METABOLOMICS APPROACH FOR GAINING INSIGHTS INTO PATHOLOGICAL MECHANISMS OF IRRITABLE BOWEL SYNDROME AND INFLAMMATORY BOWEL DISEASE

METABOLOMICS APPROACH FOR GAINING INSIGHTS INTO PATHOLOGICAL MECHANISMS OF IRRITABLE BOWEL SYNDROME AND INFLAMMATORY BOWEL DISEASE

By MAI YAMAMOTO, B.Sc.

A Thesis Submitted to the School of Graduate Studies in Partial Fulfilment of the Requirements for the Degree

Doctor of Philosophy

December 2018

DOCTOR OF PHILOSOPHY (2018)   McMaster University

(Chemistry and Chemical Biology)   Hamilton, ON

TITLE: Metabolomics approach for gaining insights into pathological

                    mechanisms of irritable bowel syndrome and inflammatory

                    bowel disease

AUTHOR: Mai Yamamoto, BSc.   (Vancouver Island University)

SUPERVISOR: Professor Dr. Philip Britz-McKibbin

PAGES: xix, 288

Abstract

Irritable bowel syndrome (IBS) and inflammatory bowel disease (IBD) are two of the most commonly diagnosed chronic digestive disorders in Western countries with increasing prevalence among Canadians. However, the etiology of IBS and IBD remain poorly understood due to a complex interplay of genetic, psychosocial and environmental factors, which hampers efforts at early detection/screening, accurate diagnosis and effective treatments notably in children. This thesis aims to reveal new biochemical insights into the pathophysiology underlying IBS and IBD when using an untargeted metabolite profiling (*i.e.,* metabolomics) approach on urine and stool specimens based on multisegment injection-capillary electrophoresis-mass spectrometry (MSI-CE-MS). *Chapter I* reviews brief history and current challenges in diagnosis and treatment, as well as current metabolomics literature of IBS and IBD. *Chapter II* first develops a robust method for high throughput profiling of anionic metabolites in human urine samples when using MSI-CE-MS. For the first time, we demonstrate that incidental capillary fractures are caused by irreversible aminolysis of the outer polyimide coating due to the frequent use of volatile ammonia based buffers under alkaline conditions (pH > 9) in electrospray ionization-MS. *Chapter III* subsequently applies this validated method to investigate differentially excreted urinary metabolites between adult IBS patients and healthy controls, which indicated significantly accelerated rates of collagen degradation and cell turn-over in IBS patients. *Chapter IV* later develops a novel stool extraction protocol for characterization of the fecal metabolome together with meta-genomic data for elucidating complex host-gut microflora interactions from a cohort of pediatric IBD patients, including Crohn's disease and ulcerative colitis. In this pilot study, a panel of discriminating metabolites in urine is shown to allow for differential diagnosis of major pediatric IBD sub-types as an alternative to colonoscopy and histopathology that are invasive, expensive and prone to ambiguous test results. Finally, *Chapter V* involves a longitudinal metabolomics study that aims to identify metabolic trajectories that predict treatment responses of a cohort of pediatric Crohn's disease patients following initiation of exclusive enteral nutrition (EEN) therapy. In the end, *Chapter VI* highlights major outcomes of thesis and future direction of metabolomics in IBS and IBD with a specific focus on improved stool specimen collection and validation of biomarker specificity relative to other related gastrointestinal disorders. In summary, this thesis has demonstrated metabolic processes that are associated with exacerbation of symptoms or remission in subset of IBS and pediatric IBD patients. With follow up studies with larger cohort of patients, potential biomarkers identified in this thesis will contribute the development of more accurate and non-invasive decision making process for diagnosis and treatment, resulting in long-lasting remission and improved quality of life of patients suffering from chronic digestive disorders.

Acknowledgements

First of all, I would like to thank my supervisor, Dr. Philip Britz-McKibbin for his constructive feedback and supervision. His boundless optimism and enthusiasm in scientific endeavour was always helpful whenever problems arose. I am thankful of his passion in rigorously ensuring the quality of each step toward the final data and having a faith in me that I can finish this thesis. I have learnt many skills and acquired in-depth technical knowledge during the 5 years, which will for sure help me thrive in the next step in my career.

I have been extremely fortunate to have supportive, smart and hard-working lab members who cheered me up when I was down and kept me going. All of you made this Ph.D. process far more enjoyable. Thank you for all your technical and mental support.

I would like to thank all of my collaborators involved in studies described in this thesis for their precious time and constructive feedback. It has been a great learning experience to be involved in clinical studies with an access to samples of patients. Thank you for your patience whenever I needed extra information and clarification on medical aspects in the studies. Committee members, Dr. Potter, Dr. Moran-Mirabal, and Dr. Pai have also given me constructive feedback from technical and medical perspectives and always reminded me of aspects in studies that I had overlooked.

I also would like to thank professors who taught me the joy and excitement of biology and chemistry in my undergraduate study at Vancouver Island University. I have never forgotten their enthusiastic eyes when they talk about what they love to teach and learn. Especially Dr. Timothy Goater, when he talks about the amazing diversity of parasites and complex interaction with hosts for survival, Dr. Caroline Josefsson when she talks about plants and their biochemistry, Dr. Erik Krogh when he talks about chemical reactions in nature and how to better educate community members about climate change, and Dr. Chris Gill, my undergraduate thesis supervisor, when he talks about mass spectrometry. I was infected with their passion and curiosity, which kept me going to this date.

Table of Contents

**Chapter V: Elucidating Treatment Responses of Pediatric Crohn's Disease
Patients Following Exclusive Enteral Nutrition Therapy: A Longitudinal
Metabolomics Study**…………..………………………………………………… 210

List of Figures

Supporting Figures

List of Tables

Supporting Tables

List of Abbreviations and Symbols

$[M+H]^+$	Protonated molecule
$[M-H]^-$	Deprotonated molecule
μ_0	Absolute electrophoretic mobility
z_0	Effective charge
^{1}H-NMR	Proton nuclear magnetic resonance
AHR	Aryl hydrocarbon receptor
ANCOVA	Analysis of covaraince
Asp-Gln	Aspartyl-glutamine
AUC	Area under the curve
BGE	Background electrolyte
CASMI	Critical assessment of small molecule identification
CD	Crohn's disease
CEC	Electrochromatography
CE-MS	Capillary electrophoresis-mass spectrometry
CI	Confidence interval
CID	Collision induced dissociation
Cl-Tyr	3-Chrolo-L-tyrosine
CNS	Central nervous system
CRP	C-reactive protein
CV	Coefficient of variance
EDTA	Ethylenediaminetetraacetic acid
EEN	Exclusive enteral nutrition
EIE	Extracted ion electropherogram
EI-MS	Electron impact ionization mass spectrometry

ELISA	Enzyme-linked immunosorbent assay
EOF	Electroosmotic flow
ESI	Electrospray ionization
ESI-MS	Electrospray ionization-mass spectrometry
ESR	Erythrocyte sedimentation rate
FAIMS	Field asymmetric ion mobility spectrometry
FC	Fold-change
FCP	Fecal calprotectin
FDR	False discovery rate
FMT	Fecal microbiome transplantation
FODMAP	Fermentable oligosaccharides, disaccharides, monosaccharides and polyol
F-Phe	4-Fluoro-L-phenylalanine
Gal-OHLys	O-Galactosyl hydroxylysine
GC	Gas chromatography
GC-MS	Gas chromatography-mass spectrometry
GI	Gastrointestinal
Glc-Gal_OHLys	O-Glucosyl-galactosyl-hydroxylysine
Gln	L-Glutamine
GWAS	Genome-wide association studies
HADS-A	Hospital anxiety and depression scale - anxiety
HADS-D	Hospital anxiety and depression scale - depression
HC	Healthy control
HGPRT	Hypoxanthine-guanine phosphoribosyltransferase
HILIC	Hydrophilic interaction chromatography
HMDB	Human Metabolome Database
IBD	Inflammatory bowel disease
IBS	Irritable bowel syndrome
IBS-C	Irritable bowel syndrome-constipation dominant
IBS-D	Irritable bowel syndrome-diarrhea-dominant
IBS-M	Irritable bowel syndrome-mixed symptoms
ID	Inner diameter
IDO	Indoleamine 2,3-dioxygenase
IMS	Ion mobility spectrometry
KEGG	Kyoto encyclopedia of genes and genomes
KO	KEGG orthologous groups
LC	Liquid chromatography
LC-MS	Liquid chromatography-mass spectrometry
LEfSe	Linear discriminant analysis effect size
Lys	L-Lysine
MALDI-TOF-MS	Matrix-assisted laser desorption time of flight mass spectrometry

Man-Trp	C-Mannopyranosyl-tryptophan
MeOH	Methanol
MS/MS	Tandem mass spectrometry
MSI	Multisegment injection
MSI-CE-MS	Multisegment injection capillary electrophoresis-mass spectrometry
MV	Molecular volume
NIH	National institute of health
NMS	2-Naphthalinesulfonate
NSAIDs	Non-steroidal anti-inflammatory drugs
OD	Outer diameter
OG	Octanoyl-glucuronide
OPLS-DA	Orthogonal partial least squares-discriminant analysis
Orn	L-Ornithine
OTU	Operational taxonomic unit
PCA	Principal component analysis
PCDAI	Pediatric Crohn's disease activity index
PCR	Polymerase chain reaction
PCs	Principal components
PICRUSt	Phylogenetic investigation of communities by reconstruction of unobserved states
PLS	Partial least square
PLS-DA	Partial least square discriminant analysis
PUCAI	Pediatric ulcerative colitis activity index
PURE	Prospective Urban and Rural Epidemiological
QC/QA	Quality control/quality assuarance
Q-TOF	quadrupole time of flight
RMT	Relative migration time
RPA	Relative peak area
R-square	Coefficient of determination
S/N	Signal-to-noise ratio
SCFA	Short chain fatty acid
SD	Standard deviation
Ser	L-Serine
TCA	Tricyclic acid
TEA	Triethylamine
TNF-alpha	Tumor necrosis factor-alpha
TSP	Standard polyimide
TSU	Polytetrafluorinated
UC	Ulcerative colitis
UGI	Upper gastrointestinal

UV	Ultraviolet
VOCs	Volatile organic compounds
Xyl-Ser	Xylosyl-serine

Declaration of Academic Achievement

The following material has been previously published and is reprinted with written permission:

Chapter II. Reprinted and adapted from Yamamoto, M.; Ly, R.; Gill, B.;Zhu, Y.; Moran-Mirabal, J.; and Britz-McKibbin, P. Robust and High-Throughput Method for Anionic Metabolite Profiling: Preventing Polyimide Aminolysis and Capillary Breakages under Alkaline Conditions in Capillary Electrophoresis-Mass Spectrometry *Anal. Chem.* **88,** 10710-10719, 2016. Copyright (2016) American Chemical Society.

Chapter I
Metabolomics for New Insights of
Irritable Bowel Syndrome and Inflammatory Bowel Disease:
A Review

Authors of this work are Mai Yamamoto, Lara Hart, Nikhil Pai, Philip Britz-McKibbin

This section was prepared for submission to Metabolomics as a review article. M.Y. wrote an initial manuscript draft for publication under the supervision of P.B.M. Other co-authors provided constructive feedback and edits on the manuscript draft.

Chapter I
Metabolomics for New Insights of Irritable Bowel Syndrome and Inflammatory Bowel Disease: A Review

1.1 Abstract

Metabolomics offers a powerful systemic approach to decipher the molecular mechanisms underlying human diseases, including pathologies of unknown etiology. Due to the complex interplay between diet, host metabolism and gut microbiota, there is growing interest in applying non-targeted metabolite profiling for biomarker discovery as a way to improve the diagnosis, prognosis and treatment of chronic digestive disorders. Irritable bowel syndrome (IBS) and inflammatory bowel disease (IBD) are common yet poorly understood gastrointestinal (GI) disorders with alarming increases in prevalence worldwide, including among children. Treatment strategies for IBS and IBD remain largely ineffective despite the major socioeconomic impacts on public health of these chronic and debilitating health conditions. Additionally, diagnosis is further complicated by other GI disorders that display similar symptoms while requiring invasive colonic imaging methods for accurate assessment. This review is aimed at providing a comprehensive update on recent progress (2013-2018) in applying metabolomics for IBS or IBD-related clinical research, including animal models and patient cohort studies. Major technical challenges in metabolomics, including best practices in experimental design, robust analytical platforms and data workflows, unknown metabolite identification and biochemical interpretation, as well as biomarker validation with QC/QA are discussed. Also, new advances towards the integration of metabolomics and microbiome data sets are proposed for guiding the design of new strategies for the treatment and prevention of chronic digestive diseases on individual patients, such as validation of optimal dietary interventions and/or probiotic supplementation regimes. With integration with other –omics data, metabolomics holds promise in resolving longstanding diagnostic dilemmas while identifying novel molecular targets for therapy as required for new advances in screening and treatment of IBS and IBD.

1.2 Introduction to Irritable Bowel Syndrome and Inflammatory Bowel Disease

1.2.1   Overview and brief history of IBS and IBD

Irritable bowel syndrome (IBS) and inflammatory bowel disease (IBD) are two of the more commonly diagnosed chronic gastrointestinal (GI) disorders. IBS, a functional GI condition, is characterized by chronically recurring abdominal pain and altered bowel habits for which no clear anatomical abnormalities are found. Altered bowel habits include changes in frequency, consistency (*i.e.,* diarrhea and/or constipation) and caliber of stool. These symptoms can be triggered by many factors such as certain foods from the diet, psychosocial stress and menstruation (Salt 1997). Conversely, immunologically mediated inflammation is found in the GI tract of IBD patients. This auto-inflammatory state results in nutritional malabsorption, bleeding, weight loss, diarrhea and profound fatigue (Charles N. Bernstein et al. 2010). The early onset of IBD significantly increases the risk of adverse health complications, including intestinal strictures, and fibrosis, bowel perforation, fistulae and growth stunting. This constellation of deleterious side effects significantly increases the financial, physical and emotional burden on affected individuals and their families (Cosnes et al. 2012). While IBS and IBD differ in their underlying pathophysiology, they also share many similarities in symptoms, and most importantly, are characterized by poorly understood etiologies. Although genes associated with IBS (Gazouli et al. 2016) and IBD (Liu et al. 2015b) have been identified in genome-wide association studies, both disorders are considered as multifactorial disorders resulting from a complex interaction between genetic and environmental factors and thus, genes themselves do not provide the comprehensive picture.

The first formal definition of the IBS first appeared in 1962, which was described as an "irritable colon syndrome" (**Figure 1.1**). Further back, similar symptoms were alluded to in publications dating as far back as 1871 (Chaudhary and Truelove 1962). The exact worldwide incidence of IBS is difficult to accurately measure due to a lack of recognition for its diagnosis among health professionals, and patients not seeking medical advice (Hungin et al. 2005).
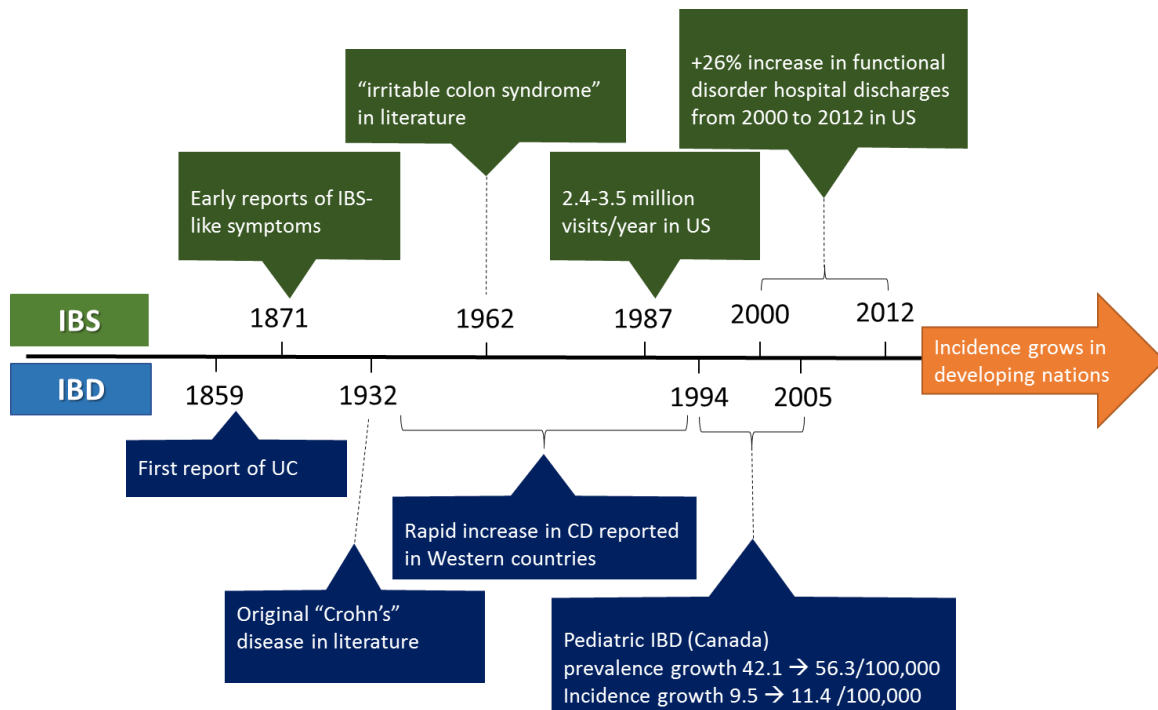
Figure 1.1. Historical appearance and currently reported prevalence of IBS and IBD.

Nevertheless, available data suggests that about 11.2 % of the worldwide population meets criteria for IBS based on a meta-analysis published in 2012 (Lovell and Ford 2012). A greater incidence of IBS has been associated with higher degree of stress in professional and managerial positions, underscoring the involvement of psychosocial aspects in its etiology (Howell et al. 2004). In fact, the reported prevalence is particularly high in North America as IBS is the most common disorder diagnosed by gastroenterologists with 2.4-3.5 million physician reported visits per year in the US (Sandler 1990). IBD also follows a similar trend as IBS with higher incidence reported in North America and Europe. Due to its greater severity and consequent need for medical interventions, there are clear records of IBD-like illnesses as far back as 1859 (Wilks 1859). Interestingly, inflammation isolated in colonic region was more frequently observed in the early 20th century, followed by rapidly increasing incidences of inflammation in other gastrointestinal regions including the ileum and upper GI tract in the latter two thirds of last century (Kirsner 1988). The first case was later termed as ulcerative colitis (UC), whereas the latter phenotype was coined

as Crohn's disease (CD) named after one of the leading authors of the first report published in 1932, which was followed by numerous case reports on similar conditions (Crohn et al. 1932). There is a historically higher prevalence reported in Caucasian and Jewish populations, but the incidence is currently increasing among Asian and Hispanic populations (Bonnevie et al. 1968) likely due to higher levels of urbanization and the widespread adopting of western lifestyle, such as a higher intake of processed foods. Indeed, epidemiological studies have shown that populations emigrating from rural and low prevalence regions to urban and higher prevalence regions are at increased risk of developing IBD (Langholz et al. 1991), which has been well documented in Canada where greater immigration patterns and urbanization exists (Benchimol et al. 2009). The increasing incidence of both IBS and IBD also show consistent associations with environmental changes associated with socioeconomical development, which are linked to alterations in habitual diet, chemical exposures, and increased stress. The growing prevalence of IBS and IBD is anticipated to impact populations in other developing countries with expanding globalization and industrialization activities.

1.2.2 Diagnosis of IBS and IBD

One of major challenges in the diagnosis of IBS and IBD includes reliable differentiation of these related GI disorders given their overlapping symptoms. Indeed, there are currently no diagnostic tests available that definitely diagnose IBS (M. Camilleri and Choi 1997). Instead, diagnosis is made based on a clinical assessment using the Rome IV criteria (**Figure 1.2**), and excluding other digestive disorders such as IBD, colon cancer and food intolerance (Lacy et al. 2016). Since these conditions often have analogous symptoms with variable severity, accurate diagnosis is challenging, time-consuming and often expensive. As a result, it is likely that patients with IBS have previously been falsely diagnosed with another unrelated disease and may have even undergone unnecessary pharmaceutical or surgical interventions (Canavan et al. 2014).
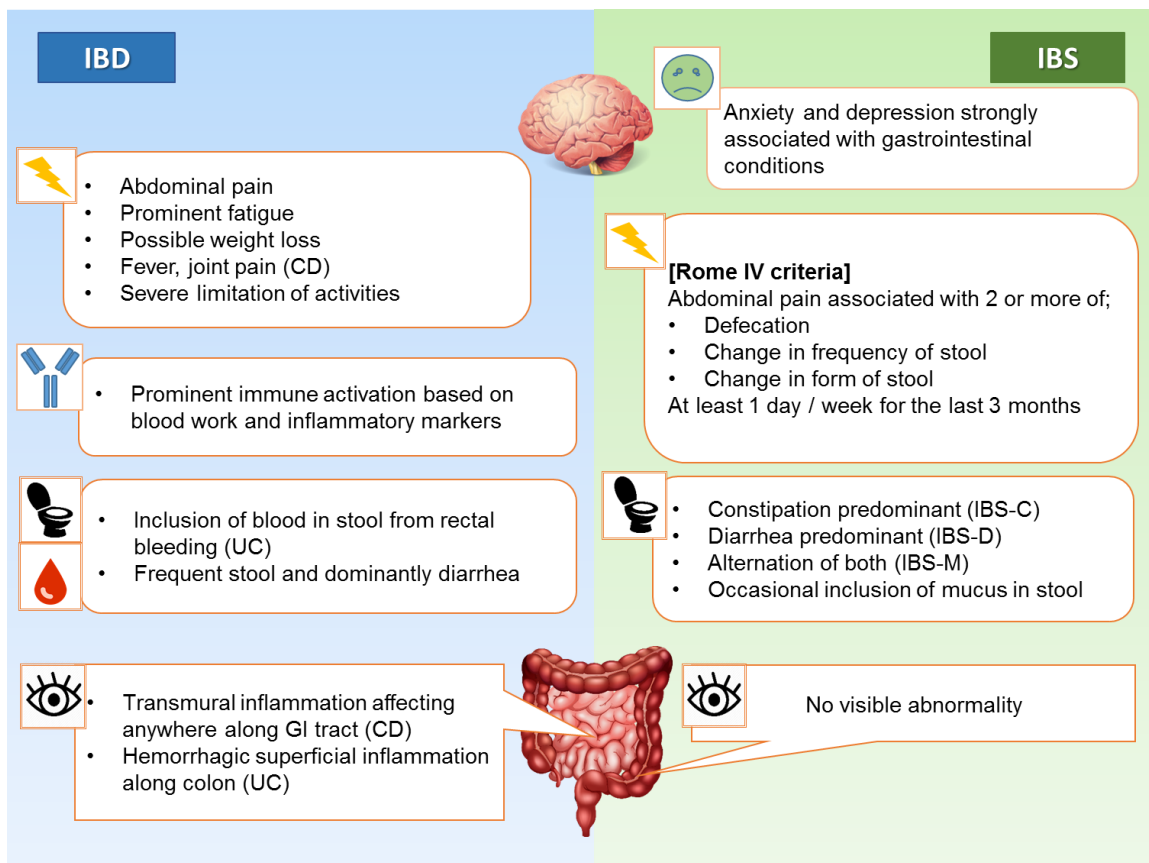
Figure 1.2. Symptoms and characteristics of IBS and IBD used for diagnostic process.

