Probing predilection to Crohn's disease and Crohn's disease flares: A crowd-sourced bioinformatics approach

Jihad Aljabban a,⁎, Michael Rohr b,1, Vincent J. Borkowski a,a, Mary Nemer a,a, Eli Cohen c, Naima Hashi d, Hisham Aljabban e, Emmanuel Boateng c, Saad Syed f, Mohammed Mohammed g, Ali Mukhtar h, Dexter Hadley b, Maryam Panahiazari a

a University of Wisconsin Hospitals and Clinics, Madison, WI, United States
b University of Central Florida College of Medicine, Orlando, FL, United States
c Vanderbilt University Medical Center, Nashville, TN, United States
d Mayo Clinic Minnesota, Rochester, MN, United States
e Barry University, Miami Shores, FL, United States
f Northwestern Memorial Hospital, Chicago, IL, United States
g Windsor University School of Medicine, Saint Kitts and Nevis, Cayon
h Columbia University Vagelos College of Physicians and Surgeons, New York, NY, United States
i University of California San Francisco, San Francisco, CA, United States

ABSTRACT

Background: Crohn's Disease (CD) is an inflammatory disease of the gastrointestinal tract that affects millions of patients. While great strides have been made in treatment, namely in biologic therapy such as anti-TNF drugs, CD remains a significant health burden.

Method: We conducted two meta-analyses using our STARGEO platform to tag samples from Gene Expression Omnibus. One analysis compares inactive colonic biopsies from CD patients to colonic biopsies from healthy patients as a control and the other compares colonic biopsies from active CD lesions to inactive lesions. Separate tags were created to tag colonic samples from inflamed biopsies (total of 65 samples) and quiescent tissue in CD patients (total of 39 samples), and healthy tissue from non-CD patients (total of 30 samples). Results from the two meta-analyses were analyzed using Ingenuity Pathway Analysis.

Results: For the inactive CD vs healthy tissue analysis, we noted FXR/RXR and LXR/RXR activation, superpathway of citrulline metabolism, and atherosclerosis signaling as top canonical pathways. The top upstream regulators include genes implicated in innate immunity, such as TLR3 and HNRNPA2B1, and sterol regulation through SREBF2. In addition, the sterol regulator SREBF2, lipid metabolism was the top disease network identified in IPA (Fig. 1). Top upregulated genes hold implications in innate immunity (DUOX2, REG1A/1B/3A) and cellular transport and absorption (ABCG5, NPC1L1, FOLH1, and SLC6A14). Top downregulated genes largely held roles in cell adhesion and integrity, including claudin 8, PAQR5, and PRKACB.

For the active vs inactive CD analysis, we found immune cell adhesion and diapedesis, hepatic fibrosis/hepatic stellate cell activation, LPS/IL-1 inhibition of RXR function, and atherosclerosis as top canonical pathways. Top upstream regulators included inflammatory mediators LPS, TNF, IL1B, and TGFβ1. Top upregulated genes function in the immune response such as IL6, CXCL1, CXCR2, MMP1/7/12, and PTGS2. Downregulated genes dealt with cellular metabolism and transport such as GPO, RBP2, G6PC, PKC1, GSTA1, and MEP1B.

Conclusion: Our results build off established and recently described research in the field of CD. We demonstrate the use of our user-friendly platform, STARGEO, in investigating disease and finding therapeutic avenues.

Keywords: Crohn’s disease, Inflammatory bowel disease, Bioinformatics, Pathology, Genomics

http://dx.doi.org/10.1016/j.jpi.2022.100094
Available online 01 May 2022
2153-3539/Association for Pathology Informatics. Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).
1. Introduction

Inflammatory bowel disease (IBD) is a chronic relapsing disease that impacts the entire gastrointestinal tract. IBD has two main categories: Crohn’s disease (CD) and ulcerative colitis (UC). CD is the more common of the two. In North America, the annual prevalence of CD is 201 per 100,000 population. Crohn’s disease is increasing in prevalence worldwide, the highest incidence being in Europe and North America. CD is bimodal when it comes to disease onset, the first peak is around age 30 years and the second is around age 50 years. CD is classified as a heterogeneous disease in respect to location, and severity. CD affects any part of the gastrointestinal tract with discontinuous transmural penetrating lesions that lead to strictures and fistulas in both the small and large intestines. CD is a complex etiology that likely involves both environmental triggers as well as genetics. There are several environmental risk factors for CD such as smoking, diet, gastrointestinal infections, nonsteroidal anti-inflammatory drug use, and antibiotic exposure. However, the true cause is unknown.

Recent work has implicated a host of susceptibility gene defects. NOD2, a pattern recognition receptor that participates in intracellular bacterial immunity, was among the first identified. Defective NOD2 activity has been tied to the dysfunction of innate immune responses and bactericidal activity altering gut dysbiosis and epithelial barrier function. Other canonical pathways involve IL-17 signaling and TREM1 expression. Inappropriately activated immune defenses result from the aberration of intestinal cells’ immune-regulatory capacity, leading to this condition’s hallmark of tumor necrosis factor alpha (TNF-α) and tumor necrosis factor alpha (TNF-α).

The recent discovery of specific genes such as CARD15/NOD2 is a step in the right direction towards targeted immunotherapies. Finding gene variants could also help in predicting which patients will have a more aggressive disease course, in that case a more aggressive treatment is needed upfront.

Importantly, Crohn’s disease introduces serious and potentially life-threatening sequelae ranging from intestinal abscesses and obstructions to bleeding, fistulas, and perforations (mortality rate of up to 4.5%). About 1 in 5 patients develop fistulas/abscesses or strictures. These debilitating sequelae underscore the importance of elucidating causes of CD development and progression. Discerning the pathogenesis of CD will lead to comprehensive approaches to both patients as well as those identified for higher risk of CD development. Our STARGEO platform provides promise in the effort to understand the influence of diet and genetics on gut microbiome and immune response, furthering our efforts in creating holistic and sustainable treatment through genomic analysis.

