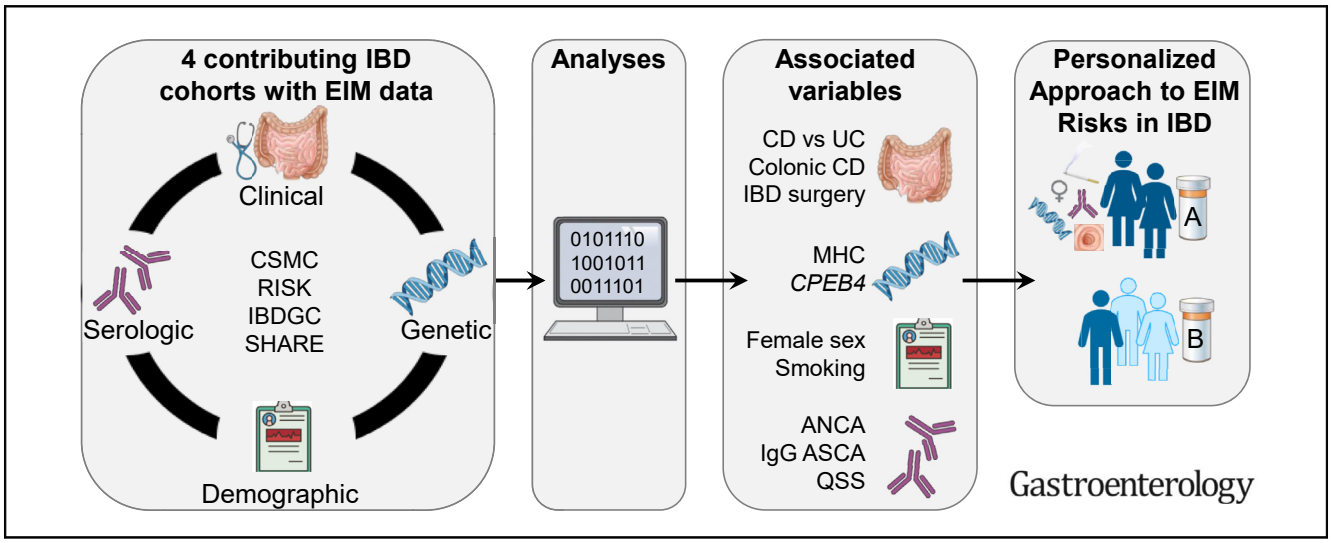
Gastroenterology 2024;167:315–332

Comprehensive Association Analyses of Extraintestinal Manifestations in Inflammatory Bowel Disease

Michelle Khrom,[1] Millie Long,[2,a] Shishir Dube,[1] Lori Robbins,[3] Gregory J. Botwin,[1] Shaohong Yang,[1] Emebet Mengesha,[1] Dalin Li,[1] Takeo Naito,[1] Nirupama N. Bonthala,[1] Christina Ha,[4] Gil Melmed,[4] Shervin Rabizadeh,[5] Gaurav Syal,[4] Eric Vasiliauskas,[4] David Ziring,[4] Steven R. Brant,[6,b] Judy Cho,[7,b] Richard H. Duerr,[8,b] John Rioux,[9,b] Phil Schumm,[10,b] Mark Silverberg,[11,b] Ashwin N. Ananthakrishnan,[12,a] William A. Faubion,[13,a] Bana Jabri,[14,a] Sergio A. Lira,[15,a] Rodney D. Newberry,[16,a] Robert S. Sandler,[2,a] Ramnik J. Xavier,[17,a] Subra Kugathasan,[18,19,c] David Hercules,[19,c] Stephan R. Targan,[1] R. Balfour Sartor,[2,a] Talin Haritunians,[1,b,§] and Dermot P. B. McGovern[1,a,b,§]

[1]F. Widjaja Inflammatory Bowel Disease Institute, Cedars-Sinai Medical Center, Los Angeles, California; [2]Center for Gastrointestinal Biology and Disease, University of North Carolina, Chapel Hill, North Carolina; [3]Palmetto Digestive Health Specialists, Charleston, South Carolina; [4]Inflammatory Bowel Disease Center, Cedars-Sinai Medical Center, Los Angeles, California; [5]Department of Pediatrics, Pediatric Inflammatory Bowel Disease Program, Cedars-Sinai Medical Center, Los Angeles, California; [6]Division of Gastroenterology and Hepatology, Department of Medicine, Rutgers Robert Wood Johnson Medical School, New Brunswick, New Jersey; [7]Icahn School of Medicine at Mount Sinai, Dr Henry D. Janowitz Division of Gastroenterology, New York, New York; [8]Division of Gastroenterology, Hepatology and Nutrition, Department of Medicine, University of Pittsburgh School of Medicine, Pittsburgh, Pennsylvania; [9]Department of Medicine, Université de Montréal and Research Center, Montreal Heart Institute, Montréal, Québec, Canada; [10]Department of Public Health Sciences, University of Chicago, Chicago, Illinois; [11]University of Toronto, Samuel Lunenfeld-Tanenbaum Research Institute, Mount Sinai Hospital, Toronto, Ontario, Canada; [12]Gastroenterology, Massachusetts General Hospital, Boston, Massachusetts; [13]Mayo Clinic, Rochester, Minnesota; [14]University of Chicago, Pritzker School of Medicine, Chicago, Illinois; [15]Immunology Institute, Icahn School of Medicine at Mount Sinai, New York, New York; [16]Department of Internal Medicine, Washington University School of Medicine, St. Louis, Missouri; [17]Broad Institute of MIT and Harvard, Cambridge, Massachusetts; [18]Children's Healthcare of Atlanta Combined Center for Pediatric Inflammatory Bowel Disease, Atlanta, Georgia; and [19]Emory School of Medicine, Atlanta, Georgia



See editorial on page 205.

BACKGROUND & AIMS: Patients with inflammatory bowel disease (IBD) frequently develop extraintestinal manifestations (EIMs) that contribute substantially to morbidity. We assembled the largest multicohort data set to date to investigate the clinical, serologic, and genetic factors associated with EIM complications in IBD. **METHODS:** Data were available in 12,083 unrelated European ancestry IBD cases with presence or absence of EIMs (eg, ankylosing spondylitis [ankylosing spondylitis and sacroiliitis], primary sclerosing cholangitis [PSC], peripheral arthritis, and skin and ocular manifestations) across 4 cohorts (Cedars-Sinai Medical Center, National Institute for Diabetes and Digestive and Kidney Diseases IBD Genetics Consortium, Sinai Helmsley Alliance for Research Excellence Consortium, and Risk Stratification and Identification of Immunogenetic and Microbial Markers of Rapid Disease

Progression in Children with Crohn's Disease cohort). Clinical and serologic parameters were analyzed by means of univariable and multivariable regression analyses using a mixed-effects model. Within-case logistic regression was performed to assess genetic associations. **RESULTS:** Most EIMs occurred more commonly in female subjects (overall EIM: $P = 9.0E-05$, odds ratio [OR], 1.2; 95% CI, 1.1–1.4), with CD (especially colonic disease location; $P = 9.8E-09$, OR, 1.7; 95% CI, 1.4–2.0), and in subjects who required surgery (both CD and UC; $P = 3.6E-19$, OR, 1.7; 95% CI, 1.5–1.9). Smoking increased risk of EIMs except for PSC, where there was a "protective" effect. Multiple serologic associations were observed, including with PSC (anti-nuclear cytoplasmic antibody; IgG and IgA, anti-*Saccharomyces cerevisiae* antibodies; and anti-flagellin) and any EIM (anti-nuclear cytoplasmic antibody; IgG and IgA, anti-*Saccharomyces cerevisiae* antibodies; and anti–*Pseudomonas fluorescens*-associated sequence). We identified genome-wide significant associations within major histocompatibility complex (ankylosing spondylitis and sacroiliitis, $P = 1.4E-15$; OR, 2.5; 95% CI, 2.0–3.1; PSC, $P = 2.7E-10$; OR, 2.8; 95% CI, 2.0–3.8; ocular, $P = 2E-08$, OR, 3.6; 95% CI, 2.3–5.6; and overall EIM, $P = 8.4E-09$; OR, 2.2; 95% CI, 1.7–2.9) and *CPEB4* (skin, $P = 2.7E-08$; OR, 1.5; 95% CI, 1.3–1.8). Genetic associations implicated tumor necrosis factor, JAK-STAT, and IL6 as potential targets for EIMs. Contrary to previous reports, only 2% of our subjects had multiple EIMs and most co-occurrences were negatively correlated. **CONCLUSIONS:** We have identified demographic, clinical, and genetic associations with EIMs that revealed underlying mechanisms and implicated novel and existing drug targets—important steps toward a more personalized approach to IBD management.

Keywords: Inflammatory Bowel Disease; Extraintestinal Manifestations; Genetics; Serology.

WHAT YOU NEED TO KNOW

BACKGROUND AND CONTEXT

Extraintestinal manifestations (EIMs) occur frequently in inflammatory bowel disease (IBD), but their causes are unknown and understanding the underlying mechanisms may help target therapeutic choices for IBD in the future.

NEW FINDINGS

Only 2% of subjects had multiple EIMs and most of the co-occurrences associations were negatively correlated. Anti-nuclear cytoplasmic antibody (ANCA)+/anti-*Saccharomyces cerevisiae* antibody (ASCA)– is associated with primary sclerosing cholangitis in Crohn's disease and *CPEB4* genetic variation is associated with skin manifestations.

LIMITATIONS

This study was restricted to subjects of European ancestry and serology was not available across all cohorts.

CLINICAL RESEARCH RELEVANCE

Our findings implicate mechanisms and potential therapeutic targets and move us a step closer to a more personalized approach to IBD management.

BASIC RESEARCH RELEVANCE

Here, in the largest multicenter study of IBD EIMs, we increased our understanding of the causes of EIMs. We identified key demographic (ie, female sex and smoking), clinical (ie, IBD subtype, surgery, and colonic Crohn's disease), serologic (ie, ASCA and ANCA), and genetic (ie, major histocompatibility complex and *CPEB4*) associations.

Inflammatory bowel disease (IBD), including Crohn's disease (CD) and ulcerative colitis (UC), occurs in genetically susceptible individuals after exposure to environmental or microbial triggers.[1] IBD can be a systemic disorder because many patients develop extraintestinal manifestations (EIMs) that contribute substantially to morbidity.[2] EIMs may be present in up to 47% of patients and can affect joints (ankylosing spondylitis [AS], sacroiliitis, and peripheral arthritis [PA]), skin (erythema nodosum [EN], pyoderma gangrenosum [PG]), eyes (ie, uveitis, iritis, episcleritis, scleritis), and the hepatobiliary tract (primary sclerosing cholangitis [PSC]).[3–5] Previous reports have suggested that patients who have 1 EIM are at increased risk for developing additional EIMs.[3–6] However, the underlying mechanisms of developing EIMs are not fully understood.

Musculoskeletal manifestations are common, affecting 20%–30% of patients with IBD[7] and 10%–20% of patients with IBD have sacroiliac changes, with approximately 7%–12% of patients having a concurrent diagnosis of AS.[8] Skin involvement has been observed in 10%–15% of patients with IBD; EN is the most common,[9] with a higher prevalence observed in female patients and in CD,[10] while PG occurs in <5% of IBD cases.[3] Ocular manifestations occur in 2%–6% of patients with IBD and are seen more commonly in patients with CD with colonic involvement.[11] PSC is the best characterized hepatobiliary manifestation, affecting 1%–8% of

[a] SHARE Consortium; [b] NIDDK IBD Genetics Consortium; [c] RISK Consortium; [§] Authors share co-senior authorship.