In contrast, in the case of a false-negative diagnosis, where a physician misses a diagnosis of IBD or cancerous growth and incorrectly identifies the condition as IBS, this can have even greater consequences, including higher mortality and morbidity (C. N. Bernstein 2017). Therefore, it is essential to perform a series of investigations to accurately diagnose IBS or IBD, such as laboratory testing, imaging and colonoscopy. The symptomatic criteria in IBS were developed to eliminate the need for this highly variable exclusion process. Within the realm of IBS, physicians use other clinical findings, such as standardized questionnaires to assess psychological stress and bloating when evaluating predominant symptoms to further categorize the patient's condition into subgroups, such as constipation dominant (IBS-C), diarrhea dominant (IBS-D), and mixed bowel habits (IBS-M) (Spiegel et al. 2010).

In the case of laboratory tests, serum C-reactive protein (CRP) and fecal calprotectin (FCP) are two common inflammatory markers applied in the diagnosis of gastrointestinal disorders. Low levels of both have been especially useful in excluding IBD (Menees et al. 2014), and shifting a focus towards functional GI disorders, such as IBS. However, these inflammatory markers only confirm the presence of inflammation without providing more specific information, such as the possibility of tumour growth or exact site of inflammation. As a result, imaging colonoscopy with a tissue biopsy for histopathology remains the gold standard for accurate diagnosis (Akarsu and Akarsu 2018). This procedure is also the primary means of diagnosing subtypes of IBD, namely Crohn's disease (CD) and ulcerative colitis (UC). In CD, inflammation appears anywhere along the digestive tract, but most often in the ileum and colon. Inflammation can be transmural, affecting deep layers of epithelial tissue, which can lead to deep ulcerations or a cobble stone-like appearance of the mucosa (Charles N. Bernstein et al. 2010). On the other hand, UC affects only the colonic epithelial layer or superficial mucosa, and presents with continuous inflammation from the rectum proximally (Charles N. Bernstein et al. 2010). The invasiveness of colonoscopy contributes to greater risks for health complications while representing a financial burden on public healthcare with considerable stress on patients, especially for children (Cash et al. 2002).

1.2.3   Traditional and emerging therapeutic treatment options

The unknown etiology IBS and IBD, and the lack of validated biomarkers for their reliable diagnosis translates to poor understanding of differential treatment responses to therapy that vary widely between patients. Treatment for IBS generally involves a wide range of therapies from pharmacological agents to psychosocial therapy for managing comorbid anxiety and depression (Deechakawan et al. 2011; K. Gwee et al. 1999; Bercik et al. 2009). Mild symptoms can be managed by medications that regulate bowel habits and reduce abdominal pain; however, severe cases remain challenging to treat because of poor response and serious side effects associated with available medications (Emeran A. Mayer

2008). For example, in patients with a IBS-C subtype, tegaserod (a partial 5-hydroxytryptamin-receptor agonist) has shown positive effects in clinical trials (W et al. 2007), but its use was later restricted by US Food and Drug Administration due to increased risk for cardiovascular events, such as stroke (Administration 2012). For patients diagnosed with a severe IBS-D subtype, alosetron (5-hydroxytryptamin-receptor antagonist) can be used to reduce defecation frequency and relieve abdominal discomfort (Cremonini et al. 2003; Emeran A Mayer and Bradesi 2003) at the expense of serious side effects, such as ischemic colitis and bowel obstruction in some cases (Chang et al. 2006). Tricyclic antidepressants are administered at low dosage to relieve abdominal pain by normalizing GI transit and improving the quality of sleep, and at high dosage to treat anxiety and depression (Clouse and Lustman 2005); however, the efficacy of these medications are inconsistent and usually provide only temporary relief (Drossman et al. 2002; L. J. Brandt et al. 2002; Jackson et al. 2000; Lesbros-Pantoflickova et al. 2004). Aside from pharmacological therapies, cognitive behavioral therapies can be used to change maladaptive thinking patterns and dysfunctional behaviors underlying the perception and occurrence of somatic symptoms (Drossman et al. 2002; Lackner et al. 2008; P. J. Whorwell 2005). The interaction between psychological stressors and symptoms are significant in IBS (Mikocka-Walus et al. 2016), but appears less pronounced in IBD, likely due to underlying immune dysfunction. To reduce the inflammatory response and induce a state of remission in IBD, the most effective options are either immune modulators, corticosteroids or nutritional therapy (Escher 2013). To maintain remission, patients may be prescribed aminosalicylate-based therapy, immune modulators or biologics (M. Stephens and Marvis 2013). As in IBS, these pharmacological therapies are accompanied by side effects, such as potential damage to the kidneys and liver, skin and immune system. Long term use (or frequent recurrent courses) of corticosteroids is particularly associated with significant adverse effects, including osteoporosis and growth retardation (Hart and Ng 2010). This is especially detrimental for pediatric patients who are at critical stages of physical and mental growth.

Beside pharmacological treatment, modulation of the microbiome has been an area of active research to effectively treat symptoms of IBS and IBD without significant deleterious health outcomes. The essential role of gut bacteria in nutrition and human health has been reported in numerous studies. Their influences are not limited to the GI tract, but also in the overall immune system and even to neurological development (Iyer et al. 2004; Finegold et al. 2012; O'Mahony et al. 2005). Oral intake of probiotics is the most accessible and practical way of introducing active microorganisms that potentially exert health benefits but do not persist in the gut (P Moayyedi et al. 2010). Single strain or a combination of *Lactobacilli* and *Bifidobacteria* are most commonly used and their putative benefits have been reported for relief of IBS symptoms in some studies (Peter J. Whorwell et al. 2006; P Moayyedi et al. 2010; O'Mahony et al. 2005). However, a recent systematic review of 35 probiotic randomized control trials reported that there is not adequate evidence to support the efficacy of specific probiotics for improvement of IBS symptoms due to heterogeneity of IBS symptoms and probiotics (McKenzie et al. 2016). In addition to the heterogeneity, lack of thorough investigations in other important factors, such as habitual diet and ingredients in the vehicles used to introduce probiotics, further contributes to current contradictory results seen in studies of probiotics efficacy for IBS symptoms. Similarly, a review of seven recent studies concluded that probiotics are ineffective in maintaining remission and preventing relapse in CD (Rolfe et al. 2006). For UC patients, some strains of probiotics have shown efficacy in maintaining remission if the disease severity is mild to moderate (Mallon et al. 2007). For severe cases of IBD, recent studies have demonstrated the efficacy of fecal microbiota transplantation (FMT) therapy, whereby 'healthy' microbiota from a suitable healthy donor are introduced in the form of fecal matter to a patient. The cocktail of healthy commensal bacteria is considered to restore homeostasis of gut microbiota and it has shown remarkable efficacy in the treatment of refractory and recurrent *Clostridium difficile* colitis (Lawrence J. Brandt et al. 2012). FMT has also led to improvement or resolution of symptoms in UC patients (Paul Moayyedi et al. 2015), but has not been adequately studied for IBS and CD patients (Rossen et al. 2015). Unlike probiotics, the use of FMT is less widely established due to lack of regulation for conditions

other than *Clostridium difficile*, and the costs associated with screening donors for potential infectious diseases (Glauser 2011). Well-designed randomized clinical trials in large populations are still required to confirm the clinical benefits and better understand the potential risks of FMT for treatment of IBS and IBD patients.

In addition to the introduction of bacterial community, modulation of gut microbial activities can be achieved by modification in diet whereby certain fermentable food by bacteria are specifically removed from diet. Among many dietary managements, the low fermentable oligosaccharides, disaccharides, monosaccharides and polyols (FODMAPs) diet has shown evidence for its efficacy in mitigating IBS symptoms (Staudacher et al. 2014; Halmos et al. 2014; R and J 2010). FODMAPs in the diet are considered to augment luminal water and induce more gas production in the colon as a result of increased colonic bacterial fermentation. This results in the luminal distention, bloating sensation and associated abdominal pain that is frequently observed in IBS patients (Major et al. 2017; Murray et al. 2013). In one randomized controlled cross-over trial on IBD patients, a low FODMAP diet significantly improved functional-like GI symptoms, such as abdominal pain (Gearry et al. 2009; Prince et al. 2016), however no change in inflammatory marker (fecal calprotectin), pH and short chain fatty acid concentrations were observed (Halmos et al. 2016). It is clear that dietary restriction of fermentable carbohydrates reduce discomfort in both IBS and IBD patients, but not without potential adverse effects. For example, studies conducted with healthy controls, as well as IBS and IBD patients, consistently demonstrate reduced population of bacteria that are considered to be beneficial for our gut health, such as *Bifidobacteria* (Staudacher et al. 2012) and the butyrate-producing *Clostridium* cluster XIVa (Halmos et al. 2015). Additionally, concurrent increase of the species abundant in IBD patients, namely *Ruminococcus spp.* was also observed, implying poorly understood and unanticipated perturbations in gut microbiota populations especially for IBD patients (Halmos et al. 2016; Halmos et al. 2015). Thus, modification in dietary FODMAPs may be an effective option to cope with symptoms in IBS, but may be detrimental for long-term GI health among IBD patients. Alternatively,

exclusive enteral nutrition (EEN) therapy represents a unique dietary intervention to improve symptoms in IBD patients by healing intestinal mucosa, and it has become the first-line therapy for induction of remission for pediatric CD patients. During EEN, patients exclusively take liquid formulation orally or through a nasogastric tubing for their entire caloric needs without any solid food intake for the period of 6 to 8 weeks. Regardless the types of formulation, studies have shown higher rate of mucosal healing in EEN as compared to corticosteroid therapy in CD patients (Otley et al. 2013). However, the exact mechanism(s) of its action are still under active investigation, but may include removal of potential allergens from the diet, the presence of anti-inflammatory compounds in the formula, and modulation of gut microbiota (Otley et al. 2013). Meister *et al.* (Meister et al. 2002) demonstrated a significant increase in anti-inflammatory cytokines and a corresponding decrease in pro-inflammatory cytokines *in vitro* using tissue cultures collected from CD patients following incubation with EEN formula. The effects on gut microbiota was also examined in pediatric CD patients where a reduction of bacterial diversity and sustained alteration in microbial community was achieved (Leach et al. 2008). The mechanism behind the association between certain bacteria and mucosal healing, which leads to symptom improvement, is still not completely understood. Additionally, a lack of studies investigating the efficacy in UC has been preventing the use of EEN, leaving corticosteroid as the first-line treatment for UC even in children despite underlying concerns on growth and development with prolonged usage. As a result, further studies are needed to validate the suitability of dietary interventions, such as low FODMAPs and EEN for effective treatment of chronic digestive disorders and their impact on modulating gut microbiota and intestinal mucosal healing as an alternative to conventional pharmacological therapies that do not treat the underlying causes of the disease.

1.2.4 Challenges in monitoring therapy responses

Successful treatments of IBS and IBD are defined by a prolonged period of remission, making regular examination during and after a treatment period necessary to fully evaluate the long-term clinical outcomes of a dietary, pharmacological and/or surgical intervention.

The need for follow-up therapeutic monitoring is also important in chronic GI disorders especially IBD, which is associated with a higher risk of developing colorectal cancer (Pinczowski et al. 1994; L. A. García Rodríguez 2000). Similar to the major challenges related to accurate diagnosis of related GI disorders, there are currently no validated, non-invasive biomarkers to monitor differential treatment responses of IBS and IBD patients to therapeutic intervention. In place of disease-specific biomarkers, symptom-based clinical activity scores, and systemic inflammatory biomarkers from blood or stool are generally employed to evaluate IBD disease status (Schoepfer et al. 2012), whereas symptom-based severity score (R. Spiller et al. 2007) and psychological assessments (Bjelland et al. 2002) are often used for monitoring IBS patients. Indeed, endoscopic and histological evaluation remain the most conclusive procedures to determine mucosal healing while confirming the lack of tumorigenic growth, but these procedures are both costly and highly invasive. As a result, fecal inflammatory markers, such as FCP and fecal lactoferrin, have shown strong agreement with endoscopic activity and demonstrated their reliability as relatively non-invasive biomarkers for mucosal healing in place of colonoscopy (Røseth et al. 2004; Taina Sipponen et al. 2010; T. Sipponen et al. 2008). However, these biomarkers lack specificity among related GI disorders and do not assess disease severity or location. Furthermore, the cut-off value for remission is undergoing frequent reassessment and changes which reflects challenges in stool specimen collection and sample heterogeneity (Papay et al. 2013). In the case of IBS, symptom-based monitoring poses greater challenges in monitoring of therapy response, which is assessed entirely by questionnaires that are subjective and prone to high placebo response (Ford and Moayyedi 2010). For example, symptomatic assessment mostly by standardized questionnaires/surveys contributes to recall bias and "fake" compliance (Z. Mujagic et al. 2015). Consequently, the discovery and validation of specific yet reliable biomarkers are urgently needed for more objective assessment of treatment responses and disease management of IBS and IBD patients. In particular, serum/plasma or notably urine represent ideal human biofluids as they are convenient for repeat collection and therapeutic monitoring during routine clinical visits.

1.3 Introduction to Metabolomics

1.3.1   Overview of 'omics' approach and relevance to IBS and IBD

Since IBS and IBD are distinct yet multifactorial GI disorders, systemic approaches are needed when evaluating the complex interactions involving genes and environment as related to disease pathophysiology and progression among individual patients. In this context, new advances in systems biology using complementary 'omics' technology provide an opportunity to improve understanding of chronic diseases of unknown etiology, including comprehensive screening of genes (genomics), mRNAs (transcriptomics), proteins (proteomics) and metabolites (metabolomics) from host as well as symbiotic gut microbiota. To date, there has been impressive progress in genomics research on IBD as related to disease risk assessment when using genome-wide association studies (GWAS) due to recent advances in next generation sequence technology and international efforts at data curation (Huang et al. 2014). As expected from the efficacy of gut microbiota-modulating therapies, genes associated with recognition and responses to bacteria and immune systems are associated with IBD (Jostins et al. 2012). Genomic research on IBS is lagging behind of IBD, however several studies and meta-analysis identified *TNFSF15* polymorphism that influences regulation of immune and inflammatory responses as an important risk factor of IBS (Czogalla et al. 2015; Zucchelli et al. 2011; Swan et al. 2013). Interestingly, the same gene was also found to be associated with IBD in a large population of Europeans and Asians (Liu et al. 2015a), suggesting that a defect in host immune system may be involved in the etiology of both disorders as triggered by changes in environmental exposures. Similarly, proteomics studies also found overexpression of proteins associated with immune activation in IBS and IBD (Ding et al. 2010; Buhner et al. 2018), as well as structural proteins involved in cell adhesion and migration (Buhner et al. 2018; Ding et al. 2010; Shkoda et al. 2007). Out of all 'omics' disciplines, the field of metagenomics of gut microbiota (*i.e.,* microbiome) is growing at the fastest rate (Huang et al. 2014). The essential role of gut bacteria in nutrition and health, as well as GI pathology, has been known for decades as early demonstrated by the observation that the most inflamed location in the gut is the largest number of bacteria are located (Sellon et al. 1998). More recently,

gene sequencing technology and bioinformatics tools have developed to allow for rapid screening and identification of gut bacterial composition and functions. Frequently used technique in gut microbiome studies is the use of 16S rRNA gene, which is a conserved yet species-specific region of bacterial DNA. Technological advances in this high-throughput, culture-independent sequencing has led to rapid increase in the number of studies investigating microbiome research as reflected by the NIH funded Human Microbiome Project (Methé et al. 2012). However, this technique lacks taxonomic resolution due to only a partial 16S gene being sequenced (Lagier et al. 2016). Alternatively, identification of bacteria by matrix-assisted laser desorption ionization-time of flight mass spectrometry (MALDI-TOF-MS) is emerging as a new high throughput screening technology for microbiome research that is easily automated and less expensive to operate. This technique relies on identification of bacterial ribosomal protein based on mass spectral pattern recognition when comparing matching spectra in reference databases. Identification down to genus level can be achieved rapidly, even species and strain level of identification is possible provided that pure strains and good reference databases are available (Singhal et al. 2015), which is essential in determining strain-specific virulence factors of pathogenic bacteria as seen in multiple pathotypes identified for *Escherichia coli* (Kaper et al. 2004). Based on these high throughput microbiome screening techniques, a higher abundance of certain pathogenic bacterial strains have been consistently measured in IBD patients and their functional roles have been associated with immune activation (Boland et al. 2017; Hamilton et al. 2017). Similarly in IBS, a significantly higher abundance of certain opportunistic pathogens have been a consistent outcome reported in IBS patients as compared to healthy controls (Saulnier et al. 2011). This state of altered or abnormal bacterial community is termed 'dysbiosis' and has been frequently associated with pathological conditions including IBS and IBD. Nevertheless, genomic and proteomic studies to date have not yet resulted in a validated biomarker(s) to enable differential diagnosis of IBD or IBS, as well as allow for prediction of disease progression or treatment responses to therapy (Drucker and Krapfenbauer 2013). Indeed, this is a reflection of the major technical hurdles associated with rigorous biomarker translation in clinical medicine,

as well as sources of bias that contribute to false discoveries, including underpowered studies, flawed experimental designs, poor replication in independent cohorts, and inadequate method validation or quality control (Ioannidis 2005).

Metabolomics is defined as the comprehensive analysis of low molecular weight (< 1,500 Da) compounds within a biological sample, such as a tissue or biofluid (D. S. Wishart et al. 2013). Metabolites are not only real-time molecular end-points of gene expression and protein/enzyme activity, but also function as factors regulating their activity that is closely associated with disease phenotype and physiology. Additionally, the human metabolome encompasses both endogenous metabolites of cellular metabolism, as well as a diverse array of exogenous compounds derived from lifelong exposures, including diet/nutrients, drugs and environmental toxins (Rattray et al. 2018). As a result, metabolomics offers a powerful systemic approach for biomarker discovery while revealing underlying disease mechanisms, which have already been adapted in clinical diagnostics for decades (Patti et al. 2012). One of the most important examples of metabolite screening is universal newborn screening program, which is recognized internationally and practiced across North America as a preventive public health program for early detection of treatable genetic disorders that reduces mortality and morbidity with improved long-term clinical outcomes as compared to symptomatic diagnosis (Therrell and Adams 2007). In this case, high throughput screening of a large panel of biomarkers are analyzed simultaneously using direct infusion-tandem mass spectrometry (MS/MS) at incremental costs from a single dried blood spot punch from neonates, including amino acids and acylcarnitines for pre-symptomatic identification of dozens of inborn error of metabolism, such as phenylketonuria and medium-chain acyl-coenzyme A dehydrogenase deficiency (Scolamiero et al. 2015). Additionally, metabolomics is widely applied in a range of fields including nutrition (Gibney et al. 2005), toxicology (Robertson et al. 2011), agriculture (Dixon et al. 2006), and drug development (Shyur and Yang 2008). Metabolomics does not target only human metabolism but can be applied to gut microbial metabolism and co-metabolism between human host and gut bacteria that is considered to

play an essential role in the pathology of chronic GI disorders. In this context, profiling of gut microbiome in conjunction with metabolomics offers a complementary strategy for better understanding the underlying causes of disease, as well as improving the diagnosis and treatment of dysbiosis of gut microbiota that are modulated by diet, lifestyle and pharmacological interventions. In addition to the integration of metabolomics and bacterial genomics, investigation of interaction among genes, proteins, immunological processes and environmental exposures is indispensable to understand multifactorial diseases, such as IBS and IBD. Traditionally each of these disciplines has generated knowledge in isolation, neglecting the interconnected role and contribution of other disciplines that can be directly linked to disease phenotype (**Figure 1.3. a**) (de Souza et al. 2017). With the rapid development of technology discussed earlier, scientific knowledge doubles every nine years (Van Noorden 2014) but does not necessarily result in knowledge integration that is critical for holistic understanding of complex processes relevant to human health. For instance, human-microbial co-metabolites found in a hypothetical patient with IBD may indicate the presence of certain bacterial pathogens in the gut, but the research findings never leave the realm of supposition unless supported by complementary bacterial metagenomics data, proteomics data of virulence factors, and evidence of immune activation. For the integrated knowledge synthesis, coined as 'interactome' in the recent perspective paper (**Figure 1.3. b**) (de Souza et al. 2017), a number of systems biology and bioinformatics tools have been developed, including the Ingenuity Pathway Knowledge Base (Calvano et al. 2005) that holds updated information of literature and public databases, and Cytoscape (Shannon et al. 2003) that allows data integration and visualization of the interactome. The next step required for truly applicable knowledge synthesis for systematic understanding of IBS and IBD and development of effective, personalized therapy is increased collaboration among researchers in different disciplines, as well as clinical researchers with awareness of ever-changing landscape of these multifactorial diseases.
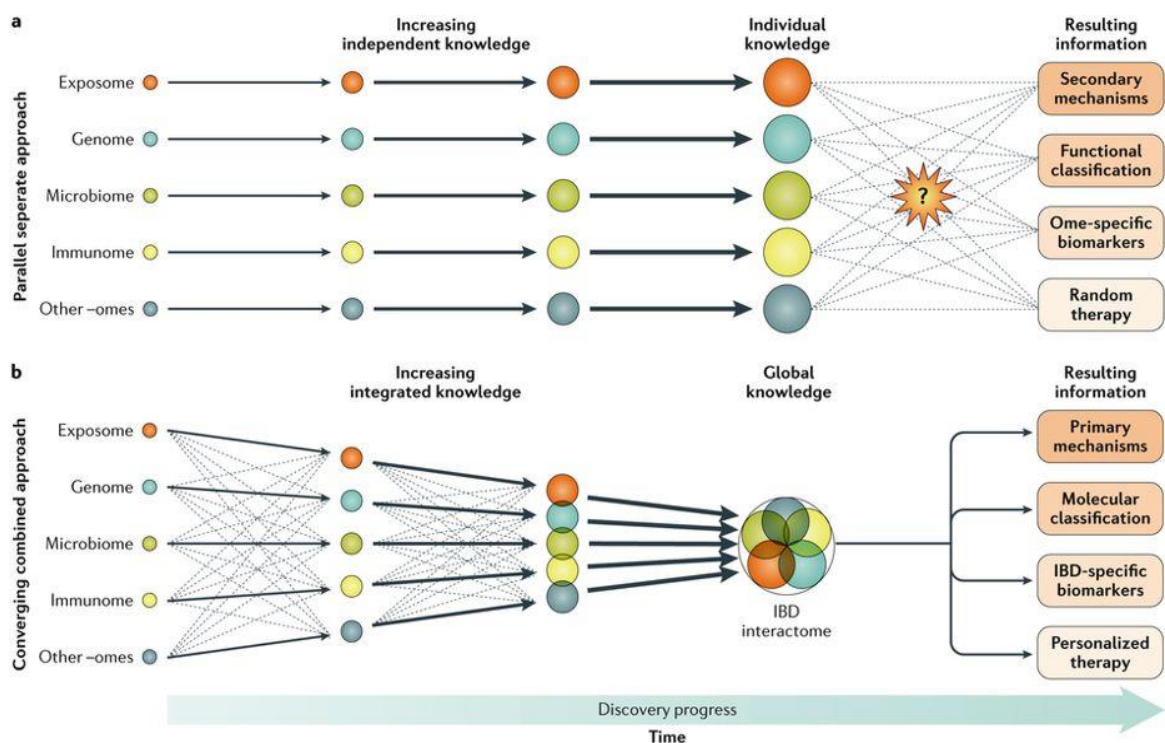
Nature Reviews | Gastroenterology & Hepatology

Figure 1.3. **a**. Parallel knowledge synthesis in each discipline in isolation, which results in only a partial understanding of biological mechanisms and biomarkers that are applicable in specific discipline. **b.** Integrated knowledge synthesis which results in powerful recognition of primary mechanisms underlying a disease that lead to effective, personalized therapy. (Source: Reproduced from (de Souza et al. 2017) with permission)

