



US 20180224448A1

(19) **United States**

(12) **Patent Application Publication**
Wang et al.

(10) **Pub. No.: US 2018/0224448 A1**

(43) **Pub. Date: Aug. 9, 2018**

(54) **METHODS FOR DETECTING NOVEL
AUTOANTIBODIES IN CROHN'S DISEASE**

(71) Applicants: **ARIZONA BOARD OF REGENTS
ON BEHALF OF ARIZONA STATE
UNIVERSITY**, Scottsdale, AZ (US);
**MAYO FOUNDATION FOR
MEDICAL EDUCATION AND
RESEARCH**, Rochester, MN (US)

(72) Inventors: **Haoyu Wang**, New York, NY (US); **Ji
Qiu**, Chandler, AZ (US); **Joshua
LaBaer**, Chandler, AZ (US); **Jonathan
Leighton**, Scottsdale, AZ (US);
Shabana Pasha, Scottsdale, AZ (US)

(21) Appl. No.: **15/892,981**

(22) Filed: **Feb. 9, 2018**

Related U.S. Application Data

(60) Provisional application No. 62/457,040, filed on Feb. 9, 2017.

Publication Classification

(51) **Int. Cl.**
G01N 33/564 (2006.01)
G01N 33/50 (2006.01)

G01N 33/52 (2006.01)
G01N 33/543 (2006.01)
C12N 15/10 (2006.01)
C07K 16/18 (2006.01)

(52) **U.S. Cl.**
CPC **G01N 33/564** (2013.01); **G01N 33/5023**
(2013.01); **G01N 33/521** (2013.01); **G01N**
33/54353 (2013.01); **G01N 33/5436** (2013.01);
C12N 2310/3519 (2013.01); **C12N 15/1065**
(2013.01); **C07K 16/18** (2013.01); **G01N**
2800/065 (2013.01); **C07K 2317/30** (2013.01);
G01N 33/54393 (2013.01)

(57) **ABSTRACT**

Identification of autoantibodies associated with Crohn's disease useful in diagnosis and management using an innovative protein array technology, namely nucleic acid programmable protein arrays (NAPPA) and applications relating thereto. Overall, reactivity of IgG autoantibodies was stronger than that of IgA autoantibodies; however, IgA autoantibodies showed greater differential reactivity between cases and controls. Four IgA autoantibodies against SNRPB, PRPH, PTTG1 and SNAI1 were newly identified with sensitivities above 15% at 95% specificity, among which anti-SNRPB-IgA had the highest sensitivity of 24.0%. Autoantibodies associated with specific disease subtypes were also found.

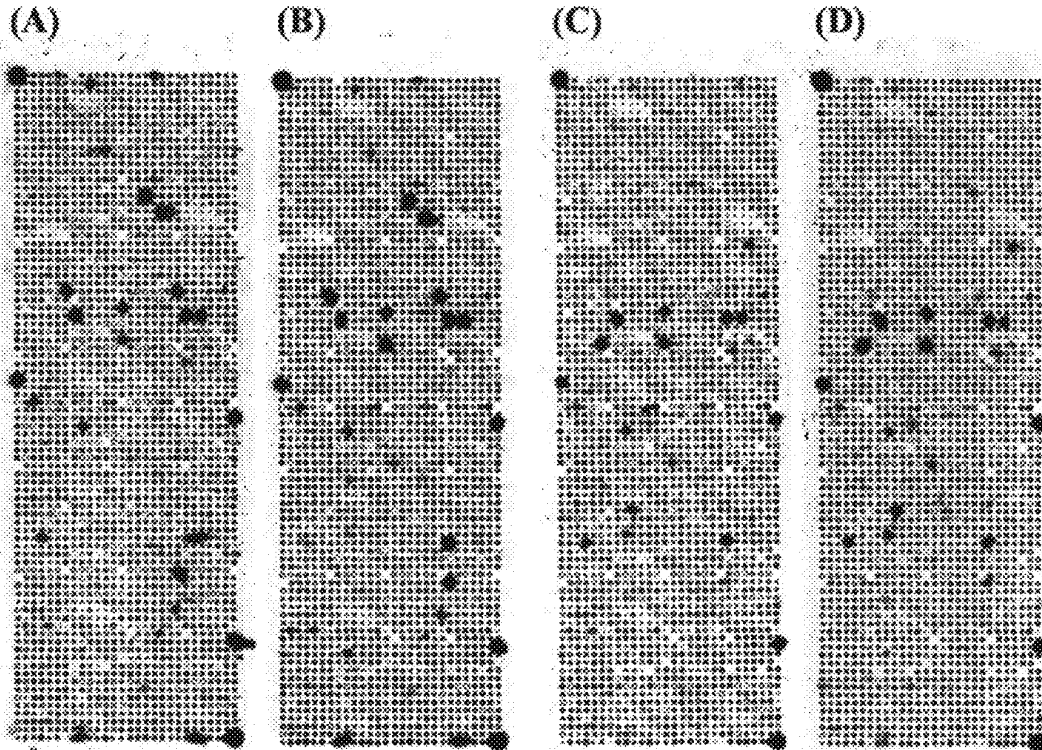


Figure 1:

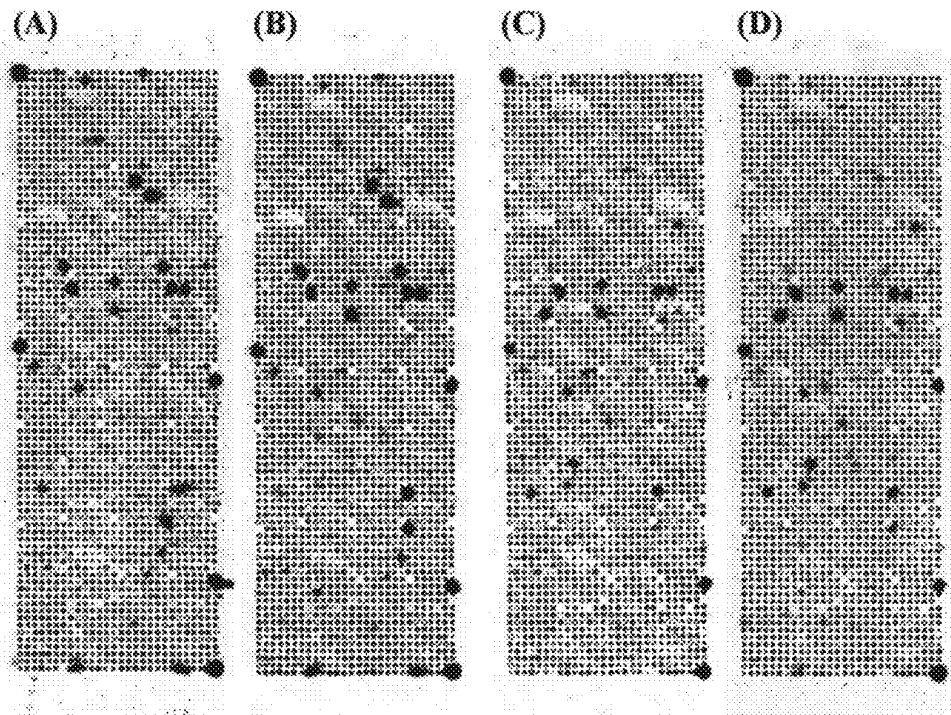


Figure 2:

