Biomarker discovery and extracellular vesicle proteomic signatures in pediatric inflammatory bowel disease

Shelley Deeke

Thesis submitted to the
Faculty of Graduate and Postdoctoral Studies
in partial fulfillment of the requirements
for the Doctorate in Philosophy degree in Biochemistry

Ottawa Institute of Systems Biology
Department of Biochemistry, Microbiology and Immunology
Faculty of Medicine
University of Ottawa

© Shelley Deeke, Ottawa, Canada, 2018

ABSTRACT

Background: Reliable biomarkers are needed to evade the risk of injury, invasiveness and discomfort of endoscopies, which are required for inflammatory bowel disease (IBD) diagnosis and extent of disease assessment in ulcerative colitis (UC) patients. The need for biomarkers is accentuated in children, wherein the most frequently used IBD biomarker yields low specificity. Proteomics of clinical samples or their enriched components is a means to evaluate and identify alterations in proteins reflective of disease, with the potential for use as biomarkers and for providing insight on disease pathogenesis.

Methods: Proteins were isolated from the intestinal mucosal-luminal interface (MLI), collected from the ascending and descending colon of pediatric treatment-naive patients. The intestinal MLI proteomes of 42 IBD and 18 control patients were analyzed by high resolution mass spectrometry (HRMS). Multivariate analysis and receiver operating characteristics curves were performed to develop protein biomarker panels to discriminate IBD from control, and for UC extent of disease. ELISAs were used to assess a subset of biomarker candidates in stool samples from an independent pediatric cohort (n=24). Extracellular vesicles (EVs) were isolated by ultracentrifugation from the intestinal MLI of 11 IBD and seven control patients, and analyzed by electron microscopy, nanoparticle tracking analysis and HRMS.

Results: A biomarker panel of four proteins classified patients as either controls or active IBD with 97.5% accuracy. A second biomarker panel correctly classified 100% of UC patients as presenting with pancolitis or non-pancolitis. The differential protein expression of two biomarker candidates (catalase and leukotriene A-4 hydrolase) identified from the intestinal MLI was

comparable in stool samples. Comparison of EV proteomes isolated from IBD patients and controls identified differential expression of processes related to host defense and immunity.

Conclusions: Proteomic analysis of clinical samples identified differentially expressed proteins that can classify IBD patients from non-IBD controls and distinguish UC patients with pancolitis from those without pancolitis; proteins identified in intestinal aspirates displayed consistent differential expression in stool. Furthermore enrichment of EVs from the intestinal MLI indicates that these may contribute to the dysregulated host response against the intestinal microbiota which is observed in IBD.

ACKNOWLEDGMENTS

Foremost I would like to express my gratitude to my supervisors Dr. Daniel Figeys and Dr. Alain Stintzi for providing me with the opportunity to be part of this incredible study. Thank you for this remarkable journey and rewarding experience in contributing to the fight against inflammatory bowel disease. I cannot imagine a more valuable doctoral experience. I am forever grateful for the mentorship, advice and guidance you both provided. I would also like to thank my thesis advisory committee members Dr. Thien-Fah Mah and Dr. Derrick Gibbings for their invaluable feedback, support and encouragement that went above and beyond their roles as TAC members.

I owe an enormous debt of gratitude to my mentor Dr. Amanda Starr. Thank you for your unwavering support which accompanied me along every step in my doctoral journey. I cannot imagine a more proficient, reliable, and talented mentor and scientist. You had a tremendous impact into shaping me into the scientist and person I am today, I am forever grateful for all your insight, guidance, advice, devotion and support.

I would like to extend the deepest thanks to Dr. David Mack and Ruth Singleton for their support, enthusiasm and passion for making a difference in the lives of children living with inflammatory bowel disease. You both continue to inspire me and I am eternally grateful for our collaboration.

My sincerest gratitude to Trevor who provided me with the confidence, encouragement and support which allowed me to pursue my doctoral studies. Without you I would have missed being part of this outstanding journey.

I would like to thank Dr. Janice Mayne for her continued support, encouragement, passion and spirit. I cannot thank you enough for the mentorship you have provided me. It was your initial confidence and support in me which set the wheels in motion for my degree; I am forever grateful for your support. I am deeply grateful to Dr. Zhibin Ning for acting as my proteomic mentor and sharing his vast knowledge of mass spectrometry with me. Your skills continue to impress and inspire me.

I am very grateful for the funding provided by the Ontario Graduate Scholarship program and the University of Ottawa for the Excellence and Admission scholarships.

I would also like to thank the past and present members of the Figeys lab and Stintzi lab for their friendship, feedback, and insight and for providing me with an outstanding support network. A special thanks to Cheng-Kang Chiang, Deeptee Seebun, Alex Star, Lioudmila Tepliakova, Xu Zhang, Rachid Daoud, Guillaume Romain, James Butcher, Jennifer Li, Kiara Ge Chu and Colleen Paquette. It was an absolute pleasure working with you, I could not have asked for a better team.

Finally I would like to thank my family and friends for their love and moral support throughout my studies.

TABLE OF CONTENTS

Contents

ABBREVIATIONS

25(OH)D: 25-hydroxy vitamin D

A

AC: Ascending colon
AIEC: Adherent-invasive *Escherichia coli*
AUC: Area under the curve (AUC)

B

B. thetaiotaomicron: *Bacteroides thetaiotaomicron*
BtMinpp: *Bacteroides thetaiotaomicron* mammalian $InsP_6$ phosphatase

C

CD: Crohn's disease
CHEO: Children's Hospital of Eastern Ontario
CoA: Area of the colon with macroscopic inflammation
CoN: Area of the colon without macroscopic inflammation
C. burnetii: Coxiella burnetii

D

DC: Descending colon
DSS: Dextran sodium sulfate
DTT: Dithiothreitol

E

EBV: Epstein– Barr virus
EEN: Exclusive enteral nutrition
ELISA: Enzyme-linked immunosorbent assay
ESI: Electrospray ionization
ESR: Erythrocyte sedimentation rate

F

FC: Fecal calprotectin

G

GI: Gastrointestinal

H

HNP1: Human neutrophil peptide 1
hucMSCs: Human umbilical cord mesenchymal stem cells

I

IBD: Inflammatory bowel disease
IBDU: Inflammatory bowel disease unclassified
IBS: Irritable bowel syndrome
IL-12 : Interleukin 12
ILV: Intraluminal vesicles

L

LC: Liquid chromatography
LFQ: Label-free quantification
LTB_4 : Leukotriene B_4

M

MALDI-MS: Matrix-assisted laser desorption/ionization mass spectrometry
MLI: Mucosal-luminal interface
MS: Mass spectrometry
MV: Membrane vesicles
MVB: Multivesicular bodies

N

NF-κB : nuclear factor kappa-light-chain-enhancer of activated B cells
NLME: Non-linear mixed-effect model
NOD2: Nucleotide-binding oligomerization domain-containing protein 2
NTA: Nanoparticle Tracking Analysis

O

OFRG: Oligonucleotide fingerprinting of rRNA genes
OMV: Outer membrane vesicles

P

pANCA: perinuclear anti-neutrophil cytoplasmic autoantibodies
PCDAI: Pediatric Crohn's Disease Activity Index
PDCD6IP: Programmed cell death 6-interacting protein
PLS-DA: Partial least squares discriminant analysis
PUCAI: Pediatric Ulcerative Colitis Activity Index

R

ROC: Receiver operating characteristics
ROS: Reactive oxygen species

S

SCX: Strong cationic exchange

Super-SILAC: Stable-isotope labeling by amino acids in cell culture

T

TLR: Toll-like receptors
TSG101: tumor susceptibility gene 101

U

UC: Ulcerative colitis

V

VPS4A: Vacuolar protein sorting-associated protein 4A

Z

ZO-1: Zonula occludens-1

LIST OF FIGURES

Chapter1

Chapter 3

Chapter 4

Chapter 5

LIST OF TABLES

Chapter 3

Chapter 4

Chapter 5

CHAPTER 1. INTRODUCTION

1.1 Inflammatory bowel disease: two subclasses with lifelong repercussions

Inflammatory bowel disease (IBD) is a lifelong condition, comprising two main subtypes, Crohn's disease (CD) and ulcerative colitis (UC), and lacking a curative therapy. CD and UC differ in many aspects including the anatomical location of the affected area, disease course and response to therapy. In CD skip lesions can localize to any area within the GI tract, whereas in UC, the inflammation is contiguous, develops proximally from the rectum and is restricted to the large intestine. Furthermore the depth of affected mucosal layers differs between subtypes; in CD the inflammation is transmural, affecting all three mucosa layers, whereas in UC the inflammation is restricted to the mucosa. Despite these differences in disease location, symptoms including abdominal pain, bloody stool, diarrhea and weight loss (Henderson et al., 2014), are similar for CD and UC, leading to challenges in subtype diagnosis. Difficulties in differentiating disease subtype are further accentuated by atypical IBD subtype presentation including rectal sparing and upper GI tract ulcers in UC (Levine et al., 2014). Disease of IBD patients which cannot be confidently assigned as either CD or UC are categorized as an inflammatory bowel disease unclassified (IBDU); between 5%-15% of newly diagnosed IBD patients are deemed IBDU (Meucci et al., 1999; Ng et al., 2013; Nuij et al., 2013).

The incidence and prevalence of IBD differs globally. North America, Europe and Oceania are burdened by the highest prevalence's of IBD (>0.3%) and while the incidence has stabilized in developed countries, it continues to rise in recently industrialized countries (Ng et al., 2017). In North American, the incidence per 100,000 individuals per year is estimated to range from 6.3 to 23.8 for CD and 8.8 to 23.1 for UC, with prevalence per 100,000 ranging from

96.3 to 318.5 for CD and from 139.8 to 286.3 for UC (Ng et al., 2017). Canada reports one of the highest incidence's of childhood-onset IBD, with children < 5 years of age from five provinces (Ontario, Quebec, Alberta, Manitoba, Nova Scotia) experiencing a significant increase (7%) in the incidence from 1999 to 2010 (Benchimol et al., 2017). If current incidence rates continue it is estimated that > 10 million individuals worldwide will suffer from IBD by 2050 (Kaplan and Ng, 2017). In contrast, Eastern Asia reports one of the lowest incidences per 100,000 individuals per year, estimated to range from 0.06 to 3.2 for CD and 0.42 to 4.6 for UC and one of the lowest prevalence per 100,000, estimated to span 1.05 to 18.6 for CD and 4.59 to 57.3 for UC (Ng et al., 2017).

IBD is associated with lifelong repercussions, including a significant decrease in the quality of life (Ghosh and Mitchell, 2007). Children and youth affected with IBD are at risk for height constraint and postponement of adolescence (Stephens et al., 2001). CD patients are at an increased risk of mortality (Jess et al., 2013; Jussila et al., 2014) while individuals with UC are at an increased risk of colorectal cancer. The risk of colorectal cancer increases with the duration of disease; 20 years after UC diagnosis the probability of developing cancer is 8%, whereas after 30 years the probability increases to 18% (Eaden et al., 2001). Furthermore, a considerable proportion (25-40%) of individuals affected by IBD also experience extra-intestinal manifestations (Bernstein et al., 2001). These manifestations include osteopenia/osteoporosis, athropathy, skin complications, arthritis, thromboembolic events and ocular complications (Spekhorst et al., 2017). According to a recent Dutch IBD Biobank study which recruited 3388 patients from 2007 to 2017, the most frequent extraintestinal manifestation is osteopenia/osteoporosis which was observed in 20% of their cohort (Spekhorst et al., 2017).

1.1.2 Genetics of IBD

Although the exact mechanism giving rise to IBD remains unknown, several factors are recognized as contributors to the pathogenesis of disease; in a genetically susceptible host, an environmental stimulus leads to an aggressive immune response against the microbiota. Despite the identification of more than 200 gene loci associated with IBD (Jostins et al., 2012; Liu et al., 2015) , genetics accounts for only a small portion of the disease risk; between 8 and 12% of individuals with IBD have a family history of the disease (Halme et al., 2002; Santos et al., 2018). The most commonly altered gene observed in CD patients is the Nucleotide-binding oligomerization domain-containing protein 2 (NOD2) (Barrett et al., 2008). Normally, the intracellular presence of muramyl-dipeptide, a component of the peptidoglycan layer found in both Gram-negative and Gram-positive bacteria, leads to the activation of the NOD2 receptor, which in turn activates the nuclear factor kappa-light-chain-enhancer of activated B cells (NF-κB) transcription factor and thus leads to the expression of various cytokines including interleukin 12 (IL-12). Whether the mutations in the NOD2 gene lead to a reduced activation of the innate immune response, and thus a decrease in microbial clearance and a concurrent increase in microbial invasion, remains the subject of debate (Abraham and Cho, 2006). Alternatively alterations in the gene may result in an exaggerated inflammatory response if mutations in the NOD2 gene prevent it from dampening toll-like receptors 2 (TLR2) signaling (Borm et al., 2008). Notably, several of the IBD risk gene loci identified overlap with those from other inflammatory and immune mediated diseases including ankylosing spondylitis and psoriasis (Jostins et al., 2012). Enrichment analysis of the IBD risk loci also identified genes associated with primary immunodeficiencies and mycobacterial infection, providing further support to indicate an altered immune response in individuals with IBD (Jostins et al., 2012).

1.1.3 Environmental contributors of IBD

Numerous environmental factors have been linked to the disease including sunlight exposure/Vitamin D status (Szilagyi and Xue, 2018), diet (Ananthakrishnan et al., 2014), smoking (Mahid et al., 2006), antibiotics (Ungaro et al., 2014) and industrialization.

The risk of CD is linked to vitamin D levels and thus to sun exposure. A study examining the risk of IBD associated with exposure to sunlight demonstrated that the risk of CD, but not UC, was reduced with greater exposure to sunlight. Herein, woman in France exposed to a mean daily dose of ultraviolet radiation ≥ 1.75 kJ/m^2 were at a lower risk of CD (0.49 hazard ratio) compared to woman exposed to a mean daily dose < 1.51kJ/m^2 (Jantchou et al., 2014). Accordingly 25-hydroxy vitamin D [25(OH)D] plasma levels, converted from 7-dehydrocholesterol in part by UV B radiation, are associated with the risk of CD (Ananthakrishnan et al., 2012). Insufficient (< 20ng/mL) plasma levels of 25(OH)D leads to an elevated risk of intestinal surgery (Ananthakrishnan et al., 2013) and levels below 35ng/mL can predict relapse within one year with a specificity of 74% and sensitivity of 70% (Gubatan et al., 2017). The link between Vitamin D and IBD may be due to its ability to mediate components of both the innate and adaptive immune system (Ardesia et al., 2015). Examples of its influence on the immune system includes inhibiting the production of pro-inflammatory cytokines by T cells (Jeffery et al., 2009) and inducing the mRNA expression of the intracellular pattern recognition receptor (PRR) NOD2 in a variety of cell types (Wang et al., 2010).

Dietary composition is also related to the risk of IBD onset. A systematic review including 19 studies indicated that an elevated intake of meats, omega-6 fatty acid and polyunsaturated fatty acids are associated with an increased risk of both CD and UC, with the addition of total fat and saturated fats specific to CD and UC respectively (Hou et al., 2011).

Elevated levels of fiber and fruits reduced the risk of CD, but showed no effect on UC, whereas elevated vegetable consumption decreased the risk of UC. The importance of diet in IBD can also be exemplified by the use of exclusive enteral nutrition (EEN) for induction of remission in pediatric CD. In EEN, an oral or nasogastric tube is utilized to deliver a liquid diet, typically for 6-8 weeks (Ruemmele et al., 2014). Several hypotheses have been proposed to explain the contribution of diet to the etiology of IBD; for example, dietary components can modify the intestinal microbial communities and dietary antigen can stimulate immune cells (Lewis, 2016).

Smoking impacts the risk of developing IBD and influences disease course, displaying contradictory effects for CD and UC. Smokers have a reduced risk (odds ratio of 0.58) of suffering from UC, but an increased risk of developing CD (odd ratio 1.76), based on a meta-analysis of 15 articles published between 1984 and 2006 (Mahid et al., 2006). Smoking is also linked to disease outcome as individuals with CD that currently smoke face an elevated risk (hazard ratio 1.27) of intestinal resection compared to CD subjects whom have never smoked (Kuenzig et al., 2016). The precise mechanism by which smoking influences the onset and course of IBD remains undetermined, although the increase in dendritic cell number observed upon exposure to smoke has been proposed as a possible influence (Parkes et al., 2014). Furthermore a recent meta-analysis of 10 articles, published from 2010 to 2016, highlighted the compositional alterations smoking causes to the intestinal microbiome, including a decrease in the *Firmicutes* phylum and an increase in the *Bacteroidetes* and *Proteobacteria* phylums (Savin et al., 2018). The changes in *Firmicutes* and *Proteobacteria* are consistent observations with those also occurring in IBD (Frank et al., 2011; Gevers et al., 2014; Manichanh et al., 2006).

1.1.4 Dysbiosis of IBD

Intestinal dysbiosis, defined as *"a decrease in gut microbial diversity owing to a shift in the balance between commensal and potentially pathogenic microorganisms*" (Ni et al., 2017), is consistently associated with IBD, yet whether the alterations are the cause or consequence of the disease remains unknown. Although different microbiome alterations have been reported, consistent observations (Matsuoka and Kanai, 2015) include a reduction in intestinal biodiversity (Manichanh et al., 2006), an increase in *Proteobacteria* abundance (Frank et al., 2011; Gevers et al., 2014) and a decrease in the *Firmicutes* phylum abundance (Gevers et al., 2014). Altered bacterial composition can lead to differences in bacterial processes, functions and metabolites produced and thus affect host-microbe interactions. For instance the Faecalibacterium genus, capable of producing short-chain fatty acids (SCFA), is reduced in IBD (Sokol et al., 2009). SCFA are beneficial to the host as they are a main energy source for colonocytes and also have anti-inflammatory properties through NF-kB modulation (Inan et al., 2000). The importance of the microbiome in IBD is further exemplified by the use of an antibiotic, which causes a loss of diversity and may enable to expansion of pathogens. A study conduction on children born between 1995 and 2003 demonstrated an elevated rate ratio for IBD (1.84) in children having used antibiotics compared to their non-user counterparts (Hviid et al., 2011).

1.1.5 Immune function in IBD

Components of both the innate and adaptive immune system are modified in IBD. Treatment for IBD aims to restore the immune balance by the administration of immune modulators to ultimately achieve intestinal mucosal healing. Microbial and pathogen recognition by cells of the innate immune system is achieved through pattern recognition receptors (PRR), including TLRs, C-type lectin receptors, NOD-like receptors and RIG-I-like receptors (Takeuchi

and Akira, 2010). Activation of these receptors results in the transcription of inflammatory modulators to eradicate the noxious agent. The aforementioned NOD2 is the most widely studied PRR in CD (Barrett et al., 2008). Furthermore, differences in numerous innate immune cell types, including neutrophils, macrophages and dendritic cells are observed in IBD. Neutrophils, the first cells of the innate immune system to be recruited to the inflammatory site, are elevated in the peripheral blood (Hanai et al., 2004) and colon (Bennike et al., 2015) of UC patients. In CD patients, the colonic mucosa contains an elevated number of CD163+ macrophages (Demetter et al., 2005).

1.1.6 Intestinal barrier in IBD

The host's epithelium and mucus layer act as a physical barrier to limit the exposure of the microbes and their biomolecular products to the immune cells residing in the underlying lamina propria. Several features of the intestinal barrier are compromised in IBD including mucosal thickness (Chen et al., 2014), cell junctions (Das et al., 2012) and antimicrobial peptide expression (Wehkamp et al., 2005). Tight junctions, which link adjacent cells, play a major role in limiting paracellular transport and thus influence intestinal permeability. Alterations in tight junction proteins have been observed in IBD including the decreased expression of the tight junction scaffolding protein zonula occludens-1 (ZO-1) in both CD and UC (Das et al., 2012). Furthermore several cytokines, including those which are elevated in IBD, can modulate intestinal tight junctions (Ma et al., 2005; Mankertz et al., 2009). For instance TNF-α, a cytokine elevated in both CD and UC (Reimund et al., 1996), modulates the phosphatidylinositol-3-kinase pathway to increase both the transcript and protein expression of the pore forming, tight junction protein claudin-2 (Mankertz et al., 2009). Increased claudin-2 elevates the intestinal cationic permeability (Rosenthal et al., 2010). In IBD, the intestinal barrier is also compromised

at the level of the mucus layer. The examination of mucus thickness in rectal biopsies from 59 participants demonstrated a gradual decrease in mean thickness from controls (31.5µm), to mild (23.1µm), moderate (19.4µm) and severe UC (12.0µm) (Strugala et al., 2008). Reduced thickness was also observed in CD (15.4µm) although low CD patient enrolment prevented investigation of thickness based on severity. Goblet cell density (ratio of positive staining compared to the entire mucosa area) was also investigated, demonstrating decreased density in severe UC (0.058) but no difference observed between control (0.198) and CD (0.217).

1.1.3 Pediatric IBD displays phenotypic differences compared to adult IBD

Phenotypic differences are observed in childhood-onset IBD compared to adult-onset IBD, including affected anatomical location, disease course and prevalence between sexes (Sauer and Kugathasan, 2009). CD is more prevalent in male pediatric patients, whereas this trend is not observed in adult-onset CD, nor in adult or pediatric UC, where the number of affected males and females is approximately equal (Sauer and Kugathasan, 2009). Pediatric IBD has been suggested to represent a more severe phenotype as compared to adult-onset IBD, exemplified by the proportion of colon affected at diagnosis, and the number of individuals requiring colectomies. The reason for this phenotypic difference remains unknown (Sauer and Kugathasan, 2009). The majority of pediatric studies report up to 80-90% participants with UC pancolitis (Sauer and Kugathasan, 2009), wherein the inflammation extends beyond the hepatic flexure (Ng et al., 2013) whereas in adults extensive colitis is observed in only 19-31% of new-onset patients (Molinie et al., 2004; Portela et al., 2010). Accordingly, ten years after diagnosis, a greater proportion (40%) of pediatric UC patients require a colectomy compared to adults (20%) (Sauer and Kugathasan, 2009).

1.2 IBD diagnosis remains challenging and diagnostic delays impact disease outcome

IBD symptoms are non-specific and commonly include abdominal pain, bloody stool, diarrhea and weight loss (Henderson et al., 2014). Diagnosis is challenging as these symptoms overlap with a variety of other conditions (Levine et al., 2014). A recent review (Conrad and Rosh, 2017) listed seven categories of conditions which exhibit symptoms which overlap with those observed in IBD, highlighting the difficulty in diagnosing IBD. These categories are infection (e.g. Salmonella), inflammation (e.g. celiac), malabsorption (e.g. lactose intolerance), allergies and immunology (5-50% of individuals with primary immunodeficiencies have gastrointestinal disorders (Agarwal and Mayer, 2013)), rheumatology (e.g. sarcoidosis), oncology (e.g. lymphoma) and others (e.g. irritable bowel syndrome). For instance, a patient suffering from lactose intolerance is likely to present with chronic abdominal pain and diarrhea, both of which are among the most commonly observed symptoms in IBD. Accordingly, IBD diagnosis requires multiple procedures including physical examination, imaging, laboratory testing, histology and endoscopy and therefore suspected cases of IBD must be referred to a specialist (Ashton et al., 2017).

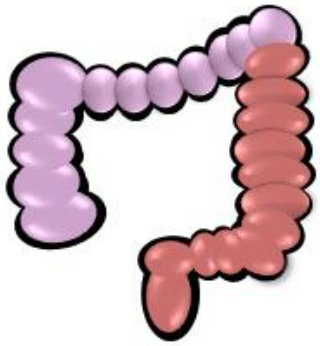
Following IBD and subtype (CD or UC) diagnosis, the activity index must be evaluated which allows physicians to monitor a patient's response to therapy. The Pediatric Crohn's Disease Activity Index (PCDAI) (Hyams et al., 2005) and the Pediatric Ulcerative Colitis Activity Index (PUCAI) (Turner et al., 2007) are calculated for CD and UC patients respectively. The PCDAI includes clinical symptoms (abdominal pain, general well-being and stool consistency), laboratory measurements (hematocrit, erythrocyte sedimentation rate (ESR) and albumin), and physical examination (changes in weight, height, abdominal tenderness, perirectal disease and extra-intestinal manifestations) (Dabritz et al., 2017). The PUCAI consists solely of

clinical symptoms: abdominal pain, rectal bleeding, stool consistency, number of stools, nocturnal stools and activity level (Dabritz et al., 2017). In UC adults, disease activity is measured according to the Mayo score which considers stool frequency, rectal bleeding, the appearance of the colonic mucosa upon endoscopy and the physician's assessment of disease activity (D'Haens et al., 2007; Lewis et al., 2008). For both adults and children with UC, the assessment of extent of disease is also necessary in order to assist in therapeutic delivery method choice. In adults this categorization is based upon the Montreal classification and consists of three possible classifications: E1 (proctitis; inflammation is restricted to the rectum), E2 (inflammation is distal to the splenic flexure) and E3 (inflammation is proximal to the splenic flexure and distal to the hepatic flexure) (Satsangi et al., 2006). A pediatric specific modification of the Montreal classification, known as the Paris Classification includes an additional category of E4 wherein the entire colon is inflamed (pancolitis; the inflammation is proximal to the hepatic flexure) (Levine et al., 2011) (**Figure 1.1**). Notably, patients with both primary sclerosing cholangitis and UC display atypical colon inflammation progression, including rectal sparing and backwash ileitis (Cleveland et al., 2018).

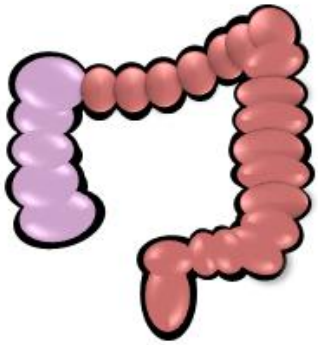
A study conducted on Canadian children with IBD demonstrated a mean of 4.5 months from symptom onset to diagnosis with 20% diagnosed more than one year after initial symptom commencement (Ricciuto et al., 2017). Diagnostic delay negatively impacts disease course by increasing the risk of intestinal surgery and bowel stenosis (Schoepfer et al., 2013). Furthermore, the importance of timely diagnosis enables early therapeutic intervention which achieves mucosal healing, remission, and thus reduces the risk of colorectal cancer (Triantafillidis et al., 2009).

E1

E2



E3



E4

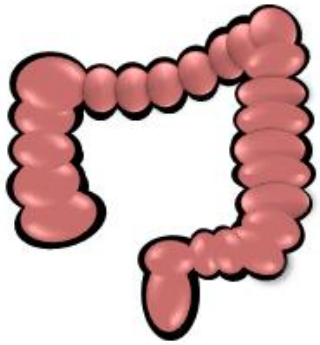


Figure 1. 1. Paris Classification for extent of disease in UC

Extent of disease classification indicates the proportion of the colon which is inflamed according to the Paris classification, which is a pediatric specific modification of the Montreal classification. Extent of disease is divided into 4 categories: E1 (proctitis; inflammation is restricted to the rectum), E2 (inflammation is distal to the splenic flexure), E3 (inflammation is proximal to the splenic flexure and distal to the hepatic flexure) and E4 (pancolitis; the inflammation is proximal to the hepatic flexure)

1.2.1 Fecal Calprotectin's biomarker performance is limited in IBD

Biomarkers currently used for IBD diagnosis simply act as adjuncts, as endoscopy is still required for diagnosis confirmation. The most routinely used biomarker in IBD diagnosis is fecal calprotectin (FC), wherein a positive (elevated) value dictates the need for further testing. Calprotectin is a calcium- and zinc-binding antimicrobial protein composed of two subunits, S100-A8 and S100-A9, accounting for 30%-60% of total cytosolic proteins in neutrophils (Hessian et al., 1993; Voganatsi et al., 2001). Calprotectin is released into the extracellular space following neutrophil degranulation. By sequestering zinc, it limits the growth of microbes (Lusitani et al., 2003) (Sohnle et al., 1991). Given that calprotectin is released upon neutrophil and monocyte stimulation, any condition leading to their intestinal infiltration results in elevated levels of FC, including infection (Gallo et al., 2017), polyps (Hodgson-Parnell et al., 2018; Mendoza and Abreu, 2009), celiac disease (Balamtekin et al., 2012), cystic fibrosis (Garg et al., 2017), colorectal carcinoma (Lehmann et al., 2014), ankylosing spondylitis (Klingberg et al., 2012), IBS (Melchior et al., 2017), Parkinson's disease (Schwiertz et al., 2018) and therapies using non-steroidal anti-inflammatory drugs (Meling et al., 1996). Furthermore, healthy children can have elevated levels of calprotectin up to the age of 9 (van Rheenen et al., 2010), possibly due to continuing digestive tract development (Song et al., 2017). Despite FC limitations in IBD diagnosis it still outperforms commonly employed blood parameters in the diagnosis of IBD including, ESR, C-reactive proteins, total white cell count, platelet count and albumin (Henderson et al., 2014). A meta-analysis of children undergoing diagnosis of suspected IBD yielded a specificity of only 0.682 (Henderson et al., 2014) whereas in adults, FC performance is superior yielding a pooled specificity of 0.93 (van Rheenen et al., 2010). Moreover, FC was recently tested in the primary care setting on adolescents and adults, yielding low specificity

(64.9%) and low sensitivity (72.7%) values (Conroy et al., 2017). Accurate and specific biomarkers for IBD diagnosis remain an unmet clinical need. This need is accentuated for pediatric patients as the performance of current clinical biomarkers, namely FC, is inferior for this particular population. Reliable, accurate and non-invasive biomarkers for IBD diagnosis would bypass the costly expense, discomfort and risk of complications which accompany colonoscopies. In 2007 the cost of a single colonoscopy in Canada (Quebec) was $340 (Sharara et al., 2008). Considering that more than 10,000 Canadians are diagnosed with IBD every year this sum amounts to more than 3.4 million in diagnostic colonoscopy fees alone. Moreover, this sum does not include suspected cases undergoing colonoscopy which are ultimately not diagnosed with IBD. This burdening expense is accentuated in the US wherein the fee for upper gastrointestinal endoscopy with colonoscopy ranges from $700-$2000 (Yang et al., 2014). Colonoscopy is also accompanied with the risk of complications. A study encompassing 12 U.S centers collected data from 2000 to 2005 yielding 7,792 colonoscopies performed on patients < 18 years of age to assess the frequency of complications associated with pediatric colonoscopy. A total of 88 (1.1%) encountered complications with gastrointestinal bleeding (34) and hypoxia (22) most frequently observed (Thakkar et al., 2008).

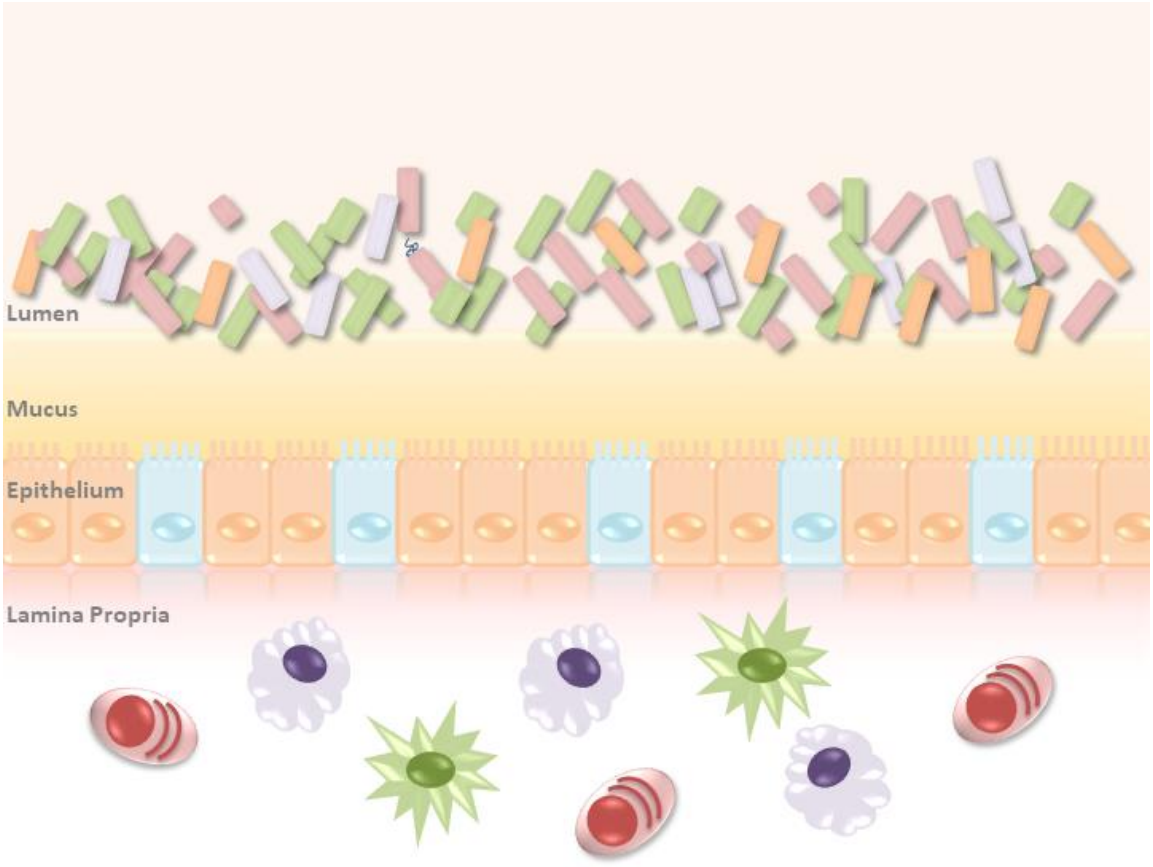
1.2.2 UC extent of disease diagnosis requires endoscopy and/or imaging

Generally in UC the inflammation begins in the rectum and progresses proximally towards the ascending colon. Extent of disease classification indicates the proportion of the colon which is inflamed according to the Paris classification, divided into 4 categories with ascending level of disease E1, E2, E3 and E4 (**Figure 1.1**). Individuals with UC must undergo assessment for extent of disease in order to aid in determining appropriate therapy in addition to conclude the rate at which colorectal cancer monitoring must ensue, in both adults and children (Rosen et

al., 2015; Stange et al., 2008). Patients with pancolitis (E4) must undergo endoscopic cancer surveillance 8 to 10 years following diagnosis whereas patients with E2 must begin monitoring 15-20 years after diagnosis (Eaden et al., 2002). Extent of disease in UC also influences the method of therapeutic delivery. For patients with limited inflammation, topical methods are preferred for treatment, suppositories for E1 (proctitis) and enemas for E2 patients (Dignass et al., 2012). For instance, studies have shown that liquid enemas are capable of delivering therapeutics to the splenic flexure, while foam enemas delivery therapies to the sigmoid and suppositories are effective for treating the rectum (Lawrance, 2011). Combination therapies consisting of both oral and topical are implemented for patients with extensive disease (E3 or E4) (Dignass et al., 2012). Extent of disease diagnosis is established using imaging and endoscopy; biomarkers capable of indicating extent of disease have not been developed. While colonoscopy is the preferred procedure for establishing extent of disease in UC (Dignass et al., 2012) x-ray barium test can also be utilized, although the exposure to radiation increases the risk of mutagenesis and thus the development of cancer (Kleinerman, 2006). Non-invasive biomarkers with the ability to delineate extent of disease would be beneficial for diagnosis in order to avoid the complications and cost associated with endoscopy and imaging procedures.

1.4 The intestinal mucosal-luminal interface is critical in IBD but has received limited investigation

The intestinal MLI is the site of interaction between the host and its commensal resident microbiota (**Figure 1.2**). The large intestine, which is commonly affected in IBD, is the most heavily colonized anatomical location in humans (Lawley and Walker, 2013). The intestinal MLI encompasses the host's epithelium lined with a thick mucus layer which is in direct contact with the microorganisms residing within the intestinal lumen. Abnormal interactions between the host


Lumen
Mucus
Epithelium
Lamina Propria

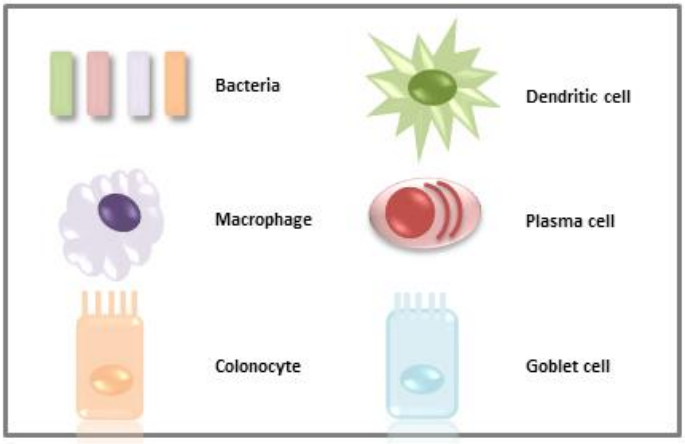

Bacteria
Dendritic cell
Macrophage
Plasma cell
Colonocyte
Goblet cell

Figure 1. 2. The intestinal mucosal-luminal interface

The intestinal MLI encompasses the host's epithelium lined with a thick mucus layer which is in direct contact with the microorganisms residing within the intestinal lumen. Abnormal interactions between the host and its commensal intestinal microbiota play a role in the pathogenesis of IBD.

and its commensal intestinal microbiota play a role in the pathogenesis of IBD and thus investigations examining alterations at the intestinal MLI are imperative to unraveling the mechanisms of disease development. Furthermore, sampling the affected site directly is critical for understanding the alterations associated with the disease, as proxy biological samples such as stool do not fully recapitulated the mucosa composition (Durban et al., 2011). A study comparing rectal biopsies to stool samples collected from nine healthy participants by 16S sequencing demonstrated that rather than grouping according to the individual, samples tended to group by sample type (biopsy or stool sample) when considering the family and genus levels. Furthermore, stool samples demonstrated reduced diversity compared to biopsies exemplified by rarefaction curves (Durban et al., 2011). Moreover, many of the limitations related to biomarker discovery can be mitigated by sampling biological fluids directly from the disease affected area (Rifai et al., 2006).

In a study applying 16S sequencing to the ascending colon MLI of new-onset pediatric patients Veillonella, Enterobacteriacea, Fusobacterium, Neisseria and Haemophilus were identified as reduced taxa in CD patients compared to non-IBD controls (Mottawea et al., 2016). Furthermore, a total of 161 operational taxonomic units (OTUs) correlated with CD disease severity (mild, moderate or severe), with OTUs capable of butyrate production and H_2S production being negatively and positively correlated with disease severity, respectively. Another study examined the MLI at the cecum and sigmoid of 47 adults by both 16S sequencing and ultra performance liquid chromatorgraphy mass spectroscopy to investigate microbial-metabolite relationship. A total of eight genera (*Bacteroides, Blautia, Clostridium, Coprococcus, Faecalibacterium, Oscillospira Roseburia* and *Ruminococcus*) correlated with the greatest number of metabolites leading the authors to hypothesis that this may indicate that these genera

as highly metabolically active or produce the greatest diversity in metabolites (McHardy et al., 2013). Despite the importance of this anatomical site in the context of IBD, there have been few proteomic investigations of the intestinal MLI, and both are limited to adults who had received therapeutic intervention (Li et al., 2016; Presley et al., 2012). The first study to perform proteomics on the intestinal MLI from IBD patients collected samples from the sigmoid and cecum from adults with inactive disease (Presley et al., 2012). In order to assess bacterial composition, oligonucleotide fingerprinting of rRNA genes (OFRG) was also implemented. In OFRG, an array of sample derived bacterial rRNA gene clones undergo serial hybridization using different single DNA probes (Valinsky et al., 2002). Using surface-enhanced laser desorption/ionization mass spectrometry and OFRG on an initial cohort of 24 MLI samples identified 87 proteins and 3,374 phylotypes respectively. From this initial analysis, 14 bacterial phylotypes which correlated with protein levels were selected for further analysis using an independent second cohort of 42 participants. Differences in 5 of the 14 bacterial phylotypes were observed for disease status (IBD versus control) and between different colon sub-regions (sigmoid versus cecum). A combination of matrix-assisted laser desorption/ionization mass spectrometry (MALDI-MS) and qPCR on the second cohort of 42 participants yielded 8 bacteria phylum which exhibited the strongest correlation with 43 proteins. Several of the 43 protein features were unassigned, with the majority of the identified proteins corresponding to haptoglobin. Interestingly some innate immune proteins were identified among the strongest correlated proteins including neutrophil defensin 3 and complement C3a. This study provided the initial proteomic MLI findings to aid in deciphering the complex host-microbe alterations which are occurring at the site of disease in IBD. Another proteomic study assessed the intestinal MLI collected from non-inflamed sites of adult IBD patients undergoing cancer surveillance (Li et al.,

2016). In a discovery cohort consisting of 51 IBD and non-IBD controls, intestinal MLI samples were collected from six colon sub-regions (cecum, ascending, transverse, descending, sigmoid and rectum). Using semi-quantitative proteomics (MALDI-TOF-MS) a total of 599 features (peaks m/z) were identified and yielded 166 colon sub-region significant features and 85 significant disease-related features. Weighted correlation network analysis was then applied in order to assess protein modules, which are defined as a collection of proteins with common expression and localization, yielding nine protein modules. Three of the nine modules were associated with colon sub-regions and six were related to disease, deemed by the Kruskal-Wallis test. A subset of protein modules was further validated, including a module which demonstrated both disease and colon sub-region differences. By *in silico* analysis and immunoprecipitation, the proteins within this module were identified as the α-defensins human neutrophil peptide 1 (HNP1), HNP2 and HNP3. Elevated expression in disease compared to controls of HNP1-3 was confirmed by immunoblotting, whereas immunohistochemistry demonstrated that the expression of HNP1-3 was due to neutrophil infiltration, a known hallmark of IBD (Ina et al., 1997). Finally mucosal whole-mount immunohistochemistry demonstrated a mosaic pattern of expression of HNP. Notably, this study demonstrated the presence of neutrophilic infiltrate in the absence of macro-inflammation in the distal colon. The authors of this study hypothesized that "*mucosal functional networks*" of millimeter scale line the intestinal mucosa rendering a heterogeneous surface, providing evidence that host-microbe interactions occurring within a few millimeters of each other may differ considerably at the intestinal MLI.

1.5 Extracellular vesicles: vehicles of intercellular communication

Extracellular vesicles (EVs) are membrane encapsulated structures that are released from cells, and participate in intercellular communication by delivering biomolecules (Zernecke et al.,

2009). EVs subtypes include exosomes, shed microvesicles, apoptotic bodies and bacterial outer membrane vesicles (OMVs) which differ in size, composition and biogenesis.

1.5.1 Mammalian (host) EVs

Exosomes range in size from 30nm-100nm and are derived from the endosomal pathway. Despite recent advances in the field that have implicated multivesicular body (MVB) and intraluminal vesicle (ILV) formation in the development of exosomes, the exact biogenesis mechanisms of exosomes remain unknown (Xu et al., 2016). The endosomal sorting complex required for transport (ESCRT) is required for both MVB and ILV formation; accordingly, several ESCRT components including tumor susceptibility gene 101 (TSG101), programmed cell death 6-interacting protein (PDCD6IP) and vacuolar protein sorting-associated protein 4A (VPS4A) are considered exosome biomarkers (Kowal et al., 2016). Other exosomal markers include the Rab GTPases members Rab27a and Rab27b (Ostrowski et al., 2010) and the Tetraspanins CD9, CD63 and CD81(van Niel et al., 2011). Similar to exosome formation, ESCRT machinery has also been linked to shed microvesicles biogenesis (Nabhan et al., 2012). Whereas exosomes arise from the endosome, shed microvesicles are derived from the outward budding of the plasma membrane, and their precis formation has yet to be fully elucidated. Apoptotic bodies are formed during apoptosis wherein the cell is fragmented into extracellular vesicles of 0.5-4μm in size (Ihara et al., 1998). Due to their formation from the entire cell, unlike exosomes or shed microvesicles, apoptotic bodies contain organelles (Elmore, 2007).