1.3.2   Methods in metabolomics – Major analytical platforms

A major technical challenge in metabolomics is the lack of a single platform amenable for comprehensive metabolite profiling in complex biological samples. This reflects the sheer chemical diversity and wide dynamic range of the human metabolome, including a large fraction of unknown compounds derived from lifelong environmental exposures (Kuehnbaum and Britz-McKibbin 2013). Among different instrumental platforms available, nuclear magnetic resonance (NMR) spectroscopy was one of the first methods used in metabolomics and is still the most commonly used method (Kuehnbaum and Britz-McKibbin 2013) due to its excellent robustness and reproducibility that allows for

unambiguous metabolite identification when comparing to NMR reference spectral databases (Keun et al. 2002). However, NMR suffers from high infrastructure/operation costs, as well as poor sensitivity that limits metabolome coverage to mainly abundant metabolites present at low micromolar to millimolar concentration levels (Psychogios et al. 2011). In contrast, high resolution mass spectrometry (MS) coupled to high efficiency separation techniques based on chromatography, electrophoresis or ion mobility, offers much higher sensitivity and specificity for expanded metabolome coverage, including the resolution of isomers/isobars in complex biological samples. For instance, gas chromatography-mass spectrometry (GC-MS) is best suited for the analysis of volatile and thermally stable metabolites, such as aromatic hydrocarbons and short-chain fatty acids with high separation efficiency, whereas extensive mass spectral database libraries facilitate compound identification when using electron impact ionization (EI)-MS (Kanani et al. 2008). Nevertheless, complicated sample workup protocols and pre-column chemical derivatization is often required when analyzing polar metabolites in biological samples prior to GC-MS (Fiehn 2016). As a result, liquid chromatography (LC) coupled to electrospray ionization (ESI)-MS is increasingly applied in metabolomic studies since it allows for detection of a wide range of polar/non-polar and involatile metabolites, including long-chain fatty acids, bile acids and other classes of lipids. However, various separation modes are needed to adequately resolve different metabolite classes using different column types/retention mechanisms (*e.g.* reversed-phase, hydrophilic interaction, ion-exchange) and gradient elution programs (Yamashita et al. 2007; Caron et al. 2009; Matějíček 2011). Alternatively, capillary electrophoresis (CE)-ESI-MS offers a complementary separation technique for resolution of the ionic metabolome which is ideal for analysis of volume-restricted or mass-limited biospecimens. Additionally, biospecimens with high salt content such as urine can be effectively analyzed by CE-MS with minimum ion suppression efefcts due to the efficient desalting mechanism of the high efficiency separation. This results in typically much sharper peaks with better resolution of ionic metabolites in CE, which are not adequately retained in reversed-phase LC demonstrated in urinary peptide analysis (Klein et al. 2014). Nevertheless, CE-MS generally suffers from poor concentration

sensitivity due to the small volume of sample that is introduced within the narrow capillary, as well as post-capillary dilution effects from the sheath liquid that is used to generate a stable electrospray for ion desorption into the gas-phase. To improve the sensitivity, a sheathless ESI design using a porous capillary tip was designed by Moini (Moini 2007) and successfully applied to cationic separation of antibody-drug-conjugates characterization (Said et al. 2016), amino acids (Ramautar et al. 2008), and anionic metabolite analysis with excellent reproducibility and limit of detection down to 10 nM level (Gulersonmez et al. 2016). Additionally, low sheath flow interface was introduced to mitigate the spray instability issue observed in the sheathless interface and minimize the dilution of sample with sheath liquid (Hsieh et al. 1999). This low-flow interface with tapered glass emitter was later modified and simplified by Dovichi *et al.* to eliminate the need for a pump and nebulizer gas by taking advantages of the electrokinetic flow (Wojcik et al. 2010). The nanoflow facilitates better desolvation, enhanced sensitivity and increased salt tolerance when compared to conventional sheath-flow interface (Cole 2000). These advantages were highlighted in the recent application of this interface for characterization of monoclonal antibodies (Han et al. 2016) and complete glycosaminoglycan mixtures (Sun et al. 2016).

Another drawback of CE-MS, that was not addressed in recent interface development, is arguably method robustness and poor migration time reproducibility, however these are likely a consequence of a lack of rigorous method development and validation, as well as poor training and vendor support given the numerous factors that impact separation and ionization performance that can contribute to incidental capillary breakage, loss of electric current and deleterious peak broadening (Zhang et al. 2017). In fact, large-scale metabolomic studies of over 8000 plasma samples collected from a large cohort of sedentary adult population demonstrates that CE-MS offers a robust platform with acceptable long-term reproducibility with a coefficient of variation < 20 % for a majority of 94 polar metabolites consistently detected in the population (Harada et al. 2018).

1.3.3 Metabolomics data workflow and major pre-analytical challenges

Metabolomics offers a powerful tool to discover new biomarkers of clinical significance that enable better screening, diagnosis and/or treatment monitoring for new breakthroughs in precision medicine. However, clinical metabolomic studies often suffer from limited study power due to small effect sizes and large biological variance when studying heterogeneous populations, which is the perfect recipe for false discoveries (Ioannidis 2005). For reproducible research findings that can be replicated across independent cohorts/multiple centres that are criticial for successful clinical biomarker translation, quality control and quality assurance (QC/QA) measures need to be implemented within each step of the data workflow in metabolomics, including pre-analytical, analytical and post-analytical processes (**Figure 1.4**). Pre-analytical steps in metabolomics generally starts with quality assurance (QA) procedures including a robust experimental design and implementation of standard operating protocols for sample collection/storage, as well as pretreatment of biological samples prior to analysis using NMR and/or one or more MS-based methodologies. Unlike "confirmatory" design of experiments of known metabolites or targeted biomarkers as employed in large phase III clinical trials, metabolomics often serve as pilot projects that are primarily hypothesis-generating, but are prone to false discoveries due to data overfitting especially when not correcting for multiple hypothesis testing (Ioannidis 2005). In addition, generally higher flexibility in experimental design and study outcomes in metabolomics can contribute to bias (Ioannidis 2005). For instance, bias may originate from a criticized practice whereby a researcher proposes a hypothesis based on available dataset and then subsequently tests the hypothesis using the same dataset (Forstmeier et al. 2017). Not surprisingly, the hypothesis tends to be proven true in this case. All or most of these issues can be addressed and evaluated at the point of study design before cohort recruitment and sample collection, such as submitting a clinical trials registry outlining goals/scope of work before study initiation.
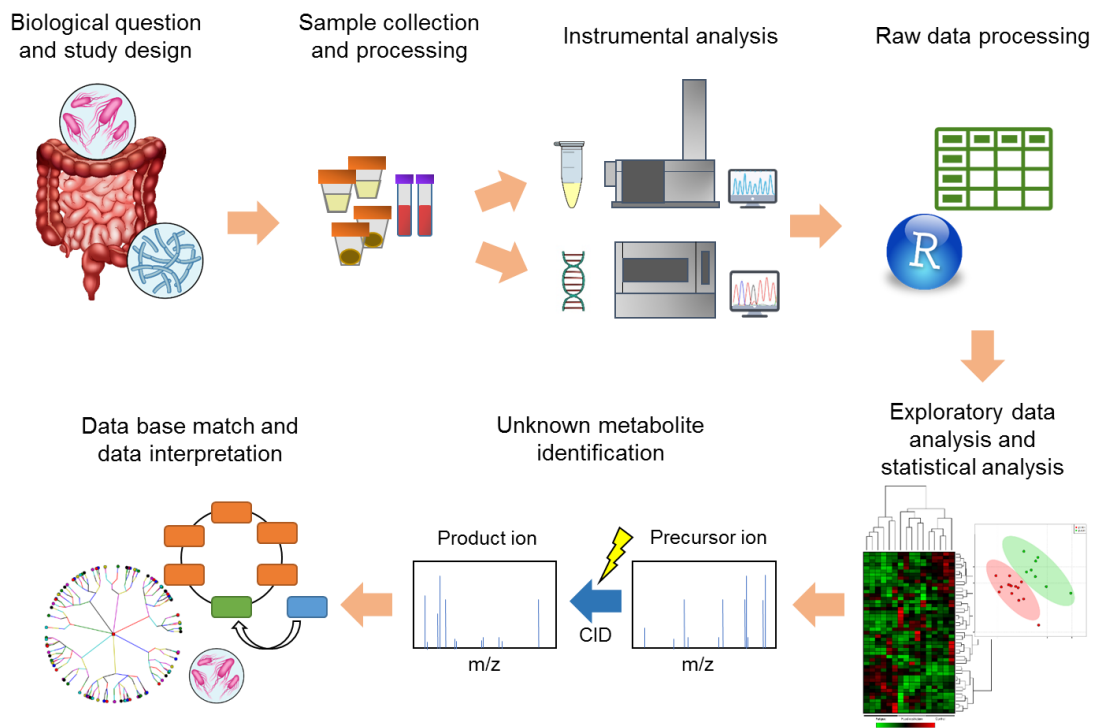
Figure 1.4. Untargeted metabolomics workflow, depicting the start of the project to the final step of data interpretation in relevance to the biological system in question. CID: Collision induced dissociation.

For instance, DNA is relatively stable over time as reflected by genetic analysis of ancient biospecimens, whereas metabolic processes are highly dynamic and are sensitive to changes in ambient conditions. Similarly, delays to samples processing, improper or inconsistent use of additives for sample stabilization (*e.g.,* EDTA for plasma collection), and repeat freeze-thaw cycles of a single aliquot of sample contribute to far greater irreproducibility and bias (Yin et al. 2013; Zivkovic et al. 2009; Lauridsen et al. 2007). For example, a wide range of urinary metabolites were found stable up to six months in a -20 °C freezer (Gika et al. 2008), whereas they were significantly influenced by bacterial biotransformation when the same urine samples were stored at room temperature without addition of a compatible antimicrobial agent, such as sodium azide or boric acid (Saude and Sykes 2007). Such effects are even more pronounced in heterogeneous stool specimens as expected from the fact that 25-54% of dry matter is composed of bacteria (Stephen and Cummings 1980). Indeed, there have been few rigorous studies examining the impact of

sample collection and storage procedures on the stool metabolome, where changes in concentrations of certain metabolites have been reported to vary up to 100-fold in fecal water samples (Gratton et al. 2016). Stool is one of the most common sample type collected from IBS and IBD patients given its importance in microbiome research, as well as measuring biomarkers associated with intestinal inflammation, such as FCP that is derived from increased turnover of leukocytes in the gut wall and migration of neutrophils into the gut lumen (Fagerhol et al. 1980).

While sample collection is one part of quality assurance, another step involves maintaining system stability of analytical instrumentation and consideration of batch effects, as well as any process that can result in unexpected contamination of samples (*e.g.* plasticizer from a tube or solvent during extraction). Unlike NMR, ESI-MS has direct contact with samples injected, which easily results in high background noise, artefact signals and potential ion suppression if the ion source is not cleaned regularly. As a result, a preventative maintenance schedule is important in MS-based metabolomic studies with daily tuning/cleaning based on a standard operating procedure to ensure robust system performance without long-term signal drift. Additionally, correction for batch effects should be considered in the experimental design when analyzing samples derived from separate batches interspersed over a long period of time when using the same instrumental platform or pooling data from different instruments or laboratories (Livera et al. 2015). High data fidelity is ensured by following the rigorous QA/QC procedures through analysis of a number of system suitability testing samples prior and during a randomized analysis of a batch/block of samples, including blanks, calibrants and a standard reference material or a representative pooled sample from cohort that serves as a QC (Broadhurst et al. 2018). For example, system and process blank samples are used to assess background noise in the instrument and detect potential contaminants resulting from reagents/solvents used in processing, respectively.
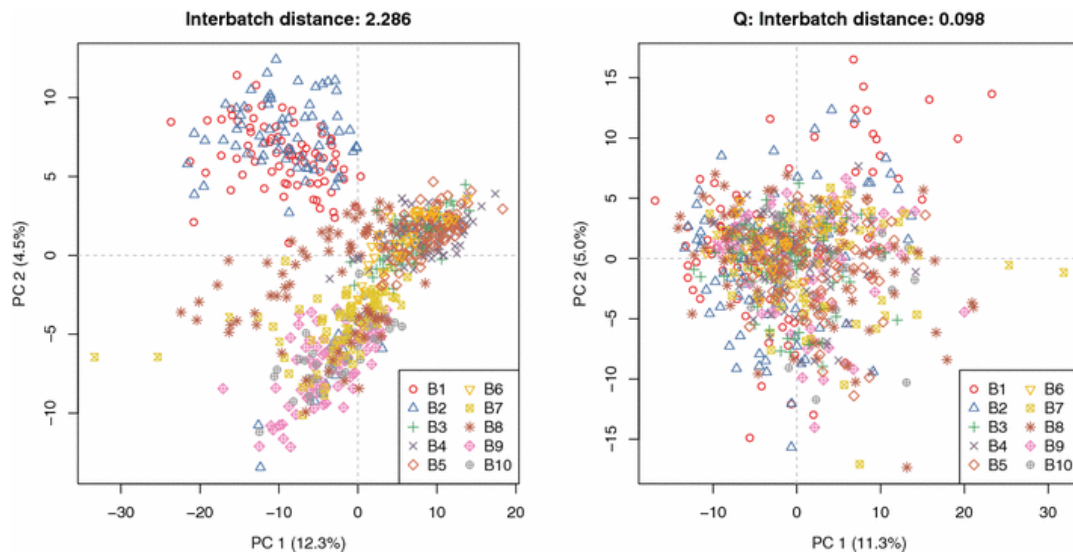
Figure 1.5. 2D scores plot from PCA of metabolites derived from *Arabidopsis* cell extracts before and after batch correction using QC samples. (Reproduced from (Wehrens et al. 2016) with permission)

Also, intermittent analysis of QC specimens within a data workflow acts as an important check to estimate technical variance while ensuring that sensitivity, selectivity and resolution are satisfactory prior to data interpretation of biological data. This is effectively visualized by using principal component analysis (PCA) on metabolomics data sets (**Figure 1.5**), which is useful when summarizing overall data trends, as well as detecting outliers and batch effects in reduced dimensionality space prior to and after application of a QC-based batch correction adjustment algorithm [139].

1.3.4 Major post-analytical challenges in MS-based metabolomics

Once overall metabolomics data fidelity is assured following implementation of a rigorous QA/QC procedure, this raw data are pre-processed prior to multivariate or univariate statistical data analysis. This process often includes missing value imputation, transformation and scaling, and has been found to greatly influence the statistical analysis outcome, and thus the entire result of a study (Di Guida et al. 2016). Missing values are frequently present in untargeted metabolomics studies that involve hundreds of metabolites with widely variable concentration levels across samples. Some metabolites are below

method detection limits and may also be completely missing from a subset of samples in a cohort. In these situations, missing value imputation is applied to replace the values with non-zero values without distorting the data structure. This step is important prior to the application of multivariate data analysis methods, such as PCA, which performs most robustly with data that do not contain missing values. Similarly, PCA is largely influenced by deviation from overall normal data distribution, which is often caused by a handful of highly concentrated metabolites in a dataset. Data transformation and scaling are thus applied to mitigate this issue by correcting for heteroscedasticity and skewed distribution, and adjusting for differences in fold changes among metabolites respectively. A comprehensive review of the effects of pre-processing step on statistical outcomes is beyond the scope of this manuscript, and authors are recommended to consult Guida *et al*. (Di Guida et al. 2016). Subsequently, supervised and unsupervised multivariate statistical analysis are applied to the final set of data consisting of metabolite levels for each sample accompanied with meta-data such as age, sex, and disease type. It is often the case in untargeted metabolomics that many molecular features in the final data matrix still remain unknown at this point due to a lengthy process involved, which is the topic of discussion in the following section. Hence, statistical analysis is an essential step to identify unknown compounds that warrant the lengthy identification process especially when using MS and MS/MS if they are deemed statistically significant (*i.e.,* differential expression in two or more groups) based on the study design. Two of the most commonly used multivariate statistical techniques are PCA and partial least square discriminant analysis (PLS-DA). PCA is as an unsupervised exploratory technique that reduces a data matrix composed of N observations and K variables down to minimum numbers of orthogonal principal components (PCs) that represent the underlying directionality of maximum variance in data (Keun et al. 2002). Typical untargeted metabolomics data often contain hundreds to thousands of molecular features (*i.e.* variables), which makes this dimensional reduction technique indispensable. As shown in the identification of a metabolomics batch effect in **Figure 1.5**, it is a valuable technique to identify outliers, clusters of groups, and trends in the data visually (Madsen et al. 2010). PLS-DA, on the other hand, is a supervised pattern

recognition technique that aims to infer the variables that maximize the discrimination between pre-defined sample groups and to predict the classification of samples in a validation set. PLS-DA has been applied in numerous studies, such as discrimination of healthy individuals from Crohn's disease patients (Jansson et al. 2009). Despite the common application of PLS-DA in metabolomic studies for metabolite selection based on variable importance in projection (VIP), it is still poorly understood by many users, and proper validation of the model tends to be overlooked, which leads to data over-fitting and false-positives (Grootveld 2015). It is in fact challenging to validate statistical models generated from a large number of metabolites (predictors) measured in a small number of samples and subjects, which is commonly seen in underpowered clinical metabolomics studies (Grootveld 2015). Therefore, researchers must be aware of the many pitfalls in applying supervised techniques, such as PLS-DA, and ensure that study is followed up with an independent validation or hold-out sets to adequately replicate study outcomes as required for lead candidate biomarker selection. For interested readers, authors refer to the educational tutorial on PLS-DA provided by Brereton *et al.* (Brereton and Lloyd 2014).

Recent advances in MS instrumentation have dramatically improved sensitivity and specificity as required for the detection and resolution of thousands of metabolites and their isomers in complex biological samples. However, these hardware developments did not accompany software tools needed to efficiently process highly complex mass spectral data sets while accurately assigning chemical identities to the large number of signals when performing full-scan MS acquisition. For example, identification of an authentic signal that corresponds to a unique metabolite requires careful spectral deconvolution and data filtering due to the prevalence of random noise, background/contaminant ion signals from matrix or solvents, as well as redundant signals derived from various in-source fragments, salt adducts and isotope signals. Indeed, non-targeted metabolite profiling applications require rigorous data filtering strategies to eliminate more than 90% of unreliable, spurious and redundant molecular features typically generated in ESI-MS experiments, which is a major contributor of false discoveries (Mahieu and Patti 2017). Removal of these non-

informative signals is by no means trivial and various vendor-specific software strategies have been developed to accelerate mass spectral deconvolution while defining each metabolite by its characteristic retention/migration time and most likely molecular formula, which is derived from its accurate mass (m/z), charge state and isotopic pattern. Open-access software tools, such as CAMERA have been developed for spectral deconvolution of raw MS data that is independent of instrument vendor (Kuhl et al. 2012), which has been integrated in the workflow of other widely used metabolomics data processing tools, such as XCMS (Mahieu et al. 2016). Once spurious and redundant signals are removed from the dataset, structural identification of an unknown metabolite ideally requires matching with an authentic chemical standard (*i.e.,* co-elution and MS/MS spectra match) since a molecular formula is often associated with hundreds of known isobaric or isomeric candidates reported within public compound/metabolome libraries, such as Human Metabolome Database, Metlin and PubChem. In the absence of authentic chemical standards and well curated MS/MS spectra databases for comparison, metabolite identification is limited to *de novo* structural elucidation and increasingly *in silico* assignment of MS/MS spectra (*i.e.,* fragment ions and neutral losses) when performing collisional-induced dissociation (CID) experiments at different energies. This process can take a few hours to months depending on structural complexity of a given metabolite and *apriori* biochemical information available and thus, represents a major bottle-neck in the metabolomics pipeline (Bowen and Northen 2010; Hegeman 2010; Neumann and Böcker 2010; David S. Wishart 2011). In an attempt to standardize the unknown metabolite identification process when using different analytical platforms, the Metabolomics Standards Initiative (L. W. Sumner et al. 2007) proposed minimum reporting standards based on four confidence level categories and recommended supporting data required to satisfy each level (**Table 1.1**). For unambiguous confirmation of chemical identity (level 1), a minimum of two orthogonal data are required, such as a high similarity score when comparing MS/MS spectra from authentic standard and unknown metabolite acquired under identical operating conditions, as well as compound co-elution when spiked in a representative sample.

Table 1.1 Metabolite identification confidence level proposed by Chemical Analysis Working Group in Metabolomics Standards Initiative. Reproduced from (Lloyd W. Sumner et al. 2007).

Confidence level	Description
1	Identified compounds based on minimum of two independent orthogonal data (*e.g.* MSMS spectra and retention/migration time match) using authentic chemical standard or synthesized and validated compound.
2	Putatively annotated compounds based on physicochemical property and spectral similarity to spectra in public or commercial data repository.
3	Putatively characterized compound classes based on physicochemical property and spectral similarity with known compound class.
4	Unknowns (*i.e.* the compound can be differentiated and quantified but its identity is completely unknown)

Additional information to support exact stereochemistry of a putative metabolite is also important biologically, such as confirmation of a specific enantiomer is often not achieved however. Attaining this level of identification is extremely challenging since MS/MS spectra available in databases or literature are still substantially smaller than the total number of expected metabolites especially when considering the large chemical diversity of the human metabolome, including drug metabolites, environmental chemicals and dietary components (Peironcely et al. 2011). For example, the total number of metabolites reported in the Human Metabolome Database exceeds 114,000, however only 867 have corresponding MS/MS spectra deposited (www.hmdb.ca/statistics; accessed on 09/15/2018) (D. S. Wishart et al. 2009). Furthermore, the variety of instrument vendors, ion optic and mass analyzer designs (*e.g.,* triple quadrupole, Orbitrap, Q-TOF, and 3D ion trap) makes MS/MS spectral comparisons difficult to perform due to highly variable spectral output even when the same magnitude of voltage is applied for CID of a precursor ion. As a result, there is growing emphasis in depositing original metabolomics meta-data and detailed experimental information, including study design and MS/MS spectra in public accessible depositories, such as MetaboLights (Haug et al. 2013) and Metabolomics Workbench (http://www.metabolomicsworkbench.org/).

A bioinformatics tool recently developed to facilitate this structural elucidation process makes effective use of known metabolites already identified within a sample in order to reduce the number of possible unknown candidates in the same sample (Yamamoto and Sasaki 2017). This strategy is based on an assumption that majority of unknown compounds structurally resemble other metabolites that have already been identified in the same sample because of shared biochemical pathways and enzymatic reactions. Peironcely *et al.* (Peironcely et al. 2013) also developed a software tool independent of the MS/MS database content by utilizing fragmentation trees of unknown compound matched with that of compounds in public databases. A fragmentation tree is created by fragmenting a molecule, isolating the resulting ion and further fragmenting the ion multiple times. The resulting pattern is generally specific to the parent ion and can be used to match with the same or similar fragmentation tree available in public databases. In this way, even a partial match can indicate the presence of certain functional groups or characteristic chemical moieties without having the exact match present in databases. The ongoing demand and interest in developing more efficient and effective method of identifying unknowns resulted in the annual contest of 'Critical Assessment of Small Molecule Identification' (CASMI) whereby participants are given mass spectra and additional meta-data in some categories and identify the chemical identity based on the data. In recent years, the contest features automated method as more interest is naturally placed on machine learning and automating unknown identification of a large number of compounds. Participants include many of developers involved in maintenance and improvement of *in-silico* mass spectra fragmentation tools, such as CSI: FingerID (Dührkop et al. 2015), MetFrag (Wolf et al. 2010) and CFM-ID (Allen et al. 2015), which aid the unknown compound identification without authentic mass spectra. Indeed, the team of researchers who developed CSI:FingerID won the second place for the best automatic structural identification category in the recent contest (Schymanski et al. 2017). Results of CASMI showcased great progress in automated structure annotation, which will greatly facilitate untargeted analysis with higher throughput. However, challenges remain identification of true unknowns that are not present in databases, as well as distinguishing stereoisomers and positional isomers

(Schymanski et al. 2017). Overcoming these challenges require increased database contents, resulting from continuous effort in unknown identification through manual annotation, and further development in MS instrumentation with improved resolution and sensitivity.

