World Journal of

Gastroenterology

World Journal of Gastroenterology

Contents

Contents

Contents

ABOUT COVER

Associate Editor of *World Journal of Gastroenterology*, Ming-Lung Yu, MD, PhD, Chair Professor, Chief, Hepatitis Research Center, Kaohsiung Medical University, No. 100 Tzyou 1st Road, Kaohsiung 807, Taiwan. fish6069@gmail.com

AIMS AND SCOPE

The primary aim of *World Journal of Gastroenterology* (*WJG*, *World J Gastroenterol*) is to provide scholars and readers from various fields of gastroenterology and hepatology with a platform to publish high-quality basic and clinical research articles and communicate their research findings online. *WJG* mainly publishes articles reporting research results and findings obtained in the field of gastroenterology and hepatology and covering a wide range of topics including gastroenterology, hepatology, gastrointestinal endoscopy, gastrointestinal surgery, gastrointestinal oncology, and pediatric gastroenterology.

INDEXING/ABSTRACTING

The *WJG* is now abstracted and indexed in Science Citation Index Expanded (SCIE, also known as SciSearch®), Current Contents/Clinical Medicine, Journal Citation Reports, Index Medicus, MEDLINE, PubMed, PubMed Central, Scopus, Reference Citation Analysis, China National Knowledge Infrastructure, China Science and Technology Journal Database, and Superstar Journals Database. The 2022 edition of Journal Citation Reports® cites the 2021 impact factor (IF) for *WJG* as 5.374; IF without journal self cites: 5.187; 5-year IF: 5.715; Journal Citation Indicator: 0.84; Ranking: 31 among 93 journals in gastroenterology and hepatology; and Quartile category: Q2. The *WJG*'s CiteScore for 2021 is 8.1 and Scopus CiteScore rank 2021: Gastroenterology is 18/149.

RESPONSIBLE EDITORS FOR THIS ISSUE

Case Control Study

Serological profiling of Crohn's disease and ulcerative colitis patients reveals anti-microbial antibody signatures

Mahasish Shome, Lusheng Song, Stacy Williams, Yunro Chung, Vel Murugan, Jin G Park, William Faubion, Shabana F Pasha, Jonathan A Leighton, Joshua LaBaer, Ji Qiu

Mahasish Shome, Lusheng Song, Stacy Williams, Yunro Chung, Vel Murugan, Jin G Park, Joshua LaBaer, Ji Qiu, Virginia G. Piper Center for Personalized Diagnostics, Biodesign Institute, Arizona State University, Tempe, AZ 85281, United States

William Faubion, Department of Internal Medicine, Mayo Clinic, Rochester, MN 55902, United States

Shabana F Pasha, Department of Internal Medicine, Division of Gastroenterology and Hepatology, Mayo Clinic Arizona, Scottsdale, AZ 85259, United States

Jonathan A Leighton, Division of Gastroenterology, Mayo Clinic School of Medicine, Scottsdale, AZ 85259, United States

Corresponding author: Ji Qiu, PhD, Research Assistant Professor, Virginia G. Piper Center for Personalized Diagnostics, Biodesign Institute, Arizona State University, 1001 S McAllister Avenue, Tempe, AZ 85281, United States. ji.qiu@asu.edu

Abstract

BACKGROUND
The healthcare burden of inflammatory bowel disease (IBD) is rising globally and there are limited non-invasive biomarkers for accurate and early diagnosis.

AIM
To understand the important role that intestinal microbiota play in IBD pathogenesis and identify anti-microbial antibody signatures that benefit clinical management of IBD patients.

METHODS
We performed serological profiling of 100 Crohn's disease (CD) patients, 100 ulcerative colitis (UC) patients and 100 healthy controls against 1173 bacterial and 397 viral proteins from 50 bacteria and 33 viruses on protein microarrays. The study subjects were randomly divided into discovery (*n* = 150) and validation (*n* = 150) sets. Statistical analysis was performed using R packages.

RESULTS
Anti-bacterial antibody responses showed greater differential prevalence among the three subject groups than anti-viral antibody responses. We identified novel

antibodies against the antigens of *Bacteroidetes vulgatus* (BVU_0562) and *Streptococcus pneumoniae* (SP_1992) showing higher prevalence in CD patients relative to healthy controls. We also identified antibodies against the antigen of *Streptococcus pyogenes* (SPy_2009) showing higher prevalence in healthy controls relative to UC patients. Using these novel antibodies, we built biomarker panels with area under the curve (AUC) of 0.81, 0.87, and 0.82 distinguishing CD *vs* control, UC *vs* control, and CD *vs* UC, respectively. Subgroup analysis revealed that penetrating CD behavior, colonic CD location, CD patients with a history of surgery, and extensive UC exhibited highest antibody prevalence among all patients. We demonstrated that autoantibodies and anti-microbial antibodies in CD patients had minimal correlation.

CONCLUSION
We have identified antibody signatures for CD and UC using a comprehensive analysis of anti-microbial antibody response in IBD. These antibodies and the source microorganisms of their target antigens improve our understanding of the role of specific microorganisms in IBD pathogenesis and, after future validation, should aid early and accurate diagnosis of IBD.

Key Words: Inflammatory bowel disease; Anti-microbial antibody; Protein microarray; Crohn's disease; Ulcerative colitis; Gut microbiome

Core Tip: We performed the largest serological profiling of anti-microbial antibodies to date in using 100 Crohn's disease (CD) and 100 ulcerative colitis (UC) patients. We identified novel anti-microbial antibodies with differential prevalence in inflammatory bowel disease (IBD) patients compared with healthy controls. There was minimal correlation between anti-microbial antibodies and our previously reported autoantibodies in CD patients. We combined novel anti-microbial antibodies to build biomarker panels distinguishing CD *vs* control, UC *vs* control and CD *vs* UC with an area under the curve of 0.81, 0.87, and 0.82, respectively. Subgroup analysis revealed that IBD patients with severe disease had the highest antibody prevalence.

INTRODUCTION

Inflammatory bowel disease (IBD) represents a group of intestinal disorders that causes chronic inflammation in the digestive tract. The two main clinical phenotypes are ulcerative colitis (UC) and Crohn's disease (CD). The public health burden of IBD is rising globally[1]. Early and accurate diagnosis is key to reducing this burden. Gastroenterologists often use a combination of relatively invasive procedures, like ileocolonoscopy with biopsy for diagnosis, and to determine the disease extent and activity. There is a need for serological biomarkers that can reveal the disease state non-invasively. Herein, our objectives were to discover anti-microbial antibody signatures in IBD patients and understand the association of microbial infection with IBD pathogenesis.

IBD is caused by a combination of genetic predisposition, faulty immune responses, and environmental factors[2]. The interaction of microbes with the gut mucosa in a genetically susceptible individual and the corresponding immune response play a pivotal role in the initiation and progression of IBD[3]. After birth, a limited diversity microbial community develops into a complex community due to the influence of diet and environmental factors[4]. During the second or third decade of life, a dysbiosis is observed in IBD patients which leads to an imbalance between commensal and potentially pathogenic microorganisms[5]. The healthy gut microbiota predominately comprises Firmicutes and Bacteroidetes, and to a lesser extent, Actinobacteria and Proteobacteria[6,7]. In IBD, dysbiosis is observed with reduced abundance of Firmicutes and either higher or similar abundance of Proteobacteria. Besides compositional changes, genetic alterations also contribute to gut dysbiosis that leads to disease initiation and progression. For example, NOD2 variants were found in 20%-40% of European and American CD patients[8,9]. NOD2 encodes an intracellular receptor for the bacterial peptidoglycan muramyl dipeptide, which helps maintain the balance of commensal bacterial flora[10].