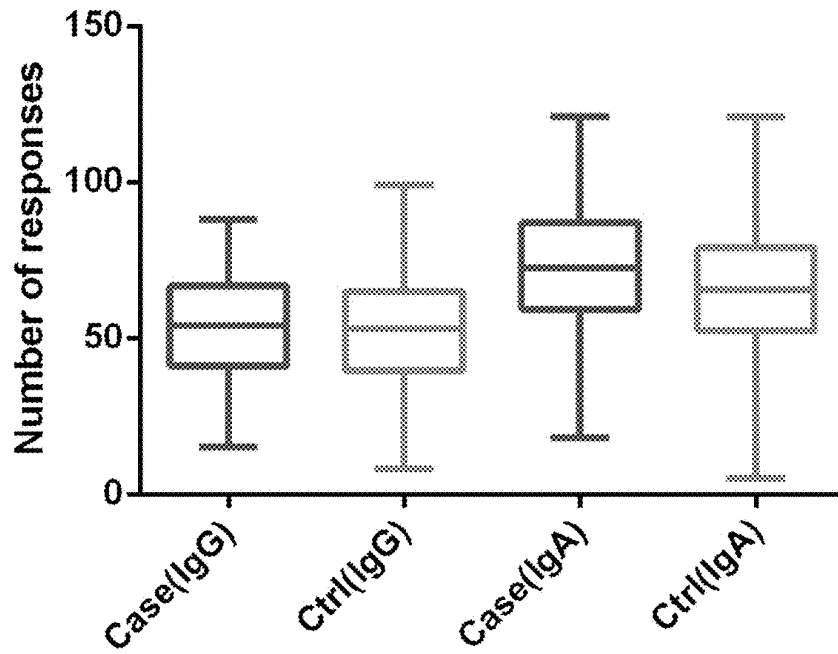


Figure 3:

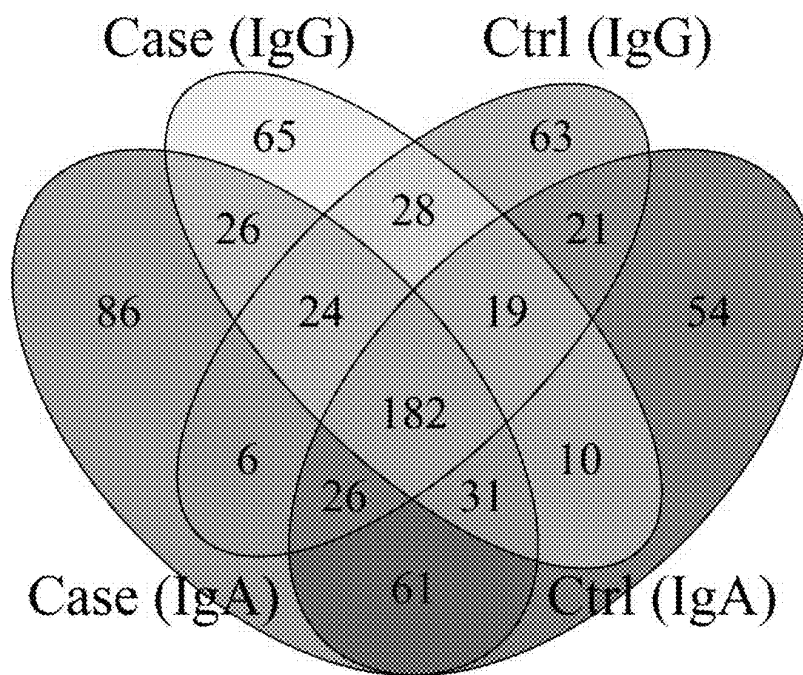


Figure 4:

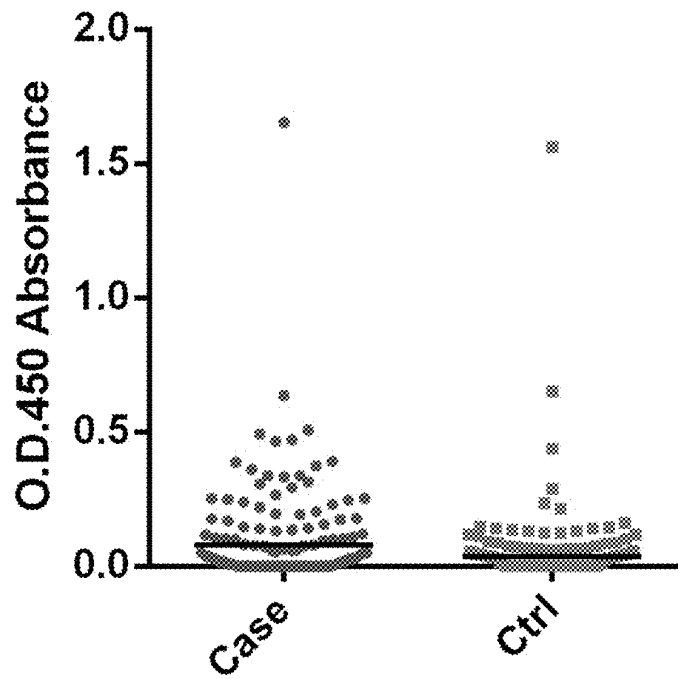


Figure 5:

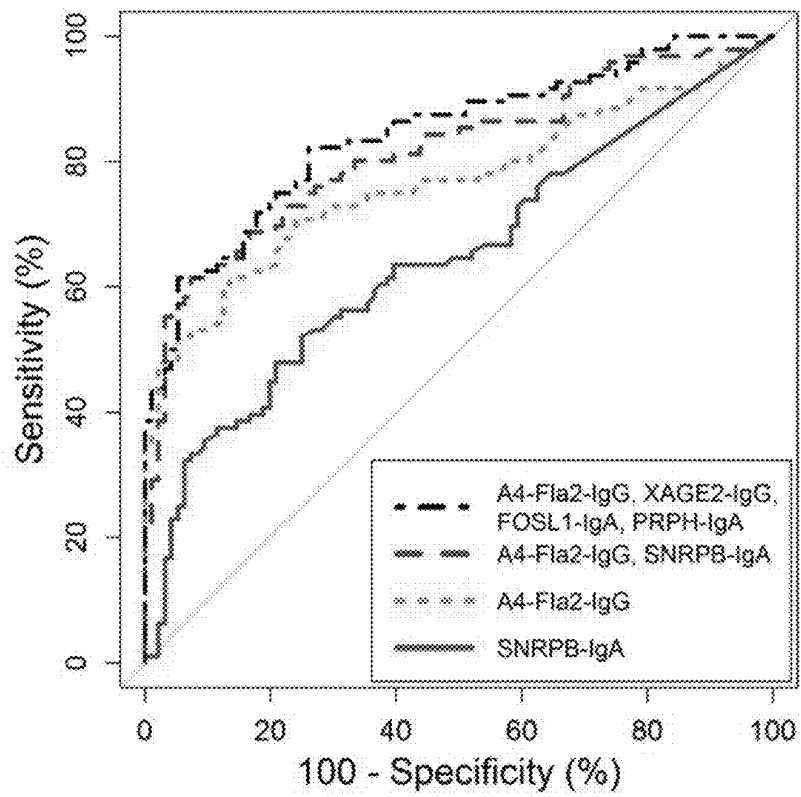


Figure 6

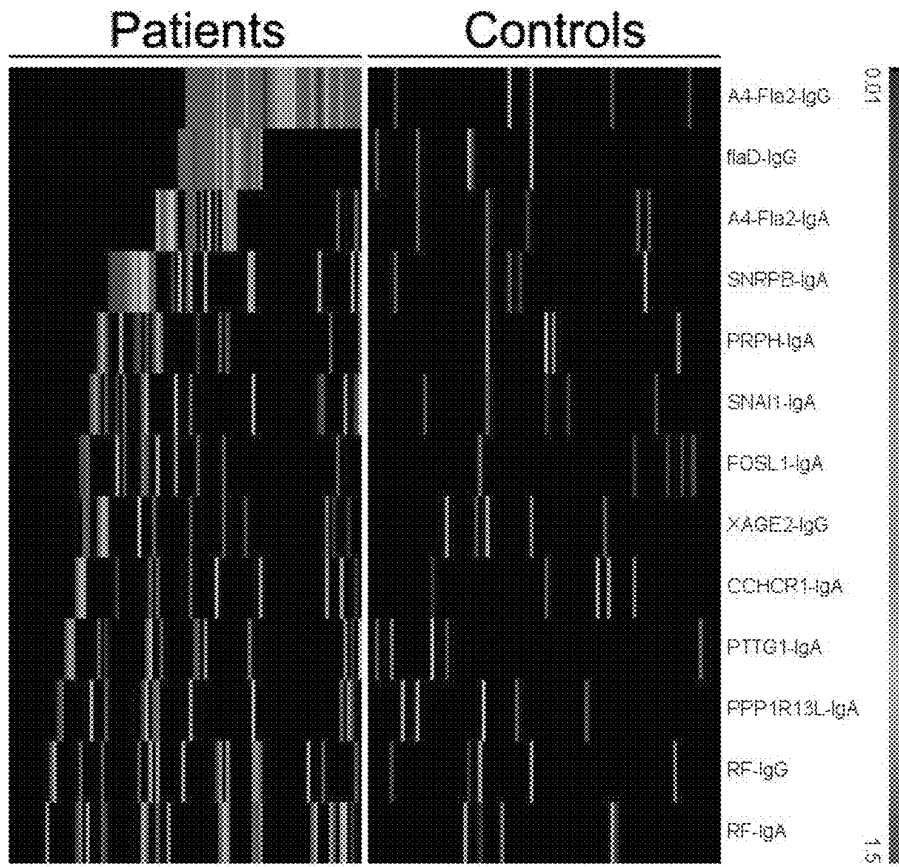
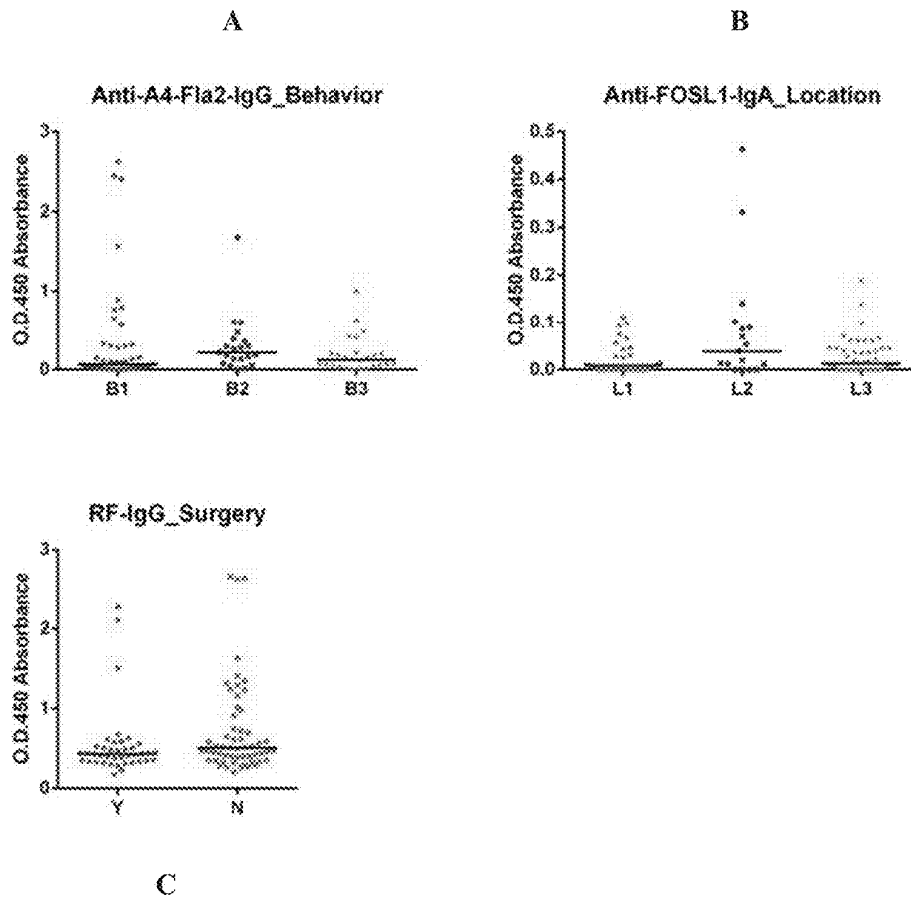


Figure 7:



METHODS FOR DETECTING NOVEL AUTOANTIBODIES IN CROHN'S DISEASE

CROSS-REFERENCE TO RELATED APPLICATIONS

[0001] This application is based on, claims a priority benefit from, and incorporates herein by reference, U.S. Provisional Patent Application No. 62/457,040, filed Feb. 9, 2017, and entitled "Novel Autoantibodies in Crohn's Disease."

FIELD OF THE INVENTION

[0002] This disclosure relates to providing compositions and methods for detecting Crohn's disease.

BACKGROUND

[0003] Crohn's disease (CD) is a form of inflammatory bowel disease (IBD) characterized by chronic transmural inflammation and skip lesions that can involve any segment of the gastrointestinal tract. The prevalence of CD has been increasing in recent years. An early accurate detection, determination of prognosis, as well as predicting appropriate therapy are therefore important to curb such increased prevalence.

[0004] The pathophysiology of CD is likely related to genetic, environmental and immune system factors. With regard to the immune system, there is evidence to suggest that an overactive intestinal mucosal immune system may be driven at least in part by a reaction to normal/altered luminal flora and autoantigens (AAGs). One manifestation of such a dysregulated immune response is the presence of CD-specific antibodies against gut microorganisms and AAGs. These include antibodies against microbial antigens (e.g., Flagellin, I2, *Saccharomyces cerevisiae* (ASCA)) and autoantibodies (AABs) such as anti-neutrophil cytoplasmic antibodies (ANCA) and pancreatic antibodies (PAb). To date, the presence of specific antibodies in the serum of patients with CD may help in detection, prognosis, and predicting response to therapy. However, most clinically used serological antibodies were discovered years ago and have a limitation by frequent false positives and an inability to accurately predict disease course. Therefore, there is a need for better diagnostic and prognostic CD markers.

[0005] Early studies on AABs in CD patients used immunofluorescence or whole cell ELISA assays on neutrophils, pancreas, goblet cells and colonic epithelial cells probed with patient sera. Early studies on AABs in CD patients used immunofluorescence or whole cell ELISA assays on neutrophils, pancreas, goblet cells and colonic epithelial cells probed with patient sera. Although whole-cell signals have been observed in CD for decades, the antigens recognized by these immunofluorescence assays remain largely unknown. This approach has proven to be ineffective, only two AAB targets of PAb have been discovered through mass spectrometry identification of reactive western blot spots. Challenges associated with methods that involve western blot include low throughput, inadequate quantification, and poor reproducibility. While all these technical challenges can be overcome, the level of effort needed to thoroughly analyze even a single sample prevents this approach from use in large clinical sample sets.

[0006] Progress in identifying individual autoantigens and autoantibodies in Crohn's disease has been challenging due to limitations of available immunoassays.

SUMMARY

[0007] Embodiments of the current disclosure describe methods for detecting an antigen-antibody complex of interest in a disease. The method comprises the steps of disposing a plurality of DNA segments corresponding to different open reading frames of genes on an array, wherein each DNA segment comprises a fusion tag; disposing a plurality of antibodies raised against the fusion tag on the array; translating the plurality of DNA segments to a plurality of polypeptides, wherein each polypeptide comprises the fusion tag and is captured by the antibodies raised against the fusion tag; and incubating the array with the plurality of captured polypeptides with samples collected from subjects having the disease and from healthy subjects to thereby form polypeptide/antibody complexes.

[0008] Further, in certain embodiments, the method comprises the steps of providing a plurality of antigens, wherein the antigens are selected from the group consisting of A4-Fla2, flaD, XAGE2, CDC42EP1, PPP1R13L, FOSL1, SNRPB, CCHCR1, LCORL, NUPL1, PRPH, PTTG1, SNAI1, and ORC2L; contacting the plurality of antigens with a sample from a subject containing a plurality of antibodies under conditions suitable to bind to the plurality of antigens to form antigen-antibody complexes; and determining the level or concentrations of said complexes.

BRIEF DESCRIPTION OF THE DRAWINGS

[0009] The technology disclosed herein will be better understood from a reading of the following detailed description taken in conjunction with the drawings in which like reference designators are used to designate like elements, and in which:

[0010] FIG. 1 shows reproducibility of serological profiling on nucleic acid programmable protein arrays (NAPPA). Slide images showed the IgG antibody reactivity toward a pool of 48 samples from CD patients (panels A and B) and a pool of 48 healthy controls (panels C and D) probed on arrays on two different days. Rings indicate strong seroreactivity.

[0011] FIG. 2 is a boxplot of number of proteins that showed antibody reactivity for samples in the discovery set on arrays. Number of responses for controls ranged between 8-99 for IgG and 5-121 for IgA. Number of responses for cases ranged between 15-88 for IgG and 18-121 for IgA. The p-values from student's t-test between CD patients and healthy controls are 0.08 for IgA and 0.42 for IgG.