1.3.5   Biomarker validation and translation into clinical practice

It is essential to link changes in the metabolome to phenotype and physiology in order to properly evaluate the clinical utility of new biomarkers identified by metabolomics for improved patient outcomes. The term 'biomarker' has been used routinely and includes everything from a pulse and blood pressure measurement to more complex biomolecules that are measured in laboratory medicine, such as serum CRP (Strimbu and Tavel 2010). A validated biomarker in a clinical setting refers to a broad range of molecular signatures that objectively reflect a patient's physiological state reproducibly and accurately, which is useful in early clinical decision making notably when symptoms are non-specific or not apparent. The diagnostic, prognostic or predictive accuracy of a biomarker(s) can be evaluated using two parameters when performing receiver operating characteristic (ROC) curves, namely sensitivity and specificity. The former is fraction of actual positive cases that are correctly assigned as positive, whereas the latter is the proportion of actual negative cases that correctly assigned as negative, respectively (Xia et al. 2013). The interest in clinical biomarker discovery using "omics" technologies continues to increase as seen in the constant growth of biomarker related publications and NIH funding in the past two decades (Ptolemy and Rifai 2010). Nevertheless, the number of successful biomarkers that have successfully passed rigorous evaluation and implemented into routine clinical applications is extremely small. Although most putative biomarkers are subsequently demonstrated to lack adequate sensitivity, specificity, robustness and/or cost effectiveness to be implemented in a routine clinical setting as compared to conventional tests, they may still contribute to better understanding of underlying molecular mechanisms of disease pathophysiology (Johnson et al. 2016). The process of biomarker evaluation can be summarized in five phases as described by Sullivan *et al* (Pepe et al. 2001). Phase I involves preclinical exploration of a population/cohort of interest to discover and prioritize a

potentially useful biomarker or panel of biomarkers, which typically involves a retrospective cross-sectional study with suitable control. Most metabolomic studies are implemented within phase I for biomarker discovery in small pilot studies. Phase II involves extensive validation of lead biomarkers within clinical settings with adequate study power while assessing the influence of confounding factors, such as age, sex ethnicity and BMI to estimate true-positive and false-negative rates. Additionally, validated assays are typically implemented at this stage for targeted and accurate analysis of well-characterized biomarker candidates using standard operating protocols. The subsequent phase focuses on a longitudinal study of the a small panel of biomarker(s) by assessing its capacity to detect preclinical disease, as well as to explore potential performance to differentiate patients with similar clinical conditions as compared to gold standard methods used for diagnosis, prognosis or treatment monitoring. In phase IV, the lead biomarker goes through a prospective screening evaluation whereby operating characteristics are determined based on detection rate, false referral rate and positive predictive value. In this phase, the potential benefits from implementing the biomarker in clinical testing is evaluated in terms of positive clinical outcomes for patients, as well as healthcare savings. Finally, phase V refers to the control study whereby the impact of the biomarker use in diagnosis or prognosis in the population is quantitatively evaluated. New biomarkers for screening, diagnosis and monitoring of therapeutic responses for patients hold great promise in many medical conditions especially chronic disorders, such as IBS and IBD, whose etiology and pathological mechanisms remain poorly understood. Indeed, there is increase in direct-to-consumer metabolite profiling services (*e.g.,* Molecular You) as applied for health evaluation and chronic disease risk assessment that also provides customized feedback on recommended changes in diet, physical activity, sleep habits and lifestyle.

1.4   Current Progress in Metabolomics of IBS and IBD

To supplement already published reviews on diagnostic potential of metabolomics for IBS and IBD by De Preter and Verbeke (De Preter and Verbeke 2013) and on overall

clinical application of metabolomics for these GI disorders by Collino *et al.* (Sebastiano et al. 2013), sections below will focus on more recently published metabolomics studies from 2013 to 2018. Representative metabolomic studies as related to the mechanistic understanding of pathophysiology, as well as biomarkers for improved diagnosis and therapeutic monitoring of IBS and IBD are summarised in **Table 1.2** and **1.3**, respectively.

1.4.1 Etiological investigation and discoveries for diagnosis of IBS

One of the key factors considered to be associated with onset and development of IBS is disruption in gut microbiota as a result of a complex interplay between infection, life-style, psychosocial stress and genetic susceptibility (G. Barbara et al. 2002; Rodríguez and Ruigómez 1999). In fact, the strongest risk factor in developing IBS is a previous history of bacterial infection with pathogens such as *Campylobacter*, *Salmonella*, *E. coli,* and *Shigella* (K.-A. Gwee et al. 2003; R. C. Spiller et al. 2000; Wang et al. 2004). Dysbiosis has been identified in stool and breath profiles of IBS patients, revealing altered intestinal microbiota with increased gastrointestinal fermentation (King et al. 1998; Lin 2004; Pimentel and Lezcano 2007). Interestingly, different microbial composition has been reported between IBS-C and IBS-D patients, indicating the strong association between microbial metabolism and variable symptoms of IBS (Malinen et al. 2005). As a result of this dysbiosis in the GI tract, excessive formation of acetate and lactate (Tana et al. 2010), imbalance of serotonin production (Coates et al. 2004) and increased production of gas (Kassinen et al. 2007) have been implicated in the characteristic symptoms of IBS, including abdominal cramps, altered bowel movement, and bloating. The influence of GI microbial communities is not limited to physiological symptoms but also to psychological symptoms such as anxiety as observed in mice models which exhibit anxiety-like behaviour upon bacterial infection with *Helicobacter pylori* (Verdu et al. 2008; Bercik et al. 2009). Also, this brain-gut axis is bidirectional in such a way that higher stress levels and predisposed neuroticism significantly increases the chance of developing post-infection IBS, further indicating the important correlation between neurological pathway and immune systems in the GI tract (K. Gwee et al. 1999). The metabolic cross-talk indicated

in these earlier studies indicate that a panel of metabolites and proteins may be useful to distinguish IBS patients from individuals without GI symptoms. Mujagic *et al*. (Zlatan Mujagic et al. 2016) conducted a targeted analysis of plasma and fecal samples collected from 196 IBS patients as confirmed by Rome III criteria and 160 healthy controls, aiming to validate the performance of 43 biomarker candidates generated from an extensive literature search. Subsequently analysis reduced the panel to eight lead biomarkers comprising cytokines, proteins associated with immune activation and neuroendocrine activity, as well as caproic acid. This plasma and fecal biomarker panel demonstrated excellent sensitivity (88.1%) and specificity (86.5%) in discriminating IBS patients from healthy controls within the confines of this retrospective case-control study. Additionally, the concentrations of two short-chain fatty acids, namely caproic acid and valeric acid, were found to be at particularly low concentrations in fecal extracts from IBS patients, suggesting the involvement of gut microbiota and host-microbial interactions on immune activation and neuromodulatory alteration. However, their optimal biomarker panel required five independent analytical methods, such as enzyme-linked immunosorbent assay (ELISA) and GC-MS, making the application of this biomarker panel challenging and cost prohibitive for routine clinical settings. On the other hand, Baranska *et al*. (Baranska et al. 2016) analyzed volatile organic compounds (VOCs) in breath condensates when using GC-MS as putative diagnostic biomarkers of IBS that demonstrated excellent sensitivity and specificity. They further validated the biomarker panel internally with 47 IBS patients and 30 healthy controls and correctly predicted group assignments for 89.4% of IBS patients and 73.3% of healthy controls. This 16 VOC panel mainly consisted of hydrocarbons, such as heptane and 1,4-cyclohexadiene that have been found to be associated with oxidative stress and immunological changes in IBS patients (Mete et al. 2013). Interestingly, these VOCs also showed moderate but significant correlation ($r = 0.54$, $p = 0.0004$) with GI symptoms recorded by the general population who frequently experience abdominal pain or discomfort, but never have sought medical advice. Since many IBS patients do not consult physicians and remain underdiagnosed (Halder et al. 2007), this finding implies a large hidden population of IBS patients in the general public. Although, they did not

explore the biochemical mechanisms associated with these VOCs nor identify all clinically significant metabolites, they demonstrated the value of breath analysis as a non-invasive and relatively fast diagnostic tool that may be applicable within a clinical setting once fully validated.

Volatile metabolite analysis has also been applied to differentially diagnose IBS patients from similar GI disorders, such as celiac disease. Arasaradnam *et al*. (Arasaradnam et al. 2014) demonstrated the application of a urine volatile compound analysis in a pilot study for differential diagnosis of 20 IBS-D from 47 celiac disease patients. Field asymmetric ion mobility spectrometry (FAIMS) and machine learning technique for data analysis were used to generate distinct chemical fingerprint for each group of patient. Metabolite fingerprinting is a powerful approach for differentiating disease groups based on metabolic patterns as it offers much wider coverage of metabolites, but it does provide information on metabolite identities or their concentration levels that is criticial for establishing references ranges and biochemical insights into mechanisms (Fernie et al. 2004). Therefore, compounds dominant in each fingerprint were subsequently identified by GC-MS as esters and phenolic metabolites at variable concentrations. Ion mobility spectrometry (IMS) is a high-throughput and sensitive technique that is also portable, which is widely applied for drug/explosive screening at airports and military operations (Borsdorf and Eiceman 2006). Validation of these findings in larger and independent cohort is still needed together with identification of clinically significant metabolites from urine, but this proof-of-concept may provide a convenient approach to allow for real-time differential diagnosis of functional GI disorders.

1.4.2   Etiological investigation and discoveries for differential diagnosis in IBD

Similar to metabolomic studies reported for IBS, the strong association between GI inflammation and dysbiosis has been a major theme when exploring the etiology and pathophysiology of IBD in affected patients. With recent innovation in sequencing technology, there have been several studies that included gut microbiome and stool

metabolome analysis in order to better understand the underlying biochemical mechanisms of IBD. Jacobs *et al*. (Jacobs et al. 2016) examined stool metabolome and microbiome collected from 36 IBD patients and their 54 first-degree relatives as healthy controls in order to determine disease-associated metabolites and bacteria. As expected from the genetic factors associated with IBD patients (Jostins et al. 2012), clinically healthy relatives were reported to have microbial and metabolic profiles similar to their diseased relatives, indicating that underlying dysbiosis may be taking place long before clinical symptoms manifest. Moreover, they identified greater amount of amino acids such as taurine and tryptophan in stool extract of IBD patients, using LC-MS. Bile acid profiles of IBD patients were also distinct, which was characterized by higher amount of primary bile acid, such as cholic acid, and sulfated bile acids, such as 3-sulfodeoxycholic acid. In this case, high excretion of amino acids is indicative of nutrient malabsorption, which is likely associated with fatigue and weight loss frequently observed in IBD patients. Consistent with Jacobs *et al*. (Jacobs et al. 2016), Santoru *et al*. (Santoru et al. 2017, 2018) also observed significantly different metabolic and microbial profile between stool samples collected from healthy controls and IBD patients. They demonstrated higher abundance of certain bacterial genera with pathogenic potential in IBD, which may have caused significantly low levels of host-microbial co-metabolites, including 3-hydroxybutyrate and phenylpropanoic acid in stool of these patients. Metabolic signatures of nutrient malabsorption were also observed in this study, further supporting earlier observations of attenuated energy metabolism caused by perturbations in serum metabolites involved in amino acid cycle and TCA cycle (Scoville et al. 2017).

The direct link between activity of microbiota and metabolome changes were not clarified in these studies but they presented a potential application of metabolomics for risk assessment of IBD based on fecal metabolites associated with gut microbiome and energy metabolism. In contrast to the clear differences measured in metabolic profiles of IBD patients and healthy controls, differentiating IBD-subtypes (*i.e.* UC vs CD) has been challenging with ambiguous results.

Table 1.2. Published metabolomic studies in 2013-2018 that focused on biomarkers for diagnosis of IBS or IBD.

Reference	Subjects	Sample	Method	Validation method	Key findings
(Zlatan Mujagic et al. 2016)	196 IBS, 160 HC	Plasma, stool	ELISA, immuno-radiography, GC-MS, LC-fluorescent, multiplex immunoassay	Double cross validation with 20% of subjects as hold-out set	Biomarker panel consisting of eight compounds including cytokines and caproate showed sensitivity (88.1 %) and specificity (86.5 %) in discriminating IBS patients from HCs.
(Baranska et al. 2016)	123 IBS, 123 HC	Breath	GC-MS	Internal validation cohort: 47 IBS, 30 HC	16 VOC panel that consists of hydrocarbons showed 89.4 % sensitivity and 73.3 % specificity in discriminating IBS from HCs.
(Ahmed et al. 2013)	30 IBS-D, 62 CD, 48 UC, 109 HC	Fecal volatiles	GC-MS	Leave-one-out cross validation	Esters, SCFAs, and cyclohexanecarboxylic acid derivatives were altered in IBS-D. Aldehydes were more abundant in IBD patients.
(Arasaradnam et al. 2014)	20 IBS-D, 47 celiac disease	Urine volatiles	FAIMS, GC-MS	Leave-one-out cross validation	Distinct chemical fingerprint abundant in phenols and esters differentiated celiac disease patients from IBS-D patients with 85 % specificity and sensitivity.
(Jacobs et al. 2016)	36 pediatric IBD (26 CD, 10 UC), 54 relatives	Stool	LC-MS, 16S rRNA	NA	IBD-associated mirobiome and metabolome types were characterised by signs of dysbiosis and altered energy metabolism. These types were identified in subset of relatives along with high fecal calprotectin.
(Santoru et al. 2017, 2018)	50 CD, 82 UC, 51 HC	Stool	GC-MS, NMR, 16S rRNA	NA	Significantly high levels of microbial metabolites (*e.g.* 3-hydroxybutyrate) along with opportunistic pathogens in phylum *Proteobacteria* and *Fusobacteria* were found at elevated levels in IBD patients.
(Scoville et al. 2017)	20 CD, 20 UC, 20 HC	Serum	LC-MS	NA	Lipids, amino acids, and energy-related metabolites were significantly altered in IBD especially in CD. Specifically, levels of fatty acids and carnitines were lower in CD than UC and HC.
(N. S. Stephens et al. 2013)	30 CD, 30 UC, 60 HC	Urine	NMR	7-fold cross validation, permutation test	Metabolites in TCA cycle (*e.g.* succinate), amino acids (*e.g.* histidine) and microbial metabolites (*e.g.* hippurate) were significantly lower in urine of IBD patients.

Abbreviations: HC: Healthy control; NMR: Nuclear magnetic resonance; FAIMS: Field asymmetric ion mobility spectrometry; SCFA: Short-chain fatty acid; FODMAP: Fermentable oligosaccharide, disaccharide, monosaccharide and polyol; EEN: Exclusive enteral nutrition; FMT: Fecal microbiota transplantation

For example, active and inactive disease states in affected IBD patients contributes to significant within subject biological variability, which often obscures effects reported in pilot studies that are difficult to replicate in other cohorts (Bjerrum et al. 2015; Topping and Clifton 2001). Similarly, earlier observations of significantly lower levels of host-microbial co-metabolites in urine of CD patients, such as *p*-cresol sulfate and hippurate (Williams et al. 2009; Schicho et al. 2012), were not observed by Stephens *et al* (N. S. Stephens et al. 2013), who took account for the individual history of surgery as a confounding factor. Furthermore, analysis of breath VOCs showed no difference between UC and CD patients even though it was effective for discriminating between IBD and other GI complications (Patel et al. 2014; Rieder et al. 2016).

Taken together, metabolomics of various biospecimens ranging from stool, urine, serum to breath have provided insights into the involvement of altered gut microbiota and concurrent changes in host metabolism based on differences between IBD patients and healthy controls. However, metabolic difference between CD and UC is still inconclusive largely due to patient heterogeneity in terms of disease severity and medical histories within an adult IBD population.

1.4.3   Metabolomics for treatment response prediction and monitoring

In comparison to retrospective case/control studies focused on differential diagnosis of GI disorders or the characterization of the metabolic phenotype of IBS/IBD patients, surprisingly few studies have conducted longitudinal investigations of metabolic changes associated with therapeutic interventions and/or disease progression. In this section, recent metabolomic studies summarized in **Table 1.3** will be discussed with a focus on biomarkers that may be used for therapeutic monitoring and prognosis of IBS and IBD patients. As discussed earlier, low FODMAP diets has shown efficacy in IBS and IBD patients as determined by symptom-based disease scores, however the underlying mechanisms of this intervention is yet to be clarified. McIntosh *et al*. (McIntosh et al. 2017) conducted a single-blinded controlled study by randomly assigning 40 IBS patients into either low or high

FODMAP diet arms and then investigated urinary and fecal microbiome changes associated with therapy responses using LC-MS and GC-MS. In this case, disease scores decreased in nearly all patients in the low FODMAP diet treatment arm while a slight mean increase was observed in high FODMAP diet group, supporting the general hypothesis of fermentable polysaccharides being causative agents in IBS symptoms. Interestingly, the urine metabolome changed significantly in low FODMAP arm following the intervention, characterized by reduction in *p*-hydroxybenzoic acid, azelaic acid, including an eight-fold average reduction in urinary histamine. Histamine is an essential signaling molecule associated with increase in mast cell activation and sensitization of nociceptive and enteric neurons, implying its direct involvement in heightened pain perception among IBS patients (Giovanni Barbara et al. 2007; Buhner et al. 2009). The exact mechanism of lower urinary histamine excretion is still unclear, but it may be applied as an objective measure of IBS disease progression and remission following therapy. The two other urinary metabolites, *p*-hydroxybenzoic acid and azelaic acid, likely originated from grains and modified by bacterial enzymes in the gut, further emphasizing importance of the dynamic interplay of habitual dietary intake and host-microbial co-metabolism. Since the motivation of a low FODMAP diet originates from observations that bacterial fermentation of certain carbohydrate is associated with production of gas, resulting in bloating and distention, it is conceivable that volatile compounds in stool may also serve as specific biomarkers of therapeutic responses to treatment for IBS patients. Building on this hypothesis, Rossi *et al.* (Rossi et al. 2018) investigated fecal VOC profiles of 93 IBS patients enrolled in a 2 X 2 factorial, multicenter, randomized, placebo-controlled trial of low FODMAP and probiotics efficacy for IBS. They used in-house GC-based sensor device called Odoreader (Raphael et al. 2016) for pattern recognition and classification of samples collected from participants before and after the 4-week therapy period. When 44 patients and 45 patients were randomly assigned to low FODMAP diet and probiotics, respectively, VOC profile of before and after the trial period clearly differentiated responders and non-responders to both therapies, especially for low FODMAP diet (**Figure 1.6**).

Table 1.3. Published metabolomic studies in 2013-2018 that involve longitudinal investigations of therapeutic interventions for IBS or IBD patients.

Reference	Subject	Sample	Method	Therapy	Key findings
(McIntosh et al. 2017)	40 IBS	Urine, stool	LC-MS, GC-MS, 16S rRNA	FODMAP	Histamine, p-hydroxybenzoic acid and azelaic acid decreased significantly in urine of low FODMAP diet group. IBS symptom score decreased by 28% in low FODMAP Higher abundance of *Actinobacteria* found in stool of low FODMAP group
(Rossi et al. 2018)	93 IBS	Stool	Odoreader (GC)	FODMAP	Fecal volatile organic compound profile predicted therapy response with 100% accuracy
(Zeber-Lubecka et al. 2016)	31 IBS-D, 11 IBS-D, 30 IBS-M, 30 HC	Stool	NMR, GC-MS	Antibiotic	Baseline differences in microbiota and lipophilic metabolite profiles between HCs and IBS patients observed. No difference observed between treatment responders and non-responders.
(Nikolaus et al. 2017)	51 CD, 36 UC	Serum	LC-MS	Infliximab (biologic)	Serum tryptophan level significantly increased only in responders and remained high during a therapy Patients who later required surgery had significantly low tryptophan level
(Gerasimidis et al. 2014)	15 pediatric CD	Stool	NMR, 16S rRNA	EEN	Butyrate decreased and sulfide rapidly increased during EEN. Overall bacterial diversity decreased. Decrease of butyrate producing *Faecalibacterium prausnitzii* correlated well with reduction in colonic inflammatory marker
(Keshteli et al. 2017)	20 UC	Urine, serum	DI-, LC-MS	NA	Model based on serum and urinary metabolites predicted relapse after 12 months from initial sample collection
(Nusbaum et al. 2018)	4 pediatric UC	Stool	GC-MS, 16S rRNA	FMT	Metabolomic profiles cluster separately among fecal donor, patients before FTM and after FMT. Butyrate and butyrate producer, F. *prausnitzii* increased in samples from FMT responders. Other microbial metabolites such as putrescine and 5-aminovaleric acid also shifted toward donor levels after FMT.

Abbreviations: FODMAP: Fermentable oligosaccharide, disaccharide, monosaccharide and polyol; EEN: Exclusive enteral nutrition; FMT: Fecal microbiota transplantation
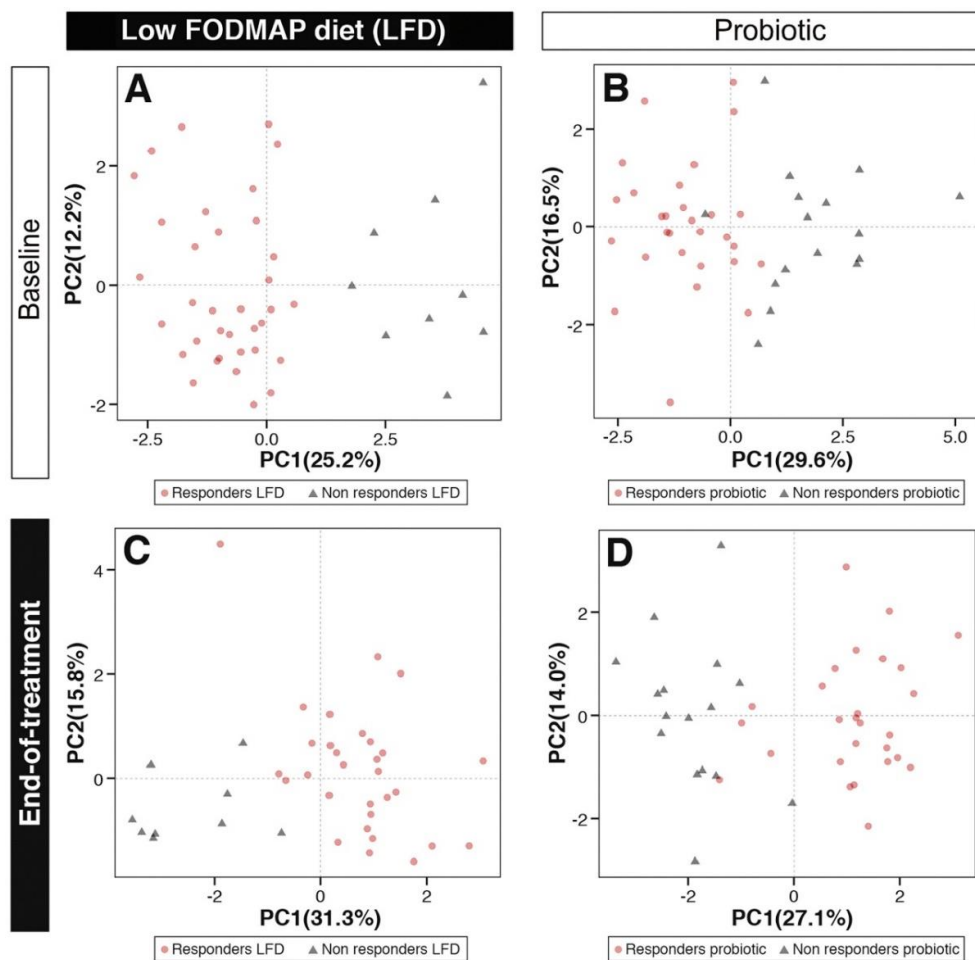
Figure 1.6. PCA plots representing fecal VOC profiles of IBS patients before and after low FODMAP diet or probiotic therapy. (Reproduced with permission from (Rossi et al. 2018))

Indeed, the PLS model developed based on the low FODMAP arm predicted the response to the diet with sensitivity and specificity of 100 %. Although the GC sensor device did not provide information about metabolite identification or absolute concentration, total analysis time was less than one hour per sample for the whole process including sample preparation, which is ideal for rapid screening in clinical settings prior to prescribing a therapy to a patient. Low FODMAP diet is known for its efficacy but clinical trials have shown that 20-50 % of patients do not respond to the therapy (Rossi et al. 2018). This may be the result of differences in host and gut microbial metabolism and/or low compliance. Therefore, appropriate education, support and selection of IBS patients who

are receptive to adhere and respond to a low FODMAP diet is needed to ensure positive long-term clinical outcomes.