In the age of precision medicine, bioinformatics is an indispensable tool for pushing care forward, but the expertise needed often makes it inaccessible to clinicians and researchers. STARGEO represents a bridge for clinicians and researchers with limited technologic expertise to conduct robust meta-analyses. Here, we demonstrate STARGEO as a tool in pathology informatics with this study on CD.

2. Methods

There is a wealth of pathological samples and data available to conduct insightful analysis but accessing and analyzing data may be a challenge for researchers and pathologists alike. The National Center for Biotechnology Information (NCBI) Gene Expression Omnibus (GEO) is one of the largest databases available to researchers. GEO is an open database containing millions of biological samples from genomics experiments. Dr. Dexter Hadley developed the Search Tag Analyze Resource for GEO (STARGEO) platform for clinicians and researchers to tag samples from GEO and perform meta-analysis of large datasets to motivate discovery. To prove the utility of this tool, we investigated CD to provide insights to pathogenesis and discern possible therapeutic avenues. STARGEO uses a standard random model for meta-analysis to generate both meta p-values and effects size across studies. Study weight percentages were calculated using the inverse variance method via the DerSimonian-Laird estimate. More information on STARGEO and can be found in our previous paper. A representative image illustrates how samples are tagged on STARGEO (Fig. 1).

We conducted two separate meta-analyses, one comparing inactive colonic biopsies from CD patients to colonic biopsies from healthy patients as a control and another comparing colonic biopsies from active lesions in CD patients compared to inactive lesions from the same patient. The purpose of the former was to understand the pathophysiology of CD and the latter to discern inflammatory changes that drive active disease and CD flares. The diagnosis of CD was according to standard clinical and pathologic criteria. Active vs inactive lesions were defined grossly and then classified by blinded gastrointestinal pathologists. The series used for the meta-analyses included GSE17594, GSE6731, and GSE75214. Separate tags were created to tag actively colonic samples from inflamed biopsies in CD patients (total of 65 samples), quiescent tissue in CD patients (total of 39 samples), and healthy tissue from non-CD patients (total of 30 samples).

We were able to extract approximately 22,000 genes for each of the meta-analyses conducted using STARGEO. We analyzed gene signature outputs with Ingenuity Pathway Analysis (IPA), to genes showing statistical significance (p < 0.05) and an absolute experimental log ratio greater than 0.1 between case and control samples. A total of 174 and 390 genes were included in the inactive vs control and active vs inactive meta-analyses, respectively. These selected genes were used for the next step analysis in IPA to dissect biological process, disease mechanisms, and potential biomarkers and therapeutic targets that were described in this manuscript. Table 4 details the most upregulated and downregulated genes for these analyses. See supplemental tables S1 and S2 for a complete list of gene p-values and experimental log ratios.

All data analyzed were taken from Gene Expression Omnibus. There was no interaction or intervention with human subjects and no involvement with access to identifiable private patient information. As such, no IRB approval was necessary.

3. Results

3.1. Canonical pathway, disease function and network analysis

Our inactive CD vs healthy colon biopsies (iHV) and active CD vs inactive CD (Avl) meta-analyses demonstrated thousands of genes with significant differences between case and control samples. We analyzed genes that demonstrated statistically significant (p < 0.05) and marked difference in expression (absolute experimental log ratios >0.1) in Ingenuity Pathway Analysis to highlight disease processes and functions that define CD risk factors and pathogenesis. A total of 174 and 390 genes from the iHV and Avl analyses, respectively, were analyzed.

The top five canonical pathways in iHV included farsenoid X receptor (FXR)/retinoid X receptor (RXR) activation, superpathway of citrulline metabolism, liver X receptor (LXR)/RXR activation, atherosclerosis signaling, and urea cycle (Table 1). For Avl, the top five canonical pathways included granulocyte adhesion and diapedesis, granulocyte adhesion and diapedesis, hepatic fibrosis/hepatic stellate cell activation, lipopolysaccharide (LPS)/interleukin-1 (IL-1) inhibition of RXR function, and atherosclerosis signaling (Table 1). Both analysis share similarities in atherosclerosis and bile acid receptor signaling (LXR/RXR/FXR function) but otherwise differ and likely represent differences in disease stage, with Avl highlighting inflammatory processes. It has been postulated that IBD and atherosclerosis share common pathogenic pathway, though much remains uncertain.

To identify the pathologic processes that define underlying CD disease and risk (iHV and CD flares (Avl), we used IPA to identify disease functions. We looked at disease functions that were predicted to be activated and with a z-score of >2. iHV disease function was mainly characterized by changes in lipid profiles and metabolism, along with recruitment of...
immune cells (Table 2). Predictably, AvI disease function largely featured inflammatory processes and recruitment of tumor cells such as “invasion on tumor cell lines”, “cell movement”, and “adhesion of immune cells” (Table 2).

Lastly, we further characterized pathogenesis in our two analyses using the IPA Network analysis.23 IPA ranks networks from the Global Molecular Network based on the number of focus genes from given networks that match with our analysis. Significance is given by the p-score ($p$-score = $-\log_{10}(p$-value)). We identified 12 networks for IvH and 25 networks for AvI. Both analyses highlighted several disease pathologies (Table 3), with the top network for IvH dealing with lipid metabolism similar to what was highlighted in our disease function analysis (Fig. 2).