Immune response to microbes results in the production of antibodies to microbial antigens[11]. Anti-*Saccharomyces cerevisiae* antibodies (ASCA) are associated with CD patients, with sensitivities and specificities ranging between 55% to 65% and 80% to 95%, respectively[12]. Perinuclear antineutrophil cytoplasmic antibodies (pANCA) are associated with UC patients, with sensitivities and specificities ranging between 50% to 71% and 75% to 98%, respectively[13]. Outer membrane protein of *Escherichia coli* (OmpC) and flagellin (CBir1) antibodies are prevalent in CD patients, with prevalence ranging between 24%-55% and 50%-56%[13]. The number and response magnitude of anti-microbial antibodies have previously been shown to indicate the presence of IBD, its severity and its clinical course; however, the clinical utility of available antibodies in diagnosis and clinical management of IBD patients has been limited. The techniques used to discover the known anti-microbial antibodies associated with IBD are of low throughput, and have only been applied to test on small number of candidate microorganisms or microbial antigens[14]. We have performed a large-scale comparative profiling of anti-microbial antibodies in CD and UC patients and healthy controls using an innovative protein microarray technology, namely Nucleic Acid Programmable Protein Array. We selected 1570 microbial proteins from our microbial protein collection (DNASU.org) from 50 bacteria and 33 viruses based on preliminary studies in our laboratory and a review of the literature, displayed them on microarrays and probed them against 100 CD, 100 UC and 100 healthy control serum samples.

MATERIALS AND METHODS

Patients and samples
All the serum samples were acquired from Serum Biobank at Mayo clinic with approval from institutional review board. CD patients were randomly selected, followed by age and gender matched healthy controls and UC patients. The samples (100 CD, 100 UC and 100 controls) were divided evenly into two non-overlapping discovery and validation sets randomly (Table 1). Disease status for study participants was assessed by clinicians at Mayo clinic.

Microbial protein array fabrication
We analyzed 1570 microbial proteins, of which 1173 proteins were from 50 different species of bacteria, 397 proteins were from 33 different species of viruses and the remaining proteins were autoantigens. These proteins were selected from our large collection of microbial antigens (DNASU.org) with reference to our anti-microbial antibody studies on other diseases (unpublished data). Microbial protein arrays were fabricated as described earlier[15,16]. Briefly, plasmids with genes of interest cloned in the pANT7_cGST expression vector were obtained from the DNASU plasmid repository, prepared, and printed into silicon nanowells using a piezoelectric dispensing system to produce microbial protein arrays. On the day of experiment, proteins were freshly expressed from printed plasmids using an *in-vitro* transcription and translation protein expression kit (Fisher Scientific) and captured by anti-GST antibody co-printed in each nanowell. After expression, microarrays were incubated with 1:100 diluted serum samples. We randomized the case and control serum samples while profiling on microarray to reduce bias. IgG and IgA anti-microbial protein antibodies were detected by Alexa-647 goat anti-human IgG (H+L) and Cy3 goat anti-human IgA (Jackson ImmunoResearch). After washing and drying, the microarrays were scanned in a Tecan PowerScanner and the raw fluorescence intensity data were extracted using the ArrayPro Analyzer Software. Raw fluorescence intensity of each protein on the microarray was divided by the median intensity of all the proteins on the microarray for normalization. The normalized value was termed as Median Normalized Intensity (MNI) and used for all analysis. Seropositivity of antibody for a particular antigen was defined as MNI ≥ 2 as we have done for our other studies[17,18].

Statistical analysis
Pairwise comparisons of numbers of IgG or IgA antibodies for each bacterial species among the 3 subject groups were performed using Chi-squared tests to assess statistical significance (Supplementary Figure 1). For each pairwise comparison, the Chi-squared *P* values were adjusted using the FDR (false discovery rate) method to reduce the likelihood of false positives. In addition to the multiple comparison adjustment at the antibody level, we performed adjustment at the species level.

For univariate analysis between two comparison groups, we calculated sensitivity for one group at the 96th percentile of the other group or the MNI of 2, whichever was larger. Antibodies with ≥ 14% sensitivity in the discovery set were selected as candidates for further validation. If an antibody had ≥ 14% sensitivity at 96% specificity in both discovery and validation sets, then it was considered as a "validated marker". Venn diagram for the overlap of microbial antigen targets were plotted using Venny[19].

We used a three-stage approach to build our multi-antibody panels. In the first stage, we selected all candidate biomarkers that passed the criteria above, *i.e.*, sensitivity was greater or equal than 14% at 96% specificity. Next, we applied the minimum redundancy maximum relevance algorithm to further select biomarkers that were possibly the most important and least correlated[20]. In the third stage, we

Table 1 Clinical information of the samples

	Discovery set			Validation set		
	CD	UC	HC	CD	UC	HC
N	50	50	50	50	50	50
Gender (female, male)	29, 21	29, 21	29, 21	28, 22	28, 22	28, 22
Age (median ± SD)	41 ± 17.66	44 ± 17.25	42 ± 18.47	39.5 ± 17.49	44.5 ± 17.23	39.5 ± 16.02
Disease behavior (B1/B2/B3)	9/10/6			16/8/2		
Disease location (L1/L2/L3/L4)	12/6/7/0			12/7/7/0		
Disease extent (E1/E2/E3)		0/32/18			0/34/16	
Surgery (Yes, No)	24, 25	8, 42		22, 27	7, 42	

Fischer's exact test *P* value is equal to 1 for the gender difference among CD, UC and HC in both discovery and validation set. Kruskal-Wallis test *P* value for the age difference among CD, UC and HC in discovery and validation set were 0.3159 and 0.1737 respectively. CD: Crohn's disease; UC: Ulcerative colitis; HC: Healthy control.

fit a logistic regression model using the selected biomarkers from the first two stages and generated its receiver operating characteristic curve and AUC value to evaluate the model's discriminatory performance between CD, UC, and healthy controls.

Pair-wise subgroup comparisons based on the Montreal classification were performed for the odds ratio (OR) of each antibody using the seropositivity threshold defined as the maximum of MNI 2 and the 75th percentile of all samples. Chi-squared tests were used to test global significance between all groups with a slight modification by adding 0.5 to each cell of the table to avoid zero cell counts. *P* values from the Chi-squared method were adjusted for each pair of comparisons and for all candidate biomarkers. The number of antibodies with significant difference in prevalence among classifications were counted based on OR > 1 and OR < 1 for each pair of classification of CD behavior, CD location, UC extent and the surgery history of CD patients. The difference in total number of antibodies with significant difference between classification groups were computed using two sample proportion test. We did not perform a subgroup analysis for the UC patients based on the surgery history because most (84 out of 100) had no surgeries (Table 1).

Spearman's rank correlation analysis was performed to assess the correlation between autoantibody and anti-microbial antibody reactivity for CD patients and healthy controls. The R "pheatmap" package was used to generate the heatmap for correlation coefficients.

Bioinformatics analysis

The NCBI Taxonomy browser was used to find the taxonomical details of all the bacteria and viruses used in our study. The taxa were downloaded as phylip tree file and was used as an input in interactive tree of life software. Two phylogenetic trees were created for bacteria and viruses with different colors distinguishing the phylum.

For sequence homology analysis, a pair-wise BLAST analysis was carried out on the antigen protein sequences of validated antibodies for CD *vs* healthy control analysis. E-values were used to generate a heatmap using Python Seaborn package.