Metabolomic studies have also explored the predictive value of certain metabolites for prevention of near-future relapse, which may allow for timely interventions before a severe case of relapse and hospitalization. In the in-depth analysis of association between IBD activity and tryptophan metabolism, Nikolaus *et al*. (Nikolaus et al. 2017) reported significantly low serum tryptophan levels in patients who later required intestinal resection within 100-200 days of initial surgical intervention. The immunomodulatory and therapeutic role of tryptophan have been shown previously (Hashimoto et al. 2012; Bettenworth et al. 2014). For the first time, their study thoroughly investigated the mechanism behind this therapeutic effect and applicability in the diagnosis, prognosis and therapeutic monitoring using a targeted metabolomics, metagenomics and transcriptomics approach. They demonstrated that serum tryptophan concentrations increased and remained high only in patients who responded to a biologic (infliximab) treatment while levels in non-responders initially increased, but dropped after the second week, suggesting its applicability in identifying responders to the same therapeutic intervention. Also, Gerasimidis *et al*. (Gerasimidis et al. 2014) conducted a longitudinal monitoring on EEN therapeutic response by analyzing fecal metabolome and microbiome of 15 pediatric CD patients by sampling at five different time points, including baseline and post-therapy when patients were back on habitual diet. Their targeted analysis of short-chain and branched-chain fatty acids by GC-MS showed an unexpected decrease in butyrate, which is considered to be beneficial for GI health, while patients showed symptomatic improvement and reduction in fecal calprotection. Concurrently, concentration of butyrate-producing and presumably beneficial *Faecalibacterium prausnitzii* declined, which was interestingly more pronounced for responders to the therapy. The study also demonstrated increase in *Bifidobacteria spp.* as well as decrease in species generally associated with active state of IBD, such as *Bacteroides/Prevotella spp.* and *E. coli.* during the therapy (Kleessen et al. 2002). Additionally, one of short-chain fatty acids, propionate, positively correlated

(Spearman rho = 0.58) with increase in fecal calprotectin when patients returned to habitual diet, indicating a potentially role of this compound or the microbial metabolism involved in its production in future relapse. Analysis of fecal microbiome and their metabolites in this study provided evidence of drastic changes taking place during dietary intervention for CD and insights in beneficial roles of certain bacterial strains without the production of butyrate but by competing against opportunistic pathogens.

1.4.4   Trends, challenges and gaps in current research

Recently published metabolomics studies on IBS and IBD were reviewed with a focus on metabolites that may serve as promising biomarkers for diagnostic, prognostic, and therapeutic monitoring applications. A common theme observed in these studies is the inclusion of gut microbiome metagenomics data in parallel or partially integrated with metabolomics data. This is a positive trend but progress in the integration of these two – omics knowledge is still in its infancy. Correlation between certain metabolites and diversity measure of bacterial community richness can be found in some studies (Nikolaus et al. 2017), but it does not lead to pertinent insights in certain species of bacteria and more importantly, their functional roles in association with gastrointestinal health. As discussed in a recent review by Chong *et al.* (Chong and Xia 2017), there are emerging computational techniques that allow integration of detailed bacterial community structure with metabolomics data, which have already been applied to understand various physiological processes and health conditions, such as digestion of fermented food (Quinn et al. 2016) and obese subjects (Greenblum et al. 2012), but has not yet been applied for IBS and IBD. Utility of these integration tools and online databases highly depend on the quality of data; however, culture-independent techniques widely used to acquire microbiome data generally lacks conclusive functional insights and does not provide enough sensitivity to sequence genes of bacteria that are present in samples with low abundance (Lau et al. 2016). This issue can be mitigated by integrating culture enrichment techniques with an existing molecular profiling workflow with 16S rRNA or MALDI-TOF-MS, as discussed earlier. Indeed, Lagier *et al.*(Lagier et al. 2016) used multiple culture conditions in anaerobic and

aerobic environment and doubled the number of prokaryotic species identified in the human gut. As bacterial metabolisms highly depend on environmental factors with large influences from competitive or symbiotic relationship with other bacterial strains, identifying low abundant strains can greatly aid our understanding of metabolic potential of a bacterial community and their effects on GI health and disease.

Additionally, there is lack of attention given to other important constituents of gut microbiome, namely, fungi and viruses. About 50 genera of fungi are generally reported in investigations of human mycobiota, and their involvement in health and disease is expected from their widely known ability to ferment carbohydrates and potential competition against commensal bacteria, but it is scarcely investigated (Paterson et al. 2017). One recent study on GI fungal population in IBS patients clearly showed difference in community of fungi between healthy controls and IBS patients (Botschuijver et al. 2017). Furthermore, commensal fungi exacerbated colitis in mice when environmental stress, such as exposure to an oxidizing agent, is imposed on them, indicating the important role of resident fungal species in GI inflammation (Chiaro et al. 2017). Adaptive ability under such environmental pressure is expected to be far greater in viruses, which mutate at the fastest rate among other microbes in the gut (Virgin 2014). The most widely investigated role of eukaryotic viruses has been chronic infection of mammals including humans, but their functions in our GI tract goes far beyond the simple model of cell invasion and tissue destruction (Foxman and Iwasaki 2011). In fact, there are multiple types of bacteriophage identified in our gut that are capable of reducing beneficial effects of commensal bacteria in mice models of IBD (Kernbauer et al. 2014). Distinct virome has also been recognised in IBD patients compared to healthy individuals (Norman et al. 2015). Remarkable adaptability and multidirectional interactions among viruses, fungi and bacteria, and their roles in pathology of IBS and IBD are consistently implicated in these studies. Their interactions hold a strong potential in advancing our current understanding of dynamic changes in GI microbiome and thus warrant further investigations with human subjects in health and disease.

Metabolomics research has enabled us to gain significant insights into the human-bacterial interaction, and it is expected to enhance the investigation of other interactions, such as bacterial-viral and bacterial-fungal interactions, in the context of IBS and IBD. Rapid development in instrumental and computational technologies has fuelled and will continue to move this field forward with more integration and collaboration between different –omics approaches. The ubiquitous presence of metabolites in complex biological systems makes metabolomics the ideal strategy to fill the gaps between different –omics and uncover multifactorial mechanisms of complex chronic conditions. With close collaborations with other –omics, metabolomics hold promise in uncovering dynamically changing molecular interactions among participating organisms in disease pathogenesis and progression, which ultimately contribute to development of effective diagnostic, prognostic and therapeutic monitoring strategies.

1.5 Thesis motivation and objectives: Metabolomics of irritable bowel syndrome and inflammatory bowel disease

1.5.1 Overview

IBS and IBD represent two of the most commonly diagnosed chronic digestive disorders in Western countries while prevalence is increasing also in other parts of the world. Both disorders are chronic debilitating GI conditions that impacts overall quality of life of affected patients as there is currently no effective cure. Therefore, increasing incidence indicates significant long-term burden on patient's physical and mental well-being. Despite the prevalence and international efforts to deepen our understanding of these conditions, etiology and pathological mechanisms remain incompletely understood. This lack of understanding has resulted in current challenges in accurate or differential diagnosis, prognosis and therapeutic monitoring especially when using on probiotic, dietary and/or fecal transplant management interventions. In this context, metabolomics hold potential for providing mechanistic insights of these disorders through differentially expressed metabolites, which are essentially the end or intermediate products of biochemical processes involved in symptom manifestation. Additionally, these metabolites will provide

diagnostic and prognostic values once rigorously validated. For the identification of such metabolites in the context of IBS and IBD, urine is considered as an accessible and non-invasive biofluid for clinical measurements of biomarkers. Urine is abundant in highly polar and ionic metabolites, which makes CE-MS as an ideal analytical platform for biomarker discovery in metabolomics; however, the lack of robustness and reproducibility especially for anion metabolite profiling under negative ion mode conditions is frequently encountered in CE-MS when using ammonia-based alkaline buffers. Therefore, the mechanism behind the lack of robustness was investigated and a solution to mitigate this issue was proposed, followed by a rigorous validation of a multiplexed separation method for high throughput yet comprehensive metabolite profiling based on multisegment injection (MSI)-CE-MS (*Chapter II*). This validated method was subsequently applied to the analysis of urine samples collected from a cohort of IBS patients and sex-matched healthy controls to reveal underlying metabolic signatures associated with IBS pathology, including several unknown metabolites of clinical significance to accelerated collagen degradation processes (*Chapter III*). Furthermore, a stool extraction protocol was subsequently developed for untargeted analysis of polar/ionic metabolites by MSI-CE-MS and applied for a comparative analysis of sub-types of pediatric IBD patients, namely recently diagnosed children with UC and CD in which urinary metabolome and stool bacterial profiling data were analyzed to provide more holistic view of underlying metabolic and gut microbiota changes involved in different IBS pathologies (*Chapter IV*). Finally, an integrative approach with urinary and stool metabolome, as well as stool microbiome data, was applied in a longitudinal clinical intervention study to investigate the therapeutic mechanism of EEN therapy in a cohort of pediatric CD patients who largely achieved remission following intervention (*Chapter V*). In this case, metabolic trajectories were identified for CD patients as a function of time course to identify biomarkers associated with positive treatment responses, such as mucosal healing and gut microbiota activity along with an attenuation in oxidative stress and immune activation.

1.5.2 Robust analysis of anionic metabolites in human urine by MSI-CE-MS

Human urine contains a wealth of real-time information through metabolic breakdown products of foods, drinks, drugs, endogenous metabolism and human-microbial co-metabolites (Bouatra et al. 2013). Many of urinary metabolites, such as amino acids, organic acids and glucuronide / sulfate conjugates of xenobiotics are highly polar and ionic. Charge and size based separation on CE-MS offers equivalent or superior separation performance to both reverse-phase LC and hydrophilic interaction chromatography (HILIC) when analyzing these strongly ionic and poorly retained metabolites (Ramautar et al. 2011; Ibáñez et al. 2012; Saric et al. 2012; Kok et al. 2015). However, lower sensitivity and poorer reproducibility has been repeatedly observed when analysing anionic metabolites under negative ion mode detection (Büscher et al. 2009), which led current metabolomics study to primarily focus on cationic metabolites (Minami et al. 2009; Sugimoto et al. 2010; Ramautar et al. 2011; Alberice et al. 2013; Kuehnbaum et al. 2013; Kuehnbaum et al. 2015). Majority of reports on anionic metabolite analysis by CE-MS are with the use of polymer-coated capillary (Soga et al. 2009; Soga et al. 2006; Soga et al. 2002). Conversely, the use of uncoated capillary with simple alkaline buffer under normal polarity has scarcely been reported (Sato and Yanagisawa 2010). *Chapter II* describes a thorough investigation of the effect of different buffer composition on the stability and mechanical strength of the polyimide-coated fused-silica capillary, their performance in resolution and signal intensity, and finally, optimized anion analysis condition was validated by 3-day continuous analysis of pooled human urine sample, resulting in acceptable reproducibility (% CV 8.8 – 28.1) for 30 representative anionic metabolites. Aminolysis of polyimide coating in ammonia-based buffer is reported for the first time, which is linked to previously observed lack of robustness when using ammonia-based buffer for anionic metabolite analysis.

1.5.3 Urinary metabolome of irritable bowel syndrome patients

Etiology and pathological mechanisms of IBS had been proposed based on animal models, which were followed by targeted metabolomics studies on human subjects

identifying the potential involvement of serotonin-pathway, microbial metabolism and low-grade immune activation behind variable symptoms of IBS (Enck et al. 2016; Keszthelyi et al. 2013; Yano et al. 2015; Christmas et al. 2010; Ahmed et al. 2013; Bercik et al. 2005; Bercik et al. 2010). Serologic biomarkers, metabolites in serum and plasma, as well as in stool, have been studied to date(Michael Camilleri et al. 2017); however, untargeted comprehensive characterization of urine metabolome of IBS patients when compared to healthy controls is yet to be explored. *Chapter III* reports the untargeted metabolic profiling of urine samples collected from IBS patients ($n = 42$) at two separate time points and healthy controls ($n = 20$) to elucidate underlying metabolic difference between the two groups that gives us an insight into the pathogenesis of IBS. In this study, multi-segment injection (MSI)-CE-MS was used for high throughput analysis of urine samples with improved data fidelity when using dilution trend filter and inclusion of QC sample in every run to monitor technical variation. The analysis resulted in 143 authentic urine metabolites, including drug conjugates, amino acids and their derivatives, and sulfated microbial metabolites, such as indoxyl sulfate and p-cresol sulfate. Additionally, unknown metabolites that were highly elevated in IBS samples, were confidently identified to be mannopyranosyl-tryptophan, galactosyl-hyroxylysine, and glucosylgalactosyl-hydroxylysine, when using high resolution mass spectrometry MS/MS and spike experiments with authentic chemical standards. These metabolites, along with other modified amino acids (imidazole propionate, dimethylglycine) and most of the primary amino acids (glutamine, serine, lysine, and ornithine) were highly elevated in urine samples of IBS patients at both time points regardless of normalization parameter (*i.e.* osmolality, creatinine). Mannosylated tryptophan has been reported to exist in proteins responsible for immune functions, such as cytokine receptor family I protein (Ihara et al. 2015), supporting the low-grade immune activation as one of underlying mechanisms of IBS symptoms. Additionally, elevated excretion of imidazole propionate, a histidine metabolite that shares a pathway with histamine, further supports the immune activation hypothesis. Unexpected findings in this study was the high excretion of glycosylated hydroxylysine, which are highly prevalent in collagen, a key structural protein found in extracellular space and

connective tissues, as well as intestine (William 1978; Hilska et al. 1998). High excretion of these urinary metabolites in samples of IBS patients compared to healthy controls suggested that mucosal tissue damage may be associated with the low-grade immune activation in IBS pathology.

1.5.4 Urinary and stool biomarkers for differential diagnosis of IBD

There are currently no biomarkers to differentially diagnose subtypes of IBD, namely ulcerative colitis (UC) and Crohn's disease (CD) (Iskandar and Ciorba 2012). Physicians rely on multiple diagnostic tools, including serum and fecal testing, imaging, clinical history, and invasive colonoscopy (Soubières and Poullis 2016). Despite the invasiveness of the procedure, differentiation of CD and UC in pediatric population is challenging due to severe and unusual manifestation of the disease when compared to adult population, resulting in undefined category of IBD (Thierry 2013). CD and UC differ not only the location of inflammation but also response to therapies (Charles N. Bernstein et al. 2010). Therefore, inconclusive diagnosis or misdiagnosis will result in delay in effective treatment and potential exacerbation of symptoms. In this context, metabolomics hold potential in revealing the underlying metabolic differences between CD and UC, which can be applied as novel non-invasive biomarkers. In *Chapter IV*, extraction protocol of lyophilized stool samples are introduced for the analysis of polar, ionic metabolites, which complements findings of gut microbial structural and functional metagenome data, as well as metabolome from matching urine samples collected from newly diagnosed pediatric 19 CD and 11 UC patients. Sufficient extraction efficiency was achieved ($> 85 \%$) based on recovery standard, which was monitored throughout the individual sample analysis. MSI-CE-MS workflow with rigorous quality control resulted in 72 and 122 authentic metabolites in stool and urine respectively. CD patients are uniquely characterized by significantly higher excretion of host-microbial co-metabolites in urine, such as phenylacetylglutamine, phenylsulfate, indoxylsulfate and hydroxyindole sulfate when compared to UC patients. Conversely, affected UC children had higher urinary excretion of kynurenine, hypoxanthine, serine and threonine regardless of normalization parameter. Additionally, excellent discrimination of

CD from UC patients was observed based on their measured urinary serine/indoxylsulfate ratio (AUC = 0.974; p = 2.73 E-5) with stability studies demonstrating reliable measurements for over 2 days when samples are refrigerated. Additionally, significantly higher excretion of fecal butyrate and bile acids were measured in samples of CD patients, whereas lactate, choline and tryptophan excretion was lower compared to UC patients. Fecal microbiome metagenomics data indicated the presence of significantly higher abundance of pathogenic bacteria in phylum *Proteobacteria* in UC patients, while stool samples of CD patients were enriched with microbial activities in tryptophan and phenylalanine metabolism, which was consistent with the findings of higher excretion of indoles and phenols in urine of CD patients. Taken together, this study showcased a panel of novel urinary biomarkers that offers potential for non-invasive differential diagnosis of IBD patients.

1.5.5 Urinary and stool metabolome to investigate mechanism of efficacy of EEN

EEN therapy offers equivalent rate of remission to corticosteroid therapy for CD patients and promotes linear growth without the adverse side effects frequently associated with steroidal therapy, such as growth retardation and loss of bone mineralization (Lafferty et al. 2016; Kansal et al. 2013). Additionally, multiple studies have shown achievement of mucosal healing through EEN therapy, which were associated with prolonged period of remission. EEN therapy has become the first-line therapy for pediatric CD for these reasons (Grover et al. 2014; Grover and Lewindon 2015); however, the underlying mechanism of efficacy is poorly understood. *Chapter V* reports the dynamic metabolic profiling of matched urine and stool metabolome of 11 pediatric CD patients, who were largely treatment naïve, to better decipher biochemical changes associated with treatment responses to EEN therapy when using MSI-CE-MS in conjunction with fecal microbiome analysis. Correlation analysis between inflammatory markers (*i.e.* C-reactive protein and fecal calprotectin) and metabolites, and paired Wilcoxon rank-sum test identified several metabolites associated with EEN formula (e.g. pantothenate and pyridoxate). This analysis also highlighted significantly lowered metabolite levels associated with proteins in immune activation and oxidative stress, as well as alteration in gut microbiota. For instance,

mannopyranosyl-tryptophan and other glycosylated amino acids involved in glycoprotein turn-over decreased significantly, while specific microbial metabolites, *para*-cresol conjugates, increased during the therapy. Furthermore, metagenomics data of fecal microbiome showed significant decrease of opportunistic pathogens, *Haemophilus* and *Veillonella* during EEN therapy. Collectively, our findings suggested improved nutrition absorption and gut microbiota modulation as key initiators of symptom improvement during EEN therapy, which is followed by recovery in control of immune functions and cellular protective mechanisms. This study is the first report of global metabolic trajectories during EEN therapy for pediatric CD patients, which will lead to development and implementation of effective and convenient tool for monitoring, as well as predicting treatment responses, at the early stage of therapy induction.

1.6 References

Administration, U. S. F. a. D. (2012). Zelnorm (tegaserod maleate) Information. http://www.fda.gov/drugs/drugsafety/postmarketdrugsafetyinformationforpatientsandproviders/ucm103223.htm. Accessed May 18, 2015 2015.

Ahmed, I., Greenwood, R., Costello, B. d. L., Ratcliffe, N. M., & Probert, C. S. (2013). An Investigation of Fecal Volatile Organic Metabolites in Irritable Bowel Syndrome. *PLoS ONE, 8*(3), e58204, doi:10.1371/journal.pone.0058204.

Akarsu, M., & Akarsu, C. (2018). Evaluation of New Technologies in Gastrointestinal Endoscopy. *JSLS : Journal of the Society of Laparoendoscopic Surgeons, 22*(1), e2017.00053, doi:10.4293/JSLS.2017.00053.

Alberice, J. V., Amaral, A. F. S., Armitage, E. G., Lorente, J. A., Algaba, F., Carrilho, E., et al. (2013). Searching for urine biomarkers of bladder cancer recurrence using a liquid chromatography–mass spectrometry and capillary electrophoresis–mass spectrometry metabolomics approach. *Journal of Chromatography A, 1318*(0), 163-170, doi:http://dx.doi.org/10.1016/j.chroma.2013.10.002.

Allen, F., Greiner, R., & Wishart, D. (2015). Competitive fragmentation modeling of ESI-MS/MS spectra for putative metabolite identification. [journal article]. *Metabolomics, 11*(1), 98-110, doi:10.1007/s11306-014-0676-4.

Arasaradnam, R. P., Westenbrink, E., McFarlane, M. J., Harbord, R., Chambers, S., O'Connell, N., et al. (2014). Differentiating Coeliac Disease from Irritable Bowel Syndrome by Urinary Volatile Organic Compound Analysis – A Pilot Study. *PLoS ONE, 9*(10), e107312, doi:10.1371/journal.pone.0107312.

Baranska, A., Mujagic, Z., Smolinska, A., Dallinga, J. W., Jonkers, D. M. A. E., Tigchelaar, E. F., et al. (2016). Volatile organic compounds in breath as markers for irritable bowel syndrome: a metabolomic approach. *Alimentary Pharmacology & Therapeutics, 44*(1), 45-56, doi:doi:10.1111/apt.13654.

Barbara, G., De Giorgio, R., Stanghellini, V., Cremon, C., & Corinaldesi, R. (2002). A role for inflammation in irritable bowel syndrome? [10.1136/gut.51.suppl_1.i41]. *Gut, 51*(suppl 1), i41.

Barbara, G., Wang, B., Stanghellini, V., de Giorgio, R., Cremon, C., Di Nardo, G., et al. (2007). Mast Cell-Dependent Excitation of Visceral-Nociceptive Sensory Neurons in Irritable Bowel Syndrome. *Gastroenterology, 132*(1), 26-37, doi:10.1053/j.gastro.2006.11.039.

Benchimol, E. I., Guttmann, A., Griffiths, A. M., Rabeneck, L., Mack, D. R., Brill, H., et al. (2009). Increasing incidence of paediatric inflammatory bowel disease in Ontario, Canada: evidence from health administrative data. *Gut, 58*(11), 1490-1497, doi:10.1136/gut.2009.188383.

Bercik, P., Verdu, E. F., & Collins, S. M. (2005). Is Irritable Bowel Syndrome a Low-Grade Inflammatory Bowel Disease? *Gastroenterology Clinics of North America, 34*(2), 235-245, doi:http://dx.doi.org/10.1016/j.gtc.2005.02.007.

Bercik, P., Verdú, E. F., Foster, J. A., Lu, J., Scharringa, A., Kean, I., et al. (2009). Role of gut-brain axis in persistent abnormal feeding behavior in mice following eradication of Helicobacter pylori infection. [10.1152/ajpregu.90752.2008]. *American Journal of Physiology - Regulatory, Integrative and Comparative Physiology, 296*(3), R587-R594.