[0012] FIG. 3 illustrates a Venn diagram of the number of antigens that reacted with at least two subjects (case or ctrl) on the array for both IgG and IgA antibody classes.

[0013] FIG. 4 is a jitter plot of anti-SNRPB-IgA antibody response measured by ELISA in the entire sample set. The black line indicates the median.

[0014] FIG. 5 shows ROC analysis of individual antibodies and antibody panels obtained from logistic regression models.

[0015] FIG. 6 is a heat map of antibody responses of novel AABs (>10% sensitivity at 95% specificity) and RFs showing complementarity to anti-flagellin antibodies.

[0016] FIG. 7 illustrates Jitter plots of antibody responses measured by ELISA in cases showing differences in different categories based on patients' clinical parameters: panel (A) Anti-A4-Fla2-IgG measured in patients with different disease behavior B1 (non-stricturing, non-penetrating), B2 (stricturing) and B3 (penetrating); panel (B) Anti-FOSL1-IgA measured in patients with different disease location L1 (ileal), L2 (colonic) and L3 (ileocolonic); panel (C) RF-IgG measured in patients with and without surgery. The black line indicates the median.

DETAILED DESCRIPTION

[0017] This technology disclosed herein is described in one or more exemplary embodiments in the following description with reference to the Figures, in which like numbers represent the same or similar elements. Reference throughout this specification to "one embodiment," "an embodiment," or similar language means that a particular feature, structure, or characteristic described in connection with the embodiment is included in at least one embodiment of the present technology disclosed herein. Thus, appearances of the phrases "in one embodiment," "in an embodiment," and similar language throughout this specification may, but do not necessarily, all refer to the same embodiment.

[0018] The described features, structures, or characteristics of the technology disclosed herein may be combined in any suitable manner in one or more embodiments. In the following description, numerous specific details are recited to provide a thorough understanding of embodiments of the technology disclosed herein. One skilled in the relevant art will recognize, however, that the technology disclosed herein may be practiced without one or more of the specific details, or with other methods, components, materials, and so forth. In other instances, well-known structures, materials, or operations are not shown or described in detail to avoid obscuring aspects of the technology disclosed herein.

[0019] The current technology discloses a high-throughput and unbiased method to test and identify different candidate antigens in Crohn's disease. Protein arrays have enabled profiling the antibody immune response in diseases that antibodies can be tested against thousands of proteins in parallel. FAM84A was identified as a novel AAg in IBD using protein arrays. However, conventional ways of making protein arrays require purification of a large number of proteins, making them expensive to produce and difficult to apply to large sample numbers.

Definitions

[0020] The term "individual," "subject," or "patient" typically refers to humans, but also to other animals including, e.g., other primates, rodents, canines, felines, equines, ovines, porcines, and the like.

[0021] The term "sample" includes any biological specimen obtained from an individual. Suitable samples for use in the present invention include, without limitation, whole blood, plasma, serum, saliva, urine, stool (i.e., feces), tears, and any other bodily fluid, or a tissue sample (i.e., biopsy) such as a small intestine or colon sample, and cellular extracts thereof (e.g., red blood cellular extract). In a preferred embodiment, the sample is a blood, plasma, or serum sample. In a more preferred embodiment, the sample is a serum sample.

[0022] The term "capture antibody" is intended to include an immobilized antibody which is specific for (i.e., binds, is bound by, or forms a complex with) one or more analytes of interest in a sample. In particular embodiments, the capture antibody is restrained on a solid support in an array.

[0023] The term "sensitivity" means the proportion of patients who are correctly categorized as having disease among those who truly have the disease. Similarly, the term "specificity" means the proportion of patients who are correctly categorized as not having the disease among all patients who truly don't have the disease.

[0024] An innovative protein array system termed NAPPA is disclosed here to overcome the above identified problems. NAPPA circumvents many of the limitations of traditional protein microarrays, and entails programming cell free protein expression extracts with cDNAs to express the proteins in situ at the time of the assay without the need for advanced protein purification. A key advantage is that human proteins are produced by human ribosomes and chaperone proteins to increase the efficiency of translation and to encourage natural folding of the proteins. Also, antibodies of both IgG and IgA classes can be probed simultaneously by using corresponding secondary antibodies conjugated with different fluorophores.

[0025] Protein arrays were produced displaying 1906 unique human proteins encoded by genes randomly selected from our human gene collections at DNASU (DNASU.org), and 3 known bacterial flagellins as positive controls, A4-Fla2 and Fla-X, and CBir1. We also included 9 bacterial flagellin genes in our collection, flaA, flaB, flaC, flaD, flaE, flaG from *Vibrio cholera*, flgG, flgH, and flgK from *Pseudomonas aeruginosa*. Arrays were quality controlled for protein display. Forty-eight (48) CD patients and 48 age-gender matched healthy controls (Table 1) were profiled for their antibody responses against these 1906 proteins. Antibodies of both IgA and IgG classes were assayed simultaneously using secondary anti-human IgA and anti-human IgG antibodies labeled with different fluorophores. An equal-volume pool of all 48 patient samples and a pool of all 48 controls were tested every experiment day to assure the quality. The correlations between two arrays profiled with the same pooled samples were generally above 0.85. Epstein Barr Nuclear antigen (EBNA) was printed on our arrays as an experimental positive control because more than 90% adults in the United States have been infected by the Epstein Barr virus. Eighty-nine (89) out of the 96 subjects assayed on the arrays showed IgG reactivity against EBNA. Sixty-eight (68) showed much weaker IgA reactivity against EBNA.

[0026] Table 1 lists clinical information and characteristics of human subjects.

TABLE 1

| | Clinical information of the samples | | | |
|----------------------|-------------------------------------|------------------|----------------|------------------|
| | Discovery set | | Validation set | |
| | CD patients | Healthy controls | CD patients | Healthy controls |
| N (sera) | 48 | 48 | 48 | 48 |
| Gender (male/female) | 21/27 | 21/27 | 21/27 | 21/27 |
| Age, yrs (mean) | 44 ± 17 | 44 ± 17 | 40 ± 16 | 40 ± 16 |
| Age, yrs (range) | 20-77 | 20-77 | 18-74 | 18-74 |

TABLE 1-continued

| Clinical information of the samples | | | | |
|-------------------------------------|---------------|------------------|----------------|------------------|
| | Discovery set | | Validation set | |
| | CD patients | Healthy controls | CD patients | Healthy controls |
| Age at diagnosis (A1/A2/A3) | 1/24/23 | | 1/29/18 | |
| Location of diseases (L1/L2/L3/L4) | 15/9/24/0 | | 14/8/26/0 | |

[0028] Referring to FIG. 2, the average numbers of human proteins showing IgG antibody reactivity on arrays were 53 ± 19 and 53 ± 20 for cases and controls, respectively. There was no difference between cases and controls ($p=0.42$) for IgG antibody. However, CD patients showed greater IgA antibody response against proteins on arrays. The average number of proteins showing IgA reactivity for cases (73 ± 22) was greater than that for controls (66 ± 24), although it did not reach statistical significance ($p=0.08$) (FIG. 2). From the perspective of individual proteins (Table 2), 501 proteins reacted with at least two patients or controls for IgG antibody. When analyzing case and control samples separately, 385 antigens reacted with at least two cases and 369 with at least two controls. There was a big overlap of 253 proteins that reacted with both CD cases and controls (FIG. 3).

TABLE 2

| Number and percentage of human proteins which showed positive antibody response on arrays in the discovery sample set with different criteria. | | | | | | |
|--|---------------------------------|----------------------|-------------------------|----------------------------------|----------------------------------|---|
| | At least two in case or control | At least two in case | At least two in control | Two more in case than in control | Two more in control than in case | At least 15% sensitivity at 95% specificity |
| IgG | 501 (26.3%) | 385 (20.2%) | 369 (19.4%) | 156 (8.2%) | 146 (7.7%) | 13 (0.7%) |
| IgA | 546 (28.6%) | 442 (23.2%) | 404 (21.2%) | 217 (11.4%) | 135 (7.1%) | 19 (1.0%) |

TABLE 1-continued

| Clinical information of the samples | | | | |
|-------------------------------------|---------------|------------------|----------------|------------------|
| | Discovery set | | Validation set | |
| | CD patients | Healthy controls | CD patients | Healthy controls |
| Behavior (B1/B2/B3) | 23/12/13 | | 26/9/13 | |
| Surgery (Y/N) | 18/30 | | 22/26 | |

Age at diagnosis;
 A1: 0-16 yrs,
 A2: 16-40 yrs,
 A3: >40 yrs.
 Location of disease;
 L1: ileal,
 L2: colonic,
 L3: ileocolonic,
 L4: isolated upper digestive.
 Behavior;
 B1: non-stricturing, non- penetrating;
 B2: stricturing,
 B3: penetrating.