RESULTS

Anti-microbial antibody profiling in IBD on microbial protein arrays

We studied IgG and IgA anti-microbial antibody profiles of 100 CD and 100 UC patients and 100 age-gender matched healthy controls (Table 1) against 1570 microbial antigens including 1173 antigens from 50 different bacteria and 397 antigens from 33 different viruses using our protein microarray platform (Figure 1, Supplementary Table 1). This study provided a representative overview of the anti-microbial antibody response in IBD patients (Supplementary Figure 1). The numbers of IgG antibodies against bacterial proteins from *Bacteroidetes vulgatus* (*B. vulgatus*) and *Citrobacter koseri* (*C. koseri*) were significantly higher in CD patients compared with those in healthy controls (Chi-square test, *P* < 0.01) (Supplementary Figure 1). On the contrary, the numbers of IgG antibodies against proteins from several bacteria, such as *Streptococcus pneumoniae* (*S. pneumoniae*), *Haemophilus influenza* (*H. influenzae*), *Staphylococcus aureus* (*S. aureus*), *Helicobacter pylori* (*H. pylori*) and *Parvimonas micra* (*P. micra*) were significantly lower in CD and UC patients compared with those in healthy controls (Chi-square test, *P* < 0.05)
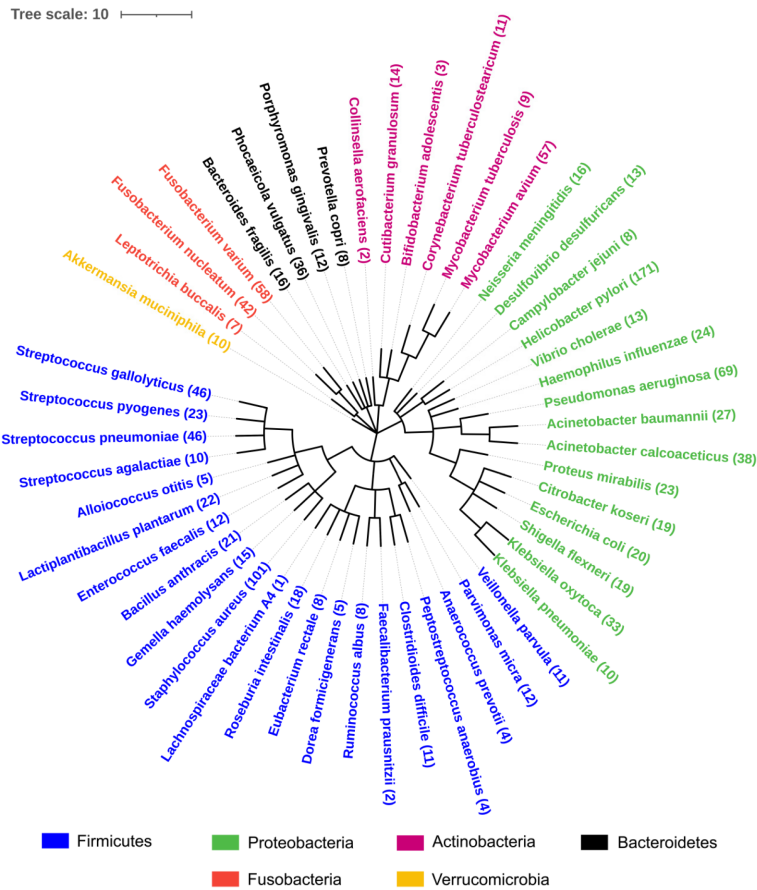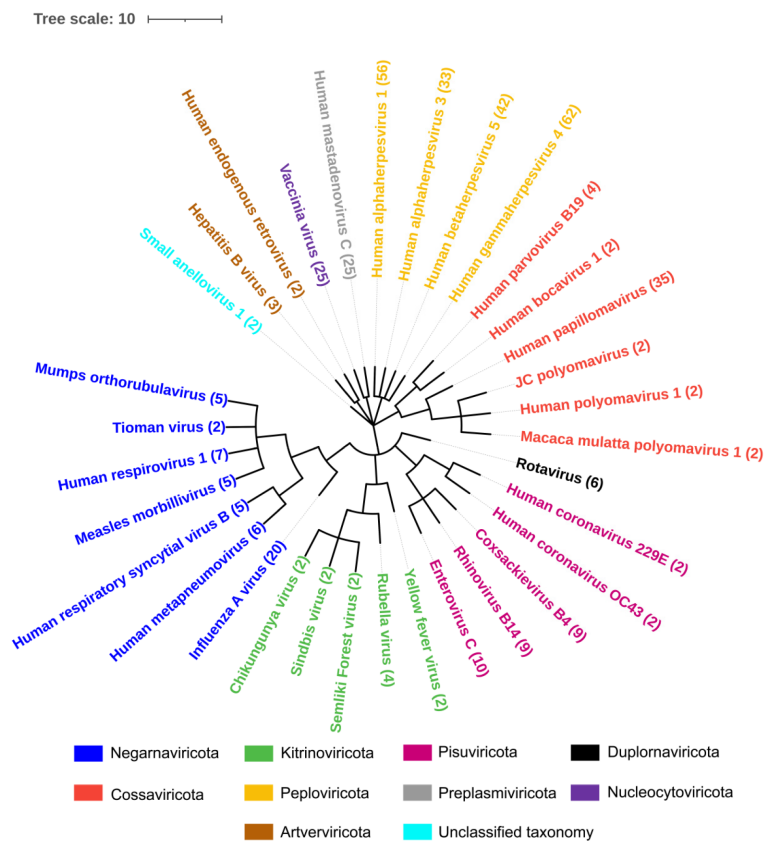
DOI: 10.3748/wjg.v28.i30.4089 **Copyright** ©The Author(s) 2022.

Figure 1 Phylogenetic tree of microbes studied with their corresponding number of proteins analyzed. A: Fifty species of bacteria with 1173 proteins were segregated into 6 phyla; B: Thirty-three species of viruses with 397 proteins were segregated into 10 phyla.

(Supplementary Figure 1). Overall, fewer IgA anti-microbial antibodies were found than IgG antibodies. The numbers of IgA antibodies against *S. pneumoniae*, *H. influenzae*, *S. aureus*, and *H. pylori* were significantly lower in UC patients compared with those in healthy controls (Chi-square test, *P* < 0.01). On the other hand, anti-viral IgG and IgA antibodies showed heterogenous prevalence with no clear trend of differences among CD, UC, healthy controls (data not shown). Therefore, we focused our analysis on anti-bacterial antibodies.

Antibodies distinguishing CD from healthy controls

We compared prevalence for individual anti-microbial antibodies between CD patients and healthy controls. We randomly split samples evenly into discovery and the validation sets (Table 1). For antibodies with elevated prevalence in CD patients, 13 IgG antibodies passed the criteria (sensitivity ≥ 14% at 96% specificity) in both discovery and validation sets (Table 2). Anti-A4-Fla2_IgG, a well-studied anti-bacterial flagellin antibody in CD, had the best performance with 47% sensitivity at 96% specificity in the full sample set (Table 2). Beside the flagellins, we found antibodies to four novel target antigens from *B. vulgatus* (BVU_0562), *P. mirabilis* (PMI_RS06815), *S. flexneri* (SF_Lpp) and *S. pneumoniae* (SP_1992) (Table 2) with no significant sequence homology to flagellins (Figure 2A).

To our surprise, 12 validated IgG antibodies showed elevated prevalence in healthy controls relative to CD patients (Supplementary Table 2). Among these 12 antibodies, anti-bacterial antibodies performed better in differentiating CD patients from healthy controls than anti-viral antibodies (Supplementary Table 2). Antibody against SPy_2009, an anchoring protein located in the cell wall of *Streptococcus pyogenes* (*S. pyogenes*), had the highest sensitivity of 24% at 96% specificity in healthy controls relative to CD patients. Seven validated IgA antibodies showed higher prevalence in healthy controls relative to CD patients (Supplementary Table 4).

