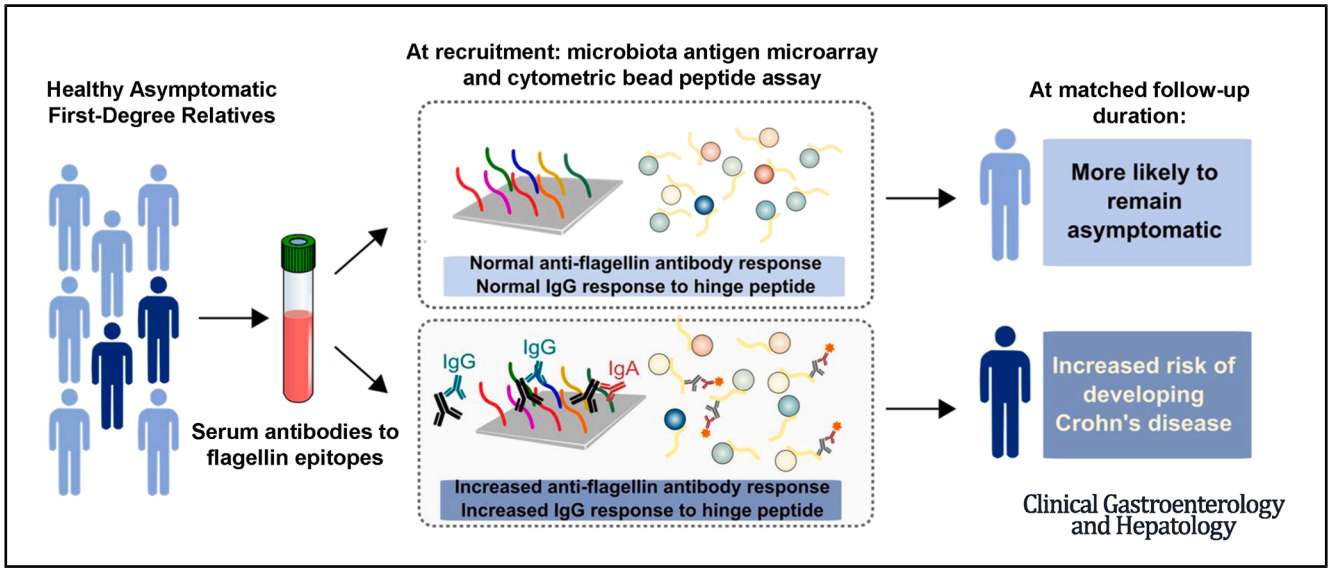
Clinical Gastroenterology and Hepatology 2026;24:2226–2236

INFLAMMATORY BOWEL DISEASE

Serum IgG Response to a Conserved Domain of Commensal Flagellins Predicts Future Risk of Crohn's Disease in First-Degree Relatives

Richard Y. Wu,[1,*] Mingyue Xue,[2,*] Qing Zhao,[3,4] Sean Jeong,[2] Anne M. Griffiths,[5] Levinus A. Dieleman,[6] A. Hillary Steinhart,[1,2] Guy Aumais,[7] Brian Bressler,[8] Remo Panaccione,[9] Colette Deslandres,[10] David R. Mack,[11] Charles N. Bernstein,[12] John K. Marshall,[13] Dan Turner,[14] Wei Xu,[15,16] Lennard W. Duck,[3] Charles O. Elson,[3,4] Williams Turpin,[1,2] Sun-Ho Lee,[1,2,§] and Kenneth Croitoru[1,2,§]

[1]Division of Gastroenterology and Hepatology, Temerty Faculty of Medicine, University of Toronto, Toronto, Ontario, Canada; [2]Division of Gastroenterology, Inflammatory Bowel Disease Centre, Mount Sinai Hospital, Toronto, Ontario, Canada; [3]Department of Medicine, University of Alabama at Birmingham, Birmingham, Alabama; [4]Department of Microbiology, University of Alabama at Birmingham, Birmingham, Alabama; [5]Division of Gastroenterology, The Hospital for Sick Children, Toronto, Ontario, Canada; [6]Division of Gastroenterology, Department of Medicine, University of Alberta, Edmonton, Alberta, Canada; [7]Department of Medicine, Hôpital Maisonneuve-Rosemont, University of Montreal, Montreal, Quebec, Canada; [8]The Inflammatory Bowel Disease Centre of British Columbia, University of British Columbia, Vancouver, British Columbia, Canada; [9]Division of Gastroenterology and Hepatology, Inflammatory Bowel Disease Clinic, University of Calgary, Calgary, Alberta, Canada; [10]Department of Hepatology and Pediatric Nutrition, Centre Hospitalier Universitaire Sainte-Justine, Montreal, Quebec, Canada; [11]Division of Gastroenterology, Hepatology, and Nutrition, Children's Hospital of Eastern Ontario and University of Ottawa, Ottawa, Ontario, Canada; [12]University of Manitoba Inflammatory Bowel Disease Clinical and Research Centre and Department of Internal Medicine, Max Rady College of Medicine, Rady Faculty of Health Sciences, University of Manitoba, Winnipeg, Manitoba, Canada; [13]Department of Medicine, Farncombe Family Digestive Health Research Institute, McMaster University, Hamilton, Ontario, Canada; [14]The Juliet Keidan Institute of Pediatric Gastroenterology and Nutrition, Shaare Zedek Medical Center, The Hebrew University of Jerusalem, Jerusalem, Israel; [15]Division of Biostatistics, Dalla Lana School of Public Health, University of Toronto, Toronto, Ontario, Canada; and [16]Biostatistics Department, Princess Margaret Cancer Centre, University Health Network, Toronto, Ontario, Canada



*Authors share co-first authorship. §Authors share co-senior authorship.

Abbreviations used in this paper: AUC, area under the ROC curve; CCC-GEM, Crohn's and Colitis Canada – Genetic, Environmental, Microbial; CD, Crohn's disease; CI, confidence interval; CRP, C-reactive protein; ELISA, enzyme-linked immunosorbent assay; FCP, fecal calprotectin; FDR, first-degree relative; IBD, inflammatory bowel disease; IgA, immunoglobulin A; IgG, immunoglobulin G; IL, interleukin; IQR, interquartile range; LMR, lactulose-to-mannitol ratio; MEME, Multiple Em for Motif Elicitation; MUSCLE, MUltiple Sequence Comparison by Log-Expectation; NCBI, National Center for Biotechnology Information; NFκB, nuclear factor kappa B; OR, odds ratio; ROC, receiver operating characteristic; UC, ulcerative colitis.

Most current article

1542-3565/$36.00
https://doi.org/10.1016/j.cgh.2025.12.006

BACKGROUND & AIMS: **Elevated antimicrobial antibodies have been reported up to 6 years before diagnosis of Crohn's disease (CD), but the specific antibody response is unclear. Here, we characterized the antimicrobial antibody responses before CD diagnosis in healthy first-degree relatives (FDRs) of patients with CD.**

METHODS: **The CCC-GEM (Crohn's and Colitis Canada – Genetic, Environmental, Microbial) Project nested case-control cohort consisted of FDRs who later developed CD (N = 77), matched 1:4 by age, sex, follow-up duration, and geographical location with healthy FDRs (N = 304). Sera at enrollment were probed for antimicrobial reactivity using a microbiota antigen microarray and a flagellin peptide cytometric bead array. Conditional logistic regression was used to assess association with CD onset, and partial Spearman was used to correlate serologic responses with lactulose-to-mannitol ratio (LMR), C-reactive protein (CRP), and fecal calprotectin (FCP). False discovery rate was controlled using the Benjamini-Hochberg method ($q < 0.05$).**

RESULTS: **Nineteen of 49 IgG antimicrobial antibody responses were significantly associated with the risk of CD; these antibodies were reactive to *Lachnospiraceae* family, particularly *Roseburia*-derived flagellins; 5 antibodies positively correlated with FCP, whereas 3 positively correlated with LMR. These IgG-seroreactive flagellins shared significant amino acid sequence homology, characterized by a conserved "hinge peptide" within D0–D1 domains of the amino-terminus. The cytometric bead array confirmed that elevated IgG seroreactivity to the hinge peptide is associated with future risk of CD independent of LMR and FCP.**

CONCLUSIONS: **Increased antibody response in healthy FDRs towards *Lachnospiraceae* flagellins is associated with future risk of CD. Importantly, pre-CD subjects shared seroreactivity toward a conserved bacterial flagellin epitope, which may represent an early preclinical biomarker of CD.**

Keywords: Anti-Flagellin Antibody; Crohn's Disease; Pre-Disease; Serology.

