Journal of Crohn's and Colitis, 2016, 209–215
doi:10.1093/ecco-jcc/jjv197
Advance Access publication October 28, 2015
Original Article

Original Article

IBD Genetic Risk Profile in Healthy First-Degree Relatives of Crohn's Disease Patients

David Kevans,[a,b] Mark S. Silverberg,[a,b] Krzysztof Borowski,[a] Anne Griffiths,[c] Wei Xu,[d] Venus Onay,[a] Andrew D. Paterson,[e] Jo Knight[f,g,h] Ken Croitoru;[a,b] on behalf of the GEM Project

[a]Zane Cohen Centre for Digestive Diseases, Mount Sinai Hospital, Toronto, ON, Canada [b]Division of Gastroenterology, Department of Medicine, University of Toronto, Toronto, ON, Canada [c]Division of Gastroenterology, Hepatology and Nutrition, Department of Paediatrics, The Hospital for Sick Children, Toronto, ON, Canada [d]Dalla Lana School of Public Health, University of Toronto, Toronto, ON, Canada [e]The Hospital for Sick Children Research Institute, The Hospital for Sick Children, Toronto, ON, Canada [f]Campbell Family Mental Health Research Institute, Centre for Addiction and Mental Health, Toronto, ON, Canada [g]Institute of Medical Science, University of Toronto, Toronto, ON, Canada [h]Department of Psychiatry, University of Toronto, Toronto, ON, Canada.

Corresponding author: Dr Kenneth Croitoru, MDCM, Zane Cohen Centre for Digestive Diseases, Division of Gastroenterology, Mount Sinai Hospital, 600 University Ave. Room 437, Toronto M5G 1X5, ON, Canada. Tel: 416-586-4800 ext. 7454; fax: 416-586-4747; email: KCroitoru@mtsinai.on.ca

Abstract

Background: Family history provides important information on risk of developing inflammatory bowel disease [IBD], and genetic profiling of first-degree relatives [FDR] of Crohn's disease [CD]-affected individuals might provide additional information. We aimed to delineate the genetic contribution to the increased IBD susceptibility observed in FDR.

Methods: $N = 976$ Caucasian, healthy, non-related FDR; $n = 4997$ independent CD; and $n = 5000$ healthy controls [HC]; were studied. Genotyping for 158 IBD-associated single nucleotide polymorphisms [SNPs] was performed using the Illumina Immunochip. Risk allele frequency [RAF] differences between FDR and HC cohorts were correlated with those between CD and HC cohorts. CD and IBD genetic risk scores [GRS] were calculated and compared between HC, FDR, and CD cohorts.

Results: IBD-associated SNP RAF differences in FDR and HC cohorts were strongly correlated with those in CD and HC cohorts, correlation coefficient 0.63 (95% confidence interval [CI] 0.53 - 0.72), $p = 9.90 \times 10^{-19}$. There was a significant increase in CD-GRS [mean] comparing HC, FDR, and CD cohorts: 0.0244, 0.0250, and 0.0257 respectively [$p < 1.00 \times 10^{-7}$ for each comparison]. There was no significant difference in the IBD-GRS between HC and FDR cohorts [$p = 0.81$]; however, IBD-GRS was significantly higher in CD compared with FDR and HC cohorts [$p < 1.00 \times 10^{-10}$ for each comparison].

Conclusion: FDR of CD-affected individuals are enriched with IBD risk alleles compared with HC. Cumulative CD-specific genetic risk is increased in FDR compared with HC. Prospective studies are required to determine if genotyping would facilitate better risk stratification of FDR.

Key Words: Crohn's disease; first-degree relatives; genotyping

1. Introduction

Inflammatory bowel diseases are a group of heterogeneous disorders that result in chronic intestinal inflammation affecting the digestive tract. The major forms of these disorders, ulcerative colitis [UC] and Crohn's disease [CD], manifest both distinct and overlapping clinical and pathological characteristics. In North America, incidence rates range from 2.2 to 14.3 per 100 000 person-years for UC and 3.1 to 14.6 per 100 000 person-years for CD. As many as 200 000 Canadians,[1] 1.4 million persons in the USA and 2.2 million Europeans suffer from these diseases.[2] The aetiology of inflammatory bowel disease [IBD] has yet to be fully elucidated; however, it is currently thought to arise as a result of a dysregulated immune response in genetically susceptible individuals to environmental triggers such as a dysbiotic host microbiota.

Epidemiological studies have provided compelling evidence that genetic factors contribute to the pathogenesis of IBD. First-degree relatives [FDR] of patients with IBD have approximately a 3 to 20-fold greater likelihood of developing the disease than the general population.[3,4,5] The siblings of patients with CD have an estimated relative risk of developing CD up to 35 times the background population risk.[4] Offspring where both parents have CD have even greater risk with approximately 36% likely to develop the disease.[6] Ashkenazi Jews also have an increased risk of IBD.[7,8] However, most significantly, twin studies have indicated that heritability is high in CD with concordance rates in monozygotic of 27–50% compared with 2–4% in dyzygotic twins.[9,10,11] Consistent with epidemiological predictions, over 70 IBD-associated genetic associations were identified in candidate gene,[12] linkage,[13] and genome-wide association studies [GWAS].[14,15,16,17,18,19,20,21] More recent meta-analysis of CD and UC genome-wide association scans, followed by validation of significant findings using Immunochip genotyping [Illumina Inc.][22] of independent case-control cohorts identified further associations, increasing the number of IBD-associated risk loci to 201.[23,24]