Components of EVs are being explored as biomarkers (Akers et al., 2013; Li et al., 2014) as their cargo is protected, and is reflective of the cell type and cell state from which they were derived. EVs have been isolated from a plethora of biological fluids including but not limited to urine (De Palma et al., 2016), saliva (Zheng et al., 2017), cerebral spinal fluid (CSF) (Akers et

al., 2013), blood (Leonetti et al., 2013) and breast milk (Karlsson et al., 2016). Urinary EVs have been proposed as biomarkers for prostate cancer (Sequeiros et al., 2017), serum EVs for biomarkers of colorectal cancer (Dong et al., 2016) and nonalcoholic fatty liver disease (Povero et al., 2014), and CSF EVs for biomarkers of glioblastoma (Akers et al., 2013).

EVs are also being harnessed for their therapeutic delivery potential as they can deliver molecular cargo to neighboring or distant cells and thus alter the recipient cells' state (Gyorgy et al., 2015). Intrinsically produced EVs, those loaded with specific cargo and artificially produced vesicles have all been employed for therapeutic delivery (Gyorgy et al., 2015). In a recent study, exosomes isolated from human umbilical cord mesenchymal stem cells (hucMSCs) were used to treat a dextran sodium sulfate (DSS) colitis mouse model (Mao et al., 2017). Exosomes injected through the tail vein localized to the liver, spleen and colon and significantly mitigated the weight loss and colonic expression of the pro-inflammatory cytokines TNF-α, IL-1β and IL-6 that are associated with the DSS treatment. Furthermore, DSS treated mice demonstrated elevated CD206+ macrophage infiltration compared to mice treated with hucMSCs derived exosomes. Moreover, co-culturing of hucMSC exosomes with peritoneal macrophages lead to a decreased mRNA expression of TNF-α, IL-1β and IL-6 in macrophages. This study provides evidence that exosomes may represent an alternative therapeutic option for IBD treatment.

1.5.2 Bacterial outer membrane vesicles

Similar to eukaryotes, bacteria also secrete membrane-bound vesicles into the extracellular environment. EVs released from Gram-negative bacteria are termed outer membrane vesicles (OMVs) whereas those released from Gram-positive bacteria and archaea are termed membrane vesicles (MVs). Bacterial OMVs range in size from 20nm-300nm (Klimentova and Stulik, 2015) and their biogenesis is influenced by a variety of factors

including temperature (McMahon et al., 2012), alterations of enzymes implicated in peptidoglycan dynamics (Park et al., 2012) and levels of misfolded proteins (Schwechheimer et al., 2014). During OMV formation the outer membrane separates from the peptidoglycan layer, expands, and is pinched off from its cell of origin. OMVs carry DNA, RNA and proteins arising from the outer membrane (OM), periplasm and cytoplasm (Kaparakis-Liaskos and Ferrero, 2015), and can thus participate in a range of physiological processes including intercellular bacterial communication, thereby defending bacteria from antimicrobial peptides and bacteriophages (Manning and Kuehn, 2011; Reyes-Robles et al., 2018), delivering virulent factors (Fiocca et al., 1999), acquiring iron and other elements required for bacterial growth, destroying competitor bacteria and providing antibiotic resistance to bacterial cells. Bacterial OMVs can also provide benefits to the host, exemplified by the vesicular trans-kingdom delivery of a mammalian $InsP_6$ phosphatase homologue, the *Bacteroides thetaiotaomicron* mammalian $InsP_6$ phosphatase (BtMinpp) (Stentz et al., 2014). *B. thetaiotaomicro* is a commensal human intestinal bacterium, that can package BtMinpp into OMVs, which interact with intestinal epithelial cells. Delivering this phosphatase to host epithelial cells can contribute to host $InsP_6$ metabolism, which in turn leads to modulating energy metabolism (Szijgyarto et al., 2011) amongst other cell function alterations. In another recent study, OMVs derived from *Akkermansia muciniphila* were shown to improve host epithelial barrier integrity by elevating the expression of the tight junction protein occludin (Chelakkot et al., 2018). Conversely bacterial OMVs can contribute to the pathogenesis of various diseases and infections by delivering virulent factors and by modulating components of the host immune system (Shah et al., 2012). Accordingly, OMVs have been isolated from infected host biological fluids and tissues (Stephens et al., 1982), including from the gastric mucosa of an individual infected with

Helicobacter pylori (Fiocca et al., 1999). Bacterial OMVs have also been linked to IBD. OMVs derived from an adherent-invasive *Escherichia coli* (AIEC) strain LF82 isolated from an ileal lesion of a CD patient influenced epithelial cell invasion, in an OmpA-dependent manner (Rolhion et al., 2010). AIEC strains are more often present in the chronic (30.4%) and early (36.4%) CD ileal lesions compared to control ileal tissue (6.2%) (Darfeuille-Michaud et al., 2004). OmpA is found on the surface of OMVs and has been implicated in invasion and adhesion (Nicholson et al., 2009). A reduction in both the fusion and invasion of OMVs with Intestine-407 epithelial cells was observed upon OmpA gene deletion in the LF82 strain (Rolhion et al., 2010). Despite recent advances in the field of bacterial OMVs, their cargo within the complex community of the human intestine remains unknown.

OMVs are also being explored as potential therapeutics. In a recent study, OMVs derived from *Pseudomonas aeruginosa* were administered to mice with acute lung infection, leading to a decrease in the release of the pro-inflammatory cytokines TNF-α, IL-1β and IL-6, a reduction in bacterial lung colonization and lessened infiltration of inflammatory cells (Zhang et al., 2018b). OMVs can be uptaken by host cells using several different pathways including the lipid raft-mediated pathway (Kesty et al., 2004) and the clathrin-mediated endocytosis pathway (Canas et al., 2016).

1.6 Limited investigation of EVs in IBD

EVs can promote or inhibit an immune response and activate immune cells (Singh et al., 2012), rendering this deliver vehicle particularly relevant in the context of IBD which is an immune mediated chronic disease. There have been few investigations of EVs isolated from patients with IBD (Leonetti et al., 2013; Leoni et al., 2015; Mitsuhashi et al., 2016; Palkovits et al., 2013; Zheng et al., 2017); with the majority of studies limited by the inclusion of patients

with established disease and ongoing therapy. In a study examining treated adult IBD patients undergoing bowel resection, serum derived EVs contained elevated levels of the protein Annexin-A1 compared to healthy participants (Leoni et al., 2015). To further evaluate the role of Annexin-A1, nanoparticles containing a mimetic Annexin-A1 peptide injected into the mucosa lead to higher wound closure compared to the peptide alone or nanoparticles which contained a scrambled peptide. Another study isolated microparticles, also known as microvesicles, from platelet-free plasma of adult CD patients with and without active disease, demonstrating that CD participants with active disease had an elevated number of microparticles compared to those with inactive disease and to healthy participants (Leonetti et al., 2013). Furthermore patients with active CD demonstrated an increase in the number of microparticles derived from platelets ($CD41^+$) and leukocytes ($CD45^+$) as assessed by flow cytometry. Another study isolated EVs from the saliva of new-onset adult IBD participants, including 48 IBD (11 CD, 37 UC) and 10 healthy controls. LC-MS/MS analysis identified eight proteins which were identified in both CD and UC but not in healthy controls. Immunoblotting was performed to validate the differential expression of the proteasome subunit alpha type 7 (PSMA7) which was among the eight proteins unique to IBD (Zheng et al., 2017). Moreover, only a single study confounded by treatment intervention has isolated EVs from the intestinal MLI of IBD patients (Mitsuhashi et al., 2016). This study demonstrated an elevated expression of IL-6, IL-8, IL-10 and TNF-α in EVs isolated from areas with active disease (Mayo score > 1) compared to healthy participants. EVs from unaffected areas (Mayo score= 0) demonstrated a similar level of expression of the aforementioned cytokines to that of the healthy individuals. This study provided the initial findings to aid in elucidating the impact of intestinal extracellular vesicles at the site of disease in IBD, although the overall proteomic landscape of EVs isolated from the intestines remains

unknown. Furthermore no studies have investigated EVs from pediatric IBD participants, a population which demonstrates a phenotypically distinct form of the disease as compared to adults.

1.7 Proteomics for biomarker discovery

Proteomics is a powerful tool, capable of quantifying the relative expression of hundreds to thousands of proteins within a single MS analysis run. The identification of differentially expressed proteins can provide valuable insight into a plethora of biological processes including the elucidation of disease pathogenesis, mechanisms of development, biomarker discovery and drug response. Protein profiling by MS relies on two main approaches: bottom-up or top-down proteomics. The former entails the digestion of proteins into peptides prior to MS analysis while the latter employs intact proteins for MS analysis. Bottom-up proteomics is compatible with complex protein mixtures whereas top-down is limited to isolated proteins or simple protein mixtures due to current instrument restrictions. Peptides are more readily fractionated, ionized and fragmented compared to intact proteins, rendering bottom-up proteomics more widely and routinely implemented compared to top-down (Zhang et al., 2013). In bottom-up proteomics, peptides are assigned to their corresponding protein by the application of a database search strategy. This search strategy relies on the comparison of experimental spectras to theoretical spectra's digested *in silico*.

Several approaches have been developed for quantitative proteomics including metabolic labeling, chemical labeling, spike-in peptides and label free quantification (LFQ), each associated with its own limitations and strengths. Stable-isotope labeling by amino acids in cell culture (SILAC) is a metabolic labeling approach wherein isotopically labeled amino acids are substituted in the growth media for incorporation into cells (Ong et al., 2002). Although this

method permits accurate quantification, it cannot be implemented on human biological specimen, as it requires isotopic labeling of the organism. Super-SILAC addresses this limitation as a representative heavy isotopic protein mixture is combined with the protein sample of interest and therefore does not require metabolic labeling of the organism of interest (Geiger et al., 2010). Given that SILAC strategies implement labeling prior to downstream sample preparation for MS (digestion, fractionation and desalting); both the sample of interest and the label are similarly affected by losses encountered throughout these procedures resulting in increased quantification accuracy. Tandem mass tags (TMT) is a chemical labeling strategy in which the label reacts with the digested peptides primary amine (Thompson et al., 2003). This costly technique can be applied to human biological specimen as the labeling occurs post protein digestion, although losses encountered throughout the digest procedure will not be accounted for in the quantification. LFQ is inexpensive and does not require the presence of an isotopically heavy counterpart for quantification although the quantification accuracy is limited by the degree of variability in sample preparation and LC-MS consistency (Xie et al., 2011).

Proteomics has been extensively used for the identification of diagnostic, prognostic and response to treatment biomarkers. Biomarker discovery is limited by multiple challenges including a wide dynamic range and low expression of potential disease related biomarkers in biological specimens or fluids remote from the area of disease (Rifai et al., 2006). The protein dynamic range of biological fluids can extend to 12 orders of magnitude (Roche et al., 2008). Furthermore, in plasma 22 proteins account for nearly 99% of the plasma protein mass (Borg et al., 2011) which imposes limitations on MS analysis as it is restricted by the number of peptides which can be selected for fragmentation, which in turn limits the number of the peptide product ions detected for identification and quantification. In data dependent acquisition, high abundant

proteins are selected for fragmentation and therefore low abundant proteins in a complex mixture may not be detected during a MS analysis. Disease-specific proteins are therefore more likely to be detected from biological specimen or fluid samples which are in direct contact with the site of disease compared to those which are distant.

1.8 Rationale, hypothesis and objectives

Reliable biomarkers are needed for IBD to allow for rapid diagnosis and thus mitigate the risk of negative outcomes which are associated with delayed diagnosis, which include intestinal surgery, bowel stenosis and colorectal cancer (Schoepfer et al., 2013; Triantafillidis et al., 2009). Furthermore non-invasive biomarkers would bypass the costly expense, discomfort and potential complications that are associated with colonoscopies. Similarly, developing biomarkers that can accurately assess UC extent of disease would avoid the negative features which accompany endoscopy and imaging procedures. Biomarker discovery is limited by multiple challenges including a wide dynamic range and low expression of potential disease related biomarker in biological specimen or fluids remote from the area of disease (Rifai et al., 2006). The evaluation of biological fluids derived from the affected area in disease can therefore mitigate these obstacles (Rifai et al., 2006). I hypothesized that specific proteins expressed at the intestinal MLI, a biological fluid which is in direct contact with the site of disease, exhibit differential expression in IBD compared to control patients and thus have utility as IBD biomarkers.

My specific hypotheses were

1- The intestinal MLI of pediatric IBD patients exhibit global proteomics alterations compared to non-IBD controls

2- Specific proteins expressed at the intestinal MLI demonstrate differential expression between non-IBD and IBD patients and could therefore serve as biomarkers for pediatric IBD diagnosis

3- Proteins expressed at the MLI at the AC display differential expression in UC patients in areas with macroscopic inflammation compared to areas without macroscopic inflammation and thus have value in differentiating pancolitis and non-pancolitis

My objectives were to

1-Investigate the proteomic alterations at the intestinal MLI in new-onset pediatric IBD as compared to non-IBD controls

2-Identify diagnostic biomarkers for pediatric IBD

3- Identify biomarkers for UC extent of disease (pancolitis versus non-pancolitis)

Additionally, considering that EVs can deliver molecular cargo to neighboring or distant cells and thus alter the recipient cells' state and have the ability to modulate immune responses I hypothesized that

4- Intestinal EV proteomic cargo is altered in IBD compared to non-IBD controls

5- Select proteins exhibit differential expression in EVs in IBD compared to non-IBD controls and do not demonstrate differential expression in their corresponding biopsies.

My objectives werer to

4- Assess the proteomic differences in new-onset pediatric IBD intestinal EVs as compared to non-IBD controls

5- Compare proteomic signatures of EVs and their corresponding intestinal biopsies

CHAPTER 2. MATERIALS AND METHODS

2.1 Patient Cohort

Intestinal MLI samples were collected from patients <18 years old undergoing diagnostic colonoscopy at the Children's Hospital of Eastern Ontario (CHEO) from November 2013 to July 2016. Clinical sample collection was performed by Dr. David Mack, Dr, Eric Benchimol and Ruth Singleton. To eliminate the confounding effects of medication, patients were omitted from study if they had used antibiotics, probiotics or immunosuppressives within 1 month prior to colonoscopy (Dethlefsen et al., 2008). Moreover, in order to study the host proteome modifications, factors known to alter the intestinal microbiome and therefore affect the host response were excluded: (1) >95th percentile body mass index (2) diabetes mellitus (3) infectious gastroenteritis within 2 months prior to colonoscopy. Only those patients with a clear diagnosis at the time of diagnostic colonoscopy were included in the intestinal MLI bioinformatic analysis. The Research Ethics Board of CHEO approved this study.

Patient data were collected and managed using REDCap electronic data capture tools hosted at the CHEO Research Institute. REDCap (Research Electronic Data Capture) is a secure, web-based application designed to support data capture for research studies (Harris et al., 2009).

2.1.2. Classification standards

Diagnosis of IBD was achieved by clinical examination, endoscopy, imaging and laboratory testing, as per standard procedure (Levine et al., 2014). Diagnosis was performed by Dr. David Mack and Dr. Eric Benchimol. The Pediatric Crohn's Disease Activity Index (PCDAI) was implemented for CD (Hyams et al., 2005) and the Pediatric Ulcerative Colitis Activity Index

(PUCAI) was implemented for UC (Turner et al., 2007). Visual appearance of mucosa by colonoscopy was utilized to assess inflammation of the ascending colon (AC) and the descending colon (DC). The Paris modification of the Montreal Classification for IBD (Levine et al., 2011) was employed to determine the extent of macroscopically inflamed mucosa.

2.2. Intestinal MLI protein processing and analysis

Proteins were isolated from the intestinal MLI of 60 new-onset treatment-naive pediatric IBD and non-IBD patients, processed for MS analysis, and bioinformatic analysis performed in order to assess the proteomic landscape changes and develop biomarker panels for IBD diagnosis and UC extent of disease (pancolitis vs non-pancolitis) as described in the subsections below.

2.2.1. Intestinal MLI sample collection

After a one day clean-out preparation (Jimenez-Rivera et al., 2009), intestinal MLI aspirates were collected at diagnostic colonoscopy. Following introduction of the colonoscope, initial fluid containing debris was aspirated away. The mucosa was then flushed with sterile water and the aspirate was collected in a sterile vial, henceforth known as the intestinal MLI sample. When collected from an area with macroscopically visible inflammation, the sample was labeled as CoA (i.e., colon affected) and upon collection from an unaffected area without the presence of macroscopic inflammation, the sample was labeled as CoN (i.e., colon non-affected). Following collection, the samples were kept on ice in the endoscopy suite and transported to the laboratory for processing. Due to variations in patients, including the extent of disease and level of precleaning, the amount of water used to flush the mucosa could not be fixed, but rather ranged from 100-125mL for the AC and 150-200mL for the DC. The recovery yield is unknown, although the approximate volume collected is 60mL for the AC and 90mL for the DC. Notably,

upon receipt into the lab the intestinal MLI samples were aliquoted to allow use for multiple studies. After the aliquot was received, a complete protease inhibitor cocktail (Roche Diagnostic GmbH, Mannheim, Germany) was added (tablets are provided as 10X, added to MLI sample to yield a 1X concentration). Debris (insoluble material) was depleted by centrifugation at 700 x g for 5 min at 4 °C. Although we did not characterise the debris, a study by Li et al., 2011 determined it consisted of bacterial cells, human cells and food components. The resulting supernatant was then centrifuged at 14,000g for 30 min at 4°C to pellet out bacteria. For intestinal MLI aspirate samples, the supernatant was filtered using a 0.22µm filter while those intended for EV isolation were filtered using a 0.45µm filter. A larger pore size filter was used for EV isolation in order to enable the inclusion of bacterial OMVs which range from 20nm-300nm (Klimentova and Stulik, 2015). Following filtration samples were stored at -80°C until use.

2.2.2. Heavy reference proteome for internal spike-in

After testing different combinations of isotopically-labeled human cell lines for complementarity to intestinal MLI samples, five were selected (Jurkat (ATCC), HEK-293 (ATCC), HCT 116 (ATCC), THP-1 (ATCC) and hepatic HuH-7 (JCRB Cell Bank)) as a Super-SILAC (Geiger et al., 2010) internal heavy reference to perform relative quantification. HuH7-7, HEK293 and HCT116 cells were grown in a 5% CO_2 humidified incubator at 37° C using a custom DMEM meida deficient in lysine, arginine and methionine, purchased from AthenaES (Baltimore, MD, USA). Media was supplemented with 10% (v/v) dialyzed FBS (GIBCO Invitrogen; Burlington, ON, Canada), 1 mM sodium pyruvate (Gibco Invitrogen), 28 g/mL gentamicin (Gibco Invitrogen), 30mg/L methionine (Sigma-Aldrich; Oakville, ON, Canada), 146mg/L [$^{13}C_6$, $^{15}N_2$]-L-lysine and 42 mg/L [$^{13}C_6$, $^{15}N_4$]-L-arginine (84mg/L for HCT116) (Sigma-Aldrich;Oakville,

ON, Canada). Jurkat and THP-1 cells were cultured in RPMI media deficient in lysine, arginine and methionine (#0422 AthenaES) that was then supplemented with 10% (v/v) dialyzed FBS (GIBCO-Invitrogen), 1 mM sodium pyruvate (Gibco-Invitrogen), 28 µg/mL gentamicin (Gibco-Invitrogen), 0.0059g/L Phenol Red (Sigma-Aldrich), 15 mg/L methionine (Sigma-Aldrich), 40 mg/L [$^{13}C_6$,$^{15}N_2$]-L-lysine and 200 mg/L [$^{13}C_6$,$^{15}N_4$]-L-arginine (Sigma Aldrich). Incorporation of heavy-labeled amino acids into the cells at a level >95% was confirmed by MS analysis. Lysis of cells was performed using 4% SDS (w/v), 50 mM Tris, pH 8.0 supplemented with a protease inhibitor cocktail tablet (Roche Diagnostic GmbH, Mannheim, Germany). Following lysis, cells were sonicated for 10 seconds followed by 30 seconds cooling on ice, which was then repeated three times. Lysates were cleared by centrifugation at 10,000g for 10 minutes at 4°C. The supernatant was collected, and protein concentration was determined by the detergent compatible protein assay (BIORAD, California, USA). Proteins were aliquoted and stored at -80°C.

2.2.3. Preparation of intestinal MLI aspirate proteins for MS analysis

Due to the majority of intestinal MLI samples below the detection limits of the detergent compatible protein assay we proceeded with precipitation. After thawing on ice, the intestinal MLI aspirates (depleted of debris and bacteria) were incubated overnight at 4°C with trichloroacetic acid (20% v/v) to precipitate proteins. The following day, the acidified intestinal MLI aspirates were centrifuged at 13,000g for 30 minutes at 4°C. The supernatant was discarded and the protein pellets were washed three times with 1mL acetone (100%) to remove acid. Air-dried protein pellets were resuspended in a lysis buffer consisting of 4% SDS (w/v), 50 mM Tris, pH 8.0 supplemented with a protease inhibitor cocktail tablet (Roche Diagnostic GmbH), estimated based on pellet size. Protein concentration was determined using the detergent compatible protein assay (BIORAD, California, USA). Intestinal MLI aspirate proteins (45ug)

were mixed 1:1 w/w with the heavy reference proteome (45ug total; 9ug of each of the 5 cell types) to enable for relative and sample-comparable quantification by Super-SILAC(Geiger et al., 2010). The combined protein mixture (i.e., intestinal MLI aspirate and heavy internal reference) was digested by filter aided sample preparation (Wisniewski et al., 2009). Briefly, 3kDa filters were prepared by washing with 300uL of ddH_2O. The protein mixture was then added to the filter along with 200uL of 8M urea, 50mM Tris pH 8.0. Filters were centrifuged at 16,000g for 13 minutes at RT, then washed with another 200uL 8M urea, 50mM Tris pH8.0. Proteins retained in the upper filter chamber were reduced using 20mM dithiothreitol (DTT) in 8M urea, 50mM Tris pH8.0 for 30 minutes at 37°C with agitation. Residual DTT was removed by centrifugation. Alkylation was performed using 20 mM iodoacetamide in 8M urea, 50mM Tris pH 8.0 for 30 minutes at RT and protected from light. Following alkylation the filter was washed with 8M urea, 50mM Tris pH 8.0, and then twice with 50mM Tris-HCl pH 8.0 to facilitate removal of residual SDS. Trypsin (1ug) diluted in 200µL of 50mM Tris-HCl pH 8.0 was added to the filter and incubated overnight at 37°C with agitation. Peptides were eluted and collected by centrifugation at 16,000g for 13 minutes at RT; a wash of 200uL of 50mM Tris-HCl pH 8.0 was added to the filter and centrifuged to maximize recovery.

Following trypsin digestion, the peptides were fractionated into 5 fractions using a strong cationic exchange (SCX) resin (Agilent Technologies, CA, USA). Briefly, peptides were first acidified with the addition of 1% (v/v) formic acid to achieve a pH of ~ 2. A total of 5mg of SCX resin was equilibrated by washing five times with 0.5% (v/v) formic acid. Following equilibration the acidified peptide sample was loaded and subsequently eluted into 5 fractions using elution buffers of increased alkalinity (Britton-Robinson buffer 0.04 M boric acid, 0.04 M phosphoric acid and 0.04 M acetic acid; adjusted to pH 4, 6, 8, 10 and 12). The resulting five

peptide fractions for each sample were desalted using a 10µm AQUA-C18 resin (Dr Maisch, GmbH, Ammerbuch, Germany). A total of 5mg of resin was activated by the addition of three washes of 100% acetonitrile. The resin was subsequently equilibrated by the addition of two washes of 0.1% (v/v) formic acid followed by the loading of the acidified peptide fraction (adjusted pH of ~ 2 by the addition 5% (v/v) formic acid). The peptides were washed twice using 0.1% (v/v) formic acid. Peptides were eluted using 80% (v/v) acetonitrile, 0.1% formic acid and dried in a SPD131DDA SpeedVac^{TM} concentrator (ThermoFisher Scientific, San Jose, CA).

2.2.4. LC-MS/MS analysis of intestinal MLI proteins

High-performance liquid chromatography electrospray ionization tandem mass spectrometry (HPLC-ESI-MS/MS) was completed using an Ekspert nanoLC 400 (Eksigent, Dublin, CA, USA) in-line with an Orbitrap ELITE MS (ThermoFisher Scientific, San Jose, CA) that was operated in the positive ion mode. Peptides, resuspended in 0.5% (v/v) formic acid, were injected onto an in-house generated analytical column (75µM internal diameter and 15cm length) packed with a 1.9 µm C18 resin (Dr Maisch, GmbH, Ammerbuch, Germany). Peptides were eluted with a 2 hour gradient of increasing acetonitrile concentration (5-30 % (v/v), 0.1 %(v/v) formic acid) at a flow rate of 300 nL/min. The heating capillary was set to 300°C and the spray voltage was fixed at 2.2kV. The MS scan ranged from 350 to 1750 m/z with subsequent selection of the 20 most intense ions for data-dependent MS/MS scan using an exclusion duration value of 90 seconds with one repeat count and a 30 second repeat duration. Data was acquired using Xcalibur software (ThermoFisher Scientific, San Jose, CA).

2.2.5. Bioinformatic analysis of intestinal MLI aspirate proteins

Peptide identification and quantification were determined from the MS-acquired data with MaxQuant version 1.5.3.30 (Cox et al., 2009) using the *Homo sapiens* database (Uniprot 2012/07/11). Search parameters included: a multiplicity of two (heavy labels: Arg10 and Lys8; if a light peptide (intestinal MLI sample) lacks it's heavy counterpart (internal spike-in) no quantification value can be provided); specific trypsin digest with a maximum of two missed cleavages; fixed modification of cysteine carbamidomethylation; variable modifications of methionine oxidation and acetylation (protein N-termini); seven amino acids selected as the minimum peptide length; 0.5 Da for the ion mass tolerance; the match between runs and requantify parameters were selected; peptide and protein false discovery rate (FDR) were fixed at 1%. The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE (Vizcaino et al., 2016) partner repository with the dataset identifier PXD006383.

While all samples were included in the database search, only samples from patients with a confident diagnosis were retained in further downstream bioinformatic analysis. In addition, following the database search, the AC and DC proteomes were evaluated separately.

Results from the MaxQuant database search were further processed in Perseus v1.5.2.6 by importation of the normalized H/L ratio. Perseus was used for inversion to light/heavy ratio, $\log_2$ transformation of light/heavy ratio, Pearson correlation coefficient calculation and hierarchical clustering of Pearson correlation values. Microsoft Excel was employed for data filtering. The internal heavy reference spike-in allows for relative quantification between samples and the fold change between non-IBD controls and IBD samples is the ratio of ratios.

For discriminate feature identification partial least squares discriminant analysis (PLS-DA) was performed in MetaboAnalyst version 4.0 (Xia et al., 2015) on the Q75 k-nearest neighbor (knn) imputed dataset selecting for proteins with a VIP score > 1.0 for components 1 and 2. In order to compare the current MLI proteomic dataset with a previous IBD intestinal biopsy dataset (Starr et al., 2017a), the same pipeline (as described above) to identify discriminant features was performed on the discovery cohort (control n=20; IBD CoA n=30) of the biopsy proteomic dataset. To perform gene ontology analysis, Uniprot accession numbers were converted to gene names using the retrieve/ID mapping tool in Uniprot and uploaded into STRING.

Receiver operating characteristics (ROC) curves were generated in MetaboAnalyst 3.0 (Xia et al., 2015) (Biomarker Analysis Module) using knn imputed data; the output values were plotted in Prism version 7 (GraphPad) to generate predictive class probability graphs and ROC curves. Prism was utilized for calculating specificity and sensitivity confidence intervals and ROC area under the curve (AUC). R studio (prcomp) was used for generating principal component analysis (PCA) plots.

For the IBD biomarker panel development, shared proteins identified by PLSDA as discriminant features in both the AC and the DC were selected for further analysis. Proteins ultimately featured in the panel were identified by iterative analysis commencing with those individual proteins with the highest combined (AC + DC) AUC from ROC curve analysis performed in MetaboAnalyst. For UC extent of disease biomarker panel development, proteins were selected for the panel based on iterative analysis of those with the highest AUC values upon comparison of samples from the AC with or without the presence of macroscopic inflammation.

Prism was used to produce the relative expression graphs in addition to performing the associated statistical calculations. The "*ROC curve based model evaluation*" tool combined with the PLS-DA algorithm in MetaboAnalyst was used to generate the predicted class probabilities and ROC curves of the biomarker panels (1-IBD biomarker panel 2- UC extent of disease biomarker panel) and of Calprotectin (S100-A8 and S100-A9).

2.3. Biomarker candidate validation on pediatric stool samples

To more readily translate the biomarker candidates to the clinical setting the expression of select proteins from the IBD biomarker panel was evaluated on stool samples collected from a cohort of new-onset patients as described in detail below.

2.3.1 Stool sample collection

A total of 41 stool samples were collected from pediatric participants prior to the initiation of pharmacological treatment, consisting of participants both independent from the biomarker discovery cohort (n=26) and participants whose MLI aspirate sample was included in the biomarker discovery cohort (n=15). Following collection, stool samples were frozen and delivered to CHEO on ice. Proteins were extracted by Sara Ahmadi using a 5:1 ratio of extraction buffer volume to stool weight. The extraction buffer consisted of 50mM Tris pH 7.2, 150mM NaCl supplemented with a protease inhibitor tablet (cOmplete^{TM}, Mini (Roche Diagnostic, Mannheim, Germany)). The mixture was then agitated by vortexing for 30 seconds, incubated by rotation for 20 minutes at 4°C and centrifuged at 10,000g at 4°C. The supernatant was collected, filtered through a 0.2µm syringe driven filter and stored at -80°C. Protein concentration was determined using the BCA protein assay kit (ThermoFisher Scientific, San Jose, CA).

2.3.2. LTA4H and CAT enzyme-linked immunosorbent assay measurements from stool samples

ELISAs were performed by Sara Ahmadi and Dr. Amanda Starr. The level of LTA4H was measured by ELISA from 50ug of protein extracted from stool (abx572445 Abbexa Ltd, Cambridge, UK) whereas the level of CAT was measured from 0.5ug of protein extracted from stool (ab171572 Abcam, Cambridge, UK). ELISAs were performed according to the manufacturer's protocol. For inter-plate normalization a reference standard was included in all plates.

2.4. Intestinal extracellular vesicle isolation, characterization and proteomic evaluation

2.4.1. Isolation of extracellular vesicles from MLI aspirates

EVs were isolated as previously described (Thery et al., 2006). Briefly, intestinal MLI aspirates depleted of debris and microbial cells as described (section 2.2.1) were thawed at 4°C, underwent ultracentrifugation at 100,000g for 70min at 4°C using a L-100XP ultracentrifuge (Beckman Coulter, California, USA) equipped with a type 70 Ti rotor (Beckman Coulter, California, USA). As described in section 2.2.1, due to variations in patients, including the extent of disease and level of precleaning, the amount of intestinal MLI sample collected varied per patient and thus the volume for EV isolation varied (ranged from 15mL – 65mL following debris and bacterial removal). Pellets were washed in ~22mL of PBS, followed by a second ultracentrifugation at 100,000g for 70min at 4°C. The EV pellet was resuspended in PBS (volume estimated based on pellet size), aliquoted and stored at -80°C.

2.4.2. Biopsy collection

Biopsies were collected from the DC following mucosal washing and taken from areas with macroscopic inflammation for IBD patients and from areas without macroscopic inflammation for non-IBD controls. Biopsies were frozen on dry-ice after collection and transferred to -80°C following the endoscopy procedure and stored until processing.

2.4.3. Electron microscopy of intestinal EVs

Negative staining with visualization by transmission electron microscopy (TEM) was performed to assess EV morphology. Following ultracentrifugation, EVs resuspended in PBS were delivered to the University of Ottawa Heart Institute, Loeb Research Centre, Electron microscopy laboratory, Cell Imaging and Histology Core Facility on a for fee for service basis. EVs were fixed in 2.5% glutaraldehyde, 0.1M sodium cacodylate, washed, resuspended in water and a drop added to a formvar grid and allowed to dry. A droplet of stain (2% uranyl acetate) was then added and allowed to incubate for 10 seconds prior to removing excesses staining solution. The grid was then dried at room temperature prior to imaging using a JEM-1230 transmission electron microscope (JEOL, Japan).

2.4.4. Nanoparticle Tracking Analysis (NTA) of Extracellular Vesicles

The overall size distribution of EVs isolated from the descending colon was assessed by nanoparticle tracking analysis (NTA) with the ZetaView ® (Particle Metrix GmbH, Germany). To ensure consistent performance, standard silica beads were assessed at the beginning of each session. In order to achieve a concentration within the instrument's accuracy range, the EV suspensions were diluted between 1:25,000 and 1:250,000 in PBS. Data was analyzed in Excel, binning EVs sizes by 30nm, and plotted in GraphPad Prism 7.0.

2.4.6. Extracellular vesicle and intestinal biopsy protein extraction, trypsin digest and desalting

Biopsies were lysed in 80µL of 4% (w/v) SDS, 50mM Tris-HCl (pH 7.8) supplemented with a cOmplete protease inhibitor cocktail tablet (Roche Diagnostic GmbH), and protein concentration was then determined using the detergent compatible protein assay (BIORAD, California, USA). A total of 25µg of protein was digested by filter aided sample preparation (Wisniewski et al., 2009) as described in section 2.2.3. The peptides were then desalted using a 5µm AQUA-C18 resin (Dr Maisch, GmbH, Ammerbuch, Germany) as described in section 2.2.3.

2.4.7. Mass spectrometry analysis of EV proteins

The EVs and biopsies were run in two separate batches. Samples were analyzed by liquid chromatography coupled with electrospray tandem mass spectrometry using a Eksigent NanoLC Ultra (Sciex, Framingham, MA) in-line with a Q Exactive™ (ThermoFisher Scientific, San Jose, CA) operated in the positive ion mode. Peptides resuspended in 0.5% (v/v) formic acid were injected onto an in-house generated analytical column (75µM internal diameter and 50cm length) packed with a 1.9µm C18 resin (Dr Maisch, GmbH, Ammerbuch, Germany). Peptides were eluted using a 200nL/min flow rate and a 4 hour gradient of increasing acetonitrile concentration (5-35%(v/v), 0.1%(v/v) formic acid). The heating capillary was set to 300°C. The MS scan ranged from 300 to 1800 m/z with subsequent selection of the 12 most intense ions for data-dependent MS/MS scan. A dynamic exclusion was implemented using a repeat count of 2 and a repeat duration of 30 seconds. Data was acquired using the Xcalibur software (ThermoFisher Scientific, San Jose, CA).

2.4.8. Database search of EV and corresponding biopsy proteins

EVs and their corresponding biopsy RAW files were combined in a single search to identify both human and microbial derived proteins. Given that searches against large databases lead to a high number of false-negative peptide matches, a two-step database search was applied in order to increase the number of confident peptide sequence matches (Jagtap et al., 2013). In the first database search, a reduced microbial database was generated using the "Sample-specific database construction option" in MetaLab1.0. The RAW files were searched against an in-house human gut microbial curated database (Zhang et al., 2018a) encompassing the human gut microbiome integrated gene catalog (IGC; 9,879,896 proteins), the NCBI virus reference sequence database consisting of 275,423 proteins, the proteomes of seven fungal species (*Candida albicans*, *Candida parapsilosis*, *Cladosporium cladosporioides*, *Clavispora lusitaniae*, *Cyberlindnera jadinii*, *Kluyveromyces marxianus*, *Saccharomyces cerevisiae*) previously identified in the gut of pediatric IBD patients (Chehoud et al., 2015; Lewis et al., 2015)) downloaded from UniprotKB and consisting of 121,347 proteins sequences; a human gut microbial gene catalog database generated in-house from shotgun metagenomics sequencing consisting of 1,088,113 predicted genes (Zhang et al., 2016). Redundant sequences were removed from the integrated databases. The reduced microbial database generated with MetaLab was combined with the complete human database downloaded from Uniprot (2017/03/05) and RAW files were searched against this custom database using MaxQuant. Acetyl (Protein N-term) and Oxidation (M) were set as variable modifications; carbamidomethyl of cysteines was set as a fixed modification; the digest was set as trypsin- specific; label-free quantification (LFQ) was enabled using a min ratio count of 1; the re-quantify and match between runs (match time

window 5min, alignment time window 20min) options were enabled; the minimum ratio count for protein quantification was set to 1 and the peptide and protein FDRs were fixed at 1%.

2.4.9. EV Bioinformatic analysis

The ExoCarta top 100 protein list was downloaded January 3rd 2018. For PCA analysis, proteins present in at least 50% of the total cohort (Q50) were analyzed; missing data was imputed by KNN in MetaboAnalyst version 3.6 (Xia et al., 2015), and the resulting dataset was evaluated by PCA with Rstudio version 3.4.3 and PCA plots generated with Prism (GraphPad). Discriminant features were identified using PLS-DA on proteins quantified in at least 50% of each group (control and IBD) with KNN imputation using MetaboAnalyst 3.5 and selecting for proteins with a VIP score >1.0 for the first component. For gene ontology analysis, protein IDs were converted to Gene names using the retrieve/ID mapping tool in Uniprot (The UniProt, 2017) and analyzed in STRING version 10.5 (Szklarczyk et al., 2017).

CHAPTER 3. Proteomic alterations at the intestinal MLI in new-onset pediatric IBD

3.1 Rationale

In addition to being the site of disease in IBD, the intestinal MLI is the location wherein host components directly interact with the commensal microbiota, an important feature for disease pathogenesis. Despite the importance of this anatomical location in IBD, there have been limited proteomic investigations completed on the intestinal MLI; those studies were restricted to adult cohorts, and confounded by treatment intervention (Li et al., 2016; Presley et al., 2012). My hypothesis was that the intestinal MLI of pediatric IBD patients exhibit global proteomics alterations compared to non-IBD controls and my objective was to investigate the proteomic alterations at the intestinal MLI in new-onset pediatric IBD as compared to non-IBD controls. To gain insight on proteomic changes that occur at disease onset, a shotgun proteomic study was performed; intestinal MLI samples were obtained from a cohort of 60 treatment-naive participants.

3.2 Results

3.2.1. Pediatric treatment-naive patient cohort

The individual intestinal MLI proteome profiles from 60 patients (42 IBD patients and 18 non-IBD controls) were analyzed by high resolution MS. Samples from both the AC and DC were analyzed for 33 patients, whereas only the AC or the DC were available for 24 patients (8 non-IBD control, 7 CD and 9 UC) and 3 patients (1 CD and 2 UC) respectively, yielding a total of 93 intestinal MLI samples for analysis. Patient characteristics are summarized in **Table 1**. There were no significant differences in age (**Figure 3.1 A, B**; AC $P = 0.5888$, DC $P = 0.9792$;

	Controls (n=18)		CD (n=22)		UC (n=20)	
Region	**AC (n=18)**	**DC (n=10)**	**AC (n=21)**	**DC (n=15)**	**AC (n=18)**	**DC (n=11)**
# Females (%)	13(72.2%)	7(70.0%)	5(23.8%)	3(20.0%)	18 (55.6%)	4(36.4%)
Median Age, years	15.4	15.7	13.8	14.1	15.4	15.8
(IQR)	(11.3-16.6)	(12.7-16.2)	(11-16.3)	(12.3-16.1)	(11.1-16.2)	(14.9–16.8)
Paris Classification*						
A1a	NA	NA	4(19.0%)	2(13.3%)	NA	NA
A1b	NA	NA	16(76.2%)	12(80.0%)	NA	NA
A2	NA	NA	1(4.8%)	1(6.7%)	NA	NA
Location						
L1	NA	NA	4(19.0%)	3(20.0%)	NA	NA
L2	NA	NA	5(23.8%)	4(26.7%)	NA	NA
L3	NA	NA	11(52.4%)	7(46.7%)	NA	NA
L4a	NA	NA	10(46.6%)	6(40.0%)	NA	NA
L4b	NA	NA	3(14.3%)	3(20.0%)	NA	NA
Behavior						
B1	NA	NA	21(100%)	14(93.3%)	NA	NA
B2	NA	NA	0	0	NA	NA
B3	NA	NA	0	1(6.7%)	NA	NA
p	NA	NA	3(14.3%)	5(33.3%)	NA	NA
Growth						
G0	NA	NA	14(66.7%)	11(73.3%)	NA	NA
G1	NA	NA	7(33.3%)	4(26.7%)	NA	NA
Extent						
E1	NA	NA	NA	NA	1(5.6%)	1(9.1%)
E2	NA	NA	NA	NA	4(22.2%)	4(36.4%)
E3	NA	NA	NA	NA	2(11.1%)	2(18.2%)
E4	NA	NA	NA	NA	11(61.1%)	4(36.4%)
Severity						
S0	NA	NA	NA	NA	15(83.3%)	9(81.8%)
S1	NA	NA	NA	NA	3(16.7%)	2(18.2%)

*(Levine et al., 2011)

Table 3. 1: Patient characteristics.

The individual intestinal MLI proteome of 60 patients was analyzed by high resolution MS. Two colon sub-regions were analyzed, the ascending colon and the descending colon. All participants were ≤ 18 years of age. In CD, the Paris classification encompasses the age, anatomical location of disease, disease behavior (stricturing/penetrating), and whether delayed growth is observed. In UC the Paris classification comprises the extent of disease and the severity. IQR=interquartile range.

This table was published in Deeke, S.A., Starr, A.E., Ning, Z., Ahmadi, S., Zhang, X., Mayne, J., Chiang, C.K., Singleton, R., Benchimol, E.I., Mack, D.R., Stintzi, A., and Figeys, D. (2018). Mucosal-luminal interface proteomics reveals biomarkers of pediatric inflammatory bowel disease-associated colitis. Am J Gastroenterol *113*, 713-724.

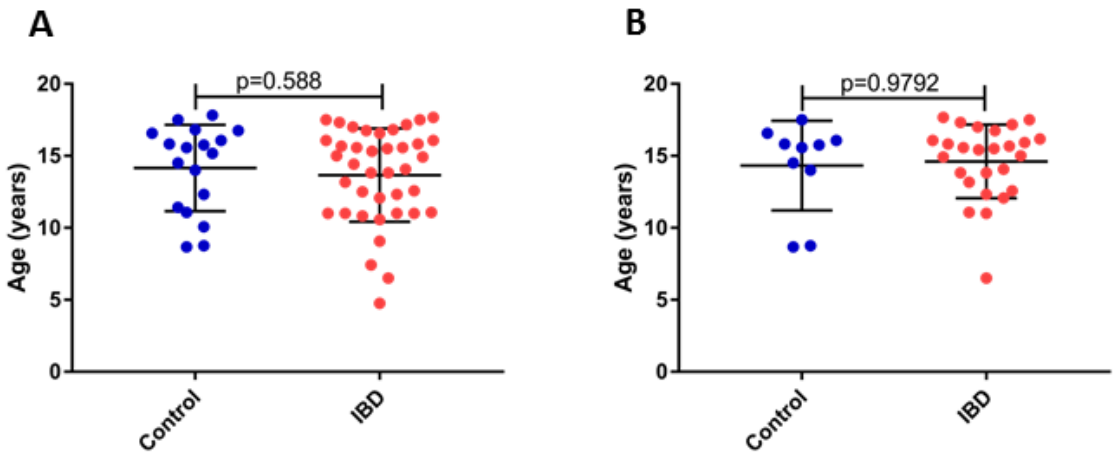

A
p=0.588
Age (years)
20
15
10
5
0
Control
IBD
B
p=0.9792
Age (years)
20
15
10
5
0
Control
IBD

Figure 3. 1: Patient cohort age

No difference in participant age was observed between the non-IBD control and IBD groups in either the AC **(A)** or DC **(B)**. Statistics performed using the Mann Whitney test.

Figure adapted from Deeke, S.A., Starr, A.E., Ning, Z., Ahmadi, S., Zhang, X., Mayne, J., Chiang, C.K., Singleton, R., Benchimol, E.I., Mack, D.R., Stintzi, A., and Figeys, D. (2018). Mucosal-luminal interface proteomics reveals biomarkers of pediatric inflammatory bowel disease-associated colitis. Am J Gastroenterol *113*, 713-724.