Antibodies distinguishing UC from healthy controls

For anti-microbial antibodies with elevated prevalence in UC patients relative to healthy controls, 4 IgG antibodies passed the criteria in both discovery and validation sets (Table 2). Antibodies to A4-Fla2_IgG and a flagellin from *C. koseri* had a sensitivity of 18% and 15% respectively. For IgG antibodies with higher prevalence in healthy controls relative to UC patients, 32 antibodies got validated (Supplementary Table 3). Source microorganisms for the target antigens of these 32 antibodies were enriched for *S. pneumoniae, S. aureus,* and *H. influenzae* (2-sample proportion test, *P* < 0.05). 2.7% of the proteins on the microbial protein microarray were from *S. pneumoniae* while 18.7% of antigens for validated antibodies were from *S. pneumoniae*, 6.1% of the proteins on the microarrays were from *S. aureus* while 18.7% of antigens for validated antibodies were from *S. aureus,* and 1.4% of the proteins on the microarrays were from *H. influenzae* while 12.5% of antigens for validated antibodies were from *H. influenzae*. Nine validated IgA antibodies showed higher prevalence in healthy controls relative to UC patients (Supplementary Table 4).

Fewer anti-viral antibodies than anti-bacterial antibodies were validated comparing CD or UC patients with healthy controls (Table 2, Supplementary Tables 2 and 3). Anti-viral antibodies to Rhinovirus B14, Enterovirus C, Influenza A virus, Human metapneumovirus had higher prevalence in healthy controls compared with CD and UC patients (Supplementary Tables 2 and 3).

Comparison of anti-microbial antibody response between CD and UC

We found 46 IgG and 22 IgA validated anti-microbial antibodies with higher prevalence in CD patients compared to UC patients while 28 IgG and 9 IgA validated anti-microbial antibodies with higher prevalence in UC patients compared to CD patients. There was minimal overlap of the target antigens of these validated IgG and IgA antibodies (Figure 2B).

Multivariate analysis to distinguish CD, UC, and healthy controls

We built multi-antibody panels that could distinguish CD *vs* control, UC *vs* control, and CD *vs* UC with an area under the curve (AUC) of 0.81, 0.87, and 0.82 respectively. For CD *vs* control, antibodies against novel flagellins (HP_0115, CK_LafA, CK_LafA.1, VC_flaD, VC_flaB, VC_flaE, VC_flaA) had an AUC of 0.73, antibodies against non-flagellins (BVU_0562, SP_1992, PMI_RS06815, and SF_Lpp) had an AUC of 0.75 and the combined AUC of antibodies against novel flagellins and non-flagellins was 0.81 (Figure 3A). For UC *vs* control, a combination of seven antibodies, four against *S. pneumoniae* and one each against *S. aureus*, *H. influenzae* and *B. vulgatus* had an AUC of 0.87 (Figure 3B). For CD *vs* UC, combination of seven antibodies, two against *H. pylori* and one each against *E. coli*, *S. pneumoniae*, *S. pyogenes*, *C. jejuni* and *L. bacterium A4* had an AUC of 0.82 (Figure 3C).

Subgroup analysis

We investigated the association of CD behavior (B1, B2, B3), CD location (L1, L2, L3), and UC extent (E1, E2, E3) based on the Montreal classification with the anti-microbial antibody prevalence. We calculated the 4th quartile odds ratio for each antibody between the two classification groups and compared the number of antibodies with significant odds ratio (*P* value < 0.05) in each group. We found B3 (penetrating) had the highest prevalence of antibodies followed by B2 (stricturing) and B1 (non-

Table 2 Sensitivities of validated IgG antibodies comparing Crohn's disease and ulcerative colitis with healthy controls in the discovery, validation, and the entire set at 96% specificity

		Antigen	Protein name	Organism	Discovery	Validation	Entire
Crohn's disease	Bacteria	HP_0115	Flagellin B	*H. pylori*	28	48	38
		BVU_0562	Uncharacterized protein	*B. vulgatus*	26	22	25
		CK_LafA	Lateral flagellin	*C. koseri*	20	22	21
		CK_LafA.1	Lateral flagellin	*C. koseri*	16	26	24
		A4-Fla2	Flagellin	*L. bacterium A4*	40	54	47
		PMI_RS06815	Hypothetical protein	*P. mirabilis*	14	16	15
		VC_flaD	Flagellin	*V. cholerae*	24	18	19
		VC_flaB	Flagellin	*V. cholerae*	28	22	24
		VC_flaE	Flagellin	*V. cholerae*	26	28	23
		VC_flaA	Flagellin	*V. cholerae*	20	22	21
		SF_Lpp	Outer membrane lipoprotein	*S. flexneri*	14	18	14
		SP_1992	Cell wall surface anchor	*S. pneumoniae*	20	16	18
	Virus	BILF2	Glycoprotein BILF2	Human herpesvirus 4	18	18	18
Ulcerative colitis	Bacteria	CK_flgG	Flagellar basal-body rod protein	*C. koseri*	14	16	15
		A4-Fla2	Flagellin	*L. bacterium A4*	22	16	18
	Virus	BVRF2	Capsid scaffolding protein	Human herpesvirus 4	14	16	14
		UL139	Membrane glycoprotein UL139	Human herpesvirus 5	14	20	17

stricturing, non-penetrating) (Table 3). For CD location, we found L2 had the highest prevalence of antibodies followed by L3 (ileocolonic) and L1 (Table 3). For UC extent, E3 (extensive UC) had higher prevalence of antibodies compared to E2 (left sided UC). In addition to the Montreal classification, we also performed subgroup analysis based on the surgery history of CD patients. We found that patients who had surgery possessed higher prevalence of antibodies compared to those without surgery (Table 3).

Correlation of anti-microbial antibodies and autoantibodies in CD patients

We previously reported novel autoantibodies in CD patients using the same set of CD patients and healthy controls[21]. We profiled both IgG and IgA autoantibodies and anti-microbial antibodies in all 100 CD and 100 healthy controls. It is interesting to note the antibodies showing differences for autoantibodies were mostly IgA, but the anti-microbial antibodies were mostly IgG[21]. Anti-SNRPB_IgA had the highest sensitivity of 20% at 96% specificity among all autoantibodies compared with 47% sensitivity at 96% specificity for the best performing anti-microbial antibody, anti-A4-Fla2_IgG.

We compared the novel autoantibodies and validated anti-microbial antibody profiles to determine if correlation existed between their reactivity. Overall, we did not observe high correlation between autoantibodies and anti-microbial antibodies in CD patients (Figure 4). Anti-microbial antibodies formed two clusters, one with anti-flagellin antibodies, and the other with SF_Lpp_IgG and PMI_RS06815_IgG. Five autoantibodies, PRPH_IgA, SNAI1_IgA, PPP1R13L_IgA, SNRPB_IgA and PTTG1_IgA, formed a cluster. The remaining antibodies had relatively unique reactivity patterns.