Abbreviations used in this paper: **ANCA, anti-nuclear cytoplasmic antibody; anti-CBir1, anti-flagellin; anti-I2, anti–*Pseudomonas fluorescens*-associated sequence; anti-OmpC, anti-outer membrane protein C precursor; AS, ankylosing spondylitis; ASCA, IgG and IgA, anti-*Saccharomyces cerevisiae* antibodies; AS-SI, ankylosing spondylitis and sacroiliitis; CD, Crohn's disease; CSMC, Cedars-Sinai Medical Center; EIM, extraintestinal manifestation; ≥2 EIMs, subjects with 2 or more of 6 extraintestinal manifestation phenotypes; EIM-6, subjects with evidence of any of the following AS-SI, EN, PG, Ps, EYE, PSC; EIM-7, subjects with evidence of any of the following AS-SI, EN, PG, Ps, EYE, PSC, PA; EN, erythema nodosum; EYE, ocular manifestations (uveitis/iritis, episcleritis/scleritis, and ocular inflammation); IBD, inflammatory bowel disease; LD, linkage disequilibrium; MHC, major histocompatibility complex; NIDDK, National Institute for Diabetes and Digestive and Kidney Diseases; OR, odds ratio; PA, peripheral arthritis; PG, pyoderma gangrenous; Ps, psoriasis; PSC, primary sclerosing cholangitis; QC, quality control; QSS, quartile sum score; RISK, Risk Stratification and Identification of Immunogenetic and Microbial Markers of Rapid Disease Progression in Children with Crohn's Disease; SHARE, Sinai Helmsley Alliance for Research Excellence; SKIN-3, skin manifestations (EN, Ps, PG); SNP, single-nucleotide polymorphism; TNF, tumor necrosis factor; UC, ulcerative colitis.**

Most current article

0016-5085/$36.00
https://doi.org/10.1053/j.gastro.2024.02.026

patients with IBD.[12] Some EIMs have been associated with mucosal inflammation, including PA, EN, and episcleritis, and others, such as AS, appear to be independent of disease activity.[13] Prior studies have reported contradictory results for sex, age, smoking status, family history, and IBD location and behavior associations with development of EIMs, at least partly due to relatively small sample sizes and underpowered studies.[5,14]

There have been significant advances in identifying IBD genetic susceptibility loci.[15] Genetic pleiotropy, particularly between immune-mediated diseases, is well established,[16] but understanding of molecular associations with clinical sub-phenotypes has not advanced in a similar manner. Approximately 70% of parent–child and >80% of sibling pairs demonstrated EIM concordance,[17] supporting a hereditary role in developing these manifestations. The greatest challenge in studying these rare IBD phenotypes is statistical power due to the relatively small sample size. To overcome this and improve our understanding of the pathogenesis of EIMs in IBD, we performed an analysis across 4 large IBD cohorts, accumulating the largest sample set to date to investigate the relationship between IBD EIMs and demographic and clinical characteristics, IBD-associated serologies, and genetic variation.

Materials and Methods

Study Population

Our study encompassed the following 4 cohorts: Cedars-Sinai Medical Center (CSMC) IBD Research Repository (Material and Information Resources for Inflammatory and Digestive Diseases [MIRIAD] Biobank; 1348 IBD cases with any EIM, 5303 IBD cases with no EIM); Sinai Helmsley Alliance for Research Excellence (SHARE; 1798 IBD cases with any EIM, 2625 IBD cases with no EIM); National Institute for Diabetes and Digestive and Kidney Diseases IBD Genetics Consortium (NIDDK IBDGC; 580 IBD cases with any EIM; 2059 IBD cases with no EIM); and Risk Stratification and Identification of Immunogenetic and Microbial Markers of Rapid Disease Progression in Children with Crohn's Disease (RISK; 87 CD cases with any EIM; 558 CD cases with no EIM). Study cohort details are provided in Supplementary Table 1*A* and Supplementary Methods/Study Population subsection.

EIM phenotypes evaluated included ankylosing spondylitis and sacroiliitis (AS-SI), PG, EN, psoriasis (Ps), ocular manifestations (ie, uveitis/iritis, episcleritis/scleritis, and ocular inflammation [EYE]), PSC, and PA, as defined by large or small joint arthritis, nonspecific joint inflammation, and arthralgias (but excluding AS-SI).

All adult patients provided informed consent approved by the Institutional Review Board at each recruiting center. Pediatric consent was obtained from a parent or legal guardian in conjunction with youth assent, when applicable. Inter- and intra-cohort duplicates were identified and resolved to avoid duplicate reporting of study subjects.

Clinical Phenotyping

Clinical data were collected from all contributing cohorts (Table 1 and Supplementary Tables 4–12) and included patient demographic characteristics (ie, sex, age at diagnosis of IBD, and self-reported race and ethnicity), family history, IBD-related surgical history, current smoking status at time of diagnosis, disease location, behavior per Montreal Classification,[18] and presence or absence of EIMs. As RISK is a pediatric inception cohort, baseline data did not include disease behavior; however, these data were collected ≥90 days after diagnosis during a 36-month follow-up window.

Subjects with IBD with documented evidence of a specific EIM are considered positive cases (EIM[+]) for a given EIM, and subjects with IBD without any evidence of a history of that specific EIM were considered non-EIM controls (EIM[–]). We defined an overall "EIM-6" phenotype for subjects with evidence of any of the following EIMs: AS-SI, EN, PG, Ps, EYE, or PSC. In SHARE, PA was the only self-reported phenotype and thus was not included in the EIM-6 definition for the combined cohorts. When we included PA, the phenotype was defined as "EIM-7." Subjects with 2 or more of the 6 EIM phenotypes (AS-SI, EN, PG, Ps, EYE, and PSC) were defined as "≥2 EIMs." We also defined a "SKIN-3" phenotype for subjects who displayed at least 1 of the following skin manifestations: PG, EN, and Ps. Total number of subjects available post quality control (QC; detailed below) per EIM phenotype is shown in Table 2. A breakdown of subjects by cohort can be found in Supplementary Table 1*B*.

Clinical Definitions of Extraintestinal Manifestations

In SHARE, PA was the only self-reported phenotype lacking confirmation by clinician medical record review and all other EIMs were diagnosed according to established clinical (including radiologic) parameters by experienced gastroenterologists with a deep knowledge of IBD and its clinical manifestations, often in conjunction with multidisciplinary input from dermatologists, rheumatologists, and ophthalmologists, as is normal in large clinical IBD centers.

Serologic Phenotyping

IBD-associated serologies, including IgG and IgA, anti-*Saccharomyces cerevisiae* antibodies (ASCA); anti-nuclear cytoplasmic antibody (ANCA); anti-flagellin (anti-CBir1); anti-outer membrane protein C precursor (anti-OmpC); and anti-*Pseudomonas fluorescens*-associated sequence (anti-I2) were represented as quantitative as well as binary (positive or negative) variables. Quartile sum score (QSS) was also calculated. Only patients from CSMC MIRIAD (any IBD) and RISK (CD only) had serology data available for evaluation (Supplementary Table 2). See Supplementary Methods/Serologic Phenotyping subsection for details.

Genotyping Quality Control

Subjects across all cohorts were genotyped using Immunochip array (Illumina, San Diego, CA). Data across the 4 cohorts were combined and stringent post-genotyping sample and single-nucleotide polymorphism (SNP) QC metrics were applied. Subjects of European ancestry, as defined by admixture or principal components, were retained. See Supplementary Methods/Genotyping Quality Control subsection for details. Post QC, 12,083 subjects with IBD had available genetic and clinical data for downstream analyses.

Table 1. Demographic and Clinical Characteristics of Subjects With Inflammatory Bowel Disease for 10,190 Extraintestinal Manifestation-6 (–) and 1704 Extraintestinal Manifestation-6 (+)

Variable	EIM-6 (–) (n = 10,190)	EIM-6 (+) (n = 1704)
CD, n (%)	6259 (84.7)	1129 (15.3)
UC, n (%)	3619 (87.3)	528 (12.7)
IBD unclassified, n (%)	312 (86.9)	47 (13.1)
Sex, n (%)		
Female	4949 (48.6)	919 (53.9)
Male	5241 (51.4)	785 (46.1)
Age at IBD diagnosis, *y*, mean (SD)	26.87 (14.34)	26.97 (13.47)
Age group at IBD diagnosis, n (%)		
A1 (<17 y)	2414 (23.7)	393 (23.1)
A2 (17–40 y)	5169 (50.7)	992 (58.2)
A3 (>40 y)	1561 (15.3)	266 (15.6)
Missing	1046 (10.3)	53 (3.1)
Family history of IBD, n (%)		
Yes	2250 (22.1)	453 (26.6)
No	6722 (66.0)	1207 (70.8)
Missing	1218 (12.0)	44 (2.6)
Smoking at diagnosis, n (%)		
Yes	1532 (15.0)	312 (18.3)
No	7777 (76.3)	1363 (80.0)
Missing	881 (8.6)	29 (1.7)
Race and ethnicity (self-declared), n (%)		
White	9853 (96.7)	1660 (97.4)
Black	5 (0.0)	0 (0.0)
Asian	35 (0.3)	6 (0.4)
Other	195 (1.9)	36 (2.1)
Missing	102 (1.0)	2 (0.1)
Jewish (self-reported), n (%)		
Yes	2546 (25)	496 (29.2)
No	7511 (73.7)	1192 (70.0)
Missing	133 (1.3)	16 (0.9)
Hispanic (self-reported), n (%)		
Yes	232 (2.3)	36 (2.1)
No	9849 (96.7)	1657 (97.2)
Missing	109 (1.1)	11 (0.6)
Disease location (CD), n (%)		
Ileal (L1)	1311 (20.9)	165 (14.6)
Colorectal (L2)	1022 (16.3)	212 (18.8)
Ileocolonic (L3)	3308 (52.9)	709 (62.8)
Missing	618 (9.9)	43 (3.8)
Upper gastrointestinal (L4)	687 (11.0)	51 (4.5)
Isolated small bowel disease (L1)	1311	165
Any colonic disease (L2+L3)	4330	921
Disease behavior (CD), n (%)		
Inflammatory (B1)	2907 (46.4)	521 (46.1)
Stricturing (B2)	1387 (22.2)	262 (23.2)
Penetrating (B3)	1442 (23.0)	323 (28.6)
Missing	523 (8.4)	23 (2.0)
Perianal (CD), n (%)		
Yes	1477 (23.6)	355 (31.4)
No	4126 (65.9)	737 (65.3)
Missing	656 (10.5)	37 (3.3)
Disease extent (UC/IBD unclassified), n (%)		
Proctitis (E1)	273 (6.9)	19 (3.3)
Left-sided (distal) (E2)	998 (25.4)	92 (16.0)
Extensive (pancolitis) (E3)	1930 (49.1)	403 (70.1)
Missing	730 (18.6)	61 (10.6)
IBD-related surgery, n (%)		
Yes	3047 (29.9)	827 (48.5)
No	4097 (40.2)	582 (34.2)
Missing	3046 (29.9)	295 (17.3)

NOTE. Data presented for EIM-6 subjects with at least 1 EIM, excluding PA. Percentages are shown in parentheses after numbers per variable. See Supplementary Tables 4–12 for demographic and clinical characteristics of patients with other EIMs.

Table 2. Total Number of Subjects of European Ancestry With Extraintestinal Manifestation (+)/Extraintestinal Manifestation (–) Post Quality Control

	Merged subjects with IBD (post QC)	
EIM	EIM (+), n (%)	EIM (–), n
AS-SI	403 (3.3)	11,624
SKIN-3[a]	783 (6.5)	11,149
EN	320 (2.6)	11,763
PG	144 (1.2)	11,939
Ps	367 (3.0)	8999
EYE	286 (2.4)	11,738
PSC	459 (3.8)	11,573
PA	2040 (17)	9874
EIM-6[b]	1704 (14)	10,190
EIM-7[c]	3255 (27)	8572
≥2 EIMs[d]	240 (2.0)	8572

NOTE. EIM (+) percentage was based on total of 12,083 subjects with IBD in our combined cohort. EIM (–) was based on subjects with IBD without any evidence of history of that specific EIM.
[a]SKIN-3 subjects with at least 1 skin manifestation (EN, PG, or Ps).
[b]EIM-6 subjects with at least 1 EIM, excluding PA.
[c]EIM-7 subjects with at least 1 EIM, including PA.
[d]≥2 EIMs subjects with at least 2 EIM-6, excluding PA.

