr>
<td></td>
<td>rs41316003</td>
<td>A</td>
<td>0.00</td>
<td>0.01</td>
</tr>
<tr>
<td>STAT3</td>
<td>rs3764147</td>
<td>G</td>
<td>0.31</td>
<td>0.27</td>
</tr>
<tr>
<td></td>
<td>rs744166</td>
<td>G</td>
<td>0.36</td>
<td>0.41</td>
</tr>
<tr>
<td></td>
<td>rs12942547</td>
<td>G</td>
<td>0.36</td>
<td>0.40</td>
</tr>
<tr>
<td>TYK2</td>
<td>rs12720356</td>
<td>C</td>
<td>0.12</td>
<td>0.14</td>
</tr>
<tr>
<td></td>
<td>rs11085735</td>
<td>A</td>
<td>0.10</td>
<td>0.10</td>
</tr>
<tr>
<td>SMAD3</td>
<td>rs17293632</td>
<td>T</td>
<td>0.25</td>
<td>0.30</td>
</tr>
<tr>
<td>TNFSF15</td>
<td>rs6478108</td>
<td>C</td>
<td>0.32</td>
<td>0.25</td>
</tr>
</tbody>
</table>

MAF: minor allele frequency.
were entered into the multivariate analysis where a p-value < 0.05 indicated independent statistical significance. Correction for multiple testing was not performed, as this was an exploratory, hypothesis generating study. Kaplan–Meier curves for time to repeat surgery were developed for each genetic variant and compared using the log-rank test.

3. Results

3.1. Patient population

One hundred and ninety-four patients who had undergone a CD-related bowel resection were identified in the PRISM database. Mean age at diagnosis was 25 ± 12 years (range 5–76 years) with mean disease duration of 19 ± 13 years (range 1–52 years). Fifty-five percent (n = 107) of patients were male. Fifty-one (26%) patients had ileal disease, 15 (8%) had colonic disease, and 128 (66%) had ileocolonic disease. Twenty-two (11%) had upper gastrointestinal involvement. Thirty-two (16%) patients had purely inflammatory disease, 67 (35%) had stricturing disease, and 95 (49%) patients had penetrating disease based on the Montreal classification of disease behavior.

3.2. Indications for surgery

Of the 194 CD patients requiring an initial CD-related resection, 69 (36%) required a second surgery. Indications for initial and repeat surgery are summarized in Fig. 1. Indication for first surgery was inflammatory in 53 patients (27%), stricturing in 59 patients (30%) and penetrating in 77 patients (40%). However, for second surgery, over half of operations were for stricturing disease (52%). Resection for tumor/dysplasia remained an uncommon indication comprising only 3% of initial and repeat operations.

3.3. Predictors for surgical recurrence

Clinical characteristics of CD patients according to surgical recurrence are shown in Table 2. In univariate analysis, significant clinical variables associated with need for repeat surgery were behavior (p = 0.005), perianal disease (p = 0.03), and smoking status (p = 0.01). Time to first surgery was similar between the two groups (single surgery 6.5 ± 6.7 years, repeat surgery 6.0 ± 6.7, p = 0.76). Patients with only a single resection also had similar follow-up time after the first surgery as those who underwent repeat resection (9 ± 9 vs. 11 ± 9 years, p = 0.25). Genetic variables significantly associated with need for repeat surgery on univariate analysis included SMAD3 homozygosity for the risk allele (HR 4.88, 95% CI 2.26–10.53) and LRRK2-MUC19 homozygosity for the risk allele (HR 7.618, 95% CI 1.00–58.81, p = 0.05). All variables reaching a p value of ≤0.10 in the univariate analysis were entered into the multivariate analysis. In the multivariate analysis, perianal disease and LRRK2-MUC19 were no longer statistically significant. Stricturing (HR 4.18, 95% CI 1.23–14.19, p = 0.022) and penetrating behavior (HR 3.97, 95% CI 1.20–13.1, p = 0.024), and SMAD3 homozygosity for the risk allele (HR 4.04, 95% CI 1.77–9.21, p = 0.001) remained significantly and independently associated with surgical recurrence. Former smoking was associated with a statistically significant reduction in repeat operation (HR 0.45, 95% CI 0.23–0.87, p = 0.018). The Kaplan–Meier survival curve for repeat surgery according to SMAD3 risk allele status is shown in Fig. 2. All other genetic variants were not associated with risk of recurrent surgery.

4. Discussion

The clinical course of CD is variable, and can be difficult to predict at time of diagnosis. The identification of numerous
susceptibility genes for IBD has provided insight into the pathogenesis of IBD. Ideally, correlation between specific genotypes and clinical phenotypes would allow for more advanced classification of patients beyond disease location and behavior. For example, certain genetic profiles may identify patients at high risk of surgery, or those at risk of rapid recurrence after first surgery. This information would help clinicians and patients with decisions regarding need for aggressive up front therapy, or initiation of post-operative prophylactic therapy. In the present study we assess the influence of innate, autophagy and IL-23/IL-17 pathway associated mutations on surgical recurrence in 194 patients who had undergone one CD-related bowel resection.