DISCUSSION

We have performed a microbiomics study to understand the connection between host anti-microbial responses and IBD and to identify antibody signatures that can aid in the accurate diagnosis of IBD. We found antibody responses to novel non-flagellin antigens with elevated prevalence in CD patients compared with healthy controls. On the contrary, we observed many anti-microbial antibodies with

Table 3 Subgroup analysis of inflammatory bowel disease patients

Classification	Comparison	Number of antibodies with		Two sample proportion test
		OR > 1	OR < 1	
Disease behavior: B1: non-stricturing, non-penetrating; B2: stricturing; B3: penetrating	B1 *vs* B2 (*P* < 0.05)	0	32	*P* < 0.001
	B2 *vs* B3 (*P* < 0.05)	0	19	*P* < 0.001
	B1 *vs* B3 (*P* < 0.05)	2	41	*P* < 0.001
Disease location: L1: ileal; L2: colonic; L3: ileocolonic	L1 *vs* L2 (*P* < 0.05)	0	38	*P* < 0.001
	L2 *vs* L3 (*P* < 0.05)	9	5	*P* = 0.131
	L1 *vs* L3 (*P* < 0.05)	5	16	*P* < 0.001
Disease extent: E2: left sided UC; E3: extensive UC	E2 *vs* E3 (*P* < 0.05)	11	39	*P* < 0.001
Surgery in CD patients	No *vs* Yes (*P* < 0.05)	6	25	*P* < 0.001

For each comparison, the number of antibodies with significant difference in prevalence between two classifications were counted based on odds ratio (OR) > 1 and OR < 1. The difference in total number of antibodies for each comparison were computed using two sample proportion test. CD: Crohn's disease; UC: Ulcerative colitis; OR: Odds ratio.
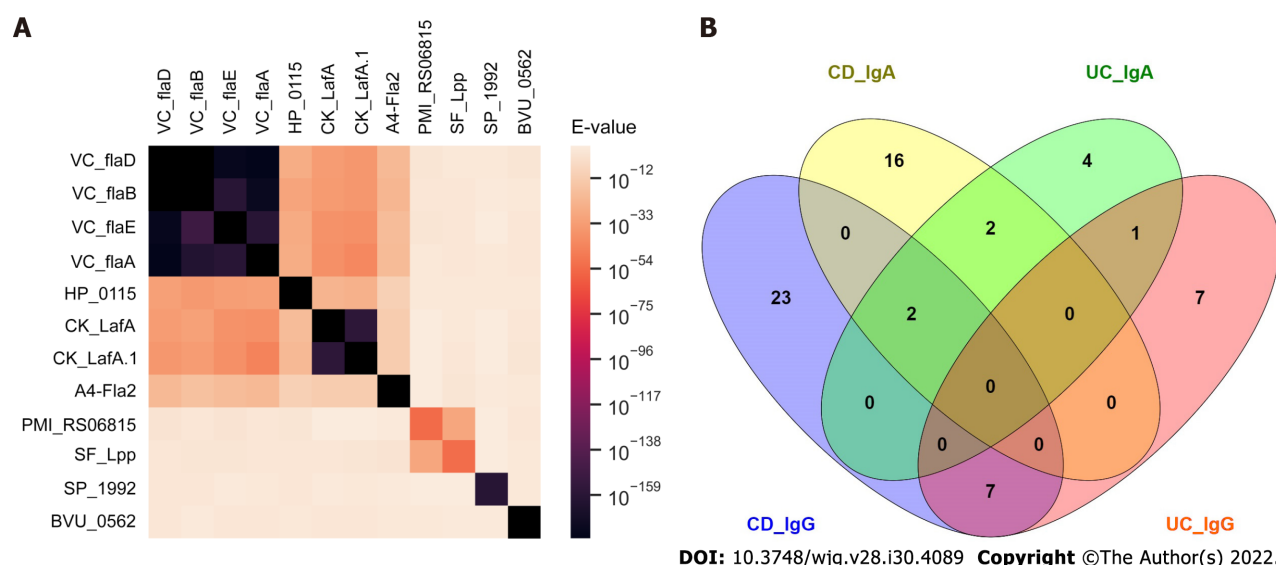
Figure 2 Sequence homology of target antigens of validated antibodies and overlap of antibodies among Crohn's disease and ulcerative colitis. A: Heatmap showing sequence homology among target antigens for antibodies with validated performance of ≥ 14% sensitivity at 96% specificity comparing Crohn's disease (CD) patients with healthy controls; B: CD_IgG, ulcerative colitis (UC)_IgG, CD_IgA and UC_IgG represent the overlap of anti-microbial antibodies of IgG and IgA isotypes in CD and UC patients with ≥ 14% sensitivity at 96% specificity against healthy controls in the discovery set. CD: Crohn's disease; UC: Ulcerative colitis.

lower prevalence in UC patients relative to healthy controls. We built antibody panels that could distinguish CD *vs* control, UC *vs* control and CD *vs* UC with AUCs of 0.81, 0.87, and 0.82, respectively. Lichtenstein *et al*[22] reported an integrated serological (ASCA-IgA, ASCA-IgG, anti-OmpC, anti-CBir1, anti-I2, pANCA) and genetic (SNP8, SNP12, SNP13) marker panel with an AUC of 0.80 to distinguish CD *vs* control. A panel of serological markers (ASCA-IgA, ASCA-IgG, ANCA, pANCA, OmpC, and CBir1) built by Plevy *et al*[23] yields an AUC of 0.78 to distinguish CD *vs* UC. Our novel antibody marker panels had comparable or better performance in IBD diagnosis or distinguishing CD from UC subtypes. We observed a stronger anti-microbial antibody response with more aggressive disease in both CD and UC patients. Finally, we demonstrated that anti-microbial antibodies and autoantibodies had different reactivity patterns in CD patients.