Mann Whitney test) or sex within the non-IBD control group (DC $P = 0.1719$; Binomial one-tailed test) and IBD group (AC $P = 0.0998$, DC $P = 0.1002$; Binomial one-tailed test), however a significantly elevated number of females were present in the AC non-IBD control group ($P = 0.048$; Binomial one-tailed). The majority (70%) of UC patients within the cohort displayed extensive disease (E3 or E4 Paris classification) at the time of diagnostic colonoscopy, which is consistent with other pediatric UC cohorts (Sauer and Kugathasan, 2009). For CD patients, the majority (77%) presented with colonic involvement (Paris classification of L2 or L3), in accordance with previous pediatric CD studies (Vernier-Massouille et al., 2008).

3.2.2 MLI proteomic data evaluation

Due to ongoing patient recruitment, intestinal MLI samples were processed and analyzed by MS spanning a period of 22 months. For accurate quantification over time, an internal heavy reference was spiked into the isolated MLI proteins, an approach known as super-SILAC (Geiger et al., 2010). The reference proteome was aliquoted and stored at -80°C prior to any analysis for stability over time. The median ratio (normalized MLI proteins/internal heavy reference proteome) of the majority of MLI proteins in both colon sub-regions (AC: 87.7%, DC: 88.5%) was within the 10-fold range (**Figure 3.2 A, B**). Overall a mean of 1447 ± 267 and 1421 ± 176 proteins were quantified per intestinal MLI sample in the AC and DC, respectively, with no significant differences between groups (**Figure 3.2 C, D**; AC $P = 0.2508$; DC $P = 0.4330$). Of those proteins which were quantified at the AC, an average of 81.99% and 82.06% had two or more unique peptides for non-IBD controls and IBD samples respectively. Similar results were obtained for the DC, averages of 82.74% and 82.87% of quantified proteins had ≥ 2 unique peptides for non-IBD control and IBD samples respectively. An average of 49 ± 13 proteins, representing 2% of the total number of proteins identified per intestinal MLI sample, lacked a

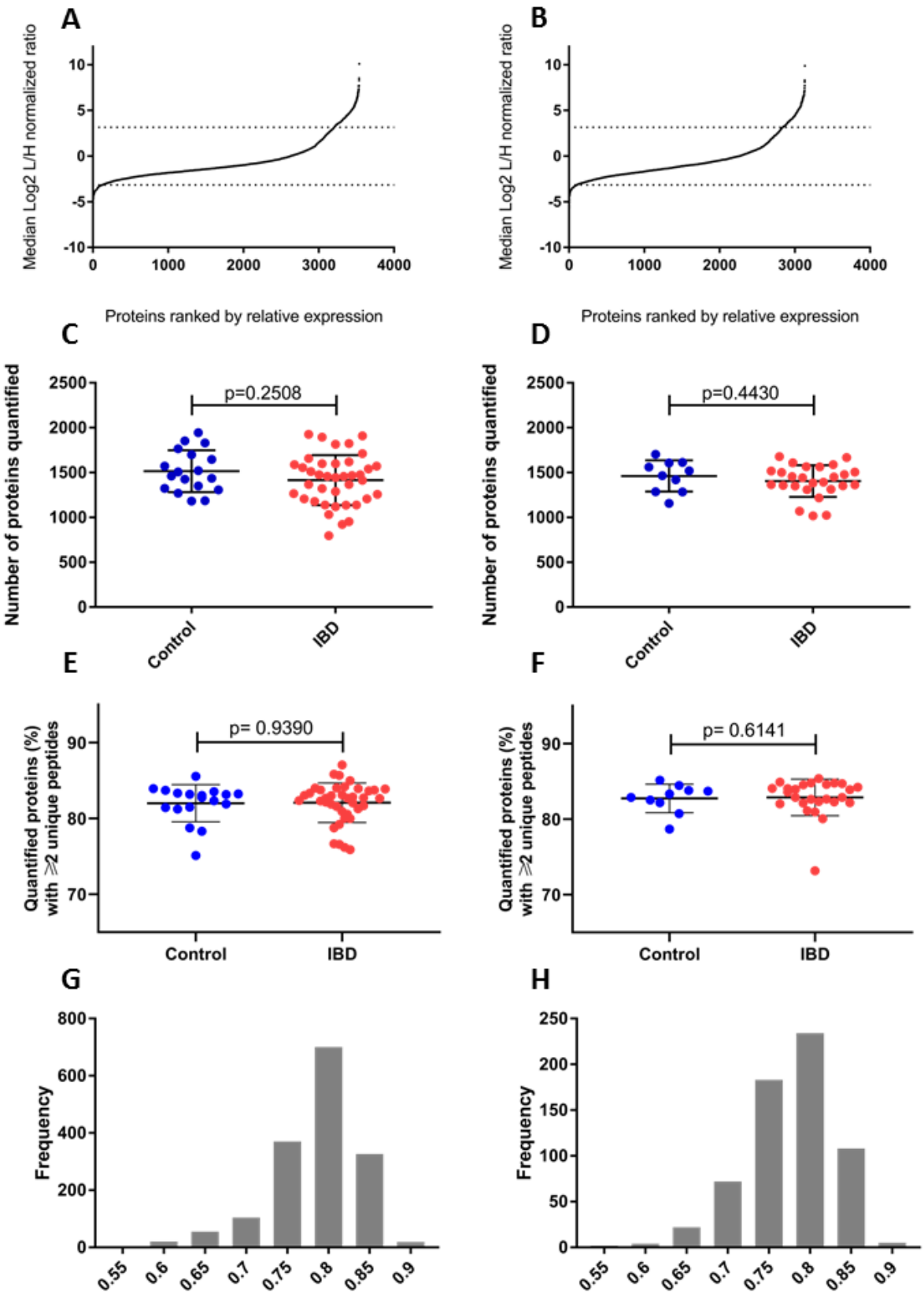

A
Median Log2 L/H normalized ratio
10
5
0
-5
-10
0
1000
2000
3000
4000
Proteins ranked by relative expression
B
Median Log2 L/H normalized ratio
10
5
0
-5
-10
0
1000
2000
3000
4000
Proteins ranked by relative expression
C
Number of proteins quantified
2500
2000
1500
1000
500
0
p=0.2508
Control
IBD
D
Number of proteins quantified
2500
2000
1500
1000
500
0
p=0.4430
Control
IBD
E
Quantified proteins (%) with ⩾2 unique peptides
90
80
70
p= 0.9390
Control
IBD
F
Quantified proteins (%) with ⩾2 unique peptides
90
80
70
p= 0.6141
Control
IBD
G
Frequency
800
600
400
200
0
0.55
0.6
0.65
0.7
0.75
0.8
0.85
0.9
H
Frequency
250
200
150
100
50
0
0.55
0.6
0.65
0.7
0.75
0.8
0.85
0.9

Figure 3. 2. Mass Spectrometry data evaluation.

The proteomic profiles of 93 intestinal MLI samples were assessed over 22 months. The median ratio (normalized MLI proteins/internal heavy reference proteome) of the majority of MLI proteins in both colon sub-regions (AC: 87.7% **(A)**, DC: 88.5% **(B)**) was within the 10-fold range, as indicated by the dotted horizontal lines. The number of proteins identified per intestinal MLI sample in the AC **(C)** and DC **(D)**. Of those proteins which were quantified, the percentage which had two or more unique peptides in the AC **(E)** and DC **(F)**. The majority of Pearson correlation values derived from the Q75 Log2 normalized light/heavy ratios were > 0.75 **(G)** AC (89%) and **(H)** DC (84%). Statistics performed using the Mann Whitney test.

Figure adapted from Deeke, S.A., Starr, A.E., Ning, Z., Ahmadi, S., Zhang, X., Mayne, J., Chiang, C.K., Singleton, R., Benchimol, E.I., Mack, D.R., Stintzi, A., and Figeys, D. (2018). Mucosal-luminal interface proteomics reveals biomarkers of pediatric inflammatory bowel disease-associated colitis. Am J Gastroenterol *113*, 713-724.

heavy reference counterpart and thus no quantification values could be provided for those proteins. Taken together, a total of 3,537 (AC) and 3,132 (DC) intestinal MLI proteins were quantified, which was reduced to 972 (AC) and 995 (DC) proteins upon the exclusion of proteins that were not quantified in ≥75% of the intestinal MLI samples (Q75). To evaluate consistency over time, Pearson correlation analysis was performed (Q75 of the Log2 normalized light/heavy ratios) resulting in the majority of values with a correlation value ≥0.75 (AC=89%; DC=84%) (**Figure 3.2 G, H**).

3.2.3. Proteomic landscape alterations of host proteins at the MLI

In order to assess factors that are influencing the variance between patient proteomes, principal component analysis (PCA) was applied. Labeling the samples according to either disease (non-IBD control or IBD) and inflammatory status (CoN or CoA) (**Figure 3.3 A, B**), batch (**Figure 3.3 C, D**) or sex (**Figure 3.3 E, F**) appears to separate based on the former, indicating that disease and inflammatory status impact the variance to a greater degree than that of sex or batch. PLS-DA was performed to identify discriminant features of active IBD, yielding 201 and 237 proteins with a VIP score > 1.0 in the AC and DC respectively (**Appendix 1 Tables 1 and 2**). Proteins previously shown to be altered in IBD were among the list of discriminant features (**Table 3.2**), including the neutrophilic proteins S100-A8 (Calprotectin), Cathepsin G and Myeloperoxidase (Bennike et al., 2015; Han et al., 2013). Consistent with the immune-mediated nature of IBD, several of the discriminant features were enriched for biological processes related to immunity including immune response, defense response and immune system process (**Figure 3.4**). During inflammation, the release of various cytokines can mediate cytoskeletal rearrangements leading to defects in barrier integrity (Lopez-Posadas et al., 2017). Accordingly, we observed several cytoskeletal organization processes enriched from IBD discriminant features

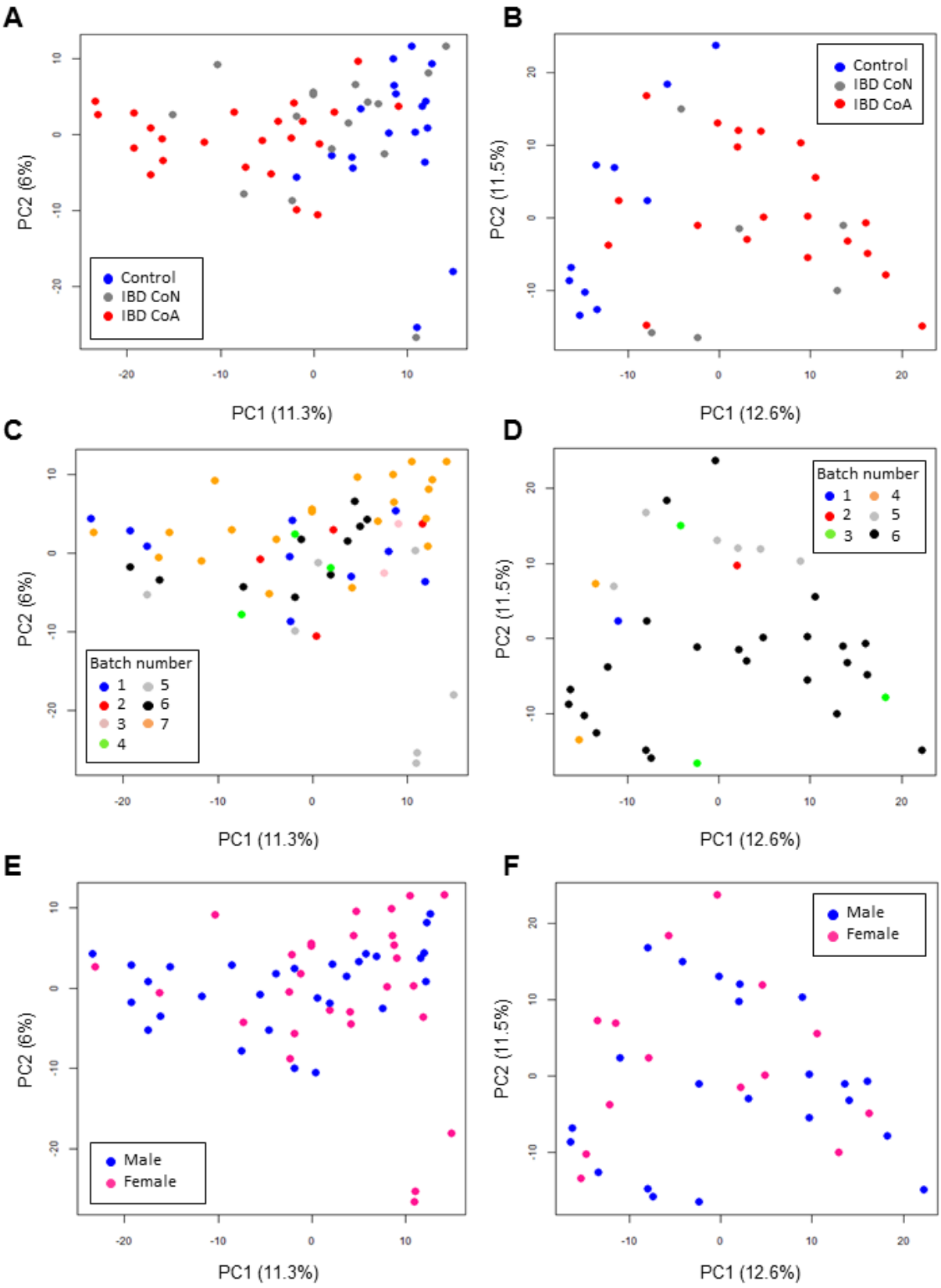

A
PC2 (6%)
PC1 (11.3%)
Control
IBD CoN
IBD CoA
B
PC2 (11.5%)
PC1 (12.6%)
Control
IBD CoN
IBD CoA
C
PC2 (6%)
PC1 (11.3%)
Batch number
1
2
3
4
5
6
7
D
PC2 (11.5%)
PC1 (12.6%)
Batch number
1
2
3
4
5
6
E
PC2 (6%)
PC1 (11.3%)
Male
Female
F
PC2 (11.5%)
PC1 (12.6%)
Male
Female

Figure 3. 3. PCA of intestinal MLI samples.

Samples were labeled according to disease and inflammatory status **(A)** and **(B)**, MS batch **(C)** and **(D)** and according to sex **(E)** and **(F)** in the ascending colon **(A), (C)**, and **(E)** and the descending colon **(B)**, **(D)** and **(F)**. CoN= aspirate from unaffected area, without the presence of macroscopic inflammation, CoA= aspirate from affected area, with the presence of macroscopic inflammation.

Figure adapted from Deeke, S.A., Starr, A.E., Ning, Z., Ahmadi, S., Zhang, X., Mayne, J., Chiang, C.K., Singleton, R., Benchimol, E.I., Mack, D.R., Stintzi, A., and Figeys, D. (2018). Mucosal-luminal interface proteomics reveals biomarkers of pediatric inflammatory bowel disease-associated colitis. Am J Gastroenterol *113*, 713-724.

Protein name	AC relative average expression IBD CoA/ control	DC relative average expression IBD CoA/ control	AC PLS-DA VIP value component 1	DC PLS-DA VIP value component 1	Reference
Protein S100-A8	8.7	3.3	3.16	1.86	(Bennike et al., 2015; Han et al., 2013)
Cathepsin G	11.7	6.4	3.54	2.57	(Bennike et al., 2015; Han et al., 2013)
Myeloperoxidase	5.7	4.5	4.39	1.53	(Bennike et al., 2015; Han et al., 2013)

Table 3. 2: Proteins identified as discriminant features in IBD compared to non-IBD control are consistent with previous reports.

Discriminant features were identified by PLS-DA. AC = ascending colon; DC= descending colon

Data further analyzed from Deeke, S.A., Starr, A.E., Ning, Z., Ahmadi, S., Zhang, X., Mayne, J., Chiang, C.K., Singleton, R., Benchimol, E.I., Mack, D.R., Stintzi, A., and Figeys, D. (2018). Mucosal-luminal interface proteomics reveals biomarkers of pediatric inflammatory bowel disease-associated colitis. Am J Gastroenterol *113*, 713-724.

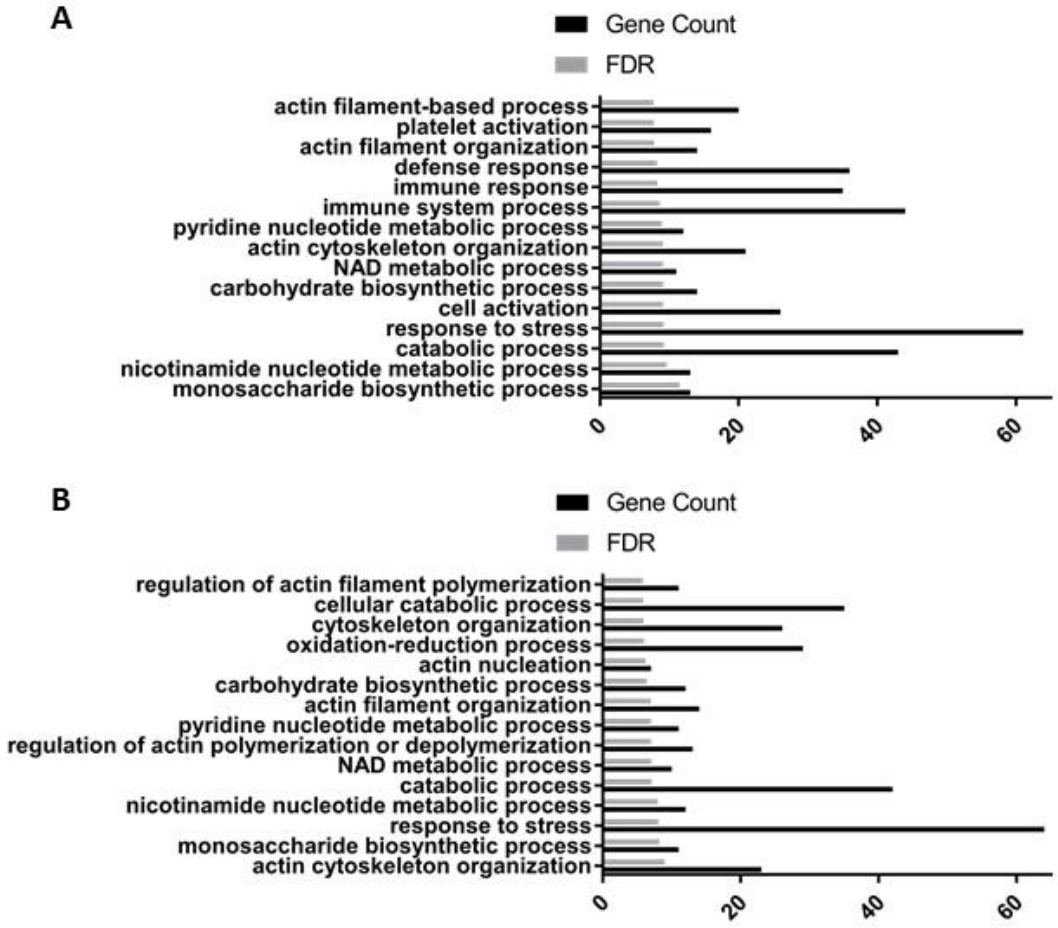

A
Gene Count
FDR
actin filament-based process
platelet activation
actin filament organization
defense response
immune response
immune system process
pyridine nucleotide metabolic process
actin cytoskeleton organization
NAD metabolic process
carbohydrate biosynthetic process
cell activation
response to stress
catabolic process
nicotinamide nucleotide metabolic process
monosaccharide biosynthetic process
0
20
40
60
B
Gene Count
FDR
regulation of actin filament polymerization
cellular catabolic process
cytoskeleton organization
oxidation-reduction process
actin nucleation
carbohydrate biosynthetic process
actin filament organization
pyridine nucleotide metabolic process
regulation of actin polymerization or depolymerization
NAD metabolic process
catabolic process
nicotinamide nucleotide metabolic process
response to stress
monosaccharide biosynthetic process
actin cytoskeleton organization
0
20
40
60

Figure 3. 4. Biological processes of IBD discriminant features.

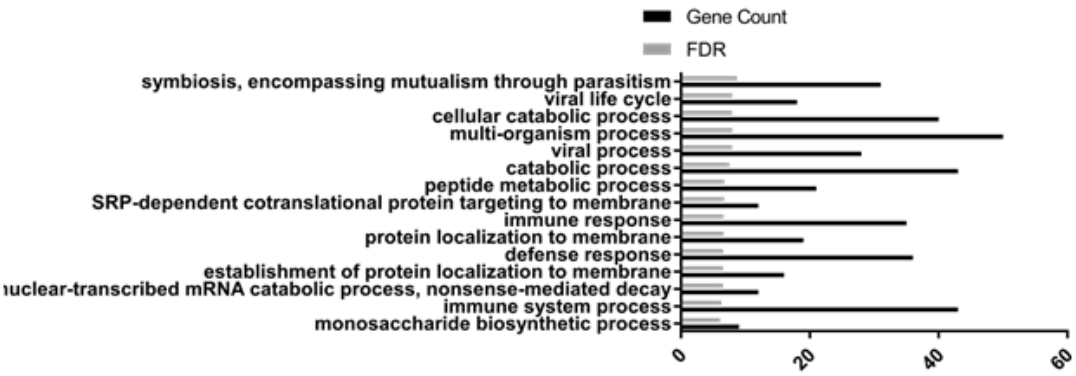
The top 15 biological processes corresponding to discriminant features identified from intestinal MLI samples by the comparison of non-IBD controls and IBD CoA using PLSDA in the AC **(A)** and DC **(B)**. Gene ontology was performed using STRING. FDR values were –Log10 transformed.

Data further analyzed from Deeke, S.A., Starr, A.E., Ning, Z., Ahmadi, S., Zhang, X., Mayne, J., Chiang, C.K., Singleton, R., Benchimol, E.I., Mack, D.R., Stintzi, A., and Figeys, D. (2018). Mucosal-luminal interface proteomics reveals biomarkers of pediatric inflammatory bowel disease-associated colitis. Am J Gastroenterol *113*, 713-724.

including actin filament organization, regulation of actin filament polymerization and regulation of actin polymerization or depolymerization (**Figure 3.4**). Comparison of non-IBD controls with IBD samples collected from areas without macroscopic inflammation (CoN) yielded 215 and 280 protein as discriminant features for the AC and DC, respectively. Similarly to the comparison of CoA, several of the top biological processes related to immunity and host defense including immune response, defense response and immune system process (**Figure 3.5**), potentially suggesting either early stages of inflammation prior to evidence of macroscopic inflammation or microscopic inflammation.

For comparison, the discriminant feature identification pipeline was implemented on a dataset obtained from AC IBD biopsies (control n=20, IBD CoA n=30) (Starr et al., 2017a), yielding 315 discriminant feature with a VIP score > 1.0. In total, 31 proteins were identified as common non-IBD controls vs IBD discriminant features between the two (MLI and biopsy) datasets (**Table 3.3**). For 26 of the 31 proteins (83.9%) the relative expression of IBD CoA compared to non-IBD controls demonstrated similar directionality (elevated or decreased in both datasets) with 13 (41.9%) proteins elevated and 13 proteins decreased in IBD CoA compared to non-IBD controls. The residual 5 (16.1%) proteins demonstrated contrasting relative expression between the two datasets (MLI and biopsy), with two proteins demonstrating relative elevated expression in IBD CoA at the MLI and a relative decrease in IBD CoA versus non-IBD controls in biopsies.

A


Gene Count
FDR
symbiosis, encompassing mutualism through parasitism
viral life cycle
cellular catabolic process
multi-organism process
viral process
catabolic process
peptide metabolic process
SRP-dependent cotranslational protein targeting to membrane
immune response
protein localization to membrane
defense response
establishment of protein localization to membrane
nuclear-transcribed mRNA catabolic process, nonsense-mediated decay
immune system process
monosaccharide biosynthetic process
0
20
40
60

B

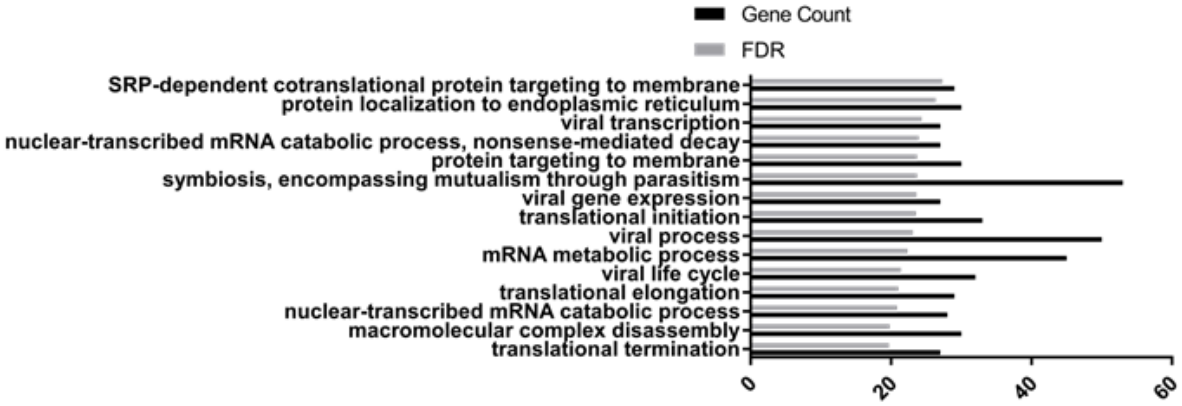

Gene Count
FDR
SRP-dependent cotranslational protein targeting to membrane
protein localization to endoplasmic reticulum
viral transcription
nuclear-transcribed mRNA catabolic process, nonsense-mediated decay
protein targeting to membrane
symbiosis, encompassing mutualism through parasitism
viral gene expression
translational initiation
viral process
mRNA metabolic process
viral life cycle
translational elongation
nuclear-transcribed mRNA catabolic process
macromolecular complex disassembly
translational termination
0
20
40
60

Figure 3. 5. Biological processes of IBD discriminant features of non-IBD controls versus IBD CoN

The top 15 biological processes corresponding to discriminant features identified from intestinal MLI samples by the comparison of non-IBD controls and IBD CoN using PLSDA in the AC **(A)** and DC **(B)**. Gene ontology was performed using STRING. FDR values were –Log10 transformed.

Data further analyzed from Deeke, S.A., Starr, A.E., Ning, Z., Ahmadi, S., Zhang, X., Mayne, J., Chiang, C.K., Singleton, R., Benchimol, E.I., Mack, D.R., Stintzi, A., and Figeys, D. (2018). Mucosal-luminal interface proteomics reveals biomarkers of pediatric inflammatory bowel disease-associated colitis. Am J Gastroenterol *113*, 713-724.

Gene names	MLI AC LOG2 (IBD CoA/Control)	Biopsy AC LOG2 (IBD CoA/Control)
PGD	2.4	0.7
LDHA	1.3	0.5
S100A8	3.1	1.6
CTSG	3.5	0.6
ANXA3	4.2	1.6
PDIA4	0.2	0.5
LCP1	2.3	0.5
RAC2	1.9	0.6
WARS	1.4	1.7
S100P	2.5	2.3
S100A11	1.8	0.8
NAMPT	2.6	1.0
DEFA3	0.7	0.4
CLCA1	-1.3	-0.6
AKR1B10	-2.8	-1.1
CA2	-1.8	-1.4
IGHA2	-1.7	-0.5
VIL1	-1.2	-1.1
ADH5	-1.2	-0.4
ANPEP	-2.0	-0.6
CBR1	-2.1	-0.6
LGALS3	-2.6	-1.1
LRPPRC	-1.6	-0.7
ATP5O	-3.6	-0.4
HADHB	-1.0	-0.5
LGALS3BP	-1.2	-0.5
TF	1.4	-0.3
NAPRT	1.1	-0.7
IGHV3-23	-2.0	0.5
IGHG2	-0.1	1.4
LAP3	-1.4	0.9

Table 3. 3. Common IBD CoA vs non-IBD control discriminant features identified in both the current MLI proteome dataset and a previous intestinal biopsy proteomic dataset.

A total of 31 proteins were commonly identified as discriminant features by PLS-DA between both (MLI and biopsy (Starr et al., 2017b)) datasets. The first 13 rows correspond to proteins which were similarly elevated in IBD CoA compared to non-IBD controls in both the MLI and the biopsy datasets, the following 13 rows correspond to proteins which were decreased in IBD CoA compared to non-IBD controls in both datasets, while the last five rows corresponds to proteins which exhibited different relative expression between IBD CoA and non-IBD controls between the two datasets (MLI and biopsy).

Data further analyzed from Deeke, S.A., Starr, A.E., Ning, Z., Ahmadi, S., Zhang, X., Mayne, J., Chiang, C.K., Singleton, R., Benchimol, E.I., Mack, D.R., Stintzi, A., and Figeys, D. (2018). Mucosal-luminal interface proteomics reveals biomarkers of pediatric inflammatory bowel disease-associated colitis. Am J Gastroenterol *113*, 713-724.

CHAPTER 4. Biomarker discovery for pediatric inflammatory bowel disease

4.1 Rationale

IBD diagnosis is challenging, as symptoms are non-specific, and requires multiple methods including an invasive endoscopy procedure. Accurate and specific biomarkers for IBD diagnosis remain an unmet clinical need. This need is accentuated for the pediatric population, where the performance of the most commonly used IBD biomarker is inferior (Henderson et al., 2014). Furthermore, individuals with UC must undergo assessment for extent of disease in order to aid in determining appropriate therapy in addition to conclude the rate at which colorectal cancer monitoring must ensue (Rosen et al., 2015; Stange et al., 2008). Currently, no biomarkers exist for the evaluation of this particular feature of disease, relying instead on endoscopy to determine extent of disease. Reliable, accurate and non-invasive biomarkers for IBD diagnosis would bypass the costly expense, discomfort and complications that are associated with colonoscopies. My hypotheses are (1) that specific proteins expressed at the intestinal MLI demonstrate differential expression between non-IBD and IBD patients and could therefore serve as biomarkers for pediatric IBD diagnosis. (2) Proteins expressed at the MLI at the AC display differential expression in UC patients in areas with macroscopic inflammation compared to areas without macroscopic inflammation and thus have value in differentiating pancolitis and non-pancolitis. Accordingly, my objectives are (1) identify diagnostic biomarkers for pediatric IBD and (2) identify biomarkers for UC extent of disease (pancolitis versus non-pancolitis).

To address this, the intestinal MLI proteomic profile of samples collected during diagnostic colonoscopy from a cohort of treatment-naive pediatric patients was analyzed and two panels of biomarkers were identified for 1) diagnosis of active pediatric IBD and 2) UC pancolitis versus non-pancolitis (extent of disease).

4.2 Results

4.2.1. Biomarker panel development for pediatric active IBD

To develop a biomarker panel for suspected pediatric IBD, discriminant features (non-IBD control vs IBD CoA) identified from both colon sub-regions (AC and DC) by PLS-DA in Chapter 3.2.3 (**Appendix 1 Tables 1, 2**) were further investigated. The biomarker panel generation pipeline is summarized in **Figure 4.1**. Iterative analysis using proteins with the highest combined area under the curve (AUC) values from both the AC and DC was performed to achieve the maximum sensitivity and specificity for both colon regions. For the DC, a minimum of two proteins was required whereas in the AC a minimum of four proteins was required to achieve maximal specificity and sensitivity (**Figure 4.2 A**). Ultimately, the IBD biomarker panel for differentiation of active disease from non-IBD controls consists of four proteins: Leukotriene A-4 hydrolase, Catalase, Transketolase and Annexin A3 (**Figure 4.2 B** and **Table 4.1**).

Utilizing the panel of four proteins listed in **Table 4.1** yielded an AUC of 0.989 (95% CI: 0.967 to 1.0) for the AC and 0.999 (95% CI: 0.999 to 1.0) for the DC (**Figure 4.3 A**). The predictive class probabilities of the AC achieves a sensitivity of 0.954 (95% CI: 0.7716 to 0.9988) and specificity >0.999 (95% CI: 0.8147 to 1.0) (**Figure 4.3 C**). For the DC, this yields a

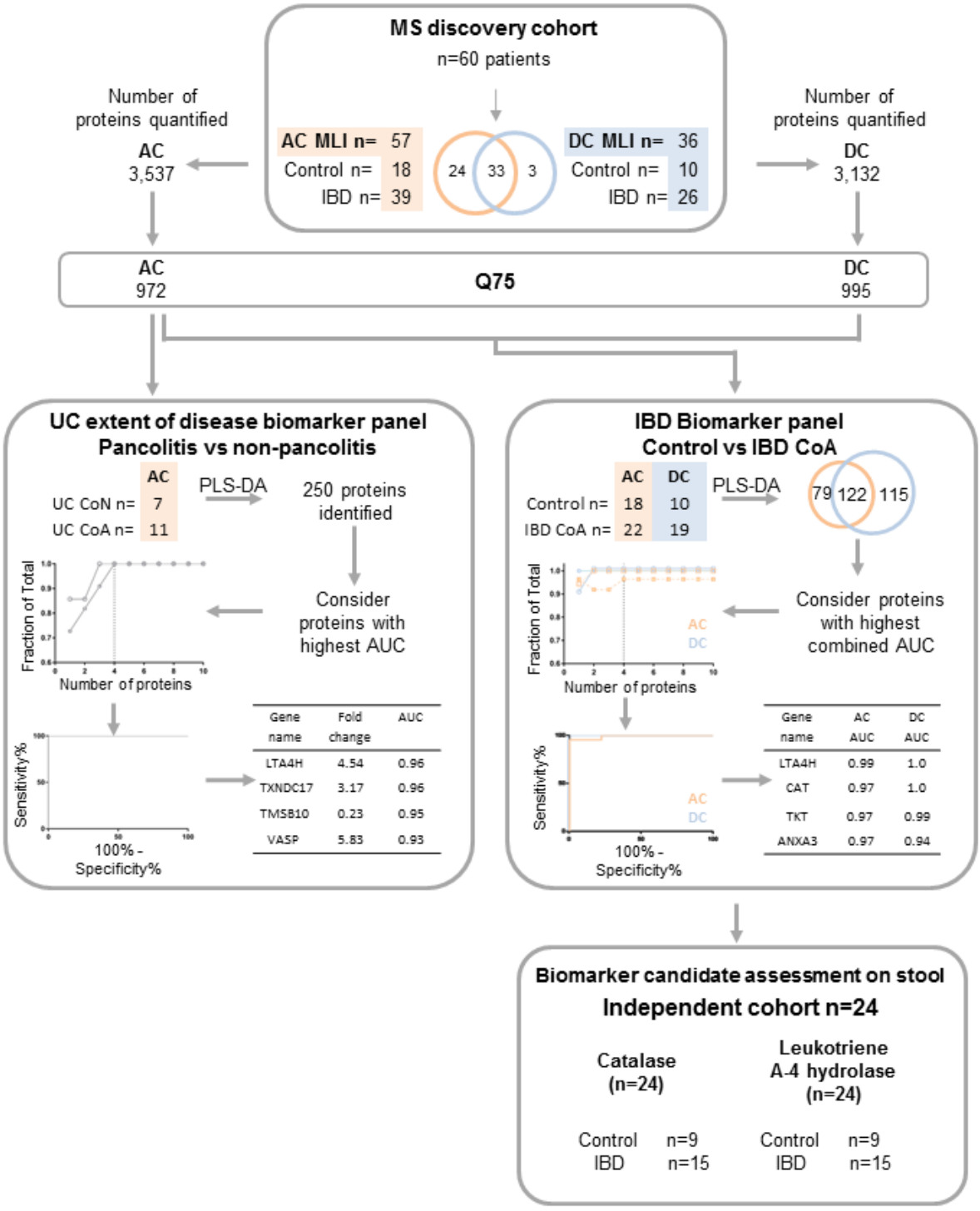
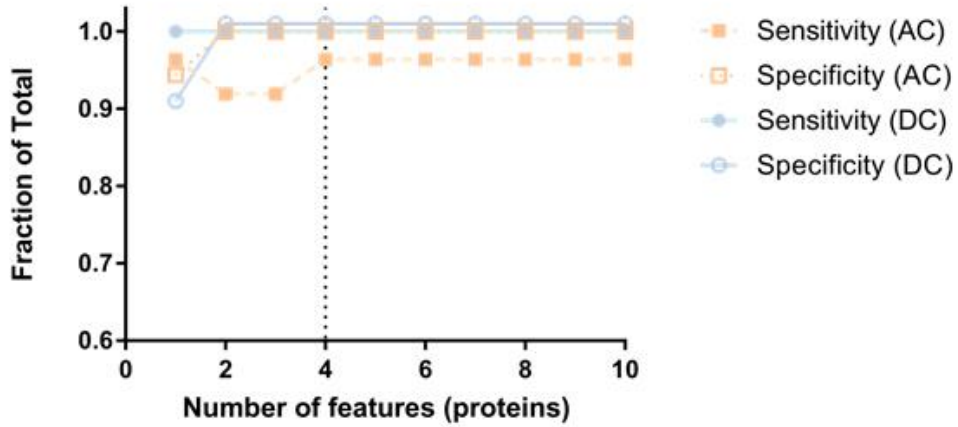

MS discovery cohort
n=60 patients
Number of proteins quantified
AC
3,537
AC MLI n= 57
Control n= 18
IBD n= 39
24
33
3
DC MLI n= 36
Control n= 10
IBD n= 26
Number of proteins quantified
DC
3,132
AC
972
Q75
DC
995
UC extent of disease biomarker panel
Pancolitis vs non-pancolitis
AC
UC CoN n= 7
UC CoA n= 11
PLS-DA
250 proteins identified
Consider proteins with highest AUC
Fraction of Total
Number of proteins
Sensitivity%
100% - Specificity%
Gene name
Fold change
AUC
LTA4H 4.54 0.96
TXNDC17 3.17 0.96
TMSB10 0.23 0.95
VASP 5.83 0.93
IBD Biomarker panel
Control vs IBD CoA
AC
DC
Control n= 18 10
IBD CoA n= 22 19
PLS-DA
79
122
115
Consider proteins with highest combined AUC
Fraction of Total
Number of proteins
AC
DC
Sensitivity%
100% - Specificity%
Gene name
AC AUC
DC AUC
LTA4H 0.99 1.0
CAT 0.97 1.0
TKT 0.97 0.99
ANXA3 0.97 0.94
Biomarker candidate assessment on stool
Independent cohort n=24
Catalase (n=24)
Control n=9
IBD n=15
Leukotriene A-4 hydrolase (n=24)
Control n=9
IBD n=15

Figure 4. 1. Biomarker panel development workflow.

Two separate biomarker panels were generated, the first for IBD diagnosis and the second for UC extent of disease (pancolitis vs non-pancolitis). Proteins identified in ≥ 75% of samples were considered for biomarker panel generation. Discriminant features were identified using PLS-DA and proteins with the highest ROC AUC were considered. Select proteins (Catalase and Leukotriene A-4 hydrolase) from the controls vs. IBD CoA biomarker panel were tested on an independent stool cohort in order to render the findings more readily translatable to the clinic.

Figure adapted from Deeke, S.A., Starr, A.E., Ning, Z., Ahmadi, S., Zhang, X., Mayne, J., Chiang, C.K., Singleton, R., Benchimol, E.I., Mack, D.R., Stintzi, A., and Figeys, D. (2018). Mucosal-luminal interface proteomics reveals biomarkers of pediatric inflammatory bowel disease-associated colitis. Am J Gastroenterol *113*, 713-724.

A


Fraction of Total
1.0
0.9
0.8
0.7
0.6
0
2
4
6
8
10
Number of features (proteins)
Sensitivity (AC)
Specificity (AC)
Sensitivity (DC)
Specificity (DC)

B

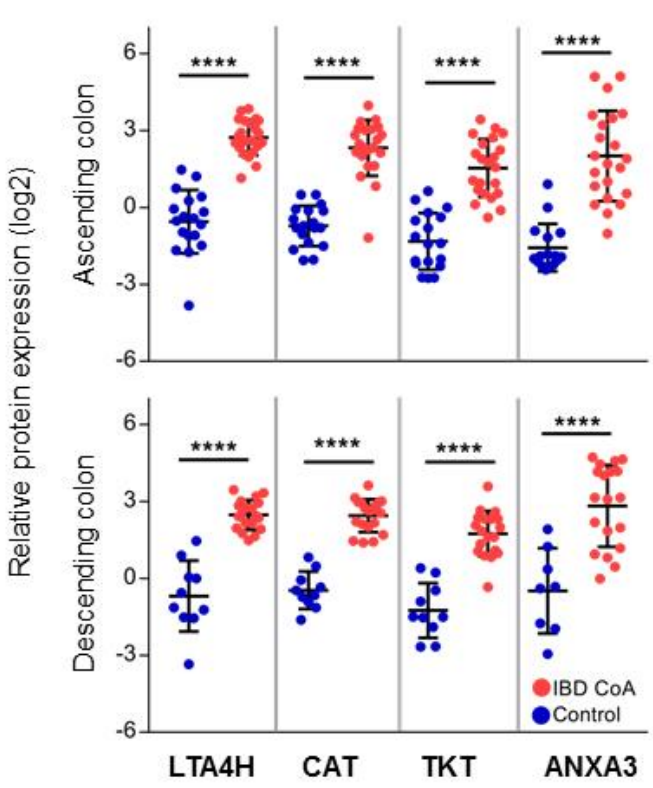

Relative protein expression (log2)
Ascending colon
Descending colon
6
3
0
-3
-6
****
IBD CoA
Control
LTA4H
CAT
TKT
ANXA3

Figure 4. 2. Biomarker panel for pediatric IBD diagnosis.

(A) Iterative analysis of proteins with the highest combined area under the curve values from both the AC and DC were performed to achieve the maximum sensitivity and specificity for both colon regions, denoted by the dotted line.

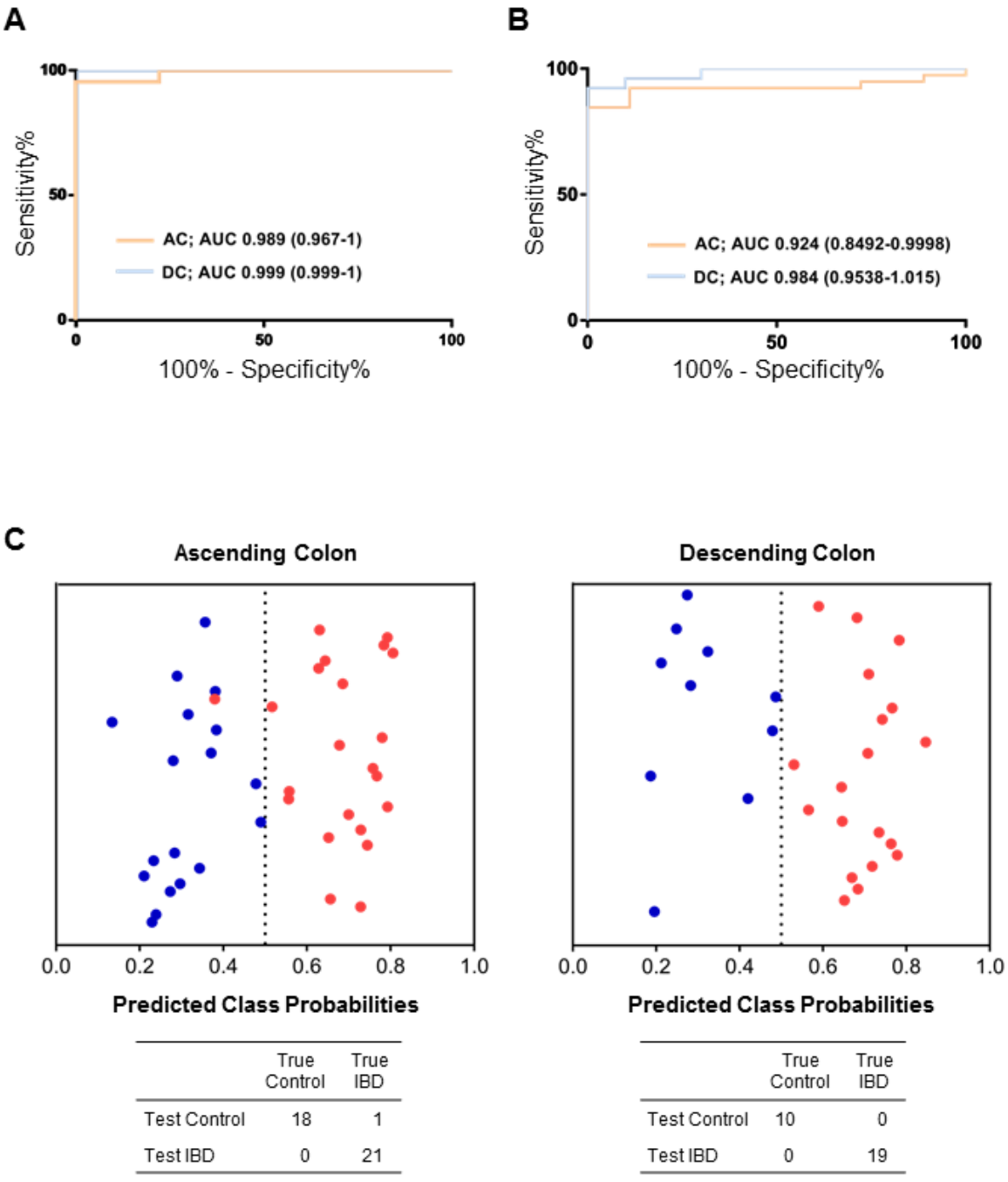
(B) Expression levels of proteins included in IBD Biomarker panel (Log2 normalized light/heavy ratio).