Statistical Analyses

Univariable regression analyses were performed in our combined data set using a mixed-effects model to evaluate the association of each clinical parameter, comparing subjects with IBD with and without the EIM phenotypes (ie, AS-SI, PG, EN, Ps, EYE, PSC, PA, SKIN-3, EIM-6, and EIM-7). The mixed-effect model allowed us to include a cohort to control for potential heterogeneity. Similarly, mixed-effects logistic regression analyses were performed to evaluate the association of IBD serologic positivity and levels with EIM status; for the serologies, CD and UC were analyzed separately, given the different serologic profiles observed in IBD subtypes. For clinical variables with $P < .05$ in the univariable analysis, we performed mixed-effect multivariable logistic regression comparing all subjects with IBD with and without each specific EIM, as well as separately for subjects with CD and UC/IBD unclassified. Serologic factors were included in a separate multivariable model, as data were limited to CSMC and RISK. χ^2 test was used to evaluate co-occurrence of EIMs. Multiple testing thresholds for the clinical ($n = 21$) and serologic ($n = 14$) tests across all 11 EIMs were defined at $P < 2.2 \times 10^{-4}$ and 3.2×10^{-4}, respectively.

Within-case logistic regression was performed to investigate associations between various EIM phenotypes (as noted above) and autosomal SNP in our subjects with IBD (IBD-EIM [+] cases compared with IBD-EIM[–] controls). The first 5 principal components and cohort were included as covariates to control for potential confounding due to population substratification and genotyping batch effects. Population substratification for the combined data set (ie, CSMC, SHARE, NIDDK, and RISK) was evaluated by means of calculating the genomic inflation factor for IBD EIM(+) cases and EIM(−) controls. Test statistics showed negligible inflation (Supplementary Table 3). Only PG displayed genomic inflation >1.1 (Supplementary Table 3), likely reflecting the small number of PG cases (Table 2) in comparison with other EIM subtypes. For genetic association analyses, genome-wide significance was defined at $P < 5 \times 10^{-8}$ and nominal level of significance at $P < 1 \times 10^{-4}$. Associated SNPs with allele frequency differences >10% between cohorts and in comparison with gnomAD (non-Finnish European population), and variants with poor genotype clustering observed on manual review, where available, were excluded from further study.

All analyses were performed using PLINK and R software unless otherwise noted.[19,20] Missing data were coded accordingly and omitted from analyses, when applicable. Linkage disequilibrium (LD) was calculated for CEU population using LDlink.[21] Variant annotation was performed as described in the Supplementary Methods/Statistical Analyses subsection.

Gene Enrichment Analysis

Genomic regions corresponding to variants exhibiting association ($P < .001$) in each of the EIM phenotypes were used as input for annotation with nearby genes (Genomic Regions Enrichment of Annotations Tool[22]). Gene lists were then evaluated for associated pathways and ontologies using Enrichr[23] (Supplementary Methods/Gene Enrichment Analysis subsection).

Results

Among a total of 12,083 subjects with IBD, we identified 14% with EIM-6, 27% with EIM-7, and only 2% with ≥2 EIMs. For individual EIMs, we observed 17% with PA, 6.5% with SKIN-3, 3.8% with PSC, 3.3% with AS-SI, 3.0% with Ps, 2.6% with EN, 2.4% with EYE, and 1.2% with PG (Table 2). EIM breakdown per cohort is shown in Supplementary Table 1. Clinical and demographic variables for each EIM phenotype are shown in Table 1 and Supplementary Tables 4–12. A summary of our multivariable clinical associations comparing subjects with IBD, CD, and UC with and without EIM in our combined data set is provided in Table 3. With the exception of PSC (more common in UC), all EIMs were more commonly observed in CD than UC. We also observed increased EIM risk in female patients, with the exception of AS-SI and PSC. Jewish ancestry was associated with Ps and overall SKIN-3. Increased risk of multiple EIMs (ie, Ps, SKIN-3, PA, and EIM-7) was associated with smoking in IBD, but smoking was observed to be protective for PSC. A positive family history of IBD was associated with increased risk of PG, PA, and EIM-7. Older age at IBD diagnosis demonstrated increased risk for AS-SI, PA, and EIM-7, yet was protective for EN. We observed an increased risk for numerous EIMs (ie, EN, PG, SKIN-3, PSC, PA, EIM-6, EIM-7, and ≥2 EIMs) in subjects with CD with any colonic disease, and any small bowel disease was protective for PG. In UC, proximal disease extent was associated with increased risk for PSC, EIM-6, and EIM-7. An increased risk with any IBD-related surgery was associated with all EIM phenotypes except for Ps and EYE.

Serologic data were limited to CSMC and RISK cohorts (subject breakdown in Supplementary Table 2). A summary of serologic associations with EIMs is shown in Table 4. In CD, ANCA positivity was associated with an increased risk for PG and PSC, and increased ANCA levels were associated with PG, PSC, EIM-6, EIM-7, and ≥2 EIMs. Anti-CBir1 positivity was associated with AS-SI and PSC in CD and UC, respectively (higher prevalence), as well as PG and PSC in CD (lower prevalence). In UC, there was higher risk of PSC, EIM-6, and EIM-7 with increased anti-CBir1 levels, and high anti-CBir1 levels were associated with decreased PSC in CD. Anti-I2 positivity was associated with increased risk for Ps, SKIN-3, PA, EIM-6, and EIM-7 in CD, and ≥2 EIMs in UC. Increased anti-I2 levels were associated with higher prevalence of PA, EIM-6, and EIM-7 in CD, and an increased risk of PSC in UC. Anti-OmpC positivity was associated with Ps, SKIN-3, PA, EIM-6, EIM-7 in CD and EN in UC, and anti-OmpC levels were associated with EN, SKIN-3, EIM-6, and EIM-7 in CD only. Overall ASCA positivity in subjects with CD was decreased for PSC, EIM-6, and EIM-7. Increased QSS was associated with EIM-7 in CD and with EYE, PSC, and EIM-6/7 in UC, and increased QSS was observed to be protective for PG and PSC in CD.

Genetic associations with the specific EIMs are detailed below. All EIMs, with the noted exception of PG, demonstrated an association with variants in the major histocompatibility complex (MHC) (Table 5 and Supplementary Tables 13–23). Of the 6 genome-wide associations observed across the EIMs, 5 were within MHC loci, in addition to an association tagging known IBD locus *CPEB4* and SKIN-3 (Table 5).

We observed a total of 240 subjects with evidence of 2 or more EIMs (not including PA) (Supplementary Results/

Table 3. Summary of Multivariable Clinical Associations Comparing Inflammatory Bowel Disease With Extraintestinal Manifestation to Inflammatory Bowel Disease Without Extraintestinal Manifestation

	AS-SI		EN		PG		Ps		SKIN-3[a]	
Variable	*P* value	OR (95% CI)	*P* value	OR (95% CI)	*P* value	OR (95% CI)	*P* value	OR (95% CI)	*P* value	OR (95% CI)
CD (yes)	**≤.01**	**1.57 (1.19–2.07)**	**≤.0001**	**2.78 (1.96–3.93)**	NS	1.52 (0.99–2.35)	**≤.0001**	**2.09 (1.62–2.71)**	**≤.0001**	**2.33 (1.88–2.87)**
Sex (female)										
IBD	**≤.01**	**0.70 (0.55–0.89)**	**≤.0001**	**3.57 (2.64–4.82)**	**≤.01**	**1.81 (1.24–2.65)**	**≤.01**	**1.34 (1.08–1.66)**	**≤.0001**	**2.06 (1.72–2.46)**
CD	**<.05**	**0.69 (0.49–0.96)**	**≤.0001**	**4.34 (2.71–6.93)**	NS	1.56 (1.00–2.43)	**≤.01**	**1.53 (1.14–2.05)**	**≤.0001**	**2.52 (1.75–3.61)**
UC	NS	0.67 (0.42–1.07)	**≤.01**	**2.46 (1.27–4.78)**	**<.05**	**2.61 (1.18–5.76)**	NS	1.07 (0.68–1.68)	**≤.01**	**1.83 (1.26–2.67)**
Jewish (yes)										
IBD							**≤.01**	**1.48 (1.17–1.87)**	**<.05**	**1.30 (1.06–1.59)**
CD							NS	1.36 (0.99–1.88)	NS	0.99 (0.65–1.51)
UC							**<.05**	**1.82 (1.15–2.87)**	NS	1.55 (1.01–2.39)
Age at IBD diagnosis										
IBD	**≤.01**	**1.01 (1.00–1.02)**	**≤.01**	**0.98 (0.97–0.99)**						
CD	**<.05**	**1.01 (1.00–1.03)**	**<.05**	**0.98 (0.96–1.0)**						
UC	NS	1.01 (1.00–1.02)	NS	0.98 (0.96–1.0)						
Smoking at diagnosis[b] (yes)										
IBD							**≤.01**	**1.46 (1.14–1.88)**	**<.05**	**1.27 (1.03–1.56)**
CD							NS	1.32 (0.94–1.86)	NS	1.08 (0.73–1.62)
UC							**≤.01**	**2.10 (1.24–3.54)**	**≤.01**	**1.92 (1.18–3.12)**
Family history (yes)										
IBD					**<.05**	**1.54 (1.05–2.24)**				
CD					NS	1.32 (0.84–2.08)				
UC					NS	2.08 (0.99–4.38)				
CD: any colonic (L2L3 vs L1)			**<.05**	**1.80 (1.04–3.13)**	**≤.01**	**6.94 (2.16–22.30)**			**≤.01**	**2.59 (1.43–4.71)**
CD: any small bowel (L1L3 vs L2)					**≤.01**	**0.46 (0.29–0.74)**	NS	1.41 (0.94–2.09)		
UC: proximal disease vs proctitis (E2E3 vs E1)										
IBD-related surgery (all subjects)	**≤.01**	**1.43 (1.11–1.83)**	**<.05**	**1.32 (1.02–1.72)**	**≤.0001**	**2.62 (1.74–3.93)**			**≤.001**	**1.34 (1.13–1.60)**
IBD-related surgery (subjects with CD only)	NS	1.10 (0.74–1.64)	NS	1.54 (0.89–2.66)	**≤.0001**	**3.30 (1.96–5.56)**			NS	1.17 (0.74–1.85)
IBD-related surgery (subjects with UC only)	**≤.001**	**2.64 (1.61–4.35)**	NS	1.42 (0.75–2.70)	**<.05**	**2.13 (1.02–4.44)**			NS	1.19 (0.78–1.82)