Our comprehensive study of anti-microbial antibodies in IBD patients provided interesting insight into its pathogenesis. Antibody responses to proteins from *B. vulgatus*, *P. mirabilis*, *S. flexneri* and *S. pneumoniae* were elevated in CD patients. *B. vulgatus* has been reported to induce colitis in IBD-susceptible mice[24,25]. *P. mirabilis* in gut can induce inflammation in cells and a colitis mouse model and has been associated with CD pathogenesis[26]. Our results suggest that *B. vulgatus* and *P. mirabilis*
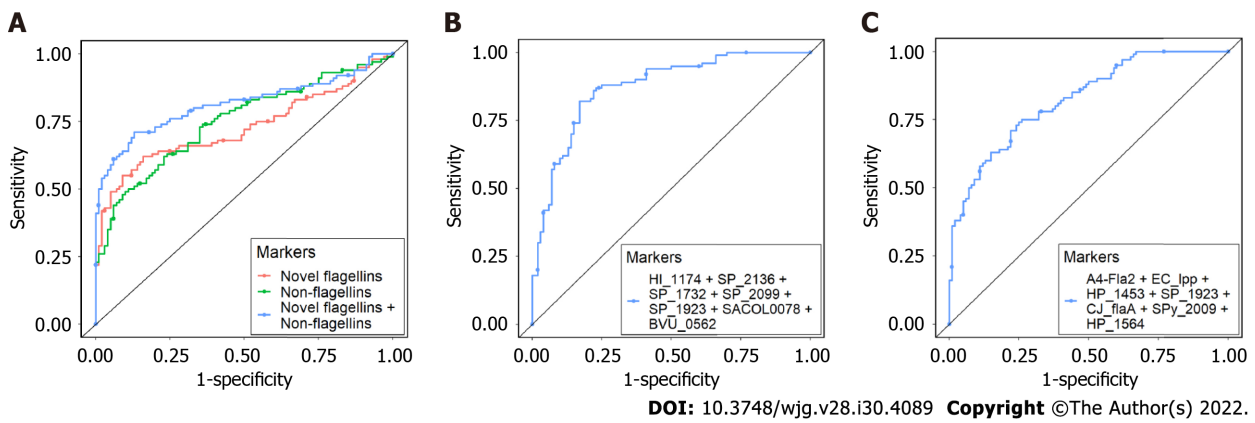
Figure 3 Receiver operating characteristic curves to discriminate Crohn's disease, ulcerative colitis, and healthy controls. A: Receiver operating characteristic (ROC) curve for Crohn's disease (CD) *vs* healthy controls. Area under the curve (AUC) values of novel anti-flagellin antibodies (HP_0115, CK_LafA, CK_LafA.1, VC_flaD, VC_flaB, VC_flaE, VC_flaA) and anti-non-flagellin antibodies (BVU_0562, SP_1992, PMI_RS06815, SF_Lpp) was 0.73 and 0.75, respectively. The AUC value obtained with a combination of novel anti-flagellin and anti-non-flagellin antibodies was 0.81; B: ROC curve for ulcerative colitis (UC) *vs* healthy controls. The AUC value obtained with a combination of 7 markers was 0.87; C: ROC curve for CD *vs* UC. The AUC value obtained with a combination of 7 markers was 0.82.
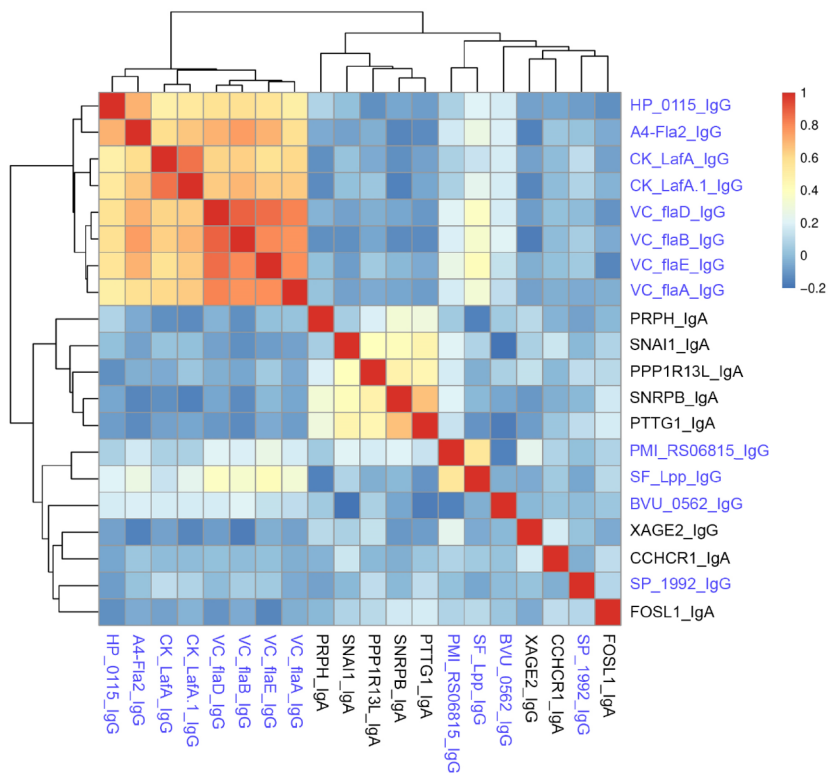
Figure 4 Spearman's rank correlation coefficient heatmap of anti-microbial antibodies and autoantibodies in Crohn's disease patients. The names of anti-microbial antibodies are colored in blue while autoantibodies are colored in black.

may also play a role in human CD development. We observed reduced antibody responses in UC patients to several genera of the Firmicutes phylum including *P. micra, S. pyogenes, S. aureus,* which were often reduced in abundance in UC patients' gut microbiota[27]. For several genera belonging to Proteobacteria phylum, such as *H. influenzae, H. pylori, K. oxytoca,* we observed overall reduced antibody responses; however, their abundance in the gut microbiota of UC patients has been reported to be either increased or remained the same compared with healthy controls[6,28].

Beyond exposure alone, antibody response requires functional immunological interaction between a microorganism and the host; however, anti-microbial antibodies by themselves do not prove causality. As such, source microorganisms whose antibodies show significant changes between IBD patients and healthy controls warrant future confirmation and functional assessment in causing IBD. A4-Fla2

flagellin included in our study showed IBD-specific prevalence with performance similar to that reported in the literature[14,29]. We also identified several antibodies to flagellins with higher prevalence in CD patients relative to healthy controls.

Previous studies mostly focused on antibodies with higher prevalence in IBD patients[30]. Our unbiased data-driven approach revealed the existence of many anti-microbial antibodies with higher prevalence in healthy controls relative to CD and especially UC patients (Supplementary Tables 3 and 4). The reduction observed in CD and UC patients may be attributed to the dysbiosis and reduced diversity of gut microbiota in CD and UC patients[28,31]. It is also possible that the reduction in anti-microbial antibodies in some CD and UC patients was in part because of immunosuppressive therapies they received. The greater number of antibodies having high prevalence in CD patients compared with UC patients indicates stronger anti-microbial humoral immunity in CD than in UC, which is consistent with reports in the literature that most known anti-microbial antibodies, such as ASCA, anti-OmpC, anti-Cbir1, and anti-I2, had higher prevalence in CD patients than in UC patients[30]. This agreement, together with comparable performance of anti-flagellin antibodies in our study and that reported in the literature, suggests that our results reflect the microbial association of IBD etiopathology. However, the use of samples from patients with established disease and the lack of information on immunosuppressive therapies of these patients limited the interpretation of the results.

We studied the association between anti-microbial antibody prevalence and various disease classifications based on Montreal classification and surgery history. We found more antibodies with significantly higher prevalence in patients with more aggressive disease behaviors relative to those with milder disease behavior. We also found more antibodies with significantly higher prevalence in colonic CD patients relative to those in ileal CD patients. Our results were consistent with previous reports that increasing diversity and magnitude of anti-microbial immune response was correlated with increased frequency of penetrating and/or stricturing disease behavior[13,32]. It is known that the colon has a microbial density of 10^{11}-10^{12} anaerobic bacteria/gram while the ileum is colonized by 10^7-10^8 anaerobic bacteria/gram[7]. Kleessen *et al*[33] found higher percentage of bacterial invasion of mucosa in colon compared to ileum. CD patients requiring surgery usually had more severe disease compared with those who did not need surgery. Stronger anti-microbial immune response in patients with severe CD or UC suggests a higher abundance of the source microorganisms for the target antigens of the differential antibodies and/or a stronger more conducive immune microenvironment at the disease site in severe disease.

Both autoantibodies and anti-microbial antibodies associated with IBD have been reported[30]. One popular hypothesis for the autoantibody elicitation is molecular mimicry, where anti-microbial antibodies cross react with human proteins. However, we found minimal correlation between the anti-microbial antibodies and the autoantibody profiles in the same set of CD samples. The lack of correlation suggests that IBD-specific autoantibodies and anti-microbial antibodies are elicited independently through different underlying mechanisms, and cross-reactivity may play less of a role in eliciting CD-associated autoantibodies. The breakdown of immune tolerance to human proteins might have occurred due to the damaged gut epithelial cells and the faulty immunological microenvironment partly caused by microbial infections. In addition, the elicitation of autoantibodies may be associated with the infections of multiple microorganisms, and the correlation with individual anti-microbial antibodies may not be great.

Strengths of our study include the broadest analysis to date of IgG and IgA antibodies against individual antigens from many different microorganisms in both CD and UC patients and the use of a two-stage approach with discovery and independent validation of antibody markers. There are some limitations to our study. Except for a few microbes, the number of proteins studied for each species is small, which might limit our interpretation of antibody response in IBD at the species level. Furthermore, many samples used in studies were collected from patients with established disease. Future studies with access to more clinical information, such as information about the use of immunosuppressive or antibiotics treatment, could aid in the interpretation of our results. In addition, samples collected from newly diagnosed patients will strengthen our ability to identify diagnostic markers and microbial connections for IBD development.