### 3.2. Top up and downregulated genes and causal analysis

Next, we focus on the top upstream regulators and top up and downregulated genes in our analyses. IPA Upstream Regulator analysis identifies upstream transcription regulators that best reflect our observed genetic expression dataset.23 The $p$-values are based on the degree of overlap of

| Table 1 |
| --- |
| Top five canonical pathways associated with genetic differences in inactive CD colon samples vs normal controls and active vs inactive CD samples. CD (Crohn’s Disease), FXR (farnesoid X receptor), RXR (retinoid X receptor), LXR (liver X receptor), LPS (lipopolysaccharide), and IL-1 (interleukin-1). |

| Overlap | p-Value |
| --- | --- |
| FXR/RXR activation | 7/137 | 4.62E-05 |
| Superpathway of citrulline metabolism | 4/39 | 1.46E-04 |
| Atherosclerosis signaling | 6/128 | 2.61E-04 |
| Urea cycle | 3/20 | 3.33E-04 |

| Top canonical pathways in active vs inactive CD |
| --- |
| Granulocyte adhesion and diapedesis | 27/180 | 5.62E-18 |
| Agranulocyte adhesion and diapedesis | 25/193 | 3.15E-15 |
| Hepatic fibrosis/hepatic stellate cell activation | 24/186 | 1.25E-14 |
| LPS/IL-1 mediated inhibition of RXR function | 26/224 | 1.26E-14 |
| Atherosclerosis signaling | 20/127 | 4.44E-14 |

| Table 2 |
| --- |
| Top ten disease functions associated with genetic differences in inactive CD colon samples vs normal controls and active vs inactive CD samples. Activation z-score and p-values shown. |

| Activation Z-Score | p-Value |
| --- | --- |
| Concentration of fatty acid | 6.73E-04 | 2.907 |
| Cell movement of neutrophils | 1.91E-04 | 2.586 |
| Fatty acid metabolism | 2.65E-04 | 2.527 |
| Extravasation | 5.30E-04 | 2.395 |
| Transport of lipid | 7.40E-05 | 2.323 |
| Adhesion of immune cells | 5.28E-03 | 2.218 |
| Mass of fat pad | 4.52E-03 | 2.213 |
| Synthesis of lipid | 7.37E-04 | 2.198 |
| Concentration of lipid | 6.19E-08 | 2.190 |
| Top disease functions in active vs inactive CD |
| Invasion of tumor cell lines | 7.49E-18 | 5.764 |
| Cell movement of tumor cell lines | 8.25E-22 | 5.698 |
| Cell movement | 1.15E-32 | 5.335 |
| Invasion of cells | 1.81E-26 | 5.130 |
| Migration of cells | 1.23E-35 | 5.103 |
| Adhesion of blood cells | 1.11E-26 | 4.952 |
| Cell movement of myeloid cells | 9.86E-27 | 4.833 |
| Proliferation of muscle cells | 5.65E-16 | 4.823 |
| Adhesion of immune cells | 2.16E-26 | 4.793 |
| Migration of tumor cell lines | 1.41E-17 | 4.755 |

Fig. 1. A screen capture of how to use STARGEO. Relevant experiments can be searched and samples from those experiments can be tagged using the regex function as shown.
known effector targets and our gene list. The top up and downregulated genes for our analyses are detailed in Table 4 and in-text.

For the IvH analysis, top upstream regulators included transcription receptor STAT3 (p = 2.48E-08), the pattern recognition receptor toll-like receptor 3 or TLR3 (p = 6.48E-08, predicted activation), the ribonucleoprotein HNRNPA2B1 (p = 1.22E-07), and sterol regulatory element binding transcription factor 2 or SREBF2 (p = 1.50E-07). The top upregulated genes have activity namely in the innate immune response and included the reactive oxygen cell trafficking

Table 3

| Top molecular networks in inactive CD vs healthy control | Score |
|-----------------------------------------------|-------|
| Lipid metabolism, molecular transport, small molecule biochemistry | 47 |
| Antimicrobial response, humoral immune response, inflammatory disease | 34 |
| Endocine system disorders, gastrointestinal disease, immunological disease | 52 |
| Cellular movement, hematological system development and function, immune cell trafficking | 29 |
| Amino acid metabolism, cell signaling, developmental disorder | 29 |

Table 4

| Top upregulated genes | Top downregulated genes |
|-----------------------|-------------------------|
| Inactive CD vs Healthy | Inactive CD vs Healthy |
| Active vs Inactive CD  | Active vs Inactive CD  |

ABC25 (p = 4.86E-07) and Niemann-Pick C1-like 1 or NPC1L1 (p = 1.06E-08). Other top upregulated molecules are important in cellular metabolism and homeostasis including the transmembrane glycoprotein and glutamate carboxypeptidase FOLH1 (p = 1.70E-03) and the neutral and cationic amino acid transporter SLC6A14 (p = 6.55E-06).

Lastly, our analysis revealed inference of estrogen signaling in CD with upregulation of the g protein-coupled estrogen receptor SULT1E1 (p = 3.15E-04). Gut-epithelial integrity is influenced by abnormal lipid profiles so we studied the downstream effects of the upstream regulator SREBF2.24 IPA implicated SREBF2 signaling with several of the genes described above including ABCG5, REG1A, REG1B, and NPC1L1 (Fig. 3). We also identified links with genes relevant to anti-bacterial activity and the immune response including the bacteriostatic lipocalin-2 or LCN2 (p = 3.62E-11), the several cytokines including CXCL1 (p = 1.55E-03), CXCL8 (p = 3.16E-02), and CXCL11 (p = 1.35E-07). This suggests a relationship between disruption of cholesterol homeostasis and a contributory pro-inflammatory response to CD.

For the IvH analysis, top upstream regulators included transcription receptor STAT3 (p = 2.48E-08), the pattern recognition receptor toll-like receptor 3 or TLR3 (p = 6.48E-08, predicted activation), the ribonucleoprotein HNRNPA2B1 (p = 1.22E-07), and sterol regulatory element binding transcription factor 2 or SREBF2 (p = 1.50E-07). The top upregulated genes have activity namely in the innate immune response and included the reactive oxygen cell trafficking