Table 3. Continued

Variable	EYE		PSC		PA		EIM–6[c]		EIM–7[d]	
	P value	OR (95% CI)	*P* value	OR (95% CI)	*P* value	OR (95% CI)	*P* value	OR (95% CI)	*P* value	OR (95% CI)
CD (yes)	**≤.0001**	**1.85 (1.38–2.50)**	**≤.0001**	**0.18 (0.15–0.23)**	**≤.0001**	**1.33 (1.18–1.50)**	NS	0.90 (0.79–1.02)	NS	1.05 (0.94–1.16)
Sex (female)										
IBD	**≤.0001**	**1.76 (1.35–2.28)**	**<.05**	**0.76 (0.62–0.94)**	**≤.0001**	**1.59 (1.42–1.77)**	**≤.0001**	**1.27 (1.13–1.43)**	**≤.0001**	**1.42 (1.29–1.57)**
CD	**≤.001**	**1.79 (1.31–2.44)**	**<.05**	**0.54 (0.33–0.89)**	**≤.0001**	**1.69 (1.47–1.94)**	**≤.0001**	**1.53 (1.31–1.79)**	**≤.0001**	**1.52 (1.32–1.76)**
UC	NS	1.66 (1.01–2.73)	NS	0.77 (0.59–1.00)	**≤.0001**	**1.52 (1.25–1.84)**	NS	0.94 (0.76–1.16)	**<.05**	**1.20 (1.02–1.42)**
Jewish (yes)										
Age at IBD diagnosis										
IBD					**≤.0001**	**1.02 (1.01–1.02)**			**≤.0001**	**1.01 (1.01–1.01)**
CD					**≤.0001**	**1.02 (1.01–1.02)**			**≤.001**	**1.01 (1.00–1.02)**
UC					**≤.0001**	**1.02 (1.02–1.03)**			**≤.001**	**1.01 (1.00–1.02)**
Smoking at diagnosis[b] (yes)										
IBD			**≤.01**	**0.63 (0.46–0.87)**	**≤.0001**	**1.51 (1.31–1.74)**			**≤.0001**	**1.36 (1.20–1.55)**
CD			**<.05**	**0.37 (0.16–0.86)**	**≤.0001**	**1.61 (1.37–1.90)**			**≤.001**	**1.38 (1.16–1.65)**
UC			NS	0.94 (0.63–1.39)	NS	1.23 (0.90–1.69)			NS	1.07 (0.82–1.41)
Family history (yes)										
IBD					**≤.01**	**1.21 (1.07–1.37)**			**≤.001**	**1.21 (1.09–1.35)**
CD					**<.05**	**1.19 (1.02–1.40)**			**<.05**	**1.21 (1.03–1.43)**
UC					**<.05**	**1.30 (1.04–1.61)**			**≤.01**	**1.33 (1.10–1.60)**
CD: any colonic (L2L3 vs L1)			**≤.01**	**7.89 (1.90–32.71)**	**≤.0001**	**1.50 (1.26–1.77)**	**≤.0001**	**1.91 (1.54–2.38)**	**≤.0001**	**1.53 (1.27–1.85)**
CD: any small bowel (L1L3 vs L2)			NS	0.88 (0.53–1.45)						
UC: proximal disease vs proctitis (E2E3 vs E1)			**≤.01**	**3.76 (1.37–10.27)**			**<.05**	**2.07 (1.18–3.63)**	**<.05**	**1.45 (1.05–2.02)**
IBD-related surgery (all subjects)			**≤.0001**	**3.08 (2.46–3.87)**	**≤.0001**	**1.39 (1.24–1.56)**	**≤.0001**	**1.73 (1.53–1.96)**	**≤.0001**	**1.65 (1.49–1.82)**
IBD-related surgery (CD subjects only)			**≤.0001**	**3.80 (2.35–6.16)**	**≤.0001**	**1.32 (1.15–1.52)**	**≤.01**	**1.29 (1.10–1.51)**	**≤.0001**	**1.46 (1.23–1.73)**
IBD-related surgery (UC subjects only)			**≤.0001**	**3.51 (2.64–4.67)**	**≤.0001**	**1.58 (1.27–1.97)**	**≤.0001**	**2.51 (2.01–3.15)**	**≤.0001**	**2.11 (1.75–2.55)**

NOTE. Significant associations ($P < .05$) are in bold. Associations are shown for IBD; CD or UC-specific associations are noted where applicable. Empty cells denote variables $P > .05$ in univariable analysis that did not move forward to multivariable analysis for a given EIM. See Supplementary Tables 24, 28, 30, 32, 34, 36, 38, 40, 42, 44, and 46 for detailed results for respective EIMs. NS, not significant.

[a]SKIN-3 subjects with at least 1 skin manifestation (EN, PG, or Ps).

[b]Current smoking status at time of diagnosis.

[c]EIM-6 subjects with at least 1 EIM, excluding PA.

[d]EIM-7 subjects with at least 1 EIM including PA. Table 3 highlights an increased risk for most EIMs in CD (except for PSC and UC) and in female subjects (except for AS-SI and PSC and male subjects). Smoking was associated with increased risk for several EIMs in IBD (Ps, SKIN-3, PA, and EIM-7) but protective for PSC. In CD, any colonic disease was associated with an increased risk for multiple EIMs (except AS-SI, Ps, and EYE). In UC, proximal disease extent was associated with increased risk for PSC and EIM-6/7. Increased risk with any IBD-related surgery was observed with all EIM phenotypes (except for Ps and EYE). Associations with colonic CD and surgery showed the strongest ORs in PG and PSC.

Table 4. Summary of Univariable Serologic Associations With Extraintestinal Manifestation Status in Crohn's Disease or Ulcerative Colitis

	Univariable serologic positivity									
	AS-SI		EN		PG		Ps		SKIN-3[a]	
Serology analysis	*P* value	OR (95% CI)	*P* value	OR (95% CI)	*P* value	OR (95% CI)	*P* value	OR (95% CI)	*P* value	OR (95% CI)
ANCA										
CD					**<.05**	**2.245 (1.206–4.179)**				
UC					NS	0.912 (0.244–3.410)				
Anti–CBir1										
CD	**≤.01**	**1.559 (1.166–2.084)**			**<.05**	**0.483 (0.253–0.924)**				
UC	NS	1.010 (0.449–2.274)			NS	1.906 (0.393–9.251)				
Anti–I2										
CD							**<.05**	**1.509 (1.016–2.241)**	**<.05**	**1.369 (1.014–1.850)**
UC							NS	1.057 (0.418–2.673)	NS	1.580 (0.806–3.094)
Anti–OmpC										
CD			NS	1.488 (0.970–2.284)			**<.05**	**1.536 (1.064–2.216)**	**<.05**	**1.427 (1.074–1.895)**
UC			**<.05**	**3.238 (1.031–10.174)**			NS	1.066 (0.374–3.038)	NS	1.636 (0.790–3.386)
ASCA–IgA										
ASCA–IgG										
Overall ASCA										
QSS[b]										
CD					**<.05**	**0.882 (0.797–0.977)**				
UC					NS	1.132 (0.860–1.490)				

	Univariable serologic positivity									
	EYE		PSC		PA		EIM–6[c]		EIM–7[d]	
Serology analysis	*P* value	OR (95% CI)	*P* value	OR (95% CI)	*P* value	OR (95% CI)	*P* value	OR (95% CI)	*P* value	OR (95% CI)
ANCA										
CD			**≤.0001**	**3.334 (2.174–5.111)**						
UC			NS	1.275 (0.886–1.833)						
Anti–CBir1										
CD			**≤.01**	**0.466 (0.293–0.739)**						
UC			**<.05**	**1.559 (1.013–2.401)**						
Anti–I2										
CD					**<.05**	**1.393 (1.056–1.838)**	**≤.01**	**1.359 (1.091–1.694)**	**≤.001**	**1.476 (1.213–1.796)**
UC					NS	1.473 (0.806–2.691)	NS	1.111 (0.769–1.605)	NS	1.172 (0.826–1.661)

Table 4. Continued

	Univariable serologic positivity									
	EYE		PSC		PA		EIM–6[c]		EIM–7[d]	
Serology analysis	*P* value	OR (95% CI)	*P* value	OR (95% CI)	*P* value	OR (95% CI)	*P* value	OR (95% CI)	*P* value	OR (95% CI)
Anti–OmpC										
CD					**<.05**	**1.385 (1.069–1.796)**	**<.05**	**1.264 (1.017–1.571)**	**≤.01**	**1.322 (1.089–1.604)**
UC					NS	1.309 (0.661–2.594)	NS	1.462 (0.953–2.243)	NS	1.338 (0.886–2.021)
ASCA–IgA										
CD			**≤.01**	**0.443 (0.253–0.777)**						
UC			NS	1.801 (0.827–3.925)						
ASCA–IgG										
CD			**≤.0001**	**0.214 (0.107–0.428)**			**≤.01**	**0.733 (0.594–0.906)**	**≤.01**	**0.759 (0.631–0.911)**
UC			NS	1.100 (0.427–2.832)			NS	1.196 (0.569–2.512)	NS	1.112 (0.545–2.270)
Overall ASCA										
CD			**≤.0001**	**0.284 (0.162–0.498)**			**≤.01**	**0.750 (0.614–0.915)**	**<.05**	**0.818 (0.688–0.972)**
UC			NS	1.672 (0.880–3.174)			NS	1.598 (0.935–2.732)	NS	1.484 (0.884–2.490)
QSS[b]										
CD	NS	0.978 (0.909–1.052)	**≤.01**	**0.906 (0.849–0.968)**			NS	1.015 (0.985–1.046)	**<.05**	**1.033 (1.005–1.061)**
UC	**<.05**	**1.269 (1.021–1.577)**	**≤.01**	**1.105 (1.038–1.175)**			**≤.0001**	**1.111 (1.053–1.171)**	**≤.001**	**1.104 (1.050–1.161)**

	Univariable serologic levels									
	AS–SI		EN		PG		Ps		SKIN–3[a]	
Serology analysis	*P* value	OR (95% CI)	*P* value	OR (95% CI)	*P* value	OR (95% CI)	*P* value	OR (95% CI)	*P* value	OR (95% CI)
ANCA										
CD					**<.05**	**1.008 (1.000–1.016)**				
UC					NS	0.995 (0.979–1.012)				
Anti–CBir1										
Anti–I2										
Anti–OmpC										
CD			**<.05**	**1.008 (1.002–1.015)**					**<.05**	**1.006 (1.001–1.011)**
UC			NS	1.013 (0.996–1.031)					NS	1.006 (0.992–1.020)
ASCA–IgA										
ASCA–IgG										
CD					**<.05**	**0.988 (0.977–0.999)**				
UC					NS	0.976 (0.902–1.057)				

Table 4. Continued

	Univariable serologic levels									
	EYE		PSC		PA		EIM–6[c]		EIM–7[d]	
Serology analysis	*P* value	OR (95% CI)	*P* value	OR (95% CI)	*P* value	OR (95% CI)	*P* value	OR (95% CI)	*P* value	OR (95% CI)
ANCA										
CD			**≤.0001**	**1.017 (1.012–1.021)**			**≤.01**	**1.004 (1.001–1.007)**	**<.05**	**1.003 (1.000–1.006)**
UC			NS	1.001 (0.997–1.005)			NS	1.001 (0.998–0.998)	NS	1.001 (0.998–1.004)
Anti–CBir1										
CD			**≤.01**	**0.988 (0.979–0.996)**			NS	1.000 (0.997–1.002)	NS	1.000 (0.998–1.002)
UC			**≤.01**	**1.008 (1.003–1.014)**			**≤.01**	**1.008 (1.003–1.013)**	**<.05**	**1.006 (1.001–1.011)**
Anti–I2										
CD			NS	0.999 (0.993–1.004)	**<.05**	**1.003 (1.000–1.006)**	**<.05**	**1.003 (1.001–1.005)**	**≤.001**	**1.004 (1.002–1.006)**
UC			**<.05**	**1.007 (1.001–1.013)**	NS	1.005 (0.997–1.014)	NS	1.005 (1.000–1.011)	NS	1.005 (0.999–1.010)
Anti–OmpC										
CD							**<.05**	**1.005 (1.001–1.009)**	**≤.01**	**1.006 (1.003–1.010)**
UC							NS	1.005 (0.996–1.014)	NS	1.003 (0.994–1.012)
ASCA–IgA										
CD			**≤.01**	**0.981 (0.970–0.993)**						
UC			NS	1.010 (0.993–1.027)						
ASCA–IgG										
CD			**≤.0001**	**0.978 (0.968–0.988)**			**≤.01**	**0.996 (0.993–0.998)**	**≤.01**	**0.996 (0.994–0.999)**
UC			NS	1.003 (0.991–1.016)			UC NS	1.004 (0.994–1.014)	NS	1.001 (0.992–1.011)

NOTE. Univariable seropositivity and serology level associations are shown. Significant associations ($P < .05$) are in bold. ORs (95% CIs) for respective CD or UC analysis. Empty fields denote variables $P > .05$ in univariable analysis for both CD and UC. See Supplementary Tables 25, 26, 29, 31, 33, 35, 37, 39, 41, 43, 45, and 47 for detailed results for respective EIMs. NS, not significant.

[a]SKIN-3 subjects with at least 1 skin manifestation (EN, PG, or Ps).

[b]QSS only available for CSMC cohort.

[c]EIM-6 subjects with at least 1 EIM, excluding PA.

[d]EIM-7 subjects with at least 1 EIM including PA. Table 4 highlights serologic differences between subjects with IBD with and without EIMs as shown. We observed the strongest antibody associations with EIM-6 and PSC, predominantly in CD. The most significant univariable associations were with PSC in CD and ANCA seropositivity and increased ANCA levels. A strong PSC association with ASCA seronegativity, specifically with ASCA IgG, was also observed in CD.