Figure adapted from Deeke, S.A., Starr, A.E., Ning, Z., Ahmadi, S., Zhang, X., Mayne, J., Chiang, C.K., Singleton, R., Benchimol, E.I., Mack, D.R., Stintzi, A., and Figeys, D. (2018). Mucosal-luminal interface proteomics reveals biomarkers of pediatric inflammatory bowel disease-associated colitis. Am J Gastroenterol *113*, 713-724.

		IBD CoA / Control Fold change		IBD CoA vs Control AUC (95% CI)	
Protein	**Gene name**	**AC**	**DC**	**AC**	**DC**
Leukotriene A-4 hydrolase	LTA4H	8.05	6.66	0.995 (0.982-1.008)	1.0 (1.0-1.0)
Catalase	CAT	8.83	7.31	0.967 (0.903-1.031)	1.0 (1.0-1.0)
Transketolase	TKT	7.23	7.39	0.967 (0.922-1.012)	0.989 (0.962-1.017)
Annexin A3	ANXA3	18.59	9.35	0.972 (0.929-1.015)	0.937 (0.849-1.025)

Table 4. 1. Proteins included in biomarker panel for pediatric IBD diagnosis.

Table adapted from Deeke, S.A., Starr, A.E., Ning, Z., Ahmadi, S., Zhang, X., Mayne, J., Chiang, C.K., Singleton, R., Benchimol, E.I., Mack, D.R., Stintzi, A., and Figeys, D. (2018). Mucosal-luminal interface proteomics reveals biomarkers of pediatric inflammatory bowel disease-associated colitis. Am J Gastroenterol *113*, 713-724.



	True Control	True IBD
Test Control	18	1
Test IBD	0	21

	True Control	True IBD
Test Control	10	0
Test IBD	0	19

Figure 4. 3. IBD biomarker panel performance

(A) Utilizing the relative expression values from the panel of four proteins (LTA4H, CAT, TKT and ANXA3) yielded an ROC area under the curve of 0.989 (95% CI: 0.967-1.0) for the AC and 0.999 (95% CI: 0.999-1.0) for the DC.

(B) Utilizing both the CoN and CoA relative expression values from the panel of four proteins yields ROC AUC values of 0.9245 (95% CI: 0.8492 to 0.9998) for the AC and 0.9846 (95% CI: 0.9538 to 1.015) for the DC.

(C) Predictive class probabilities of the biomarker panel (LTA4H, CAT, TKT and ANXA3) in the AC and DC as indicated. Samples classified as IBD are located to the right of 0.5 (indicated by dotted line) and controls are located to the left of 0.5, producing a classification accuracy of 97.5% and 100% for the AC and DC respectively. IBD samples are displayed red and control samples are presented in blue.

Figure adapted from Deeke, S.A., Starr, A.E., Ning, Z., Ahmadi, S., Zhang, X., Mayne, J., Chiang, C.K., Singleton, R., Benchimol, E.I., Mack, D.R., Stintzi, A., and Figeys, D. (2018). Mucosal-luminal interface proteomics reveals biomarkers of pediatric inflammatory bowel disease-associated colitis. Am J Gastroenterol *113*, 713-724.

sensitivity >0.999 (95% CI: 0.8235 to 1.0) and a specificity >0.999 (95% CI: 0.6915 to 1.0). Taken together this produces a classification accuracy of 97.5% and 100% for the AC and DC respectively. With the inclusion of samples collected from colon regions without macroscopic inflammation (CoN), the IBD biomarker panel achieves ROC AUC values of 0.9245 (95% CI: 0.8492 to 0.9998) for the AC and 0.9846 (95% CI: 0.9538 to 1.015) for the DC (**Figure 4.3 B**). The relative expression of all four biomarker panel proteins was significantly elevated in CoN samples compared to non-IBD controls, apart from TKT in the AC (**Figure 4.4 A**). No difference was observed in the relative expression of the biomarker panel proteins between IBD subtypes (CD vs UC) (**Figure 4.4 B**).

4.2.1.1. Comparison of novel biomarker panel with Calprotectin at the intestinal MLI

Considering that calprotectin is the current clinical biomarker, the classification accuracy of the novel IBD biomarker panel was compared with that of calprotectin detected at the intestinal MLI. The ROC AUC of calprotectin at the MLI was 0.7551 (95% CI: 0.5944 to 0.9157) and 0.6632 (95% CI: 0.4567 to 0.8696) for the AC and DC respectively (**Figure 4.5 A**). In the AC, calprotectin (S100-A8 and S100-A9) produced a sensitivity of 0.682 (95% CI: 0.4513 to 0.8614) and specificity of 0.833 (95% CI: 0.5858 to 0.9642) with similar performance also observed in the DC (sensitivity of 0.684 (95% CI: 0.4345 to 0.8742) and specificity of 0.700 (95% CI: 0.3475 to 0.9333)). The classification accuracy of calprotectin was 75.0% for the AC and 69.0% for the DC, highlighting the improved performance of our novel biomarker panel. Notably, the poor classification performance observed herein for calprotectin could be influenced by the finding that the majority of quantification values were beyond the threshold (light/heavy ratio > 10 fold) of accurate detection.

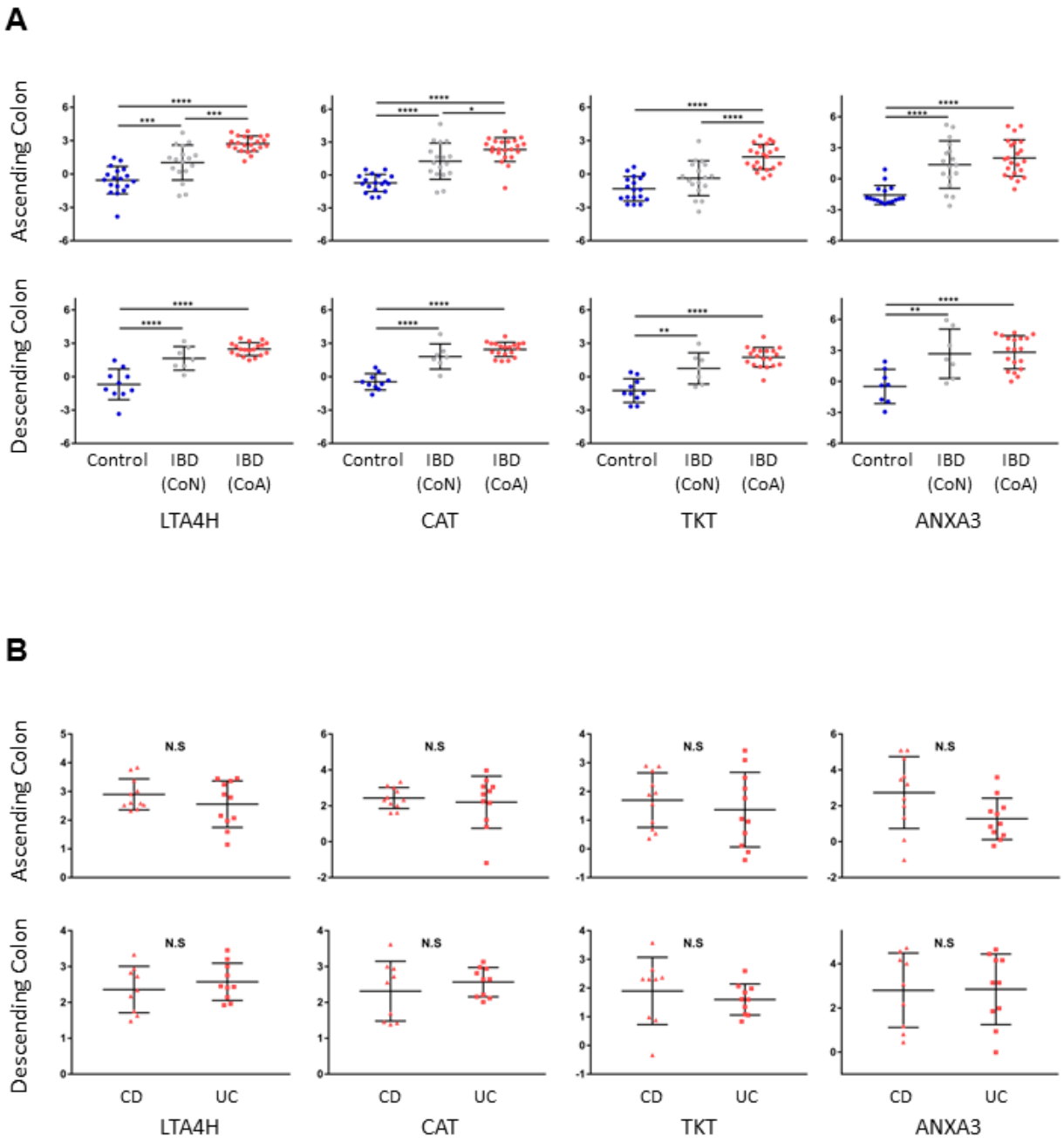

A
Ascending Colon
Descending Colon
Control
IBD (CoN)
IBD (CoA)
LTA4H
CAT
TKT
ANXA3
B
Ascending Colon
Descending Colon
N.S
CD
UC
LTA4H
CAT
TKT
ANXA3

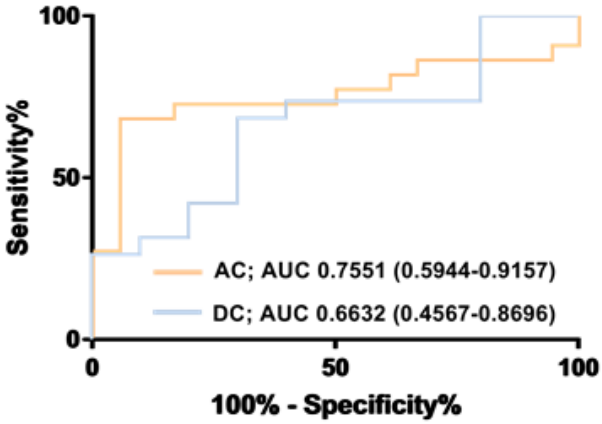
Figure 4. 4. IBD Biomarker proteins relative expression

(A) Comparison of protein relative expression of non-IBD controls with samples collected from colon areas with (CoA) and without (CoN) macroscopic inflammation. One-way ANOVA, Tukey's multiple comparison test; * denotes $P < 0.05$, ** denotes $P < 0.01$, *** denotes $P < 0.001$ and **** denotes $P < 0.0001$.

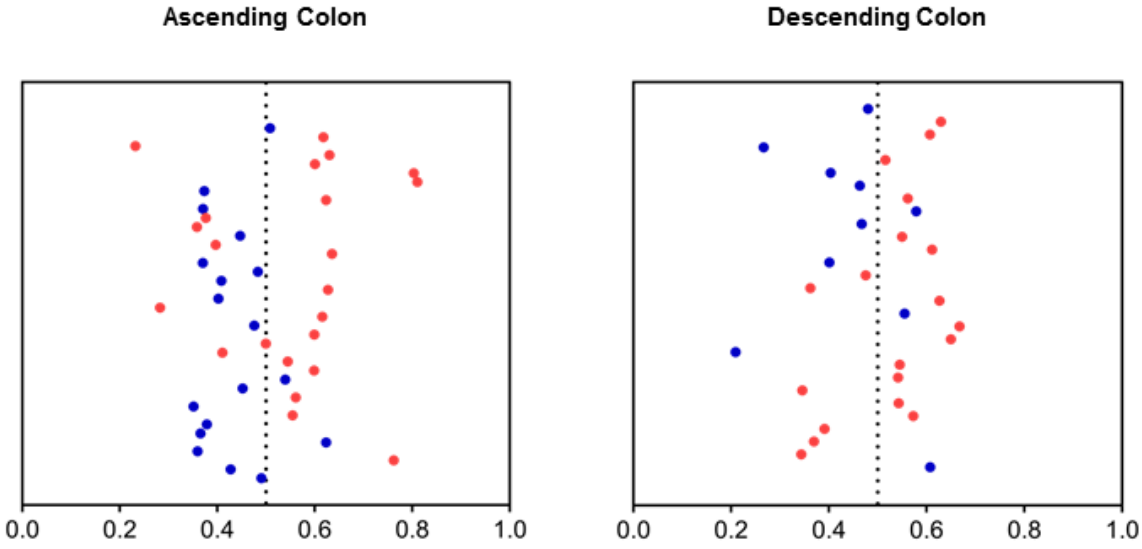
(B) Comparison of protein relative expression between IBD subtypes; CD (CoA) and UC (CoA). One-way ANOVA, Tukey's multiple comparison.

Figure published in Deeke, S.A., Starr, A.E., Ning, Z., Ahmadi, S., Zhang, X., Mayne, J., Chiang, C.K., Singleton, R., Benchimol, E.I., Mack, D.R., Stintzi, A., and Figeys, D. (2018). Mucosal-luminal interface proteomics reveals biomarkers of pediatric inflammatory bowel disease-associated colitis. Am J Gastroenterol *113*, 713-724.

A



B



	True Control	True IBD
Test Control	15	7
Test IBD	3	15

	True Control	True IBD
Test Control	7	6
Test IBD	3	13

Figure 4. 5. Calprotectin (S100-A8 and S100-A9) classification performance at the intestinal MLI

(A) ROC curve of intestinal MLI calprotectin (S100-A8 and S100-A9) from the AC (orange) and DC (blue).

(B) Predictive class probabilities of Calprotectin (S100-A8 and S100-A9) in the AC and DC as indicated. Samples classified as IBD are located to the right of 0.5 (indicated by dotted line) and non-IBD controls are located to the left of 0.5, producing a classification accuracy of 75.0% and 69.0% for the AC and DC respectively. IBD samples are displayed in red and control samples are presented in blue.

Figure adapted from Deeke, S.A., Starr, A.E., Ning, Z., Ahmadi, S., Zhang, X., Mayne, J., Chiang, C.K., Singleton, R., Benchimol, E.I., Mack, D.R., Stintzi, A., and Figeys, D. (2018). Mucosal-luminal interface proteomics reveals biomarkers of pediatric inflammatory bowel disease-associated colitis. Am J Gastroenterol *113*, 713-724.

4.2.2. Validation of select biomarkers on a non-invasive biological specimen

To aid in narrowing the bench to bedside gap, we validated a subset of biomarkers on a non-invasive biological specimen (stool). Using an independent cohort of children (n=24) (**Table 4.2**), the absolute quantity of CAT and LTA4H were assessed by ELISA. The differential expression observed at the intestinal MLI was recapitulated in stool, wherein both CAT ($P < 0.0001$) and LTA4H ($P = 0.0002$) demonstrated significant elevated expression in IBD samples compared to non-IBD controls (**Figure 4.6**). The cohort was then expanded to also include stool samples from patients for which their intestinal MLI aspirate samples were used for developing the biomarker panel (**Table 4.2**). Using this expanded cohort, the correlation of LTA4H and CAT expression in stool was compared with disease activity (PUCAI and PCDAI) and laboratory results (ESR, HCT and albumin) used for PCDAI establishment. Stool expression of LTA4H correlated with the PUCAI (Spearman r = 0.567 (95% CI: 0.1378 to 0.817), $P = 0.0114$) and albumin (r = -0.3696 (95% CI: 0.1378 to 0.03671), $P = 0.0265$) but not with ESR, HCT or PCDAI (**Table 4.3**). Stool expression of CAT correlated with albumin (r = -0.4293 (95% CI: -0.6669 to -0.1125); $P = 0.008$) but not with but not with ESR, HCT, PCDAI or PUCAI.

4.2.3. Biomarker panel development for UC pancolitis vs non-pancolitis (extent of disease)

By definition, the AC is inflamed in a UC patient with pancolitis whereas the AC does not display macroscopic inflammation in a UC patient without pancolitis. Therefore, to develop a biomarker panel which delineates UC pancolitis from non-pancolitis, the AC proteome of UC patients was further analyzed. Proteins considered for biomarker panel generation were identified by PLS-DA and displayed the highest ROC AUC values upon evaluation of intestinal MLI samples collected from AC inflamed areas (CoA) compared to intestinal MLI samples from non-

Study ID	Gender	Diagnosis	Age at endoscopy	Independent from discovery cohort	MLI proteomic sample used in discovery cohort	Stool collected pre or post endoscopy	Days between stool collection to endoscopy
HM308	Male	Control	10.2	*		POST	16
HM309	Female	Control	5.7	*		PRE	2
HM311	Female	CD	15.2	*		PRE	3
HM326	Female	UC	12.9	*		PRE	0
HM341	Male	UC	4.5	*		PRE	0
HM351	Male	UC	17.2	*		PRE	1
HM366	Female	UC	13.9	*		PRE	1
HM371	Male	Control	7.6	*		PRE	12
HM376	Female	Control	10.8	*		PRE	0
HM377	Female	UC	10.8	*		POST	11
HM382	Female	UC	13.2	*		PRE	2
HM384	Male	UC	14.3	*		PRE	0
HM387	Male	Control	11.4	*		PRE	1
HM402	Female	UC	7.4		*	PRE	1
HM411	Male	UC	15.3		*	POST	2
HM418	Male	CD	10.8		*	PRE	2
HM419	Male	UC	10.6		*	POST	3
HM426	Male	UC	15.5		*	PRE	4
HM435	Male	UC	12.8	*		PRE	1
HM465	Male	CD	14.4		*	PRE	11
HM469	Female	UC	12.5		*	PRE	39
HM481	Male	UC	14.9		*	PRE	51
HM482	Female	UC	17.7		*	PRE	0
HM483	Male	CD	13.2		*	PRE	0
HM491	Female	UC	17.5		*	PRE	2
HM494	Female	Control	17.5		*	POST	2
HM503	Female	Control	15.7	*		POST	7
M505	Male	UC	15.8		*	PRE	0
HM513	Female	Control	8.8		*	PRE	1
HM514	Male	UC	16.2		*	POST	2
HM540	Male	Control	16.9	*		PRE	2
HM578	Male	Control	6.5	*		PRE	1
HM581	Female	Control	16.2	*		PRE	3
HM582	Male	UC	14.2	*		PRE	2
HM588	Male	CD	16.8	*		PRE	1
HM601	Female	CD	17.6	*		PRE	2
HM602	Female	UC	10.2	*		PRE	2
HM609	Female	Control	14.7	*		PRE	1
HM617	Female	UC	16.7	*		PRE	1
HM618	Male	CD	10.7	*		PRE	2
HM629	Male	CD	14.4	*		PRE	2

Table 4. 2: **Patient characteristics for validation of select biomarkers on a non-invasive biological specimen (stool).**

Stool samples collected from new-onset IBD and non-IBD controls prior to the initiation of pharmacological treatment.

Table published in Deeke, S.A., Starr, A.E., Ning, Z., Ahmadi, S., Zhang, X., Mayne, J., Chiang, C.K., Singleton, R., Benchimol, E.I., Mack, D.R., Stintzi, A., and Figeys, D. (2018). Mucosal-luminal interface proteomics reveals biomarkers of pediatric inflammatory bowel disease-associated colitis. Am J Gastroenterol *113,* 713-724.

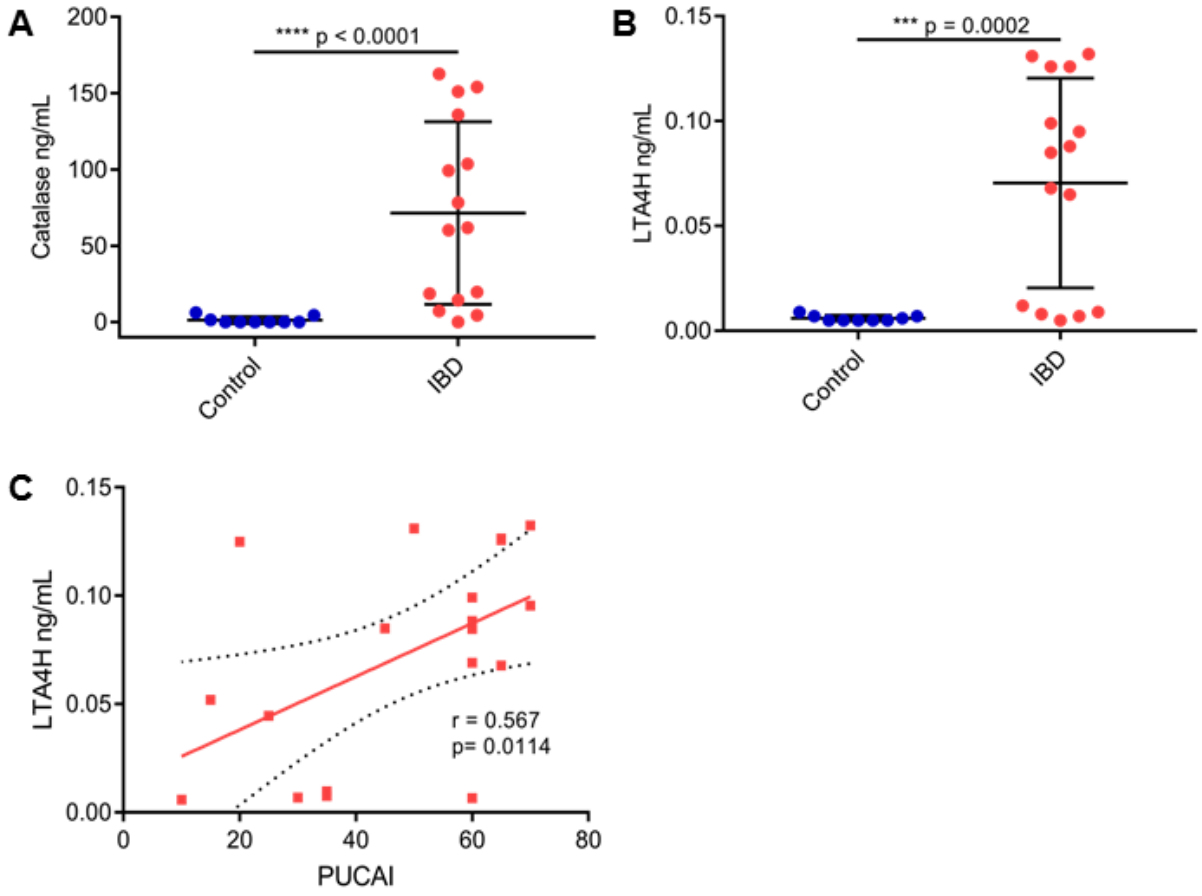

A
Catalase ng/mL
200
150
100
50
0
**** p < 0.0001
Control
IBD
B
LTA4H ng/mL
0.15
0.10
0.05
0.00
*** p = 0.0002
Control
IBD
C
LTA4H ng/mL
0.15
0.10
0.05
0.00
r = 0.567
p= 0.0114
0
20
40
60
80
PUCAI

Figure 4. 6. Evaluation of CAT and LTA4H levels in stool

The expression of **(A)** CAT (ng/mL) and **(B)** LTA4H (ng/mL) was evaluated in stool by ELISA on an independent cohort of 24 pediatric patients. Mann Whitney test, *** denotes *P* values < 0.001 and **** denotes *P* value < 0.0001.

(C) Correlation of LTA4H stool expression with the PUCAI considering both independent patients and patients for which their intestinal aspirate samples were used for developing the biomarker panel (n=38). Spearman r = 0.567 (95% CI: 0.1378 to 0.817), *P* = 0.0114.

ELISAs performed by Sara Ahmadi and Amanda Starr. Inter-plate normalization performed by Amanda Starr. Statistical analysis and patient eligibility performed by Shelley Deeke.

Figure adapted from Deeke, S.A., Starr, A.E., Ning, Z., Ahmadi, S., Zhang, X., Mayne, J., Chiang, C.K., Singleton, R., Benchimol, E.I., Mack, D.R., Stintzi, A., and Figeys, D. (2018). Mucosal-luminal interface proteomics reveals biomarkers of pediatric inflammatory bowel disease-associated colitis. Am J Gastroenterol *113*, 713-724.

	PCDAI	PUCAI	HCT (%)	ESR (mm/hr)	Albumin (g/dL)
			LTA4H		
Spearman r	0.5238	0.567	-0.1106	0.2759	-0.3696
95% confidence interval	NA	0.1378 to 0.817	-0.4278 to 0.2308	-0.06794 to 0.5611	-0.6287 to -0.03671
***P* (two-tailed)**	0.1966	0.0114	0.5146	0.1034	0.0265
Significant? (alpha = 0.05)	No	Yes	No	No	Yes
			CAT		
Spearman r	0.7066	0.22	-0.2503	0.2243	-0.4293
95% confidence interval	NA	-0.2597 to 0.6126	-0.5347 to 0.0852	-0.1173 to 0.5185	-0.6669 to -0.1125
***P* (two-tailed)**	0.0595	0.3513	0.1297	0.182	0.008
Significant? (alpha = 0.05)	No	No	No	No	Yes

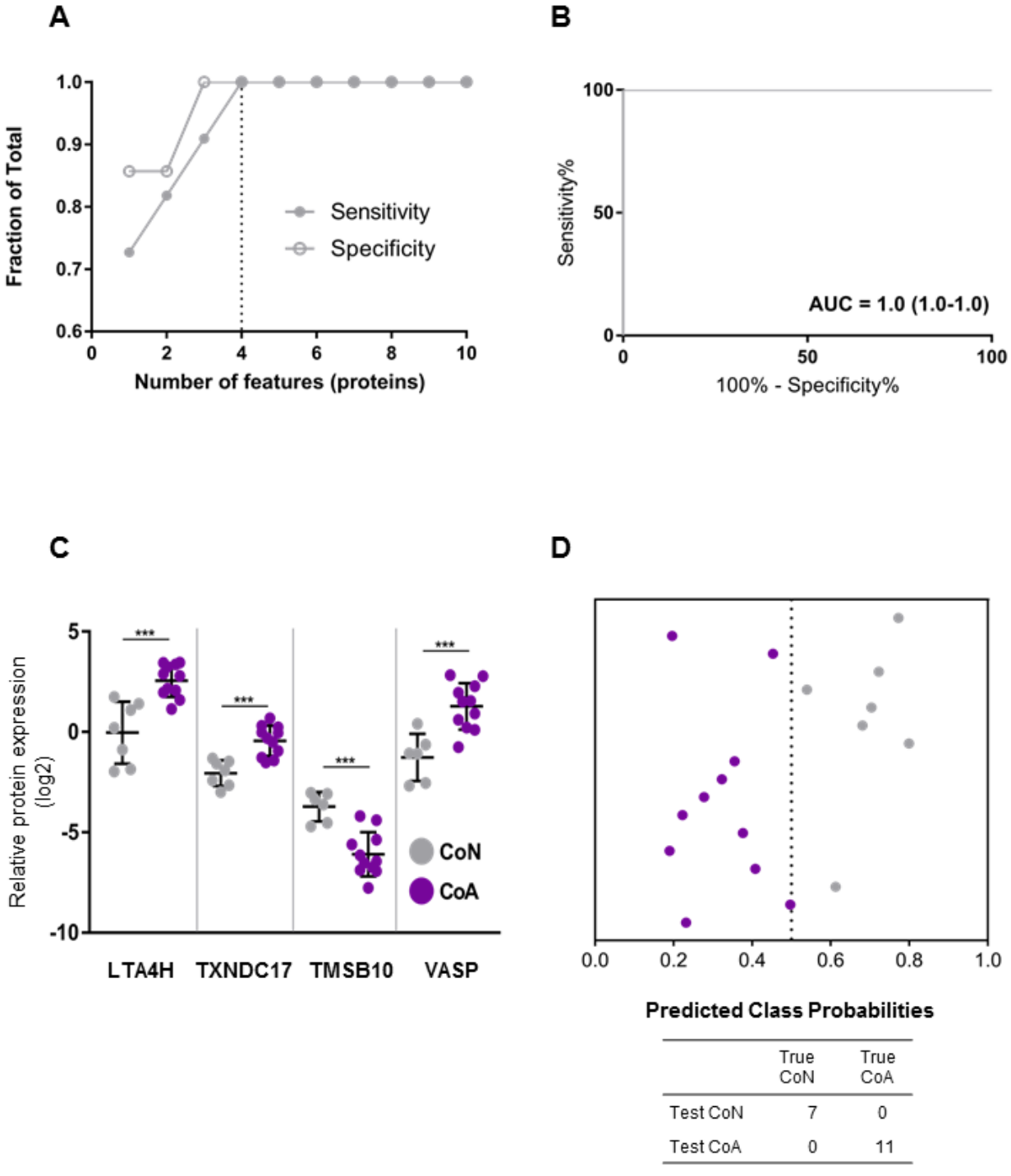
Table 4. 3: Correlation of LTA4H and CAT stool expression with disease activity.

The correlation of LTA4H and CAT expression in stool was compared with disease activity (PUCAI and PCDAI) and laboratory results (ESR, HCT and albumin).

ELISAs performed by Sara Ahmadi and Amanda Starr. Inter-plate normalization performed by Amanda Starr. Statistical analysis and patient eligibility performed by Shelley Deeke.

Table adapted from Deeke, S.A., Starr, A.E., Ning, Z., Ahmadi, S., Zhang, X., Mayne, J., Chiang, C.K., Singleton, R., Benchimol, E.I., Mack, D.R., Stintzi, A., and Figeys, D. (2018). Mucosal-luminal interface proteomics reveals biomarkers of pediatric inflammatory bowel disease-associated colitis. Am J Gastroenterol *113*, 713-724.

inflamed areas (CoN). Evaluation by iterative analysis deemed four proteins as the minimum number required to achieve the maximum sensitivity and specificity (**Figure 4.7 A**). The relative expression of proteins featuring in the panel is shown in **Figure 4.7 C** and biomarker protein characteristics listed in **Table 4.4**. ROC curve analysis of the relative expression of the proteins featuring in the biomarker panel yielded a sensitivity of >0.9999 (95% CI: 0.5904 to 1.0), a specificity of >0.9999 (95% CI: 0.7151to 1.0), a ROC AUC value of 1.0 and 100% classification accuracy (**Figure 4.7 B, D**).



	True CoN	True CoA
Test CoN	7	0
Test CoA	0	11

Figure 4. 7. UC pancolitis versus non-pancolitis biomarker panel generation and performance

(A) By iterative analysis, four proteins were deemed the minimum number required to achieve maximal sensitivity and specificity. Proteins identified by PLS-DA and with the highest ROC AUC values by comparison of intestinal MLI samples collected from areas with (CoA) and without (CoN) inflammation were considered for biomarker panel development.

(B) Applying the panel of four proteins (LTA4H, TXNDC17, TMBS10 and VASP) to a ROC analysis produced a ROC AUC of 1.0 in discriminating pancolitis and non-pancolitis.

(C) Relative expression (Log2 normalized light/heavy) of UC pancolitis vs non-pancolitis protein biomarkers. T-test was performed, *** denotes $P < 0.001$

(D) Predictive class probabilities of the biomarker panel (LTA4H, TXNDC17, TMBS10 and VASP) in differentiating pancolitis from non-pancolits. Samples classified as uninflamed are located to the right of 0.5 (indicated by dotted line) and sample deemed inflamed are located to the left of 0.5, producing a classification accuracy of 100%. Samples collected from the AC without evidence of macroscopic inflammation are displayed in grey and AC samples from areas with macroscopic inflammation are presented in purple.

Figure adapted from Deeke, S.A., Starr, A.E., Ning, Z., Ahmadi, S., Zhang, X., Mayne, J., Chiang, C.K., Singleton, R., Benchimol, E.I., Mack, D.R., Stintzi, A., and Figeys, D. (2018). Mucosal-luminal interface proteomics reveals biomarkers of pediatric inflammatory bowel disease-associated colitis. Am J Gastroenterol *113*, 713-724.

Protein	Gene name	UC CoA/ UC CoN Fold change	UC CoA vs UC CoN AUC (CI)
Leukotriene A-4 hydrolase	LTA4H	4.54	0.961 (0.881-1.041)
Thioredoxin domain-containing protein 17	TXNDC17	3.17	0.961 (0.881-1.041)
Thymosin beta-10	TMSB10	0.23	0.948 (0.851-1.045)
Vasodilator-stimulated phosphoprotein	VASP	5.83	0.935 (0.824-1.046)

Table 4. 4: Proteins included in biomarker panel for pediatric extent of disease in UC (pancolitis vs non-pancolitis).

Fold change and individual ROC AUC values for proteins featuring in the UC pancolitis vs non-pancolitis biomarker panel consisting of four proteins (LTA4H, TXNDC17, TMBS10 and VASP).

Table adapted from Deeke, S.A., Starr, A.E., Ning, Z., Ahmadi, S., Zhang, X., Mayne, J., Chiang, C.K., Singleton, R., Benchimol, E.I., Mack, D.R., Stintzi, A., and Figeys, D. (2018). Mucosal-luminal interface proteomics reveals biomarkers of pediatric inflammatory bowel disease-associated colitis. Am J Gastroenterol *113*, 713-724.

CHAPTER 5. Extracellular vesicle proteomic signature in new-onset IBD

5.1 Rationale

EVs can deliver molecular cargo to neighboring or distant cells and thus alter the recipient cells' state. Furthermore, EVs can promote or inhibit immune responses and activate immune cells (Singh et al., 2012), rendering this deliver vehicle particularly relevant in IBD, which is an immune mediated chronic disease. Few studies have investigated EVs in the context of IBD (Leonetti et al., 2013; Leoni et al., 2015; Mitsuhashi et al., 2016; Palkovits et al., 2013; Zheng et al., 2017), and none have thus far performed proteomics of EVs isolated from the intestinal MLI of individuals with IBD. Furthermore, no studies have investigated intestinal EVs from pediatric IBD participants, a population which demonstrates a phenotypically distinct form of the disease as compared to adults.

Similarly to eukaryotes, bacteria also secrete membrane-bound vesicles into the extracellular environment and can thus influence a wide range of physiological processes including intercellular bacterial communication, defending bacterial from antimicrobial peptides and bacteriophages (Manning and Kuehn, 2011; Reyes-Robles et al., 2018), delivering virulent factors (Fiocca et al., 1999), destroying competitor bacteria and providing antibiotic resistance to bacterial cells (Ciofu et al., 2000). Despite their capacity for modulating host-microbe interactions, the proteomic composition of bacterial OMVs within the human intestine remains unidentified.

My hypotheses are (1) that intestinal EV proteomic cargo is altered in IBD compared to non-IBD controls and (2) select proteins exhibit differential expression in EVs in IBD compared to non-IBD controls and do not demonstrate differential expression in their corresponding biopsies. Accordinlgy my objectives are (1) to assess the proteomic differences in new-onset

pediatric IBD intestinal EVs as compared to non-IBD controls and (2) Compare proteomic signatures of EVs and their corresponding intestinal biopsies.

5.2 Results

5.2.1. Patient characteristics of isolated EVs

EVs were isolated from intestinal MLI aspirates collected from the DC of 11 treatment-naive IBD patients (7 UC, 3 CD and 1 IBDU) and 7 non-IBD controls during diagnostic colonoscopy. IBD samples were collected from affected areas with macroscopic inflammation whereas non-IBD control samples were collected from areas without macroscopic inflammation. No difference in age was observed between the non-IBD control and IBD groups ($P = 0.2553$, Mann-Whitney test). Moreover, no significant difference was observed between sexes in either group (non-IBD control $P = 0.0635$, IBD $P = 0.500$, Binomial one-tailed test).

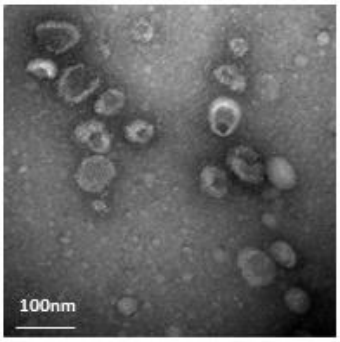
5.2.2. Pediatric descending colon EV characteristics

EV morphology was assessed by TEM, and overall size distribution was assessed using nanoparticle tracking analysis (NTA), demonstrating membrane-enclosed vesicles with sizes in agreement with those of exosomes, microvesicles and bacterial OMVs (**Figure 5.1 A, B**). Comparison of the mean size (nm) of a subset of participant EVs yielded no differences between the mean sizes of EVs isolated from non-IBD controls, and those isolated from IBD patients (Mann-Whitney test, $P = 0.6857$) (**Figure 5.1 C**).

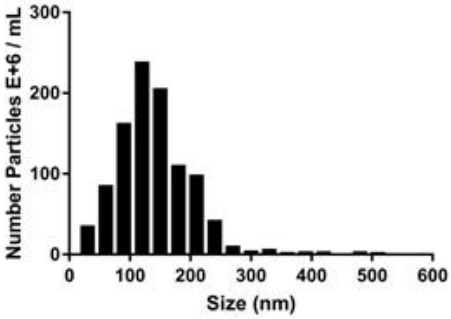
5.2.3. Proteomic overview of EVs isolated from the intestinal MLI of children with and without IBD

In total 3,399 proteins were identified: 2,645 proteins were of human origin, 752 proteins were of microbial origin, and 2 proteins matched to both human and microbe databases. The

A


100nm

B


Number Particles E+6 / mL
300
200
100
0
0
100
200
300
400
500
600
Size (nm)

C

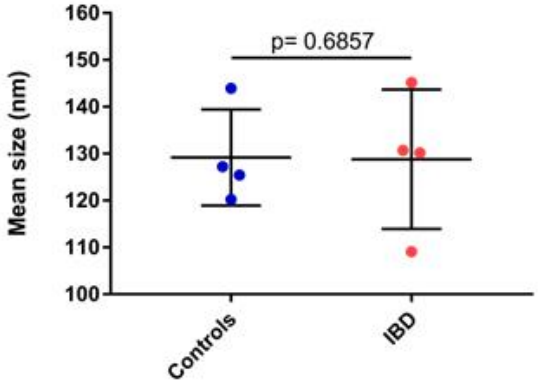

Mean size (nm)
160
150
140
130
120
110
100
p= 0.6857
Controls
IBD

Figure 5. 1.Characteristics of EVs isolated by ultracentrifugation from the intestinal MLI of pediatric IBD patients.

(A) Transmission electron microscopy image of extracellular vesicles isolated from the descending colon. Fee for service sample performed at the University of Ottawa Heart Institute, Loeb Research Centre, Electron microscopy laboratory, Cell Imaging and Histology Core Facility.

(B) Size distribution profile from EVs isolated from an intestinal MLI sample collected from the descending colon of a non-IBD patient (ID 537) assessed by Nanoparticle Tracking Analysis

(C) Comparison of the mean size (nm) of EVs isolated from non-IBD controls and IBD participants. Mann-Whitney test was performed to evaluate statistical difference.

Figure adapted from Zhang, X., S.A. Deeke, Z.B. Ning, A.E. Starr, J. Butcher, J. Li, J. Mayne, K. Cheng, B. Liao, L.Y. Li, R. Singleton, D. Mack, A. Stintzi, and D. Figeys. 2018a. Metaproteomics reveals associations between microbiome and intestinal extracellular vesicle proteins in pediatric inflammatory bowel disease. *Nature Communications*

This work is licensed under a Creative Commons Attribution 4.0 International License (https://creativecommons.org/licenses/by/4.0/)

intensity-based contribution of human proteins to the overall LFQ intensity ranged between 82.8% and 99.1%, and no significant difference was observed between IBD and non-IBD controls (P = 0.1042, Mann Whitney test) (**Figure 5.2 A**). Of the Top 100 protein list from ExoCarta (Keerthikumar et al., 2016), an exosomal database repository, 85 were identified in the EV proteomic dataset (**Table 5.1**). This list includes proteins involved in exosomal biogenesis such as CD9, CD63, TSG101 and PDCD6IP, which are commonly used exosomal markers that are involved in multivesicular body formation and intraluminal vesicle sorting (Baietti et al., 2012; Kalra et al., 2016; van Niel et al., 2011).

5.2.4. IBD proteomic signature of intestinal EV

To compare the proteomic profile of EVs isolated from the intestinal MLI of children with IBD to their non-IBD counterpart, PLS-DA was performed on the combined host- and microbial-derived proteins. A total of 191 proteins with a VIP > 1.0 were identified, eight of which are of microbial origin (**Table 5.2)**. Following gene ontology enrichment analysis of the remaining 183 proteins of human origin, it was found that several of the top 10 biological processes were related to immunity and host defense, namely defense response to bacterium, antibacterial humoral response, and immune system process (**Figure 5.3 A**). Interestingly, of the 17 proteins corresponding to the defense response to bacterium process, the majority (82%) were elevated in IBD EVs compared to non-IBD control EVs, suggesting the potential contribution of intestinal EVs to the dysregulated immune response which is observed in IBD (**Figure 5.3 B**).