Fig. 2. Top network (Lipid Metabolism, Molecular Transport, Small Molecule Biochemistry) in the inactive Crohn’s disease vs healthy control meta-analysis identified by IPA Network analysis. Legend illustrates class of the gene. Red indicates upregulation and green downregulation, with shade depicting magnitude of change. Solid and dashed lines depict direct and indirect, respectively, relationship between genes. Figure was generated using Ingenuity Pathway Analysis. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)
For the top downregulated genes in iAVI analysis, we found genes namely involved in cell adhesion and epithelial integrity. The most downregulated gene was the tight junction protein claudin 8 (*p* = 1.11E-14). Additionally, we found downregulation of progesterin and adiponectin receptor 5 or PAQR5 (*p* = 2.129E-03), with recent studies suggesting a role in progesterin in epithelial barrier function.\(^{56}\) Furthermore, the cAMP-dependent kinase PRKACB (*p* = 2.79E-5), which helps maintain epithelial homeostasis through the nuclear receptor LRH-1, was starkly downregulated.\(^{37}\) Other top downregulated genes deal with cellular metabolism, transport, and physiology. In terms of cellular transport and metabolism, we found downregulation of the cationic transporter SLC38A4 (*p* = 9.74E-07), the monocarboxylate transporter SLC16A9 (*p* = 1.28E-10), and the pH regulator carbonic anhydrase CA12 (*p* = 4.06E-02). Also, we found downregulation of the GPI mannoslytransferase PIGZ (*p* = 1.06E-02), which is a part of glycolipid biosynthesis.\(^{38}\) Other top downregulated genes are not as well described but may be aberrantly expressed in IBD. We noted downregulation of the metallothionein MT1M (*p* = 7.38E-09), which have high content of cysteine residues and bind heavy metals and may have some relation to IBD.\(^{58}\) Lastly, we found downregulation of the recently described mucolipin 2 or MCOLN2 (*p* = 1.95E-06), an ion channel whose function remains obscure but may be critical to normal immune cell function.\(^{40}\)

For the AvI analysis, top upstream regulators included LPS (*p* = 3.79E-57, with predicted activation), tumor necrosis factor or TNF (*p* = 5.40E-46, with predicted activation), IL1B (*p* = 3.62E-40, with predicted activation), and transforming growth factor B1 or TGB1 (*p* = 4.33E-39, with predicted activation). Expectedly, the top upregulated genes have pro-inflammatory properties. The top upregulated gene was the pro-inflammatory cytokine interleukin-6 or IL6 (*p* = 4.49E-02). Among other top upregulated genes, we found chemokine and chemokine receptors like chemokine ligand 1 or CXCL1 (*p* = 4.86E-03) and CXC receptor 2 or CXCR2 (*p* = 4.37E-02). Several matrix metalloproteinases were represented in our top downregulated genes, such as MMP1 (*p* = 6.67E-04), MMP7 (*p* = 6.13E-04), and MMP12 (*p* = 2.82E-08). Of note, the pro-inflammatory cyclooxygenase 2 or COX2/PTGS2 (*p* = 2.76E-03) was upregulated. Lastly, we found upregulation of the novel secretory necrosis factor-inducible protein TNFAP16 (*p* = 4.44E-16), whose exact function is still poorly described but serves as a biomarker in IBD.\(^{41}\)

The top downregulated genes for the AvI analysis largely functioned in cellular metabolism and transport. Interestingly, the top downregulated gene carboxypeptidase O or CPO (*p* = 3.47E-02) is an enzyme with largely unknown function but its relative carboxypeptidase E or CPE may be protective against colitis.\(^{42,43}\) Other downregulated genes hold essential properties in metabolism such as retinoid binding protein 2 or RBP2 (*p* = 6.19E-05), a protein involved in the uptake and metabolism of vitamin A.\(^{44}\) Other downregulated genes are required in glucose metabolism, counting glucose-6-phosphatase or G6PC (*p* = 3.29E-03) and phosphoenolpyruvate carboxykinase 1 or PCK1 (*p* = 3.06E-05).\(^{45,46}\) Additional genes of note have transporter functions like the potassium channel KCNJ13 (*p* = 1.38E-02), the sodium-sulfate cotransporter SLC13A1 (*p* = 8.13E-04), and the sodium-coupled solute transporter SLC5A12 (*p* = 1.52E-08).\(^{47}\) Other genes highlighted detoxification function, for instance the glutathione S-transferases GSTA1 (*p* = 2.91E-03) and GSTA2 (*p* = 1.28E-03).\(^{48}\) The last downregulated gene worth mentioning was meprin IB or MEP1B (*p* = 3.86E-06), a zinc metalloprotease expressed in the intestinal brush border that may protect against pathologic gut microbes.\(^{49}\) Since TNF has a pivotal role in CD, as evidenced by the success of anti-TNF inhibitors, we examined the relationship between TNF as an upstream regulator in our analysis with the top up and downregulated genes above. Fig. 4 illustrates the relationship between TNF and several of the genes described above, demonstrating the various effects TNF plays in CD pathogenesis.

### 4. Discussion

With the currently unknown pathogenesis of CD, a genomic source is very likely given the change in regulation of the genes discussed above. This is further explained by the strong family history component of the disease.\(^{50}\) A strong next step would be to trace CD-diagnosed family members and see if the genes are active between them as well. Using the above data, we can already postulate that with the genes being active more in CD patients than in others, it likely plays a role in disease manifestation or progression or is altered as a result of other CD pathogenesis.