Table 5. Genome-Wide Extraintestinal Manifestation Associations ($P < 5 \times 10^{-8}$) for Inflammatory Bowel Disease Autosomes

Phenotype	Chromosome	SNP[a]	Position-hg19, *Mb*	Candidate gene	Additional SNPs in LD[b]	Effect allele	*P* value	OR	95% CI	
									Lower	Upper
AS-SI	6	rs6905036	31.27	*HLA-C, HLA-B*	3	G	1.36E-15	2.499	1.996	3.127
AS-SI	6	rs2844510	31.41	*LINC1149, MICA, HCP5*	—	A	9.89E-13	2.107	1.717	2.586
PSC	6	rs9276456	32.72	*HLA-DQA2, HLA-DQB2, MIR3135B*	36	G	2.71E-10	2.757	2.012	3.773
EIM-6	6	rs4349859	31.37	*HLA-B, MICA*	—	A	8.35E-09	2.198	1.682	2.875
EYE	6	rs4349859	31.37	*HLA-B, MICA*	—	A	1.99E-08	3.607	2.305	5.646
SKIN-3	5	rs80079682	173.29	*LINC1485, BOD1, CPEB4*	—	A	2.71E-08	1.519	1.311	1.760

[a]SNPs in LD $r^2 < 0.5$.
[b]Number of additional SNPs ($P < .0001$) in LD ($r^2 > 0.5$) with listed SNP.

Table 6. Extraintestinal Manifestation Co-occurrence Analyses

EIM 1	EIM-2	χ^2	*P* value	OR	95% CI	
					Lower	Upper
EN	PSC	38.4	5.66E-10	0.14	0.06	0.28
Ps	PSC	25.3	4.75E-07	0.33	0.2	0.52
AS-SI	EN	21.94	2.82E-06	0.27	0.14	0.48
EYE	PSC	21.1	4.35E-06	0.29	0.16	0.5
EN	Ps	11.97	5.00E-04	0.40	0.22	0.68
EN	EYE	10.5	1.00E-03	1.78	1.24	2.54
EYE	Ps	7.97	4.70E-03	0.42	0.21	0.76
PSC	PG	7.04	7.00E-03	0.39	0.17	0.78
EN	PG	5.8	1.50E-02	1.80	1.09	2.9

NOTE. Co-occurrence analysis for EIM pairs as shown.

Subjects With 2 or More EIM-6 Phenotypes subsection). We demonstrated presence of EN and either EYE ($P = 1.00 \times 10^{-3}$; OR, 1.78; 95% CI, 1.24–2.54) or PG ($P = 1.50 \times 10^{-2}$; OR, 1.80; 95% CI, 1.09–2.90) to be positively correlated (Table 6). However, most co-occurrence associations revealed negatively correlated EIMs, including EN and PSC ($P = 5.66 \times 10^{-10}$; OR, 0.14; 95% CI, 0.06–0.28), as well as Ps and PSC ($P = 4.75 \times 10^{-7}$; OR, 0.33; 95% CI, 0.20–0.52) (Table 6).

EIM-6

Characteristics associated with EIM-6 (Table 1) included CD diagnosis ($P = 1.66 \times 10^{-5}$; OR, 1.27; 95% CI, 1.14–1.42), female sex ($P = 9.00 \times 10^{-5}$; OR, 1.23; 95% CI, 1.11–1.36), any colonic disease in CD ($P = 9.78 \times 10^{-9}$; OR, 1.69; 95% CI, 1.41–2.02), proximal disease extent in UC ($P = 4.08 \times 10^{-4}$; OR, 2.36; 95% CI, 1.47–3.80), and a history of any IBD-related surgery ($P = 3.63 \times 10^{-19}$; OR, 1.72; 95% CI, 1.52–1.93). Female sex ($P = 6.74 \times 10^{-8}$; OR, 1.53; 95% CI, 1.31–1.79), any colonic disease ($P = 4.22 \times 10^{-9}$; OR, 1.91; 95% CI, 1.54–2.38), and surgery ($P = 1.32 \times 10^{-3}$; OR, 1.29; 95% CI, 1.10–1.51) in CD, and extensive disease ($P = 0.01$; OR, 2.07; 95% CI, 1.18–3.63) and surgery ($P = 9.06 \times 10^{-16}$; OR, 2.51; 95% CI, 2.01–3.15) in UC, remained associated after multivariable analysis (Supplementary Table 24).

Although no significant associations with serologies were seen after multivariable analyses (Supplementary Table 26), in univariable analyses, ANCA; anti-I2; and anti-OmpC in CD (Supplementary Table 26) and anti-CBir1 levels and QSS in UC (Supplementary Table 25) were associated with EIM-6.

Genome-wide significance with known IBD susceptibility locus *HLA-B/MICA* at rs4349859 ($P = 8.35 \times 10^{-9}$; OR, 2.20; 95% CI, 1.68–2.88) (Table 5) and nominal significance

($P < 1 \times 10^{-4}$) at *IL6* with EIM-6 were observed (Supplementary Tables 13 and 27).

Ankylosing Spondylitis and Sacroiliitis

Factors associated with AS-SI (Supplementary Table 4) included CD ($P = 1.80 \times 10^{-10}$; OR, 2.14; 05% CI, 1.69–2.70), older age at IBD diagnosis ($P = 9.44 \times 10^{-3}$; OR, 1.01; 95% CI, 1.00–1.02), and colectomy in UC ($P = 2.14 \times 10^{-4}$; OR, 2.56; 95% CI, 1.56–4.20). AS-SI was observed less frequently in female subjects ($P = 8.38 \times 10^{-3}$; OR, 0.76; 95% CI, 0.62–0.93). In multivariable analysis, female sex ($P = 0.03$; OR, 0.69; 95% CI, 0.49–0.96) and older age at diagnosis ($P = 0.03$; OR, 1.01; 95% CI, 1.00–1.03) in CD and colectomy ($P = 1.33 \times 10^{-4}$; OR, 2.64; 95% CI, 1.61–4.35) in UC remained significantly associated (Supplementary Table 28). Anti-CBir1 positivity was associated with AS-SI in subjects with CD (Supplementary Tables 25 and 29). MHC associations ($r^2 < 0.1$) in *HLA-C/HLA*-B (rs6905036; $P = 1.36 \times 10^{-15}$; OR, 2.50; 95% CI, 2.00–3.13) and *MICA/HCP5* (rs2844510; $P = 9.89 \times 10^{-13}$; OR, 2.11; 95% CI, 1.72–2.59) demonstrated genome-wide significance with AS-SI (Table 5). We observed association at $P < 1 \times 10^{-4}$ for numerous additional variants within MHC and other loci (Supplementary Tables 14 and 27).

Psoriasis

Most significant associations with Ps (Supplementary Table 5) included CD ($P = 2.19 \times 10^{-8}$; OR, 2.05; 95% CI, 1.59–2.63), female sex ($P = 0.03$; OR, 1.27; 95% CI, 1.03–1.57), Jewish ancestry ($P = 4.41 \times 10^{-4}$; OR, 1.51; 95% CI, 1.20–1.90), and smoking ($P = 5.60 \times 10^{-4}$; OR, 1.55; 95% CI, 1.21–1.99). In multivariable analysis, Jewish ancestry ($P = 0.01$; OR, 1.82; 95% CI, 1.15–2.87) and smoking ($P = 5.48 \times 10^{-3}$; OR, 2.10; 95% CI, 1.24–3.54) in UC and female sex ($P = 4.54 \times 10^{-3}$; OR, 1.53; 95% CI, 1.14–2.05) in CD remained significantly associated (Supplementary Table 30). Anti-I2 and anti-OmpC positivity were associated with Ps in subjects with CD in univariable analysis only (Supplementary Table 31). No variants achieved genome-wide significance with Ps, although variants in established Ps loci *PSORS1C1* (rs28732100; $P = 4.24 \times 10^{-7}$; OR, 1.04; 95% CI, 0.64–1.45) and *HLA-C/HLA-B* (rs12199223; $P = 8.69 \times 10^{-7}$; OR, 0.56; 95% CI, 0.34–0.78) approached genome-wide significance. Novel associations ($P < 1 \times 10^{-4}$) were observed for *DYTN* and *CEP128* (Supplementary Tables 15 and 27).

Pyoderma Gangrenosum

CD overall ($P = 1.67 \times 10^{-4}$; OR, 2.13; 95% CI, 1.44–3.17), female sex ($P = 5.68 \times 10^{-4}$; OR, 1.82; 95% CI, 1.30–2.57), any colonic disease ($P = 2.76 \times 10^{-4}$; OR, 4.16; 95% CI, 1.93–8.96), perianal disease in CD ($P = 3.53 \times 10^{-3}$; OR, 1.78; 95% CI, 1.21–2.63), and any IBD-related surgery ($P = 1.15 \times 10^{-7}$; OR, 2.85; 95% CI, 1.94–4.20) were associated with PG (Supplementary Tables 6 and 32). Small bowel disease ($P = 2.26 \times 10^{-4}$; OR, 0.47; 95% CI, 0.31–0.70) was less common in patients with PG. In multivariable analysis in CD, any colonic disease ($P = 1.15 \times 10^{-3}$; OR, 6.94; 95% CI, 2.16–22.30), any small bowel disease ($P = 1.47 \times 10^{-3}$; OR, 0.46; 95% CI, 0.29–0.74), and surgery ($P = 7.68 \times 10^{-6}$; OR, 3.30; 95% CI, 1.96–5.56) remained associated; female sex ($P = 0.02$; OR, 2.61; 95% CI, 1.18–5.76) and surgery ($P = 0.04$; OR, 2.13; 95% CI, 1.02–4.44) remained associated in UC (Supplementary Table 32). Increased risk of PG was associated with both ANCA positivity and level in CD, while anti-CBir1 positivity, IgG ASCA level, and QSS exhibited protective associations (Supplementary Tables 25 and 33). We observed nominal significance ($P < 1 \times 10^{-4}$) at several SNPs (Supplementary Table 16).

Erythema Nodosum

Increased risk of EN (Supplementary Table 7) was associated with CD ($P = 3.46 \times 10^{-15}$; OR, 3.43; 95% CI, 2.52–4.66), female sex ($P = 2.39 \times 10^{-18}$; OR, 3.11; 95% CI, 2.41–4.01), any colonic disease ($P = 4.96 \times 10^{-4}$; OR, 1.93; 95% CI, 1.33–2.79), and any IBD-related surgery ($P = 8.34 \times 10^{-5}$; OR, 1.66; 95% CI, 1.29–2.14). Older age at diagnosis was associated with decreased risk of EN in IBD ($P = 3.89 \times 10^{-5}$; OR, 0.98; 95% CI, 0.97–0.99). In multivariable analysis, female sex ($P = 8.42 \times 10^{-10}$; OR, 4.34; 95% CI, 2.71–6.93), age at diagnosis ($P = 0.02$; OR, 0.98; 95% CI, 0.96–1.0), and any colonic disease ($P = 0.04$; OR, 1.80; 95% CI, 1.04–3.13) remained significant in CD. Only female sex ($P = 7.65 \times 10^{-3}$; OR, 2.46; 95% CI, 1.27–4.78) remained associated in UC (Supplementary Table 34). Risk of EN was observed with anti-OmpC level in CD and positivity in UC (Supplementary Tables 25 and 35). Several known IBD susceptibility loci demonstrated nominal significance ($P < 1 \times 10^{-4}$) with EN, including *TSPAN14*, *HLA-DRB1/HLA-DQA1*, and PTPN2 (Supplementary Tables 17 and 27).

We observed genome-wide association in known IBD locus *BOD1/CPEB4* (rs80079682: $P = 2.71 \times 10^{-8}$; OR, 1.52; 95% CI, 1.31–1.76) (Table 6) for any subjects with any of Ps, PG, or EN (SKIN-3). Detailed associations with SKIN-3 are discussed in the Supplementary Results/Skin Manifestations subsection (see also Supplementary Tables 8, 18, 36, and 37).