Bercik, P., Verdu, E. F., Foster, J. A., Macri, J., Potter, M., Huang, X., et al. (2010). Chronic Gastrointestinal Inflammation Induces Anxiety-Like Behavior and Alters Central Nervous System Biochemistry in Mice. *Gastroenterology, 139*(6), 2102-2112.e2101, doi:http://dx.doi.org/10.1053/j.gastro.2010.06.063.

Bernstein, C. N. (2017). The Natural History of Inflammatory Bowel Disease In D. C. Baumgart (Ed.), *Crohn's Disease and Ulcerative Colitis: From Epidemiology and Immunology to a Rational Diagnostic and Therapeutic Approach* (2nd ed.): Springer.

Bernstein, C. N., Fried, M., Krabshuis, J. H., Cohen, H., Eliakim, R., Fedail, S., et al. (2010). World Gastroenterology Organization Practice Guidelines for the Diagnosis and Management of IBD in 2010. *Inflammatory Bowel Diseases, 16*(1), 112-124, doi:10.1002/ibd.21048.

Bettenworth, D., Nowacki, T. M., Ross, M., Kyme, P., Schwammbach, D., Kerstiens, L., et al. (2014). Nicotinamide treatment ameliorates the course of experimental colitis mediated by enhanced neutrophil-specific antibacterial clearance. *Molecular Nutrition & Food Research, 58*(7), 1474-1490, doi:doi:10.1002/mnfr.201300818.

Bjelland, I., Dahl, A. A., Haug, T. T., & Neckelmann, D. (2002). The validity of the Hospital Anxiety and Depression Scale: An updated literature review. *Journal of Psychosomatic Research, 52*(2), 69-77, doi:https://doi.org/10.1016/S0022-3999(01)00296-3.

Bjerrum, J. T., Wang, Y., Hao, F., Coskun, M., Ludwig, C., Günther, U., et al. (2015). Metabonomics of human fecal extracts characterize ulcerative colitis, Crohn's disease and healthy individuals. [journal article]. *Metabolomics, 11*(1), 122-133, doi:10.1007/s11306-014-0677-3.

Boland, K., Turpin, W., Mohammadi, A., Borowski, K., Stempak, J. M., Kabakchiev, B., et al. (2017). Microbiome Composition is Altered in Patients with IBD Independent of Endoscopic Activity. *Gastroenterology, 152*(5), S991, doi:10.1016/S0016-5085(17)33357-7.

Bonnevie, O., Riis, P., & Ànthonisen, P. (1968). An Epidemiological Study of Ulcerative Colitis in Copenhagen County. *Scandinavian Journal of Gastroenterology, 3*(4), 432-438, doi:10.3109/00365526809180141.

Borsdorf, H., & Eiceman, G. A. (2006). Ion Mobility Spectrometry: Principles and Applications. *Applied Spectroscopy Reviews, 41*(4), 323-375, doi:10.1080/05704920600663469.

Botschuijver, S., Roeselers, G., Levin, E., Jonkers, D. M., Welting, O., Heinsbroek, S. E. M., et al. (2017). Intestinal Fungal Dysbiosis Is Associated With Visceral Hypersensitivity in Patients With Irritable Bowel Syndrome and Rats. *Gastroenterology, 153*(4), 1026-1039, doi:https://doi.org/10.1053/j.gastro.2017.06.004.

Bouatra, S., Aziat, F., Mandal, R., Guo, A. C., Wilson, M. R., Knox, C., et al. (2013). The human urine metabolome. *PLoS One, 8*(9), e73076, doi:10.1371/journal.pone.0073076.

Bowen, B. P., & Northen, T. R. (2010). Dealing with the Unknown: Metabolomics and Metabolite Atlases. *Journal of the American Society for Mass Spectrometry, 21*(9), 1471-1476, doi:http://dx.doi.org/10.1016/j.jasms.2010.04.003.

Brandt, L. J., Aroniadis, O. C., Mellow, M., Kanatzar, A., Kelly, C., Park, T., et al. (2012). Long-Term Follow-Up of Colonoscopic Fecal Microbiota Transplant for Recurrent Clostridium difficile Infection. [Colon/Small Bowel]. *The American Journal Of Gastroenterology, 107*, 1079, doi:10.1038/ajg.2012.60.

Brandt, L. J., Bjorkman, D., Fennerty, M. B., Locke, G. R., Olden, K., Peterson, W., et al. (2002). Systematic review on the management of irritable bowel syndrome in North America. *The American Journal Of Gastroenterology, 97*(11 Suppl), S7-26.

Brereton, R. G., & Lloyd, G. R. (2014). Partial least squares discriminant analysis: taking the magic away. *Journal of Chemometrics, 28*(4), 213-225, doi:doi:10.1002/cem.2609.

Broadhurst, D., Goodacre, R., Reinke, S. N., Kuligowski, J., Wilson, I. D., Lewis, M. R., et al. (2018). Guidelines and considerations for the use of system suitability and quality control samples in mass spectrometry assays applied in untargeted clinical metabolomic studies. *Metabolomics, 14*(6), 72, doi:10.1007/s11306-018-1367-3.

Buhner, S., Hahne, H., Hartwig, K., Li, Q., Vignali, S., Ostertag, D., et al. (2018). Protease signaling through protease activated receptor 1 mediate nerve activation by mucosal supernatants from irritable bowel syndrome but not from ulcerative colitis patients. *PLOS ONE, 13*(3), e0193943, doi:10.1371/journal.pone.0193943.

Buhner, S., Li, Q., Vignali, S., Barbara, G., De Giorgio, R., Stanghellini, V., et al. (2009). Activation of Human Enteric Neurons by Supernatants of Colonic Biopsy Specimens From Patients With Irritable Bowel Syndrome. *Gastroenterology, 137*(4), 1425-1434, doi:https://doi.org/10.1053/j.gastro.2009.07.005.

Büscher, J. M., Czernik, D., Ewald, J. C., Sauer, U., & Zamboni, N. (2009). Cross-Platform Comparison of Methods for Quantitative Metabolomics of Primary Metabolism. *Analytical Chemistry, 81*(6), 2135-2143, doi:10.1021/ac8022857.

Calvano, S. E., Xiao, W., Richards, D. R., Felciano, R. M., Baker, H. V., Cho, R. J., et al. (2005). A network-based analysis of systemic inflammation in humans. *Nature, 437*, 1032, doi:10.1038/nature03985

https://www.nature.com/articles/nature03985#supplementary-information.

Camilleri, M., & Choi, M. G. (1997). Review article: irritable bowel syndrome. *Alimentary Pharmacology & Therapeutics, 11*(1), 3-15, doi:10.1046/j.1365-2036.1997.84256000.x.

Camilleri, M., Halawi, H., & Oduyebo, I. (2017). Biomarkers as a diagnostic tool for irritable bowel syndrome: where are we? *Expert Review of Gastroenterology & Hepatology, 11*(4), 303-316, doi:10.1080/17474124.2017.1288096.

Canavan, C., West, J., & Card, T. (2014). The epidemiology of irritable bowel syndrome. *Clinical Epidemiology, 6*, 71-80, doi:10.2147/CLEP.S40245.

Caron, P., Audet-Walsh, E., Lépine, J., Bélanger, A., & Guillemette, C. (2009). Profiling Endogenous Serum Estrogen and Estrogen-Glucuronides by Liquid

Chromatography−Tandem Mass Spectrometry. *Analytical Chemistry, 81*(24), 10143-10148, doi:10.1021/ac9019126.

Cash, B. D., Schoenfeld, P., & Chey, W. D. (2002). The utility of diagnostic tests in irritable bowel syndrome patients: a systematic review. *The American Journal Of Gastroenterology, 97*(11), 2812-2819.

Chang, L., Chey, W. D., Harris, L., Olden, K., Surawicz, C., & Schoenfeld, P. (2006). Incidence of ischemic colitis and serious complications of constipation among patients using alosetron: systematic review of clinical trials and post-marketing surveillance data. *The American Journal Of Gastroenterology, 101*(5), 1069-1079, doi:10.1111/j.1572-0241.2006.00459.x.

Chaudhary, N. A., & Truelove, S. C. (1962). THE IRRITABLE COLON SYNDROMEA Study of the Clinical Features, Predisposing Causes, and Prognosis in 130 Cases. *QJM: An International Journal of Medicine, 31*(3), 307-322, doi:10.1093/oxfordjournals.qjmed.a066971.

Chiaro, T. R., Soto, R., Zac Stephens, W., Kubinak, J. L., Petersen, C., Gogokhia, L., et al. (2017). A member of the gut mycobiota modulates host purine metabolism exacerbating colitis in mice. [10.1126/scitranslmed.aaf9044]. *Science Translational Medicine, 9*(380).

Chong, J., & Xia, J. (2017). Computational Approaches for Integrative Analysis of the Metabolome and Microbiome. *Metabolites, 7*(4), 62.

Christmas, D. M., Badawy, A. A. B., Hince, D., Davies, S. J. C., Probert, C., Creed, T., et al. (2010). Increased serum free tryptophan in patients with diarrhea-predominant irritable bowel syndrome. *Nutrition Research, 30*(10), 678-688, doi:http://dx.doi.org/10.1016/j.nutres.2010.09.009.

Clouse, R. E., & Lustman, P. J. (2005). USE OF PSYCHOPHARMACOLOGICAL AGENTS FOR FUNCTIONAL GASTROINTESTINAL DISORDERS. *Gut, 54*(9), 1332-1341, doi:10.1136/gut.2004.048884.

Coates, M. D., Mahoney, C. R., Linden, D. R., Sampson, J. E., Chen, J., Blaszyk, H., et al. (2004). Molecular defects in mucosal serotonin content and decreased serotonin reuptake transporter in ulcerative colitis and irritable bowel syndrome. *Gastroenterology, 126*(7), 1657-1664, doi:http://dx.doi.org/10.1053/j.gastro.2004.03.013.

Cole, R. B. (2000). Some tenets pertaining to electrospray ionization mass spectrometry. *Journal of Mass Spectrometry, 35*(7), 763-772, doi:doi:10.1002/1096-9888(200007)35:7<763::AID-JMS16>3.0.CO;2-#.

Cosnes, J., Bourrier, A., Nion-Larmurier, I., Sokol, H., Beaugerie, L., & Seksik, P. (2012). Factors affecting outcomes in Crohn's disease over 15 years. *Gut, 61*(8), 1140-1145, doi:10.1136/gutjnl-2011-301971.

Cremonini, F., Delgado-Aros, S., & Camilleri, M. (2003). Efficacy of alosetron in irritable bowel syndrome: a meta-analysis of randomized controlled trials. *Neurogastroenterology & Motility, 15*(1), 79-86, doi:10.1046/j.1365-2982.2003.00389.x.

Crohn, B. B., Ginzburg, L., & Oppenheimer, G. D. (1932). Regional ileitis: A pathologic and clinical entity. *Journal of the American Medical Association, 99*(16), 1323-1329, doi:10.1001/jama.1932.02740680019005.

Czogalla, B., Schmitteckert, S., Houghton, L. A., Sayuk, G. S., Camilleri, M., Olivo-Diaz, A., et al. (2015). A meta-analysis of immunogenetic Case–Control Association Studies in irritable bowel syndrome. *Neurogastroenterology & Motility, 27*(5), 717-727, doi:doi:10.1111/nmo.12548.

De Preter, V., & Verbeke, K. (2013). Metabolomics as a diagnostic tool in gastroenterology. *World Journal of Gastrointestinal Pharmacology and Therapeutics, 4*(4), 97-107, doi:10.4292/wjgpt.v4.i4.97.

de Souza, H. S. P., Fiocchi, C., & Iliopoulos, D. (2017). The IBD interactome: an integrated view of aetiology, pathogenesis and therapy. [Perspective]. *Nature Reviews Gastroenterology & Hepatology, 14*, 739, doi:10.1038/nrgastro.2017.110

https://www.nature.com/articles/nrgastro.2017.110#supplementary-information.

Deechakawan, W., Cain, K. C., Jarrett, M. E., Burr, R. L., & Heitkemper, M. M. (2011). Effect of Self-Management Intervention on Cortisol and Daily Stress Levels in Irritable Bowel Syndrome. *Biological Research For Nursing, 15*(1), 26-36, doi:10.1177/1099800411414047.

Di Guida, R., Engel, J., Allwood, J. W., Weber, R. J. M., Jones, M. R., Sommer, U., et al. (2016). Non-targeted UHPLC-MS metabolomic data processing methods: a comparative investigation of normalisation, missing value imputation, transformation and scaling. *Metabolomics, 12*, 93, doi:10.1007/s11306-016-1030-9.

Ding, Y., Lu, B., Chen, D., Meng, L., Shen, Y., & Chen, S. (2010). Proteomic analysis of colonic mucosa in a rat model of irritable bowel syndrome. *PROTEOMICS, 10*(14), 2620-2630, doi:doi:10.1002/pmic.200900572.

Dixon, R. A., Gang, D. R., Charlton, A. J., Fiehn, O., Kuiper, H. A., Reynolds, T. L., et al. (2006). Applications of Metabolomics in Agriculture. *Journal of Agricultural and Food Chemistry, 54*(24), 8984-8994, doi:10.1021/jf061218t.

Drossman, D. A., Camilleri, M., Mayer, E. A., & Whitehead, W. E. (2002). AGA technical review on irritable bowel syndrome. *Gastroenterology, 123*(6), 2108-2131, doi:http://dx.doi.org/10.1053/gast.2002.37095.

Drucker, E., & Krapfenbauer, K. (2013). Pitfalls and limitations in translation from biomarker discovery to clinical utility in predictive and personalised medicine. *The EPMA Journal, 4*(1), 7-7, doi:10.1186/1878-5085-4-7.

Dührkop, K., Shen, H., Meusel, M., Rousu, J., & Böcker, S. (2015). Searching molecular structure databases with tandem mass spectra using CSI:FingerID. *Proceedings of the National Academy of Sciences of the United States of America, 112*(41), 12580-12585, doi:10.1073/pnas.1509788112.

Enck, P., Aziz, Q., Barbara, G., Farmer, A. D., Fukudo, S., Mayer, E. A., et al. (2016). Irritable bowel syndrome. [Primer]. *Nature Reviews Disease Primers, 2*, 16014, doi:10.1038/nrdp.2016.14.

Escher, J. C. (2013). Corticosteroids. In P. Mamula, J. E. Markowitz, & R. Baldassano (Eds.), *Pediatric Inflammatory Bowel Disease* (2nd Edition ed., pp. 325-330): Springer.

Fagerhol, M. K., Dale, I., & Anderson, T. (1980). Release and Quantitation of a Leucocyte Derived Protein (L1). *Scandinavian Journal of Haematology, 24*(5), 393-398, doi:doi:10.1111/j.1600-0609.1980.tb02754.x.

Fernie, A. R., Trethewey, R. N., Krotzky, A. J., & Willmitzer, L. (2004). Metabolite profiling: from diagnostics to systems biology. [10.1038/nrm1451]. *Nat Rev Mol Cell Biol, 5*(9), 763-769.

Fiehn, O. (2016). Metabolomics by Gas Chromatography-Mass Spectrometry: the combination of targeted and untargeted profiling. *Current protocols in molecular biology / edited by Frederick M. Ausubel ... [et al.], 114*, 30.34.31-30.34.32, doi:10.1002/0471142727.mb3004s114.

Finegold, S. M., Downes, J., & Summanen, P. H. (2012). Microbiology of regressive autism. *Anaerobe, 18*(2), 260-262, doi:http://dx.doi.org/10.1016/j.anaerobe.2011.12.018.

Ford, A. C., & Moayyedi, P. (2010). Meta-analysis: factors affecting placebo response rate in the irritable bowel syndrome. *Alimentary Pharmacology & Therapeutics, 32*(2), 144-158, doi:doi:10.1111/j.1365-2036.2010.04328.x.

Forstmeier, W., Wagenmakers, E.-J., & Parker, T. H. (2017). Detecting and avoiding likely false-positive findings – a practical guide. *Biological Reviews, 92*(4), 1941-1968, doi:10.1111/brv.12315.

Foxman, E. F., & Iwasaki, A. (2011). Genome–virome interactions: examining the role of common viral infections in complex disease. [Review Article]. *Nature Reviews Microbiology, 9*, 254, doi:10.1038/nrmicro2541.

Gazouli, M., Wouters, M. M., Kapur-Pojskić, L., Bengtson, M.-B., Friedman, E., Nikčević, G., et al. (2016). Lessons learned — resolving the enigma of genetic factors in IBS. [Review Article]. *Nature Reviews Gastroenterology &Amp; Hepatology, 13*, 77, doi:10.1038/nrgastro.2015.206

https://www.nature.com/articles/nrgastro.2015.206#supplementary-information.

Gearry, R. B., Irving, P. M., Barrett, J. S., Nathan, D. M., Shepherd, S. J., & Gibson, P. R. (2009). Reduction of dietary poorly absorbed short-chain carbohydrates (FODMAPs) improves abdominal symptoms in patients with inflammatory bowel disease—a pilot study. *Journal of Crohn's and Colitis, 3*(1), 8-14, doi:10.1016/j.crohns.2008.09.004.

Gerasimidis, K., Bertz, M., Hanske, L., Junick, J., Biskou, O., Aguilera, M., et al. (2014). Decline in Presumptively Protective Gut Bacterial Species and Metabolites Are Paradoxically Associated with Disease Improvement in Pediatric Crohn's Disease During Enteral Nutrition *Inflammatory Bowel Diseases, 20*(5), 861-871.

Gibney, M. J., Walsh, M., Brennan, L., Roche, H. M., German, B., & van Ommen, B. (2005). Metabolomics in human nutrition: opportunities and challenges. *Am J Clin Nutr, 82*(3), 497-503.

Gika, H. G., Theodoridis, G. A., & Wilson, I. D. (2008). Liquid chromatography and ultra-performance liquid chromatography–mass spectrometry fingerprinting of human urine: Sample stability under different handling and storage conditions for metabonomics studies. *Journal of Chromatography A, 1189*(1–2), 314-322, doi:http://dx.doi.org/10.1016/j.chroma.2007.10.066.

Glauser, W. (2011). Risk and rewards of fecal transplants. *CMAJ : Canadian Medical Association Journal, 183*(5), 541-542, doi:10.1503/cmaj.109-3806.

Gratton, J., Phetcharaburanin, J., Mullish, B. H., Williams, H. R. T., Thursz, M., Nicholson, J. K., et al. (2016). Optimized Sample Handling Strategy for Metabolic Profiling of Human Feces. *Analytical Chemistry, 88*(9), 4661-4668, doi:10.1021/acs.analchem.5b04159.

Greenblum, S., Turnbaugh, P. J., & Borenstein, E. (2012). Metagenomic systems biology of the human gut microbiome reveals topological shifts associated with obesity and inflammatory bowel disease. [10.1073/pnas.1116053109]. *Proceedings of the National Academy of Sciences, 109*(2), 594.

Grootveld, M. (2015). CHAPTER 1 Introduction to the Applications of Chemometric Techniques in 'Omics' Research: Common Pitfalls, Misconceptions and 'Rights and Wrongs'. In *Metabolic Profiling: Disease and Xenobiotics* (pp. 1-34): The Royal Society of Chemistry.

Grover, Z., & Lewindon, P. (2015). Two-Year Outcomes After Exclusive Enteral Nutrition Induction Are Superior to Corticosteroids in Pediatric Crohn's Disease Treated Early with

Thiopurines. [journal article]. *Digestive Diseases and Sciences, 60*(10), 3069-3074, doi:10.1007/s10620-015-3722-9.

Grover, Z., Muir, R., & Lewindon, P. (2014). Exclusive enteral nutrition induces early clinical, mucosal and transmural remission in paediatric Crohn's disease. [journal article]. *Journal of Gastroenterology, 49*(4), 638-645, doi:10.1007/s00535-013-0815-0.

Gulersonmez, M. C., Lock, S., Hankemeier, T., & Ramautar, R. (2016). Sheathless capillary electrophoresis-mass spectrometry for anionic metabolic profiling. *ELECTROPHORESIS, 37*(7-8), 1007-1014, doi:doi:10.1002/elps.201500435.

Gwee, K.-A., Collins, S. M., Read, N. W., Rajnakova, A., Deng, Y., Graham, J. C., et al. (2003). Increased rectal mucosal expression of interleukin 1β in recently acquired post-infectious irritable bowel syndrome. *Gut, 52*(4), 523-526, doi:10.1136/gut.52.4.523.

Gwee, K., Leong, Y., Graham, C., McKendrick, M., Collins, S., Walters, S., et al. (1999). The role of psychological and biological factors in postinfective gut dysfunction. *Gut, 44*(3), 400-406.

Halder, S. L. S., Locke, G. R., III, Schleck, C. D., Zinsmeister, A. R., Melton, L. J., III, & Talley, N. J. (2007). Natural History of Functional Gastrointestinal Disorders: A 12-year Longitudinal Population-Based Study. *Gastroenterology, 133*(3), 799-807.e791, doi:10.1053/j.gastro.2007.06.010.

Halmos, E. P., Christophersen, C. T., Bird, A. R., Shepherd, S. J., Gibson, P. R., & Muir, J. G. (2015). Diets that differ in their FODMAP content alter the colonic luminal microenvironment. [10.1136/gutjnl-2014-307264]. *Gut, 64*(1), 93.

Halmos, E. P., Christophersen, C. T., Bird, A. R., Shepherd, S. J., Muir, J. G., & Gibson, P. R. (2016). Consistent Prebiotic Effect on Gut Microbiota With Altered FODMAP Intake in Patients with Crohn's Disease: A Randomised, Controlled Cross-Over Trial of Well-Defined Diets. [Original Contributions]. *Clinical And Translational Gastroenterology, 7*, e164, doi:10.1038/ctg.2016.22

https://www.nature.com/articles/ctg201622#supplementary-information.

Halmos, E. P., Power, V. A., Shepherd, S. J., Gibson, P. R., & Muir, J. G. (2014). A Diet Low in FODMAPs Reduces Symptoms of Irritable Bowel Syndrome. *Gastroenterology, 146*(1), 67-75.e65, doi:10.1053/j.gastro.2013.09.046.

Hamilton, A. L., Kamm, M. A., Teo, S. M., De Cruz, P., Wright, E. K., Feng, H., et al. (2017). P773 Post-operative Crohn's disease recurrence is associated with specific changes in the faecal microbiome – potential pathogenic and protective roles. *Journal of Crohn's and Colitis, 11*(suppl_1), S476-S476, doi:10.1093/ecco-jcc/jjx002.894.

Han, M., Rock, B. M., Pearson, J. T., & Rock, D. A. (2016). Intact mass analysis of monoclonal antibodies by capillary electrophoresis—Mass spectrometry. *Journal of Chromatography B, 1011*, 24-32, doi:https://doi.org/10.1016/j.jchromb.2015.12.045.

Harada, S., Hirayama, A., Chan, Q., Kurihara, A., Fukai, K., Iida, M., et al. (2018). Reliability of plasma polar metabolite concentrations in a large-scale cohort study using capillary electrophoresis-mass spectrometry. *PLoS ONE, 13*(1), e0191230, doi:10.1371/journal.pone.0191230.

Hart, A. L., & Ng, S. C. (2010). Review article: the optimal medical management of acute severe ulcerative colitis. *Alimentary Pharmacology & Therapeutics, 32*(5), 615-627, doi:doi:10.1111/j.1365-2036.2010.04392.x.

Hashimoto, T., Perlot, T., Rehman, A., Trichereau, J., Ishiguro, H., Paolino, M., et al. (2012). ACE2 links amino acid malnutrition to microbial ecology and intestinal inflammation. *Nature, 487*, 477, doi:10.1038/nature11228

https://www.nature.com/articles/nature11228#supplementary-information.