5.2.5. Comparing the EV and biopsy proteome alterations in IBD

Intestinal biopsy samples corresponding to the same patients from which the EV were isolated (7 non-IBD controls and 11 IBD) were analyzed by high resolution MS. In total, 3972 proteins were identified in biopsies, with no difference observed in the number of proteins from

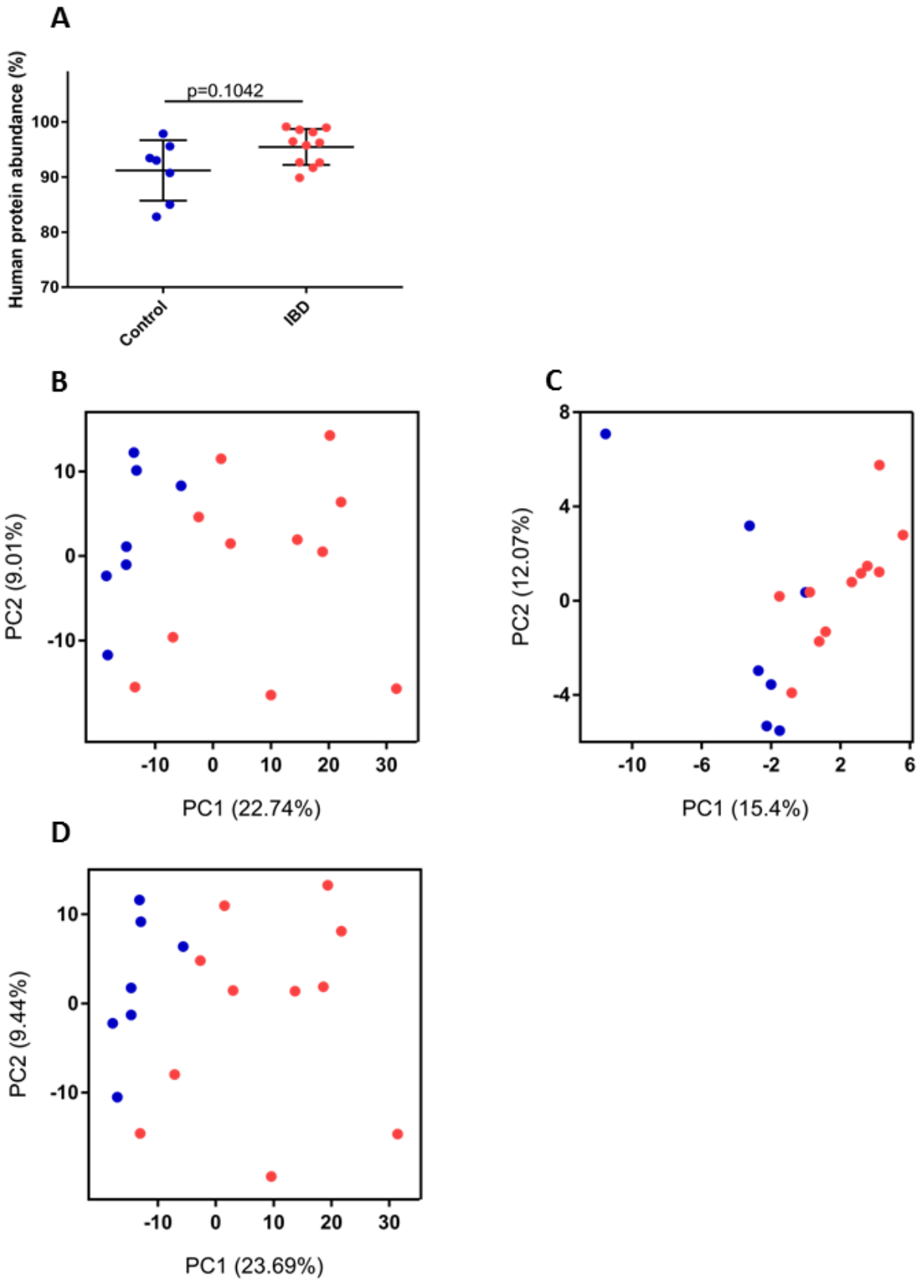

A
Human protein abundance (%)
p=0.1042
100
90
80
70
Control
IBD
B
PC2 (9.01%)
10
0
-10
-10
0
10
20
30
PC1 (22.74%)
C
PC2 (12.07%)
8
4
0
-4
-10
-6
-2
2
6
PC1 (15.4%)
D
PC2 (9.44%)
10
0
-10
-10
0
10
20
30
PC1 (23.69%)

Figure 5. 2. Proteomic overview of EVs isolated from the intestinal MLI collected from children with and without IBD

(A) Percentage of human protein contribution to overall LFQ intensity. IBD samples are indicated in red whereas non-IBD controls are presented in blue. Mann-Whitney test was performed to evaluate statistical difference.

(B) PCA of host and microbial EV proteins identified in at least 50% of samples. IBD samples are indicated in red whereas non-IBD controls are presented in blue.

(C) PCA of microbial EV proteins identified in at least 50% of samples. IBD samples are indicated in red whereas non-IBD controls are presented in blue.

(D) PCA of host EV proteins identified in at least 50% of samples. IBD samples are indicated in red whereas non-IBD controls are presented in blue.

Data further analyzed from Zhang, X., S.A. Deeke, Z.B. Ning, A.E. Starr, J. Butcher, J. Li, J. Mayne, K. Cheng, B. Liao, L.Y. Li, R. Singleton, D. Mack, A. Stintzi, and D. Figeys. 2018a. Metaproteomics reveals associations between microbiome and intestinal extracellular vesicle proteins in pediatric inflammatory bowel disease. *Nature Communications*

This work is licensed under a Creative Commons Attribution 4.0 International License (https://creativecommons.org/licenses/by/4.0/)

Uniprot Protein ID	Gene Name	Average number of peptides per patient	Uniprot Protein ID	Gene Name	Average number of peptides per patient
A0A087WVQ6	CLTC	13.4	P22314	UBA1	2.0
A0A087X2B5	BSG	0.5	P23229	ITGA6	0.5
A0A0A0MSI0	PRDX1	5.5	P23526	AHCY	1.9
A0A0U1RQF0	FASN	2.9	P23528	CFL1	4.5
A6NNI4	CD9	3.4	P26038	MSN	10.2
B0AZS6	YWHAZ	4.1	P27105	STOM	9.9
B4DUR8	CCT3	2.1	P27348	YWHAQ	2.8
C9JIS1	GNB2	3.6	P29401	TKT	6.4
E7EQR4	EZR	14.8	P31946	YWHAB	4.4
E7ESV4	RAP1B	3.6	P32119	PRDX2	6.0
E9PK54	HSPA8	3.9	P35579	MYH9	25.2
E9PS23	CFL1	4.1	P48643	CCT5	1.9
F5H442	TSG101	2.5	P50395	GDI2	5.4
F6X3N5	GNB1	3.1	P50995	ANXA11	8.0
F8VWK8	CD63	2.4	P51148	RAB5C	2.2
O00299	CLIC1	4.8	P51149	RAB7A	3.1
O00560	SDCBP	5.2	P53396	ACLY	2.7
O43707	ACTN4	5.4	P55072	VCP	2.1
O75955	FLOT1	1.8	P60174	TPI1	4.1
P00338	LDHA	9.4	P60953	CDC42	2.7
P00558	PGK1	6.0	P61006	RAB8A	3.0
P01023	A2M	15.3	P61020	RAB5B	1.6
P02768	ALB	22.9	P61106	RAB14	1.3
P02786	TFRC	1.3	P61586	RHOA	3.2
P04075	ALDOA	6.5	P61981	YWHAG	3.8
P04083	ANXA1	7.4	P62258	YWHAE	4.4
P04406	GAPDH	5.1	P62805	HIST1H4A	6.4
P04899	GNAI2	4.6	P62820	RAB1A	2.8
P05023	ATP1A1	9.2	P62826	RAN	2.7
P05556	ITGB1	3.3	P62873	GNB1	4.0
P06733	ENO1	7.6	P62937	PPIA	2.9
P07195	LDHB	5.0	P63000	RAC1	6.8
P07355	ANXA2	9.1	P63092	GNAS	3.5
P07737	PFN1	4.8	P63104	YWHAZ	11.2
P07900	HSP90AA1	7.5	P78371	CCT2	1.3
P08133	ANXA6	8.0	P84077	ARF1	3.3
P08238	HSP90AB1	6.6	Q04917	YWHAH	8.4
P08758	ANXA5	4.0	Q08380	LGALS3BP	6.9
P09525	ANXA4	7.5	Q14764	MVP	16.5
P11021	HSPA5	2.4	Q14974	KPNB1	4.1
P11142	HSPA8	7.7	Q60FE5	FLNA	13.8
P13473	LAMP2	0.4	Q71U36	TUBA1A	3.9
P13639	EEF2	6.9	Q8WUM4	PDCD6IP	23.2
P14618	PKM	9.4	Q9H223	EHD4	4.1
P17987	TCP1	3.2	Q9P2B2	PTGFRN	0.3
			V9GZ54	MSN	2.8

Table 5. 1. Proteins in the EV dataset that match to the ExoCarta Top100 protein list.

Of the Top 100 protein list from ExoCarta (Keerthikumar et al., 2016), an exosomal database repository, 85 were identified in the EV proteomic dataset.

Data further analyzed from Zhang, X., S.A. Deeke, Z.B. Ning, A.E. Starr, J. Butcher, J. Li, J. Mayne, K. Cheng, B. Liao, L.Y. Li, R. Singleton, D. Mack, A. Stintzi, and D. Figeys. 2018a. Metaproteomics reveals associations between microbiome and intestinal extracellular vesicle proteins in pediatric inflammatory bowel disease. *Nature Communications*

This work is licensed under a Creative Commons Attribution 4.0 International License (https://creativecommons.org/licenses/by/4.0/)

Protein ID	COG name	Taxonomy
435590.BVU_0563	Outer membrane receptor for ferrienterochelin and colicins	Bacteroides
FungiAET14829.1	heat shock protein 70	Cladosporium cladosporioides
MH0002_GL0000402	NADPH-dependent 2,4-dienoyl-CoA reductase, sulfur reductase, or a related oxidoreductase	Bacteroides vulgatus
MH0089_GL0033973	Hypothetical protein	Faecalibacterium prausnitzii
MH0187_GL0031525	Hypothetical protein	Faecalibacterium prausnitzii
MH0192_GL0134142	Tetratricopeptide (TPR) repeat	Prevotella sp. CAG:279
MH0243_GL0008858	Serine/threonine protein kinase	Prevotella
MH0288_GL0054326	Outer membrane receptor for monomeric catechols	Bacteroides vulgatus

Table 5. 2. **EV Proteins of microbial origin identified as discriminant features by PLS-DA in IBD compared to non-IBD controls.**

A total of 191 proteins with a VIP > 1.0 were identified by PLS-DA, eight of which are of microbial origin.

Data further analyzed from Zhang, X., S.A. Deeke, Z.B. Ning, A.E. Starr, J. Butcher, J. Li, J. Mayne, K. Cheng, B. Liao, L.Y. Li, R. Singleton, D. Mack, A. Stintzi, and D. Figeys. 2018a. Metaproteomics reveals associations between microbiome and intestinal extracellular vesicle proteins in pediatric inflammatory bowel disease. *Nature Communications*

This work is licensed under a Creative Commons Attribution 4.0 International License (https://creativecommons.org/licenses/by/4.0/)

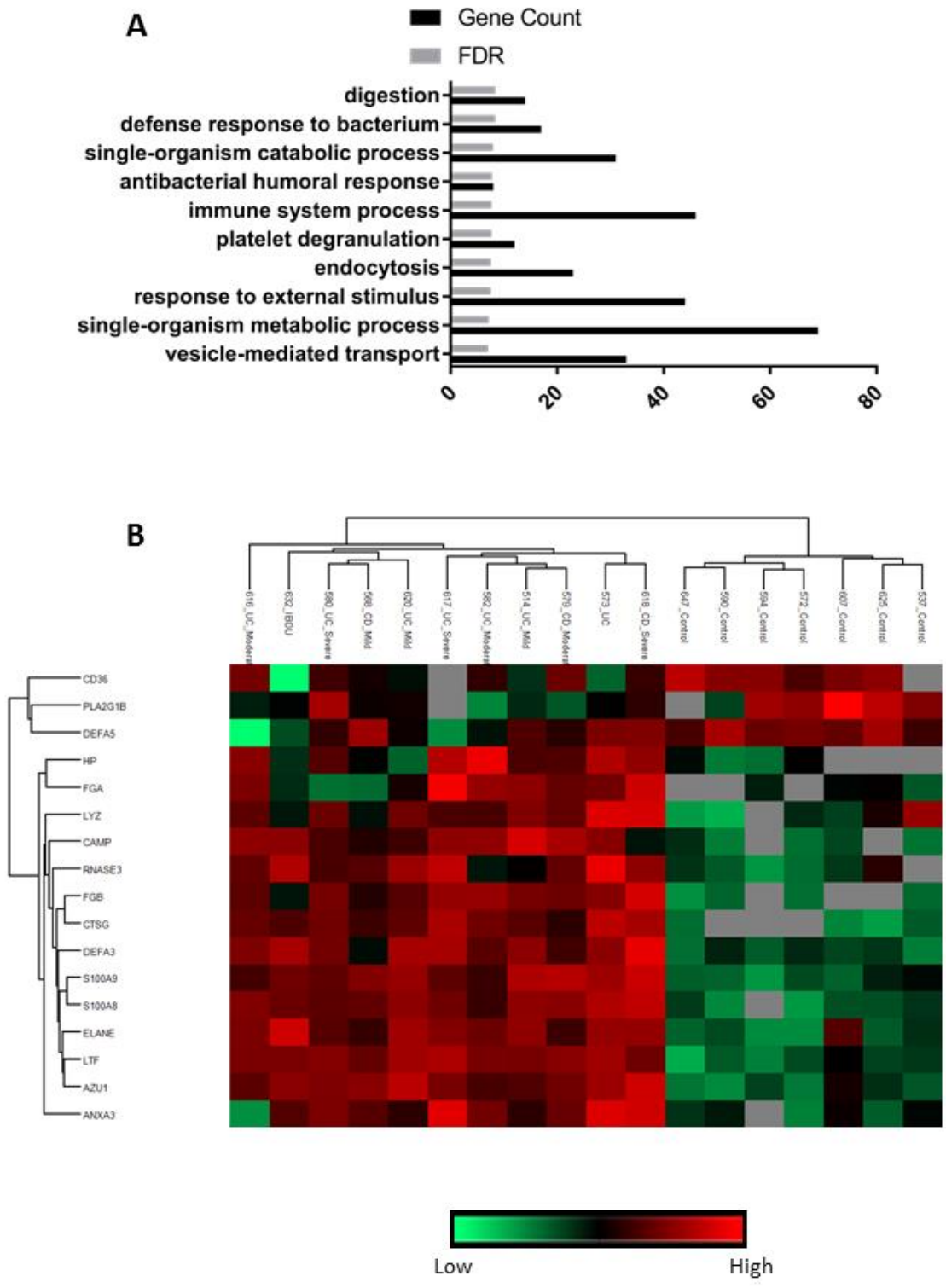

A
Gene Count
FDR
digestion
defense response to bacterium
single-organism catabolic process
antibacterial humoral response
immune system process
platelet degranulation
endocytosis
response to external stimulus
single-organism metabolic process
vesicle-mediated transport
0
20
40
60
80
B
616_UC_Moderate
632_IBDU
580_UC_Severe
569_CD_Mild
620_UC_Mild
617_UC_Severe
582_UC_Moderate
514_UC_Mild
579_CD_Moderate
573_UC
618_CD_Severe
647_Control
590_Control
594_Control
572_Control
607_Control
625_Control
537_Control
CD36
PLA2G1B
DEFA5
HP
FGA
LYZ
CAMP
RNASE3
FGB
CTSG
DEFA3
S100A9
S100A8
ELANE
LTF
AZU1
ANXA3
Low
High

Figure 5. 3. Discriminant features identified upon the comparison of non-IBD controls and IBD intestinal EVs include those associated with host defense response.

(A) Top10 biological processes of host proteins differentially expressed in IBD EVs compared to non-IBD controls. FDR values were –Log10 transformed.

(B) Hierarchical clustering of proteins identified as discriminant feature corresponding to the defense response to bacterium category. Log10 (LFQ) values underwent Z score normalization. Grey indicates lack of MS detection.

Data further analyzed from Zhang, X., S.A. Deeke, Z.B. Ning, A.E. Starr, J. Butcher, J. Li, J. Mayne, K. Cheng, B. Liao, L.Y. Li, R. Singleton, D. Mack, A. Stintzi, and D. Figeys. 2018a. Metaproteomics reveals associations between microbiome and intestinal extracellular vesicle proteins in pediatric inflammatory bowel disease. *Nature Communications*

This work is licensed under a Creative Commons Attribution 4.0 International License (https://creativecommons.org/licenses/by/4.0/)

non-IBD controls compared to those from IBD patients (P = 0.8601, Mann Whitney test). On average, 2.3 ± 0.45 fold additional proteins were identified in biopsies compared to their EV counterparts. Similarly to EVs, the majority of the intensity-based contribution in intestinal biopsies was attributed to human proteins, with the human-derived protein LFQ intensity accounting for 94.5% to 96.0% of the overall LFQ intensity. No significant difference was observed between the human protein intensity-based contribution of IBD patients and non-IBD controls (P = 0.1509, Mann Whitney test) (**Figure 5.4 A**). To compare proteins which are differentially expressed in IBD compared to non-IBD controls in EVs but not in biopsies, t-test with Benjamini-Hochberg FDR adjustment was performed on both the EV and biopsy datasets. There were 27 proteins with a q value < 0.05 and > 1.0 in the EV and biopsy datasets respectively (**Table 5.3**). Several of the top biological processed related to these 27 proteins were associated with actin cytoskeleton remodeling including actin nucleation, actin filament organization, Arp2/3 complex-mediated actin nucleation and actin cytoskeleton organization (**Figure 5.4 B**). Notably, several proteins were also linked to the gene ontology biological process of "killing of cells in other organism involved in symbiotic interaction" further supporting the evidence that intestinal EVs are contributing to the dysregulated immune response in IBD.

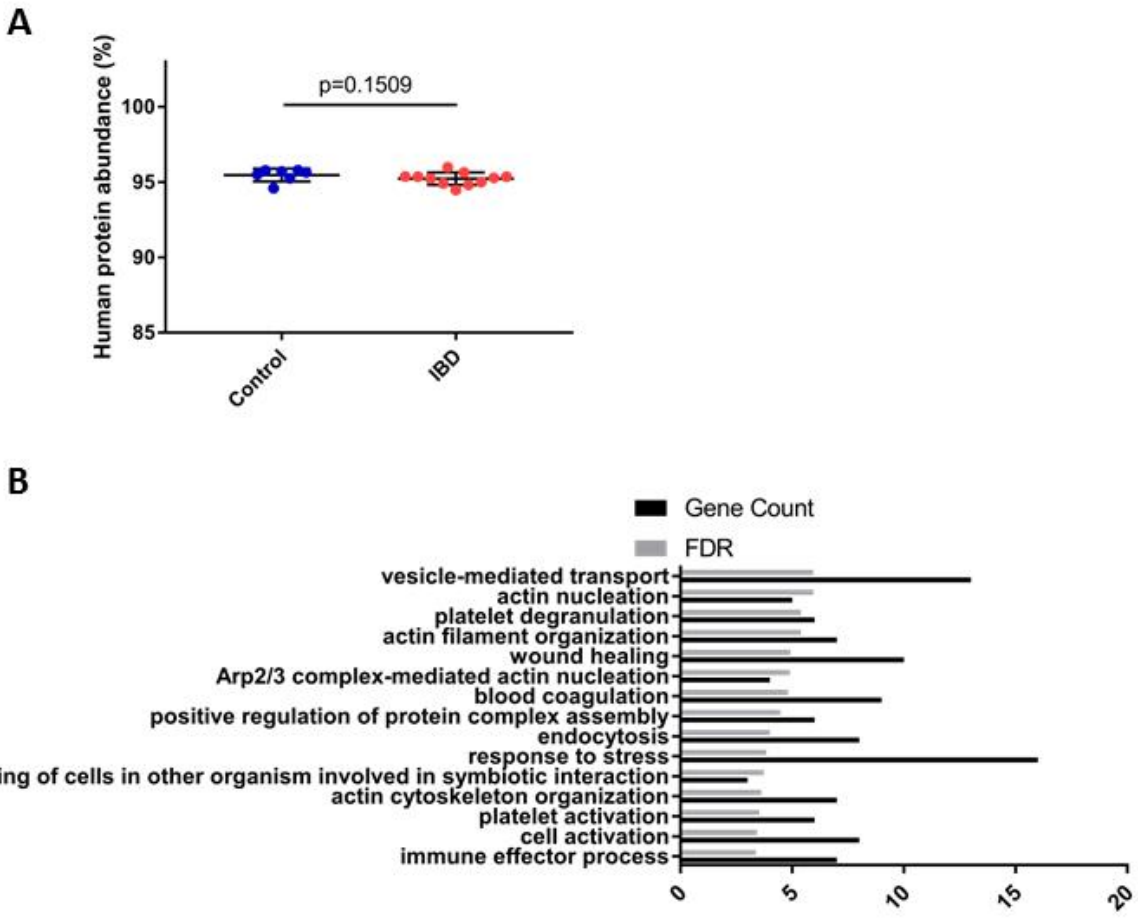

A
Human protein abundance (%)
p=0.1509
100
95
90
85
Control
IBD
B
Gene Count
FDR
vesicle-mediated transport
actin nucleation
platelet degranulation
actin filament organization
wound healing
Arp2/3 complex-mediated actin nucleation
blood coagulation
positive regulation of protein complex assembly
endocytosis
response to stress
killing of cells in other organism involved in symbiotic interaction
actin cytoskeleton organization
platelet activation
cell activation
immune effector process
0
5
10
15
20

Figure 5. 4. Comparing the EV and biopsy proteome alterations in IBD

(A) Intensity-based contribution of proteins of human origin to the overall LFQ intensity. No significant difference was observed between IBD patients and non-IBD controls ($P = 0.1509$, Mann Whitney test).

(B) Proteins differentially expressed in IBD compared to non-IBD controls in EVs but not in biopsies. T-test with Benjamini-Hochberg FDR adjustment was performed on both the EV and biopsy datasets. Proteins which yielded a q value < 0.05 and > 1.0 in the EV and biopsy datasets respectively were considered.

Data further analyzed from Zhang, X., S.A. Deeke, Z.B. Ning, A.E. Starr, J. Butcher, J. Li, J. Mayne, K. Cheng, B. Liao, L.Y. Li, R. Singleton, D. Mack, A. Stintzi, and D. Figeys. 2018a. Metaproteomics reveals associations between microbiome and intestinal extracellular vesicle proteins in pediatric inflammatory bowel disease. *Nature Communications*

This work is licensed under a Creative Commons Attribution 4.0 International License (https://creativecommons.org/licenses/by/4.0/)

Uniprot protein ID	Gene name	Control EV Log10 (LFQ Average)	IBD EV Log10 (LFQ Average)	Control Biopsy Log10 (LFQ Average)	IBD Biopsy Log10 (LFQ Average)	EV q value	Biopsy q value
A0A0A0MS51	GSN	6.68	8.74	8.66	8.63	0.006	0.855
F5H6E2	MYO1C	7.75	7.25	7.97	7.83	0.036	0.222
O15143	ARPC1B	6.61	8.24	8.02	8.16	0.023	0.134
O60218	AKR1B10	7.95	7.20	7.63	7.34	0.043	0.138
P01009	SERPINA1	8.11	8.55	8.44	8.63	0.045	0.221
P01023	A2M	7.56	9.20	8.49	8.49	0.002	0.831
P01024	C3	7.48	8.75	8.41	8.61	0.003	0.621
P02679	FGG	6.63	8.83	8.16	8.80	0.003	0.224
P02768	ALB	8.61	9.60	10.44	10.36	0.038	0.909
P02787	TF	6.61	8.12	8.99	9.04	0.005	0.565
P04114	APOB	6.70	8.95	6.85	7.21	0.022	1.000
P08311	CTSG	6.87	8.96	7.62	8.35	0.000	0.144
P11215	ITGAM	6.67	8.14	5.88	7.44	0.004	0.614
P12724	RNASE3	6.96	8.20	7.84	8.27	0.022	0.385
P13796	LCP1	8.24	9.06	8.78	8.96	0.021	0.190
P38919	EIF4A3	6.51	6.99	7.51	7.50	0.030	1.000
P49913	CAMP	6.59	9.08	5.97	7.86	0.003	0.222
P59998	ARPC4	7.49	8.78	7.69	7.72	0.002	0.513
P61158	ACTR3	6.58	8.33	8.16	8.27	0.035	0.482
P61160	ACTR2	6.54	8.32	7.89	7.98	0.002	1.000
P62714	PPP2CB	6.38	7.45	7.13	7.09	0.021	0.972
P62807	HIST1H2BC	7.87	9.66	9.84	9.88	0.022	0.513
P69905	HBA1	7.26	9.43	9.67	9.54	0.002	0.896
Q14764	MVP	7.43	9.04	8.68	8.71	0.034	0.787
Q60FE5	FLNA	6.35	8.34	9.07	9.06	0.020	0.945
Q99878	HIST1H2AJ	7.14	9.03	9.32	9.34	0.031	0.513

Table 5. 3. Differentially expressed proteins in IBD vs non-IBD controls in EVs (*q* value < 0.05) but not in biopsies (*q* > 1.0).

A total of 27 human proteins displayed differential expression between IBD and non-IBD controls in EVs with concurrent similar levels of expression between IBD and non-IBD controls in intestinal biopsies.

Data further analyzed from Zhang, X., S.A. Deeke, Z.B. Ning, A.E. Starr, J. Butcher, J. Li, J. Mayne, K. Cheng, B. Liao, L.Y. Li, R. Singleton, D. Mack, A. Stintzi, and D. Figeys. 2018a. Metaproteomics reveals associations between microbiome and intestinal extracellular vesicle proteins in pediatric inflammatory bowel disease. *Nature Communications*

This work is licensed under a Creative Commons Attribution 4.0 International License (https://creativecommons.org/licenses/by/4.0/)

CHAPTER 6. DISCUSSION

Herein, the proteomic evaluation of clinical samples, and subsequent enrichment of components thereof, identified proteins with differential expression representing potential biomarkers and provided insight on the pathogenesis of IBD. This study yielded the first proteomic evaluation of pediatric host proteins (Chapters 3 and 4) and EVs (Chapter 5) from the intestinal MLI of IBD and non-IBD control patients. From this study, two distinct biomarker panels were generated; the first to classify IBD from non-IBD patients and the second to delineate pancolitis from non-pancolitis (Chapter 4). A subset of proteins featuring in the IBD biomarker panel, namely CAT and LTA4H, were validated in a non-invasive biological specimen (stool) obtained from an independent cohort, suggesting that the findings can be translated to the clinical setting. Additionally, proteomic analysis of the enrichment of EVs from the intestinal MLI of IBD and non-IBD control patients identified differential protein cargo, including an elevation in proteins related to host defense in IBD.

As an initial analysis, the global proteomic changes from intestinal MLI IBD samples were compared to non-IBD controls. This analysis yielded several immune related biological processes when IBD samples collected from both areas with and without evidence of macroscopic inflammation were included (**Figure 3.4** and **Figure 3.5**). The identification of altered immune related processes in the absence of intestinal macroscopic inflammation is in accordance with another proteomic study (Li et al., 2016) wherein several α-defensins, members of the innate immunity, were elevated in both inactive UC and CD compared to non-IBD controls. In the study by Li et al., 2016 intestinal MLI samples were collected in adults from areas without macroscopic inflammation and compared to non-IBD controls. Whether the

immune related alterations from areas lacking macroscopic inflammation represent the early stages preceding the development of macroscopic inflammation requires further investigation. Interestingly, also among the top biological processes upon the comparison of non-IBD controls and IBD samples collected from areas without inflammation were processes related to viral life cycle (**Figure 3.5**). Interestingly viruses have previously been linked to IBD with differences observed within the enteric virome of IBD patients (Norman et al., 2015). Metagenomic sequencing of stool samples has demonstrated elevated bacteriophage richness in IBD patients compared with non-IBD controls (Norman et al., 2015); while the richness of *Caudovirales* (order of viruses) was increased, no difference was observed in the richness of *Microviridae*. Furthermore, Epstein– Barr virus (EBV) has been more frequently identified in the intestinal mucosa of IBD patients (67%) compared to healthy controls (18%), with similar levels observed between inflamed and non-inflamed areas (Lopes et al., 2017). Whether viruses play a role in the pathogenesis of disease or play a role in the stages prior to the development of macro-inflammation remains unknown and warrants further study.

In comparing the current intestinal MLI dataset with that of a previously published intestinal biopsy dataset (Starr et al., 2017b) the value of evaluating different biological specimen was highlighted. Of the 315 proteins which were deemed as discriminant features in the biopsy dataset a mere 31overlapped with the discriminant features identified at the intestinal MLI (**Table 3.3**). Considering the differences between these biological specimens (e.g. origin of subcellular location) this finding is not unexpected as the majority (187/315; 60%) of proteins identified as discriminant features from the biopsy dataset are not present in the Q75 MLI dataset. Given the distinct subcellular locations under investigation in these two datasets, protein processing can also explain differences observed. Because the exact mechanism leading to IBD

has yet to be elucidated, assessing alterations at multiple levels can aid in shedding light on the etiology of IBD. Future plans aim to combine additional layers of molecular information including the integration of microbial proteomic data with the existing MLI and biopsy datasets.

The proteome evaluation of the intestinal MLI permitted the development of biomarker panels for IBD diagnosis. The four proteins featuring in the active IBD biomarker panel (**Table 4.1**) have functions which are consistent with the pathophysiology of the disease. Both Transketolase and Catalase are increased at the AC and DC in patients with active IBD compared to non-IBD controls (**Figure 4.2**) and have roles in minimizing reactive oxygen species (ROS) damage. Given that ROS are increased in the colonic tissue of IBD patients (Cao et al., 2004; Lih-Brody et al., 1996), the elevated expression of these enzymes may be a response intended to mitigate damage that is associated with the elevated levels of ROS. Transketolase can indirectly drive production of the antioxidant NAD(P)H (Lassen et al., 2008), which protects cellular components by scavenging damage-causing free radicals (Kirsch and De Groot, 2001). Interestingly, a recent study identified TKT as one of eight proteins in salivary exosomes of UC and CD patients that was not identified in their healthy control counterparts (Zheng et al., 2017). Furthermore mice deficient in Transketolase-like protein 1 (TKTL1), which shares 60% percentage identity (NCBI BLAST) with TKT, demonstrated elevated susceptibility to Dextran Sodium Sulfate (DSS)-induced colitis (Bentz et al., 2011). In particular, TKTL1 deficient mice displayed increased mucosal damage with loss of crypt structure and shorter colon lengths as compared to their wild-type (WT) DSS-treated counterparts. Similarly, catalase is an ROS scavenger also featuring in the IBD biomarker panel, specifically protecting cells from the damaging effects of hydrogen peroxide which is elevated in the colonic tissue of UC patients (Cao et al., 2004). Interestingly catalase has been identified as one of the antigens of the

perinuclear anti-neutrophil cytoplasmic autoantibodies (pANCA) (Roozendaal et al., 1998), which has been proposed as a serological biomarker to assist in differentiating IBD subtypes. A meta-analysis encompassing 60 studies determined that a pANCA positive result yielded a sensitivity of 55.3% and a specificity of 88.5% for diagnosis of UC compared to CD (Reese et al., 2006). Due to low sensitivity, this biomarker has not been implemented for common use in the clinic. Herein, IBD patients with and without macroscopic inflammation demonstrated significantly elevated expression of catalase compared to non-IBD controls in both colon sub-regions (**Figure 4.4 A**).

The IBD biomarker panel protein Annexin A3 was previously identified as a potential blood biomarker for colorectal cancer (Marshall et al., 2010). When combined with the expression of six other genes (CLEC4D, LMNB1, PRRG4, TNFAIP6, VNN1 and IL2RB), the blood biomarker panel which features Annexin A3 yields a ROC AUC of 0.80 in diagnosing colorectal cancer. Furthermore, the expression of Annexin A3 is elevated in colorectal cancer tissue (Bai et al., 2017). Most Annexin A3 studies have investigated its' role in carcinogenesis with altered expression of Annexin A3 reported in tumour development and metastasis for several cancers types (Du et al., 2018; Kollermann et al., 2008; Wu et al., 2013; Yan et al., 2010). Interestingly, UC patients are at an elevated risk of colorectal cancer, a risk that increases with the duration of disease. Moreover, a recent study provided evidence that Annexin A3's effect on cancer cell invasion and proliferation is modulated by the NF-kB pathway. In this study, Annexin A3 knockdown lead to the inhibition of the NF-κB pathway, modulated through the upregulation of IκBα (Du et al., 2018). Interestingly NF- κB activation is increased in the colon of IBD patients (Schreiber et al., 1998)**.** Futher investigation is required in order to

evaluate whether the elevated intestinal MLI expression of Annexin A3 in IBD patients is contributing to the observed increase in NF-kB activation.

Leukotriene A4 hydrolase is the only protein to appear in both the IBD diagnosis and UC extent of disease biomarker panels. This enzyme displays both aminopeptidase and hydrolase activity, and catalyzes the last step in leukotriene B_4 (LTB_4) synthesis pathway. LTB_4 is a potent chemoattractant which mediates leukocyte adherence and aggregation. The elevated expression of LTA4H observed in both the AC and DC in IBD compared to non-IBD controls (**Figure 4.2 B**) and its elevated expression at the AC of UC patients with macroscopic inflammation compared to those without inflammation (**Figure 4.7 C**) is consistent with the intestinal neutrophilic infiltration seen in IBD. Furthermore LTB_4 can stimulate the production of ROS (Woo et al., 2000) also consistent with the known elevated levels of ROS in colon tissue of IBD patients (Cao et al., 2004).

Like LKTA4H, VASP also features in the extent of disease (pancolitis vs non-pancolitis) biomarker panel (**Figure 4.7**). VASP is an actin associated protein which plays a role in a variety of processes which require cytoskeleton remodeling including cell migration (Barzik et al., 2005). VASP localization has been reported as altered in CD compared to non-IBD controls and UC patients; by immunohistochemistry VASP was shown to localize to apical tight junctions in non-IBD controls without inflammation (10/10) and in the majority (6/7) of inactive UC patients, whereas the majority (7/10) of inactive CD patients demonstrated basolateral positioning (Ohira et al., 2009). Whether the localization of VASP in patients with active disease is also altered remains undetermined. Interestingly VASP has been implicated in the bacterial invasion and replication of *Coxiella burnetii.* This pathogenic bacterium replicates in a parasitophorous vacuole (PV), a membrane which protects the intracellular pathogen from the host

phagolysosome. VASP is required for optimum PV assembly, and thus enables *C. burnetii* replication (Colonne et al., 2016). Although *C. burnetti* has not been directly involved in UC, most parasites belonging to the phylum Apicomplexa assemble a PV in which to inhabit and replicate (Clough and Frickel, 2017). The expression of VASP was elevated in the AC of UC patients with macroscopic inflammation compared to those lacking macroscopic inflammation (**Figure 4.7**). Future studies could evaluate whether the elevated expression of VASP promotes the expansion of the Apicomplexa phylum and whether its elevated expression contributes to the onset of macroscopic inflammation or is a consequence of inflammation.

The UC extent of disease biomarker panel also includes TXNDC17, a disulfide reductase. TXNDC17 has been shown to affect NF-κB activation, which in turn regulates the expression of inflammatory genes (Jeong et al., 2004). The cytokine TNF-α is a powerful stimulator of the NF-κB transcription factor and anti-TNF- α agents have transformed the therapeutic landscape in IBD treatment. Infliximab is a TNF- α monoclonal antibody utilized for inducing and maintaining remission in pediatric IBD (Hyams et al., 2007). Given that NF- κB activation is increased in the colon of IBD patients (Schreiber et al., 1998) and overexpression of TXNDC17 inhibits TNF-α induced NF-κB activation (Jeong et al., 2004) the elevated expression of TXNDC17 in areas with macroscopic inflammation compared those without macroscopic inflammation (**Figure 4.7**) may serve as a response in an attempt to dampen the NF-κB activation and thus mitigate the expression of inflammatory genes at the affected area. Thymosin beta-10 (TMSB10), also among the features in the pancolitis vs non-pancolitis panel, has a role in cytoskeleton organization. Its expression is increased in many cancers (Zhang et al., 2017; Zhang et al., 2014) although herein we observe a decreased expression in areas of macroscopic inflammation compared to those without macroscopic inflammation (**Figure 4.7**). TMSB10

inhibits the polymerization of actin (Yu et al., 1993) which in turn would affect cell structure alterations. Given that cytoskeletal rearrangement is involved in many immune cell processes including migration to the affected area (Kopecki et al., 2016) and that TMSB10 inhibits actin polymerization, the observed decreased expression in areas with macroscopic inflammation may be to allow cytoskeletal rearrangement required for immune cells which are present at the affected area.

The expression of the two IBD biomarker panel proteins CAT and LTA4H was assessed in stool, to evaluate the potential use of these as biomarkers from a non-invasive biological specimen, thereby narrowing the bench to bedside gap. In addition to recapitulating the elevated expression in IBD compared to non-IBD controls that was observed by MS in the MLI aspirate, LTA4H correlated with the UC activity index (PUCAI). PUCAI is used in monitoring patients for response to therapy (Turner et al., 2018). The approach for treating children with acute severe colitis relies on the PUCAI score to dictate whether after three days a second line therapy is required (Turner et al., 2018). Further patient recruitement will enable the evaluation of the expression in stool of the remaining proteins featuring in the biomarker panel and thus will permit the assessment of the overall biomarker panel performance in stool compared to it's performance at the intestinal MLI.

One of the strengths of this study is the inclusion of control patients which present with symptoms similar to that of individuals with IBD including abdominal pain, diarrhea and weight loss. Our biomarker panel has thus been developed using patients which are suspected of having IBD and thus represent valuable IBD clinical controls. In particular, two patients with IBS were present within the discovery non-IBD control cohort; both IBS patients correctly classified with the biomarker panel as non-IBD controls. IBS, defined as "*a functional bowel disorder in which*

abdominal pain or discomfort is associated with defecation or a change in bowel habit, and with features of disordered defecation" (Longstreth et al., 2006) can present with symptoms similar to that of IBD including abdominal pain and diarrhea (Halpin and Ford, 2012). Recently calprotectin, which is the most routinely used biomarker in the clinic for IBD diagnosis, was shown to be elevated in 36.6% (34/93) of recruited adult IBS participants (Melchior et al., 2017). A positive (elevated) calprotectin result would thus lead to further examination including endoscopy, highlighting the need for accurate biomarkers in IBD diagnosis. Furthermore a study considering 20,193 IBD cases suggested that ~10% of patients with IBD are initially misdiagnosed with IBS prior to their ultimate diagnosis of IBD (Card et al., 2014). Given that the treatment approaches for IBS are different from IBD (Oswiecimska et al., 2017) a misdiagnosis could prevent appropriate therapeutic intervention. The ultimate therapeutic goal in IBD is to achieve intestinal mucosal healing, whereas in IBS the objective is to restore intestinal motility and lessen symptoms (Oswiecimska et al., 2017). Therefore a patient with IBD initially misdiagnosed with IBS may suffer from intestinal mucosal inflammation for a longer duration which in turn could increase the risk of colorectal cancer (Triantafillidis et al., 2009), underscoring the need for improved biomarkers capable of differentiating IBD and IBS. Although testing on additional IBS patients is needed to assess the capacity of the biomarker panel in distinguishing IBD from IBS, initial results are promising.

Our study was however limited by the lack of control patients presenting with infectious colitis, who can also present with symptoms similar to that of IBD, including abdominal pain and diarrhea (Shivashankar and Lichtenstein, 2018). Ongoing recruitment will enable future studies to assess the performance of the biomarker panel in these patients. Furthermore, studies are needed to evaluate the performance of the biomarker panels on pediatric patients from other

institutions, since all recruitment was performed at CHEO, and whether the panel also demonstrates value in the adult population. Finally, due to low ROC values, we did not pursue biomarkers for differentiating IBD subtypes (CD vs UC).

Given the ability of EVs to deliver molecular cargo and modify the response of a recipient cell, the proteomic alterations of EVs were compared between pediatric IBD patients and those from non-IBD controls. This comparison identified proteins associated with biological processes related to immunity and host defense including defense response to bacterium (**Figure 5.3 A**). Interestingly, it has been shown that microvesicles (MVs) derived from neutrophils can participate in the defense against invading bacteria. MVs have been isolated from human serum and shown to exhibit antibacterial effects which are dependent on actin remodeling and Integrin Beta 2 (Timar et al., 2013). Integrin Beta 2 was identified as a discriminant feature by PLS-DA (**Appendix 2, Table 1**) and was significantly elevated ($p=0.0005$) in EVs from IBD patients compared to non-IBD controls. Therefore EVs from IBD patients possibly demonstrate increased ability to participate in antibacterial activities compared to those from non-IBD patients. Furthermore, upon comparing differentially expressed proteins from IBD with non-IBD controls identified in EVs but not in biopsies (**Figure 5.4 B**) several biological processes related to cytoskeleton remodeling were enriched, including actin cytoskeleton organization, Arp2/3 complex-mediated actin nucleation and actin filament organization. Given that antibacterial effects of MVs was dependent upon actin remodeling (Timar et al., 2013), these results thus provided further evidence that intestinal EVs may be participating in antibacterial processes.

In the EV dataset the leukocyte marker PTPRC (CD45) was identified in 100% of EVs samples isolated from IBD patients whereas it was only identified in 43% of non-IBD control patients. This result is consistent with a study wherein microparticles isolated from platelet-free

plasma of adults with active CD displayed an elevated number of microparticles with CD45 expression compared to microparticles isolated from healthy subjects (Leonetti et al., 2013). This elevated CD45 is likely due to the increase in neutrophils in the colon and blood of IBD patients (Bennike et al., 2015; Hanai et al., 2004). Additional proteins identified in the EVs that were previously described in neutrophil-derived MVs include myeloperoxidase, lactoferrin, azurocidin and cathelicidin antimicrobial peptide precursor (Timar et al., 2013), providing further support of an elevated number of neutrophilic derived EVs within the colon of patients with IBD.

MVs isolated from neutrophils have also been shown to have anti-inflammatory effects by mediating the release of TGF-β by macrophages in an annexin-A1 dependent manner (Rhys et al., 2018). Notably, in the intestinal EV dataset, annexin-A1 was identified as a discriminant feature (**Appendix 2, Table 1**) and elevated in IBD EVs compared to control EVs. This observation suggests that although defense responses are among the top biological processes identified in IBD EVs compared to control EVs (**Figure 5.3**), EVs within the intestine of IBD patients may also participate in anti-inflammatory responses.

In addition to evaluating host protein alterations within the EV preparation, differences in microbial proteins from OMVs were assessed. The outer membrane receptor for ferrienterochelin and colicins was identified as a discriminant feature between non-IBD control EVs and IBD EVs (**Table 5.2**). Bacteria have evolved a variety of mechanism by which to acquire iron, an essential bacterial nutrient, including through ferrientrochelin (ferric iron complexed to Enterochelin) uptake. Studies demonstrating the ability of OMVs to uptake and deliver iron to bacteria continue to emerge (Lin et al., 2017; Prados-Rosales et al., 2014). The finding of the outer membrane receptor for ferrientrochelin and colicins in OMVs suggests that iron availability in the intestinal MLI environment differs between non-IBD controls and IBD patients.

The EV study was limited by the number of patients, which was insufficient to enable comparison of the EV proteomic content between IBD subtypes (CD vs UC). Furthermore consistent with the study by Mitsuhashi et al., 2016, EVs were isolated from the intestinal MLI of non-IBD controls and IBD patients by ultracentrifugation and therefore contaminant proteins may be present. Ongoing recruitment will enable future studies to examine proteomic differences between the IBD subtypes and larger clinical sample collection volumes will provide sufficient sample amount to be compatible with alternative EV isolation methods. Notably, immune-based isolation procedures were incompatible with this study as there are no known universal cell surface antigens capable of capturing OMVs released by all bacterial species present within the gut. Moreover, given that proteins identified rely on their presence within the database (matching spectra to database) unknown microbial proteins which may be present in the intestinal MLI would not be identified in this study. Although the quantification accuracy of the EV study could be improved by using a heavy internal standard, it could not be implemented as a representative reference proteome could not be generated for bacterial OMVs. An internal heavy reference would require an isotopically labeled sample which is representative of both the human EVs and the intestinal bacterial OMVs. Due to inter-individual diversity of the intestinal microbiota and the intestinal microbiota's complexity a single reference proteome is unlikely to represent an entire cohort of patients and thus a label-free quantification approach was used.

The enrichment of EVs from the intestinal MLI identified host defense proteins as elevated in IBD patients compared to non-IBD controls. These EVs therefore represent potential contributors of the exaggerated host response which is characteristic of IBD and thus may represent potential therapeutic targets. Regulating aspects of the exosomal pathway has previously been proposed as a possible treatment in IBD (Xu et al., 2016) and herein we provide

further evidence that modulating EVs could mitigate the exaggerated immune response seen in IBD. These differentially expressed EV proteins may also provide utility as IBD biomarkers. Ongoing recruitment leading to an elevated sample size will enable the future evaluation of EVs as potential IBD biomarkers and possibly also as biomarkers of microbiome alterations.

To conclude, herein the first pediatric intestinal MLI proteomic study enabled the identification of biomarker panels for the diagnosis of inflammatory bowel disease and UC extent of disease. The expansion to evaluate select biomarkers identified from the intestinal MLI on a non-invasive biological specimen (stool) by ELISA demonstrated consistent differential protein expression between non-IBD controls and IBD. Further assessment of the intestinal MLI by enrichment of EVs with subsequent proteome assessment identified host defense proteins elevated in EVs isolated from IBD patients compared to non-IBD controls. The proteomic evaluation of clinical samples within this study thus provides part of the initial work required to evolve beyond invasive procedures for the diagnosis of IBD and also aids in the elucidation of pathogenesis of the disease.

REFERENCES

Abraham, C., and Cho, J.H. (2006). Functional consequences of NOD2 (CARD15) mutations. Inflamm Bowel Dis *12*, 641-650.