Family history alone provides important information regarding an individual's risk of developing IBD, and genetic profiling of FDR of CD-affected individuals might provide additional information on disease susceptibility risk. Since the genetic risk of IBD has been defined through GWAS in large case-control studies, the representation of these IBD risk alleles in FDR needs to be defined. Genetic risk in FDR of CD-affected individuals is assumed to be increased based on previous epidemiological data,[3,4,5] but it is not known to what degree this is the case. In this study we aimed to define the genetic contribution to the increased IBD susceptibility observed in FDRs of CD-affected individuals by comparing the carriage of 163 known IBD-associated risk loci,[23] between FDR, CD-affected individuals, and a healthy control cohort [HC]. In addition we examined differences in cumulative genetic risk between HC, FDR, and CD cohorts by comparing IBD and CD genetic risk scores between these cohorts.

2. Material and Methods

2.1. Study population

FDR were enrolled by identifying probands with CD and recruiting their healthy FDR siblings and offspring from sites around Canada and the USA as part of the Genetics, Environmental, Microbial [GEM] Project [www.gemproject.ca]. Only FDR between 6 and 35 years of age were enrolled. The diagnosis of CD in probands was established by review of history and clinical information available to the recruiting physicians. FDR were defined as full sibling or offspring as declared by the proband and subject. For FDR, the criteria of a 'healthy' or 'disease-free' state was defined by the GEM Project clinical sub-committee and includes lack of any history of gastrointestinal diseases. At study entry, each subject completed a standard questionnaire to ensure the absence of symptoms related to gastrointestinal illness. Eligibility included a lack of gastrointestinal symptoms considered significant, such as: unintentional weight loss in the past 3 months greater than 15% of baseline; recurring abdominal pain more than once weekly for greater than 3 months in the past year; diarrhoea more than three times per day of greater than 3 months' duration in the past year; and blood in stools. Other exclusion criteria include diagnosis with diabetes or pregnancy at time of enrolment. As of January 2014, there were 1037 FDR with available genotyping data. From this cohort, only Caucasian FDR who were the first recruited individual in a family were included, resulting in a final study cohort of 978 [72% siblings, 28% offspring]. The baseline demographics of the FDR cohort are detailed in Supplementary Table 1, available as Supplementary data at *ECCO-JCC* online.

A healthy control [HC] cohort and Crohn's disease [CD] cohort from the International IBD Genetics Consortium [IIBDGC] repository were also utilised in analyses. Subjects in the CD cohort were not related to individuals in the FDR cohort. From the CD cohort a random sample of 5000 CD Caucasian subjects were selected from the 14 763 CD-affected individuals, with Immunochip genotyping, examined by Jostins *et al.*[23] Similarly for the HC cohort, 5000 Caucasian HC were randomly selected from a total of 15 977 healthy controls, with Immunochip genotyping, examined by Jostins *et al.*[23] Data on age at enrolment / diagnosis and proportion of subjects of Jewish ethnicity in the HC and CD cohorts were not available [Supplementary Table 1].

2.2. Genotyping—technique and quality control

For FDR, one purple-top [EDTA] vacutainer [BD Inc., NJ, USA] containing whole blood was collected from each subject and genomic DNA extracted using the Gentra Puregene Blood Kit [Quiagen, CA, USA]. Each DNA sample was quantified by Nanodrop at 20 ng/µl and aliquoted into 96-well reaction plates. Single nucleotide polymorphism [SNP] genotyping was performed using the ImmunoChip [Illumina Inc., San Diego, CA, USA]. SNP quality control was performed using PLINK:[25] SNPs with missingness >5 % and Hardy-Weinberg equilibrium [HWE] outliers [$p \leq 1 \times 10^{-6}$] were removed. All pairs of subjects were tested for an identity-by-descent value greater than 0.1875 to avoid relatedness in the dataset. Two subjects from FDR and three subjects from the CD cohort were removed due to an identity-by-descent value greater than 0.1875. The remaining samples were used to analyse population structure in the cohorts. Population structure was assessed using HapMap Caucasians and multiple dimension scaling [MDS]. MDS was performed on 10 725 SNPs with r2 < 0.2 with a cutoff of 8 standard deviations [SD] for outliers. No outliers were identified and therefore no samples were removed based on population structure analysis. [Supplementary Figure 1, available as Supplementary data at *ECCO-JCC* online] Following quality control as described above, 4997 of the CD cohort, 976 of the FDR cohort, and 5000 of the HC cohort were available for analysis. In this report we focused analyses on the 163 recently reported IBD-associated Immunochip SNPs of which 158 passed quality control assessments [Supplementary Table 2, available as Supplementary data at *ECCO-JCC* online].[23] As an extra validation step, the insertion C polymorphism [rs2066847] in nucleotide-binding oligomerisation domain-containing protein 2 [*NOD2*] was genotyped using the TaqMan 5'Nuclease Allelic Discrimination assay [Applied Biosystems, CA, USA]. TaqMan genotyping of

rs2066847 was in complete agreement with Immunochip genotyping of rs5743293; therefore these were considered to be identical variants. The variant on the Immunochip with id rs2066847 did not agree with TaqMan genotyping of this variant and was thus discarded.[13,23]