Crohn's disease (CD) is a chronic relapsing and remitting inflammatory condition affecting the gastrointestinal tract.[1] It is hypothesized that CD results from the loss of tolerance towards commensal bacteria, which triggers a sustained inflammatory response in genetically susceptible individuals.[2] Previous studies reported patients with CD exhibit altered antibody responses towards bacterial antigens.[3–5] These immune responses are also predictive of CD complications,[6] and are minimally affected by host genetics,[7] treatment,[8] surgery[9] and disease activity.[10]

It has been shown that like other immune-mediated diseases, CD is preceded by a preclinical phase.[11–14] Through the Danish National Patient Registry data, changes in serum biochemical parameters are detected up to 8 years before diagnosis.[12] The PREDICTS (Proteomic Evaluation and Discovery in an IBD Cohort of Tri-service Subjects) cohort, which monitored the serum samples of United States military personnel who later developed CD, showed that select antimicrobial antibodies and protein biomarkers were present up to 5 years prior to their diagnoses.[11] Likewise, through the CCC-GEM (Crohn's and Colitis Canada – Genetic, Environmental, Microbial) cohort, we also showed that preclinical subjects exhibit altered antimicrobial response towards a small panel of commensal microbial antigens years before diagnosis,[15] but the specific antigens are still unclear.

Recently, using a microbiota protein antigen array, patients with established CD were found to have increased immunoglobulin G (IgG) against a panel of commensal flagellins.[16] Interestingly, this IgG response was found to target a conserved domain of commensal flagellins known as the "hinge peptide," which may reflect a key epitope in CD pathogenesis.[17] However, it is unclear whether this antibody response is also present in the preclinical antibody repertoire, and if so, whether this has any correlation with the future risk for CD, the gut epithelial barrier dysfunction, or the host inflammatory response. To address this, here we characterized the antimicrobial antibody responses to the commensal flagellin antigens in a pre-CD nested case-control cohort within the GEM cohort using a microbiota antigen protein microarray and validated using a flagellin peptide cytometric bead array.

Materials and Methods

Cohort Recruitment

Healthy first-degree relatives (FDRs) of patients with CD were recruited between the years 2008 and 2017 as part of the prospective GEM study. The GEM study was designed to uncover potential risk factors of CD by comparing the baseline pre-disease characteristics of subjects who eventually developed CD with those who remained healthy.[18] This large multicenter cohort study was approved by the Mount Sinai Hospital Research Ethics

Board (Toronto-Managing Center: REB #07-0322-E) as well as the local recruitment centers involved in the cohort enrollment.

A total of 5176 subjects between the age of 6 and 35 years were recruited and screened using a standardized questionnaire at enrollment.[18] The inclusion and exclusion criteria, as well as demographic and environmental risk questionnaires, were collected as previously described.[15,18] Written informed consents were obtained from all subjects and/or their guardians prior to participating within the GEM study. All subjects were followed every 6 months via telephone call. If a diagnosis of CD was made, the subject is determined as a "pre-CD FDR" and the diagnosis of CD was then confirmed by their treating physician based on clinical, endoscopic, radiographic, and/or histological reports to verify the medical diagnosis. On the other hand, "healthy control" is defined as absence of CD diagnosis at follow-up. The current study was based on the analysis of pre-disease samples (serum, urine, and stool samples) of a nested case-control cohort within the GEM project. Each subject (as of February 2021 at the time of analysis) was matched with up to 4 healthy controls by age, sex, geographic location (using postal codes and country codes), and follow-up duration.

What You Need to Know

Background

Previous studies showed an elevated antimicrobial antibody response years before the diagnosis of Crohn's disease (CD), although the specific antibody responses and their association with CD risk factors are unknown.

Findings

We found that elevated anti-flagellin immunoglobulin G (IgG) responses were associated with future onset of CD, and this response is driven by selective seroreactivity towards the hinge epitope of commensal flagellin.

Implications for patient care

IgG response to the hinge peptide is associated with future onset of CD independent of host inflammation and abnormal intestinal permeability. Early anti-flagellin IgG response may represent a pre-clinical biomarker for future onset of CD.

Microbiota Protein Antigen Microarray

At enrollment, serum samples were collected from subjects via venipuncture and were stored in serum separating tubes.[15] The microbiota antigen microarray assay was processed, performed, and analyzed as previously described.[19] For details, see the Supplementary Materials.

Protein Antigen Identification

The protein sequences were inputted as FASTA sequences and searched against the National Center for Biotechnology Information (NCBI) database (http://blast.ncbi.nlm.nih.gov/Blast.cgi, latest update: October 29, 2024).[19] For details, see the Supplementary Materials.

Sequence Alignment Analysis

The protein sequences were aligned via MUSCLE algorithm (MUltiple Sequence Comparison by Log-Expectation) using software Mega version 11.0.13. MUSCLE-aligned sequences were used to construct phylogenetic tree using Neighbor-Joining Tree algorithm. Sequence alignments were compared and visualized using BioEdit 5.0.9.

Motif Enrichment Analysis

Given that CD immune responses are often directed against common microbial antigens, we analyzed the conserved antigen motifs, which are defined as peptide sequences shared across multiple species using MEME (Multiple Em for Motif Elicitation: https://meme-suite.org/meme/tools/meme).[20] Candidate motifs were identified by aligning peptide sequences from the microbiota protein microarray and selected based on frequency of sequence recurrences across the peptides. The motif "enrichment" or E-value measures of the likelihood that the repeated sequence occurrence is enriched purely by artefact. Coverage percentage indicates the percentage of matched amino acids to the motif of interest.

Flagellin Peptide Cytometric Bead Assay

Detailed methods can be found in the study previously described, and details can be found in the Supplementary Materials.[17]

C-Reactive Protein Measurement

The baseline C-reactive protein (CRP) levels from all subjects were measured using Prometheus assays as per manufacturer's protocol (Prometheus Laboratory) as outlined in previous studies.[15] The upper limit of detection was 203.8 mg/L, and the lower limit of detection was set at 0.1 mg/L.

Fecal Calprotectin Measurement

The fecal calprotectin (FCP) levels were measured using methods as previously described.[18]

Gut Epithelial Barrier Function

The intestinal epithelial barrier function was measured in urine using ratio of fractional excretions of

lactulose-to-mannitol ratio (LMR) as previously described.[15]

Statistical Analysis

Baseline characteristics were summarized as counts (percentages) for categorical variables and as medians (interquartile ranges [IQRs]) for continuous variables. To evaluate the association between antibody responses and future risk of CD, we performed conditional logistic regression (clogit) models with strata defined by the matched nested case-control sets. Each antibody (log-transformed and standardized) was modeled separately, adjusting for the matching factors through stratification. Odds ratios (ORs) with 95% confidence intervals (CIs) and *P* values were reported. To assess discriminative ability, linear predictors from each fitted model were used to generate receiver operating characteristic (ROC) curves using the pROC package in R. We calculated the area under the ROC curve (AUC) with 95% CIs based on 2000 bootstrap resamples. For clinical interpretability, we further derived: (1) the optimal threshold based on Youden's index with corresponding sensitivity and specificity; and (2) sensitivities at approximate fixed specificities of 80% and 90% (nearest possible values from ROC). To account for multiple testing across antibodies, false discovery rate was controlled using the Benjamini–Hochberg procedure, with $q < 0.05$ considered statistically significant.