Agarwal, S., and Mayer, L. (2013). Diagnosis and Treatment of Gastrointestinal Disorders in Patients With Primary Immunodeficiency. Clin Gastroenterol H *11*, 1050-1063.

Akers, J.C., Ramakrishnan, V., Kim, R., Skog, J., Nakano, I., Pingle, S., Kalinina, J., Hua, W., Kesari, S., Mao, Y.*, et al.* (2013). MiR-21 in the extracellular vesicles (EVs) of cerebrospinal fluid (CSF): a platform for glioblastoma biomarker development. PLoS One *8*, e78115.

Ananthakrishnan, A.N., Cagan, A., Gainer, V.S., Cai, T.X., Cheng, S.C., Savova, G., Chen, P., Szolovits, P., Xia, Z.Q., De Jager, P.L.*, et al.* (2013). Normalization of Plasma 25-Hydroxy Vitamin D Is Associated with Reduced Risk of Surgery in Crohn's Disease. Inflamm Bowel Dis *19*, 1921-1927.

Ananthakrishnan, A.N., Khalili, H., Higuchi, L.M., Bao, Y., Korzenik, J.R., Giovannucci, E.L., Richter, J.M., Fuchs, C.S., and Chan, A.T. (2012). Higher Predicted Vitamin D Status Is Associated With Reduced Risk of Crohn's Disease. Gastroenterology *142*, 482-489.

Ananthakrishnan, A.N., Khalili, H., Konijeti, G.G., Higuchi, L.M., de Silva, P., Fuchs, C.S., Willett, W.C., Richter, J.M., and Chan, A.T. (2014). Long-term intake of dietary fat and risk of ulcerative colitis and Crohn's disease. Gut *63*, 776-784.

Ardesia, M., Ferlazzo, G., and Fries, W. (2015). Vitamin D and inflammatory bowel disease. Biomed Res Int *2015*, 470805.

Ashton, J.J., Harden, A., and Beattie, R.M. (2017). Paediatric inflammatory bowel disease: improving early diagnosis. Arch Dis Child.

Bai, Z., Wang, J., Wang, T., Li, Y., Zhao, X., Wu, G., Yang, Y., Deng, W., and Zhang, Z. (2017). The MiR-495/Annexin A3/P53 Axis Inhibits the Invasion and EMT of Colorectal Cancer Cells. Cell Physiol Biochem *44*, 1882-1895.

Baietti, M.F., Zhang, Z., Mortier, E., Melchior, A., Degeest, G., Geeraerts, A., Ivarsson, Y., Depoortere, F., Coomans, C., Vermeiren, E.*, et al.* (2012). Syndecan-syntenin-ALIX regulates the biogenesis of exosomes. Nat Cell Biol *14*, 677-685.

Balamtekin, N., Baysoy, G., Uslu, N., Orhan, D., Akcoren, Z., Ozen, H., Gurakan, F., Saltik-Temizel, I.N., and Yuce, A. (2012). Fecal calprotectin concentration is increased in children with celiac disease: relation with histopathological findings. Turk J Gastroenterol *23*, 503-508.

Barrett, J.C., Hansoul, S., Nicolae, D.L., Cho, J.H., Duerr, R.H., Rioux, J.D., Brant, S.R., Silverberg, M.S., Taylor, K.D., Barmada, M.M.*, et al.* (2008). Genome-wide association defines more than 30 distinct susceptibility loci for Crohn's disease. Nature Genetics *40*, 955-962.

Barzik, M., Kotova, T.I., Higgs, H.N., Hazelwood, L., Hanein, D., Gertler, F.B., and Schafer, D.A. (2005). Ena/VASP proteins enhance actin polymerization in the presence of barbed end capping proteins. J Biol Chem *280*, 28653-28662.

Benchimol, E.I., Bernstein, C.N., Bitton, A., Carroll, M.W., Singh, H., Otley, A.R., Vutcovici, M., El-Matary, W., Nguyen, G.C., Griffiths, A.M.*, et al.* (2017). Trends in Epidemiology of Pediatric Inflammatory Bowel Disease in Canada: Distributed Network Analysis of Multiple Population-Based Provincial Health Administrative Databases. Am J Gastroenterol *112*, 1120-1134.

Bennike, T.B., Carlsen, T.G., Ellingsen, T., Bonderup, O.K., Glerup, H., Bogsted, M., Christiansen, G., Birkelund, S., Stensballe, A., and Andersen, V. (2015). Neutrophil Extracellular Traps in Ulcerative Colitis: A Proteome Analysis of Intestinal Biopsies. Inflamm Bowel Dis *21*, 2052-2067.

Bentz, S., Pesch, T., Wolfram, L., de Valliere, C., Leucht, K., Fried, M., Coy, J.F., Hausmann, M., and Rogler, G. (2011). Lack of transketolase-like (TKTL) 1 aggravates murine experimental colitis. Am J Physiol Gastrointest Liver Physiol *300*, G598-607.

Bernstein, C.N., Blanchard, J.F., Rawsthorne, P., and Yu, N. (2001). The prevalence of extraintestinal diseases in inflammatory bowel disease: a population-based study. Am J Gastroenterol *96*, 1116-1122.

Borg, J., Campos, A., Diema, C., Omenaca, N., de Oliveira, E., Guinovart, J., and Vilaseca, M. (2011). Spectral counting assessment of protein dynamic range in cerebrospinal fluid following depletion with plasma-designed immunoaffinity columns. Clin Proteomics *8*, 6.

Borm, M.E., van Bodegraven, A.A., Mulder, C.J., Kraal, G., and Bouma, G. (2008). The effect of NOD2 activation on TLR2-mediated cytokine responses is dependent on activation dose and NOD2 genotype. Genes Immun *9*, 274-278.

Canas, M.A., Gimenez, R., Fabrega, M.J., Toloza, L., Baldoma, L., and Badia, J. (2016). Outer Membrane Vesicles from the Probiotic Escherichia coli Nissle 1917 and the Commensal ECOR12 Enter Intestinal Epithelial Cells via Clathrin-Dependent Endocytosis and Elicit Differential Effects on DNA Damage. PLoS One *11*, e0160374.

Cao, W., Vrees, M.D., Kirber, M.T., Fiocchi, C., and Pricolo, V.E. (2004). Hydrogen peroxide contributes to motor dysfunction in ulcerative colitis. Am J Physiol Gastrointest Liver Physiol *286*, G833-843.

Card, T.R., Siffledeen, J., and Fleming, K.M. (2014). Are IBD patients more likely to have a prior diagnosis of irritable bowel syndrome? Report of a case-control study in the General Practice Research Database. United Eur Gastroent *2*, 505-512.

Chehoud, C., Albenberg, L.G., Judge, C., Hoffmann, C., Grunberg, S., Bittinger, K., Baldassano, R.N., Lewis, J.D., Bushman, F.D., and Wu, G.D. (2015). Fungal Signature in the Gut Microbiota of Pediatric Patients With Inflammatory Bowel Disease. Inflamm Bowel Dis *21*, 1948-1956.

Chelakkot, C., Choi, Y., Kim, D.K., Park, H.T., Ghim, J., Kwon, Y., Jeon, J., Kim, M.S., Jee, Y.K., Gho, Y.S.*, et al.* (2018). Akkermansia muciniphila-derived extracellular vesicles influence gut permeability through the regulation of tight junctions. Exp Mol Med *50*, e450.

Chen, S.J., Liu, X.W., Liu, J.P., Yang, X.Y., and Lu, F.G. (2014). Ulcerative colitis as a polymicrobial infection characterized by sustained broken mucus barrier. World J Gastroentero *20*, 9468-9475.

Ciofu, O., Beveridge, T.J., Kadurugamuwa, J., Walther-Rasmussen, J., and Hoiby, N. (2000). Chromosomal beta-lactamase is packaged into membrane vesicles and secreted from Pseudomonas aeruginosa. J Antimicrob Chemother *45*, 9-13.

Cleveland, N.K., Rubin, D.T., Hart, J., Weber, C.R., Meckel, K., Tran, A.L., Aelvoet, A.S., Pan, I., Gonsalves, A., Gaetano, J.N.*, et al.* (2018). Patients With Ulcerative Colitis and Primary Sclerosing Cholangitis Frequently Have Subclinical Inflammation in the Proximal Colon. Clin Gastroenterol H *16*, 68-74.

Clough, B., and Frickel, E.M. (2017). The Toxoplasma Parasitophorous Vacuole: An Evolving Host-Parasite Frontier. Trends Parasitol *33*, 473-488.

Colonne, P.M., Winchell, C.G., Graham, J.G., Onyilagha, F.I., MacDonald, L.J., Doeppler, H.R., Storz, P., Kurten, R.C., Beare, P.A., Heinzen, R.A.*, et al.* (2016). Vasodilator-Stimulated Phosphoprotein Activity Is Required for Coxiella burnetii Growth in Human Macrophages. PLoS Pathog *12*, e1005915.

Conrad, M.A., and Rosh, J.R. (2017). Pediatric Inflammatory Bowel Disease. Pediatr Clin North Am *64*, 577-591.

Conroy, S., Hale, M.F., Cross, S.S., Swallow, K., Sidhu, R.H., Sargur, R., and Lobo, A.J. (2017). Unrestricted faecal calprotectin testing performs poorly in the diagnosis of inflammatory bowel disease in patients in primary care. J Clin Pathol.

Cox, J., Matic, I., Hilger, M., Nagaraj, N., Selbach, M., Olsen, J.V., and Mann, M. (2009). A practical guide to the MaxQuant computational platform for SILAC-based quantitative proteomics. Nat Protoc *4*, 698-705.

D'Haens, G., Sandborn, W.J., Feagan, B.G., Geboes, K., Hanauer, S.B., Irvine, E.J., Lemann, M., Marteau, P., Rutgeerts, P., Scholmerich, J.*, et al.* (2007). A review of activity indices and efficacy end points for clinical trials of medical therapy in adults with ulcerative colitis. Gastroenterology *132*, 763-786.

Dabritz, J., Gerner, P., Enninger, A., Classen, M., and Radke, M. (2017). Inflammatory Bowel Disease in Childhood and Adolescence. Dtsch Arztebl Int *114*, 331-+.

Darfeuille-Michaud, A., Boudeau, J., Bulois, P., Neut, C., Glasser, A.L., Barnich, N., Bringer, M.A., Swidsinski, A., Beaugerie, L., and Colombel, J.F. (2004). High prevalence of adherent-invasive Escherichia coli associated with ileal mucosa in Crohn's disease. Gastroenterology *127*, 412-421.

Das, P., Goswami, P., Das, T.K., Nag, T., Sreenivas, V., Ahuja, V., Panda, S.K., Gupta, S.D., and Makharia, G.K. (2012). Comparative tight junction protein expressions in colonic Crohn's disease, ulcerative colitis, and tuberculosis: a new perspective. Virchows Arch *460*, 261-270.

De Palma, G., Sallustio, F., and Schena, F.P. (2016). Clinical Application of Human Urinary Extracellular Vesicles in Kidney and Urologic Diseases. Int J Mol Sci *17*.

Demetter, P., De Vos, M., Van Huysse, J.A., Baeten, D., Ferdinande, L., Peeters, H., Mielants, H., Veys, E.M., De Keyser, F., and Cuvelier, C.A. (2005). Colon mucosa of patients both with spondyloarthritis and

Crohn's disease is enriched with macrophages expressing the scavenger receptor CD163. Ann Rheum Dis *64*, 321-324.

Dethlefsen, L., Huse, S., Sogin, M.L., and Relman, D.A. (2008). The pervasive effects of an antibiotic on the human gut microbiota, as revealed by deep 16S rRNA sequencing. PLoS Biol *6*, e280.

Dignass, A., Eliakim, R., Magro, F., Maaser, C., Chowers, Y., Geboes, K., Mantzaris, G., Reinisch, W., Colombel, J.F., Vermeire, S.*, et al.* (2012). Second European evidence-based consensus on the diagnosis and management of ulcerative colitis part 1: definitions and diagnosis. J Crohns Colitis *6*, 965-990.

Dong, L., Lin, W., Qi, P., Xu, M.D., Wu, X., Ni, S., Huang, D., Weng, W.W., Tan, C., Sheng, W.*, et al.* (2016). Circulating Long RNAs in Serum Extracellular Vesicles: Their Characterization and Potential Application as Biomarkers for Diagnosis of Colorectal Cancer. Cancer Epidemiol Biomarkers Prev *25*, 1158-1166.

Du, R., Liu, B., Zhou, L., Wang, D., He, X., Xu, X., Zhang, L., Niu, C., and Liu, S. (2018). Downregulation of annexin A3 inhibits tumor metastasis and decreases drug resistance in breast cancer. Cell Death Dis *9*, 126.

Durban, A., Abellan, J.J., Jimenez-Hernandez, N., Ponce, M., Ponce, J., Sala, T., D'Auria, G., Latorre, A., and Moya, A. (2011). Assessing gut microbial diversity from feces and rectal mucosa. Microb Ecol *61*, 123-133.

Eaden, J.A., Abrams, K.R., and Mayberry, J.F. (2001). The risk of colorectal cancer in ulcerative colitis: a meta-analysis. Gut *48*, 526-535.

Eaden, J.A., Mayberry, J.F., British Society for, G., Association of Coloproctology for Great, B., and Ireland (2002). Guidelines for screening and surveillance of asymptomatic colorectal cancer in patients with inflammatory bowel disease. Gut *51 Suppl 5*, V10-12.

Elmore, S. (2007). Apoptosis: a review of programmed cell death. Toxicol Pathol *35*, 495-516.

Fiocca, R., Necchi, V., Sommi, P., Ricci, V., Telford, J., Cover, T.L., and Solcia, E. (1999). Release of Helicobacter pylori vacuolating cytotoxin by both a specific secretion pathway and budding of outer membrane vesicles. Uptake of released toxin and vesicles by gastric epithelium. J Pathol *188*, 220-226.

Frank, D.N., Robertson, C.E., Hamm, C.M., Kpadeh, Z., Zhang, T., Chen, H., Zhu, W., Sartor, R.B., Boedeker, E.C., Harpaz, N.*, et al.* (2011). Disease phenotype and genotype are associated with shifts in intestinal-associated microbiota in inflammatory bowel diseases. Inflamm Bowel Dis *17*, 179-184.

Gallo, A., Vallone, C., Sabatelli, L., Ventura, G., Covino, M., Cammarota, G., Gasbarrini, A., Landolfi, R., and Montalto, M. (2017). Fecal calprotectin in management of Clostridium difficile infection: a longitudinal study. Scand J Gastroenterol, 1-6.

Garg, M., Leach, S.T., Coffey, M.J., Katz, T., Strachan, R., Pang, T., Needham, B., Lui, K., Ali, F., Day, A.S.*, et al.* (2017). Age-dependent variation of fecal calprotectin in cystic fibrosis and healthy children. J Cyst Fibros *16*, 631-636.

Geiger, T., Cox, J., Ostasiewicz, P., Wisniewski, J.R., and Mann, M. (2010). Super-SILAC mix for quantitative proteomics of human tumor tissue. Nat Methods *7*, 383-385.

Gevers, D., Kugathasan, S., Denson, L.A., Vazquez-Baeza, Y., Van Treuren, W., Ren, B., Schwager, E., Knights, D., Song, S.J., Yassour, M.*, et al.* (2014). The treatment-naive microbiome in new-onset Crohn's disease. Cell Host Microbe *15*, 382-392.

Ghosh, S., and Mitchell, R. (2007). Impact of inflammatory bowel disease on quality of life: Results of the European Federation of Crohn's and Ulcerative Colitis Associations (EFCCA) patient survey. J Crohns Colitis *1*, 10-20.

Gubatan, J., Mitsuhashi, S., Zenlea, T., Rosenberg, L., Robson, S., and Moss, A.C. (2017). Low Serum Vitamin D During Remission Increases Risk of Clinical Relapse in Patients With Ulcerative Colitis. Clin Gastroenterol H *15*, 240-+.

Gyorgy, B., Hung, M.E., Breakefield, X.O., and Leonard, J.N. (2015). Therapeutic applications of extracellular vesicles: clinical promise and open questions. Annu Rev Pharmacol Toxicol *55*, 439-464.

Halme, L., Turunen, U., Helio, T., Paavola, P., Walle, T., Miettinen, A., Jarvinen, H., Kontula, K., and Farkkila, M. (2002). Familial and sporadic inflammatory bowel disease: comparison of clinical features and serological markers in a genetically homogeneous population. Scand J Gastroenterol *37*, 692-698.

Halpin, S.J., and Ford, A.C. (2012). Prevalence of Symptoms Meeting Criteria for Irritable Bowel Syndrome in Inflammatory Bowel Disease: Systematic Review and Meta-Analysis. American Journal of Gastroenterology *107*, 1474-1482.

Han, N.Y., Choi, W., Park, J.M., Kim, E.H., Lee, H., and Hahm, K.B. (2013). Label-free quantification for discovering novel biomarkers in the diagnosis and assessment of disease activity in inflammatory bowel disease. J Dig Dis *14*, 166-174.

Hanai, H., Takeuchi, K., Iida, T., Kashiwagi, N., Saniabadi, A.R., Matsushita, I., Sato, Y., Kasuga, N., and Nakamura, T. (2004). Relationship between fecal calprotectin, intestinal inflammation, and peripheral blood neutrophils in patients with active ulcerative colitis. Digest Dis Sci *49*, 1438-1443.

Harris, P.A., Taylor, R., Thielke, R., Payne, J., Gonzalez, N., and Conde, J.G. (2009). Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform *42*, 377-381.

Henderson, P., Anderson, N.H., and Wilson, D.C. (2014). The diagnostic accuracy of fecal calprotectin during the investigation of suspected pediatric inflammatory bowel disease: a systematic review and meta-analysis. Am J Gastroenterol *109*, 637-645.

Hessian, P.A., Edgeworth, J., and Hogg, N. (1993). MRP-8 and MRP-14, two abundant Ca(2+)-binding proteins of neutrophils and monocytes. J Leukoc Biol *53*, 197-204.

Hodgson-Parnell, L., Spence, O., and Chapple, K. (2018). Solitary juvenile polyp as a cause of elevated faecal calprotectin in an adult. BMJ Case Rep *2018*.

Hou, J.K., Abraham, B., and El-Serag, H. (2011). Dietary Intake and Risk of Developing Inflammatory Bowel Disease: A Systematic Review of the Literature. American Journal of Gastroenterology *106*, 563-573.

Hviid, A., Svanstrom, H., and Frisch, M. (2011). Antibiotic use and inflammatory bowel diseases in childhood. Gut *60*, 49-54.

Hyams, J., Crandall, W., Kugathasan, S., Griffiths, A., Olson, A., Johanns, J., Liu, G., Travers, S., Heuschkel, R., Markowitz, J*., et al.* (2007). Induction and maintenance infliximab therapy for the treatment of moderate-to-severe Crohn's disease in children. Gastroenterology *132*, 863-873.

Hyams, J., Markowitz, J., Otley, A., Rosh, J., Mack, D., Bousvaros, A., Kugathasan, S., Pfefferkorn, M., Tolia, V., Evans, J*., et al.* (2005). Evaluation of the pediatric crohn disease activity index: a prospective multicenter experience. J Pediatr Gastroenterol Nutr *41*, 416-421.

Ihara, T., Yamamoto, T., Sugamata, M., Okumura, H., and Ueno, Y. (1998). The process of ultrastructural changes from nuclei to apoptotic body. Virchows Arch *433*, 443-447.

Ina, K., Kusugami, K., Yamaguchi, T., Imada, A., Hosokawa, T., Ohsuga, M., Shinoda, M., Ando, T., Ito, K., and Yokoyama, Y. (1997). Mucosal interleukin-8 is involved in neutrophil migration and binding to extracellular matrix in inflammatory bowel disease. American Journal of Gastroenterology *92*, 1342-1346.

Inan, M.S., Rasoulpour, R.J., Yin, L., Hubbard, A.K., Rosenberg, D.W., and Giardina, C. (2000). The luminal short-chain fatty acid butyrate modulates NF-kappa B activity in a human colonic epithelial cell line. Gastroenterology *118*, 724-734.

Jagtap, P., Goslinga, J., Kooren, J.A., McGowan, T., Wroblewski, M.S., Seymour, S.L., and Griffin, T.J. (2013). A two-step database search method improves sensitivity in peptide sequence matches for metaproteomics and proteogenomics studies. Proteomics *13*, 1352-1357.

Jantchou, P., Clavel-Chapelon, F., Racine, A., Kvaskoff, M., Carbonnel, F., and Boutron-Ruault, M.C. (2014). High Residential Sun Exposure Is Associated With a Low Risk of Incident Crohn's Disease in the Prospective E3N Cohort. Inflamm Bowel Dis *20*, 75-81.

Jeffery, L.E., Burke, F., Mura, M., Zheng, Y., Qureshi, O.S., Hewison, M., Walker, L.S.K., Lammas, D.A., Raza, K., and Sansom, D.M. (2009). 1,25-Dihydroxyvitamin D-3 and IL-2 Combine to Inhibit T Cell Production of Inflammatory Cytokines and Promote Development of Regulatory T Cells Expressing CTLA-4 and FoxP3. J Immunol *183*, 5458-5467.

Jeong, W., Chang, T.S., Boja, E.S., Fales, H.M., and Rhee, S.G. (2004). Roles of TRP14, a thioredoxin-related protein in tumor necrosis factor-alpha signaling pathways. J Biol Chem *279*, 3151-3159.

Jess, T., Frisch, M., and Simonsen, J. (2013). Trends in Overall and Cause-Specific Mortality Among Patients With Inflammatory Bowel Disease From 1982 to 2010. Clin Gastroenterol H *11*, 43-48.

Jimenez-Rivera, C., Haas, D., Boland, M., Barkey, J.L., and Mack, D.R. (2009). Comparison of two common outpatient preparations for colonoscopy in children and youth. Gastroenterol Res Pract *2009*, 518932.

Jostins, L., Ripke, S., Weersma, R.K., Duerr, R.H., McGovern, D.P., Hui, K.Y., Lee, J.C., Schumm, L.P., Sharma, Y., Anderson, C.A*., et al.* (2012). Host-microbe interactions have shaped the genetic architecture of inflammatory bowel disease. Nature *491*, 119-124.

Jussila, A., Virta, L.J., Pukkala, E., and Farkkila, M.A. (2014). Mortality and causes of death in patients with inflammatory bowel disease: a nationwide register study in Finland. J Crohns Colitis *8*, 1088-1096.

Kalra, H., Drummen, G.P., and Mathivanan, S. (2016). Focus on Extracellular Vesicles: Introducing the Next Small Big Thing. Int J Mol Sci *17*, 170.

Kaparakis-Liaskos, M., and Ferrero, R.L. (2015). Immune modulation by bacterial outer membrane vesicles. Nat Rev Immunol *15*, 375-387.

Kaplan, G.G., and Ng, S.C. (2017). Understanding and Preventing the Global Increase of Inflammatory Bowel Disease. Gastroenterology *152*, 313-321 e312.

Karlsson, O., Rodosthenous, R.S., Jara, C., Brennan, K.J., Wright, R.O., Baccarelli, A.A., and Wright, R.J. (2016). Detection of long non-coding RNAs in human breastmilk extracellular vesicles: Implications for early child development. Epigenetics, 0.

Keerthikumar, S., Chisanga, D., Ariyaratne, D., Al Saffar, H., Anand, S., Zhao, K., Samuel, M., Pathan, M., Jois, M., Chilamkurti, N.*, et al.* (2016). ExoCarta: A Web-Based Compendium of Exosomal Cargo. J Mol Biol *428*, 688-692.

Kesty, N.C., Mason, K.M., Reedy, M., Miller, S.E., and Kuehn, M.J. (2004). Enterotoxigenic Escherichia coli vesicles target toxin delivery into mammalian cells. Embo J *23*, 4538-4549.

Kirsch, M., and De Groot, H. (2001). NAD(P)H, a directly operating antioxidant? Faseb J *15*, 1569-1574.

Kleinerman, R.A. (2006). Cancer risks following diagnostic and therapeutic radiation exposure in children. Pediatr Radiol *36 Suppl 2*, 121-125.

Klimentova, J., and Stulik, J. (2015). Methods of isolation and purification of outer membrane vesicles from gram-negative bacteria. Microbiol Res *170*, 1-9.

Klingberg, E., Carlsten, H., Hilme, E., Hedberg, M., and Forsblad-d'Elia, H. (2012). Calprotectin in ankylosing spondylitis--frequently elevated in feces, but normal in serum. Scand J Gastroenterol *47*, 435-444.

Kollermann, J., Schlomm, T., Bang, H., Schwall, G.P., von Eichel-Streiber, C., Simon, R., Schostak, M., Huland, H., Berg, W., Sauter, G.*, et al.* (2008). Expression and prognostic relevance of annexin A3 in prostate cancer. Eur Urol *54*, 1314-1323.

Kopecki, Z., Ludwig, R.J., and Cowin, A.J. (2016). Cytoskeletal Regulation of Inflammation and Its Impact on Skin Blistering Disease Epidermolysis Bullosa Acquisita. Int J Mol Sci *17*.

Kowal, J., Arras, G., Colombo, M., Jouve, M., Morath, J.P., Primdal-Bengtson, B., Dingli, F., Loew, D., Tkach, M., and Thery, C. (2016). Proteomic comparison defines novel markers to characterize heterogeneous populations of extracellular vesicle subtypes. Proc Natl Acad Sci U S A *113*, E968-977.

Kuenzig, M.E., Lee, S.M., Eksteen, B., Seow, C.H., Barnabe, C., Panaccione, R., and Kaplan, G.G. (2016). Smoking influences the need for surgery in patients with the inflammatory bowel diseases: a systematic review and meta-analysis incorporating disease duration. BMC Gastroenterol *16*, 143.

Lassen, N., Black, W.J., Estey, T., and Vasiliou, V. (2008). The role of corneal crystallins in the cellular defense mechanisms against oxidative stress. Semin Cell Dev Biol *19*, 100-112.

Lawley, T.D., and Walker, A.W. (2013). Intestinal colonization resistance. Immunology *138*, 1-11.

Lawrance, I.C. (2011). Topical agents for idiopathic distal colitis and proctitis. J Gastroenterol Hepatol *26*, 36-43.

Lehmann, F.S., Trapani, F., Fueglistaler, I., Terracciano, L.M., von Flue, M., Cathomas, G., Zettl, A., Benkert, P., Oertli, D., and Beglinger, C. (2014). Clinical and histopathological correlations of fecal calprotectin release in colorectal carcinoma. World J Gastroenterol *20*, 4994-4999.

Leonetti, D., Reimund, J.M., Tesse, A., Viennot, S., Martinez, M.C., Bretagne, A.L., and Andriantsitohaina, R. (2013). Circulating microparticles from Crohn's disease patients cause endothelial and vascular dysfunctions. PLoS One *8*, e73088.

Leoni, G., Neumann, P.A., Kamaly, N., Quiros, M., Nishio, H., Jones, H.R., Sumagin, R., Hilgarth, R.S., Alam, A., Fredman, G.*, et al.* (2015). Annexin A1-containing extracellular vesicles and polymeric nanoparticles promote epithelial wound repair. J Clin Invest *125*, 1215-1227.

Levine, A., Griffiths, A., Markowitz, J., Wilson, D.C., Turner, D., Russell, R.K., Fell, J., Ruemmele, F.M., Walters, T., Sherlock, M.*, et al.* (2011). Pediatric modification of the Montreal classification for inflammatory bowel disease: the Paris classification. Inflamm Bowel Dis *17*, 1314-1321.

Levine, A., Koletzko, S., Turner, D., Escher, J.C., Cucchiara, S., de Ridder, L., Kolho, K.L., Veres, G., Russell, R.K., Paerregaard, A.*, et al.* (2014). ESPGHAN revised porto criteria for the diagnosis of inflammatory bowel disease in children and adolescents. J Pediatr Gastroenterol Nutr *58*, 795-806.

Lewis, J.D. (2016). The Role of Diet in Inflammatory Bowel Disease. Gastroenterol Hepatol (N Y) *12*, 51-53.

Lewis, J.D., Chen, E.Z., Baldassano, R.N., Otley, A.R., Griffiths, A.M., Lee, D., Bittinger, K., Bailey, A., Friedman, E.S., Hoffmann, C.*, et al.* (2015). Inflammation, Antibiotics, and Diet as Environmental Stressors of the Gut Microbiome in Pediatric Crohn's Disease. Cell Host & Microbe *18*, 489-500.

Lewis, J.D., Chuai, S., Nessel, L., Lichtenstein, G.R., Aberra, F.N., and Ellenberg, J.H. (2008). Use of the Noninvasive Components of the Mayo Score to Assess Clinical Response in Ulcerative Colitis. Inflamm Bowel Dis *14*, 1660-1666.

Li, L., Masica, D., Ishida, M., Tomuleasa, C., Umegaki, S., Kalloo, A.N., Georgiades, C., Singh, V.K., Khashab, M., Amateau, S.*, et al.* (2014). Human bile contains microRNA-laden extracellular vesicles that can be used for cholangiocarcinoma diagnosis. Hepatology *60*, 896-907.

Li, X., LeBlanc, J., Elashoff, D., McHardy, I., Tong, M., Roth, B., Ippoliti, A., Barron, G., McGovern, D., McDonald, K.*, et al.* (2016). Microgeographic Proteomic Networks of the Human Colonic Mucosa and Their Association With Inflammatory Bowel Disease. Cell Mol Gastroenterol Hepatol *2*, 567-583.

Lih-Brody, L., Powell, S.R., Collier, K.P., Reddy, G.M., Cerchia, R., Kahn, E., Weissman, G.S., Katz, S., Floyd, R.A., McKinley, M.J*., et al.* (1996). Increased oxidative stress and decreased antioxidant defenses in mucosa of inflammatory bowel disease. Dig Dis Sci *41*, 2078-2086.

Lin, J., Zhang, W., Cheng, J., Yang, X., Zhu, K., Wang, Y., Wei, G., Qian, P.Y., Luo, Z.Q., and Shen, X. (2017). A Pseudomonas T6SS effector recruits PQS-containing outer membrane vesicles for iron acquisition. Nat Commun *8*, 14888.

Liu, J.Z., van Sommeren, S., Huang, H., Ng, S.C., Alberts, R., Takahashi, A., Ripke, S., Lee, J.C., Jostins, L., Shah, T*., et al.* (2015). Association analyses identify 38 susceptibility loci for inflammatory bowel disease and highlight shared genetic risk across populations. Nat Genet *47*, 979-986.

Longstreth, G.F., Thompson, W.G., Chey, W.D., Houghton, L.A., Mearin, F., and Spiller, R.C. (2006). Functional bowel disorders. Gastroenterology *130*, 1480-1491.

Lopes, S., Andrade, P., Conde, S., Liberal, R., Dias, C.C., Fernandes, S., Pinheiro, J., Simoes, J.S., Carneiro, F., Magro, F*., et al.* (2017). Looking into Enteric Virome in Patients with IBD: Defining Guilty or Innocence? Inflamm Bowel Dis *23*, 278-284.

Lopez-Posadas, R., Sturzl, M., Atreya, I., Neurath, M.F., and Britzen-Laurent, N. (2017). Interplay of GTPases and Cytoskeleton in Cellular Barrier Defects during Gut Inflammation. Front Immunol *8*, 1240.

Lusitani, D., Malawista, S.E., and Montgomery, R.R. (2003). Calprotectin, an abundant cytosolic protein from human polymorphonuclear leukocytes, inhibits the growth of Borrelia burgdorferi. Infect Immun *71*, 4711-4716.

Ma, T.Y., Boivin, M.A., Ye, D., Pedram, A., and Said, H.M. (2005). Mechanism of TNF-{alpha} modulation of Caco-2 intestinal epithelial tight junction barrier: role of myosin light-chain kinase protein expression. Am J Physiol Gastrointest Liver Physiol *288*, G422-430.

Mahid, S.S., Minor, K.S., Soto, R.E., Hornung, C.A., and Galandiuk, S. (2006). Smoking and inflammatory bowel disease: a meta-analysis. Mayo Clin Proc *81*, 1462-1471.

Manichanh, C., Rigottier-Gois, L., Bonnaud, E., Gloux, K., Pelletier, E., Frangeul, L., Nalin, R., Jarrin, C., Chardon, P., Marteau, P*., et al.* (2006). Reduced diversity of faecal microbiota in Crohn's disease revealed by a metagenomic approach. Gut *55*, 205-211.

Mankertz, J., Amasheh, M., Krug, S.M., Fromm, A., Amasheh, S., Hillenbrand, B., Tavalali, S., Fromm, M., and Schulzke, J.D. (2009). TNFalpha up-regulates claudin-2 expression in epithelial HT-29/B6 cells via phosphatidylinositol-3-kinase signaling. Cell Tissue Res *336*, 67-77.

Manning, A.J., and Kuehn, M.J. (2011). Contribution of bacterial outer membrane vesicles to innate bacterial defense. Bmc Microbiol *11*, 258.

Mao, F., Wu, Y., Tang, X., Kang, J., Zhang, B., Yan, Y., Qian, H., Zhang, X., and Xu, W. (2017). Exosomes Derived from Human Umbilical Cord Mesenchymal Stem Cells Relieve Inflammatory Bowel Disease in Mice. Biomed Res Int *2017*, 5356760.

Marshall, K.W., Mohr, S., Khettabi, F.E., Nossova, N., Chao, S., Bao, W., Ma, J., Li, X.J., and Liew, C.C. (2010). A blood-based biomarker panel for stratifying current risk for colorectal cancer. Int J Cancer *126*, 1177-1186.

Matsuoka, K., and Kanai, T. (2015). The gut microbiota and inflammatory bowel disease. Semin Immunopathol *37*, 47-55.

McHardy, I.H., Goudarzi, M., Tong, M., Ruegger, P.M., Schwager, E., Weger, J.R., Graeber, T.G., Sonnenburg, J.L., Horvath, S., Huttenhower, C.*, et al.* (2013). Integrative analysis of the microbiome and metabolome of the human intestinal mucosal surface reveals exquisite inter-relationships. Microbiome *1*, 17.

McMahon, K.J., Castelli, M.E., Vescovi, E.G., and Feldman, M.F. (2012). Biogenesis of Outer Membrane Vesicles in Serratia marcescens Is Thermoregulated and Can Be Induced by Activation of the Rcs Phosphorelay System. J Bacteriol *194*, 3241-3249.

Melchior, C., Aziz, M., Aubry, T., Gourceroi, G., Quillard, M., Zalar, A., Coeffier, M., Dechelotte, P., Leroi, A.M., and Ducrotte, P. (2017). Does calprotectin level identify a subgroup among patients suffering from irritable bowel syndrome? Results of a prospective study. United Eur Gastroent *5*, 261-269.

Meling, T.R., Aabakken, L., Roseth, A., and Osnes, M. (1996). Faecal calprotectin shedding after short-term treatment with non-steroidal anti-inflammatory drugs. Scand J Gastroenterol *31*, 339-344.

Mendoza, J.L., and Abreu, M.T. (2009). Biological markers in inflammatory bowel disease: practical consideration for clinicians. Gastroenterol Clin Biol *33 Suppl 3*, S158-173.

Meucci, G., Bortoli, A., Riccioli, F.A., Girelli, C.M., Radaelli, F., Rivolta, R., and Tatarella, M. (1999). Frequency and clinical evolution of indeterminate colitis: a retrospective multi-centre study in northern Italy. GSMII (Gruppo di Studio per le Malattie Infiammatorie Intestinali). Eur J Gastroenterol Hepatol *11*, 909-913.

Mitsuhashi, S., Feldbrugge, L., Csizmadia, E., Mitsuhashi, M., Robson, S.C., and Moss, A.C. (2016). Luminal Extracellular Vesicles (EVs) in Inflammatory Bowel Disease (IBD) Exhibit Proinflammatory Effects on Epithelial Cells and Macrophages. Inflamm Bowel Dis *22*, 1587-1595.

Molinie, F., Gower-Rousseau, C., Yzet, T., Merle, V., Grandbastien, B., Marti, R., Lerebours, E., Dupas, J.L., Colombel, J.F., Salomez, J.L.*, et al.* (2004). Opposite evolution in incidence of Crohn's disease and ulcerative colitis in Northern France (1988-1999). Gut *53*, 843-848.

Mottawea, W., Chiang, C.K., Mulbauer, M., Starr, A.E., Butcher, J., Abujamel, T., Deeke, S.A., Brandel, A., Zhou, H., Shokralla, S.*, et al.* (2016). Altered intestinal microbiota-host mitochondria crosstalk in new onset Crohn's disease. Nat Commun *7*.

Nabhan, J.F., Hu, R., Oh, R.S., Cohen, S.N., and Lu, Q. (2012). Formation and release of arrestin domain-containing protein 1-mediated microvesicles (ARMMs) at plasma membrane by recruitment of TSG101 protein. Proc Natl Acad Sci U S A *109*, 4146-4151.

Ng, S.C., Shi, H.Y., Hamidi, N., Underwood, F.E., Tang, W., Benchimol, E.I., Panaccione, R., Ghosh, S., Wu, J.C.Y., Chan, F.K.L.*, et al.* (2017). Worldwide incidence and prevalence of inflammatory bowel disease in the 21st century: a systematic review of population-based studies. Lancet.

Ng, S.C., Tang, W., Ching, J.Y., Wong, M., Chow, C.M., Hui, A.J., Wong, T.C., Leung, V.K., Tsang, S.W., Yu, H.H.*, et al.* (2013). Incidence and phenotype of inflammatory bowel disease based on results from the Asia-pacific Crohn's and colitis epidemiology study. Gastroenterology *145*, 158-165 e152.

Ni, J., Wu, G.D., Albenberg, L., and Tomov, V.T. (2017). Gut microbiota and IBD: causation or correlation? Nat Rev Gastroenterol Hepatol *14*, 573-584.

Nicholson, T.F., Watts, K.M., and Hunstad, D.A. (2009). OmpA of Uropathogenic Escherichia coli Promotes Postinvasion Pathogenesis of Cystitis. Infect Immun *77*, 5245-5251.

Norman, J.M., Handley, S.A., Baldridge, M.T., Droit, L., Liu, C.Y., Keller, B.C., Kambal, A., Monaco, C.L., Zhao, G., Fleshner, P.*, et al.* (2015). Disease-specific alterations in the enteric virome in inflammatory bowel disease. Cell *160*, 447-460.

Nuij, V.J.A.A., Zelinkova, Z., Rijk, M.C.M., Beukers, R., Ouwendijk, R.J.T., Quispel, R., van Tilburg, A.J.P., Tang, T.J., Smalbraak, H., Bruin, K.F.*, et al.* (2013). Phenotype of Inflammatory Bowel Disease at Diagnosis in the Netherlands: A Population-based Inception Cohort Study (the Delta Cohort). Inflamm Bowel Dis *19*, 2215-2222.

Ohira, M., Oshitani, N., Hosomi, S., Watanabe, K., Yamagami, H., Tominaga, K., Watanabe, T., Fujiwara, Y., Maeda, K., Hirakawa, K.*, et al.* (2009). Dislocation of Rab13 and vasodilator-stimulated phosphoprotein in inactive colon epithelium in patients with Crohn's disease. Int J Mol Med *24*, 829-835.

Ong, S.E., Blagoev, B., Kratchmarova, I., Kristensen, D.B., Steen, H., Pandey, A., and Mann, M. (2002). Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. Mol Cell Proteomics *1*, 376-386.

Ostrowski, M., Carmo, N.B., Krumeich, S., Fanget, I., Raposo, G., Savina, A., Moita, C.F., Schauer, K., Hume, A.N., Freitas, R.P.*, et al.* (2010). Rab27a and Rab27b control different steps of the exosome secretion pathway. Nat Cell Biol *12*, 19-30; sup pp 11-13.

Oswiecimska, J., Szymlak, A., Roczniak, W., Girczys-Poledniok, K., and Kwiecien, J. (2017). New insights into the pathogenesis and treatment of irritable bowel syndrome. Adv Med Sci-Poland *62*, 17-30.

Palkovits, J., Novacek, G., Kollars, M., Hron, G., Osterode, W., Quehenberger, P., Kyrle, P.A., Vogelsang, H., Reinisch, W., Papay, P.*, et al.* (2013). Tissue factor exposing microparticles in inflammatory bowel disease. J Crohns Colitis *7*, 222-229.

Park, J.S., Lee, W.C., Yeo, K.J., Ryu, K.S., Kumarasiri, M., Hesek, D., Lee, M., Mobashery, S., Song, J.H., Il Kim, S.*, et al.* (2012). Mechanism of anchoring of OmpA protein to the cell wall peptidoglycan of the gram-negative bacterial outer membrane. Faseb J *26*, 219-228.

Parkes, G.C., Whelan, K., and Lindsay, J.O. (2014). Smoking in inflammatory bowel disease: Impact on disease course and insights into the aetiology of its effect. Journal of Crohns & Colitis *8*, 717-725.

Portela, F., Magro, F., Lago, P., Cotter, J., Cremers, I., de Deus, J., Vieira, A., Lopes, H., Caldeira, P., Barros, L*., et al.* (2010). Ulcerative colitis in a Southern European country: a national perspective. Inflamm Bowel Dis *16*, 822-829.

Povero, D., Eguchi, A., Li, H., Johnson, C.D., Papouchado, B.G., Wree, A., Messer, K., and Feldstein, A.E. (2014). Circulating extracellular vesicles with specific proteome and liver microRNAs are potential biomarkers for liver injury in experimental fatty liver disease. PLoS One *9*, e113651.

Prados-Rosales, R., Weinrick, B.C., Pique, D.G., Jacobs, W.R., Jr., Casadevall, A., and Rodriguez, G.M. (2014). Role for Mycobacterium tuberculosis membrane vesicles in iron acquisition. J Bacteriol *196*, 1250-1256.

Presley, L.L., Ye, J., Li, X., Leblanc, J., Zhang, Z., Ruegger, P.M., Allard, J., McGovern, D., Ippoliti, A., Roth, B*., et al.* (2012). Host-microbe relationships in inflammatory bowel disease detected by bacterial and metaproteomic analysis of the mucosal-luminal interface. Inflamm Bowel Dis *18*, 409-417.

Reese, G.E., Constantinides, V.A., Simillis, C., Darzi, A.W., Orchard, T.R., Fazio, V.W., and Tekkis, P.P. (2006). Diagnostic precision of anti-Saccharomyces cerevisiae antibodies and perinuclear antineutrophil cytoplasmic antibodies in inflammatory bowel disease. Am J Gastroenterol *101*, 2410-2422.

Reimund, J.M., Wittersheim, C., Dumont, S., Muller, C.D., Kenney, J.S., Baumann, R., Poindron, P., and Duclos, B. (1996). Increased production of tumour necrosis factor-alpha, interleukin-1 beta, and interleukin-6 by morphologically normal intestinal biopsies from patients with Crohn's disease. Gut *39*, 684-689.

Reyes-Robles, T., Dillard, R.S., Cairns, L.S., Silva-Valenzuela, C.A., Housman, M., Ali, A., Wright, E.R., and Camilli, A. (2018). Vibrio cholerae outer membrane vesicles inhibit bacteriophage infection. J Bacteriol.

Rhys, H.I., Dell'Accio, F., Pitzalis, C., Moore, A., Norling, L.V., and Perretti, M. (2018). Neutrophil Microvesicles from Healthy Control and Rheumatoid Arthritis Patients Prevent the Inflammatory Activation of Macrophages. EBioMedicine.

Ricciuto, A., Fish, J.R., Tomalty, D.E., Carman, N., Crowley, E., Popalis, C., Muise, A., Walters, T.D., Griffiths, A.M., and Church, P.C. (2017). Diagnostic delay in Canadian children with inflammatory bowel disease is more common in Crohn's disease and associated with decreased height. Arch Dis Child.

Rifai, N., Gillette, M.A., and Carr, S.A. (2006). Protein biomarker discovery and validation: the long and uncertain path to clinical utility. Nat Biotechnol *24*, 971-983.

Roche, S., Gabelle, A., and Lehmann, S. (2008). Clinical proteomics of the cerebrospinal fluid: Towards the discovery of new biomarkers. Proteomics Clin Appl *2*, 428-436.

Rolhion, N., Barnich, N., Bringer, M.A., Glasser, A.L., Ranc, J., Hebuterne, X., Hofman, P., and Darfeuille-Michaud, A. (2010). Abnormally expressed ER stress response chaperone Gp96 in CD favours adherent-invasive Escherichia coli invasion. Gut *59*, 1355-1362.