Ocular Manifestations

CD ($P = 3.11 \times 10^{-8}$; OR, 2.18; 95% CI, 1.65–2.87) and female sex ($P = 1.75 \times 10^{-6}$; OR, 1.81; 95% CI, 1.42–2.32) were independently associated with increased risk of ocular manifestations (Supplementary Tables 9 and 38). Only female sex remained significant in multivariable analysis in CD ($P = 2.24 \times 10^{-4}$; OR, 1.79; 95% CI, 1.31–2.44) (Supplementary Table 38). No significant serologic associations were seen in CD or UC (Supplementary Table 39); however, QSS was associated with increased risk of ocular manifestations in UC (Supplementary Table 25). Known IBD susceptibility locus *HLA-B/MICA* demonstrated genome-wide significance at rs4349859 for EYE ($P = 1.99 \times 10^{-8}$; OR, 3.61; 95% CI, 2.31–5.65) (Table 5). Additional nominal associations ($P < 1 \times 10^{-4}$), including *CFB* and *XKR6/MTMR9*, were observed (Supplementary Tables 19 and 27).

Primary Sclerosing Cholangitis

Variables associated with increased risk of PSC (Supplementary Table 10) included any colonic disease in CD ($P = 1.29 \times 10^{-4}$; OR, 3.52; 95% CI, 1.85–6.70), proximal disease extent in UC ($P = 2.06 \times 10^{-4}$; OR, 6.54; 95% CI, 2.43–17.62), and any IBD-related surgery ($P = 9.95 \times 10^{-14}$; OR, 2.22; 95% CI, 1.80–2.73). CD ($P = 1.65 \times 10^{-35}$; OR, 0.28; 95% CI, 0.23–0.34), female sex ($P = 1.80 \times 10^{-4}$; OR, 0.70; 95% CI, 0.58–0.84), smoking ($P = 1.0 \times 10^{-3}$; OR, 0.62; 95% CI, 0.47–0.83), and complicated disease behavior in CD ($P = 3.25 \times 10^{-4}$; OR, 0.53; 95% CI, 0.38–0.75) were less common in patients with PSC. In multivariable analysis, sex (female; $P = 0.01$; OR, 0.54; 95% CI, 0.33–0.89), smoking ($P = 0.02$; OR, 0.37; 95% CI, 0.16–0.86), any colonic disease ($P = 4.43 \times 10^{-3}$; OR, 7.89; 95% CI, 1.90–32.71), and surgery ($P = 5.77 \times 10^{-8}$; OR, 3.80; 95% CI, 2.35–6.16) in CD and proximal disease extent ($P = 9.89 \times 10^{-3}$; OR, 3.76; 95% CI, 1.37–10.27), and surgery ($P = 5.15 \times 10^{-18}$; OR, 3.51; 95% CI, 2.64–4.67) in UC remained significant (Supplementary Table 40).

In CD, very significant associations between PSC and ANCA (positivity: $P = 3.37 \times 10^{-8}$; OR, 3.33; 95% CI, 2.17–5.11; and levels: $P = 7.27 \times 10^{-14}$; OR, 1.02; 95% CI, 1.01–1.02) were observed in univariable analysis. Other serologic associations are shown in Supplementary Tables 25 and 41. In multivariable analysis, ANCA levels remained significant in CD ($P = 4.68 \times 10^{-4}$; OR, 1.01; 95% CI, 1.01–1.02) and anti-I2 levels remained significantly associated in UC ($P = 3.28 \times 10^{-3}$; OR, 1.01; 95% CI, 1.00–1.02) (Supplementary Table 41).

A genome-wide genetic association was observed with PSC and rs9276456 near *HLA-DQA2/HLA-DQB2* ($P = 2.71 \times 10^{-10}$; OR, 2.76; 95% CI, 2.01–3.77) (Table 5), with an additional 36 SNPs in LD with rs9276456 ($r^2 > 0.5$) also reaching genome-wide significance (data not shown). Numerous additional SNPs within MHC were nominally significant at $P < 1 \times 10^{-4}$ (Supplementary Table 20). Additional nominal genetic associations observed outside the MHC included *LGALS9/NOS2*, *MUC19/LRRK2*, and *CLEC16A* (Supplementary Table 20). Associations with PSC demonstrated a strong enrichment of genes involved in allograft rejection ($P_{adjusted} = 1.17 \times 10^{-10}$), cell adhesion molecules ($P_{adjusted} = 1.47 \times 10^{-10}$), graft-vs-host disease ($P_{adjusted} = 2.18 \times 10^{-10}$), type 1 diabetes mellitus ($P_{adjusted} = 3.08 \times 10^{-10}$), and antigen processing and presentation ($P_{adjusted} = 1.59 \times 10^{-9}$), in addition to other immune-mediated pathways (Supplementary Table 27).

Peripheral Arthritis

The most frequent EIM observed was PA (n = 2040; Supplementary Table 11), with the highest percentage of PA observed in SHARE probably reflecting the self-reporting of this phenotype (Supplementary Table 1). PA was more frequent in CD ($P = 8.93 \times 10^{-17}$; OR, 1.58; 95% CI, 1.42–1.76), female sex ($P = 5.65 \times 10^{-17}$; OR, 1.54; 95% CI, 1.39–1.70), older age at IBD diagnosis ($P = 1.59 \times 10^{-11}$; OR, 1.01; 95% CI, 1.01–1.02), smoking ($P = 5.24 \times 10^{-13}$; OR, 1.6; 95% CI, 1.41–1.82), any colonic disease in CD ($P = 2.18 \times 10^{-6}$; OR, 1.45; 95% CI, 1.24–1.69) and any IBD-related surgery ($P = 3.73 \times 10^{-13}$; OR, 1.49; 95% CI, 1.34–1.66). Younger age at diagnosis was protective for PA. All variables remained associated with PA in CD with multivariable analyses, while all but smoking remained significant in UC (Supplementary Table 42). Anti-I2 positivity and level and anti-OmpC positivity were associated with PA in CD (Supplementary Tables 25 and 43). Nominal genetic associations ($P < 1 \times 10^{-4}$) with PA were identified, including *JAK2* and *GPR35/AQP12B* (Supplementary Table 21).

Given the high prevalence of PA, we created EIM-7 when PA was included with the other EIM phenotypes, and associations with EIM-7 are detailed in the Supplementary Results/EIM-7 subsection.

Discussion

EIMs represent a challenge in clinical practice. Gaining better understanding of the underlying mechanisms of EIMs may help identify patients at increased risk, allowing early therapeutic intervention and improved quality of life. Here, we present the largest investigation to date of characteristics associated with the development of EIMs in IBD to better understand the underlying pathogenesis and ultimately develop a more personalized approach to managing IBD.

We observed a much higher PA prevalence in the SHARE cohort (50%) in comparison with the other cohorts in which PA was diagnosed by experienced IBD physicians. PA prevalence was lowest in the CSMC (adult and pediatric) and RISK (pediatric) cohorts. Distinguishing inflammatory arthritis from arthralgia and osteoarthritis can be challenging and, for this reason, we excluded PA from our analyses of "any" (EIM-6) or multiple EIMs. This example highlights the challenges of studying rare and complex phenotypes and the balance of increasing power through combining cohorts with the concern about definitions across different cohorts. Nevertheless, with the notable exception of PA, our prevalence of EIMs across the cohorts is consistent with those reported previously and, reassuringly, we confirmed previously identified EIM associations.[3,4,8,24–26]

In multivariable analyses across all EIMs, we observed increased risks associated with CD and female sex consistent with previous reports.[3,4] Notable exceptions include increased risk of PSC with UC, and increased risk of AS-SI and PSC and male sex, also consistent with prior studies.[24] Multivariable analyses further highlighted robust associations between several EIMs and colonic CD, as well as IBD-related surgery, again, as reported previously[27]; these findings were the most robust in PSC. In contrast to the increased risk of EIM-6, PG, EN, SKIN-3, PSC, PA, EIM-7, and ≥2 EIMs with colonic disease, any small bowel disease involvement was only associated with an increased risk for Ps in univariable analyses. Differences in underlying biology between small bowel and colonic inflammation have recently been highlighted, with some suggesting a need for a new classification of CD location based on the presence or absence of small bowel disease.[28] Our data lend further

weight to these arguments. If we believe that these EIMs are gut-"derived," then elucidation of differences in small bowel vs colonic biology may shed light on the underlying mechanisms that lead to manifestations beyond the gut. Prior studies have suggested that EIMs co-occur and the presence of 1 EIM increases a patient's risk of developing additional EIMs.[6] In contrast to this dogma, only 2% of our cohort had multiple EIMs and, except for EN co-occurring with EYE or PG, our data demonstrated a strong negative correlation for most EIMs. Our findings are supported by a lack of common underlying etiology shared across all EIMs. Prior observations of high EIM co-occurrence may be due to ascertainment bias and/or smaller number of IBD cases studied in comparison with this study.

We and others have reported extensively on the role of circulating antibodies to a subset of microbial antigens (ie, ASCA, anti-I2, anti-OmpC, and anti-CBir1) and autoantigens (ie, ANCA) in IBD.[29,30] Although there are no specific biomarkers for EIM activity in IBD, studies have reported on associations with these IBD serologies and various EIMs. Our group has previously examined IBD-associated serologies in AS, PSC, skin, and ocular manifestations.[8,25,26,31] In this study, which represents a larger cohort than was included in previous analyses, we observed the strongest antibody associations with EIM-6 and PSC, predominantly in CD. We demonstrated the most robust univariable associations with PSC in CD, including ANCA seropositivity and ASCA seronegativity. It has been reported that a subset of ANCA-positive subjects with CD exhibited clinical characteristics commonly seen in UC, such as left-sided colitis and left-sided colonic inflammation.[30,32] In CD, increased levels of ANCA with absence of ASCA (ANCA+/ASCA–) was associated with noncomplicated disease behavior.[32] Subsequently, it was noted that ANCA+/ASCA– was associated with nonresponse to anti–tumor necrosis (anti-TNF) factor in subjects with CD and additional studies have observed ANCA association with poor response to anti-TNF treatments in both CD and UC.[33,34] Furthermore, subjects with IBD with PSC treated with anti-TNF therapy had an increased risk of developing acute cholangitis.[35] Our ANCA+/ASCA– signal in PSC may be indicative of a homogeneous subgroup of subjects with CD with a poor prognosis with standard biologic treatments.

AS-IBD has been associated with IgG ASCA, anti-OmpC, and anti-CBir1,[8] and people with AS have demonstrated elevated anti-I2 and ANCA levels.[8,31] ASCA IgG and anti-CBir1 associations have also been demonstrated in AS, but specifically in subjects with fecal calprotectin–positive AS not known to have IBD.[31] In this study, we observed an association with anti-CBir1 positivity in patients with CD with AS-SI. Anti-CBir1 is associated with complicated disease behavior,[36] as well as earlier-onset IBD (younger than 7 years).[37] Given that the most common EIM observed in pediatric IBD involves joints, our anti-CBir1 association with AS may be indicative of an age-dependent novel CD subgroup with poor prognosis. We did not observe any significant serologic associations with ocular EIMs, consistent with our earlier observations.[26] Ocular EIMs occur infrequently and our sample size, despite being the largest to-date, may still not be appropriately powered to demonstrate any serologic differences here.

During the past decade, the role of genetics in the development of IBD has been well-described.[15] Although the genetic contribution to EIMs in IBD has not been investigated as intensively, genetic involvement in EIM complications is supported by reports of strong familial concordance in EIM[17] and studies demonstrating overlap in genetic risk loci and shared biologic pathways, particularly for PSC and AS.[38] Genetic risk factors, notably with HLA alleles, have also been demonstrated for various EIMs.[27] We observed, for the first time, genome-wide significance for 5 EIM phenotypes and variants in the MHC and also *CPEB4*. We identified an association for SKIN-3 manifestation and a variant approximately 23 kb 5′ of known IBD locus *CPEB4*,[15] encoding an RNA-binding protein that regulates activation of the unfolded protein response and required for cell cycle progression, specifically for cytokinesis. Our associated variant ($r^2 < 0.2$ with IBD-associated *CPEB4* variants[15,39]) is a strong expression quantitative trait locus (eQTL) for *CPEB4* in whole blood and resides in a genomic region demonstrating histone modifications consistent with enhancer elements in rectal mucosa.