The correlations between antimicrobial antibody responses and known CD-related biomarkers were calculated using partial Spearman correlations controlling for nested case control clustering including age, sex, geographic location (using postal codes and country codes), and follow-up duration. Missing data were handled by complete-case analysis: participants with missing values for CRP, FCP, and LMR were excluded from the respective analyses. All statistical analyses were performed using R-studio (R version 4.3.1 [2023-06-16 Universal C runtime]).

Results

The Baseline Characteristics of GEM Nested Case-Control Cohort

The nested case-control cohort consists of 77 pre-CD subjects (who later developed CD) and 304 healthy control (FDRs who remained healthy) (Table 1). The median age of recruitment was 15.0 years (IQR, 12.0–21.0 years), and the sex distribution was 58.4% female. Interestingly, 80.5% of the subjects were siblings, whereas only 19.5% were an offspring to a patient diagnosed with CD—suggesting predominance of sibling pairs that may reflect shared early-life exposures. Among the pre-CD group, the median duration from the collection of baseline sample in relation to timing of CD diagnosis was 2.46 years (IQR, 1.25–4.66 years). Among the pre-CD subjects, 80.5% of the subjects were recruited in Canada, 13.0% in Israel, and 6.5% in the United States (Table 1). Compared with healthy FDRs, pre-CD subjects had more subjects with CRP >10 mg/L (26.3% vs 3.6%), FCP >100 μg/g (66.7% vs 19.2%), and LMR >0.025 (39.7% vs 22.9%) (Table 1).

Pre-Disease Antimicrobial Antibody Responses Are Associated With the Future Onset of CD

The baseline blood samples of pre-CD subjects and healthy controls were tested for antimicrobial responses using an established microbiota antigen microarray (Figure 1*A*).[17,19] Compared with healthy controls, pre-CD FDRs had 28 elevated antimicrobial antibody responses that were associated with future risk of CD (Figure 1*B*; peptide annotations in Table 2). Among these, 19 were IgG and 9 were immunoglobulin A (IgA) responses, with the top 8 signals all IgGs targeting commensal flagellins: anti-3-1-157Fla (OR, 2.37; 95% CI, 1.56–3.61; false discovery rate = 0.0028), anti-RinuF5 (OR, 2.19; 95% CI, 1.50–3.22; false discovery rate = 0.003), anti-CBir11 (OR, 2.18; 95% CI, 1.47–3.23; false discovery rate = 0.004), anti-MDR254 (OR, 2.00; 95% CI, 1.39–2.87; false discovery rate = 0.004), anti-RinuF2 (OR, 2.39; 95% CI, 1.52–3.76; false discovery rate = 0.004), anti-A4Fla4 (OR, 2.29; 95% CI, 1.47–3.59; false discovery rate = 0.004), anti-FlaX (OR, 2.04; 95% CI, 1.39–2.99; false discovery rate = 0.004), and anti-RhomF1 (OR, 1.70; 95% CI, 1.27–2.27; false discovery rate = 0.005) (Supplementary Figure 1; Supplementary Table 1). Four of the antigens elicited both IgG and IgA response: anti-RintF1 (OR, 1.49; 95% CI, 1.17–1.90; false discovery rate = 0.007), anti-RintF2 (OR, 1.37; 95% CI, 1.08–1.73; false discovery rate = 0.039), anti-A4Fla4 (OR, 1.42; 95% CI, 1.09–1.84; false discovery rate = 0.039), and anti-RhomF1 (OR, 1.38; 95% CI, 1.07–1.78; false discovery rate = 0.046) (Figure 1*C*; Supplementary Table 1). The discriminative ability of individual antibodies was modest, with AUCs ranging from 0.60 to 0.69 (Supplementary Table 1). At the optimal Youden's cutoff, anti-RinuF5 IgG achieved balanced sensitivity/specificity of 66%/66%, whereas anti-3-1-57Fla IgG showed high sensitivity (82%) but low specificity (48%) (Supplementary Table 1).

Pre-Disease IgG Response to Lachnospiraceae-Derived Flagellins Is Associated With Host Systemic and Intestinal Inflammation and Gut Barrier Dysfunction

To characterize the antigens, the peptide sequences of all 49 proteins were compared with human microbiome signatures using NCBI Gene Bank and the metagenome gene catalog.[19] Surprisingly, all IgG responses associated with CD onset were directed against closely

Table 1. Demographics of the Nested Case-Control Cohort

	Healthy FDRs (n = 304)	Pre-CD FDRs (n = 77)	P^a
Age at recruitment, *y*	15.0 (11.0–22.0)	15.0 (12.0–21.0)	.600
Sex			1.000
Female	177 (58.2)	45 (58.4)	
Male	127 (41.8)	32 (41.6)	
Relation to proband			.040
Sibling	204 (67.1)	62 (80.5)	
Offspring	100 (32.9)	15 (19.5)	
Follow-up duration, *y*	2.46 (1.25–4.66)	2.46 (1.25–4.66)	1.000
Country of recruitment			1.000
Canada	259 (85.2)	62 (80.5)	
Israel	37 (12.2)	10 (13.0)	
United States	8 (2.6)	5 (6.5)	
CRP			6.71e−7
CRP >10 mg/L	11 (3.6)	20 (26.3)	
FCP			7.41e−10
FCP >100 μg/g	54 (19.2)	42 (66.7)	
LMR			.00602
LMR >0.025	66 (22.9)	25 (39.7)	

NOTE. Categorical variables were presented as number (%); continuous variables were presented as median (interquartile range).
CD, Crohn's disease; CRP, C-reactive protein; FCP, fecal calprotectin; FDR, first degree relatives; LMR, lactulose-to-mannitol ratio.
[a]Univariate conditional logistic regression accounting for nested groups comparing healthy and pre-CD FDRs.

clustered members of the *Lachnospiraceae* family (Supplementary Figure 2*A*). Among these members, we identified 5 conserved motifs shared across the CD-associated antigens, with the most highly conserved sequence being a 34-amino-acid "hinge peptide" sequence localized within the D0-D1 domains of *Lachnospiraceae* flagellin (Figure 2*A*; Supplementary Figure 2*B*).

To determine the association of this "hinge peptide" with known risk factors for CD, we correlated the antibody responses with CRP, FCP, and LMR.[21] To do this, we first ranked all 49 commensal antigens based on the enrichment of Motif 1 ("hinge peptide") within the protein sequence (Figure 2*B*). When this was aligned with future risks for CD, pre-CD subjects showed elevated IgG response only to peptides highly enriched in the "hinge peptide" (Figure 2*C*). Among these, 5 of the 21 IgGs positively correlated with FCP (RintF1 $\rho = 0.15$; AgF2 $\rho = 0.14$; MDR254 $\rho = 0.14$; RinuF5 $\rho = 0.15$; and anti-CBir66 $\rho = 0.15$); and only 3 of the 21 anti-flagellin IgGs (RinuF5 $\rho = 0.17$; 3-1-57Fla $\rho = 0.15$; and MDR254 $\rho = 0.16$) positively correlated with LMR (Figure 2*D*). No significant correlations were seen between serum CRP and the anti-flagellin IgG responses.

On the other hand, the IgA responses towards RintF2, RintF1, RhomF1, A4Fla4, CBir5, riB2, yidx, riB20 and CBir45 significantly correlated with future onset of CD, despite having low enrichment of the hinge peptide in the protein sequence (Supplementary Figure 3*A–B*). Furthermore, none of the IgA responses correlated with FCP and LMR, but IgA responses to RintF1, RhomF1, yidx, and RiB20 correlated with CRP (Supplementary Figure 3*C*). Taken together, these findings indicate that the pre-disease anti-flagellin IgG responses are shared by a common consensus motif in the "hinge peptide," and this response is correlated with the future risk of developing CD.