2.3. Comparisons of IBD-associated SNP allele frequency in FDR, CD, and HC cohorts

PLINK was applied to provide summary statistics for each of the SNPs, including allele frequency, genotype distribution, and HWE test for each cohort. Further analyses were performed using SAS v.9.2 [SAS Institute, NC, USA]. Logistic regression models were applied for each of the SNPs to compare FDR, HC, and CD cohorts using an additive genetic model. Odds ratios [OR] were estimated for each comparison with accompanying *p*-values. In this analysis, correction for multiple testing was performed using Bonferroni correction [based on the number of SNPs included in the analysis] with a threshold for significance of $p = 3.2 \times 10^{-4}$. For the 158 IBD-associated SNPs, the difference in risk allele frequencies [RAF] between HC and CD cohorts, and HC and FDR cohorts, was calculated. The correlation between these RAF differences was then assessed using Pearson's correlation with an accompanying *p*-value.

2.4. Genetic risk scores—calculation and comparison across study cohorts

A weighted GRS was calculated, which utilised the 158 available IBD risk loci on the Illumina Immunochip. For each risk allele, the genetic burden conferred was calculated, using PLINK, by multiplying the log transformation odds ratio [OR] of its association with disease by the allele dose [wild type, heterozygote or homozygote]. ORs were taken from the report by Jostins *et al.*[23] Wild type, heterozygote and homozygote genotype for a given allele were assigned weights of 0, 1 and 2 respectively. The cumulative genetic risk score [GRS] was then calculated by adding together the contribution of each of the risk loci [log OR x allele dose] and then dividing by the number of non-missing SNPs. Two risk scores were calculated: an IBD-GRS which summed the contribution of 158 IBD risk loci; and a CD-GRS which summed the contribution of specifically CD-associated risk loci [$n = 30$]. The mean IBD-GRS and CD-GRS were compared between HC, FDR, and CD cohorts using t tests with uncorrected *p*-values reported in these analyses and a threshold for significance set at $p < 0.05$. Comparisons of IBD-GRS and CD-GRS between various subgroups of the FDR cohort were also made. Although CD probands related to recruited FDR were not included in the study, baseline demographic data on these individuals were available. Using these data, the correlation of FDR genetic risk scores with the age at diagnosis of their corresponding CD proband was performed [Pearson correlation coefficient with accompanying *p*-value]. The mean IBD-GRS and CD-GRS were also compared between: sibling and offspring FDR; and FDR whose age at enrolment was less than 17 years compared with those of 17 years or greater; and t tests with uncorrected *p*-values were reported for these comparisons [threshold for significance set at $p < 0.05$] (SAS v.9.2 [SAS Institute, NC, USA]).

2.5. IBD and CD genetic risk categories in FDR and HC cohorts

For IBD genetic risk, FDR and HC cohorts were grouped into four genetic risk categories [category 1, lowest genetic risk to category 4, highest genetic risk]. For these IBD genetic risk category assignments, the CD cohort was considered a reference cohort, and IBD-GRS reference ranges were developed by dividing this cohort in quartiles based on IBD-GRS values. The reference range for each IBD-GRS quartile derived from the CD cohort was then applied to HC and FDR cohorts, allowing subjects in these cohorts to be assigned to one of the four IBD genetic risk categories. A similar approach was taken to categorise subjects in the FDR and HC cohorts into four categories of CD genetic risk categories [category 1, lowest genetic risk to category 4, highest genetic risk]; however, in this instance, reference ranges for each genetic risk category were derived by dividing the CD cohort into quartiles based on their CD-GRS values. For IBD and CD genetic risk categories, the proportion of subjects in HC and FDR cohorts in each of these categories was compared using the chi-square test for trend with *p*-values < 0.05 considered significant (SAS v.9.2 [SAS Institute, NC, USA]).

2.6. Ethical considerations

The study was reviewed and approved by the research ethics boards at all recruitment sites, and all subjects provided written informed consent for inclusion in the study.

3. Results

3.1. Comparison of IBD-associated allele frequencies in HC, FDRs, and CD

Comparing FDR and HC cohorts, the RAF of two of the 158 IBD-associated SNPs assessed differed significantly in RAF: rs2188962 [IBD5 locus], OR 1.21, $p = 9.2 \times 10^{-5}$ and rs3764147 9*Laccase Domain Containing 1 [LACC1]*], OR 1.23, $p = 1.9 \times 10^{-4}$). Comparing FDR and CD cohorts, one of the 158 IBD-associated SNPs assessed differed significantly in RAF: rs6863411 (*Sprouty Homolog 4* [*SPRY4*], *Nedd4 family interacting protein 1* [*NDFIP1*]), OR 0.77, $p = 2.3 \times 10^{-7}$). Aside from those described above, none of the other SNPs assessed differed in RAF, comparing FDR with HC or CD cohorts, at a level which reached the pre-defined *p*-value threshold for significance [$p = 3.2 \times 10^{-4}$] [Supplementary Tables 3 and 4, available as Supplementary data at *ECCO-JCC* online]. For the 158 IBD-associated SNPs evaluated, the differences in RAF between FDR and HC cohorts were strongly correlated with the differences in RAF between CD and HC cohorts; correlation coefficient 0.63 (95% confidence interval [CI] 0.53 - 0.72), $p = 9.90 \times 10^{-19}$; and RAF differences were generally in the same direction but of a lower magnitude[Figure 1]. In contrast, where the same comparison was made using 140 randomly selected non-IBD associated SNPs, the differences in minor allele frequency [MAF] between FDR and HC cohorts versus CD and HC cohorts were less strongly correlated (correlation coefficient 0.23 [95% CI 0.07 - 0.38], $p = 0.003$) [Supplementary Figure 2, available as Supplementary data at *ECCO-JCC* online].