Roozendaal, C., Zhao, M.H., Horst, G., Lockwood, C.M., Kleibeuker, J.H., Limburg, P.C., Nelis, G.F., and Kallenberg, C.G. (1998). Catalase and alpha-enolase: two novel granulocyte autoantigens in inflammatory bowel disease (IBD). Clin Exp Immunol *112*, 10-16.

Rosen, M.J., Dhawan, A., and Saeed, S.A. (2015). Inflammatory Bowel Disease in Children and Adolescents. JAMA Pediatr *169*, 1053-1060.

Rosenthal, R., Milatz, S., Krug, S.M., Oelrich, B., Schulzke, J.D., Amasheh, S., Gunzel, D., and Fromm, M. (2010). Claudin-2, a component of the tight junction, forms a paracellular water channel. J Cell Sci *123*, 1913-1921.

Ruemmele, F.M., Veres, G., Kolho, K.L., Griffiths, A., Levine, A., Escher, J.C., Dias, J.A., Barabino, A., Braegger, C.P., Bronsky, J*., et al.* (2014). Consensus guidelines of ECCO/ESPGHAN on the medical management of pediatric Crohn's disease. Journal of Crohns & Colitis *8*, 1179-1207.

Santos, M.P.C., Gomes, C., and Torres, J. (2018). Familial and ethnic risk in inflammatory bowel disease. Ann Gastroenterol *31*, 14-23.

Satsangi, J., Silverberg, M.S., Vermeire, S., and Colombel, J.F. (2006). The Montreal classification of inflammatory bowel disease: controversies, consensus, and implications. Gut *55*, 749-753.

Sauer, C.G., and Kugathasan, S. (2009). Pediatric inflammatory bowel disease: highlighting pediatric differences in IBD. Gastroenterol Clin North Am *38*, 611-628.

Savin, Z., Kivity, S., Yonath, H., and Yehuda, S. (2018). Smoking and the intestinal microbiome. Arch Microbiol.

Schoepfer, A.M., Dehlavi, M.A., Fournier, N., Safroneeva, E., Straumann, A., Pittet, V., Peyrin-Biroulet, L., Michetti, P., Rogler, G., Vavricka, S.R*., et al.* (2013). Diagnostic Delay in Crohn's Disease Is Associated With a Complicated Disease Course and Increased Operation Rate. American Journal of Gastroenterology *108*, 1744-1753.

Schreiber, S., Nikolaus, S., and Hampe, J. (1998). Activation of nuclear factor kappa B inflammatory bowel disease. Gut *42*, 477-484.

Schwechheimer, C., Kulp, A., and Kuehn, M.J. (2014). Modulation of bacterial outer membrane vesicle production by envelope structure and content. Bmc Microbiol *14*.

Schwiertz, A., Spiegel, J., Dillmann, U., Grundmann, D., Burmann, J., Fassbender, K., Schafer, K.H., and Unger, M.M. (2018). Fecal markers of intestinal inflammation and intestinal permeability are elevated in Parkinson's disease. Parkinsonism Relat Disord.

Sequeiros, T., Rigau, M., Chiva, C., Montes, M., Garcia-Grau, I., Garcia, M., Diaz, S., Celma, A., Bijnsdorp, I., Campos, A*., et al.* (2017). Targeted proteomics in urinary extracellular vesicles identifies biomarkers for diagnosis and prognosis of prostate cancer. Oncotarget *8*, 4960-4976.

Shah, B., Sullivan, C.J., Lonergan, N.E., Stanley, S., Soult, M.C., and Britt, L.D. (2012). Circulating Bacterial Membrane Vesicles Cause Sepsis in Rats. Shock *37*, 621-628.

Sharara, N., Adam, V., Crott, R., and Barkun, A.N. (2008). The costs of colonoscopy in a Canadian hospital using a microcosting approach. Can J Gastroenterol *22*, 565-570.

Shivashankar, R., and Lichtenstein, G.R. (2018). Mimics of Inflammatory Bowel Disease. Inflamm Bowel Dis.

Singh, P.P., Smith, V.L., Karakousis, P.C., and Schorey, J.S. (2012). Exosomes isolated from mycobacteria-infected mice or cultured macrophages can recruit and activate immune cells in vitro and in vivo. J Immunol *189*, 777-785.

Sohnle, P.G., Collins-Lech, C., and Wiessner, J.H. (1991). The zinc-reversible antimicrobial activity of neutrophil lysates and abscess fluid supernatants. J Infect Dis *164*, 137-142.

Sokol, H., Seksik, P., Furet, J.P., Firmesse, O., Nion-Larmurier, L., Beaugerie, L., Cosnes, J., Corthier, G., Marteau, P., and Dore, J. (2009). Low Counts of Faecalibacterium prausnitzii in Colitis Microbiota. Inflamm Bowel Dis *15*, 1183-1189.

Song, J.Y., Lee, Y.M., Choi, Y.J., and Jeong, S.J. (2017). Fecal calprotectin level in healthy children aged less than 4 years in South Korea. J Clin Lab Anal *31*.

Spekhorst, L.M., Imhann, F., Festen, E.A.M., van Bodegraven, A.A., de Boer, N.K.H., Bouma, G., Fidder, H.H., d'Haens, G., Hoentjen, F., Hommes, D.W.*, et al.* (2017). Cohort profile: design and first results of the Dutch IBD Biobank: a prospective, nationwide biobank of patients with inflammatory bowel disease. BMJ Open *7*, e016695.

Stange, E.F., Travis, S.P., Vermeire, S., Reinisch, W., Geboes, K., Barakauskiene, A., Feakins, R., Flejou, J.F., Herfarth, H., Hommes, D.W.*, et al.* (2008). European evidence-based Consensus on the diagnosis and management of ulcerative colitis: Definitions and diagnosis. J Crohns Colitis *2*, 1-23.

Starr, A.E., Deeke, S.A., Li, L., Zhang, X., Daoud, R., Ryan, J., Ning, Z., Cheng, K., Nguyen, L.V.H., Abou-Samra, E.*, et al.* (2017a). Proteomic and Metaproteomic Approaches to Understand Host-Microbe Interactions. Anal Chem.

Starr, A.E., Deeke, S.A., Ning, Z., Chiang, C.K., Zhang, X., Mottawea, W., Singleton, R., Benchimol, E.I., Wen, M., Mack, D.R.*, et al.* (2017b). Proteomic analysis of ascending colon biopsies from a paediatric inflammatory bowel disease inception cohort identifies protein biomarkers that differentiate Crohn's disease from UC. Gut *66*, 1573-1583.

Stentz, R., Osborne, S., Horn, N., Li, A.W., Hautefort, I., Bongaerts, R., Rouyer, M., Bailey, P., Shears, S.B., Hemmings, A.M.*, et al.* (2014). A bacterial homolog of a eukaryotic inositol phosphate signaling enzyme mediates cross-kingdom dialog in the mammalian gut. Cell Rep *6*, 646-656.

Stephens, D.S., Edwards, K.M., Morris, F., and McGee, Z.A. (1982). Pili and outer membrane appendages on Neisseria meningitidis in the cerebrospinal fluid of an infant. J Infect Dis *146*, 568.

Stephens, M., Batres, L.A., Ng, D., and Baldassano, R. (2001). Growth failure in the child with inflammatory bowel disease. Semin Gastrointest Dis *12*, 253-262.

Strugala, V., Dettmar, P.W., and Pearson, J.P. (2008). Thickness and continuity of the adherent colonic mucus barrier in active and quiescent ulcerative colitis and Crohn's disease. Int J Clin Pract *62*, 762-769.

Szijgyarto, Z., Garedew, A., Azevedo, C., and Saiardi, A. (2011). Influence of inositol pyrophosphates on cellular energy dynamics. Science *334*, 802-805.

Szilagyi, A., and Xue, X. (2018). Comparison of geographic distributions of Irritable Bowel Syndrome with Inflammatory Bowel Disease fail to support common evolutionary roots: Irritable Bowel Syndrome and Inflammatory Bowel Diseases are not related by evolution. Med Hypotheses *110*, 31-37.

Szklarczyk, D., Morris, J.H., Cook, H., Kuhn, M., Wyder, S., Simonovic, M., Santos, A., Doncheva, N.T., Roth, A., Bork, P.*, et al.* (2017). The STRING database in 2017: quality-controlled protein-protein association networks, made broadly accessible. Nucleic Acids Res *45*, D362-D368.

Takeuchi, O., and Akira, S. (2010). Pattern Recognition Receptors and Inflammation. Cell *140*, 805-820.

Thakkar, K., El-Serag, H.B., Mattek, N., and Gilger, M. (2008). Complications of pediatric colonoscopy: a five-year multicenter experience. Clin Gastroenterol Hepatol *6*, 515-520.

The UniProt, C. (2017). UniProt: the universal protein knowledgebase. Nucleic Acids Res *45*, D158-D169.

Thery, C., Amigorena, S., Raposo, G., and Clayton, A. (2006). Isolation and characterization of exosomes from cell culture supernatants and biological fluids. Curr Protoc Cell Biol *Chapter 3*, Unit 3 22.

Thompson, A., Schafer, J., Kuhn, K., Kienle, S., Schwarz, J., Schmidt, G., Neumann, T., Johnstone, R., Mohammed, A.K., and Hamon, C. (2003). Tandem mass tags: a novel quantification strategy for comparative analysis of complex protein mixtures by MS/MS. Anal Chem *75*, 1895-1904.

Timar, C.I., Lorincz, A.M., Csepanyi-Komi, R., Valyi-Nagy, A., Nagy, G., Buzas, E.I., Ivanyi, Z., Kittel, A., Powell, D.W., McLeish, K.R.*, et al.* (2013). Antibacterial effect of microvesicles released from human neutrophilic granulocytes. Blood *121*, 510-518.

Triantafillidis, J.K., Nasioulas, G., and Kosmidis, P.A. (2009). Colorectal cancer and inflammatory bowel disease: epidemiology, risk factors, mechanisms of carcinogenesis and prevention strategies. Anticancer Res *29*, 2727-2737.

Turner, D., Otley, A.R., Mack, D., Hyams, J., de Bruijne, J., Uusoue, K., Walters, T.D., Zachos, M., Mamula, P., Beaton, D.E.*, et al.* (2007). Development, validation, and evaluation of a pediatric ulcerative colitis activity index: a prospective multicenter study. Gastroenterology *133*, 423-432.

Turner, D., Ruemmele, F.M., Orlanski-Meyer, E., Griffiths, A.M., Carpi, J.M., Bronsky, J., Veres, G., Aloi, M., Strisciuglio, C., Braegger, C.P.*, et al.* (2018). Management of Paediatric Ulcerative Colitis, Part 2: Acute Severe Colitis; An Evidence-based Consensus Guideline from ECCO and ESPGHAN. J Pediatr Gastroenterol Nutr.

Ungaro, R., Bernstein, C.N., Gearry, R., Hviid, A., Kolho, K.L., Kronman, M.P., Shaw, S., Van Kruiningen, H., Colombel, J.F., and Atreja, A. (2014). Antibiotics Associated With Increased Risk of New-Onset Crohn's Disease But Not Ulcerative Colitis: A Meta-Analysis. American Journal of Gastroenterology *109*, 1728-1738.

Valinsky, L., Della Vedova, G., Scupham, A.J., Alvey, S., Figueroa, A., Yin, B., Hartin, R.J., Chrobak, M., Crowley, D.E., Jiang, T.*, et al.* (2002). Analysis of bacterial community composition by oligonucleotide fingerprinting of rRNA genes. Appl Environ Microbiol *68*, 3243-3250.

van Niel, G., Charrin, S., Simoes, S., Romao, M., Rochin, L., Saftig, P., Marks, M.S., Rubinstein, E., and Raposo, G. (2011). The tetraspanin CD63 regulates ESCRT-independent and -dependent endosomal sorting during melanogenesis. Dev Cell *21*, 708-721.

van Rheenen, P.F., Van de Vijver, E., and Fidler, V. (2010). Faecal calprotectin for screening of patients with suspected inflammatory bowel disease: diagnostic meta-analysis. BMJ *341*, c3369.

Vernier-Massouille, G., Balde, M., Salleron, J., Turck, D., Dupas, J.L., Mouterde, O., Merle, V., Salomez, J.L., Branche, J., Marti, R.*, et al.* (2008). Natural history of pediatric Crohn's disease: a population-based cohort study. Gastroenterology *135*, 1106-1113.

Vizcaino, J.A., Csordas, A., Del-Toro, N., Dianes, J.A., Griss, J., Lavidas, I., Mayer, G., Perez-Riverol, Y., Reisinger, F., Ternent, T.*, et al.* (2016). 2016 update of the PRIDE database and its related tools. Nucleic Acids Res *44*, 11033.

Voganatsi, A., Panyutich, A., Miyasaki, K.T., and Murthy, R.K. (2001). Mechanism of extracellular release of human neutrophil calprotectin complex. J Leukoc Biol *70*, 130-134.

Wang, T.T., Dabbas, B., Laperriere, D., Bitton, A.J., Soualhine, H., Tavera-Mendoza, L.E., Dionne, S., Servant, M.J., Bitton, A., Seidman, E.G.*, et al.* (2010). Direct and indirect induction by 1,25-dihydroxyvitamin D3 of the NOD2/CARD15-defensin beta2 innate immune pathway defective in Crohn disease. J Biol Chem *285*, 2227-2231.

Wehkamp, J., Salzman, N.H., Porter, E., Nuding, S., Weichenthal, M., Petras, R.E., Shen, B., Schaeffeler, E., Schwab, M., Linzmeier, R.*, et al.* (2005). Reduced Paneth cell alpha-defensins in ileal Crohn's disease. Proc Natl Acad Sci U S A *102*, 18129-18134.

Wisniewski, J.R., Zougman, A., Nagaraj, N., and Mann, M. (2009). Universal sample preparation method for proteome analysis. Nat Methods *6*, 359-362.

Woo, C.H., Eom, Y.W., Yoo, M.H., You, H.J., Han, H.J., Song, W.K., Yoo, Y.J., Chun, J.S., and Kim, J.H. (2000). Tumor necrosis factor-alpha generates reactive oxygen species via a cytosolic phospholipase A2-linked cascade. J Biol Chem *275*, 32357-32362.

Wu, N., Liu, S., Guo, C., Hou, Z., and Sun, M.Z. (2013). The role of annexin A3 playing in cancers. Clin Transl Oncol *15*, 106-110.

Xia, J., Sinelnikov, I.V., Han, B., and Wishart, D.S. (2015). MetaboAnalyst 3.0--making metabolomics more meaningful. Nucleic Acids Res *43*, W251-257.

Xie, F., Liu, T., Qian, W.J., Petyuk, V.A., and Smith, R.D. (2011). Liquid chromatography-mass spectrometry-based quantitative proteomics. J Biol Chem *286*, 25443-25449.

Xu, A.T., Lu, J.T., Ran, Z.H., and Zheng, Q. (2016). Exosome in intestinal mucosal immunity. J Gastroenterol Hepatol *31*, 1694-1699.

Yan, X.D., Yin, J., Yao, H.Y., Mao, N., Yang, Y.L., and Pan, L.Y. (2010). Increased Expression of Annexin A3 Is a Mechanism of Platinum Resistance in Ovarian Cancer. Cancer Res *70*, 1616-1624.

Yang, Z., Clark, N., and Park, K.T. (2014). Effectiveness and Cost-effectiveness of Measuring Fecal Calprotectin in Diagnosis of Inflammatory Bowel Disease in Adults and Children. Clin Gastroenterol H *12*, 253-+.

Yu, F.X., Lin, S.C., Morrisonbogorad, M., Atkinson, M.A.L., and Yin, H.L. (1993). Thymosin Beta-10 and Thymosin Beta-4 Are Both Actin Monomer Sequestering Proteins. Journal of Biological Chemistry *268*, 502-509.

Zernecke, A., Bidzhekov, K., Noels, H., Shagdarsuren, E., Gan, L., Denecke, B., Hristov, M., Koppel, T., Jahantigh, M.N., Lutgens, E.*, et al.* (2009). Delivery of microRNA-126 by apoptotic bodies induces CXCL12-dependent vascular protection. Sci Signal *2*, ra81.

Zhang, X., Deeke, S.A., Ning, Z.B., Starr, A.E., Butcher, J., Li, J., Mayne, J., Cheng, K., Liao, B., Li, L.Y.*, et al.* (2018a). Metaproteomics reveals associations between microbiome and intestinal extracellular vesicle proteins in pediatric inflammatory bowel disease. Nat Commun *9*.

Zhang, X., Ning, Z., Mayne, J., Moore, J.I., Li, J., Butcher, J., Deeke, S.A., Chen, R., Chiang, C.K., Wen, M.*, et al.* (2016). MetaPro-IQ: a universal metaproteomic approach to studying human and mouse gut microbiota. Microbiome *4*, 31.

Zhang, X., Ren, D., Guo, L., Wang, L., Wu, S., Lin, C.Y., Ye, L.P., Zhu, J.R., Li, J., Song, L.B.*, et al.* (2017). Thymosin beta 10 is a key regulator of tumorigenesis and metastasis and a novel serum marker in breast cancer. Breast Cancer Res *19*.

Zhang, X., Yang, F., Zou, J., Wu, W., Jing, H., Gou, Q., Li, H., Gu, J., Zou, Q., and Zhang, J. (2018b). Immunization with Pseudomonas aeruginosa outer membrane vesicles stimulates protective immunity in mice. Vaccine *36*, 1047-1054.

Zhang, X.J., Su, Y.R., Liu, D., Xu, D.B., Zeng, M.S., and Chen, W.K. (2014). Thymosin beta 10 correlates with lymph node metastases of papillary thyroid carcinoma. J Surg Res *192*, 487-493.

Zhang, Y., Fonslow, B.R., Shan, B., Baek, M.C., and Yates, J.R., 3rd (2013). Protein analysis by shotgun/bottom-up proteomics. Chem Rev *113*, 2343-2394.

Zheng, X., Chen, F., Zhang, Q., Liu, Y., You, P., Sun, S., Lin, J., and Chen, N. (2017). Salivary exosomal PSMA7: a promising biomarker of inflammatory bowel disease. Protein Cell *8*, 686-695.

CONTRIBUTIONS OF COLLABORATORS

1. **Dr. Amanda Starr (University of Ottawa)**
   Dr. Starr contributed to the study design, analysis, and interpretation of results, oversaw my project and provided critical content to my thesis. She also developed and optimized the protein isolation from intestinal aspirates and the Super-SILAC spike-in reference. Dr. Starr also performed the inter-plate normalization for the CAT and LTA4H ELISAs.

2. **Dr. Daniel Figeys (University of Ottawa)**
   Dr. Figeys contributed to the study design and interpretation of results and provided critical content to my thesis.

3. **Dr. Alain Stintzi (University of Ottawa)**
   Dr. Stintzi contributed to the study design and interpretation of results.

4. **Ruth Singleton (Children's Hospital of Eastern Ontario)**
   Performed the patient enrolment and contributed to the clinical data acquisition.

5. **Dr. David Mack (Children's Hospital of Eastern Ontario)**
   Performed patient diagnosis, clinical sample collection and contributed to the clinical data acquisition, study design and interpretation of results.

6. **Dr. Xu Zhang (University of Ottawa)**
   Dr. Zhang curated the microbial database and established the microbial database search strategy.

7. **Sara Ahmadi (University of Ottawa)**
   Sara Ahmadi performed the ELISAs (LTA4H and CAT) from patient stool samples.

8. **Dr. Zhibin Ning (University of Ottawa)**
   Dr. Ning contributed to the interpretation of mass spectrometry results and ensured the mass spectrometry working conditions.

9. **Dr. Eric Benchimol (Children's Hospital of Eastern Ontario)**
   Dr. Benchimol performed patient diagnosis, clinical sample collection and contributed to the clinical data acquisition.

10. **Dr. Arkadiy Reunov (University of Ottawa)**
    Dr. Reunov performed the negative staining of extracellular vesicles and the visualization by transmission electron microscopy on a fee for service basis (University of Ottawa

Heart Institute, Loeb Research Centre, Electron microscopy laboratory, Cell Imaging and Histology Core Facility).

APPENDICES

Appendix 1

Table 1. Discriminant features of active IBD identified by PLS-DA in the ascending colon

Uniprot ID	Protein name	Gene name	VIP Comp. 1	VIP Comp. 2
A2ACR1	Proteasome subunit beta type;Proteasome subunit beta type-9	PSMB9	1.8463	1.6801
A8K7I4	Calcium-activated chloride channel regulator 1	CLCA1	1.5368	1.4143
B1AKD8	Rootletin	CROCC	1.2546	1.154
B4DFR2	Dynein light chain roadblock-type 1;Dynein light chain roadblock-type 2	DYNLRB1;DYNLRB2	1.6243	1.4534
B4DHX4	Rab GDP dissociation inhibitor alpha	GDI1	1.2472	1.161
B4DLN1			1.1633	1.081
B4DQI8	Ras-related protein Rap-1b;Ras-related protein Rap-1b-like protein	RAP1B	1.1289	1.0478
B4DQJ8	6-phosphogluconate dehydrogenase, decarboxylating	PGD	2.2108	2.3277
B4DTT0			1.0819	1.0066
B4DU58	Macrophage-capping protein	CAPG	2.4373	2.1792
B5MDF5	GTP-binding nuclear protein Ran	RAN	1.4771	1.7973
B9A064	Immunoglobulin lambda-like polypeptide 5;Ig lambda-1 chain C regions	IGLL5;IGLC1	1.6102	1.4333
E5RJA0	Acyl-protein thioesterase 1	LYPLA1	1.25	1.1289
E7ETB3	Aspartyl aminopeptidase	DNPEP	1.7306	1.5657
E7EU23	Rab GDP dissociation inhibitor beta	GDI2	1.1727	1.0502
E9PGT1	Translin	TSN	2.0529	1.8356
F5H2A5	Aconitate hydratase, mitochondrial	ACO2	1.2581	1.172
F5H4I0	Dihydrolipoyl dehydrogenase, mitochondrial;Dihydrolipoyl dehydrogenase	DLD	1.6307	1.4759
F5H816	Glycogen phosphorylase, liver form;Alpha-1,4 glucan phosphorylase	PYGL	3.552	3.1729
F8W1R7	Myosin light polypeptide 6	MYL6	2.2469	2.0005
F8WA45	Gelsolin	GSN	3.9791	3.5481
G3V5I3	Alpha-1-antichymotrypsin;Alpha-1-antichymotrypsin His-Pro-less	SERPINA3	1.2057	1.1533
H0YG33	Heat shock 70 kDa protein 1A	HSPA1A	1.5788	1.4071
H0YGH4	Alpha-2-macroglobulin	A2M	1.2572	1.1192
O14818	Proteasome subunit alpha type-7	PSMA7	1.5312	1.3654
O15143	Actin-related protein 2/3 complex subunit 1B	ARPC1B	2.2915	2.0736
O15144	Actin-related protein 2/3 complex subunit 2	ARPC2	2.5097	2.4231

O15145	Actin-related protein 2/3 complex subunit 3	ARPC3	1.1637	1.0873
O43707	Alpha-actinin-4	ACTN4	1.025	1.048
O60218	Aldo-keto reductase family 1 member B10	AKR1B10	2.7767	2.6013
O60506	Heterogeneous nuclear ribonucleoprotein Q	SYNCRIP	1.1035	1.1542
O75083	WD repeat-containing protein 1	WDR1	1.948	1.7345
O75131	Copine-3	CPNE3	1.1342	1.036
O75367-3	Core histone macro-H2A.1;Histone H2A	H2AFY	1.0266	1.104
O75368	SH3 domain-binding glutamic acid-rich-like protein	SH3BGRL	3.1917	2.8473
O75369-2	Filamin-B	FLNB	1.1236	1.0456
O75533	Splicing factor 3B subunit 1	SF3B1	1.2108	1.0781
O75882-2	Attractin	ATRN	1.5141	1.4473
O95336	6-phosphogluconolactonase	PGLS	1.3372	1.2357
P00338	L-lactate dehydrogenase A chain	LDHA	1.6209	1.4512
P00441	Superoxide dismutase [Cu-Zn]	SOD1	1.2311	1.1477
P00491	Purine nucleoside phosphorylase	PNP	1.6607	1.6543
P00558	Phosphoglycerate kinase 1;Phosphoglycerate kinase	PGK1	2.2985	2.0622
P00918	Carbonic anhydrase 2	CA2	1.232	1.2405
P01024	Complement C3;Complement C3 beta chain;C3-beta-c;Complement C3 alpha chain;C3a anaphylatoxin;Acylation stimulating protein;Complement C3b alpha chain;Complement C3c alpha chain fragment 1;Complement C3dg fragment;Complement C3g fragment;Complement C3d fragment;Complement C3f fragment;Complement C3c alpha chain fragment 2	C3	1.4562	1.5589
P01591	Immunoglobulin J chain	IGJ;JCHAIN	1.5416	1.4401
P01593	Ig kappa chain V-I region AG		1.6798	1.5773
P01764	Ig heavy chain V-III region 23	IGHV3-23	2.1501	2.1419
P01833	Polymeric immunoglobulin receptor;Secretory component	PIGR	1.1629	1.0494
P01859	Ig gamma-2 chain C region	IGHG2	1.1814	1.1702
P01871	Ig mu chain C region	IGHM	1.3428	1.1978
P01877	Ig alpha-2 chain C region	IGHA2	1.5395	1.4419
P02787	Serotransferrin	TF	1.4056	1.6775
P02794	Ferritin heavy chain;Ferritin heavy chain, N-terminally processed;Ferritin	FTH1	1.2312	1.1052
P02795	Metallothionein-2	MT2A	2.1661	2.0692
P04040	Catalase	CAT	3.6374	3.2579
P04075	Fructose-bisphosphate aldolase A;Fructose-bisphosphate aldolase	ALDOA	1.705	1.549

P04083	Annexin A1	ANXA1	1.3152	1.1742
P04211	Ig lambda chain V region 4A		1.1334	1.0538
P04433	Ig kappa chain V-III region VG		1.0798	1.4879
P04632	Calpain small subunit 1	CAPNS1	1.6207	1.4427
P04899	Guanine nucleotide-binding protein G(i) subunit alpha-2	GNAI2	1.4558	1.2975
P05107	Integrin beta-2;Integrin beta	ITGB2	2.4874	2.2753
P05109	Protein S100-A8;Protein S100-A8, N-terminally processed	S100A8	3.1598	3.0118
P05164-2	Myeloperoxidase;Myeloperoxidase;89 kDa myeloperoxidase;84 kDa myeloperoxidase;Myeloperoxidase light chain;Myeloperoxidase heavy chain	MPO	4.386	4.4191
P05556	Integrin beta-1	ITGB1	1.218	1.0887
P06703	Protein S100-A6	S100A6	1.2227	1.2013
P06731	Carcinoembryonic antigen-related cell adhesion molecule 5	CEACAM5	1.2493	1.1921
P06733	Alpha-enolase	ENO1	2.0816	1.8529
P06744	Glucose-6-phosphate isomerase	GPI	2.219	2.0251
P06753-2	Tropomyosin alpha-3 chain	TPM3;DKFZp686J1372	1.9104	1.7049
P06870-2	Kallikrein-1	KLK1	1.9016	1.702
P07148	Fatty acid-binding protein, liver	FABP1	3.0206	3.2388
P07384	Calpain-1 catalytic subunit	CAPN1	1.3741	1.226
P07478	Trypsin-2	PRSS2	2.2574	2.0095
P07737	Profilin-1	PFN1	2.0577	2.2476
P08246	Neutrophil elastase	ELANE	3.2086	3.0049
P08311	Cathepsin G	CTSG	3.5457	3.2234
P09211	Glutathione S-transferase P	GSTP1	1.7301	1.5404
P09327	Villin-1	VIL1	1.273	1.2053
P09429	High mobility group protein B1	HMGB1	1.2266	1.3776
P09960	Leukotriene A-4 hydrolase	LTA4H	3.9257	3.5848
P11279	Lysosome-associated membrane glycoprotein 1	LAMP1	2.5432	2.2741
P11413	Glucose-6-phosphate 1-dehydrogenase	G6PD	1.7753	1.657
P11766	Alcohol dehydrogenase class-3	ADH5	1.3005	1.3832
P11940	Polyadenylate-binding protein 1;Polyadenylate-binding protein	PABPC1	1.7757	1.7386
P12429	Annexin A3;Annexin	ANXA3	4.2175	3.7842
P12814-3	Alpha-actinin-1	ACTN1	3.773	3.3581
P12821	Angiotensin-converting enzyme;Angiotensin-converting enzyme, soluble form	ACE	1.7212	1.6198
P12955	Xaa-Pro dipeptidase	PEPD	2.6663	2.3766
P13667	Protein disulfide-isomerase A4	PDIA4	1.077	1.0845
P13796	Plastin-2	LCP1	2.809	2.5002

P14174	Macrophage migration inhibitory factor	MIF	1.1728	1.0925
P15086	Carboxypeptidase B	CPB1	1.2343	1.1009
P15121	Aldose reductase	AKR1B1	1.6013	1.5548
P15144	Aminopeptidase N	ANPEP	1.5127	1.6491
P15153	Ras-related C3 botulinum toxin substrate 2	RAC2	1.1601	1.276
P16152	Carbonyl reductase [NADPH] 1	CBR1	1.2207	1.628
P16444	Dipeptidase 1	DPEP1	1.7896	1.6268
P17066	Heat shock 70 kDa protein 6	HSPA6	1.4927	1.34
P17931	Galectin-3;Galectin	LGALS3	1.2702	1.1624
P18206-2	Vinculin	VCL	1.8523	1.686
P18564	Integrin beta-6;Integrin beta	ITGB6	2.1348	1.9002
P18669	Phosphoglycerate mutase 1	PGAM1	1.3515	1.6602
P19105	Myosin regulatory light chain 12A;Myosin regulatory light chain 12B	MYL12A;MYL12B	1.3614	1.2256
P19338	Nucleolin	NCL	1.7367	1.5675
P20160	Azurocidin	AZU1	3.6714	3.7348
P20700	Lamin-B1	LMNB1	1.1802	1.0516
P22392-2	Nucleoside diphosphate kinase B;Nucleoside diphosphate kinase	NME2;NME1-NME2	1.5253	1.3593
P23246	Splicing factor, proline- and glutamine-rich	SFPQ	1.2692	1.1614
P23381	Tryptophan--tRNA ligase, cytoplasmic;T1-TrpRS;T2-TrpRS	WARS	1.9955	1.858
P25787	Proteasome subunit alpha type-2	PSMA2	2.1276	1.9045
P25815	Protein S100-P	S100P	2.7626	2.5438
P27487	Dipeptidyl peptidase 4;Dipeptidyl peptidase 4 membrane form;Dipeptidyl peptidase 4 soluble form	DPP4	3.5175	3.1628
P28062-2	Proteasome subunit beta type-8	PSMB8	1.2403	1.3225
P28066	Proteasome subunit alpha type-5	PSMA5	1.1382	1.0255
P28838-2	Cytosol aminopeptidase	LAP3	1.7848	1.6074
P29218	Inositol monophosphatase 1	IMPA1	1.1324	1.1053
P29401	Transketolase	TKT	3.4146	3.0576
P30043	Flavin reductase (NADPH)	BLVRB	1.6952	1.5122
P30520	Adenylosuccinate synthetase isozyme 2	ADSS	1.2528	1.1901
P30740	Leukocyte elastase inhibitor	SERPINB1	2.2311	2.1283
P31146	Coronin-1A;Coronin	CORO1A	2.9882	2.7085
P31949	Protein S100-A11;Protein S100-A11, N-terminally processed	S100A11	1.9279	1.7428
P32119	Peroxiredoxin-2	PRDX2	1.3378	1.1907
P34896-2	Serine hydroxymethyltransferase, cytosolic	SHMT1	1.137	1.0178
P35237	Serpin B6	SERPINB6	2.1354	1.9202
P35754	Glutaredoxin-1	GLRX	2.4151	2.2354
P36957	Dihydrolipoyllysine-residue succinyltransferase component of 2-	DLST	1.4492	1.2899

	oxoglutarate dehydrogenase complex, mitochondrial			
P37837	Transaldolase	TALDO1	3.1005	2.7663
P39023	60S ribosomal protein L3	RPL3	1.2832	1.2216
P40227	T-complex protein 1 subunit zeta	CCT6A	1.4208	1.905
P42704	Leucine-rich PPR motif-containing protein, mitochondrial	LRPPRC	1.5057	1.3506
P43490	Nicotinamide phosphoribosyltransferase	NAMPT	2.9494	2.6729
P46779	60S ribosomal protein L28	RPL28	1.053	1.1193
P46781	40S ribosomal protein S9	RPS9	1.2618	1.1653
P46940	Ras GTPase-activating-like protein IQGAP1	IQGAP1	2.2344	2.0585
P47756-2	F-actin-capping protein subunit beta	CAPZB	1.5324	1.4321
P48047	ATP synthase subunit O, mitochondrial	ATP5O	1.8942	1.7972
P50552	Vasodilator-stimulated phosphoprotein	VASP	2.3371	2.085
P51149	Ras-related protein Rab-7a	RAB7A	1.8678	1.6626
P52272-2	Heterogeneous nuclear ribonucleoprotein M	HNRNPM	1.9274	1.7543
P52565	Rho GDP-dissociation inhibitor 1	ARHGDIA	1.3922	1.4307
P52566	Rho GDP-dissociation inhibitor 2	ARHGDIB	3.7081	3.3022
P52907	F-actin-capping protein subunit alpha-1	CAPZA1	1.5478	1.4283
P55084	Trifunctional enzyme subunit beta, mitochondrial;3-ketoacyl-CoA thiolase	HADHB	1.1128	1.0759
P55259-3	Pancreatic secretory granule membrane major glycoprotein GP2	GP2	1.2469	1.1919
P58546	Myotrophin	MTPN	2.302	2.0891
P59666	Neutrophil defensin 3;HP 3-56;Neutrophil defensin 2	DEFA3	1.3475	1.2392
P59998	Actin-related protein 2/3 complex subunit 4	ARPC4;ARPC4-TTLL3	2.0249	1.8791
P60174	Triosephosphate isomerase	TPI1	1.1594	1.0344
P61006	Ras-related protein Rab-8A	RAB8A	1.5304	1.3818
P61088	Ubiquitin-conjugating enzyme E2 N;Putative ubiquitin-conjugating enzyme E2 N-like	UBE2N;UBE2NL	1.1358	1.1848
P61160	Actin-related protein 2	ACTR2	1.6463	1.5027
P61586	Transforming protein RhoA	RHOA	1.2426	1.1218
P61626	Lysozyme C;Lysozyme	LYZ	4.4498	4.023
P62829	60S ribosomal protein L23	RPL23	1.6372	1.46
P62917	60S ribosomal protein L8	RPL8	1.1018	1.0685
P62937	Peptidyl-prolyl cis-trans isomerase A;Peptidyl-prolyl cis-trans isomerase A, N-terminally processed;Peptidyl-prolyl cis-trans isomerase	PPIA	1.2163	1.1509
P62942	Peptidyl-prolyl cis-trans isomerase FKBP1A;Peptidyl-prolyl cis-trans isomerase	FKBP1A	1.974	1.7751
P62993	Growth factor receptor-bound protein 2	GRB2	1.3133	1.2684
P63261	Actin, cytoplasmic 2;Actin, cytoplasmic 2, N-terminally processed	ACTG1	1.2268	1.1204

P63313	Thymosin beta-10	TMSB10	1.2537	1.1262
P68871	Hemoglobin subunit beta;LVV-hemorphin-7;Spinorphin	HBB	2.3637	2.3945
P69905	Hemoglobin subunit alpha	HBA1	1.2258	1.2764
P78417	Glutathione S-transferase omega-1	GSTO1	1.623	1.68
P78527	DNA-dependent protein kinase catalytic subunit	PRKDC	1.3428	1.3598
P98073	Enteropeptidase;Enteropeptidase non-catalytic heavy chain;Enteropeptidase catalytic light chain	TMPRSS15	2.2729	2.032
Q01518	Adenylyl cyclase-associated protein 1	CAP1	1.8738	1.6797
Q03154	Aminoacylase-1	ACY1;ABHD14A-ACY1	2.7541	2.4555
Q07075	Glutamyl aminopeptidase	ENPEP	1.7931	1.5993
Q08380	Galectin-3-binding protein	LGALS3BP	1.1954	1.0767
Q13228	Selenium-binding protein 1	SELENBP1	1.4204	1.287
Q13263	Transcription intermediary factor 1-beta	TRIM28	1.1118	1.081
Q14019	Coactosin-like protein	COTL1	1.1072	1.4041
Q15233	Non-POU domain-containing octamer-binding protein	NONO	1.8265	1.6982
Q15717	ELAV-like protein 1	ELAVL1	1.2289	1.1451
Q15907	Ras-related protein Rab-11B;Ras-related protein Rab-11A	RAB11B;RAB11A	1.1745	1.0488
Q16658	Fascin	FSCN1	1.3809	1.232
Q16819	Meprin A subunit alpha	MEP1A	1.2768	1.2157
Q16820	Meprin A subunit beta	MEP1B	1.3818	1.2298
Q16851-2	UTP--glucose-1-phosphate uridylyltransferase	UGP2	2.1405	1.9053
Q5HY54	Filamin-A	FLNA	1.5251	1.3629
Q5JP53	Tubulin beta chain	TUBB	1.5807	1.4614
Q6UWV6	Ectonucleotide pyrophosphatase/phosphodiesterase family member 7	ENPP7	1.5546	1.4045
Q6W4X9	Mucin-6	MUC6	1.6352	1.4881
Q6XQN6	Nicotinate phosphoribosyltransferase	NAPRT	1.1565	1.3674
Q86UP6	CUB and zona pellucida-like domain-containing protein 1	CUZD1	2.6907	2.4386
Q86UX7-2	Fermitin family homolog 3	FERMT3	1.1045	1.0101
Q93077	Histone H2A type 1-C	HIST1H2AC	1.0711	1.0869
Q96C19	EF-hand domain-containing protein D2	EFHD2	1.424	1.4163
Q96G03	Phosphoglucomutase-2	PGM2	1.747	1.5576
Q99536	Synaptic vesicle membrane protein VAT-1 homolog	VAT1	2.7258	2.4294
Q99895	Chymotrypsin-C	CTRC	1.565	1.5016
Q9BRA2	Thioredoxin domain-containing protein 17	TXNDC17	1.1633	1.1057
Q9BRF8	Serine/threonine-protein phosphatase CPPED1	CPPED1	2.3603	2.1009

Q9BYE9	Cadherin-related family member 2	CDHR2	1.2275	1.3673
Q9H0U4	Ras-related protein Rab-1B	RAB1B	1.3423	1.195
Q9ULV4	Coronin-1C;Coronin	CORO1C	1.6893	1.5078
Q9Y266	Nuclear migration protein nudC	NUDC	1.4919	1.5078
Q9Y490	Talin-1	TLN1	1.629	1.5116

Table 2. Discriminant features of active IBD identified by PLS-DA in the descending colon

Uniprot ID	Protein name	Gene name	VIP Comp. 1	VIP Comp. 2
A2ACR1	Proteasome subunit beta type;Proteasome subunit beta type-9	PSMB9	1.7976	1.6783
A6NJA2	Ubiquitin carboxyl-terminal hydrolase;Ubiquitin carboxyl-terminal hydrolase 14	USP14	1.0217	1.0822
A8K7I4	Calcium-activated chloride channel regulator 1	CLCA1	1.0364	1.0263
B1AKD8	Rootletin	CROCC	1.1071	1.0279
B4DFR2	Dynein light chain roadblock-type 1;Dynein light chain roadblock-type 2	DYNLRB1;DYNLRB2	2.76	2.624
B4DN87	Serpin H1	SERPINH1	1.4054	1.3282
B4DQJ8	6-phosphogluconate dehydrogenase, decarboxylating	PGD	2.64	2.5169
B4DR31	Dihydropyrimidinase-related protein 2	DPYSL2	1.1278	1.0377
B4DU58	Macrophage-capping protein	CAPG	1.9874	1.8225
B4E2S7	Lysosome-associated membrane glycoprotein 2	LAMP2	1.0142	1.0514
B5MCA4	Epithelial cell adhesion molecule	EPCAM	1.5529	1.4183
C9JFR7	Cytochrome c	CYCS	1.613	1.5803
C9JIF9	Acylamino-acid-releasing enzyme	APEH	1.0788	1.0055
D3YTB1	60S ribosomal protein L32	RPL32	1.8299	1.7184
D6R938	Calcium/calmodulin-dependent protein kinase type II subunit delta	CAMK2D	1.3395	1.4176
E7ER45	Maltase-glucoamylase, intestinal;Maltase;Glucoamylase	MGAM	1.6995	1.5999
E7ETB3	Aspartyl aminopeptidase	DNPEP	1.3214	1.3797
E7EU23	Rab GDP dissociation inhibitor beta	GDI2	1.308	1.2178
E9PFM1	Eukaryotic translation initiation factor 4 gamma 1	EIF4G1	1.2702	1.3163
E9PGN7	Plasma protease C1 inhibitor	SERPING1	1.1147	1.2189
F5GYN4	Ubiquitin thioesterase OTUB1	OTUB1	1.4885	1.4977

F5GYQ4	Glutamate dehydrogenase 1, mitochondrial;Glutamate dehydrogenase 2, mitochondrial	GLUD1;GLUD2	1.6284	1.6916
F5H4I0	Dihydrolipoyl dehydrogenase, mitochondrial;Dihydrolipoyl dehydrogenase	DLD	1.5647	1.4579
F5H4R9	Dynamin-2	DNM2	1.3026	1.1942
F5H816	Glycogen phosphorylase, liver form;Alpha-1,4 glucan phosphorylase	PYGL	2.9496	2.6928
F5HGG6	Dipeptidyl peptidase 1;Dipeptidyl peptidase 1 exclusion domain chain;Dipeptidyl peptidase 1 heavy chain;Dipeptidyl peptidase 1 light chain	CTSC	1.1522	1.0602
F8W1R7	Myosin light polypeptide 6	MYL6	1.0757	1.0196
F8WA45	Gelsolin	GSN	2.8215	2.5778
F8WA86	Calponin-3;Calponin	CNN3	1.0936	1.0292
H0YG33	Heat shock 70 kDa protein 1A	HSPA1A	1.2133	1.163
H0YGH4	Alpha-2-macroglobulin	A2M	1.2788	1.5369
H0YL72	Isocitrate dehydrogenase [NAD] subunit alpha, mitochondrial	IDH3A	1.2573	1.1571
H0YN18		PSMA4	1.4059	1.2863
H3BP91	Enoyl-CoA delta isomerase 1, mitochondrial	ECI1;DCI	1.349	1.3284
H3BPE7	RNA-binding protein FUS	FUS	1.545	1.4353
H7BXD5	Grancalcin	GCA	4.8671	4.5503
H9KV70	Neutrophil gelatinase-associated lipocalin	LCN2	1.1387	1.1556
O00151	PDZ and LIM domain protein 1	PDLIM1	1.3915	1.4657
O14818	Proteasome subunit alpha type-7	PSMA7	1.1995	1.1005
O15143	Actin-related protein 2/3 complex subunit 1B	ARPC1B	2.5142	2.3325
O15144	Actin-related protein 2/3 complex subunit 2	ARPC2	2.2976	2.1581
O15145	Actin-related protein 2/3 complex subunit 3	ARPC3	1.1704	1.3406
O15511	Actin-related protein 2/3 complex subunit 5	ARPC5	2.5055	2.3885
O43390	Heterogeneous nuclear ribonucleoprotein R	HNRNPR	1.1233	1.0615
O60218	Aldo-keto reductase family 1 member B10	AKR1B10	2.9106	2.7561
O60506	Heterogeneous nuclear ribonucleoprotein Q	SYNCRIP	1.0235	1.0317
O60664	Perilipin-3	PLIN3	1.013	1.2977
O60701	UDP-glucose 6-dehydrogenase	UGDH	1.1366	1.1692
O75083	WD repeat-containing protein 1	WDR1	2.0582	1.8878
O75131	Copine-3	CPNE3	1.1965	1.134
O75367-3	Core histone macro-H2A.1;Histone H2A	H2AFY	1.1247	1.0334
O75368	SH3 domain-binding glutamic acid-rich-like protein	SH3BGRL	2.4495	2.2345
O75533	Splicing factor 3B subunit 1	SF3B1	1.6482	1.5003
O75882-2	Attractin	ATRN	1.52	1.3933
O76021	Ribosomal L1 domain-containing protein 1	RSL1D1	1.003	1.0342