Pre-Disease Anti-Flagellin Antibody Responses Are Driven by Seroreactivity to a Dominant Epitope on Lachnospiraceae *Flagellin Peptide*

To validate our findings, we applied an established flagellin peptide cytometric bead array to measure IgG reactivity to 8 conserved *Lachnospiraceae* flagellin peptides (Figure 3*A*).[17] Indeed, IgG concentrations specific to the hinge peptide (D0-1N p25-59) and its sub-epitopes (D1N p41-59 and D0N p25-44) were significantly elevated in pre-CD subjects, but not towards the "non-hinge" regions (Figure 3*B*), suggesting a highly specific response to the "hinge peptide." In addition, IgG concentrations of the hinge peptide (D0-1N p25-59) was correlated with FCP ($\rho = 0.15$; $P = .004$) and LMR ($\rho = 0.18$; $P = .0009$), but not with CRP ($\rho = -0.05$; $P = .36$) (Figure 3*C–D*). In fact, elevated IgG responses to both hinge sub-epitopes (D1N p41-59 and D0N p25-44) also correlated with FCP, and D1N p41-59 correlated with LMR (Figure 3*C–D*).

Given there may be potential associations between anti-hinge peptide IgG response vs CD risk and CD risk factors, next we performed conditional logistic regression adjusting for CRP, LMR, and FCP, as well as combination of all 3 risk factors. Interestingly, when adjusted for CRP, FCP, or LMR individually or together (Supplementary Figure 4*A–D*), IgG response to the hinge peptide (D0-1N p25-59) remained significantly associated with future onset of CD. Taken together, this demonstrates that the IgG response to the hinge peptide is associated with future onset of CD independent of baseline changes in CRP, FCP, or intestinal permeability.

Discussion

Patients with inflammatory bowel disease (IBD) are known to have distinct antimicrobial antibody responses compared with the general population.[3,4,16,22–24] Recent evidence shows that select antibody responses are already present in the pre-disease phase of CD,[5,11,13–15] suggesting that mucosal and systemic immune activations may be occurring years well in advance of CD onset. However, characterization of the pre-disease antigen and antibody

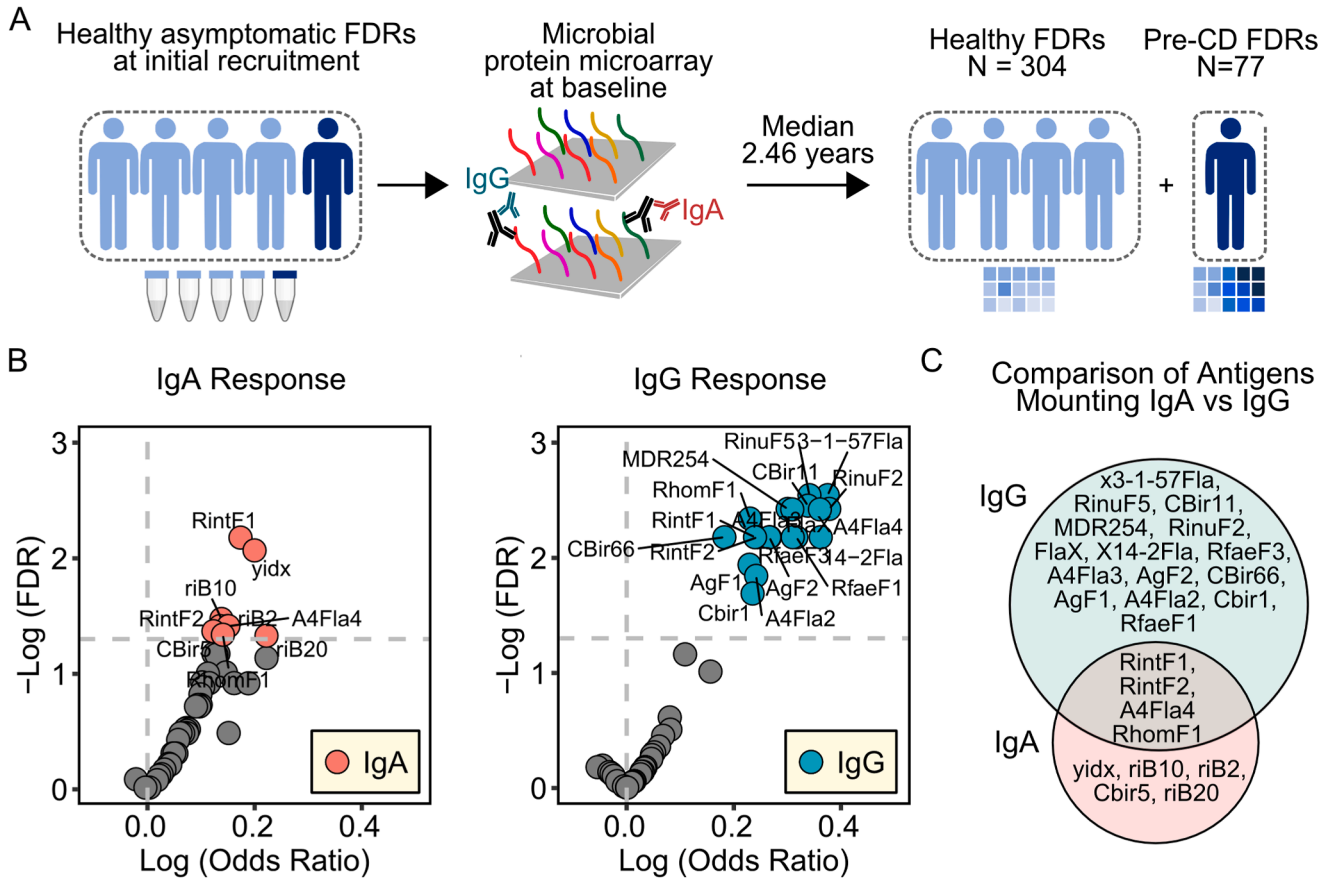


Figure 1. Study schematic of the nested cohort and the antimicrobial antibody response in pre-CD FDRs vs healthy FDRs. (*A*) Schematic diagram illustrating the GEM nested case-control cohort (*dark blue* denotes baseline samples collected from pre-CD FDRs; *light blue* denotes baseline samples collected from healthy FDRs). Patient sera were processed for microbial protein microarray at baseline and were followed for a median of 2.46 years for potential diagnosis of CD. (*B*) Volcano plots demonstrating a total of 28 antimicrobial antibody responses that were significantly increased in pre-CD FDRs (n = 77) compared with healthy controls (n = 304). Data presented as scattered volcano plots showing IgA in *red*, and IgG in *green*, and analyzed statistically using conditional logistic regression with false discovery rate using Benjamini-Hochberg method with $q < 0.05$ defined as statistically significant. *Vertical dotted line* indicates an OR of 0, and *horizontal dotted line* indicates q of 0.05. (*C*) Venn diagram comparing the IgG- and IgA-seroreactive antigens that were significantly associated with future onset of CD. *Top circle* indicates antigens mounting IgGs, and *bottom circle* indicates antigens mounting IgAs in pre-CD FDRs.

repertoires have been challenging due to the difficulties in following pre-CD cohorts and the limited number of antigens profiled in conventional assays.[11,15] Here, we applied a novel microbiota antigen microarray and flagellin peptide cytometric bead array to the well-established GEM pre-CD cohort.[15,18,21] We now extend on the previous work by others[11,13,17,25] and demonstrate the preclinical anti-flagellin antibody responses target a conserved hinge peptide within the D0-1 domains of *Lachnospiraceae* flagellin. This response occurs independent of baseline CRP, FCP, and LMR. This suggests that early anti-flagellin antibody responses selectively target a conserved domain of commensal flagellins, which may serve as an early biomarker to predict future CD onset.