We conducted meta-analysis of colon samples from inactive Crohn’s patients compared to healthy controls to discern the baseline changes that characterize CD. Our results also showcase the role of innate immunity in predisposition to CD. Immune function was represented in our top disease function and network analysis (Tables 2 and 3). The gut mucosa maintains a delicate balance through interactions of the mucosal immune system and the intestinal microbiota.\(^{51}\) Intestinal epithelial cells hold both barrier and innate immune function for gut homeostasis that is disrupted by loss of gut epithelial integrity and chronic inflammation in IBD. TLR3 was one of the top upstream regulators in our analysis. TLR3 is a pattern recognition receptor of the innate immune system that is classically activated by double-
stranded RNA. In the context of IBD, its activation may lead to predilection of disease through the chemokine CCL20. CCL20 expression is linked to pathogenic bacteria and can be attenuated through treatment of commensal bacteria in vitro. Targeting CCL20 has been proposed to target inflammation and may be promising in treatment of IBD. TLR3 may also influence IBD pathogenesis through induction of the alternate complement pathway in the mucosa. The most upregulated and downregulated in our analysis suggest compromise of both barrier and enterocyte function, along with maladaptive inflammatory changes. For innate immunity, we found stark upregulation of the oxidase DUOX2, bactericidal c-type lectin and REG family genes REG1A/1B/3A, and the alpha defensin DEFA6. DUOX2 is expressed in gastrointestinal cells and produces reactive oxygen species. DUOX2 has been shown to shape the microbiota in pediatric CD patients. Likewise, the antibacterial peptide DEFA6 may shape the microbiome and its upregulation in IBD patients is correlated with cytokine induction. Additionally, DEFA6, along with REG1A/1/3A, is expression correlated with cytokine induction in IBD. The c-lectins encoded by REG1A, REG1B, and REG3A play similar roles to DEFA6 and DUOX2 in enterocyte innate immune function through bactericidal action. Restoring immune homeostasis is a promising direction in the treatment of IBD. One avenue to accomplish this is through probiotics, which has been shown to improve intestinal barrier function and protect against inflammatory changes.

Our results highlight various cellular pathways and functions, such as innate immune cell signaling, lipid homeostasis, intestinal–epithelial integrity, and ionic transport. We begin this discussion with the top canonical pathways identified by IPA (Table 1). We found FXR/RXR and LXR/RXR activation. FXR is a ligand of bile acids and is paramount to bile acid synthesis and homeostasis. The purpose of bile acid-mediated activation of FXR in CD are just being described with CD patients showing higher signaling of FXR compared to control patients. The implications of this increase are still being explored but FXR has been implicated in colitis as a modulator of intestinal immunity in murine models. Additionally, bile acids are toxic in excess, with dysregulated FXR activity linked to progression in IBD. While FXR is responsible for bile acid regulation, LXR and RXR regulate lipid metabolism. It is through regulation of lipid metabolism and other facets of gastrointestinal homeostasis that LXR and RXR are involved in CD. RXR and LXR largely mediates anti-inflammatory effects with deficiency linked to colitis susceptibility in murine models and reduced expressions noted in IBD patients. Thus modulation of bile acid signaling may serve therapeutic role in the treatment of IBD, which may be accomplished by bile acid sequestrants or by other means.

In addition to immune dysregulation, our results implicate changes to cholesterol metabolism (Tables 2 and 3). We discussed some changes to lipid metabolism through bile acid receptor activity above. We found upregulation of NPC1L1, which encodes a protein responsible for cholesterol absorption. Acute high cholesterol in animal models can lead to an inflammatory response that is dependent on NPC1L1. Similarly, the ABC transporter ABC5 limits intestinal absorption and promotes excretion of sterols. Interestingly, ABCG5 and ABCG8 may regulate TLR-induced inflammation. There has been recent evidence for the utility of statin therapy in the treatment of colitis. Statin use in animal models of colitis has been shown to have disease modifying effects. Recent clinical studies expand on murine models and suggest statin therapy can protect against development of CD.

To predict genetic changes during CD flares, we did a meta-analysis comparing colonic biopsies of active Crohn’s patients to inactive Crohn’s patients. As with our analysis of healthy colon samples compared to inactive
Crohn’s, the results of our analysis illustrated changes to immune activity. Top canonical pathways included granulocyte and agranulocyte adhesion and diapedesis, and LPS/IL-1 mediated inhibition of RXR function as top pathways (Table 1). Unsurprisingly, the top two canonical pathways as well as top disease functions and networks highlighted immune cell recruitment (Tables 1–3). Additionally, top upstream regulators included immune mediators such as LPS and TNF. As discussed above, CD patients have higher serum levels of LPS and LPS itself can alter the microbiome in IBD. Additionally, LPS induces inflammatory mediators such as TNF. TNF is classically described in CD, with anti-TNF therapy being approved in 1998. TNF is involved in several disease processes in CD including neutrophil recruitment, granuloma formation, and epithelial permeability (Fig. 4). Additionally, top upstream regulators and genes also implicated in inflammation. The most upregulated gene is IL6, a pro-inflammatory cytokine shown to be increased in both mesenteric and serum samples in active CD patients. Likewise, the chemokine CXCL8, or IL8, has roles in both gastrointestinal inflammation and malignancy, with antagonism relieving colitis in murine IBD models. Additional, we found upregulation of another chemokine, CXCL1, that also contributes to immune cell recruitment and inflammation in CD. Lastly, increased expression of the chemokine receptor CXCR2 is linked with chronic inflammation in IBD, and CXCR2 knockout murine models demonstrated less neutrophil infiltration with limited mucosal damage and clinical symptoms.

Aside from inflammation, our results reveal fibrosis and extracellular changes as promoters of CD progression and flares. We identified hepatic fibrosis/hepatic stellate cell activation as a top canonical pathway. This pathway is characterized by cellular proliferation and extracellular matrix deposition. There is an evolving understanding of intestinal fibrosis being independent, rather than a consequence of, inflammation in CD and a possible route for disease modifying therapy. Furthermore, the top upregulated genes included several matrix metalloproteinases or MMPs. MMPs shape the extracellular matrix of tissue and exert various biologic effects through cleavage of bioactive proteins. MMPs have long been shown to be elevated in IBD and shape epithelial barrier function, the immune response, fibrosis, and other pathogenic activity. MMP expression have also been linked to increased risk for fibrosis and malignancy in IBD. Our analysis illustrated upregulation of several MMPs including MMP1, MMP7, and MMP12. These results support previous literature and its value in future treatment and risk stratifying strategies for CD.