CONCLUSION

In summary, we have demonstrated the power of a microbiomics study of anti-microbial antibodies in IBD for the identification of anti-microbial antibody signatures to improve early accurate diagnosis and help understand IBD etiology. The elucidation of the source microorganisms of the target antigens of the antibody biomarkers could lead to novel strategies for the prevention of IBD.

ARTICLE HIGHLIGHTS

Research background

Early and accurate diagnosis of inflammatory bowel disease (IBD) can benefit the clinical management of IBD patients. The clinical utility of available non-invasive biomarkers is limited.

Research motivation

The gut microbiota plays an important role in the pathogenesis of IBD. Antibody responses to microbial antigens can be exploited to identify better diagnosis markers and improve our understanding of IBD pathology.

Research objectives

This study aimed to compare anti-microbial antibody profiles in Crohn's disease (CD) and ulcerative colitis (UC) patients and healthy controls.

Research methods

We employed an innovative protein microarray platform and profiled antibodies against 1570 microbial antigens in 100 CD, 100 UC, and 100 healthy controls.

Research results

Antibodies to bacterial proteins were better in distinguishing IBD patients from healthy controls compared with antibodies to viral proteins. We identified a set of novel anti-microbial antibodies against the antigens of *Bacteroidetes vulgatus* (BVU_0562) and *Streptococcus pneumoniae* (SP_1992) elevated in CD patients relative to healthy controls. In addition, anti-microbial antibodies against the antigen of *Streptococcus pyogenes* (SPy_2009) were found to be elevated in healthy controls relative to UC patients. We constructed antibody panels that could distinguish CD *vs* control, UC *vs* control, and CD *vs* UC with an AUC of 0.81, 0.87 and 0.82 respectively. Patients with severe disease had higher prevalence of anti-microbial antibodies. There was minimal correlation among the occurrence of autoantibodies and anti-microbial antibodies in CD patients.

Research conclusions

This study discovered novel anti-microbial antibodies with differential prevalence in CD and UC patients relative to healthy controls. In addition, this study revealed the potentially different roles of gut microbiota in CD and UC pathology.

Research perspectives

Our study demonstrated the power of a microbiomics approach to identify biomarkers that could aid in the early and accurate detection of IBD non-invasively, and shed light into the role of various microbes in IBD etiology.

ORCID number: William Faubion 0000-0003-1291-5745; Ji Qiu 0000-0002-7913-9042.

REFERENCES

1 **GBD 2017 Inflammatory Bowel Disease Collaborators**. The global, regional, and national burden of inflammatory bowel disease in 195 countries and territories, 1990-2017: a systematic analysis for the Global Burden of Disease Study 2017. *Lancet Gastroenterol Hepatol* 2020; **5**: 17-30 [PMID: 31648971 DOI: 10.1016/S2468-1253(19)30333-4]

2 **de Souza HS**, Fiocchi C. Immunopathogenesis of IBD: current state of the art. *Nat Rev Gastroenterol Hepatol* 2016; **13**: 13-27 [PMID: 26627550 DOI: 10.1038/nrgastro.2015.186]

3 **Fakhoury M**, Negrulj R, Mooranian A, Al-Salami H. Inflammatory bowel disease: clinical aspects and treatments. *J Inflamm Res* 2014; **7**: 113-120 [PMID: 25075198 DOI: 10.2147/JIR.S65979]

4 **Zuo T**, Ng SC. The Gut Microbiota in the Pathogenesis and Therapeutics of Inflammatory Bowel Disease. *Front Microbiol* 2018; **9**: 2247 [PMID: 30319571 DOI: 10.3389/fmicb.2018.02247]

5 **Ni J**, Wu GD, Albenberg L, Tomov VT. Gut microbiota and IBD: causation or correlation? *Nat Rev Gastroenterol Hepatol* 2017; **14**: 573-584 [PMID: 28743984 DOI: 10.1038/nrgastro.2017.88]

6 **Frank DN**, St Amand AL, Feldman RA, Boedeker EC, Harpaz N, Pace NR. Molecular-phylogenetic characterization of microbial community imbalances in human inflammatory bowel diseases. *Proc Natl Acad Sci USA* 2007; **104**: 13780-13785 [PMID: 17699621 DOI: 10.1073/pnas.0706625104]

7 **Sartor RB**. Microbial influences in inflammatory bowel diseases. *Gastroenterology* 2008; **134**: 577-594 [PMID: 18242222 DOI: 10.1053/j.gastro.2007.11.059]

8 **Ahmad T**, Armuzzi A, Bunce M, Mulcahy-Hawes K, Marshall SE, Orchard TR, Crawshaw J, Large O, de Silva A, Cook JT, Barnardo M, Cullen S, Welsh KI, Jewell DP. The molecular classification of the clinical manifestations of Crohn's disease. *Gastroenterology* 2002; **122**: 854-866 [PMID: 11910336 DOI: 10.1053/gast.2002.32413]

9 **Hugot JP**, Chamaillard M, Zouali H, Lesage S, Cézard JP, Belaiche J, Almer S, Tysk C, O'Morain CA, Gassull M, Binder V, Finkel Y, Cortot A, Modigliani R, Laurent-Puig P, Gower-Rousseau C, Macry J, Colombel JF, Sahbatou M, Thomas G. Association of NOD2 leucine-rich repeat variants with susceptibility to Crohn's disease. *Nature* 2001; **411**: 599-603 [PMID: 11385576 DOI: 10.1038/35079107]

10 **Ramanan D**, Tang MS, Bowcutt R, Loke P, Cadwell K. Bacterial sensor Nod2 prevents inflammation of the small intestine by restricting the expansion of the commensal Bacteroides vulgatus. *Immunity* 2014; **41**: 311-324 [PMID: 25088769 DOI: 10.1016/j.immuni.2014.06.015]

11 **Elkadri AA**, Stempak JM, Walters TD, Lal S, Griffiths AM, Steinhart AH, Silverberg MS. Serum antibodies associated with complex inflammatory bowel disease. *Inflamm Bowel Dis* 2013; **19**: 1499-1505 [PMID: 23702714 DOI: 10.1097/MIB.0b013e318281f2a1]

12 **Vermeire S**, Joossens S, Peeters M, Monsuur F, Marien G, Bossuyt X, Groenen P, Vlietinck R, Rutgeerts P. Comparative study of ASCA (Anti-Saccharomyces cerevisiae antibody) assays in inflammatory bowel disease. *Gastroenterology* 2001; **120**: 827-833 [PMID: 11231936 DOI: 10.1053/gast.2001.22546]

13 **Kuna AT**. Serological markers of inflammatory bowel disease. *Biochem Med (Zagreb)* 2013; **23**: 28-42 [PMID: 23457764 DOI: 10.11613/bm.2013.006]

14 **Lodes MJ**, Cong Y, Elson CO, Mohamath R, Landers CJ, Targan SR, Fort M, Hershberg RM. Bacterial flagellin is a dominant antigen in Crohn disease. *J Clin Invest* 2004; **113**: 1296-1306 [PMID: 15124021 DOI: 10.1172/JCI20295]

15 **Takulapalli BR**, Qiu J, Magee DM, Kahn P, Brunner A, Barker K, Means S, Miersch S, Bian X, Mendoza A, Festa F, Syal K, Park JG, LaBaer J, Wiktor P. High density diffusion-free nanowell arrays. *J Proteome Res* 2012; **11**: 4382-4391 [PMID: 22742968 DOI: 10.1021/pr300467q]

16 **Song L**, Wallstrom G, Yu X, Hopper M, Van Duine J, Steel J, Park J, Wiktor P, Kahn P, Brunner A, Wilson D, Jenny-Avital ER, Qiu J, Labaer J, Magee DM, Achkar JM. Identification of Antibody Targets for Tuberculosis Serology using High-Density Nucleic Acid Programmable Protein Arrays. *Mol Cell Proteomics* 2017; **16**: S277-S289 [PMID: 28223349 DOI: 10.1074/mcp.M116.065953]