The anti-flagellin antibody patterns are known to be associated with penetrating and stricturing complications as well as higher surgical rates in adults and children with CD.[3,26,27] Previously, the PREDICTS study showed that individuals with CD displayed seropositivity in ASCA and ANCA antibodies up to 8 years prior to diagnosis.[7] At 5 years prior to onset, this cohort expressed increased seroreactivity to commensal flagellins CBir1, as well as the *Lachnospiraceae* flagellins Fla-X and A4Fla2.[11] Similarly, we found that early antimicrobial antibodies have been associated with future onset of CD.[15] This study extends on those prior findings by showing a specific target epitope within the conserved hinge peptide (D0-1N p25-59) of *Lachnospiraceae* flagellins. This is not surprising given that *Lachnospiraceae* is an abundant family of anerobic butyrate-producers within the Bacillota phylum,[28] and have similar pattern of colonization in the small intestine and proximal colon seen in CD ileocolitis.[29,30] *Lachnospiraceae*-derived flagellins are also strongly immunogenic, likely owing to the TLR-5 epitope expressed in its D1 domain[31] and optimization of B-cell responses.[32] Indeed, IgG response to *Lachnospiraceae* flagellins are observed in patients with CD but not ulcerative colitis (UC).[33]

Table 2. Peptide Identity of Seroreactive Antigens in Pre-CD FDRs

Protein	Description of antigen	Query cover	E-value	Percent identity
RinuF5	*R. inulinivorans* Fla5	100%	0	100%
RinuF2	*R. inulinivorans* Fla2	100%	0	100%
RintF2	*R. intestinalis* Fla2	100%	0	100%
RintF1	*R. intestinalis* Fla1	100%	0	100%
RhomF1	*R. hominis* Fla	100%	0	100%
RfaeF3	*R. faecis* Fla3	100%	0	100%
RfaeF1	*R. faecis* Fla1	100%	0	100%
MDR254	MDR254 Fla	97%	6e−91	63%
riB10	Lachnospiraceae SSB	92%	7e−133	87%
AgF2	Lachnospiraceae Fla2	100%	0	100%
AgF1	Lachnospiraceae Fla1	100%	0	100%
3-1-57Fla	Lachnospiraceae Fla	100%	0	83%
CBir11	Lachnospiraceae Fla	98%	2e−145	82%
FlaX	Lachnospiraceae Fla	99%	0	97%
Cbir1	Lachnospiraceae Fla	99%	0	81%
CBir66	Lachnospiraceae CBir66 Fla	80%	8e−106	100%
A4Fla4	Lachnospiraceae A4 Fla4	100%	0	100%
A4Fla3	Lachnospiraceae A4 Fla3	100%	0	100%
A4Fla2	Lachnospiraceae A4 Fla2	100%	0	100%
14-2Fla	Lachnospiraceae 14-2 Fla	100%	0	100%
CBir5	*H. spp* TIpB	100%	0	100%
riB20	Firmicutes uncharacterized	76%	2e−33	43%
riB2	*F. plautii* Fla	76%	4e−33	62%
yidx	*E. coli* FliC	100%	4e−161	100%

Fla, flagellin; FlicC, flagellar filament structural protein; SSB, single-standed binding protein.

Flagellins share a conserved amino and carboxy D0-D1 domains separated by a hypervariable region.[31] The epitope we identified also confirms the recent study by Zhao et al, which showed a strong IgG response to the hinge peptide in patients with established CD.[17] Furthermore, using next-generation epitope profiling techniques,[25,34] several studies have also performed in-depth characterization of the antigen repertoire in IBD cohorts. For instance, Bourgonje et al showed that the dominant bacterial antigens in patients with established CD as well as chronic fatigue syndrome were localized to the conserved regions of *Lachnospiraceae* flagellins.[25,34] These responses against bacterial flagellins were further associated with ileal disease and fibrostenotic disease types.[34] The findings here support those prior work and indicate that this antibody response to the commensal flagellin epitope may occur before the onset of CD in healthy FDRs. When adjusted for CRP and LMR, this select response was independently associated with the future onset of CD, suggesting that this early anti-flagellin response may serve as a preclinical biomarker of future disease prior to mucosal inflammation. Despite these observations, validation of our findings in other preclinical cohorts such as PREDICTS and the Israeli Army cohort are warranted. Similarly, it would be important to compare anti-flagellin responses with existing clinical serological markers in future studies.

Although our findings may suggest a role of preclinical anti-flagellin antibody responses in the development of IBD, it raises several important unanswered questions on the mechanisms by which this may play a role in IBD. Previously, antibacterial IgG antibodies have been shown to target bacteria taxa that translocate across the mucosal barrier function.[33] This may lead to adaptive immune activation and therefore loss of tolerance to commensals.[33] For instance, commensal flagellin IgG responses have been shown to correlate with altered T cell and B cell responses.[4,16,24,35,36] In patients with established CD,