3.2. Genetic risk score comparisons in HC, FDR, and CD cohorts

CD-associated genetic risk was significantly greater in the FDR compared with the HC cohort, CD-GRS [mean] 0.0250 versus 0.0244 respectively, $p = 8.0 \times 10^{-8}$. Similarly, CD-associated genetic risk was significantly greater in the CD compared with the HC cohort, CD-GRS [mean] 0.0257 versus 0.0244 respectively, $p = 1.1 \times 10^{-91}$. CD-associated genetic risk was also significantly greater in the CD compared with the FDR cohort, CD-GRS [mean] 0.0257 versus 0.0250 respectively, $p = 2.0 \times 10^{-9}$. In contrast IBD-associated genetic risk was similar comparing FDR and HC cohorts, IBD-GRS [mean] 0.0231 versus 0.0231 respectively, $p = 0.80$. IBD-associated genetic risk was greater in the CD compared with the HC cohort, IBD-GRS

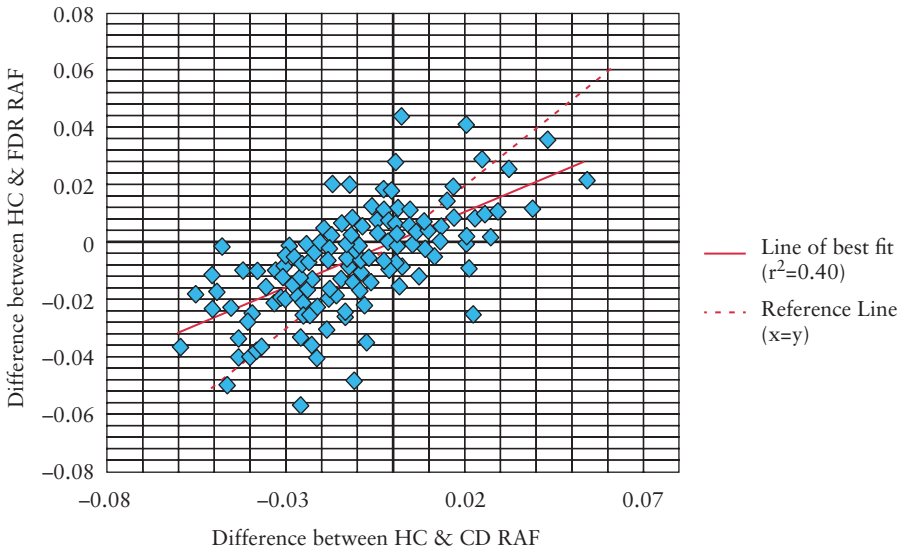

Figure 1. Correlation between the difference in risk allele frequency of 158 IBD-associated SNPs in FDR and HC versus CD and HC cohorts. For the 158 IBD-associated SNPs evaluated, the differences in RAF between FDR and HC cohorts were strongly correlated with the differences in RAF between CD and HC cohorts; correlation coefficient 0.63 [95% confidence 0.53 - 0.72], $p = 9.90 \times 10^{-19}$. IBD, inflammatory bowel disease; HC, healthy controls; FDR, first-degree relatives; CD, Crohn's disease; SNPs, single nucleotide polymorphisms; RAF, risk allele frequency.

[mean] 0.0240 versus 0.0231 respectively, $p = 1.1 \times 10^{-280}$, and in the CD compared with the FDR cohort, IBD-GRS [mean] 0.0240 versus 0.0231 respectively $p = 7.1 \times 10^{-82}$. [Table 1 and Figure 2].

3.3 Evaluation of differences in genetic risk scores between subgroups of FDRs

There was a modest but significant inverse correlation between CD proband age of diagnosis and the IBD-GRS of their corresponding FDR, correlation coefficient -0.08 [95% CI -0.14 - 0.01], $p = 0.02$ [Figure 3]. An inverse correlation was also observed for the CD-GRS but this did not reach statistical significance: correlation coefficient -0.04 [95% CI -0.11 - 0.024], $p = 0.21$ [Supplementary Figure 3, available as Supplementary data at *ECCO-JCC* online]. There was a trend toward a modestly higher IBD-GRS in FDR less than 17 years old at study enrolment compared with those of 17 years or greater [$p = 0.08$]; however CD-GRS was similar between these subgroups of FDR. IBD-GRS and CD-GRS in the FDR cohort were similar comparing siblings and offspring [Supplementary Table 5, available as Supplementary data at *ECCO-JCC* online].