Multiple novel variants showed associations with various EIMs at nominal significance. Of particular interest, we identified a variant in *IL6* (rs2069835) associated with SKIN and EIM-6. IL6 has a key role in modulating inflammation, skin fibrosis, and wound healing. Given that uninhibited intestinal inflammation is a critical aspect of IBD, efforts have focused on drugs targeting the IL6 pathway for various inflammatory immune-mediated diseases.[40] In addition, we observed a nominally significant SNP in *ERBB4* (rs6711391) associated with SKIN-3. ERBB4 expression is induced in the colon in IBD and contributes to the homeostatic maintenance of the intestinal tract.

Our genome-wide associations within MHC are particularly interesting, given its well-established role in IBD.[41] Specifically, we observed associations for EIM-6, AS-SI, and eye complications with variants intergenic to *HLA-B* and *MICA*. One of the best-known genetic risk factors for EIM is *HLA-B*27*, widely reported to be associated with skin, eye, and joint manifestations, particularly AS, which has demonstrated the strongest association with polymorphisms in *HLA-B*.[6,42,43] We observed a second, independent ($r^2 < 0.1$) genome-wide association with AS-SI and rs2844510 intergenic to *MICA*, an MHC Class I Chain-Related Protein A gene whose protein mediates activation of natural killer cells, $\gamma\delta$ T cells, and a subset of $CD8^+$ T cells, is located near *HLA-B*, and has been implicated in AS susceptibility in both European and East Asian ancestry populations.[44] We detected nominal associations for AS and additional independent variants ($r^2 < 0.3$ with genome-wide–associated AS-SI SNPs) in the genomic region harboring *HLA-B* and *MICA*, lending further support for a role for these loci in AS in IBD.

Multiple *HLA* susceptibility factors have been reported previously with PSC, including *HLA-B*, *HLA-DRB3*0101, HLA-DRB1*0301, and HLA-DQA1*01:03*, among others.[45] We demonstrated genome-wide significance for PSC and

rs9276456 residing <1 kb 3′ of *HLA-DQA2*. We also observed nominal association with rs9266669 intergenic to *HLA-B*. This variant is in high LD ($r^2 = 0.80$) with a PSC risk variant rs4143332 reported to be in near-perfect LD ($r^2 = 0.996$) with *HLA-B*08:01*.[45] rs9266669 is also in high LD ($r^2 = 0.77$) with previously reported PSC genome-wide association study variants rs3099844 and rs2844559 located near *HLA-B*.[46] We detected additional nominal associations with PSC and multiple SNPs ($r^2 < 0.5$) near other *HLA* genes, including *HLA-DQB1*, a locus implicated in PSC and autoimmune hepatitis, further strengthening the role of HLA genes and PSC complications in IBD. Aside from MHC, we did not observe associations for PSC in previously reported PSC risk loci, apart from *CLEC16A*, a gene implicated in both PSC and IBD.[16,47] Interestingly, our associated intronic variant rs887864 is in high LD with a PSC-associated (rs725613; $r^2 = 0.91$), but not the IBD-associated *CLEC16A* (rs8061882; $r^2 = 0.1$) variant.[16,47]

Dysregulation of immune homeostasis and a role for T helper 17 cells are both central to IBD and intestinal allograft rejection.[48] Interestingly, allograft rejection was highlighted in gene enrichment analyses of multiple EIMs, including EN, PSC, and EIM-6 (Supplementary Table 27). Genetic associations with EN, PSC, and EIM-6 were also enriched for genes involved in graft-vs-host disease, which shares key characteristics with IBD, such as intestinal tissue damage and loss of intestinal barrier function.[49] We further observed an enrichment of genes involved in autoimmune thyroid disease with PSC and EIM-6, which is interesting, given that thyroid, gastrointestinal, and liver dysfunction are closely intertwined.[50]

We have previously published studies examining EIMs that were performed on older genotyping arrays with more limited sample sizes.[25,26] These current analyses not only used the immune-focused Immunochip array, but also included an expanded CSMC study cohort in addition to NIDDK, SHARE, and RISK cohorts, allowing for the investigation of IBD EIMs in the largest sample size to date. As mentioned previously, studies such as these face the challenge of balancing a need for increasing study size for improved statistical power, with the potential for variation in phenotyping and lack of specificity around EIM diagnoses across cohorts. This study was not population-based, and subjects included were recruited at major IBD centers and EIM diagnoses (except PA in 1 cohort) were verified by experienced IBD clinicians and specialists. Thorough inter- and intra-cohort QC metrics were also applied. Nevertheless, heterogeneity in data collection methods is 1 key limitation of large multicenter studies. A specific limitation of our study includes incomplete disease duration data. Accounting for disease duration would enable an adjustment for our non-EIM control subjects with IBD recruited early in their disease course, who may subsequently develop EIM over time. We believe that our analysis has adopted a more conservative approach with the inclusion of subjects with IBD in the non-EIM control group that may have subsequently developed EIM. Some of our observed clinical and genetic associations were consistent with previously published studies. However, others differ from previous reports, including that only approximately 2% of our subjects had ≥2 EIMs, which is lower than previous reports and may be reflective of our recruitment centers or the large size of our study in comparison with other studies. We acknowledge some additional limitations, including restricting our analyses to subjects of predominantly European ancestry. Our efforts incorporating non-European ancestry cohorts in future trans-ethnic studies will increase study power. In addition, serology was not available across all cohorts. Future studies will likely use platforms with denser and broader "coverage" moving from targeted arrays, such as Immunochip to genome-wide arrays and ultimately whole genome sequencing. A critical role for MHC in EIMs was supported by our genome-wide SNP associations within MHC and numerous additional associations at nominal significance levels, although we were unable to assess the association with specific HLA alleles. Imputation of the MHC to 4-digit HLA allele resolution will allow for more in-depth investigation of this region.

IBD is characterized by variable degrees of disease heterogeneity, including the presence or absence of EIMs, which can be a significant source of morbidity and poor quality of life.[2–5] Here we have reported the largest multicenter study of clinical, serologic, and genetic associations with EIMs in IBD. We have identified novel and confirmed previously known associations between sex, disease type, location, severity, and need for surgery and IBD-associated serologies. In contrast to previous smaller studies, only a small fraction of our cohort presented with multiple EIMs, with strong negative correlation observed for multiple EIM pairs. In addition to confirming known genetic associations (but now at genome-wide level) within MHC, we have identified a novel genome-wide association for skin manifestations and *CPEB4.* Our genetic findings also implicated pathways and genes that are either targets for existing drugs (eg, anti-TNFs, cell adhesion molecules, and JAK inhibitors) or therapies in development (eg, anti-IL6), as well as implicating IL6 in the development of EIMs. These findings contribute to a more complete understanding of the underlying pathogenesis of EIMs and the molecular and other associations implicated in the clinical heterogeneity of IBD. These are important steps in the path to more personalized approaches to the management of IBD, which may be of particular importance in patients with UC and CD with more systemic phenotypes.

Supplementary Material

Note: To access the supplementary material accompanying this article, visit the online version of *Gastroenterology* at www.gastrojournal.org, and at http://doi.org/10.1053/j.gastro.2024.02.026.

References

1. McGovern DP, Kugathasan S, Cho JH. Genetics of inflammatory bowel diseases. Gastroenterology 2015; 149:1163–1176.e2.

2. Vavricka SR, Brun L, Ballabeni P, et al. Frequency and risk factors for extraintestinal manifestations in the Swiss inflammatory bowel disease cohort. Am J Gastroenterol 2011;106:110–119.
3. Ott C, Schölmerich J. Extraintestinal manifestations and complications in IBD. Nat Rev Gastroenterol Hepatol 2013;10:585–595.
4. Bernstein CN, Blanchard JF, Rawsthorne P, et al. The prevalence of extraintestinal diseases in inflammatory bowel disease: a population-based study. Am J Gastroenterol 2001;96:1116–1122.
5. Repiso A, Alcántara M, Muñoz-Rosas C, et al. Extraintestinal manifestations of Crohn's disease: prevalence and related factors. Rev Esp Enferm Dig 2006; 98:510–517.
6. Greuter T, Vavricka SR. Extraintestinal manifestations in inflammatory bowel disease - epidemiology, genetics, and pathogenesis. Expert Rev Gastroenterol Hepatol 2019;13:307–317.
7. Bernstein CN, Wajda A, Blanchard JF. The clustering of other chronic inflammatory diseases in inflammatory bowel disease: a population-based study. Gastroenterology 2005;129:827–836.
8. Mundwiler ML, Mei L, Landers CJ, et al. Inflammatory bowel disease serologies in ankylosing spondylitis patients: a pilot study. Arthritis Res Ther 2009;11:R177.
9. Greenstein AJ, Janowitz HD, Sachar DB. The extraintestinal complications of Crohn's disease and ulcerative colitis: a study of 700 patients. Medicine (Baltimore) 1976;55:401–412.
10. Freeman HJ. Erythema nodosum and pyoderma gangrenosum in 50 patients with Crohn's disease. Can J Gastroenterol 2005;19:603–606.
11. Levine JS, Burakoff R. Extraintestinal manifestations of inflammatory bowel disease. Gastroenterol Hepatol (N Y) 2011;7:235–241.
12. Broomé U, Bergquist A. Primary sclerosing cholangitis, inflammatory bowel disease, and colon cancer. Semin Liver Dis 2006;26:31–41.
13. Rothfuss KS, Stange EF, Herrlinger KR. Extraintestinal manifestations and complications in inflammatory bowel diseases. World J Gastroenterol 2006;12:4819–4831.
14. Isene R, Bernklev T, Høie O, et al. Extraintestinal manifestations in Crohn's disease and ulcerative colitis: results from a prospective, population-based European inception cohort. Scand J Gastroenterol 2015;50:300–305.
15. **de Lange KM**, **Moutsianas L**, **Lee JC**, et al. Genome-wide association study implicates immune activation of multiple integrin genes in inflammatory bowel disease. Nat Genet 2017;49:256–261.
16. Ellinghaus D, Jostins L, Spain SL, et al. Analysis of five chronic inflammatory diseases identifies 27 new associations and highlights disease-specific patterns at shared loci. Nat Genet 2016;48:510–518.
17. Satsangi J, Grootscholten C, Holt H, et al. Clinical patterns of familial inflammatory bowel disease. Gut 1996; 38:738–741.
18. Satsangi J, Silverberg MS, Vermeire S, et al. The Montreal classification of inflammatory bowel disease: controversies, consensus, and implications. Gut 2006; 55:749–753.
19. Chang CC, Chow CC, Tellier LC, et al. Second-generation PLINK: rising to the challenge of larger and richer datasets. Gigascience 2015;4:7.
20. R Development Team. R. A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, 2018.
21. Machiela MJ, Chanock SJ. LDlink: a web-based application for exploring population-specific haplotype structure and linking correlated alleles of possible functional variants. Bioinformatics 2015;31:3555–3557.
22. McLean CY, Bristor D, Hiller M, et al. GREAT improves functional interpretation of cis-regulatory regions. Nat Biotechnol 2010;28:495–501.
23. Kuleshov MV, Jones MR, Rouillard AD, et al. Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. Nucleic Acids Res 2016;44:W90–W97.
24. Landi M, Maldonado-Ficco H, Perez-Alamino R, et al. Gender differences among patients with primary ankylosing spondylitis and spondylitis associated with psoriasis and inflammatory bowel disease in an iberoamerican spondyloarthritis cohort. Medicine (Baltimore) 2016;95:e5652.
25. Weizman A, Huang B, Berel D, et al. Clinical, serologic, and genetic factors associated with pyoderma gangrenosum and erythema nodosum in inflammatory bowel disease patients. Inflamm Bowel Dis 2014;20:525–533.
26. Taleban S, Li D, Targan SR, et al. Ocular manifestations in inflammatory bowel disease are associated with other extra-intestinal manifestations, gender, and genes implicated in other immune-related traits. J Crohns Colitis 2016;10:43–49.
27. Hedin CRH, Vavricka SR, Stagg AJ, et al. The pathogenesis of extraintestinal manifestations: implications for IBD research, diagnosis, and therapy. J Crohns Colitis 2019;13:541–554.
28. **Potdar AA**, **Dube S**, Naito T, et al. Altered intestinal ACE2 levels are associated with inflammation, severe disease, and response to anti-cytokine therapy in inflammatory bowel disease. Gastroenterology 2021; 160:809–822.e7.
29. Mitsuyama K, Niwa M, Takedatsu H, et al. Antibody markers in the diagnosis of inflammatory bowel disease. World J Gastroenterol 2016;22:1304–1310.
30. Vasiliauskas EA, Plevy SE, Landers CJ, et al. Perinuclear antineutrophil cytoplasmic antibodies in patients with Crohn's disease define a clinical subgroup. Gastroenterology 1996;110:1810–1819.
31. Matzkies FG, Targan SR, Berel D, et al. Markers of intestinal inflammation in patients with ankylosing spondylitis: a pilot study. Arthritis Res Ther 2012;14:R261.
32. Targan SR. The utility of ANCA and ASCA in inflammatory bowel disease. Inflamm Bowel Dis 1999;5:61–63; discussion 66–67.
33. Esters N, Vermeire S, Joossens S, et al. Serological markers for prediction of response to anti-tumor necrosis factor treatment in Crohn's disease. Am J Gastroenterol 2002;97:1458–1462.