Haug, K., Salek, R. M., Conesa, P., Hastings, J., de Matos, P., Rijnbeek, M., et al. (2013). MetaboLights—an open-access general-purpose repository for metabolomics studies and associated meta-data. *Nucleic Acids Research, 41*(D1), D781-D786, doi:10.1093/nar/gks1004.

Hegeman, A. D. (2010). Plant metabolomics—meeting the analytical challenges of comprehensive metabolite analysis. *Briefings in Functional Genomics, 9*(2), 139-148, doi:10.1093/bfgp/elp053.

Hilska, M., Collan, Y., Peltonen, J., Gullichsen, R., Paajanen, H., & Laato, M. (1998). The distribution of collagen types I, III, and IV in normal and malignant colorectal mucosa. *European Journal of Surgery, 164*(6), 457-464, doi:10.1080/110241598750004274.

Howell, S., Talley, N. J., Quine, S., & Poulton, R. (2004). The Irritable Bowel Syndrome has Origins in the Childhood Socioeconomic Environment. [Original Contribution]. *American Journal Of Gastroenterology, 99*, 1572, doi:10.1111/j.1572-0241.2004.40188.x.

Hsieh, F., Baronas, E., Muir, C., & Martin, S. A. (1999). A novel nanospray capillary zone electrophoresis/mass spectrometry interface. *Rapid Communications in Mass Spectrometry, 13*(1), 67-72, doi:doi:10.1002/(SICI)1097-0231(19990115)13:1<67::AID-RCM453>3.0.CO;2-F.

Huang, H., Vangay, P., McKinlay, C. E., & Knights, D. (2014). Multi-omics analysis of inflammatory bowel disease. *Immunology Letters, 162*(2, Part A), 62-68, doi:https://doi.org/10.1016/j.imlet.2014.07.014.

Hungin, A. P. S., Chang, L., Locke, G. R., Dennis, E. H., & Barghout, V. (2005). Irritable bowel syndrome in the United States: prevalence, symptom patterns and impact. *Alimentary Pharmacology & Therapeutics, 21*(11), 1365-1375, doi:doi:10.1111/j.1365-2036.2005.02463.x.

Ibáñez, C., Simó, C., García-Cañas, V., Gómez-Martínez, Á., Ferragut, J. A., & Cifuentes, A. (2012). CE/LC-MS multiplatform for broad metabolomic analysis of dietary polyphenols effect on colon cancer cells proliferation. *ELECTROPHORESIS, 33*(15), 2328-2336, doi:doi:10.1002/elps.201200143.

Ihara, Y., Inai, Y., Ikezaki, M., Matsui, I.-S. L., Manabe, S., & Ito, Y. (2015). C-Mannosylation: Modification on Tryptophan in Cellular Proteins. In N. Taniguchi, T. Endo, W. G. Hart, H. P. Seeberger, & C.-H. Wong (Eds.), *Glycoscience: Biology and Medicine* (pp. 1091-1099). Tokyo: Springer Japan.

Ioannidis, J. P. A. (2005). Why Most Published Research Findings Are False. *PLoS Medicine, 2*(8), e124, doi:10.1371/journal.pmed.0020124.

Iskandar, H. N., & Ciorba, M. A. (2012). Biomarkers in inflammatory bowel disease: current practices and recent advances. *Translational Research, 159*(4), 313-325, doi:http://dx.doi.org/10.1016/j.trsl.2012.01.001.

Iyer, L. M., Aravind, L., Coon, S. L., Klein, D. C., & Koonin, E. V. (2004). Evolution of cell–cell signaling in animals: did late horizontal gene transfer from bacteria have a role? *Trends in Genetics, 20*(7), 292-299, doi:http://dx.doi.org/10.1016/j.tig.2004.05.007.

Jackson, J. L., O'Malley, P. G., Tomkins, G., Balden, E., Santoro, J., & Kroenke, K. (2000). Treatment of functional gastrointestinal disorders with antidepressant medications: a meta-analysis1. *The American Journal of Medicine, 108*(1), 65-72, doi:10.1016/s0002-9343(99)00299-5.

Jacobs, J. P., Goudarzi, M., Singh, N., Tong, M., McHardy, I. H., Ruegger, P., et al. (2016). A Disease-Associated Microbial and Metabolomics State in Relatives of Pediatric

Inflammatory Bowel Disease Patients. *Cellular and Molecular Gastroenterology and Hepatology, 2*(6), 750-766, doi:https://doi.org/10.1016/j.jcmgh.2016.06.004.

Jansson, J., Willing, B., Lucio, M., Fekete, A., Dicksved, J., Halfvarson, J., et al. (2009). Metabolomics Reveals Metabolic Biomarkers of Crohn's Disease. *PLoS One, 4*(7), e6386, doi:10.1371/journal.pone.0006386.

Johnson, C. H., Ivanisevic, J., & Siuzdak, G. (2016). Metabolomics: beyond biomarkers and towards mechanisms. *Nature Review Molecular Cell Biology, 17*(7), 451-459, doi:10.1038/nrm.2016.25.

Jostins, L., Ripke, S., Weersma, R. K., Duerr, R. H., McGovern, D. P., Hui, K. Y., et al. (2012). Host–microbe interactions have shaped the genetic architecture of inflammatory bowel disease. *Nature, 491*, 119, doi:10.1038/nature11582

https://www.nature.com/articles/nature11582#supplementary-information.

Kanani, H., Chrysanthopoulos, P. K., & Klapa, M. I. (2008). Standardizing GC–MS metabolomics. *Journal of Chromatography B, 871*(2), 191-201, doi:http://dx.doi.org/10.1016/j.jchromb.2008.04.049.

Kansal, S., Wagner, J., Kirkwood, C. D., & Catto-Smith, A. G. (2013). Enternal Nutritin in Crohn's Disease: An Underused Therapy. *Gastroenterology Research and Practice, 2013*.

Kaper, J. B., Nataro, J. P., & Mobley, H. L. T. (2004). Pathogenic Escherichia coli. [Review Article]. *Nature Reviews Microbiology, 2*, 123, doi:10.1038/nrmicro818.

Kassinen, A., Krogius-Kurikka, L., Mäkivuokko, H., Rinttilä, T., Paulin, L., Corander, J., et al. (2007). The Fecal Microbiota of Irritable Bowel Syndrome Patients Differs Significantly From That of Healthy Subjects. *Gastroenterology, 133*(1), 24-33, doi:http://dx.doi.org/10.1053/j.gastro.2007.04.005.

Kernbauer, E., Ding, Y., & Cadwell, K. (2014). An enteric virus can replace the beneficial function of commensal bacteria. *Nature, 516*, 94, doi:10.1038/nature13960

https://www.nature.com/articles/nature13960#supplementary-information.

Keshteli, A. H., van den Brand, F. F., Madsen, K. L., Mandal, R., Valcheva, R., Kroeker, K. I., et al. (2017). Dietary and metabolomic determinants of relapse in ulcerative colitis patients: A pilot prospective cohort study. *World Journal of Gastroenterology, 23*(21), 3890-3899, doi:10.3748/wjg.v23.i21.3890.

Keszthelyi, D., Troost, F. J., Jonkers, D. M., Kruimel, J. W., Leue, C., & Masclee, A. A. M. (2013). Decreased levels of kynurenic acid in the intestinal mucosa of IBS patients: Relation to serotonin and psychological state. *Journal of Psychosomatic Research, 74*(6), 501-504, doi:http://dx.doi.org/10.1016/j.jpsychores.2013.01.008.

Keun, H. C., Ebbels, T. M. D., Antti, H., Bollard, M. E., Beckonert, O., Schlotterbeck, G., et al. (2002). Analytical Reproducibility in 1H NMR-Based Metabonomic Urinalysis. *Chemical Research in Toxicology, 15*(11), 1380-1386, doi:10.1021/tx0255774.

King, T. S., Elia, M., & Hunter, J. O. (1998). Abnormal colonic fermentation in irritable bowel syndrome. *The Lancet, 352*(9135), 1187-1189, doi:http://dx.doi.org/10.1016/S0140-6736(98)02146-1.

Kirsner, J. B. (1988). Historical aspects of inflammatory bowel disease. *Journal of clinical gastroenterology, 10*(3), 286-297.

Kleessen, B., Kroesen, A. J., Buhr, H. J., & Blaut, M. (2002). Mucosal and Invading Bacteria in Patients with Inflammatory Bowel Disease Compared with Controls. *Scandinavian Journal of Gastroenterology, 37*(9), 1034-1041, doi:10.1080/003655202320378220.

Klein, J., Papadopoulos, T., Mischak, H., & Mullen, W. (2014). Comparison of CE-MS/MS and LC-MS/MS sequencing demonstrates significant complementarity in natural peptide identification in human urine. *ELECTROPHORESIS, 35*(7), 1060-1064, doi:doi:10.1002/elps.201300327.

Kok, M. G. M., Somsen, G. W., & de Jong, G. J. (2015). Comparison of capillary electrophoresis–mass spectrometry and hydrophilic interaction chromatography–mass spectrometry for anionic metabolic profiling of urine. *Talanta, 132*(0), 1-7, doi:http://dx.doi.org/10.1016/j.talanta.2014.08.047.

Kuehnbaum, N. L., & Britz-McKibbin, P. (2013). New Advances in Separation Science for Metabolomics: Resolving Chemical Diversity in a Post-Genomic Era. *Chemical Reviews, 113*(4), 2437-2468, doi:10.1021/cr300484s.

Kuehnbaum, N. L., Gillen, J. B., Kormendi, A., Lam, K. P., DiBattista, A., Gibala, M. J., et al. (2015). Multiplexed separations for biomarker discovery in metabolomics: Elucidating adaptive responses to exercise training. *ELECTROPHORESIS, 36*(18), 2226-2236, doi:doi:10.1002/elps.201400604.

Kuehnbaum, N. L., Kormendi, A., & Britz-McKibbin, P. (2013). Multisegment Injection-Capillary Electrophoresis-Mass Spectrometry: A High-Throughput Platform for Metabolomics with High Data Fidelity. *Analytical Chemistry, 85*(22), 10664-10669, doi:10.1021/ac403171u.

Kuhl, C., Tautenhahn, R., Böttcher, C., Larson, T. R., & Neumann, S. (2012). CAMERA: An integrated strategy for compound spectra extraction and annotation of LC/MS data sets. *Analytical Chemistry, 84*(1), 283-289, doi:10.1021/ac202450g.

L. A. García Rodríguez, A. R., M.-A. Wallander, S. Johansson, L. Olbe (2000). Detection of Colorectal Tumor and Inflammatory Bowel Disease during Follow-up of Patients with Initial Diagnosis of Irritable Bowel Syndrome. *Scandinavian Journal of Gastroenterology, 35*(3), 306-311, doi:doi:10.1080/003655200750024191.

Lackner, J. M., Jaccard, J., Krasner, S. S., Katz, L. A., Gudleski, G. D., & Holroyd, K. (2008). Self administered cognitive behavior therapy for moderate to severe IBS: Clinical efficacy, tolerability, feasibility. *Clinical gastroenterology and hepatology : the official clinical practice journal of the American Gastroenterological Association, 6*(8), 899-906, doi:10.1016/j.cgh.2008.03.004.

Lacy, B. E., Mearin, F., Chang, L., Chey, W. D., Lembo, A. J., Simren, M., et al. (2016). Bowel Disorders. *Gastroenterology, 150*(6), 1393-1407.e1395, doi:10.1053/j.gastro.2016.02.031.

Lafferty, L., Tuohy, M., Carey, A., Sugrue, S., Hurley, M., & Hussey, S. (2016). Outcomes of exclusive enteral nutrition in paediatric Crohn's disease. [Original Article]. *European Journal Of Clinical Nutrition, 71*, 185, doi:10.1038/ejcn.2016.210.

Lagier, J.-C., Khelaifia, S., Alou, M. T., Ndongo, S., Dione, N., Hugon, P., et al. (2016). Culture of previously uncultured members of the human gut microbiota by culturomics. [Letter]. *Nature Microbiology, 1*, 16203, doi:10.1038/nmicrobiol.2016.203

https://www.nature.com/articles/nmicrobiol2016203#supplementary-information.

Langholz, E., Munkholm, P., Nielsen, O. H., Kreiner, S., & Binder, V. (1991). Incidence and prevalence of ulcerative colitis in Copenhagen county from 1962 to 1987. *Scandinavian Journal of Gastroenterology, 26*(12), 1247-1256.

Lau, J. T., Whelan, F. J., Herath, I., Lee, C. H., Collins, S. M., Bercik, P., et al. (2016). Capturing the diversity of the human gut microbiota through culture-enriched molecular profiling. [journal article]. *Genome Medicine, 8*(1), 72, doi:10.1186/s13073-016-0327-7.

Lauridsen, M., Hansen, S. H., Jaroszewski, J. W., & Cornett, C. (2007). Human Urine as Test Material in 1H NMR-Based Metabonomics: Recommendations for Sample Preparation and Storage. *Analytical Chemistry, 79*(3), 1181-1186, doi:10.1021/ac061354x.

Leach, S. T., Mitchell, H. M., Eng, W. R., Zhang, L., & Day, A. S. (2008). Sustained modulation of intestinal bacteria by exclusive enteral nutrition used to treat children with Crohn's disease. *Alimentary Pharmacology & Therapeutics, 28*(6), 724-733, doi:10.1111/j.1365-2036.2008.03796.x.

Lesbros-Pantoflickova, D., Michetti, P., Fried, M., Beglinger, C., & Blum, A. L. (2004). Meta-analysis: the treatment of irritable bowel syndrome. *Alimentary Pharmacology & Therapeutics, 20*(11-12), 1253-1269, doi:10.1111/j.1365-2036.2004.02267.x.

Lin, H. C. (2004). Small intestinal bacterial overgrowth: A framework for understanding irritable bowel syndrome. *Journal of the American Medical Association, 292*(7), 852-858, doi:10.1001/jama.292.7.852.

Liu, J. Z., van Sommeren, S., Huang, H., Ng, S. C., Alberts, R., Takahashi, A., et al. (2015a). Association analyses identify 38 susceptibility loci for inflammatory bowel disease and highlight shared genetic risk across populations. [Article]. *Nature Genetics, 47*, 979, doi:10.1038/ng.3359

https://www.nature.com/articles/ng.3359#supplementary-information.

Liu, J. Z., van Sommeren, S., Huang, H., Ng, S. C., Alberts, R., Takahashi, A., et al. (2015b). Association analyses identify 38 susceptibility loci for inflammatory bowel disease and highlight shared genetic risk across populations. *Nature genetics, 47*(9), 979-986, doi:10.1038/ng.3359.

Livera, A. M. D., Sysi-Aho, M., Jacob, L., Gagnon-Bartsch, J. A., Castillo, S., Simpson, J. A., et al. (2015). Statistical Methods for Handling Unwanted Variation in Metabolomics Data. *Analytical Chemistry, 87*(7), 3606-3615, doi:10.1021/ac502439y.

Lovell, R. M., & Ford, A. C. (2012). Global Prevalence of and Risk Factors for Irritable Bowel Syndrome: A Meta-analysis. *Clinical Gastroenterology and Hepatology, 10*(7), 712-721.e714, doi:https://doi.org/10.1016/j.cgh.2012.02.029.

Madsen, R., Lundstedt, T., & Trygg, J. (2010). Chemometrics in metabolomics—A review in human disease diagnosis. *Analytica Chimica Acta, 659*(1–2), 23-33, doi:http://dx.doi.org/10.1016/j.aca.2009.11.042.

Mahieu, N. G., Genenbacher, J. L., & Patti, G. J. (2016). A Roadmap for the XCMS Family of Software Solutions in Metabolomics. *Current Opinion in Chemical Biology, 30*, 87-93, doi:10.1016/j.cbpa.2015.11.009.

Mahieu, N. G., & Patti, G. J. (2017). Systems-Level Annotation of a Metabolomics Data Set Reduces 25 000 Features to Fewer than 1000 Unique Metabolites. *Analytical Chemistry, 89*(19), 10397-10406, doi:10.1021/acs.analchem.7b02380.

Major, G., Pritchard, S., Murray, K., Alappadan, J. P., Hoad, C. L., Marciani, L., et al. (2017). Colon Hypersensitivity to Distension, Rather Than Excessive Gas Production, Produces Carbohydrate-Related Symptoms in Individuals With Irritable Bowel Syndrome. *Gastroenterology, 152*(1), 124-133.e122, doi:10.1053/j.gastro.2016.09.062.

Malinen, E., Rinttila, T., Kajander, K., Matto, J., Kassinen, A., Krogius, L., et al. (2005). Analysis of the Fecal Microbiota of Irritable Bowel Syndrome Patients and Healthy Controls with Real-Time PCR. *The American Journal Of Gastroenterology, 100*(2), 373-382.

Mallon, P. T., McKay, D., Kirk, S. J., & Gardiner, K. (2007). Probiotics for induction of remission in ulcerative colitis. *Cochrane Database of Systematic Reviews*(4), doi:10.1002/14651858.CD005573.pub2.

Matějíček, D. (2011). On-line two-dimensional liquid chromatography–tandem mass spectrometric determination of estrogens in sediments. *Journal of Chromatography A, 1218*(16), 2292-2300, doi:http://dx.doi.org/10.1016/j.chroma.2011.02.041.

Mayer, E. A. (2008). Irritable Bowel Syndrome. *New England Journal of Medicine, 358*(16), 1692-1699, doi:doi:10.1056/NEJMcp0801447.

Mayer, E. A., & Bradesi, S. (2003). Alosetron and irritable bowel syndrome. *Expert Opinion on Pharmacotherapy, 4*(11), 2089-2098, doi:doi:10.1517/14656566.4.11.2089.

McIntosh, K., Reed, D. E., Schneider, T., Dang, F., Keshteli, A. H., De Palma, G., et al. (2017). FODMAPs alter symptoms and the metabolome of patients with IBS: a randomised controlled trial. [10.1136/gutjnl-2015-311339]. *Gut, 66*(7), 1241.

McKenzie, Y. A., Thompson, J., Gulia, P., & Lomer, M. C. E. (2016). British Dietetic Association systematic review of systematic reviews and evidence-based practice guidelines for the use of probiotics in the management of irritable bowel syndrome in adults (2016 update). *Journal of Human Nutrition and Dietetics, 29*(5), 576-592, doi:doi:10.1111/jhn.12386.

Meister, D., Bode, J., Shand, A., & Ghosh, S. (2002). Anti-inflammatory effects of enteral diet components on Crohn's disease-affected tissues in vitro. *Digestive and Liver Disease, 34*(6), 430-438, doi:10.1016/S1590-8658(02)80041-X.

Menees, S. B., Kurlander, J., Goel, A., Powell, C. C., & Chey, W. D. (2014). Sa1079 A Meta-Analysis of the Utility of Common Serum and Fecal Biomarkers in Adults With IBS. *Gastroenterology, 146*(5), S-194, doi:10.1016/S0016-5085(14)60683-1.

Mete, R., Tulubas, F., Oran, M., Yilmaz, A., Avci, B. A., Yildiz, K., et al. (2013). The role of oxidants and reactive nitrogen species in irritable bowel syndrome: A potential etiological explanation. *Medical Science Monitor : International Medical Journal of Experimental and Clinical Research, 19*, 762-766, doi:10.12659/MSM.889068.

Methé, B. A., Nelson, K. E., Pop, M., Creasy, H. H., Giglio, M. G., Huttenhower, C., et al. (2012). A framework for human microbiome research. *Nature, 486*(7402), 215-221, doi:10.1038/nature11209.

Mikocka-Walus, A., Knowles, S. R., Keefer, L., & Graff, L. (2016). Controversies Revisited: A Systematic Review of the Comorbidity of Depression and Anxiety with Inflammatory Bowel Diseases. *Inflammatory Bowel Diseases, 22*(3), 752-762, doi:10.1097/MIB.0000000000000620.

Minami, Y., Kasukawa, T., Kakazu, Y., Iigo, M., Sugimoto, M., Ikeda, S., et al. (2009). Measurement of internal body time by blood metabolomics. [10.1073/pnas.0900617106]. *Proceedings of the National Academy of Sciences, 106*(24), 9890.

Moayyedi, P., Ford, A. C., Talley, N. J., Cremonini, F., Foxx-Orenstein, A. E., Brandt, L. J., et al. (2010). The efficacy of probiotics in the treatment of irritable bowel syndrome: a systematic review. *Gut, 59*(3), 325-332, doi:10.1136/gut.2008.167270.

Moayyedi, P., Surette, M. G., Kim, P. T., Libertucci, J., Wolfe, M., Onischi, C., et al. (2015). Fecal Microbiota Transplantation Induces Remission in Patients With Active Ulcerative Colitis in a Randomized Controlled Trial. *Gastroenterology, 149*(1), 102-109.e106, doi:https://doi.org/10.1053/j.gastro.2015.04.001.

Moini, M. (2007). Simplifying CE−MS Operation. 2. Interfacing Low-Flow Separation Techniques to Mass Spectrometry Using a Porous Tip. *Analytical Chemistry, 79*(11), 4241-4246, doi:10.1021/ac0704560.

Mujagic, Z., Keszthelyi, D., Aziz, Q., Reinisch, W., Quetglas, E. G., De Leonardis, F., et al. (2015). Systematic review: instruments to assess abdominal pain in irritable bowel

syndrome. *Alimentary Pharmacology & Therapeutics, 42*(9), 1064-1081, doi:doi:10.1111/apt.13378.

Mujagic, Z., Tigchelaar, E. F., Zhernakova, A., Ludwig, T., Ramiro-Garcia, J., Baranska, A., et al. (2016). A novel biomarker panel for irritable bowel syndrome and the application in the general population. [Article]. *Scientific Reports, 6*, 26420, doi:10.1038/srep26420.

Murray, K., Wilkinson-Smith, V., Hoad, C., Costigan, C., Cox, E., Lam, C., et al. (2013). Differential Effects of FODMAPs (Fermentable Oligo-, Di-, Mono-Saccharides and Polyols) on Small and Large Intestinal Contents in Healthy Subjects Shown by MRI. [Colon/Small Bowel]. *The American Journal Of Gastroenterology, 109*, 110, doi:10.1038/ajg.2013.386.

Neumann, S., & Böcker, S. (2010). Computational mass spectrometry for metabolomics: Identification of metabolites and small molecules. [journal article]. *Analytical and Bioanalytical Chemistry, 398*(7), 2779-2788, doi:10.1007/s00216-010-4142-5.

Nikolaus, S., Schulte, B., Al-Massad, N., Thieme, F., Schulte, D. M., Bethge, J., et al. (2017). Increased Tryptophan Metabolism Is Associated With Activity of Inflammatory Bowel Diseases. *Gastroenterology, 153*(6), 1504-1516.e1502, doi:10.1053/j.gastro.2017.08.028.

Norman, Jason M., Handley, Scott A., Baldridge, Megan T., Droit, L., Liu, Catherine Y., Keller, Brian C., et al. (2015). Disease-Specific Alterations in the Enteric Virome in Inflammatory Bowel Disease. *Cell, 160*(3), 447-460, doi:https://doi.org/10.1016/j.cell.2015.01.002.

Nusbaum, D. J., Sun, F., Ren, J., Zhu, Z., Ramsy, N., Pervolarakis, N., et al. (2018). Gut microbial and metabolomic profiles after fecal microbiota transplantation in pediatric ulcerative colitis patients. *FEMS Microbiology Ecology, 94*(9), fiy133-fiy133, doi:10.1093/femsec/fiy133.

O'Mahony, L., McCarthy, J., Kelly, P., Hurley, G., Luo, F., Chen, K., et al. (2005). Lactobacillus and bifidobacterium in irritable bowel syndrome: Symptom responses and relationship to cytokine profiles. *Gastroenterology, 128*(3), 541-551, doi:http://dx.doi.org/10.1053/j.gastro.2004.11.050.