3.4. IBD and CD genetic risk categories in HC and FDR cohorts

There was a modest difference in the proportion of subjects assigned to each IBD genetic risk category comparing the FDR and HC cohorts [$p = 0.04$] with, for example, 8% versus 7% of each cohort respectively assigned to IBD genetic risk category 4. There was a more marked difference in the proportion of subjects assigned to each CD genetic risk category comparing the FDR and HC cohorts [$p = 5.1 \times 10^{-15}$] with, for example, 21% versus 14% of each cohort respectively assigned to the CD genetic risk category 4 [Table 2].

4. Discussion

We genotyped healthy FDR of CD-affected individuals for 158 IBD risk loci, which is a comprehensive assessment of IBD genetic risk in FDR based on known IBD risk alleles. We demonstrate an enrichment of IBD-associated risk alleles in FDR compared with HC as evidenced by the significant correlation in risk allele frequency differences between FDR and HC, and CD and HC, cohorts. Cumulative CD-specific genetic risk, expressed as a weighted genetic risk score, is increased in FDR compared with healthy controls, whereas cumulative IBD genetic risk is similar between these groups.

Table 1. Comparison of CD-GRS and IBD-GRS between HC, FDR, and CD cohorts.

	CD genetic risk score			*IBD genetic risk score*		
	Mean	**SD**	**Range**	**Mean**	**SD**	**Range**
HC	0.0244	0.003	0.013 – 0.040	0.0231	0.001	0.018 – 0.027
FDR	0.0250	0.003	0.014 – 0.035	0.0231	0.001	0.019 – 0.027
CD	0.0257	0.003	0.012 – 0.039	0.0240	0.001	0.020 – 0.029

SD, standard deviation; HC, healthy controls; FDR, first-degree relatives; CD, Crohn's disease; GRS, genetic risk score.

CD-GRS pairwise comparisons: FDR versus HC cohort, $p = 8.0 \times 10^{-8}$; CD versus HC cohort, $p = 1.1 \times 10^{-91}$; CD versus FDR cohort, $p = 2.0 \times 10^{-9}$. IBD-GRS pairwise comparisons: FDR versus HC cohort, $p = 0.8$; CD versus HC cohort, $p = 1.1 \times 10^{-280}$; and CD versus FDR cohort, $p = 7.1 \times 10^{-82}$.

We demonstrated that two CD-associated variants occurred more frequently in FDR than HC, rs2188962 in the region of the IBD5 locus and rs3764147 in the region of *LACC1*, both of unknown function. One variant was enriched in the CD compared with the FDR cohort, rs6863411 [*SPRY4, NDFIP1*]. *NDFIP1* is known to modulate T helper-17 cell differentiation,[26] therefore it is possible that the increased representation of this variant in individuals with CD compared with FDR contributes to the dysregulated immune response observed in individuals with CD; however, this finding requires further validation.

In addition to evaluating allele frequencies of individual SNPs, we also performed a global assessment to determine whether the RAF of each of the 158 IBD-associated SNPs were similar in FDR and CD cohorts. This assessment showed that FDR are enriched with IBD risk alleles but not to the same degree as CD-affected

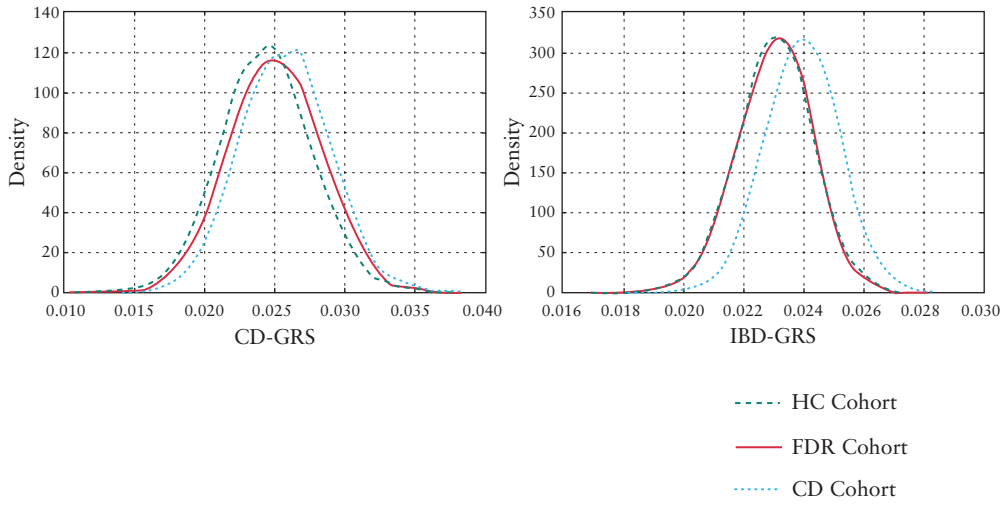

Figure 2. Distribution of IBD-GRS and CD-GRS in FDR, HC and CD cohorts; distributions of the CD-GRS and IBD-GRS in HC, FDR and CD cohorts. IBD, inflammatory bowel disease; HC, healthy controls; FDR, first-degree relatives; CD, Crohn's disease; GRS, genetic risk score.