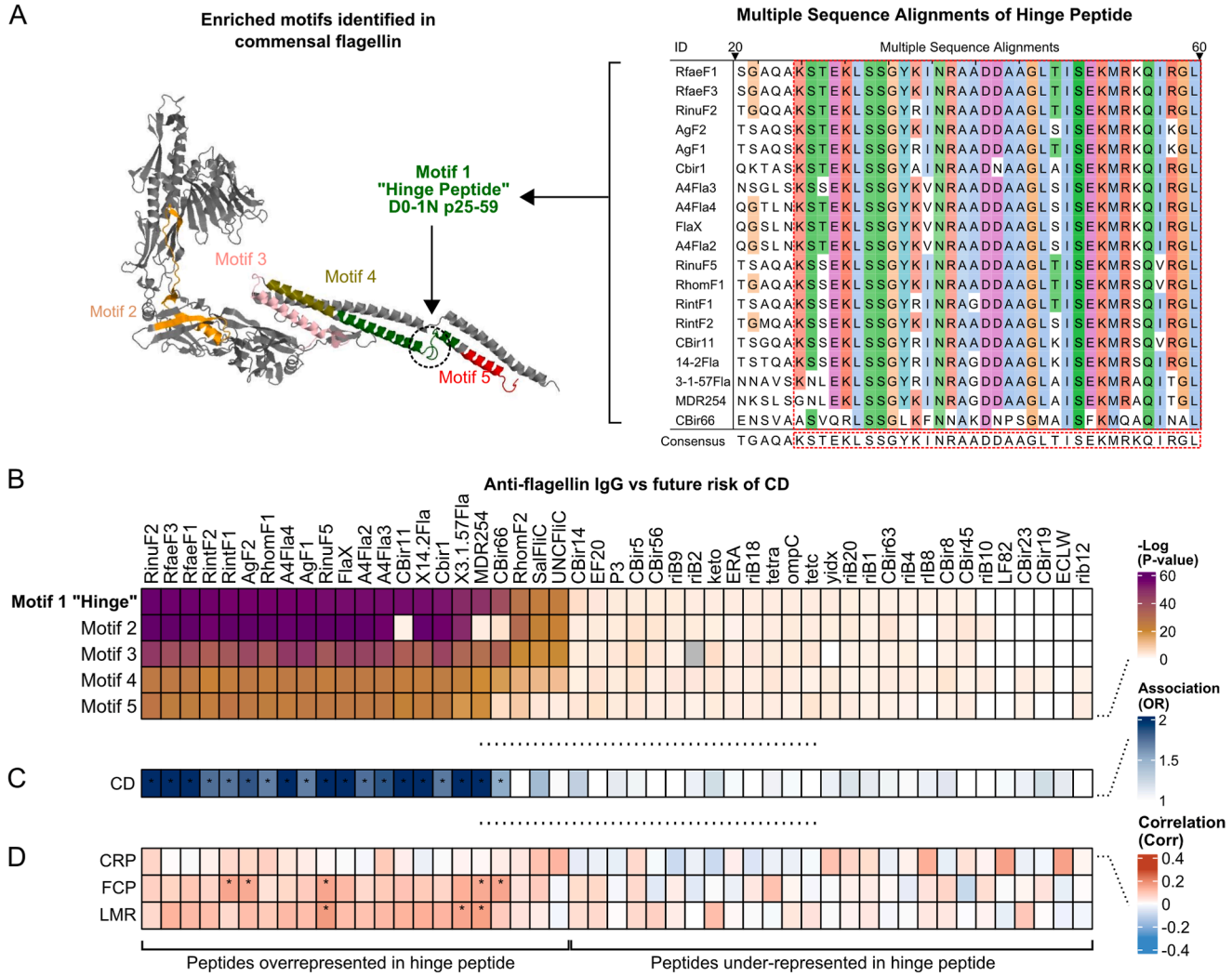


Figure 2. Anti-flagellin IgG responses target a conserved hinge peptide and is correlated with independent risk factors of CD: FCP, CRP, and LMR. (*A*) Five conserved motifs annotated using *Lachnospiraceae*-derived A4Fla2 flagellin (PyMOL version 3.0.4) and multiple sequence alignment of the most highly enriched "hinge peptide" in the CD-associated antibody responses (MEGA version 11.0.13, comparison by log-expectation [MUSCLE] algorithm). (*B*) Heatmap of the commensal antigens ranked by enrichment of "hinge peptide" expressed using multiple sequence enrichment *P* values. (*C*) IgG fluorescence intensities were compared with risk of future CD onset. Data were expressed as ORs, analyzed using conditional logistic regression with false discovery rate using Benjamini-Hochberg method with $q < 0.05$ as statistical significance. (*D*) IgG fluorescence intensities were correlated with risk factors of CD: CRP, FCP, and LMR. Data were expressed as partial Spearman correlation ρ coefficients, with a *P* value <.05 deemed as statistically significant.

elevated anti-flagellin IgGs were associated with flagellin-specific T cell populations.[16] T cell transfer of flagellin (CBir1)-specific $CD4^+$ T cells has been shown to induce T-cell-mediated colitis in mice.[4] On the other hand, it has also been shown in vivo that augmenting mucosal anti-flagellin IgA response by vaccinating purified flagellin protects spontaneous colitis and alters the gut microbiome in interleukin (IL)-10 deficiency mice.[36] Interestingly, this study found that this response appears to be dependent on B-cell responses. Taking this to consideration, although beyond the scope of our study, further functional mechanistic studies to study the peripheral blood mononuclear cells derived from pre-CD FDRs, and careful interrogation of flagellin-specific B cell and T cell responses are much needed. Such information would also help explain the apparent differences in anti-flagellin IgG vs IgA responses as noted in our findings, as this could represent difference on innate vs adaptive responses to commensal flagellins. Lastly, in a recent study by Gaifem et al, it has been shown that preclinical ASCA IgGs form immune complexes to activate innate immune cells, which enabled downstream nuclear factor kappa B (NFκB) activation.[37] This raises the possibility that select antimicrobial IgGs could impact innate immune responses, and further characterization of macrophage and dendritic cell populations can be explored in future studies.

The strength of our study is that all clinical samples were derived from a well-characterized multinational at-risk cohort of healthy FDRs, which allows for the analysis of biomarkers before disease onset.[15,18,21] Although traditional serological profiling techniques use

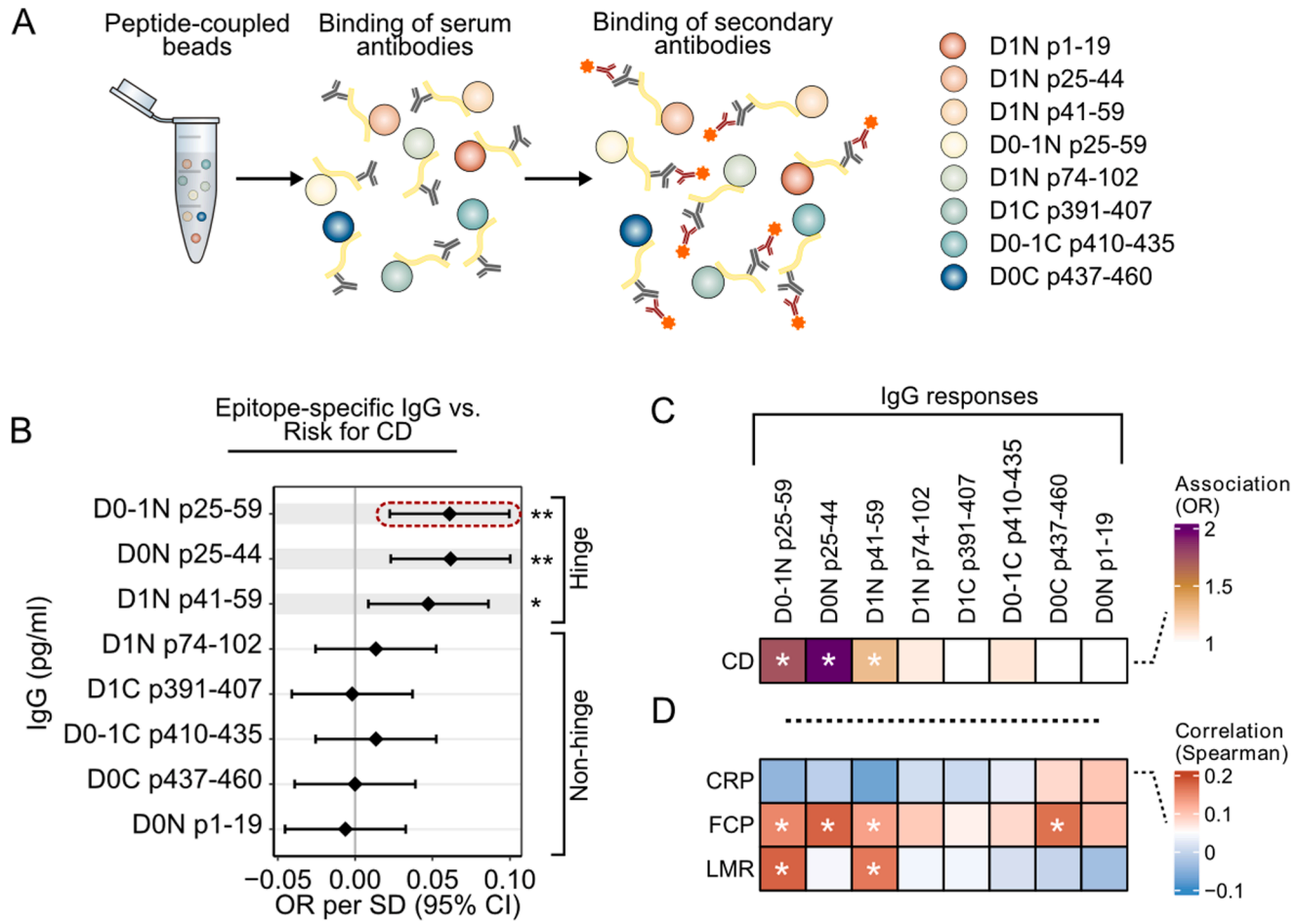


Figure 3. Seroreactivity to the hinge peptide of D0-1 domains of Lachnospiraceae flagellin correlates with future risk of CD. (*A*) Schematic diagram of the workflow for the flagellin peptide cytometric bead array and the associated conserved motifs. (*B*) Forest plot comparing epitope-specific IgG concentrations to the risk of developing CD. Data was analyzed using conditional logistic regression with false discovery rate using Benjamini-Hochberg. * denotes $q < 0.05$ and ** denotes $q < 0.005$. (*C*) Heatmap showing the correlation between epitope-specific IgG concentrations (pg/mL) and ORs of developing CD. Data were expressed as ORs and analyzed using conditional logistic regression with false discovery rate using Benjamini-Hochberg method. * denotes $q < 0.05$. (*D*) Epitope-specific IgG concentration were correlated with CRP, FCP, and LMR. Data were expressed as partial Spearman correlation ρ coefficients and * denotes *P* value <.05.