17 **Song L**, Song M, Rabkin CS, Williams S, Chung Y, Van Duine J, Liao LM, Karthikeyan K, Gao W, Park JG, Tang Y, Lissowska J, Qiu J, LaBaer J, Camargo MC. *Helicobacter pylori* Immunoproteomic Profiles in Gastric Cancer. *J Proteome Res* 2021; **20**: 409-419 [PMID: 33108201 DOI: 10.1021/acs.jproteome.0c00466]

18   **Song L**, Song M, Camargo MC, Van Duine J, Williams S, Chung Y, Kim KM, Lissowska J, Sivins A, Gao W, Karthikeyan K, Park J, Leja M, Cohen JI, LaBaer J, Qiu J, Rabkin CS. Identification of anti-Epstein-Barr virus (EBV) antibody signature in EBV-associated gastric carcinoma. *Gastric Cancer* 2021; **24**: 858-867 [PMID: 33661412 DOI: 10.1007/s10120-021-01170-z]

19   **Oliveros JC**.   VENNY. An interactive tool for comparing lists with Venn Diagrams. Available from: https://bioinfogp.cnb.csic.es/tools/venny/index.html

20   **Ding C**, Peng H. Minimum redundancy feature selection from microarray gene expression data. *J Bioinform Comput Biol* 2005; **3**: 185-205 [PMID: 15852500 DOI: 10.1142/s0219720005001004]

21   **Wang H**, Demirkan G, Bian X, Wallstrom G, Barker K, Karthikeyan K, Tang Y, Pasha SF, Leighton JA, Qiu J, LaBaer J. Identification of Antibody Against SNRPB, Small Nuclear Ribonucleoprotein-Associated Proteins B and B', as an Autoantibody Marker in Crohn's Disease using an Immunoproteomics Approach. *J Crohns Colitis* 2017; **11**: 848-856 [PMID: 28204086 DOI: 10.1093/ecco-jcc/jjx019]

22   **Lichtenstein GR**, Targan SR, Dubinsky MC, Rotter JI, Barken DM, Princen F, Carroll S, Brown M, Stachelski J, Chuang E, Landers CJ, Stempak JM, Singh S, Silverberg MS. Combination of genetic and quantitative serological immune markers are associated with complicated Crohn's disease behavior. *Inflamm Bowel Dis* 2011; **17**: 2488-2496 [PMID: 21391291 DOI: 10.1002/ibd.21661]

23   **Plevy S**, Silverberg MS, Lockton S, Stockfisch T, Croner L, Stachelski J, Brown M, Triggs C, Chuang E, Princen F, Singh S. Combined serological, genetic, and inflammatory markers differentiate non-IBD, Crohn's disease, and ulcerative colitis patients. *Inflamm Bowel Dis* 2013; **19**: 1139-1148 [PMID: 23518807 DOI: 10.1097/MIB.0b013e318280b19e]

24   **Bloom SM**, Bijanki VN, Nava GM, Sun L, Malvin NP, Donermeyer DL, Dunne WM Jr, Allen PM, Stappenbeck TS. Commensal Bacteroides species induce colitis in host-genotype-specific fashion in a mouse model of inflammatory bowel disease. *Cell Host Microbe* 2011; **9**: 390-403 [PMID: 21575910 DOI: 10.1016/j.chom.2011.04.009]

25   **Glassner KL**, Abraham BP, Quigley EMM. The microbiome and inflammatory bowel disease. *J Allergy Clin Immunol* 2020; **145**: 16-27 [PMID: 31910984 DOI: 10.1016/j.jaci.2019.11.003]

26   **Zhang J**, Hoedt EC, Liu Q, Berendsen E, Teh JJ, Hamilton A, O' Brien AW, Ching JYL, Wei H, Yang K, Xu Z, Wong SH, Mak JWY, Sung JJY, Morrison M, Yu J, Kamm MA, Ng SC. Elucidation of Proteus mirabilis as a Key Bacterium in Crohn's Disease Inflammation. *Gastroenterology* 2021; **160**: 317-330.e11 [PMID: 33011176 DOI: 10.1053/j.gastro.2020.09.036]

27   **Zakerska-Banaszak O**, Tomczak H, Gabryel M, Baturo A, Wolko L, Michalak M, Malinska N, Mankowska-Wierzbicka D, Eder P, Dobrowolska A, Slomski R, Skrzypczak-Zielinska M. Dysbiosis of gut microbiota in Polish patients with ulcerative colitis: a pilot study. *Sci Rep* 2021; **11**: 2166 [PMID: 33495479 DOI: 10.1038/s41598-021-81628-3]

28   **Manichanh C**, Rigottier-Gois L, Bonnaud E, Gloux K, Pelletier E, Frangeul L, Nalin R, Jarrin C, Chardon P, Marteau P, Roca J, Dore J. Reduced diversity of faecal microbiota in Crohn's disease revealed by a metagenomic approach. *Gut* 2006; **55**: 205-211 [PMID: 16188921 DOI: 10.1136/gut.2005.073817]

29   **Targan SR**, Landers CJ, Yang H, Lodes MJ, Cong Y, Papadakis KA, Vasiliauskas E, Elson CO, Hershberg RM. Antibodies to CBir1 flagellin define a unique response that is associated independently with complicated Crohn's disease. *Gastroenterology* 2005; **128**: 2020-2028 [PMID: 15940634 DOI: 10.1053/j.gastro.2005.03.046]

30   **Mitsuyama K**, Niwa M, Takedatsu H, Yamasaki H, Kuwaki K, Yoshioka S, Yamauchi R, Fukunaga S, Torimura T. Antibody markers in the diagnosis of inflammatory bowel disease. *World J Gastroenterol* 2016; **22**: 1304-1310 [PMID: 26811667 DOI: 10.3748/wjg.v22.i3.1304]

31   **Furrie E**, Macfarlane S, Cummings JH, Macfarlane GT. Systemic antibodies towards mucosal bacteria in ulcerative colitis and Crohn's disease differentially activate the innate immune response. *Gut* 2004; **53**: 91-98 [PMID: 14684582 DOI: 10.1136/gut.53.1.91]

32   **Dubinsky MC**, Lin YC, Dutridge D, Picornell Y, Landers CJ, Farrior S, Wrobel I, Quiros A, Vasiliauskas EA, Grill B, Israel D, Bahar R, Christie D, Wahbeh G, Silber G, Dallazadeh S, Shah P, Thomas D, Kelts D, Hershberg RM, Elson CO, Targan SR, Taylor KD, Rotter JI, Yang H; Western Regional Pediatric IBD Research Alliance. Serum immune responses predict rapid disease progression among children with Crohn's disease: immune responses predict disease progression. *Am J Gastroenterol* 2006; **101**: 360-367 [PMID: 16454844 DOI: 10.1111/j.1572-0241.2006.00456.x]

33   **Kleessen B**, Kroesen AJ, Buhr HJ, Blaut M. Mucosal and invading bacteria in patients with inflammatory bowel disease compared with controls. *Scand J Gastroenterol* 2002; **37**: 1034-1041 [PMID: 12374228 DOI: 10.1080/003655202320378220]

Published by **Baishideng Publishing Group Inc**
7041 Koll Center Parkway, Suite 160, Pleasanton, CA 94566, USA
Telephone: +1-925-3991568
E-mail: bpgoffice@wjgnet.com
Help Desk: https://www.f6publishing.com/helpdesk
https://www.wjgnet.com