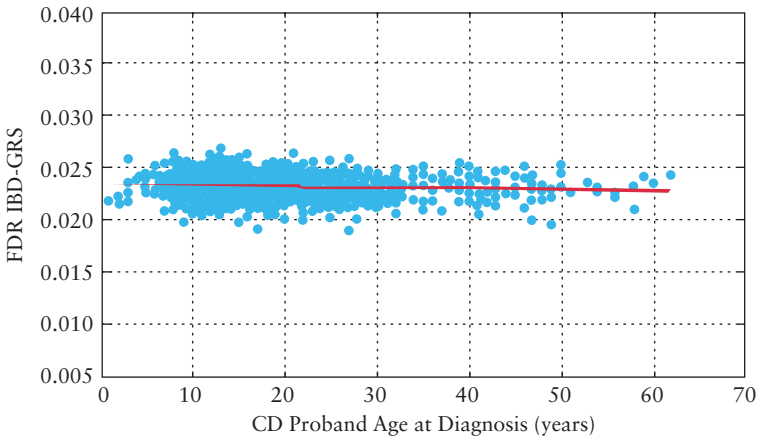

Figure 3. Scatterplot of FDR IBD-GRS and age of diagnosis of their corresponding CD proband. Each point [blue] represents a subject within the FDR cohort. The red line represents a line of best fit [r^2 0.006]. There was a modest but significant inverse correlation between FDR IBD-GRS and the age of diagnosis of their corresponding CD proband; correlation coefficient -0.08 [95% confidence interval: -0.14 - -0.01], $p = 0.02$. IBD, inflammatory bowel disease; HC, healthy controls; FDR, first-degree relatives; CD, Crohn's disease; GRS, genetic risk score.

individuals; differences in RAF between FDR and HC, and CD & HC, cohorts were strongly correlated, generally in the same direction but of a lower magnitude. The finding has two possible explanations; the first is that IBD genetic risk is genuinely elevated to a greater degree in CD-affected individuals compared with FDR. The other possibility is that a significant proportion of FDR with high IBD genetic risk develop overt disease at a younger age, resulting in an under-representation of FDR with this genetic profile, in cohorts such as ours, which recruited both paediatric and adult FDR. Some support for this second hypothesis was provided by the finding that there was a trend toward a modestly higher IBD-GRS in younger [< 17 years] versus older [≥ 17 years] FDR. On balance we believe that both the aforementioned factors contribute to the higher IBD genetic risk observed in CD compared with the FDR.

As IBD is a complex polygenic disease, cumulative, subtle genetic effects may contribute more to overall disease susceptibility than the carriage of individual risk alleles. To reflect this cumulative genetic risk, IBD- and CD-weighted genetic risk scores [IBD-GRS and CD-GRS] were calculated utilising published ORs.[23] There was a highly significant increase in CD-GRS comparing HC and FDR cohorts. In contrast, IBD-GRS was similar comparing these cohorts. This finding demonstrates that FDR of CD-affected individuals have increased cumulative CD-specific genetic susceptibility. Notably, the IBD-GRS and CD-GRS were significantly lower in the FDR compared with the CD cohort, which—as we discuss above—may reflect a true difference between the FDR and CD cohorts or a selection bias in the FDR cohort. Using the CD cohort as a 'reference' disease cohort, we also derived four incremental categories of IBD and CD genetic risk which were then used to assign individuals in the FDR and HC cohorts to one of four genetic risk categories. As expected, a significantly greater proportion of FDRs than HC were assigned to upper genetic risk categories [categories 3 & 4]. This segregation was much more marked for CD-GRS compared with IBD-GRS, reflecting the greater burden of CD-specific cumulative genetic risk compared with more general IBD genetic risk in the FDR cohort. Notably however, a proportion of HC were assigned to higher genetic risk categories with, for example, 14% of healthy controls assigned to CD genetic risk category 4. This finding illustrates that higher IBD genetic risk occurs in the general population, contributing to the incidence of 'sporadic' IBD in individuals of European ancestry without a defined family history of IBD.

Table 2. IBD and CD genetic risk categories in HC and FDR Cohorts.

	IBD genetic risk score categories				*CD genetic risk score categories*			
	Category 1	**Category 2**	**Category 3**	**Category 4**	**Category 1**	**Category 2**	**Category 3**	**Category 4**
HC cohort [n=5000]	2587 [52%]	1274 [25%]	786 [16%]	353 [7%]	1980 [40%]	1317 [26%]	982 [20%]	721 [14%]
FDR cohort [n=976]	475 [49%]	261 [27%]	164 [17%]	76 [8%]	310 [32%]	241 [25%]	217 [22%]	208 [21%]

IBD, inflammatory bowel disease; HC, healthy controls; FDR, first-degree relatives; CD, Crohn's disease; GRS, genetic risk score.

IBD genetic risk score categories were defined using IBD-GRS reference ranges derived from quartiles of IBD-GRS in the CD cohort. References ranges for IBD-GRS in each category were as follows: category 1, ≤ 0.0231555; category 2, 0.0231556–0.0240036; category 3, 0.0240037–0.0248114; category 4, ≥ 0.0248115.

CD genetic risk score categories were defined using CD-GRS reference ranges derived from quartiles of CD-GRS in the CD cohort. References ranges for CD-GRS in each category were as follows: category 1 ≤ 0.0235405; category 2, 0.0235406–0.0257224; category 3, 0.0257225–0.0278698; category 4, ≥ 0.0278699.