As with the IVH analysis, the top downregulated genes in the AVI analysis mostly dictate cellular metabolism, transport, and cellular detoxification. The most downregulated gene was the membrane-bound carboxypeptidase or CPO. This peptidase cleaves acidic and polar C-terminal amino acids but, otherwise, much remains unknown. Its relative carboxypeptidase E has anti-inflammatory activity that is protective against experimental colitis, so CPO may share similar properties in maintaining intestinal immune homeostasis. We also observed impairment of detoxification through the downregulation of glutathione S-transferases GSTA1 and GSTA2. The glutathione transferases are enzymes that function in detoxification of electrophilic compounds such as exogenous toxins and products of oxidative stress. Downregulation of these enzymes are seen in IBD and may contribute to pathogenesis. Additionally, there is downregulation of other genes involved in metabolism including retinol-binding protein 2 (RBP2). RBP2 is present in intestinal epithelium and regulates uptake and metabolism of vitamin A. Decrease of mRNA levels of RBP2 was identified in both experimental murine models and IBD patient tissue, suggesting an unfound role of RBP2 in maintaining intestinal homeostasis. Thus, these genes may represent future biomarkers to disease activity and progression.

The results from our AVH analysis propose genetic drivers to CD progression and possible predictors to CD flares. Some of these genes may have utility in stratifying aggressive disease and predict which patients may benefit from earlier biologic therapy. Additionally, molecular and gene-directed remain a high-priority for pharmaceutical targets. As noted above with the mentioned genes, there are already viable targets for which we have directed therapies, with examples being Tumor Necrosis Factor Alpha and Infliximab. We anticipate therapies could be directed towards, for example, TL83 or the oxidase DUOX2 given the importance they have with regulation of cell signaling and production of reactive oxygen species, respectively. Potential next steps exist for gene therapy related to the aforementioned genes such as FOLH1 and REG1A. Future studies would require animal knock-out or knock-down models to examine full gene ability. With upregulation of these genes noted in CD, a potential gene therapy via a viral or other vector could be possible.

5. Conclusion
Crohn’s disease (CD) is an inflammatory bowel disease with a range of severity in both intestinal and extra-intestinal symptoms. The exact cause of CD remains unknown. Our use of STARGEO aimed to elucidate the genetic changes that define predisposition to CD and to flares of disease. From our inactive CD vs healthy tissue analysis, we found genetic changes in innate immune function through such immune mediators as STAT3 and TL83 and upregulation of DUOX2, bactericidal c-type lectins, and REG family genes. Such changes appear maladaptive and may compromise intestinal homeostasis. We also noted changes in genes involved in cholesterol absorption and regulation, which have implications in both the immune response and intestinal barrier function. Expectedly, our active vs inactive CD analyses highlighted pro-inflammatory changes that characterize disease flares. Overall, this study is an effective demonstration of STARGEO, a promising, accessible tool for researchers and clinicians alike.

Author contributions
Conception or design of the work (CD), Data collection (DC), drafting manuscript (DM), critical revision of manuscript (CM), final edits and approval (FE). Jihad Aljabban (CD, DC, DM, CM, FE); Michael Rohr (CD, DC, DM, CM, FE); Saad Syed (DM, CM, FE); Eli Cohen MD (DM, CM, FE); Naima Hashim (DM, CM, FE); Hajrah Khan (DM, CM, FE); Ali Mukhtar (DM, CM, FE); Hisham Aljabban (DM, CM, FE); Emmanuel Boataeng (CM, FE); Mary Nemer (CM, FE); Maryam Panahiazar (CM, FE); Dexter Hadley (DC, CM, FE).

Funding
This study did not use any funding.

Conflicts of interests
None of the authors have any conflict of interests to declare.

Acknowledgments
We would like to acknowledge the National Center for Biotechnology and Information (NCBI) for the establishment of the public data repository Gene Expression Omnibus (GEO).

Appendix A. Supplementary data
Supplementary data to this article can be found online at https://doi.org/10.1016/j.jpi.2022.100094.

References
1. Torres J, Mehandru S, Colombel JF, et al. Crohn’s disease. Lancet 2017;389(10080): 1741–1755. https://doi.org/10.1016/S0140-6736(16)31711-1. Epub 2016 Dec 1. PMID: 27914655.
2. Cosnes J, Goujard Rousseau C, Seksik P, et al. Epidemiology and natural history of inflammatory bowel diseases. Gastroenterology 2011. https://doi.org/10.1053/j.gastro.2011.01.055.
33. Cole GB, Keum G, Liu J, et al. Speciﬁc estrogen sulfotransferase (SULT1E1) substrates help regulate the triglyceride/fatty acid cycle and development of insulin resistance in mice. J Lipid Res 2010. https://doi.org/10.1194/jlr.M005436.

34. Park SH, Kim J, Yu M, et al. Epithelial cholesterol deﬁciency attenuates human antigen R-linked diabetes and induces an SREBP2-linked circuit. J Biol Chem 2016. https://doi.org/10.1074/jbc.M115.729293.

35. Stallhofer J, Friedrich M, Knezev-Zema A, et al. Lipopolid-2 is a disease activity marker in inﬂammatory bowel disease regulated by IL-17A, IL-22, and TGF-α and modulated by L2Rβ2 gene status. Inflamm Bowel Dis 2015. https://doi.org/10.1097/MIB.0000000000000151.

36. van der Giesen J, van der Woude C, Peppelenschm bosch M, et al. A direct effect of sex hormones on epithelial barrier function in inﬂammatory bowel disease models. Cells 2019. https://doi.org/10.3390/cells8040026.

37. Bayer JR, Wang H, Nattiv R, et al. LRH-1 mitigates intestinal inﬂammatory disease by maintaining epithelial homeostasis and cell survival. Nat Commun 2018. https://doi.org/10.1038/s41467-018-06137-w.

38. Kimotshi T, Inoue N. GBD2: a genome-wide extension (PGM, PIGV, PIGB, PIGZ). Handbook of Glycoinferases and Related Genes2nd ed.; 2014.

39. Waeytens A, De Vos M, Laukens D. Evidence for a potential role of metallothioneins in inﬂammatory bowel diseases. Mediat Inﬂamm 2020. https://doi.org/10.1155/2020.1371/journal.pone.0102347.

40. Cuajungco MP, Silva J, Habibi A, et al. The mucolipin-2 (TRPML2) ion channel: a tissue-speciﬁc protein crucial to normal cell function. Pﬂugers Arch Eur J Physiol 2016. http://dx.doi.org/10.1007/s00424-015-2322-0.