34. Nguyen DL, Nguyen ET, Bechtold ML. pANCA positivity predicts lower clinical response to infliximab therapy among patients with IBD. South Med J 2015;108:139–143.
35. Kulkarni C, Murag S, Cholankeril G, et al. Association of anti-TNF therapy with increased risk of acute cholangitis in patients with primary sclerosing cholangitis. Inflamm Bowel Dis 2021;27:1602–1609.
36. Papadakis KA, Yang H, Ippoliti A, et al. Anti-flagellin (CBir1) phenotypic and genetic Crohn's disease associations. Inflamm Bowel Dis 2007;13:524–530.
37. Markowitz J, Kugathasan S, Dubinsky M, et al. Age of diagnosis influences serologic responses in children with Crohn's disease: a possible clue to etiology? Inflamm Bowel Dis 2009;15:714–719.
38. van Sommeren S, Janse M, Karjalainen J, et al. Extra-intestinal manifestations and complications in inflammatory bowel disease: from shared genetics to shared biological pathways. Inflamm Bowel Dis 2014; 20:987–994.
39. **Huang H**, **Fang M**, **Jostins L**, et al. Fine-mapping inflammatory bowel disease loci to single-variant resolution. Nature 2017;547:173–178.
40. Choy EH, De Benedetti F, Takeuchi T, et al. Translating IL-6 biology into effective treatments. Nat Rev Rheumatol 2020;16:335–345.
41. **Goyette P**, **Boucher G**, Mallon D, et al. High-density mapping of the MHC identifies a shared role for HLA-DRB1*01:03 in inflammatory bowel diseases and heterozygous advantage in ulcerative colitis. Nat Genet 2015;47:172–179.
42. Orchard TR, Chua CN, Ahmad T, et al. Uveitis and erythema nodosum in inflammatory bowel disease: clinical features and the role of HLA genes. Gastroenterology 2002;123:714–718.
43. Brown MA, Pile KD, Kennedy LG, et al. HLA class I associations of ankylosing spondylitis in the white population in the United Kingdom. Ann Rheum Dis 1996; 55:268–270.
44. Zhou X, Wang J, Zou H, et al. MICA, a gene contributing strong susceptibility to ankylosing spondylitis. Ann Rheum Dis 2014;73:1552–1557.
45. **Liu JZ**, **Hov JR**, **Folseraas T**, et al. Dense genotyping of immune-related disease regions identifies nine new risk loci for primary sclerosing cholangitis. Nat Genet 2013; 45:670–675.
46. Karlsen TH, Franke A, Melum E, et al. Genome-wide association analysis in primary sclerosing cholangitis. Gastroenterology 2010;138:1102–1111.
47. **Ji SG**, **Juran BD**, Mucha S, et al. Genome-wide association study of primary sclerosing cholangitis identifies new risk loci and quantifies the genetic relationship with inflammatory bowel disease. Nat Genet 2017; 49:269–273.
48. Belyayev L, Loh K, Fishbein TM, et al. The parallel paradigm between intestinal transplant inflammation and inflammatory bowel disease. Curr Opin Organ Transplant 2019;24:207–211.
49. Haring E, Zeiser R, Apostolova P. Interfering with inflammation: heterogeneous effects of interferons in graft-vs-host disease of the gastrointestinal tract and inflammatory bowel disease. Front Immunol 2021;12: 705342.
50. Kyriacou A, McLaughlin J, Syed AA. Thyroid disorders and gastrointestinal and liver dysfunction: a state of the art review. Eur J Intern Med 2015;26:563–571.

Author names in bold designate shared co-first authorship.

Received July 16, 2022. Accepted February 13, 2024.

Correspondence
Address correspondence to: Dermot P. B. McGovern, MD, PhD, F. Widjaja Inflammatory Bowel Disease Institute, Cedars-Sinai Medical Center, 8730 Alden Drive, Suite 2E242, Los Angeles, California 90048. e-mail: Dermot.McGovern@cshs.org.

Acknowledgments
The authors are thankful to all of the clinicians, researchers, coordinators, and patients who have contributed time, data, and samples to the MIRIAD Biobank, as well as the National Institute for Diabetes and Digestive and Kidney Diseases Inflammatory Bowel Disease Genetics Consortium, Sinai Helmsley Alliance for Research Excellence, and Risk Stratification and Identification of Immunogenetic and Microbial Markers of Rapid Disease Progression in Children with Crohn's Disease Consortia.

CRediT Authorship Contributions
Michelle Khrom, MSN (Formal analysis: Lead; Investigation: Lead; Methodology: Equal; Writing – original draft: Lead)
Millie Long, MD, MPH (Data curation: Lead; Writing – review & editing: Supporting)
Shishir Dube, PhD (Formal analysis: Lead; Investigation: Supporting; Methodology: Equal; Writing – original draft: Supporting)
Lori Robbins, MD (Resources: Equal; Writing – review & editing: Supporting)
Gregory J. Botwin, BS (Methodology: Supporting; Resources: Supporting; Writing – review & editing: Supporting)
Shaohong Yang, MD (Data curation: Lead; Writing – review & editing: Supporting)
Emebet Mengesha, BS (Data curation: Lead; Investigation: Supporting; Project administration: Lead; Writing – review & editing: Supporting)
Dalin Li, PhD (Formal analysis: Supporting; Investigation: Supporting; Methodology: Equal; Writing – review & editing: Supporting)
Takeo Naito, MD, PhD (Formal analysis: Supporting; Methodology: Supporting; Writing – review & editing: Supporting)
Nirupama N. Bonthala, MD (Resources: Equal; Writing – review & editing: Supporting)
Christina Ha, MD (Resources: Equal; Writing – review & editing: Supporting)
Gil Melmed, MD (Resources: Equal; Writing – review & editing: Supporting)
Shervin Rabizadeh, MD, MBA (Resources: Equal; Writing – review & editing: Supporting)
Gaurav Syal, MD, MHDS (Resources: Equal; Writing – review & editing: Supporting)
Eric Vasiliauskas, MD (Resources: Equal; Writing – review & editing: Supporting)
David Ziring, MD (Resources: Equal; Writing – review & editing: Supporting)
Steven R. Brant, MD (Funding acquisition: Supporting; Resources: Equal; Writing – review & editing: Supporting)
Judy Cho, MD (Funding acquisition: Supporting; Resources: Equal; Writing – review & editing: Supporting)
Richard H. Duerr, MD (Funding acquisition: Supporting; Resources: Equal; Writing – review & editing: Supporting)
John Rioux, PhD (Funding acquisition: Supporting; Resources: Equal; Writing – review & editing: Supporting)
Phil Schumm, MA (Data curation: Lead; Resources: Supporting; Writing – review & editing: Supporting)
Mark Silverberg, MD, PhD (Funding acquisition: Supporting; Resources: Equal; Writing – review & editing: Supporting)
Ashwin N. Ananthakrishnan, MD, MPH (Resources: Equal; Writing – review & editing: Supporting)
William A. Faubion, MD (Funding acquisition: Supporting; Resources: Equal; Writing – review & editing: Supporting)
Bana Jabri, MD, PhD (Funding acquisition: Supporting; Resources: Equal; Writing – review & editing: Supporting)
Sergio A. Lira, MD, PhD (Funding acquisition: Supporting; Resources: Equal; Writing – review & editing: Supporting)
Rodney D. Newberry, MD (Funding acquisition: Supporting; Resources: Equal; Writing – review & editing: Supporting)
Robert S. Sandler, MD, MPH (Funding acquisition: Supporting; Resources: Equal; Writing – review & editing: Supporting)
Ramnik J. Xavier, MD, PhD (Funding acquisition: Supporting; Resources: Equal; Writing – review & editing: Supporting)

Subra Kugathasan, MD (Funding acquisition: Supporting; Resources: Equal; Writing – review & editing: Supporting)

David Hercules, N/A (Data curation: Lead; Writing – review & editing: Supporting)

Stephan R. Targan, MD (Funding acquisition: Equal; Resources: Equal; Writing – review & editing: Supporting)

R. Balfour Sartor, MD (Funding acquisition: Supporting; Resources: Equal; Writing – review & editing: Supporting)

Talin Haritunians, PhD (Data curation: Supporting; Formal analysis: Lead; Investigation: Lead; Methodology: Equal; Supervision: Equal; Writing – original draft: Lead)

Dermot P. B. McGovern, MD, PhD (Conceptualization: Lead; Funding acquisition: Lead; Investigation: Supporting; Methodology: Supporting; Supervision: Equal; Writing – original draft: Lead)

Conflicts of interest

These authors disclose the following: Dalin Li, Nirupama N. Bonthala, Christina Ha, Gil Melmed, Shervin Rabizadeh, Gaurav Syal, Eric Vasiliauskas, David Ziring, Stephan R. Targan, Talin Haritunians, and Dermot P. B. McGovern are faculty members at Cedars-Sinai Medical Center. Cedars-Sinai has the right to receive future royalty payments from Prometheus Biosciences, Inc, a wholly owned subsidiary of Merck, which has access to the data and specimens in Cedars-Sinai's MIRIAD Biobank (including a proportion of the data used in this study). Millie Long is a consultant for AbbVie, Janssen, Pfizer, Takeda, BMS, Target PharmaSolutions, Prometheus, Lilly, Salix, Valeant, Genentech, Roche, and Calibr, and received research support from Pfizer and Takeda. Gil Melmed is a consultant for AbbVie, Arena, Boomerang, Bristol-Myers Squibb, Dieta, Ferring, Janssen, Lilly, Oshi Health, Pfizer, Shield, Shionogi, Takeda, and Techlab, and received research support from Pfizer. Ramnik J. Xavier is the co-founder of Celsius Therapeutics and Jnana Therapeutics and SAB of Moonlake Immunotherapeutics, Senda Biosciences, and Nestle. Dermot P. B. McGovern has consulted for Bridge Biotherapeutics, Takeda, and Palatin Technologies. Dalin Li, Shervin Rabizadeh, Stephan R. Targan, and Dermot P. B. McGovern are consultants for Merck. R. Balfour Sartor has consulted for and received grant support from Takeda, Janssen, Second Genome, Vedanta, BiomX, Biomica, SERES, and Artizan. The remaining authors disclose no conflicts.

Funding

This work was supported by internal funds from the F. Widjaja Foundation Inflammatory Bowel Disease Institute. The Cedars-Sinai MIRIAD Biobank is supported by the F. Widjaja Foundation Inflammatory Bowel Disease Institute, National Institutes of Health/National Institute of Diabetes and Digestive and Kidney Diseases grants (P01 DK046763 and U01 DK062413), and The Leona M. and Harry B. Helmsley Charitable Trust (Dermot P. B. McGovern and Sinai Helmsley Alliance for Research Excellence Consortium).

Data Availability

Data underlying this article are available in the article and in the Supplementary Materials. Summary statistics will be made available through the National Institute for Diabetes and Digestive and Kidney Diseases Inflammatory Bowel Disease Genetics Consortium portal, currently under development.