[0027] Referring to FIG. 1, a high-throughput study using a protein array screens all proteins simultaneously and is probed with samples from Crohn's disease and controls to identify novel autoantigens against antibodies that are present in samples. Studies that use an immunoproteomics approach to identify CD-specific AAbs of both IgG and IgA classes were conducted. The current NAPPA technology was applied to screen about 1900 human proteins in CD patients and age- and gender-matched controls. Candidate AAbs identified from antibody profiling on protein arrays were first confirmed by Enzyme Linked Immunosorbent Assay (ELISA) and further validated by ELISA in an independent sample set. With the advantage of NAPPA, especially in obtaining thousands of "fresh" proteins with natural folding on the array, characteristics of the humoral immune response were identified, which may have been missed before.

[0029] Bacterial flagellin antibodies showed the strongest reactivity and the highest prevalence on arrays. Quantitatively, anti-A4-Fla2-IgG achieved a sensitivity of 46% at 95% specificity among the 48 pairs of samples profiled on arrays, which matched the performance reported in the literature. Anti-A4-Fla2-IgA did not differentiate the cases from control as well as IgG, with a sensitivity of 19% at 95% specificity, which was also comparable to the literature. Other known flagellin antibodies anti-Fla-X-IgG and anti-CBir1-IgG showed very similar performance to anti-A4-Fla2-IgG. Thus, only A4-Fla2 was included for the subsequent experiments.

[0030] All novel AAbs discovered on arrays were ranked according to their sensitivity at 95% specificity for IgG and IgA antibodies respectively based on array data. Nine (9) IgG candidate antibodies and 19 IgA candidate antibodies that had at least 10% sensitivity at 95% specificity were selected (Table 3), together with two flagellins (A4-Fla2 and flaD). All these antibodies were verified by ELISA using the discovery sample set listed in Table 1. Three (3) (out of 9) of the IgG candidates and 14 (out of 19) IgA candidates had comparable performance on ELISA and were selected for further validation (Table 3).

TABLE 3

| List of candidate antibodies selected by protein array and their sensitivity at 95% specificity based on ELISA result. | | | | |
|--|----------|------------------------------------|----------------|------------|
| Antibody type | Antigen | Sensitivity (%) at 95% specificity | | |
| | | Discovery Set | Validation Set | Entire Set |
| IgG | A4-Fla2 | 50 | 50 | 50 |
| | flaD | 29 | 17 | 24 |
| | WWTR1 | 6 | | N/A |
| | XAGE2 | 29 | 2 | 14 |
| | RAD51AP1 | 8 | | N/A |

TABLE 3-continued

| List of candidate antibodies selected by protein array and their sensitivity at 95% specificity based on ELISA result. | | | | |
|--|----------|------------------------------------|----------------|------------|
| Antibody | | Sensitivity (%) at 95% specificity | | |
| type | Antigen | Discovery Set | Validation Set | Entire Set |
| IgA | STAU1 | 6 | | N/A |
| | MAP7 | 8 | | N/A |
| | CDC42EP1 | 10 | 10 | 10 |
| | PML | 10 | 4 | 7 |
| | FOS | 4 | | N/A |
| | KLF4 | 8 | | N/A |
| | RF | 21 | 15 | 18 |
| | A4-Fla2 | 21 | 21 | 20 |
| | PPP1R13L | 25 | 6 | 14 |
| | FOSL1 | 13 | 23 | 13 |
| | MACC1 | 4 | | N/A |
| | RAD51AP1 | 21 | 0 | 7 |
| | NUFIP2 | 13 | 23 | 8 |
| | SNRBP | 25 | 19 | 24 |
| | ZGPAT | 2 | | N/A |
| | CCHCR1 | 15 | 10 | 13 |
| | LCORL | 10 | 10 | 10 |
| | NUPL1 | 19 | 4 | 10 |
| | PAPOLG | 6 | | N/A |
| | PRPH | 13 | 15 | 16 |
| | PTTG1 | 17 | 8 | 16 |
| | SNAI1 | 10 | 27 | 17 |
| | C1orf106 | 8 | | N/A |
| | FILIP1L | 13 | 4 | 7 |
| NEUROD6 | 15 | 2 | 1 | |
| ORC2L | 15 | 8 | 10 | |
| STAU1 | 8 | | N/A | |
| RF | 33 | 15 | 24 | |

[0031] We further validated the performance of the 3 IgG and 14 IgA AAGs showing differential antibody responses between cases and controls by ELISA using an independent sample set (validation sample set, Table 1). Validation was also performed on one known flagellin (A4-Fla2) and one novel flagellin (flaD). IgA AAbs against small nuclear ribonucleoprotein-associated proteins B and B' (SNRBP), peripherin (PRPH), securin (PTTG1) and zinc finger protein SNAI1 (SNAI1) had sensitivity above 15% at 95% specificity in the entire sample set (Table 3). Among the newly identified AAb, anti-SNRBP IgA had the highest sensitivity of 24.0% (FIG. 4). We also estimated the anti-SNRBP-IgA titers for all samples by fitting their ELISA absorbance data onto a standard curve generated using serially diluted commercially available human IgA. Significant difference ($p=0.01$) was also observed between cases and controls based on their anti-SNRBP-IgA concentrations.

[0032] Novel autoantibodies such as anti A4-Fla2, flaD, XAGE2, and CDC42EP1 were identified for IgG antibody with sensitivity of 50%, 24%, 14%, and 10% respectively at 95% specificity. We also identified novel autoantibodies for IgA antibody for anti A4-Fla2, PPP1R13L, FOSL1, SNRBP, CCHCR1, LCORL, NUPL1, PRPH, PTTG1, SNAI1, and ORC2L with sensitivity of 20%, 14%, 13%, 24%, 13%, 10%, 10%, 16%, 16%, 17%, and 10% respectively at 95% specificity.

[0033] We also assayed one special type of AAb, rheumatoid factor (RF), in CD patients and controls. All three classes of RFs, IgG, IgA and IgM, were assayed by ELISA in all 96 cases and controls. Similar to those reported in the literature, at 95% specificity, the sensitivity of the RF-IgA was 24.0%. IgM only had a sensitivity of 7.3%, while IgG had a sensitivity of 17.7%. They did show complementarity

in CD patients, but their combination did not yield a better performance than that for RF-IgA alone.

[0034] A receiver operating characteristic (ROC) analysis is applied to the antibodies against A4-Fla2, RF and all AAbs with a performance greater than 10% sensitivity at 95% specificity in the full sample set. Anti-A4-Fla2-IgG has the best performance in the 96 pairs of samples with 50% sensitivity at 95% specificity and an AUC of 0.767. The best novel AAb, anti-SNRBP-IgA, which is a core component of small nuclear ribonucleoprotein, provides a sensitivity of 24.0% at 95% specificity with an AUC of 0.645 (FIG. 5). The heat map (FIG. 6) showed the complementarity of the newly found antibodies and the RF to the known anti-A4-Fla2. Logistic regression was applied to these antibodies to generate panels with better performance. The logistic regression model of anti-A4-Fla2-IgG and anti-SNRBP-IgA had a sensitivity of 57.3% and increased the AUC to 0.817. The best logistic regression model included A4-Fla2-IgG, FOSL1-IgA (Fos-related antigen 1), PRPH-IgA (peripherin), and XAGE2-IgG (X antigen family member 2) and the sensitivity reached 60.4% at 95% specificity and AUC 0.844 (FIG. 4). The second best model included anti-SNRBP-IgA with a slightly higher AUC 0.852 and a slightly lower sensitivity of 59.4% at 95% specificity.