O95373	Importin-7	IPO7	1.1022	1.0095
P00338	L-lactate dehydrogenase A chain	LDHA	1.3466	1.2474
P00491	Purine nucleoside phosphorylase	PNP	1.4413	1.5415
P00492	Hypoxanthine-guanine phosphoribosyltransferase	HPRT1	1.7927	1.645
P00558	Phosphoglycerate kinase 1;Phosphoglycerate kinase	PGK1	1.2463	1.2625
P00738	Haptoglobin;Haptoglobin alpha chain;Haptoglobin beta chain	HP	1.2176	1.1151
P00918	Carbonic anhydrase 2	CA2	1.651	1.5429
P01024	Complement C3;Complement C3 beta chain;C3-beta-c;Complement C3 alpha chain;C3a anaphylatoxin;Acylation stimulating protein;Complement C3b alpha chain;Complement C3c alpha chain fragment 1;Complement C3dg fragment;Complement C3g fragment;Complement C3d fragment;Complement C3f fragment;Complement C3c alpha chain fragment 2	C3	1.8844	1.8301
P01591	Immunoglobulin J chain	IGJ;JCHAIN	2.7305	2.4857
P01764	Ig heavy chain V-III region 23	IGHV3-23	2.4884	2.2975
P01772	Ig heavy chain V-III region KOL		1.2672	1.2126
P01781	Ig heavy chain V-III region GAL		1.0295	1.0015
P01834	Ig kappa chain C region	IGKC	1.6278	1.4809
P01877	Ig alpha-2 chain C region	IGHA2	2.0333	1.9322
P02763	Alpha-1-acid glycoprotein 1	ORM1	1.0732	1.1131
P02787	Serotransferrin	TF	2.0429	1.8731
P02794	Ferritin heavy chain;Ferritin heavy chain, N-terminally processed;Ferritin	FTH1	1.5183	1.3882
P04040	Catalase	CAT	3.4417	3.1636
P04075	Fructose-bisphosphate aldolase A;Fructose-bisphosphate aldolase	ALDOA	1.2194	1.2156
P04083	Annexin A1	ANXA1	1.9386	1.7815
P05091	Aldehyde dehydrogenase, mitochondrial	ALDH2	1.5834	1.5558
P05107	Integrin beta-2;Integrin beta	ITGB2	1.5535	1.4638
P05109	Protein S100-A8;Protein S100-A8, N-terminally processed	S100A8	1.8583	1.8
P05164-2	Myeloperoxidase;Myeloperoxidase;89 kDa myeloperoxidase;84 kDa myeloperoxidase;Myeloperoxidase light chain;Myeloperoxidase heavy chain	MPO	1.5262	1.3974
P05451	Lithostathine-1-alpha	REG1A	1.0111	1.025
P05556	Integrin beta-1	ITGB1	1.2632	1.1948
P06702	Protein S100-A9	S100A9	1.326	2.1379

P06731	Carcinoembryonic antigen-related cell adhesion molecule 5	CEACAM5	1.7189	1.9041
P06733	Alpha-enolase	ENO1	1.3034	1.1867
P06744	Glucose-6-phosphate isomerase	GPI	2.5958	2.3911
P07148	Fatty acid-binding protein, liver	FABP1	3.4238	3.1218
P07384	Calpain-1 catalytic subunit	CAPN1	1.1861	1.1054
P07478	Trypsin-2	PRSS2	2.4668	2.2499
P07737	Profilin-1	PFN1	2.6549	2.4885
P08217	Chymotrypsin-like elastase family member 2A	CELA2A	1.379	1.4621
P08311	Cathepsin G	CTSG	2.5686	2.3374
P09093	Chymotrypsin-like elastase family member 3A	CELA3A	1.0134	1.4024
P09211	Glutathione S-transferase P	GSTP1	1.1994	1.1283
P09327	Villin-1	VIL1	2.054	1.902
P09923	Intestinal-type alkaline phosphatase	ALPI	1.8653	1.7849
P09960	Leukotriene A-4 hydrolase	LTA4H	3.7419	3.4179
P09972	Fructose-bisphosphate aldolase C;Fructose-bisphosphate aldolase	ALDOC	1.2979	1.4493
P10153	Non-secretory ribonuclease	RNASE2	2.4335	2.3114
P10412	Histone H1.4	HIST1H1E	1.2122	1.1262
P10809	60 kDa heat shock protein, mitochondrial	HSPD1	1.1286	1.0856
P11279	Lysosome-associated membrane glycoprotein 1	LAMP1	2.376	2.1819
P11413	Glucose-6-phosphate 1-dehydrogenase	G6PD	2.4098	2.2288
P11766	Alcohol dehydrogenase class-3	ADH5	1.2397	1.1445
P11940	Polyadenylate-binding protein 1;Polyadenylate-binding protein	PABPC1	1.2832	1.1695
P12277	Creatine kinase B-type	CKB	2.4258	2.4499
P12429	Annexin A3;Annexin	ANXA3	4.1043	3.7723
P12814-3	Alpha-actinin-1	ACTN1	3.4722	3.1593
P12821	Angiotensin-converting enzyme;Angiotensin-converting enzyme, soluble form	ACE	1.1252	1.2104
P12955	Xaa-Pro dipeptidase	PEPD	1.7979	1.8097
P13010	X-ray repair cross-complementing protein 5	XRCC5	1.1527	1.2528
P13667	Protein disulfide-isomerase A4	PDIA4	1.6048	1.4678
P13796	Plastin-2	LCP1	3.1818	2.9051
P13797	Plastin-3	PLS3	1.1886	1.1181
P14618	Pyruvate kinase PKM;Pyruvate kinase	PKM	1.4993	1.48
P15153	Ras-related C3 botulinum toxin substrate 2	RAC2	1.7948	1.8525
P15924	Desmoplakin	DSP	1.2144	1.136
P16152	Carbonyl reductase [NADPH] 1	CBR1	1.3981	1.5522
P17844	Probable ATP-dependent RNA helicase	DDX5	1.5096	1.3837

	DDX5			
P17931	Galectin-3;Galectin	LGALS3	1.1252	1.0275
P18085	ADP-ribosylation factor 4	ARF4	1.3453	1.2508
P18135	Ig kappa chain V-III region HAH;Ig kappa chain V-III region HIC		1.5931	1.4887
P18206-2	Vinculin	VCL	1.6788	1.5851
P18564	Integrin beta-6;Integrin beta	ITGB6	1.7055	1.5632
P21796	Voltage-dependent anion-selective channel protein 1	VDAC1	1.1294	1.112
P22695	Cytochrome b-c1 complex subunit 2, mitochondrial	UQCRC2	1.2022	1.094
P23381	Tryptophan--tRNA ligase, cytoplasmic;T1-TrpRS;T2-TrpRS	WARS	1.9274	1.7593
P25398	40S ribosomal protein S12	RPS12	1.1638	1.0971
P25685	DnaJ homolog subfamily B member 1	DNAJB1	1.268	1.3582
P25787	Proteasome subunit alpha type-2	PSMA2	1.1375	1.0348
P25788-2	Proteasome subunit alpha type-3	PSMA3	1.461	1.3321
P25815	Protein S100-P	S100P	2.7324	2.503
P26447	Protein S100-A4	S100A4	1.1452	1.1481
P27487	Dipeptidyl peptidase 4;Dipeptidyl peptidase 4 membrane form;Dipeptidyl peptidase 4 soluble form	DPP4	3.4936	3.1791
P28062-2	Proteasome subunit beta type-8	PSMB8	1.2964	1.1929
P28799	Granulins;Acrogranin;Paragranulin;Granulin-1;Granulin-2;Granulin-3;Granulin-4;Granulin-5;Granulin-6;Granulin-7	GRN	1.1887	1.1782
P28838-2	Cytosol aminopeptidase	LAP3	1.2033	1.2113
P29401	Transketolase	TKT	3.5371	3.2218
P30040	Endoplasmic reticulum resident protein 29	ERP29	1.1196	1.0233
P30043	Flavin reductase (NADPH)	BLVRB	1.523	1.4803
P30520	Adenylosuccinate synthetase isozyme 2	ADSS	1.3283	1.404
P30740	Leukocyte elastase inhibitor	SERPINB1	1.8026	1.7059
P31040	Succinate dehydrogenase [ubiquinone] flavoprotein subunit, mitochondrial	SDHA	1.0776	1.07
P31146	Coronin-1A;Coronin	CORO1A	3.6647	3.3402
P31947-2	14-3-3 protein sigma	SFN	1.9245	1.9685
P31949	Protein S100-A11;Protein S100-A11, N-terminally processed	S100A11	2.6716	2.454
P33176	Kinesin-1 heavy chain	KIF5B	1.6421	1.5066
P35237	Serpin B6	SERPINB6	2.2093	2.0478
P35754	Glutaredoxin-1	GLRX	3.4591	3.2191
P36222	Chitinase-3-like protein 1	CHI3L1	2.9302	2.8158
P37837	Transaldolase	TALDO1	2.654	2.4272

P38159	RNA-binding motif protein, X chromosome;RNA-binding motif protein, X chromosome, N-terminally processed;RNA binding motif protein, X-linked-like-1	RBMX;RBMXL1	1.0883	1.415
P38919	Eukaryotic initiation factor 4A-III;Eukaryotic initiation factor 4A-III, N-terminally processed	EIF4A3	1.3188	1.1998
P40939	Trifunctional enzyme subunit alpha, mitochondrial;Long-chain enoyl-CoA hydratase;Long chain 3-hydroxyacyl-CoA dehydrogenase	HADHA	1.1724	1.1227
P41091	Eukaryotic translation initiation factor 2 subunit 3;Putative eukaryotic translation initiation factor 2 subunit 3-like protein	EIF2S3;EIF2S3L	1.2315	1.1282
P42704	Leucine-rich PPR motif-containing protein, mitochondrial	LRPPRC	1.2106	1.1312
P42765	3-ketoacyl-CoA thiolase, mitochondrial	ACAA2	1.1285	1.0631
P43490	Nicotinamide phosphoribosyltransferase	NAMPT	2.9098	2.6712
P46087	Probable 28S rRNA (cytosine(4447)-C(5))-methyltransferase	NOP2	1.3831	1.3003
P46779	60S ribosomal protein L28	RPL28	1.2528	1.521
P47756-2	F-actin-capping protein subunit beta	CAPZB	1.4801	1.4125
P49368	T-complex protein 1 subunit gamma	CCT3	1.2951	1.1797
P50552	Vasodilator-stimulated phosphoprotein	VASP	1.8538	1.9073
P50991	T-complex protein 1 subunit delta	CCT4	1.4279	1.3845
P51149	Ras-related protein Rab-7a	RAB7A	1.3389	1.2394
P52272-2	Heterogeneous nuclear ribonucleoprotein M	HNRNPM	1.265	1.2124
P52565	Rho GDP-dissociation inhibitor 1	ARHGDIA	1.5477	1.5761
P52566	Rho GDP-dissociation inhibitor 2	ARHGDIB	4.1978	3.841
P52907	F-actin-capping protein subunit alpha-1	CAPZA1	1.8952	2.0101
P53597	Succinyl-CoA ligase [ADP/GDP-forming] subunit alpha, mitochondrial	SUCLG1	1.2718	1.1687
P55259-3	Pancreatic secretory granule membrane major glycoprotein GP2	GP2	1.0199	1.3097
P58546	Myotrophin	MTPN	1.9267	1.818
P59666	Neutrophil defensin 3;HP 3-56;Neutrophil defensin 2	DEFA3	1.1357	1.1736
P59998	Actin-related protein 2/3 complex subunit 4	ARPC4;ARPC4-TTLL3	1.8823	1.8552
P60709	Actin, cytoplasmic 1;Actin, cytoplasmic 1, N-terminally processed	ACTB	1.855	1.7017
P60981	Destrin	DSTN	1.4475	1.3442
P61026	Ras-related protein Rab-10	RAB10	1.0681	1.0409
P61160	Actin-related protein 2	ACTR2	1.2844	1.3552
P61247	40S ribosomal protein S3a	RPS3A	1.2719	1.1798

P61586	Transforming protein RhoA	RHOA	1.0413	1.0336
P61626	Lysozyme C;Lysozyme	LYZ	3.3884	3.4296
P62269	40S ribosomal protein S18	RPS18	1.1029	1.0055
P62328	Thymosin beta-4;Hematopoietic system regulatory peptide	TMSB4X	1.0927	1.0645
P62424	60S ribosomal protein L7a	RPL7A	1.6331	1.5743
P62829	60S ribosomal protein L23	RPL23	1.1773	1.0827
P62942	Peptidyl-prolyl cis-trans isomerase FKBP1A;Peptidyl-prolyl cis-trans isomerase	FKBP1A	1.5813	1.4437
P63104	14-3-3 protein zeta/delta	YWHAZ	1.0276	1.168
P63261	Actin, cytoplasmic 2;Actin, cytoplasmic 2, N-terminally processed	ACTG1	1.2092	1.1979
P68363	Tubulin alpha-1B chain	TUBA1B	1.3006	1.1892
P68871	Hemoglobin subunit beta;LVV-hemorphin-7;Spinorphin	HBB	1.2272	1.1205
P69905	Hemoglobin subunit alpha	HBA1	1.0846	1.0309
P78417	Glutathione S-transferase omega-1	GSTO1	1.1368	1.0931
P84095	Rho-related GTP-binding protein RhoG	RHOG	1.5189	1.5894
Q00325-2	Phosphate carrier protein, mitochondrial	SLC25A3	1.2246	1.149
Q00796	Sorbitol dehydrogenase	SORD	1.0141	1.0515
Q00839-2	Heterogeneous nuclear ribonucleoprotein U	HNRNPU	1.633	1.5724
Q01082	Spectrin beta chain, non-erythrocytic 1	SPTBN1	1.4146	1.2948
Q01518	Adenylyl cyclase-associated protein 1	CAP1	1.9465	1.7708
Q03154	Aminoacylase-1	ACY1;ABHD14A-ACY1	1.1108	1.0458
Q04828	Aldo-keto reductase family 1 member C1	AKR1C1	1.7774	1.7884
Q07075	Glutamyl aminopeptidase	ENPEP	1.3041	1.1921
Q08209-2	Serine/threonine-protein phosphatase 2B catalytic subunit alpha isoform	PPP3CA	1.5138	1.4275
Q08380	Galectin-3-binding protein	LGALS3BP	2.0571	1.9705
Q08752	Peptidyl-prolyl cis-trans isomerase D	PPID	1.8033	1.6561
Q08945	FACT complex subunit SSRP1	SSRP1	2.7081	2.6963
Q12874	Splicing factor 3A subunit 3	SF3A3	1.0953	1.0188
Q13347	Eukaryotic translation initiation factor 3 subunit I	EIF3I	1.3691	1.2624
Q13630	GDP-L-fucose synthase	TSTA3	1.1227	1.1041
Q14103-3	Heterogeneous nuclear ribonucleoprotein D0	HNRNPD	2.3785	2.1638
Q14152	Eukaryotic translation initiation factor 3 subunit A	EIF3A	1.2719	1.4401
Q14444-2	Caprin-1	CAPRIN1	1.1785	1.1797
Q14847	LIM and SH3 domain protein 1	LASP1	1.0958	1.1259

Q14CN2	Calcium-activated chloride channel regulator 4;Calcium-activated chloride channel regulator 4, 110 kDa form;Calcium-activated chloride channel regulator 4, 30 kDa form	CLCA4	1.7376	1.5938
Q15392	Delta(24)-sterol reductase	DHCR24	1.861	1.7035
Q15717	ELAV-like protein 1	ELAVL1	1.4928	1.4923
Q15942	Zyxin	ZYX	1.3806	1.3947
Q16819	Meprin A subunit alpha	MEP1A	1.9989	1.9237
Q16820	Meprin A subunit beta	MEP1B	1.5388	1.5158
Q5JR08	Rho-related GTP-binding protein RhoC	RHOC;RHOA	1.114	1.0579
Q5VWC4	26S proteasome non-ATPase regulatory subunit 4	PSMD4	1.6723	1.591
Q6XQN6	Nicotinate phosphoribosyltransferase	NAPRT	1.7374	1.7262
Q7L2H7	Eukaryotic translation initiation factor 3 subunit M	EIF3M	1.5944	1.4618
Q86UX7-2	Fermitin family homolog 3	FERMT3	2.2232	2.0239
Q93009	Ubiquitin carboxyl-terminal hydrolase 7;Ubiquitin carboxyl-terminal hydrolase	USP7	1.3519	1.2304
Q96C19	EF-hand domain-containing protein D2	EFHD2	1.8751	1.7399
Q96G03	Phosphoglucomutase-2	PGM2	1.4695	1.351
Q99536	Synaptic vesicle membrane protein VAT-1 homolog	VAT1	1.6632	1.5243
Q99729-3	Heterogeneous nuclear ribonucleoprotein A/B	HNRNPAB	1.1126	1.1953
Q9BRF8	Serine/threonine-protein phosphatase CPPED1	CPPED1	3.2827	2.9872
Q9BVC6	Transmembrane protein 109	TMEM109	1.2684	1.1898
Q9BYE9	Cadherin-related family member 2	CDHR2	1.1511	1.1862
Q9BZZ5-2	Apoptosis inhibitor 5	API5	1.2193	1.3007
Q9HAT2	Sialate O-acetylesterase	SIAE	1.3015	1.2151
Q9NUQ9	Protein FAM49B	FAM49B	1.8544	1.692
Q9UHX1-6	Poly(U)-binding-splicing factor PUF60	PUF60	1.0709	1.0011
Q9Y266	Nuclear migration protein nudC	NUDC	1.7004	1.5489
Q9Y5B9	FACT complex subunit SPT16	SUPT16H	1.6893	1.545
Q9Y6R7	IgGFc-binding protein	FCGBP	2.3875	2.2967

Appendix 2

Table 1. Discriminant features of IBD EVs compared to non-IBD control EVs identified by PLS-DA

Uniprot ID	Protein name	Gene name	VIP Comp. 1
A0A024R4M0	40S ribosomal protein S9	RPS9	1.3775
A0A075B6N7	Ig alpha-2 chain C region (Fragment)	IGHA2	1.8066
A0A075B6N8	Ig gamma-3 chain C region (Fragment)	IGHG3	1.9225
A0A087WVI4	Transmembrane channel-like protein	TMC4	1.1075
A0A087WVQ6	Clathrin heavy chain	CLTC	1.7524
A0A087WWU8	Tropomyosin alpha-3 chain	TPM3	1.5044
A0A087WZW1	Colipase	CLPS	1.1697
A0A0A0MS51	Gelsolin	GSN	3.2269
A0A0A0MSA6	Chymotrypsin-like elastase family member 3B (Fragment)	CELA3B	1.0945
A0A0C4DFQ8	Neutral ceramidase	ASAH2	1.5167
A0A0C4DGG6	Niemann-Pick C1-like protein 1	NPC1L1	1.0639
A0A0G2JMB2	Ig alpha-2 chain C region (Fragment)	IGHA2	1.0507
A0A0G2JP44	Cadherin-related family member 5	CDHR5	1.5346
A0A0G2JPR0	Complement C4-A	C4A	2.1659
A0A0J9YVY4	Disks large-associated protein 2 (Fragment)	DLGAP2	1.0565
A0A0J9YVZ6	Unconventional myosin-XVB	MYO15B	1.2593
A0A0J9YY01	Unconventional myosin-XVB	MYO15B	2.0666
A0PJZ3	Glucoside xylosyltransferase 2	GXYLT2	1.1811
A4D1E9	GTP-binding protein 10	GTPBA	1.0618
A6NNI4	Tetraspanin	CD9	1.031
A8K7I4	Calcium-activated chloride channel regulator 1	CLCA1	1.0681
B7ZKQ9	SCARB1 protein	SCARB1	1.4051
D6RB89	Retinol-binding protein 2 (Fragment)	RBP2	1.1516
D6RF35	Vitamin D-binding protein	GC	1.2753
E7EU05	Platelet glycoprotein 4 (Fragment)	CD36	1.5436
E9PF55	Tensin-1	TNS1	1.5172
F5H6E2	Unconventional myosin-Ic	MYO1C	1.0873
F8W062	Phospholipase A(2)	PLA2G1B	1.2822
F8W1R7	Myosin light polypeptide 6	MYL6	1.1962
F8WEC7	Zinc finger protein 473	ZNF473	1.2048
H3BMY1	Chymotrypsinogen B2 (Fragment)	CTRB2	1.0651
H3BVC8	Prostasin (Fragment)	PRSS8	1.009
H7BYV1	Interferon-induced transmembrane protein 2 (Fragment)	IFITM2	1.0971

K7ENI6	Uncharacterized protein; TMEM256-PLSCR3 readthrough (NMD candidate)	TMEM256-PLSCR3	1.003
O14745	Na(+)/H(+) exchange regulatory cofactor NHE-RF1	NHRF1	1.4886
O14894	Transmembrane 4 L6 family member 5	T4S5	1.7518
O15143	Actin-related protein 2/3 complex subunit 1B	ARC1B	2.3715
O15144	Actin-related protein 2/3 complex subunit 2	ARPC2	1.8011
O15212	Prefoldin subunit 6	PFD6	1.1934
O15484	Calpain-5	CAN5	1.1133
O43280	Trehalase	TREA	1.177
O43760	Synaptogyrin-2	SNG2	1.104
O43895	Xaa-Pro aminopeptidase 2	XPP2	1.6191
O60218	Aldo-keto reductase family 1 member B10	AK1BA	1.4042
O60234	Glia maturation factor gamma	GMFG	1.0867
O60488	Long-chain-fatty-acid--CoA ligase 4	ACSL4	1.2764
O94832	Unconventional myosin-Id	MYO1D	1.7646
O95926	Pre-mRNA-splicing factor SYF2	SYF2	1.0474
P00338	L-lactate dehydrogenase A chain	LDHA	1.452
P00558	Phosphoglycerate kinase 1	PGK1	1.0154
P00738	Haptoglobin	HPT	1.9878
P01023	Alpha-2-macroglobulin	A2MG	2.8462
P01024	Complement C3	CO3	2.8236
P01857	Ig gamma-1 chain C region	IGHG1	1.2063
P01859	Ig gamma-2 chain C region	IGHG2	2.2563
P02647	Apolipoprotein A-I	APOA1	1.7683
P02671	Fibrinogen alpha chain	FIBA	1.2568
P02675	Fibrinogen beta chain	FIBB	3.5854
P02679	Fibrinogen gamma chain	FIBG	3.3887
P02748	Complement component C9	CO9	1.2021
P02768	Serum albumin	ALBU	2.0228
P02787	Serotransferrin	TRFE	2.1776
P02788	Lactotransferrin	TRFL	3.7163
P04004	Vitronectin	VTNC	1.3614
P04083	Annexin A1	ANXA1	2.0669
P04114	Apolipoprotein B-100	APOB	3.6976
P05107	Integrin beta-2	ITB2	3.7647
P05109	Protein S100-A8	S10A8	4.645
P05164	Myeloperoxidase	PERM	2.368
P06702	Protein S100-A9	S10A9	2.6459
P06733	Alpha-enolase	ENOA	1.0416
P07195	L-lactate dehydrogenase B chain	LDHB	1.8821
P07900	Heat shock protein HSP 90-alpha	HS90A	1.1967
P08133	Annexin A6	ANXA6	1.7008

P08183	Multidrug resistance protein 1	MDR1	1.5522
P08217	Chymotrypsin-like elastase family member 2A	CEL2A	1.8012
P08246	Neutrophil elastase	ELNE	3.8066
P08311	Cathepsin G	CATG	3.1443
P08473	Neprilysin	NEP	2.1209
P08590	Myosin light chain 3	MYL3	1.2279
P08603	Complement factor H	CFAH	1.9093
P09093	Chymotrypsin-like elastase family member 3A	CEL3A	1.3586
P09210	Glutathione S-transferase A2	GSTA2	1.5101
P09327	Villin-1	VILI	1.4732
P09848	Lactase-phlorizin hydrolase	LPH	3.0575
P09923	Intestinal-type alkaline phosphatase	PPBI	1.5342
P09960	Leukotriene A-4 hydrolase	LKHA4	1.9724
P11215	Integrin alpha-M	ITAM	2.14
P11413	Glucose-6-phosphate 1-dehydrogenase	G6PD	1.4423
P11678	Eosinophil peroxidase	PERE	1.419
P12277	Creatine kinase B-type	KCRB	1.3359
P12429	Annexin A3	ANXA3	1.5462
P12724	Eosinophil cationic protein	ECP	1.8809
P13639	Elongation factor 2	EF2	1.1382
P13796	Plastin-2	PLSL	1.5314
P13866	Sodium/glucose cotransporter 1	SC5A1	1.3588
P14618	Pyruvate kinase PKM	KPYM	2.9786
P14780	Matrix metalloproteinase-9	MMP9	2.1123
P15086	Carboxypeptidase B	CBPB1	1.2271
P15144	Aminopeptidase N	AMPN	1.0118
P16444	Dipeptidase 1	DPEP1	1.1823
P17987	T-complex protein 1 subunit alpha	TCPA	1.7924
P18206	Vinculin	VINC	1.2811
P19105	Myosin regulatory light chain 12A	ML12A	2.0705
P20160	Azurocidin	CAP7	3.4905
P22748	Carbonic anhydrase 4	CAH4	1.0961
P27216	Annexin A13	ANX13	1.6068
P27487	Dipeptidyl peptidase 4	DPP4	1.2799
P27797	Calreticulin	CALR	1.0394
P29144	Tripeptidyl-peptidase 2	TPP2	1.2811
P29401	Transketolase	TKT	1.9197
P29992	Guanine nucleotide-binding protein subunit alpha-11	GNA11	1.1803
P30086	Phosphatidylethanolamine-binding protein 1	PEBP1	1.0492
P31949	Protein S100-A11	S10AB	1.1723
P32119	Peroxiredoxin-2	PRDX2	1.0019

P32320	Cytidine deaminase	CDD	2.0003
P35579	Myosin-9	MYH9	2.8122
P36871	Phosphoglucomutase-1	PGM1	1.0059
P40879	Chloride anion exchanger	S26A3	2.1971
P42685	Tyrosine-protein kinase FRK	FRK	1.4759
P46059	Solute carrier family 15 member 1	S15A1	1.3841
P49913	Cathelicidin antimicrobial peptide	CAMP	3.7153
P52790	Hexokinase-3	HXK3	2.0237
P55060	Exportin-2	XPO2	2.8896
P59666	Neutrophil defensin 3	DEF3	2.763
P59998	Actin-related protein 2/3 complex subunit 4	ARPC4	2.1463
P60174	Triosephosphate isomerase	TPIS	1.0954
P61158	Actin-related protein 3	ARP3	2.4544
P61160	Actin-related protein 2	ARP2	2.6029
P61626	Lysozyme C	LYSC	1.1212
P62070	Ras-related protein R-Ras2	RRAS2	1.1911
P62081	40S ribosomal protein S7	RS7	1.1347
P62195	26S protease regulatory subunit 8	PRS8	1.0681
P62805	Histone H4	H4	2.4964
P62807	Histone H2B type 1-C/E/F/G/I	H2B1C	3.4552
P62826	GTP-binding nuclear protein Ran	RAN	1.0548
P68871	Hemoglobin subunit beta	HBB	2.6865
P69905	Hemoglobin subunit alpha	HBA	3.7512
Q01518	Adenylyl cyclase-associated protein 1	CAP1	1.4494
Q01523	Defensin-5	DEF5	1.0219
Q02505	Mucin-3A	MUC3A	1.0861
Q04917	14-3-3 protein eta	1433F	1.088
Q07075	Glutamyl aminopeptidase	AMPE	2.0545
Q13228	Selenium-binding protein 1	SBP1	1.1208
Q14002	Carcinoembryonic antigen-related cell adhesion molecule 7	CEAM7	1.1807
Q14204	Cytoplasmic dynein 1 heavy chain 1	DYHC1	1.3205
Q14764	Major vault protein	MVP	2.988
Q14CN2	Calcium-activated chloride channel regulator 4	CLCA4	1.2061
Q16539	Mitogen-activated protein kinase 14	MK14	1.3584
Q16820	Meprin A subunit beta	MEP1B	1.1826
Q16851	UTP--glucose-1-phosphate uridylyltransferase	UGPA	1.6658
Q53GD3	Choline transporter-like protein 4	CTL4	1.0374
Q60FE5	Filamin A	Q60FE5	2.6828
Q6P2I7	Endogenous Bornavirus-like nucleoprotein 2	EBLN2	1.2989
Q6PIF6	Unconventional myosin-VIIb	MYO7B	1.0832
Q6UXY8	Transmembrane channel-like protein 5	TMC5	1.2874

Q71U36	Tubulin alpha-1A chain	TBA1A	1.4805
Q86U38	Nucleolar protein 9	NOP9	1.3708
Q86UP6	CUB and zona pellucida-like domain-containing protein 1	CUZD1	1.8059
Q86VP6	Cullin-associated NEDD8-dissociated protein 1	CAND1	1.0619
Q8IVL8	Carboxypeptidase O	CBPO	1.5536
Q8NFJ5	Retinoic acid-induced protein 3	RAI3	1.0391
Q8TE67	Epidermal growth factor receptor kinase substrate 8-like protein 3	ES8L3	1.0202
Q92485	Acid sphingomyelinase-like phosphodiesterase 3b	ASM3B	1.3126
Q92626	Peroxidasin homolog	PXDN	1.5618
Q96K78	Adhesion G-protein coupled receptor G7	AGRG7	1.0904
Q96NT5	Proton-coupled folate transporter	PCFT	1.2652
Q99878	Histone H2A type 1-J	H2A1J	2.5677
Q99895	Chymotrypsin-C	CTRC	1.2827
Q9BYF1	Angiotensin-converting enzyme 2	ACE2	1.1282
Q9H444	Charged multivesicular body protein 4b	CHM4B	1.063
Q9H4A4	Aminopeptidase B	AMPB	1.0017
Q9H4H8	Protein FAM83D	FA83D	1.2052
Q9H6S3	Epidermal growth factor receptor kinase substrate 8-like protein 2	ES8L2	1.3067
Q9HBH0	Rho-related GTP-binding protein RhoF	RHOF	1.1872
Q9HD89	Resistin	RETN	1.0241
Q9NZ20	Group 3 secretory phospholipase A2	PA2G3	1.6807
Q9UBC5	Unconventional myosin-Ia	MYO1A	1.0582
Q9UKN1	Mucin-12	MUC12	1.4811
Q9UNQ0	ATP-binding cassette sub-family G member 2	ABCG2	1.1691
Q9UQQ1	N-acetylated-alpha-linked acidic dipeptidase-like protein	NALDL	1.3256
Q9Y678	Coatomer subunit gamma-1	COPG1	1.3093
X6RAN8	Dual oxidase 2	X6RAN8	1.3171

CURRICULUM VITAE

Academic and Training Background

Ph.D. Biochemistry **2014-2018**
Biomarker discovery and extracellular vesicle proteomic signatures in pediatric inflammatory bowel disease
Supervisors: Dr. Daniel Figeys & Dr. Alain Stintzi
Faculty of Medicine
University of Ottawa, Ontario, Canada

M.Sc. Biochemistry **2008-2011**
Faculty of Medicine, University of Ottawa Heart Institute, Ontario, Canada

Baccalauréat science spécialisé en biochimie **2004-2008**
Faculty of Science
University of Ottawa, Ontario, Canada

Publications

1. **Metaproteomics reveals associations between microbiome and intestinal extracellular vesicle proteins in pediatric inflammatory bowel disease.** *Nat Comm* **2018**. Xu Zhang,* **Shelley A Deeke**,* Zhibin Ning, Amanda E Starr, James Butcher, Jennifer Li, Janice Mayne, Kai Cheng, Bo Liao, Leyuan Li, Ruth Singleton, David Mack, Alain Stintzi, Daniel Figeys

2. **Mucosal-luminal interface proteomics reveals biomarkers of pediatric inflammatory bowel disease-associated colitis**. *American J Gastroenterol* **2018**. **Deeke, S. A**.*; Starr, A. E.*; Ning, Z.; Ahmadi, S.; Zhang, X.; Mayne, J.; Chiang, C. K.; Singleton, R.; Benchimol, E.I.; Mack, D.R.; Stintzi, A.; Figeys, D.

3. **Proteomic and Metaproteomic Approaches to Understand Host-Microbe Interactions**. *Anal Chem* **2017**. Starr, A. E.; **Deeke, S. A.**; Li, L.; Zhang, X.; Daoud, R.; Ryan, J.; Ning, Z.; Cheng, K.; Nguyen, L. V. H.; Abou-Samra, E.; Lavallee-Adam, M.; Figeys, D.

4. **Proteomic analysis of ascending colon biopsies from a paediatric inflammatory bowel disease inception cohort identifies protein biomarkers that differentiate Crohn's disease from UC**. *Gut* **2017,** *66* (9), 1573-1583. Starr, A. E.; **Deeke, S. A**.; Ning, Z.; Chiang, C. K.; Zhang, X.; Mottawea, W.; Singleton, R.; Benchimol, E. I.; Wen, M.; Mack, D. R.; Stintzi, A.; Figeys, D.

5. **Altered intestinal microbiota-host mitochondria crosstalk in new onset Crohn's disease**. *Nat Commun* **2016,** *7*, 13419. Mottawea, W.; Chiang, C. K.; Muhlbauer, M.; Starr, A. E.; Butcher, J.; Abujamel, T.; **Deeke, S. A**.; Brandel, A.; Zhou, H.; Shokralla, S.;

Hajibabaei, M.; Singleton, R.; Benchimol, E. I.; Jobin, C.; Mack, D. R.; Figeys, D.; Stintzi, A.

6. **MetaPro-IQ: a universal metaproteomic approach to studying human and mouse gut microbiota**. *Microbiome* **2016,** *4* (1), 31. Zhang, X.; Ning, Z.; Mayne, J.; Moore, J. I.; Li, J.; Butcher, J.; **Deeke, S. A.**; Chen, R.; Chiang, C. K.; Wen, M.; Mack, D.; Stintzi, A.; Figeys, D.

7. **In Vitro Metabolic Labeling of Intestinal Microbiota for Quantitative Metaproteomics.** *Anal Chem* **2016,** *88* (12), 6120-5. Zhang, X.; Ning, Z.; Mayne, J.; **Deeke, S. A.**; Li, J.; Starr, A. E.; Chen, R.; Singleton, R.; Butcher, J.; Mack, D. R.; Stintzi, A.; Figeys, D.

8. **Bottom-Up Proteomics (2013-2015): Keeping up in the Era of Systems Biology**. *Anal Chem* **2016,** *88* (1), 95-121. Mayne, J.; Ning, Z.; Zhang, X.; Starr, A. E.; Chen, R.; **Deeke, S.**; Chiang, C. K.; Xu, B.; Wen, M.; Cheng, K.; Seebun, D.; Star, A.; Moore, J. I.; Figeys, D.

9. **Identification of a novel muscle A-type lamin-interacting protein (MLIP).** *J Biol Chem* **2011,** *286* (22), 19702-13. Ahmady, E.; **Deeke, S. A.**; Rabaa, S.; Kouri, L.; Kenney, L.; Stewart, A. F.; Burgon, P. G.

10. **IRF2BP2 is a skeletal and cardiac muscle-enriched ischemia-inducible activator of VEGFA expression**. *FASEB J* **2010,** *24* (12), 4825-34. Teng, A. C.; Kuraitis, D.; **Deeke, S. A.**; Ahmadi, A.; Dugan, S. G.; Cheng, B. L.; Crowson, M. G.; Burgon, P. G.; Suuronen, E. J.; Chen, H. H.; Stewart, A. F.

Oral presentations

1. **Deeke SA**, Starr AE, Ning Z, Ahmadi S, Zhang X, Mayne J, Chiang C-K, Singleton R, Benchimol EI, Mack D, Stintzi A and Figeys D. (2018). Proteomics of intestinal aspirates identifies biomarkers for pediatric inflammatory bowel disease diagnosis. Canadian National Proteomics Networks, 10th Annual Symposium, Vancouver BC.

2. **Deeke SA**, Starr A, Zhang X, Mottawea W, Singleton R, Benchimol E, Mack D, Stintzi A, Figeys D. (2015). Host protein biogeographical alterations at the intestinal mucosal-luminal interface in new-onset pediatric inflammatory bowel disease. Systems Biology of infection symposium, 2nd Edition, Ascona, Switzerland

Poster presentations

1. **Deeke SA**, Starr AE, Ning Z, Ahmadi S, Zhang X, Mayne J, Chiang C-K, Singleton R, Benchimol EI, Mack D, Stintzi A and Figeys D. (2017). Mucosal-luminal interface proteomics reveals biomarkers of pediatric inflammatory bowel disease-associated colitis. Poster Presentation. 16th Human Proteome Organization World Congress, Dublin, Ireland.

2. **Deeke SA**, Starr AE, Singleton R, Mayne J, Benchimol EI, Mack D, Stintzi A and Figeys D. (2016). Proteomic profiling of intestinal extracellular vesicle isolated from new-onset pediatric inflammatory bowel disease. Gordon Research Conference on Extracellular Vesicles, Sunday River, Newry, ME, United States.

3. **Deeke SA**, Starr A, Mack D, Stintzi A, Figeys D. (2015). Characterizing the host and microbial proteome at the mucosal-luminal interface in new-onset pediatric inflammatory bowel disease. Ottawa Institute of Systems Biology retreat, Mont-Tremblant, Canada

4. **Deeke SA**, Starr A, Mack D, Stintzi A, Figeys D. (2014). Characterizing the host and microbial proteome at the mucosal-luminal interface in new-onset pediatric inflammatory bowel disase. 7th Annual CHEO Research Day, Ottawa, Canada

5. **Deeke SA**, Starr A, Mack D, Stintzi A , Figeys D. (2014). Characterizing the host and microbial proteome at the mucosal-luminal interface in new-onset pediatric inflammatory bowel disease. 13th Human Proteome Organization World Congress, Madrid, Spain

6. **Deeke SA**, Mayne J, Ning Z, Steele S, Wesolowski M, Chrétien M and Figeys D. (2012). PCSK9 Membrane Systems Biology. 10th Gordon Research Conference, Proprotein Processing, Trafficking & Secretion, New London, United States

7. **Deeke SA**, Mayne J, Ning Z, Steele S, Wesolowski M, Chrétien M and Figeys D. (2012). PCSK9 Membrane Systems Biology. 2nd China-Canada Systems Biology Conference, Ottawa, Canada

8. **Deeke SA**, Burgon PG, Stewart AFR. (2010). Developmental and molecular bases of LITF in cell fate determination. The Ottawa Conference on New Directions in Biology and Disease of Skeletal Muscle, Ottawa, Canada

Work Experience

Student, Science Strategy, Canadian Institutes of Health Research **March –August 2018**

- Collection of evidence to inform on policy, devising solutions and recommendations.

Mass Spectrometry Facility Assistant **2014- 2018**
Ottawa Institute of Systems Biology

- Operating and maintaining hybrid high performance mass spectrometers
    - Troubleshooting running conditions, calibration of instrument
    - Maintaining instrument activity log of MS utilization and status
    - Preparing analytical columns for MS analysis.
- Preparation of clients samples, mass spectrometry analysis and database search

Teaching Assistant, Lab Demonstrator **2015-2018**
University of Ottawa

- Guide undergrad students through experimental workflow and assess quality of their results. Correction of laboratory reports.

Laboratory Research Assistant **2011-2014**
University of Ottawa, Ottawa Institute of Systems Biology

- Proteomic analysis of mucosal biopsies from patients with inflammatory bowel disease
- Investigating protein changes caused by gain-of-function and loss-of-function PCSK9 mutants (regulator of plasma LDL cholesterol levels) in cell culture by shotgun proteomics.

Customer Service representative, part-time **2007-2012**
Gamma-Dynacare Medical Laboratories

- Contacting clients to report STAT, critical and abnormal results while maintaining confidentiality and accuracy of appropriate documentation.
- Answering inquiries from clients and Patient Service Centres.
- Solving and documenting incoming problems.

Select Awards and Honors

Award, 3 minute thesis oral presentation, 2nd place	**2018**
Award, 1st place, seminar presentation, BMI, University of Ottawa	**2018**
Award, 1st place poster presentation, BMI, University of Ottawa	**2017**
Leadership award - Graduate Studies	**2016**
Travel grant - Office of Graduate Studies	**2017, 2018**
Ontario Graduate Scholarship (OGS)	**2016-2018**
Ph.D. Excellence Scholarship, University of Ottawa	**2016-2018**
Biochemistry Graduate Program Travel Grant	**2009, 2014**
FGPS Travel award	**2015**
Dean's Scholarship, Faculty of Graduate and Postdoctoral Studies	**2011**
M.Sc. Admission scholarship, University of Ottawa	**2008-2010**
Graduated cum laude, University of Ottawa	**2008**
Dean's list, University of Ottawa	**2007-2008**
Fellowship for Studying in French scholarship	**2004-2005**

Leadership/Science Outreach

Volunteer, Science Outreach **2016, 2017**
IBD Foundation Youth Gut Together

- Educating, inspiring and empowering youth with insight into the inflammatory bowel disease research laboratory facilities.
- Involved in planning experiments to engage children and youth in hands-on activities regarding research directly related to the study in which they are enrolled.

VP Academic, Representative of BMIGSA **2015-2017**

(Biochemistry, microbiology & immunology graduate student association)

- Represent the graduate student body and ensure that the academic needs of students are effectively conveyed to the faculty of medicine graduate studies council in order to provide the best experience for student during their studies.
- Organize academic and professional events for graduate students.

Volunteer, Science Outreach **2010**

Elementary school outreach program - Brain Awareness Week, The Dana Foundation

Increase public awareness of the progress and benefits of brain research through an interactive presentation to elementary school students.

RIGHTS AND PERMISSIONS

For Deeke, S.A., A.E. Starr, Z. Ning, S. Ahmadi, X. Zhang, J. Mayne, C.K. Chiang, R. Singleton, E.I. Benchimol, D.R. Mack, A. Stintzi, and D. Figeys. 2018. Mucosal-luminal interface proteomics reveals biomarkers of pediatric inflammatory bowel disease-associated colitis. *Am J Gastroenterol*

Ownership of copyright in in original research articles remains with the Author, and provided that, when reproducing the contribution or extracts from it or from the Supplementary Information, the Author acknowledges first and reference publication in the Journal, the Author retains the following non-exclusive rights:

a) To reproduce the contribution in whole or in part in any printed volume (book or thesis) of which they are the author(s).

b) The author and any academic institution where they work at the time may reproduce the contribution for the purpose of course teaching.

c) To reuse figures or tables created by the Author and contained in the Contribution in oral presentations and other works created by them.

d) To post a copy of the contribution as accepted for publication after peer review (in locked Word processing file, of a PDF version thereof) on the Author's own web site, or the Author's institutional repository, or the Author's funding body's archive, six months after publication of the printed or online edition of the Journal, provided that they also link to the contribution on the publisher's website.

The above use of the term 'Contribution' refers to the author's own version, not the final version as published in the Journal.

Thesis reuse applies to both print and electronic reuse.

For Zhang, X., S.A. Deeke, Z.B. Ning, A.E. Starr, J. Butcher, J. Li, J. Mayne, K. Cheng, B. Liao, L.Y. Li, R. Singleton, D. Mack, A. Stintzi, and D. Figeys. 2018a. Metaproteomics reveals associations between microbiome and intestinal extracellular vesicle proteins in pediatric inflammatory bowel disease. ***Nature Communications***

This work is licensed under a Creative Commons Attribution 4.0 International License, which permits unrestricted use, distribution, and reproduction in any medium provided the appropriate credit to the original author(s) and the source is given, a link provided to the Creative Commons license, and changes are indicated. **You are not required to obtain permission to reuse this article.** The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.