An important question for clinicians counselling relatives about the risk of developing IBD is whether their IBD genetic risk is related in any way to their own phenotypic characteristics or those of their IBD-affected family member. It has long been suggested that IBD genetic risk is enriched in individuals diagnosed with IBD at a younger age; however ,it is not known if this is also true of their first-degree relatives. A recent report by Ananthakrishnan *et al*., evaluating genetic risk in a CD cohort using Immunochip genotyping, demonstrating that earlier age of diagnosis was associated with a modest increase in IBD genetic risk burden in CD-affected individuals.[27] Our data extend these findings by demonstrating that IBD genetic risk [IBD-GRS] is also modestly increased in healthy first-degree relatives whose CD-affected family member is diagnosed at a younger age. We did not find a similar correlation when specifically CD-associated genetic risk [CD-GRS] was examined. Finally, we examined whether any information on IBD genetic risk could be deduced from an individual's relationship to their CD-affected family member. We compared IBD and CD genetic risk between siblings and offspring of CD-affected individuals and found there to be no significant difference between these two groups of FDR.

Whereas epidemiological studies have long suggested increased IBD genetic risk in relatives of CD-affected individuals,[3,4,5] we have demonstrated this conclusively. An obvious question posed by these data is whether genotyping provides any additional benefit over family history for assessment of IBD risk. There are a number of arguments which can be made against the utility of genotyping for IBD-associated SNPs in FDR. Family history, it can be argued, provides similar information to genotyping, and is less costly and easier to obtain. In addition, the problem of 'missing heritability' in IBD has resulted in concern regarding the practicality of using genetic risk factors in the context of disease susceptibility risk prediction.[28] We believe, however, that there are potential limitations to solely using family history as a risk prediction tool. Since significant genetic variance can occur within families,[29] this poses substantial limits on the degree to which family history can be informative of disease risk. It has been demonstrated that the predictive power of family history for disease susceptibility diminishes quickly for lower-frequency diseases such as CD, and SNP-based models have been shown not to have the same dependence on disease frequency.[29] Our data demonstrate that the cumulative IBD and CD genetic risk observed in FDR is variable. Therefore even in individuals with a family history of CD, a differential IBD-related genetic burden might be of importance in determining progression to overt disease. Determining whether IBD genetic risk characterisation might facilitate better risk stratification of FDR requires a longitudinal study and therefore cannot be addressed in this report. However, such studies in FDR are required as the discovery of biomarkers, including genetic signatures, of disease susceptibility would be an important clinical advance, as the increasing understanding of environmental influences on IBD susceptibility are likely to result in novel preventive interventions which ameliorate IBD risk.

We acknowledge that this study has a number of limitations. The GEM project recruited FDR between the ages of 6 and 35 years, which limits the extrapolation of these data to the general population of FDR of CD-affected individuals. The analysis was restricted to a solely Caucasian FDR cohort, but this was justified in order to minimise population stratification; and only the first recruited FDR from each family was used to avoid the non-independence of related subjects. The proportion of the FDR cohort of Jewish ancestry was known, but that of the CD and HC groups used for comparison was unknown, which was a limitation also. The GEM project did not co-recruit CD probands along with their healthy FDR and therefore direct comparisons between these groups could not be made. Nonetheless, these data presented represent the largest genotyping study of FDR of CD-affected individuals; our analyses are statistically powerful given the size of the cohorts included in the study; and finally, the use of the Immunochip allowed us to present comprehensive genotyping of our cohorts for known IBD risk loci.

In conclusion, we have demonstrated that healthy FDR of CD-affected individuals are significantly enriched with IBD risk alleles compared with HC. In addition, cumulative CD-specific genetic risk is increased in FDR compared with a healthy cohort. This 'at risk' cohort will be a critical group for ongoing prospective follow-up to study the various microbial and environmental risk factors for CD, and will provide important insights into the interrelationship between genetic susceptibility and such triggers in disease pathogenesis.

Funding

This study is supported by Crohn's & Colitis Canada as the Michael J. Howorth IBD Genetics, Environmental, Microbial Project [GEM Project]; by the Canadian Institutes of Health Research [CIHR], Grant number CMF108031; and by the Leona M. and Harry B. Helmsley Charitable Trust. DK is a recipient of a CIHR / Canadian Association of Gastroenterology [CAG] / Abbvie IBD Fellowship. MSS is supported by the Gale and Graham Wright Chair in Digestive Diseases at Mount Sinai Hospital and by grants from Crohn's & Colitis Canada and the National Institute of Diabetes and Digestive and Kidney Diseases [NIDDK] [DK-062423]. JK holds the Joanne Murphy Professorship in Behavioral Science. KC is supported by grants from Crohn's & Colitis Canada and CIHR.

Conflict of Interest

None.

Acknowledgments

We would like to thank the following individuals and groups for their support of the GEM Project and for their contributions to this work: the NIDDK International IBD Genetics Consortium for the contribution of genotyping data on Crohn's disease-affected individuals and healthy controls; members of the GEM Project Global Office, Steering Committee and Recruitment Site Directors [see supplement, available as Supplementary data at *ECCO-JCC* online]; and the membership and volunteers of Crohn's and Colitis Canada for their help in recruiting subjects.