enzyme-linked immunosorbent assay (ELISA)-based commercial kits, they often suffer from selection bias by containing only few antigens such as the Prometheus panel we described previously.[16,19] Here, we targeted more antigens in using a systematic microbiota antigen microarray and validated using a flagellin peptide cytometric beads array. In addition, by incorporating clinical datasets including subjects' FCP, serum CRP, and gut permeability, we were able to correlate the antimicrobial antibody responses with known risk factors of CD. Despite these strengths, any overinterpretation of the correlation to causation in the pre-CD anti-flagellin antibody response warrants caution given the lack of mechanistic experiments in our study design. Secondly, we recognize that the microbiota antigen microarray was performed at a single time-point, and is limited in the number of antigens detected,[38] which is surpassed by high-throughput approaches such as phage display immunoprecipitation sequencing and IgA sequencing methodologies. Similarly, it is also possible that the assays may also not fully capture all potential epitopes in the pre-CD cohort. Furthermore, although the GEM cohort enables the study of preclinical risk factors, the median follow-up of 2.46 years may introduce bias, as healthy FDRs can still potentially develop CD at longer follow-up. Potential confounders such as patient disease phenotypes at diagnosis, diet, and microbiome data are not included as part of our analysis. In addition, our cohort demographic is also mostly comprised of pediatric and young adults with a median age of 15 years, which therefore limits the generalizability to an older-onset IBD population. Therefore, although our observations are hypothesis-generating, further studies to corroborate our findings in other pre-clinical disease cohorts are further warranted.

Conclusion

In summary, our findings demonstrate that anti-flagellin antibody responses are associated with future onset of CD. In addition, these pre-CD IgG responses

appear to be driven by a conserved domain within the hinge peptide of D0-1 domain of *Lachnospiraceae* flagellin and is independent of known CD risk factors such as gut and systemic inflammation and intestinal permeability. The clinical implication of these findings open new perspectives in disease prevention and treatment. It raises the possibility for anti-flagellin biomarkers and the potential designing of flagellin-directed vaccines in selected high-risk individuals. Further validation and mechanistic studies are warranted.

Supplementary Material

Note: To access the supplementary material accompanying this article, visit the online version of *Clinical Gastroenterology and Hepatology* at www.cghjournal.org, and at https://doi.org/10.1016/j.cgh.2025.12.006.

References

1. Dolinger M, Torres J, Vermeire S. Crohn's disease. Lancet 2024;403:1177–1191.
2. Rudbaek JJ, Agrawal M, Torres J, et al. Deciphering the different phases of preclinical inflammatory bowel disease. Nat Rev Gastroenterol Hepatol 2024;21:86–100.
3. Targan SR, Landers CJ, Yang H, et al. Antibodies to CBir1 flagellin define a unique response that is associated independently with complicated Crohn's disease. Gastroenterology 2005;128:2020–2028.
4. Lodes MJ, Cong Y, Elson CO, et al. Bacterial flagellin is a dominant antigen in Crohn disease. J Clin Invest 2004;113:1296–1306.
5. Mow WS, Vasiliauskas EA, Lin Y-C, et al. Association of antibody responses to microbial antigens and complications of small bowel Crohn's disease. Gastroenterology 2004;126:414–424.
6. Dubinsky MC, Kugathasan S, Mei L, et al. Increased immune reactivity predicts aggressive complicating Crohn's disease in children. Clin Gastroenterol Hepatol 2008;6:1105–1111.
7. Walker LJ, Aldhous MC, Drummond HE, et al. Anti-*Saccharomyces cerevisiae* antibodies (ASCA) in Crohn's disease are associated with disease severity but not NOD2/CARD15 mutations. Clin Exp Immunol 2004;135:490–496.
8. Teml A, Kratzer V, Schneider B, et al. Anti-*Saccharomyces cerevisiae* antibodies: a stable marker for Crohn's disease during steroid and 5-aminosalicylic acid treatment. Am J Gastroenterol 2003;98:2226–2231.
9. Eser A, Papay P, Primas C, et al. The impact of intestinal resection on serum levels of anti-*Saccharomyces cerevisiae* antibodies (ASCA) in patients with Crohn's disease. Aliment Pharmacol Ther 2012;35:292–299.
10. Landers CJ, Cohavy O, Misra R, et al. Selected loss of tolerance evidenced by Crohn's disease–associated immune responses to auto- and microbial antigens. Gastroenterology 2002;123:689–699.
11. Torres J, Petralia F, Sato T, et al. Serum biomarkers identify patients who will develop inflammatory bowel diseases up to 5 years before diagnosis. Gastroenterology 2020;159:96–104.
12. Vestergaard MV, Allin KH, Poulsen GJ, et al. Characterizing the pre-clinical phase of inflammatory bowel disease. Cell Rep Med 2023;4:101263.
13. Van Schaik FDM, Oldenburg B, Hart AR, et al. Serological markers predict inflammatory bowel disease years before the diagnosis. Gut 2013;62:683–688.
14. Choung RS, Princen F, Stockfisch TP, et al. Serologic microbial associated markers can predict Crohn's disease behaviour years before disease diagnosis. Aliment Pharmacol Ther 2016;43:1300–1310.
15. Lee SH, Turpin W, Espin-Garcia O, et al. Anti-microbial antibody response is associated with future onset of Crohn's disease independent of biomarkers of altered gut barrier function, subclinical inflammation, and genetic risk. Gastroenterology 2021;161:1540–1551.
16. Alexander KL, Zhao Q, Reif M, et al. Human microbiota flagellins drive adaptive immune responses in Crohn's disease. Gastroenterology 2021;161:522–535.e6.
17. Zhao Q, Duck LW, Killian JT, et al. Crohn's patients and healthy infants share immunodominant B cell response to commensal flagellin peptide epitopes. Gastroenterology 2024;167:1415–1428.
18. Turpin W, Espin-Garcia O, Xu W, et al. Association of host genome with intestinal microbial composition in a large healthy cohort. Nat Genet 2016;48:1413–1417.
19. Christmann BS, Abrahamsson TR, Bernstein CN, et al. Human seroreactivity to gut microbiota antigens. J Allergy Clin Immunol 2015;136:1378–1386.e5.
20. Bailey TL, Johnson J, Grant CE, et al. The MEME suite. Nucleic Acids Res 2015;43:W39.
21. Turpin W, Lee S-H, Raygoza Garay JA, et al. Increased intestinal permeability is associated with later development of Crohn's disease. Gastroenterology 2020;159:2092–2100.e5.
22. Bourgonje AR, Roo-Brand G, Lisotto P, et al. Patients with inflammatory bowel disease show IgG immune responses towards specific intestinal bacterial genera. Front Immunol 2022;13:842911.
23. Cook L, Lisko DJ, Wong MQ, et al. Analysis of flagellin-specific adaptive immunity reveals links to dysbiosis in patients with inflammatory bowel disease. Cell Mol Gastroenterol Hepatol 2020;9:485–506.
24. Shen C, Landers CJ, Derkowski C, et al. Enhanced CBir1-specific innate and adaptive immune responses in Crohn's disease. Inflamm Bowel Dis 2008;14:1641–1651.
25. Bourgonje A, Hörstke N, Fehringer M, et al. Systemic antibody responses against gut microbiota flagellins implicate shared and divergent immune reactivity in Crohn's disease and chronic fatigue syndrome. Microbiome 2024;12:141.
26. 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.
27. Kugathasan S, Denson LA, Walters TD, et al. Prediction of complicated disease course for children newly diagnosed with Crohn's disease: a multicentre inception cohort study. Lancet 2017;389:1710–1718.
28. Maki-Yonekura S, Yonekura K, Namba K. Conformational change of flagellin for polymorphic supercoiling of the flagellar filament. Nat Struct Mol Biol 2010;17:417–422.
29. Patterson AM, Mulder IE, Travis AJ, et al. Human gut symbiont *Roseburia hominis* promotes and regulates innate immunity. Front Immunol 2017;8:1166.
30. Nava GM, Friedrichsen HJ, Stappenbeck TS. Spatial organization of intestinal microbiota in the mouse ascending colon. ISME J 2011;5:627–638.