[0035] The association between different clinical parameters, including age, gender, disease behavior (B1: non-stricturing, non-penetrating; B2: stricturing; B3: penetrating) and location (L1: ileal; L2: colonic; L3: ileocolonic) based on Montreal classification, and surgery status, and antibody reactivity were also evaluated, (Table 4). None of the antibodies were significantly associated with any aspect of disease behavior. Anti-A4-Fla2 antibody has been reported to be associated with the stricturing B2 phenotype. Referring to panel A of FIG. 7, a significantly higher reactivity in B2 patients relative to B1 and B3 patients was not observed, although B2 patients did have a higher median reactivity. No significant association was observed between anti-SNRBP-IgA and patient sub-grouped by clinical parameters (data not shown). Referring to panel B of FIG. 7, anti-FOSL1-IgA was found to be significantly associated with colonic disease location (L2) ($p=0.004$, Table 4). RF-IgA also showed higher reactivity in patients with colonic disease (L2) relative to patients with ileal (L1) and ileocolonic (L3) diseases, but did not reach significance ($p=0.086$) (Table 4). Anti-PPP1R13L-IgA was negatively associated with ileal disease ($p=0.03$) (data not shown). Interestingly, referring to panel C of FIG. 7, patients without surgery had significantly higher RF-IgG relative to those with surgery ($p=0.017$). It is not clear whether the decreased RF is a consequence of the surgery and other related treatment or a sign of improvement in inflammation and healing.

TABLE 4

| P-values of the hypothesis tests of association between the antibody levels and the clinical status. | | | | |
|--|----------|----------|--------|---------|
| Antibody | Severity | Location | Gender | Surgery |
| A4-Fla2-IgG | 0.536 | 0.568 | 0.685 | 1.000 |
| flaD-IgG | 0.685 | 0.507 | 0.610 | 0.202 |
| A4-Fla2-IgA | 0.286 | 0.593 | 0.873 | 0.311 |
| XAGE2-IgG | 0.212 | 0.308 | 0.854 | 0.361 |
| CDC42EP1-IgG | 0.150 | 0.272 | 0.682 | 0.911 |
| PPP1R13L-IgA | 0.643 | 0.192 | 0.446 | 0.731 |
| FOSL1-IgA | 0.828 | 0.004 | 0.650 | 0.191 |

TABLE 4-continued

| P-values of the hypothesis tests of association between the antibody levels and the clinical status. | | | | |
|---|----------|----------|--------|---------|
| Antibody | Severity | Location | Gender | Surgery |
| SNRPB-IgA | 0.623 | 0.949 | 0.318 | 0.202 |
| PTTG1-IgA | 0.458 | 0.478 | 0.371 | 0.888 |
| PRPH-IgA | 0.621 | 0.217 | 0.751 | 0.471 |
| CCHCR1-IgA | 0.637 | 0.785 | 0.650 | 1.000 |
| LCORL-IgA | 0.806 | 0.540 | 0.297 | 0.587 |
| NUPL1-IgA | 0.727 | 0.349 | 0.344 | 0.417 |
| SNAIL-IgA | 0.634 | 0.849 | 0.581 | 0.346 |
| ORC2L-IgA | 0.931 | 0.394 | 0.682 | 0.587 |
| RF-IgG | 0.918 | 0.792 | 0.184 | 0.017 |
| RF-IgA | 0.737 | 0.086 | 0.170 | 0.440 |

[0036] There are two possible functional consequences of autoantibodies in CD. One is that these antibodies reflect dysregulated immune response in the gut. It is also possible that these autoantibodies could play a pathogenic role. However, the few autoantibodies discovered so far in CD limit our exploration of their functional consequences. Our establishment of a NAPPA based immunoproteomics platform should promote such discoveries.

[0037] Although the identified candidates warrants future investigation to understand the physiological relevance, they highlight the need for identifying more individual antigens in Crohn's Disease. Future studies will need to be performed on larger human proteome and CD-related microbial organism proteomes using patient samples with accurate clinical information. This work will not only yield markers that can improve our understanding of disease pathogenesis but also facilitate better management of CD patients.

[0038] Several AAbs have been reported in CD before. Anti-neutrophil cytoplasmic antibodies (ANCA) have been found in 5-31% of patients with CD. In addition to ANCA, other AAbs have been identified in CD, such as antibodies against the exocrine pancreas (PAb). Long after the initial observation of ANCA and PAb by immunofluorescence, the advancement of mass spectrometry technology led to the identification of CUZD1 and GP2 as two contributing AAbs for PAb from reactive western blot bands. However, few studies have applied high-throughput quantitative immunoproteomics technologies in the study of serological markers in CD or IBD, in general. Our study used NAPPA that enabled proteome-level study of antibody responses without the need to purify individual proteins. We have previously applied NAPPA to the study of disease specific antibodies in various diseases. In this study, we applied NAPPA, an innovative protein array technology, to identify AAbs in CD. To our knowledge, this is one of the first studies that applied an immuno-proteomics approach in the identification of specific AAbs in CD.

[0039] We discovered a set of AAbs that had higher prevalence in CD cases compared to controls on arrays. We also validated 8 AAbs in an independent set of samples by ELISA. The classification performance of these AAb markers ranged from 13-24% sensitivity at 95% specificity. Anti-SNRPB-IgA had the best performance with a sensitivity of 24.0% at 95% specificity. Overall, the performances of these new antibodies were comparable to those of reported AAbs in CD. For example, anti-GP2-IgG was reported to have a sensitivity of 10.2% at 99% specificity by ELISA, and anti-GP2-IgA had a sensitivity of 3.4% at 99% specificity. Anti-FAM84A-IgG had a sensitivity of 18% at 95%

specificity in CD. These novel AAbs could modestly complement the known anti-A4-Fla2-IgG flagellin antibody. This demonstrates that by using novel approaches it is possible to find new AAbs that have activities comparable to those previously identified.

[0040] Serologic markers have the potential to improve the diagnostic process for CD patients and minimize the need for invasive diagnostic procedures. The Prometheus IBD serology 7 panel includes AAb and anti-microbial antigen antibody markers and uses a sophisticated algorithm to evaluate the serologic test results. It is used clinically to aid IBD diagnosis; however, the panel lacks sensitivity and specificity, and has been shown to have a lower predictive value than routine laboratory testing in the diagnosis of IBD. Identification of novel CD-specific antibodies could improve the performance of currently available markers and expand the clinical utility spectrum of serological markers beyond diagnosis. Our study focused on AAbs, which have not been previously studied as extensively as anti-microbial antigen antibodies.

[0041] In view of the above, one point of novelty is the identification of the autoantigens against which antibodies were detected as candidates in CD (Table 3). Many of them have not been previously associated with CD.

Non-Limiting Examples

Characteristics of Plasma Samples

[0042] All samples were collected with written informed consent, and the study was approved by institutional review boards at Arizona State University and Mayo Clinic. Ninety-six (96) CD serum samples from patients within four years of diagnosis were selected from the Mayo Clinic Inflammatory Bowel Disease Biobank. Characteristics of CD patients including age, gender, Montreal Classification and surgery status are shown in Table 1. Ninety-six age- and gender-matched healthy control samples without a history of inflammatory bowel disease, celiac sprue or irritable bowel syndrome were selected from the Mayo Cancer Center/Normal Serum Biobank at Mayo Clinic Rochester. The samples were evenly split into discovery and validation sets randomly. The discovery set was used for array profiling as an initial screening. The validation set was used for validating the candidate AAbs selected from the screening.

Nucleic Acid Programmable Protein Array Production

[0043] Open reading frames were obtained from DNASU (dnasu.org). All genes of interest were cloned in the nucleic acid-programmable protein array (NAPPA) compatible expression vectors with a C-terminal glutathione S-transferase (GST) fusion tag were obtained from the DNASU plasmid repository. Plasmid DNA was spotted on glass slides coated with aminosilane (APTS) (Thermo Fisher Scientific) in a printing buffer consisting of anti-GST polyclonal antibody (GE Healthcare), BSA and bisulfosuccinimidyl suberate protein cross-linker (Thermo Fisher Scientific, Waltham, Mass.) using a QArrayer2 (Genetix). At the time of usage, proteins were expressed from spotted plasmid DNAs using a HeLa cell lysate-based in vitro transcription and translation (IVTT) protein expression system (Thermo Fisher scientific, Rockford, Ill.). Printed arrays were blocked with the SuperBlock buffer (Thermo Fisher Scientific, Waltham, Mass.) at room temperature (RT) for 1 hr with

gentle shaking. Arrays were rinsed 5 times with deionized (DI) water, dried by centrifugation at 1000 rpm for 1 minute at RT, and sealed with HybriWell (Grace BIO-LABS, Bend, Oreg.). A 155 μ L human IVTT mixture was injected into the HybriWell and incubated at 30° C. for 1.5 hrs for protein expression. Expressed GST-tagged proteins were captured in situ by the anti-GST antibodies co-printed on the glass slides and displayed for subsequent serological profiling.