Supplementary Data

Supplementary data are available at *ECCO-JCC* online.

References

1. Bernstein CN, Wajda A, Svenson LW, *et al*. The epidemiology of inflammatory bowel disease in Canada: a population-based study. *Am J Gastroenterol* 2006;**101**:1559–68.
2. Loftus EV Jr. Clinical epidemiology of inflammatory bowel disease: Incidence, prevalence, and environmental influences. *Gastroenterology* 2004;**126**:1504–17.
3. Orholm M, Munkholm P, Langholz E, Nielsen OH, Sorensen TI, Binder V. Familial occurrence of inflammatory bowel disease. *N Engl J Med* 1991;**324**:84–8.
4. Probert CS, Jayanthi V, Hughes AO, Thompson JR, Wicks AC, Mayberry JF. Prevalence and family risk of ulcerative colitis and Crohn's disease: an epidemiological study among Europeans and South Asians in Leicestershire. *Gut* 1993;**34**:1547–51.
5. Fielding JF. The relative risk of inflammatory bowel disease among parents and siblings of Crohn's disease patients. *J Clin Gastroenterol* 1986;**8**:655–7.
6. Bennett RA, Rubin PH, Present DH. Frequency of inflammatory bowel disease in offspring of couples both presenting with inflammatory bowel disease. *Gastroenterology* 1991;**100**:1638–43.
7. Lowe AM, Roy PO, B-Poulin M, *et al*. Epidemiology of Crohn's disease in Quebec, Canada. *Inflamm Bowel Dis* 2009;**15**:429–35.
8. Ben-Horin S, Avidan B, Yanai H, Lang A, Chowers Y, Bar-Meir S. Familial clustering of Crohn's disease in Israel: prevalence and association with disease severity. *Inflamm Bowel Dis* 2009;**15**:171–5.
9. Halfvarson J. Genetics in twins with Crohn's disease: less pronounced than previously believed? *Inflamm Bowel Dis* 2011;**17**:6–12.
10. Tysk C, Lindberg E, Jarnerot G, Floderus-Myrhed B. Ulcerative colitis and Crohn's disease in an unselected population of monozygotic and dizygotic twins. A study of heritability and the influence of smoking. *Gut* 1988;**29**:990–6.
11. Brant SR. Update on the heritability of inflammatory bowel disease: the importance of twin studies. *Inflamm Bowel Dis* 2011;**17**:1–5.
12. Stokkers PC, Reitsma PH, Tytgat GN, van Deventer SJ. HLA-DR and -DQ phenotypes in inflammatory bowel disease: a meta-analysis. *Gut* 1999;**45**:395–401.
13. Hugot JP, Chamaillard M, Zouali H, *et al*. Association of NOD2 leucine-rich repeat variants with susceptibility to Crohn's disease. *Nature* 2001;**411**:599–603.
14. 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.
15. 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.
16. Parkes M, Barrett JC, Prescott NJ, *et al*. Sequence variants in the autophagy gene IRGM and multiple other replicating loci contribute to Crohn's disease susceptibility. *Nat Genet* 2007;**39**:830–2.
17. Wellcome Trust Case Control C. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature* 2007;**447**:661–78.
18. 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**:955–62.
19. 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.
20. Anderson CA, Boucher G, Lees CW, *et al*. Meta-analysis identifies 29 additional ulcerative colitis risk loci, increasing the number of confirmed associations to 47. *Nat Genet* 2011;**43**:246–52.
21. McGovern DP, Gardet A, Torkvist L, *et al*. Genome-wide association identifies multiple ulcerative colitis susceptibility loci. *Nat Genet* 2010;**42**:332–7.
22. Cortes A, Brown MA. Promise and pitfalls of the Immunochip. *Arthritis Res Ther* 2011;**13**:101.
23. Jostins L, Ripke S, Weersma RK, *et al*. Host-microbe interactions have shaped the genetic architecture of inflammatory bowel disease. *Nature* 2012;**491**:119–24.
24. Liu JZ, van Sommeren S, Huang H, *et al*. Association analyses identify 38 susceptibility loci for inflammatory bowel disease and highlight shared genetic risk across populations. *Nat Genet* 2015;**47**:979–86.
25. 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.
26. Ramon HE, Beal AM, Liu Y, Worthen GS, Oliver PM. The E3 ubiquitin ligase adaptor Ndfip1 regulates Th17 differentiation by limiting the production of proinflammatory cytokines. *J Immunol* 2012;**188**:4023–31.
27. Ananthakrishnan AN, Huang H, Nguyen DD, Sauk J, Yajnik V, Xavier RJ. Differential effect of genetic burden on disease phenotypes in Crohn's disease and ulcerative colitis: analysis of a North American cohort. *Am J Gastroenterol* 2014;**109**:395–400.
28. Manolio TA, Collins FS, Cox NJ, *et al*. Finding the missing heritability of complex diseases. *Nature* 2009;**461**:747–53.
29. Do CB, Hinds DA, Francke U, Eriksson N. Comparison of family history and SNPs for predicting risk of complex disease. *PLoS Genet* 2012;**8**:e1002973.