Our study confirms a high rate of repeat resection with 36% of patients requiring a second operation. Rates of re-operation at 5 and 10 years were 23% and 30% respectively. These rates are in keeping with the literature with reported rates of repeat surgery between 11–32% at 5 years and 20–44% at 10 years. Penetrating and stricturing disease behaviors were the only clinical variables associated with need for repeat surgery. Current smoking was not associated with increased risk of surgical recurrence. However, smoking cessation was identified as the only clinical or genetic factor protective for recurrent surgery. Homozygosity for the SMAD3 risk allele was associated with increased risk of recurrent surgery and earlier time to repeat operation.

We attempted to account for all variables known to influence the risk of surgery in CD. Unfortunately, as with most previously published studies assessing surgical risk, it is difficult to account for the effects of medication. As a component of our data was gathered retrospectively, medication use was not standardized, and key variables such as indication for medication use, time of medication use, and response to medication, were not recorded uniformly. As such, we were not able to include medication use in our analysis. SMAD3 homozygosity was relatively common, present in 13% of patients with repeat surgery (n = 9). The identification of SMAD3 as a risk factor for repeat surgery in CD is of particular interest as SMAD3 is a mediator of signaling via transforming growth factor-beta (TGF-β), a multifunctional cytokine regulating a variety of important biological responses. SMAD3 proteins are signal transducers and transcriptional modulators that mediate multiple signaling pathways. TGF-β plays an important role in numerous biological responses including cell growth and differentiation, apoptosis, cell migration, immune cell function and extracellular matrix production.34,35 Mice lacking TGF-β develop widespread inflammation, including in the gut, and die early in life,36 and animals whose T cells cannot respond to TGF-β also die of wasting disease and gut inflammation.37 As such, TGF-β is thought of as a master negative regulator of intestinal inflammation. However, paradoxically, TGF-β has been shown to be increased in IBD tissues, with mucosal inflammation proceeding unchecked.38 SMAD7, an inhibitor of TGF-β signaling, appears to play a key role in modulation of TGF-β signaling in IBD. Monteleone et al., demonstrated that Smad7 is overexpressed in IBD mucosa and purified mucosal T cells, and in both whole tissue and isolated cells there is defective TGF-β signaling as measured by reduced phosphor-SMAD3 immunoreactivity. Specific antisense oligonucleotides for SMAD7 reduced SMAD7 protein expression and restored TGF-β signaling.39 Animal models of several different fibrotic diseases have shown that agents that block TGF-β function, such as TGF-β antibodies or antisense oligonucleotides, reduce the fibrotic response.40 This effect appears to be SMAD3 dependent based on animal models of fibrosis where loss of SMAD3 results in a diminished fibrotic response. For example, in radiation-induced fibrosis of the skin, SMAD3 KO mice showed less influx of mast cells, neutrophils, macrophages, and myofibroblasts and decrease expression of TGF-β.41 As well, after cutaneous irradiation, less scarring was noted in SMAD3 KO mice compared to WT controls.42 Through this mechanism of pathological progressive fibrosis and scarring, the SMAD3 risk allele may result in abnormal wound healing at a site of injury, namely at the surgical site, with resultant accelerated disease recurrence and need for repeat surgery. By this mechanism, SMAD3 mutations may result in increased risk of second surgery, without affecting a patient’s risk of undergoing initial surgery. However, this mechanism of action for the SMAD3 risk allele is speculative. Functional studies on this SNP to identify tissue specific enhancer sequence are currently being conducted.

The identification of a novel association between the previously identified CD-associated SMAD3 risk variant and increased risk of repeat surgery provides insight into a possible mechanism for post-operative disease recurrence in patients with CD. These findings need to be interpreted with caution, as this study was designed as an exploratory, hypothesis generating study. These findings will need to be confirmed in a replication cohort. If this association can be confirmed, it would allow for early risk stratification of patients who may benefit from aggressive therapy and/or post-operative prophylactic treatment. As well, inhibition of SMAD3 signaling at sites of stricturing disease may be an excellent therapeutic target to interfere with the fibrotic response locally, without interfering with other TGF-β mediated signaling pathways systemically.

In summary, our data identifies stricturing and penetrating behavior as the main clinical predictors for surgical recurrence, and smoking cessation as a protective factor for surgical recurrence. We also identified a new association between SMAD3 risk alleles and increased risk of repeat operation and shorter time to repeat surgery. The identification of SMAD3 as a risk allele for repeat operation is of particular interest as it may represent a new therapeutic target specifically for the treatment of fibrotic/stricturing CD.

Figure 2 Survival curve for time to second surgery according to SMAD3 status.
Acknowledgments

Funding support: The project described was supported by Grant Number P30DK043351 from the National Institute of Diabetes and Digestive and Kidney Diseases. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institute of Diabetes and Digestive and Kidney Diseases or the National Institutes of Health.

Statement of authorship: SF, AA, and VY were involved in the conception and design of the study, acquisition of data, analysis and interpretation of the data, and drafting the article. AG carried out sample analysis, and was involved in drafting the article. CS carried out sample analysis. JK, BS, RX, and MD were involved in the conception and design of the study and revising the manuscript. All authors have read and approved the final manuscript. The authors do not have any financial conflict of interest to declare.

References

32. Silverberg MS, Satsangi J, Ahmad T, Arnott ID, Bernstein CN, Brant SR. Toward an integrated clinical, molecular and...