Serum Antibody Profiling

[0044] Expressed NAPPA arrays were washed 3 times with 5% milk-PBST (PBS with 0.2% Tween), blocked in 5% milk-PBST for 1 hr, rinsed with DI water 10 times, and dried by centrifugation at 1000 rpm for 1 min at RT. Arrays were then placed in the hybridization chambers of the HS 4800™ Pro hybridization station (Tecan, Mannedorf, Switzerland) and programmed with 5% milk-PBST wash, 16 hr incubation of 160 μ L 1:20 diluted serum at 4° C., 5% milk-PBST wash, 1 hr incubation of 160 μ L 1:500 diluted Alexa Fluor 647® Goat Anti-Human IgG and 1:300 diluted Cy3 Goat Anti-Human IgA (Jackson ImmunoResearch Laboratories, West Grove, Pa.) and 5% milk-PBST wash. Slides were rinsed with DI water, dried by centrifugation and scanned by a Tecan scanner under consistent settings. A pooled sample was prepared by mixing equal volumes of samples from 48 patients and 48 controls. The pooled sample was run as a technical replicate on every serum screening day.

Candidate Selection

[0045] Strong antibody reactivity sometimes resulted in saturated signals of the local spot with diffusion to the neighboring spots, which was defined as a ring. Antibody reactivity was qualitatively determined based on the existence of rings. Quantitative array image analysis was performed based on the signal intensity around the spot as published before. The reference, Bian X, Wallstrom G, Davis A, et al. Immunoproteomic profiling of antiviral antibodies in new-onset type 1 diabetes using protein arrays. *Diabetes* 2016; 65:285-96, is incorporated herein in its entirety. The raw intensity was defined as the median intensity around the spot. Spots without an evident ring were assigned zero intensity. For each candidate, the 95 percentile of the intensity of the control samples was used as the cutoff, and all the candidates were ranked by their sensitivity according to the arbitrary cutoff. Those with higher than 15% sensitivity were selected and then confirmed by ELISA using the same sample set.

ELISA Assays

[0046] Rapid Antigenic Protein in situ Display (RAPID) ELISA was performed to assess antigen specific serum antibodies as described. Ninety-six (96)-well ELISA plates (Corning Life Sciences, Salt Lake City, Utah) were coated with 50 μ L 10 μ g/mL anti-GST antibody (GE Healthcare, Pittsburgh, Pa.) in a coating buffer (0.5 M carbonate bicarbonate buffer, pH 9.6) overnight at 4° C. On the next day, plates were washed three times with 100 μ L PBST and blocked with 100 μ L 5% milk-PBST for 1.5 hrs. In the meantime, candidate antigens were expressed in the human HeLa cell lysate-based protein expression system at 30° C. for 1.5 hr from 200 ng/ μ L antigen encoding plasmids (pANT7-cGST), which were the same as those used to express proteins on arrays that supported the expression of

proteins with a C-terminal GST tag. Expressed proteins were diluted in milk-PBST at 1:50 dilution. 50 μ L diluted antigen was added into each well to allow the GST-tagged antigen to be captured by the coated anti-GST antibody at RT for 1 hr on a shaker at 500 rpm. All the wells in the plates were washed 5 times with PBST, incubated with diluted serum (1:50 to 1:200 serum dilution for human protein and 1:500 to 1:10000 for bacterial protein) at RT for 1 hr, washed 5 times with PBST, and incubated with 50 μ L 1:10,000 diluted HRP labeled Goat Anti-Human IgG or Goat Anti-Human IgA (Jackson ImmunoResearch Laboratories, West Grove, Pa.) at RT for 1 hr on a shaker at 500 rpm. Finally, plates were washed and incubated with 1-Step Ultra TMB-ELISA Substrate (Thermo Fisher scientific, Waltham, Mass.) for 10 min for detection and 50 μ L 2M sulfuric acid was added to stop the reaction. O.D.450 was measured immediately by Envision® Multilabel Reader (PerkinElmer, Waltham, Mass.).

[0047] Plates coated with 50 μ L 50 μ g/mL goat IgG antibody (Jackson ImmunoResearch Laboratories, West Grove, Pa.) were used to assess rheumatoid factor (RF). After incubating with serum samples, plates were incubated with 1:10,000 diluted HRP labeled Goat-Anti-Human IgG, or IgA, or IgM to probe different classes of RF.

[0048] Plates coated with serially diluted human IgA (Jackson ImmunoResearch Laboratories, West Grove, Pa.) ranging from 0 to 1 μ g/ml were used to generate a standard curve for IgA ELISA. A four parameter logistic equation was used to fit the curve. Serum concentrations of anti-SNRPB-IgA antibody in μ g/mL were estimated based on the standard curve using the O.D.450 absorbance data.

Statistics and Data Analysis

[0049] The overall qualitative reactivity between cases and controls based on the array data was compared using descriptive statistics and one-tailed Student's t-test assuming equal variance with null hypothesis that cases have stronger antibody reactivity. The performance of the individual candidates and combined classifiers were assessed by ROC analysis. For the panel construction, the bestglm package in R was used to select logistic regression model based on Akaike information criterion (AIC) while only individual candidates with higher than 10% sensitivity at 95% specificity in the entire sample set were included. For the subtype analysis, two-tailed Student's t-test assuming unequal variance was used to exam the association between the antibody responses against different antigens and clinical information including gender and surgery status. One-way ANOVA was used for the association between the antibody responses against different antigens and non-binomial clinical information such as behavior and disease location. Student's t-test and Mann-Whitney U test were used to test the differences between patient subgroups. The heat map was generated in MultiExperiment based on the ELISA data. Graphs and plots were drawn in GraphPad or in R. Identification of Candidate Antigens Associated with Crohn's Disease

[0050] To identify CD-associated candidate Abs, we first performed comprehensive profiling of antibodies against 1906 unique human proteins in serum samples from forty-eight (48) CD patients and 48 age-gender matched healthy controls on NAPPA arrays. Based on array data, we selected the following antigens A4-Fla2, flaD, XAGE2, and CDC42EP1 for IgG antibody and for IgA antibody anti

A4-Fla2, PPP1R13L, FOSL1, SNRPB, CCHCR1, LCORL, NUPL1, PRPH, PTTG1, SNAI1, and ORC2L that showed high antibody reactivity on the NAPPA arrays (Table 3). Antibodies against SNRPB-IgA had the best performance with a sensitivity of 24.0% at 95% specificity.

Validation of Candidate Antigens Associated with Crohn's Disease

[0051] To validate these novel antibodies, we performed a validation study of responses in ELISA to A4-Fla2, flaD, XAGE2, and CDC42EP1 for IgG antibody and for IgA antibody anti A4-Fla2, PPP1R13L, FOSL1, SNRPB, CCHCR1, LCORL, NUPL1, PRPH, PTTG1, SNAI1, and ORC2L in an independent set of 96 serum samples from groups of subjects with two different clinical characteristics: CD patients (48) and healthy controls (48) (Table 1). We observed significant differences for the antigens' reactivity between the CD patients and healthy control groups (Table 3, FIGS. 2, 4, and 5).

[0052] Using an immuno-proteomics approach, we applied serum samples to arrays displaying thousands of proteins in parallel. We discovered a set of AAbs that had higher prevalence in CD cases compared to controls on arrays. We also validated 8 AAbs in an independent set of samples by ELISA. The classification performance of these AAb markers ranged from 13-24% sensitivity at 95% specificity.