31. Clasen SJ, Bell MEW, Borbón A, et al. Silent recognition of flagellins from human gut commensal bacteria by toll-like receptor 5. Sci Immunol 2023;8:eabq7001.
32. Oh JZ, Ravindran R, Chassaing B, et al. TLR5-Mediated sensing of gut microbiota is necessary for antibody responses to seasonal influenza vaccination. Immunity 2014;41:478–492.
33. Vujkovic-Cvijin I, Welles HC, Ha CWY, et al. The systemic anti-microbiota IgG repertoire can identify gut bacteria that translocate across gut barrier surfaces. Sci Transl Med 2022;14:eabl3927.
34. Bourgonje AR, Andreu-Sánchez S, Vogl T, et al. Phage-display immunoprecipitation sequencing of the antibody epitope repertoire in inflammatory bowel disease reveals distinct antibody signatures. Immunity 2023;56:1393–1409.e6.
35. Feng T, Wang L, Schoeb TR, et al. Microbiota innate stimulation is a prerequisite for T cell spontaneous proliferation and induction of experimental colitis. J Exp Med 2010;207:1321–1332.
36. Tran HQ, Ley RE, Gewirtz AT, et al. Flagellin-elicited adaptive immunity suppresses flagellated microbiota and vaccinates against chronic inflammatory diseases. Nat Commun 2019;10:5650.
37. Gaifem J, Rodrigues CS, Petralia F, et al. A unique serum IgG glycosylation signature predicts development of Crohn's disease and is associated with pathogenic antibodies to mannose glycan. Nat Immunol 2024;25:1692–1703.
38. Bourgonje AR, Vogl T, Segal E, et al. Antibody signatures in inflammatory bowel disease: current developments and future applications. Trends Mol Med 2022;28:693–705.

Correspondence
Address correspondence to: Kenneth Croitoru, MDCM, Division of Gastroenterology, Zane Cohen Centre for Digestive Diseases, Temerty School of Medicine Department of Medicine, University of Toronto, Mount Sinai Hospital, 600 University Avenue, Room 437, Toronto, Ontario M5G 1X5, Canada. e-mail: ken.croitoru@sinaihealth.ca; or Sun-Ho Lee, MD, PhD, Division of Gastroenterology, Zane Cohen Centre for Digestive Diseases, Temerty School of Medicine Department of Medicine, University of Toronto, Mount Sinai Hospital, 600 University Avenue, Room 455, Toronto, Ontario M5G 1X5, Canada. e-mail: Sun-ho.lee@sinaihealth.ca.

Acknowledgments
The authors thank the members of the CCC GEM Global Project Office.

CRediT Authorship Contributions
Richard Y. Wu (Conceptualization: Equal; Formal analysis: Equal; Writing – original draft: Lead)
Mingyue Xue (Formal analysis: Lead; Writing – review & editing: Equal)
Qing Zhao (Formal analysis: Supporting; Writing – review & editing: Supporting)
Sean Jeong (Formal analysis: Supporting; Writing – review & editing: Supporting)
Anne M. Griffiths (Resources: Supporting; Writing – review & editing: Supporting)
Levinus A. Dieleman (Resources: Supporting; Writing – review & editing: Supporting)
A. Hillary Steinhart (Resources: Supporting; Writing – review & editing: Supporting)
Guy Aumais (Resources: Supporting; Writing – review & editing: Supporting)
Brian Bressler (Writing – review & editing: Supporting)
Remo Panaccione (Resources: Supporting; Writing – review & editing: Supporting)
Colette Deslandres (Resources: Supporting; Writing – review & editing: Supporting)
David R. Mack (Resources: Supporting; Writing – review & editing: Supporting)
Charles N. Bernstein (Resources: Supporting; Writing – review & editing: Supporting)
John K Marshall (Resources: Supporting; Writing – review & editing: Supporting)
Dan Turner (Resources: Supporting; Writing – review & editing: Supporting)
Wei Xu (Formal analysis: Supporting; Writing – review & editing: Supporting)
Lennard W. Duck (Formal analysis: Supporting; Methodology: Supporting)
Charles O. Elson (Formal analysis: Supporting; Methodology: Equal; Writing – review & editing: Supporting)
Williams Turpin (Conceptualization: Supporting; Formal analysis: Supporting; Supervision: Supporting; Writing – review & editing: Equal)
Sun-Ho Lee (Conceptualization: Lead; Formal analysis: Lead; Methodology: Equal; Supervision: Equal; Writing – review & editing: Lead)
Kenneth Croitoru, MDCM (Conceptualization: Lead; Formal analysis: Supporting; Funding acquisition: Lead; Methodology: Lead; Supervision: Lead; Writing – review & editing: Supporting)

Conflicts of interest
These authors disclose the following: Charles O. Elson and the University of Alabama at Birmingham hold a patent on *Lachnospiraceae* A4 Fla2, which has been licensed for clinical use by Prometheus Laboratories. Charles O. Elson, Qing Zhao, Lennard W. Duck, and the University of Alabama at Birmingham have filed a patent on the flagellin peptide cytometric bead array. Charles O. Elson is Founder, Board member, and Chief Scientific Officer of ImmPrev Bio, which is developing an antigen directed immunotherapy for Crohn's disease. The remaining authors disclose no conflicts.

Funding
The cytometric bead array was developed under a Synergy grant from the Rainin Foundation. This study was supported by grants from Crohn's and Colitis Canada Grant #CCC-GEMIII, Canadian Institutes of Health Research (CIHR) Grant #CMF108031 and The Leona M. and Harry B. Helmsley Charitable Trust. Williams Turpin is a former recipient of a Postdoctoral Fellowship Research Award from the CIHR Fellowship/Canadian Association of Gastroenterology (CAG)/Ferring Pharmaceuticals Inc. Sun-Ho Lee is a recipient of the Imagine/CAG/CIHR Fellowship Award. Sun-Ho Lee and Williams Turpin are recipients of fellowship awards from the Department of Medicine, Mount Sinai Hospital, Toronto. Kenneth Croitoru is recipient of the Canada Research Chair in Inflammatory Bowel Diseases. Of note, SciNet (Niagara supercomputer at the SciNet HPC Consortium) is funded by the Canada Foundation for Innovation; the Government of Ontario; Ontario Research Fund - Research Excellence; and the University of Toronto.

Data Availability
Requests for raw and analyzed data should follow the instructions given at http://www.gemproject.ca/data-access/. All submissions will be reviewed by the GEM Project Operating Committee to ensure that the requested samples/data will not interfere in any way with the intended GEM Project analysis of the nested cohort as per the original GEM Project Study Design and is not a duplication of analysis already ongoing. Those proposals meeting this evaluation will be distributed to all members of the GEM Project Steering Committee for review and open discussion. This review will focus on the global scientific merit of the proposal. This review will assess the basic scientific merit and the availability of requested samples and data, ensuring there is no compromise of the original intent of the GEM project. It would be of value to contact a member of the Steering Committee who could help sponsor your application. Those projects achieving a majority vote of approval at the GEM Project Steering Committee will be informed that the GEM Project will provide a letter of support stating that the requested samples or data will be made available to the applicants once the applicant receives funding from a granting agency that applies an independent peer review process to the proposal. The criteria to be used for review of all submissions will include the "scientific relevance" of the proposal and the judged availability of biological material requested. The budget to be requested from a funding agency must allow for any expenses in processing samples or in setting up the appropriate queries of the database. The intent is to allow sufficient time for applicants to consider submission for funding opportunities.