41. Yu Q, Zhang S, Wang H, et al. TNFAP6 is a potential biomarker of disease activity in inﬂammatory bowel disease. Biomark Med 2016. https://doi.org/10.2217/bmm.16.19.

42. Nowotny S. Exploring the biological function of carboxypeptidase O: analysis of expression and association with chylomicrons and lipid droplets. Digit Commun Andre Univ 2020.

43. Bär F, Fisch P, Papg R, et al. Carboxypeptidase O modulates intestinal immune homeostasis and protects against experimental colitis in mice. PLoS One 2014. https://doi.org/10.1371/journal.pone.0132467.

44. Folli C, Calderone V, Ottoleno S, et al. Identiﬁcation, retina binding, and a s analysis of a human retinol-binding protein. Proc Natl Acad Sci U S A 2001. https://doi.org/10.1073/pnas.041313998.

45. Millward CA, DeSantis D, Hsieh CW, et al. Phosphoenolpyruvate carboxykinase (Pck1) regulates the spatial segregation of microbiota and host in the intestine. Science (80- ) 2011. https://doi.org/10.1126/science.1199818.

46. Bär F, Fisch P, Papg R, et al. Carboxypeptidase O modulates intestinal immune homeostasis and protects against experimental colitis in mice. PLoS One 2014. https://doi.org/10.1371/journal.pone.0132467.

47. Hediger MA, Romero MF, Bin Peng J, et al. The ABCs of solute carriers: physiological, pathological and therapeutic implications of human membrane transport proteins. Pﬂugers Arch Eur J Physiol 2004. https://doi.org/10.1007/s00424-003-1192-y.

48. Hayes JD, Flanagan JU, Jowsey UR, Glutathione transfers, Annu Rev Pharmacol Toxicol 2005. https://doi.org/10.1146/annurev.pharmtox.45.120403.095857.

49. Vazquez E, Bringer MA, Gardarin A, et al. Role of meprin to protect ileal mucosa of Crohn’s disease patients from colonization by adherent-invasive E. coli. PLoS One 2011. https://doi.org/10.1371/journal.pone.0012199.

50. Farmer RG, Michener WM, Mortimer EA. Studies of family history among patients with inﬂammatory bowel disease. Clin Gastroenterol 19809.

51. Okumura K, Takeeda K. Maintenance of gut homeostasis by the mucosal immune system. J Biol Chem 2009. https://doi.org/10.1074/jbc.R800025455.

52. Okumura R, Takeda K. Maintenance of gut homeostasis by the mucosal immune system. J Biol Chem 2009. https://doi.org/10.1074/jbc.R800025455.

53. Skovdahl HR, Van Beelen GA, Östvik A, et al. Quantitative PCR analysis of CCL20 and its corresponding receptor CCR6R expression in human colon biopsies. Nat Commun 2015. https://doi.org/10.1038/ncomms9786.

54. Akira S, Takeda K. Toll-like receptor signalling. Nat Rev Immunol 2004. https://doi.org/10.1038/nri1391.

55. Ranasinghe R, Eri R. Modulation of the CCR6-CC120 axis: a potential therapeutic target in inflammatory bowel disease. Biomark Med 2016. https://doi.org/10.2217/bmm.16.9.

56. Brien RM. Glucose-6-phosphatase catalytic subunit gene family. J Biol Chem 2009. https://doi.org/10.1074/jbc.R800025455.

57. Hediger MA, Romero MF, Bin Peng J, et al. The ABCs of solute carriers: physiological, pathological and therapeutic implications of human membrane transport proteins. Pﬂugers Arch Eur J Physiol 2004. https://doi.org/10.1007/s00424-003-1192-y.

58. Boyd M, Thodberg M, Vitezic M, et al. Characterization of the enhancer and promoter ﬂanking regions of Crohn’s disease. Nat Commun 2012. https://doi.org/10.1038/ncomms11582.

59. Bär F, Fisch P, Papg R, et al. Carboxypeptidase O modulates intestinal immune homeostasis and protects against experimental colitis in mice. PLoS One 2014. https://doi.org/10.1371/journal.pone.0132467.

60. Hayes JD, Flanagan JU, Jowsey UR, Glutathione transfers, Annu Rev Pharmacol Toxicol 2005. https://doi.org/10.1146/annurev.pharmtox.45.120403.095857.

61. Lu TT, Makishima M, Repa JJ, et al. Molecular basis for feedback regulation of bile acid synthesis and transport. Nat Rev Gastroenterol Hepatol 2004. https://doi.org/10.1038/gastro.2003.12.008.

62. Wilson A, Almousa A, Teft WA, et al. Attenuation of bile acid-mediated FXR and PXR activation by 7a,9α,12α-trihydroxy-22β-cholesta-5,7,24-trien-3β-ol. Proc Natl Acad Sci U S A 2010.https://doi.org/10.1073/pnas.0905000107.

63. Kramer A, Green J, Heubi J, et al. Causal and molecular approaches in ingenuity pathway analysis. Bioinformatics 2014. https://doi.org/10.1093/bioinformatics/btt703.

64. Danese S, Fiocchi C. Atherosclerosis and inﬂammation in Crohn’s disease: a focus on the ATG16L1 gene. Expert Rev Mol Diagn 2008.https://doi.org/10.1586/14737159.8.2.199.

65. Gutteridge J, O’Brien RM. Glucose-6-phosphatase catalytic subunit gene family. J Biol Chem 2009. https://doi.org/10.1074/jbc.R800025455.

66. Millward CA, DeSantis D, Hsieh CW, et al. Phosphoenolpyruvate carboxykinase (Pck1) regulates the spatial segregation of microbiota and host in the intestine. Science (80- ) 2011. https://doi.org/10.1126/science.1199818.

67. Vazquez E, Bringer MA, Gardarin A, et al. Role of meprin to protect ileal mucosa of Crohn’s disease patients from colonization by adherent-invasive E. coli. PLoS One 2011. https://doi.org/10.1371/journal.pone.0011299.