[0053] Serologic markers have the potential to improve the diagnostic process for CD patients and minimize the need for invasive diagnostic procedures. The Prometheus IBD serology 7 panel includes AAb and anti-microbial antigen antibody markers and uses a sophisticated algorithm to evaluate the serologic test results. It is used clinically to aid IBD diagnosis; however, the panel lacks sensitivity and specificity, and has been shown to have a lower predictive value than routine laboratory testing in the diagnosis of IBD. Identification of novel CD-specific antibodies could improve the performance of currently available markers and expand the clinical utility spectrum of serological markers beyond diagnosis. Our study focused on AAbs, which have not been previously studied as extensively as anti-microbial antigen antibodies.

[0054] Naturally folded protein arrays like NAPPA aid in the discovery of antibodies against human proteins in CD. We profiled fewer than 2000 human proteins in this study, which represents a small fraction of human proteins and microbial proteins in our collections. The success of this study suggests excellent outcomes for future expanded studies. Furthermore, we can produce protein arrays with any cDNA sequences at a low cost. This flexibility of NAPPA combined with next gen sequencing technologies make it uniquely possible to study antibody responses to antigens encoded CD related microorganisms as they are discovered by microbiome studies in the future. It is believed that no single antigen will meet all clinical needs. Achieving the best performance will rely on a panel of antibodies. An immunoproteomics approach using protein arrays will facilitate the discovery of additional antibodies, creating the opportunity to combine them into panels that can benefit the clinical management of CD patients. Our data supported this assumption.

[0055] SNRPB is the core component of spliceosomal small nuclear ribonucleoproteins (snRNPS) and plays an important role in pre-mRNA splicing. SNRPB is an AAg in Systemic Lupus Erythematosus (SLE) patients originally

identified as the 'smith antigen' (Sm). Whereas reactivity to this antigen in SLE is primarily of the IgG subtype, we observed that in CD it is primarily IgA. It would be of great interest to assay anti-SNRPB-IgA and anti-SNRPB-IgG in SLE and CD patients to further confirm antibody subtype specificity in these two diseases.

[0056] There are two possible functional consequences of autoantibodies in CD. One is that these antibodies reflect dysregulated immune response in the gut. It is also possible that these autoantibodies could play a pathogenic role. However, the few autoantibodies discovered so far in CD limit our exploration of their functional consequences. Our establishment of a NAPPA based immunoproteomics platform should promote such discoveries. In the case of the novel autoantibodies reported in our study, future validation and association studies with clinical parameters in a larger sample set, assays of their titer changing over time during CD development, and their expression patterns in the gut, should shed light on their role of autoimmunity.

[0057] We also calculated the sensitivities and specificities of all candidate antigens that went into the validation study. In summary, we have demonstrated that serological profiling on NAPPA in the discovery of AAbs in CD has great potential. Future studies will need to be performed on larger human proteome and CD-related microbial organism proteomes using patient samples with accurate clinical information. This work will not only yield markers that can improve our understanding of disease pathogenesis but also facilitate better management of CD patients.

[0058] Although we had clinical information available on disease phenotype, duration and prior surgeries, the serum samples were procured at different stages of the patients' diseases, and not necessarily at the time of high disease activity or surgery. Second, several patients were receiving medical therapy with different agents, including steroids, immunomodulators and/or biologics, while some patients were not on maintenance therapy. All these confounding factors might have contributed to relatively weak signals for the association analysis. Third, our study focused on AAbs that can differentiate CD patients from healthy subjects and we did not study patients with ulcerative colitis. In the future, we need to assess the CD specificity of the antibodies we discovered against UC patients in addition to the healthy controls we have assessed. Fourth, we note that proteins on the arrays may lack some native post-translational modifications (PTMs) that might be important for some antibody-antigen interactions.

[0059] In summary, we have developed a platform that can assay sero-reactivity to many CD related proteins in parallel. We screened 1906 proteins, to determine reactivity in CD sera to known and novel proteins. We observed unique antibody reactivity patterns. At individual antigen levels, antibodies against A4-Fla2, flaD, XAGE2, and CDC42EP1 were identified for IgG antibody and for IgA antibody anti A4-Fla2, PPP1R13L, FOSL1, SNRPB, CCHCR1, LCORL, NUPL1, PRPH, PTTG1, SNAI1, and ORC2L.

[0060] While the preferred embodiments of the present technology have been illustrated in detail, it should be apparent that modifications and adaptations to those embodiments may occur to one skilled in the art without departing from the scope of the present technology.

What is claimed is:

1. A method for detecting an antigen-antibody complex of interest in a disease, comprising:

- disposing a plurality of DNA segments corresponding to different open reading frames of genes on an array, wherein each DNA segment comprises a fusion tag; disposing a plurality of antibodies raised against the fusion tag on the array; translating the plurality of DNA segments to a plurality of polypeptides, wherein each polypeptide comprises the fusion tag and is captured by the antibodies raised against the fusion tag; and incubating the array with the plurality of captured polypeptides with samples collected from subjects having the disease and from healthy subjects to thereby form polypeptide/antibody complexes.
2. The method of claim 2, wherein the disposing a plurality of DNA segments step further comprises coating the array with aminosilane.
 3. The method of claim 1, wherein the fusion tag is a C-terminal glutathione S-transferase (GST) fusion tag.
 4. The method of claim 1, wherein the disposing a plurality of antibodies step further comprises incubating the array in a printing buffer containing anti-GST polyclonal antibodies.
 5. The method of claim 1, wherein the translating step further comprises expressing polypeptides from the disposed plurality of DNA segments using a HeLa cell lysate-based in vitro transcription and translation (IVTT) protein expression system.
 6. The method of claim 5, wherein the expressing step further comprises incubating the array with a human IVTT mixture.
 7. The method of claim 1, wherein the plurality of captured polypeptides are selected from the group consisting of A4-Fla2, flaD, XAGE2, CDC42EP1, PPP1R13L, FOSL1, SNRPB, CCHCR1, LCORL, NUPL1, PRPH, PTTG1, SNAI1, and ORC2L.
 8. The method of claim 1, wherein the plurality of captured polypeptides comprise A4-Fla2, FOSL1, PRPH, and XAGE2.
 9. The method of claim 1, wherein the plurality of captured polypeptides comprise A4-Fla2, FOSL1, PRPH, XAGE2, and SNRPB.
 10. The method of claim 1, wherein the plurality of captured polypeptides comprise A4-Fla2 and SNRPB.
 11. The method of claim 1, wherein the plurality of captured polypeptides comprise SNRPB.
 12. The method of claim 1, wherein the plurality of DNA segments corresponding to different open reading frames of genes are obtained from DNASU plasmid repository.
 13. The method of claim 1, wherein the plurality of DNA segments corresponding to different open reading frames of genes encode about 1900 different human proteins.
 14. The method of claim 1, wherein the sample is a blood or serum sample.
 15. The method of claim 1, wherein the healthy subjects match in gender and age with the subjects having the disease.
 16. The method of claim 1, wherein the disease is Crohn's disease.
 17. A method of detecting Crohn's disease in a subject, comprising:
 - providing a plurality of antigens, wherein the antigens are selected from the group consisting of A4-Fla2, flaD, XAGE2, CDC42EP1, PPP1R13L, FOSL1, SNRPB, CCHCR1, LCORL, NUPL1, PRPH, PTTG1, SNAI1, and ORC2L;
 - contacting the plurality of antigens with a sample from a subject containing a plurality of antibodies under conditions suitable to bind to the plurality of antigens to form antigen-antibody complexes; and
 - determining the level or concentrations of said complexes.
 18. The method of claim 16, further comprising comparing the level or concentrations of the complexes to a control level or concentration of the complexes from a healthy subject, wherein an increased level or concentration of the complexes indicates an increased likelihood of said subject having Crohn's disease.
 19. The method of claim 16, wherein the sample is a blood or serum sample.
 20. The method of claim 16, wherein the antigens comprise A4-Fla2, FOSL1, PRPH, and XAGE2.
 21. The method of claim 1, wherein the plurality of captured polypeptides comprise A4-Fla2, FOSL1, PRPH, XAGE2, and SNRPB.
 22. The method of claim 16, wherein the antigens comprise A4-Fla2 and SNRPB.
 23. The method of claim 16, wherein the antigens comprise SNRPB.

* * * * *