68. Farmer RG, Michener WM, Mortimer EA. Studies of family history among patients with inﬂammatory bowel disease. Clin Gastroenterol 19809.

69. Okumura K, Takeeda K. Maintenance of gut homeostasis by the mucosal immune system. J Biol Chem 2009. https://doi.org/10.1074/jbc.R800025455.

70. Okumura R, Takeda K. Maintenance of gut homeostasis by the mucosal immune system. J Biol Chem 2009. https://doi.org/10.1074/jbc.R800025455.

71. Skovdahl HR, Van Beelen GA, Östvik A, et al. Quantitative PCR analysis of CCL20 and its corresponding receptor CCR6R expression in human colon biopsies. Nat Commun 2015. https://doi.org/10.1038/ncomms9786.
64. Gadaleta RM, Van Erpecum KJ, Oldenburg B, et al. Farnesoid X receptor activation inhibits inflammation and preserves the intestinal barrier in inflammatory bowel disease. Gut 2011. https://doi.org/10.1136/gut.2011.212159.

65. Klepsch V, Moschen AR, Tilg H, et al. Nuclear receptors regulate intestinal inflammation in the context of IBD. Front Immunol 2019. https://doi.org/10.3389/fimmu.2019.01070.

66. Kanagarajan N, Nam JH, Al Noah Z, et al. Disease modifying effect of statins in dextran sulfate sodium model of mouse colitis. Inflamm Res 2008;57.

67. Lochhead P, Khalili H, Sachs MC, et al. Association between statin use and inflammatory bowel diseases: results from a Swedish, nationwide, population-based case-control study. J Crohn’s Colitis 2021;15.

68. Bhagavathula AS, Clark C, Rahmani J. Statin use and new-onset of inflammatory bowel disease: a systematic review and meta-analysis of over ten million participants. Eur J Pharmacol 2021;891.

69. Stephens M, von der Weid PY. Lipopolysaccharides modulate intestinal epithelial permeability and inflammation in a species-specific manner. Gut Microbes 2020. https://doi.org/10.1080/19490976.2019.1629235.

70. Adegbola SO, Sahnan K, Warusavitarne J, et al. Anti-TNF therapy in Crohn’s disease. Int J Mol Sci 2018. https://doi.org/10.3390/ijms19082244.

71. Mahida YR, Kurlac L, Gallagher A, et al. High circulating concentrations of interleukin-6 in active Crohn’s disease but not ulcerative colitis. Gut 1991. https://doi.org/10.1136/gut.32.12.1531.

72. Walana W, Ye Y, Li M, et al. IL-8 antagonist, CXCL8(3-72)K11R/G31P coupled with probiotic exhibit variably enhanced therapeutic potential in ameliorating ulcerative colitis. Biomed Pharmacother 2018. https://doi.org/10.1016/j.biopha.2018.04.008.

73. Cotton JA, Platnich JM, Muruve DA, et al. Interleukin-8 in gastrointestinal inflammation and malignancy: induction and clinical consequences. Int J Interf Cytokine Mediat Res 2016. https://doi.org/10.2147/IJCIR.S63682.

74. Wang D, Dellois RN, Richmond A. The role of chemokines in intestinal inflammation and cancer. Curr Opin Pharmacol 2009. https://doi.org/10.1016/j.coph.2009.08.003.

75. Buannne P, Di Carlo E, Caputi L, et al. Crucial pathophysiological role of CXCR2 in experimental ulcerative colitis in mice. J Leukoc Biol 2007. https://doi.org/10.1189/jlb.0207118.

76. Cheng Y, Ma Xl, Wei Yq, et al. Potential roles and targeted therapy of the CXCLs/CXCR2 axis in cancer and inflammatory diseases. Biochim Biophys Acta - Rev Cancer 2019. https://doi.org/10.1016/j.bbcan.2019.01.005.

77. Rogler G, Hausmann M. Factors promoting development of fibrosis in crohn’s disease. Front Med 2017. https://doi.org/10.3389/fmed.2017.00096.

78. O’Sullivan S, Gilmer JP, Medina C. Matrix metalloproteinases in inflammatory bowel disease: an update. Mediat Inflamm 2015. https://doi.org/10.1155/2015/964131.

79. Biancheri P, Di Sabatino A, Corazza GR, et al. Proteases and the gut barrier. Cell Tissue Res 2013. https://doi.org/10.1007/s00441-012-1390-z.

80. Baugh MD, Perry MJ, Hollander AP, et al. Matrix metalloproteinase levels are elevated in inflammatory bowel disease. Gastroenterology 1999. https://doi.org/10.1001/01 gastro.1999.00580599703392.

81. Meijer MJW, Mieremet-Ooms MAC, van der Zon AM, et al. Increased mucosal matrix metalloproteinase-1, –2, –3 and –9 activity in patients with inflammatory bowel disease and the relation with Crohn’s disease phenotype. Dig Liver Dis 2007. https://doi.org/10.1016/j.dld.2007.05.010.

82. Langmann T, Moehle C, Maserer R, et al. Loss of detoxification in inflammatory bowel disease: dysregulation of pregnane X receptor target genes. Gastroenterology 2004. https://doi.org/10.1053/j.gastro.2004.04.019.

83. Lee SA, Zhang Yang KJ, Brun PJ, et al. Retinol-binding protein 2 (RBP2) binds monoaacylglycerols and modulates gut endocrine signaling and body weight. Sci Adv 2020. https://doi.org/10.1126/sciadv.aay8937.

84. Rankin CR, Theodorou E, Law IKM, et al. Identification of novel mRNAs and lncRNAs associated with mouse experimental colitis and human inflammatory bowel disease. Am J Physiol - Gastrointest Liver Physiol 2018. https://doi.org/10.1152/ajpgi.00077.2018.

85. Dryden GW. Overview of biologic therapy for Crohn’s disease. Expert Opin Biol Ther 2009;9.