Otley, A., Day, A. S., & Zachos, M. (2013). Nutritional Management of Inflammatory Bowel Disease. In P. Mamula, J. E. Markowitz, & R. N. Baldassano (Eds.), *Pediatric Inflammatory Bowel Disease* (2nd ed., pp. 295-312): Springer.

Papay, P., Ignjatovic, A., Karmiris, K., Amarante, H., Miheller, P., Feagan, B., et al. (2013). Optimising monitoring in the management of Crohn's disease: A physician's perspective. *Journal of Crohn's and Colitis, 7*(8), 653-669, doi:10.1016/j.crohns.2013.02.005.

Patel, N., Alkhouri, N., Eng, K., Cikach, F., Mahajan, L., Yan, C., et al. (2014). Metabolomic analysis of breath volatile organic compounds reveals unique breathprints in children with inflammatory bowel disease: a pilot study. *Alimentary Pharmacology & Therapeutics, 40*(5), 498-507, doi:doi:10.1111/apt.12861.

Paterson, M. J., Oh, S., & Underhill, D. M. (2017). Host–microbe interactions: commensal fungi in the gut. *Current Opinion in Microbiology, 40*, 131-137, doi:https://doi.org/10.1016/j.mib.2017.11.012.

Patti, G. J., Yanes, O., & Siuzdak, G. (2012). Innovation: Metabolomics: the apogee of the omics trilogy. [10.1038/nrm3314]. *Nat Rev Mol Cell Biol, 13*(4), 263-269.

Peironcely, J. E., Reijmers, T., Coulier, L., Bender, A., & Hankemeier, T. (2011). Understanding and Classifying Metabolite Space and Metabolite-Likeness. *PLoS ONE, 6*(12), e28966, doi:10.1371/journal.pone.0028966.

Peironcely, J. E., Rojas-Chertó, M., Tas, A., Vreeken, R., Reijmers, T., Coulier, L., et al. (2013). Automated Pipeline for De Novo Metabolite Identification Using Mass-Spectrometry-Based Metabolomics. *Analytical Chemistry, 85*(7), 3576-3583, doi:10.1021/ac303218u.

Pepe, M. S., Etzioni, R., Feng, Z., Potter, J. D., Thompson, M. L., Thornquist, M., et al. (2001). Phases of Biomarker Development for Early Detection of Cancer. *JNCI: Journal of the National Cancer Institute, 93*(14), 1054-1061, doi:10.1093/jnci/93.14.1054.

Pimentel, M., & Lezcano, S. (2007). Irritable bowel syndrome: Bacterial overgrowth—What's known and what to do. *Current Treatment Options in Gastroenterology, 10*(4), 328-337, doi:10.1007/s11938-007-0076-1.

Pinczowski, D., Ekbom, A., Baron, J., Yuen, J., & Adami, H.-O. (1994). Risk factors for colorectal cancer in patients with ulcerative colitis: A case-control study. *Gastroenterology, 107*(1), 117-120, doi:https://doi.org/10.1016/0016-5085(94)90068-X.

Prince, A. C., Myers, C. E., Joyce, T., Irving, P., Lomer, M., & Whelan, K. (2016). Fermentable Carbohydrate Restriction (Low FODMAP Diet) in Clinical Practice Improves Functional Gastrointestinal Symptoms in Patients with Inflammatory Bowel Disease. *Inflammatory Bowel Diseases, 22*(5), 1129-1136, doi:10.1097/MIB.0000000000000708.

Psychogios, N., Hau, D. D., Peng, J., Guo, A. C., Mandal, R., Bouatra, S., et al. (2011). The human serum metabolome. *PLoS One, 6*(2), e16957, doi:10.1371/journal.pone.0016957.

Ptolemy, A. S., & Rifai, N. (2010). What is a biomarker? Research investments and lack of clinical integration necessitate a review of biomarker terminology and validation schema. *Scandinavian Journal of Clinical and Laboratory Investigation, 70*(sup242), 6-14, doi:10.3109/00365513.2010.493354.

Quinn, R. A., Navas-Molina, J. A., Hyde, E. R., Song, S. J., Vázquez-Baeza, Y., Humphrey, G., et al. (2016). From Sample to Multi-Omics Conclusions in under 48 Hours. [10.1128/mSystems.00038-16]. *mSystems, 1*(2).

R, G. P., & J, S. S. (2010). Evidence-based dietary management of functional gastrointestinal symptoms: The FODMAP approach. *Journal of Gastroenterology and Hepatology, 25*(2), 252-258, doi:doi:10.1111/j.1440-1746.2009.06149.x.

Ramautar, R., Busnel, J.-M., Deelder, A. M., & Mayboroda, O. A. (2011). Enhancing the Coverage of the Urinary Metabolome by Sheathless Capillary Electrophoresis-Mass Spectrometry. *Analytical Chemistry, 84*(2), 885-892, doi:10.1021/ac202407v.

Ramautar, R., Mayboroda, O. A., Derks, R. J. E., van Nieuwkoop, C., van Dissel, J. T., Somsen, G. W., et al. (2008). Capillary electrophoresis-time of flight-mass spectrometry using noncovalently bilayer-coated capillaries for the analysis of amino acids in human urine. *ELECTROPHORESIS, 29*(12), 2714-2722, doi:10.1002/elps.200700929.

Raphael, B. M. A., Ben de Lacy, C., Paul, W., Tanzeela, K., Norman, M. R., Raj, P., et al. (2016). The use of a gas chromatography-sensor system combined with advanced statistical methods, towards the diagnosis of urological malignancies. *Journal of Breath Research, 10*(1), 017106.

Rattray, N. J. W., Deziel, N. C., Wallach, J. D., Khan, S. A., Vasiliou, V., Ioannidis, J. P. A., et al. (2018). Beyond genomics: understanding exposotypes through metabolomics. *Human Genomics, 12*, 4, doi:10.1186/s40246-018-0134-x.

Rieder, F., Kurada, S., Grove, D., Cikach, F., Lopez, R., Patel, N., et al. (2016). A Distinct Colon-Derived Breath Metabolome is Associated with Inflammatory Bowel Disease, but not its Complications. [Original Contributions]. *Clinical And Translational Gastroenterology, 7*, e201, doi:10.1038/ctg.2016.57

https://www.nature.com/articles/ctg201657#supplementary-information.

Robertson, D. G., Watkins, P. B., & Reily, M. D. (2011). Metabolomics in toxicology: preclinical and clinical applications. *Toxicological Sciences, 120 Suppl 1*, S146-170, doi:10.1093/toxsci/kfq358.

Rodríguez, L. A. G., & Ruigómez, A. (1999). Increased risk of irritable bowel syndrome after bacterial gastroenteritis: cohort study. [10.1136/bmj.318.7183.565]. *BMJ, 318*(7183), 565.

Rolfe, V. E., Fortun, P. J., Hawkey, C. J., & Bath-Hextall, F. J. (2006). Probiotics for maintenance of remission in Crohn's disease. *Cochrane Database of Systematic Reviews*(4), doi:10.1002/14651858.CD004826.pub2.

Røseth, A. G., Aadland, E., & Grzyb, K. (2004). Normalization of faecal calprotectin: a predictor of mucosal healing in patients with inflammatory bowel disease. *Scandinavian Journal of Gastroenterology, 39*(10), 1017-1020, doi:10.1080/00365520410007971.

Rossen, N. G., MacDonald, J. K., de Vries, E. M., D'Haens, G. R., de Vos, W. M., Zoetendal, E. G., et al. (2015). Fecal microbiota transplantation as novel therapy in gastroenterology: A systematic review. *World Journal of Gastroenterology : WJG, 21*(17), 5359-5371, doi:10.3748/wjg.v21.i17.5359.

Rossi, M., Aggio, R., Staudacher, H. M., Lomer, M. C., Lindsay, J. O., Irving, P., et al. (2018). Volatile Organic Compounds in Feces Associate With Response to Dietary Intervention in Patients With Irritable Bowel Syndrome. *Clinical Gastroenterology and Hepatology, 16*(3), 385-391.e381, doi:10.1016/j.cgh.2017.09.055.

Said, N., Gahoual, R., Kuhn, L., Beck, A., François, Y.-N., & Leize-Wagner, E. (2016). Structural characterization of antibody drug conjugate by a combination of intact, middle-up and bottom-up techniques using sheathless capillary electrophoresis – Tandem mass spectrometry as nanoESI infusion platform and separation method. *Analytica Chimica Acta, 918*, 50-59, doi:https://doi.org/10.1016/j.aca.2016.03.006.

Salt, W. B. (1997). *Irritable Bowel Syndrome & the Mind-Body/ Brain-Gut Connection* (1st ed.). Ohio: Parkview Publishing.

Sandler, R. S. (1990). Epidemiology of irritable bowel syndrome in the United States. *Gastroenterology, 99*(2), 409-415.

Santoru, M. L., Piras, C., Murgia, A., Palmas, V., Camboni, T., Liggi, S., et al. (2017). Cross sectional evaluation of the gut-microbiome metabolome axis in an Italian cohort of IBD patients. *Scientific Reports, 7*(1), 9523, doi:10.1038/s41598-017-10034-5.

Santoru, M. L., Piras, C., Murgia, A., Palmas, V., Camboni, T., Liggi, S., et al. (2018). Author Correction: Cross sectional evaluation of the gut-microbiome metabolome axis in an Italian cohort of IBD patients. *Scientific Reports, 8*(1), 4993, doi:10.1038/s41598-018-23330-5.

Saric, J., Want, E. J., Duthaler, U., Lewis, M., Keiser, J., Shockcor, J. P., et al. (2012). Systematic Evaluation of Extraction Methods for Multiplatform-Based Metabotyping: Application to the Fasciola hepatica Metabolome. *Analytical Chemistry, 84*(16), 6963-6972, doi:10.1021/ac300586m.

Sato, S., & Yanagisawa, S. (2010). Capillary electrophoresis–electrospray ionization-mass spectrometry using fused-silica capillaries to profile anionic metabolites. [journal article]. *Metabolomics, 6*(4), 529-540, doi:10.1007/s11306-010-0223-x.

Saude, E. J., & Sykes, B. D. (2007). Urine stability for metabolomic studies: effects of preparation and storage. [journal article]. *Metabolomics, 3*(1), 19-27, doi:10.1007/s11306-006-0042-2.

Saulnier, D. M., Riehle, K., Mistretta, T. A., Diaz, M. A., Mandal, D., Raza, S., et al. (2011). Gastrointestinal Microbiome Signatures of Pediatric Patients With Irritable Bowel

Syndrome. *Gastroenterology, 141*(5), 1782-1791, doi:https://doi.org/10.1053/j.gastro.2011.06.072.

Schicho, R., Shaykhutdinov, R., Ngo, J., Nazyrova, A., Schneider, C., Panaccione, R., et al. (2012). Quantitative Metabolomic Profiling of Serum, Plasma, and Urine by 1H NMR Spectroscopy Discriminates between Patients with Inflammatory Bowel Disease and Healthy Individuals. *Journal of Proteome Research, 11*(6), 3344-3357, doi:10.1021/pr300139q.

Schoepfer, A. M., Vavricka, S., Zahnd-Straumann, N., Straumann, A., & Beglinger, C. (2012). Monitoring inflammatory bowel disease activity: Clinical activity is judged to be more relevant than endoscopic severity or biomarkers. *Journal of Crohn's and Colitis, 6*(4), 412-418, doi:10.1016/j.crohns.2011.09.008.

Schymanski, E. L., Ruttkies, C., Krauss, M., Brouard, C., Kind, T., Dührkop, K., et al. (2017). Critical Assessment of Small Molecule Identification 2016: automated methods. [journal article]. *Journal of Cheminformatics, 9*(1), 22, doi:10.1186/s13321-017-0207-1.

Scolamiero, E., Cozzolino, C., Albano, L., Ansalone, A., Caterino, M., Corbo, G., et al. (2015). Targeted metabolomics in the expanded newborn screening for inborn errors of metabolism. [10.1039/C4MB00729H]. *Molecular BioSystems, 11*(6), 1525-1535, doi:10.1039/C4MB00729H.

Scoville, E. A., Allaman, M. M., Brown, C. T., Motley, A. K., Horst, S. N., Williams, C. S., et al. (2017). Alterations in lipid, amino acid, and energy metabolism distinguish Crohn's disease from ulcerative colitis and control subjects by serum metabolomic profiling. *Metabolomics, 14*(1), 17, doi:10.1007/s11306-017-1311-y.

Sebastiano, C., J., M. F.-P., & Serge, R. (2013). Clinical metabolomics paves the way towards future healthcare strategies. *British Journal of Clinical Pharmacology, 75*(3), 619-629, doi:doi:10.1111/j.1365-2125.2012.04216.x.

Sellon, R. K., Tonkonogy, S., Schultz, M., Dieleman, L. A., Grenther, W., Balish, E., et al. (1998). Resident Enteric Bacteria Are Necessary for Development of Spontaneous Colitis and Immune System Activation in Interleukin-10-Deficient Mice. *Infection and Immunity, 66*(11), 5224-5231.

Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., et al. (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res, 13*(11), 2498-2504, doi:10.1101/gr.1239303.

Shkoda, A., Werner, T., Daniel, H., Gunckel, M., Rogler, G., & Haller, D. (2007). Differential Protein Expression Profile in the Intestinal Epithelium from Patients with Inflammatory Bowel Disease. *Journal of Proteome Research, 6*(3), 1114-1125, doi:10.1021/pr060433m.

Shyur, L.-F., & Yang, N.-S. (2008). Metabolomics for phytomedicine research and drug development. *Current Opinion in Chemical Biology, 12*(1), 66-71, doi:http://dx.doi.org/10.1016/j.cbpa.2008.01.032.

Singhal, N., Kumar, M., Kanaujia, P. K., & Virdi, J. S. (2015). MALDI-TOF mass spectrometry: an emerging technology for microbial identification and diagnosis. *Frontiers in Microbiology, 6*, 791, doi:10.3389/fmicb.2015.00791.

Sipponen, T., Björkesten, C.-G. A. F., Färkkilä, M., Nuutinen, H., Savilahti, E., & Kolho, K.-L. (2010). Faecal calprotectin and lactoferrin are reliable surrogate markers of endoscopic response during Crohn's disease treatment. *Scandinavian Journal of Gastroenterology, 45*(3), 325-331, doi:10.3109/00365520903483650.

Sipponen, T., Savilahti, E., Kolho, K. L., Nuutinen, H., Turunen, U., & Farkkila, M. (2008). Crohn's disease activity assessed by fecal calprotectin and lactoferrin: correlation with

Crohn's disease activity index and endoscopic findings. *Inflammatory Bowel Diseases, 14*(1), 40-46, doi:10.1002/ibd.20312.

Soga, T., Baran, R., Suematsu, M., Ueno, Y., Ikeda, S., Sakurakawa, T., et al. (2006). Differential Metabolomics Reveals Ophthalmic Acid as an Oxidative Stress Biomarker Indicating Hepatic Glutathione Consumption. *Journal of Biological Chemistry, 281*(24), 16768-16776, doi:10.1074/jbc.M601876200.

Soga, T., Igarashi, K., Ito, C., Mizobuchi, K., Zimmermann, H.-P., & Tomita, M. (2009). Metabolomic Profiling of Anionic Metabolites by Capillary Electrophoresis Mass Spectrometry. *Analytical Chemistry, 81*(15), 6165-6174, doi:10.1021/ac900675k.

Soga, T., Ueno, Y., Naraoka, H., Ohashi, Y., Tomita, M., & Nishioka, T. (2002). Simultaneous Determination of Anionic Intermediates for Bacillus subtilis Metabolic Pathways by Capillary Electrophoresis Electrospray Ionization Mass Spectrometry. *Analytical Chemistry, 74*(10), 2233-2239, doi:10.1021/ac020064n.

Soubières, A. A., & Poullis, A. (2016). Emerging Biomarkers for the Diagnosis and Monitoring of Inflammatory Bowel Diseases. *Inflammatory Bowel Diseases, 22*(8), 2016-2022, doi:10.1097/mib.0000000000000836.

Spiegel, B. M. R., Farid, M., Esrailian, E., Talley, J., & Chang, L. (2010). Is Irritable Bowel Syndrome a Diagnosis of Exclusion?: A Survey of Primary Care Providers, Gastroenterologists, and IBS Experts. [Functional GI Disorders]. *The American Journal Of Gastroenterology, 105*, 848, doi:10.1038/ajg.2010.47.

Spiller, R., Aziz, Q., Creed, F., Emmanuel, A., Houghton, L., Hungin, P., et al. (2007). Guidelines on the irritable bowel syndrome: mechanisms and practical management. [10.1136/gut.2007.119446]. *Gut, 56*(12), 1770.

Spiller, R. C., Jenkins, D., Thornley, J. P., Hebden, J. M., Wright, T., Skinner, M., et al. (2000). Increased rectal mucosal enteroendocrine cells, T lymphocytes, and increased gut permeability following acuteCampylobacter enteritis and in post-dysenteric irritable bowel syndrome. *Gut, 47*(6), 804-811, doi:10.1136/gut.47.6.804.

Staudacher, H. M., Irving, P. M., Lomer, M. C. E., & Whelan, K. (2014). Mechanisms and efficacy of dietary FODMAP restriction in IBS. [Review Article]. *Nature Reviews Gastroenterology & Hepatology, 11*, 256, doi:10.1038/nrgastro.2013.259.

Staudacher, H. M., Lomer, M. C. E., Anderson, J. L., Barrett, J. S., Muir, J. G., Irving, P. M., et al. (2012). Fermentable Carbohydrate Restriction Reduces Luminal Bifidobacteria and Gastrointestinal Symptoms in Patients with Irritable Bowel Syndrome. *The Journal of Nutrition, 142*(8), 1510-1518, doi:10.3945/jn.112.159285.

Stephen, A. M., & Cummings, J. H. (1980). The microbial contribution to human faecal mass. *J Med Microbiol, 13*(1), 45-56, doi:10.1099/00222615-13-1-45.

Stephens, M., & Marvis, A. M. (2013). 5-Aminosalicylate Therapy. In P. Mamula, J. Markowitz, & R. Baldassano (Eds.), *Pediatric Inflammatory Bowel Disease* (2nd ed., pp. 281-288): Springer.

Stephens, N. S., Siffledeen, J., Su, X., Murdoch, T. B., Fedorak, R. N., & Slupsky, C. M. (2013). Urinary NMR metabolomic profiles discriminate inflammatory bowel disease from healthy. *Journal of Crohn's and Colitis, 7*(2), e42-e48, doi:http://dx.doi.org/10.1016/j.crohns.2012.04.019.

Strimbu, K., & Tavel, J. A. (2010). What are Biomarkers? *Current opinion in HIV and AIDS, 5*(6), 463-466, doi:10.1097/COH.0b013e32833ed177.

Sugimoto, M., Wong, D. T., Hirayama, A., Soga, T., & Tomita, M. (2010). Capillary electrophoresis mass spectrometry-based saliva metabolomics identified oral, breast and

pancreatic cancer-specific profiles. [journal article]. *Metabolomics, 6*(1), 78-95, doi:10.1007/s11306-009-0178-y.

Sumner, L. W., Amberg, A., Barrett, D., Beale, M., Beger, R., Daykin, C., et al. (2007). Proposed minimum reporting standards for chemical analysis. *Metabolomics, 3*, doi:10.1007/s11306-007-0082-2.

Sumner, L. W., Amberg, A., Barrett, D., Beale, M. H., Beger, R., Daykin, C. A., et al. (2007). Proposed minimum reporting standards for chemical analysis. [journal article]. *Metabolomics, 3*(3), 211-221, doi:10.1007/s11306-007-0082-2.

Sun, X., Lin, L., Liu, X., Zhang, F., Chi, L., Xia, Q., et al. (2016). Capillary Electrophoresis–Mass Spectrometry for the Analysis of Heparin Oligosaccharides and Low Molecular Weight Heparin. *Analytical Chemistry, 88*(3), 1937-1943, doi:10.1021/acs.analchem.5b04405.

Swan, C., Duroudier, N. P., Campbell, E., Zaitoun, A., Hastings, M., Dukes, G. E., et al. (2013). Identifying and testing candidate genetic polymorphisms in the irritable bowel syndrome (IBS): association with TNFSF15 and TNFα. [10.1136/gutjnl-2011-301213]. *Gut, 62*(7), 985.

Tana, C., Umesaki, Y., Imaoka, A., Handa, T., Kanazawa, M., & Fukudo, S. (2010). Altered profiles of intestinal microbiota and organic acids may be the origin of symptoms in irritable bowel syndrome. *Neurogastroenterology & Motility, 22*(5), 512-e115, doi:10.1111/j.1365-2982.2009.01427.x.

Therrell, B. L., & Adams, J. (2007). Newborn screening in North America. [journal article]. *Journal of Inherited Metabolic Disease, 30*(4), 447-465, doi:10.1007/s10545-007-0690-z.

Thierry, L. (2013). Differential Diagnosis of Inflammatory Bowel Disease. In M. Petar, B. N. Robert, & M. E. Jonathan (Eds.), *Pediatric Inflammatory Bowel Disease* (Second edition ed.): Springer.

Topping, D. L., & Clifton, P. M. (2001). Short-Chain Fatty Acids and Human Colonic Function: Roles of Resistant Starch and Nonstarch Polysaccharides. *Physiological Reviews, 81*(3), 1031-1064.

Van Noorden, R. (2014). Global scientific output doubles every nine years. *Nature News Blog*

Verdu, E. F., Bercik, P., Huang, X. X., Lu, J., Al-Mutawaly, N., Sakai, H., et al. (2008). The role of luminal factors in the recovery of gastric function and behavioral changes after chronic Helicobacter pylori infection. [10.1152/ajpgi.90316.2008]. *American Journal of Physiology - Gastrointestinal and Liver Physiology, 295*(4), G664-G670.

Virgin, Herbert W. (2014). The Virome in Mammalian Physiology and Disease. *Cell, 157*(1), 142-150, doi:https://doi.org/10.1016/j.cell.2014.02.032.

W, E. B., K, C. W., J, M. D., & J, W. P. (2007). Tegaserod for the treatment of irritable bowel syndrome and chronic constipation. *Cochrane Database of Systematic Reviews, 4*,

Wang, L.-H., Fang, X.-C., & Pan, G.-Z. (2004). Bacillary dysentery as a causative factor of irritable bowel syndrome and its pathogenesis. *Gut, 53*(8), 1096-1101, doi:10.1136/gut.2003.021154.

Wehrens, R., Hageman, J. A., van Eeuwijk, F., Kooke, R., Flood, P. J., Wijnker, E., et al. (2016). Improved batch correction in untargeted MS-based metabolomics. *Metabolomics, 12*, 88, doi:10.1007/s11306-016-1015-8.

Whorwell, P. J. (2005). Review article: the history of hypnotherapy and its role in the irritable bowel syndrome. *Alimentary Pharmacology & Therapeutics, 22*(11-12), 1061-1067, doi:10.1111/j.1365-2036.2005.02697.x.

Whorwell, P. J., Altringer, L., Morel, J., Bond, Y., Charbonneau, D., O'Mahony, L., et al. (2006). Efficacy of an Encapsulated Probiotic Bifidobacterium infantis 35624 in Women with

Irritable Bowel Syndrome. *The American Journal Of Gastroenterology, 101*(7), 1581-1590.

Wilks, S. (1859, Morbid appearances in the intestine of Miss Bankes. *London Medical Gazette* pp. 264-265.

William, T. B. (1978). Structure and Biosynthesis of Connective Tissue Proteoglycans. In H. Martin (Ed.), *THE GLYCOCONJUGATES* (1 ed., Vol. 2): Academic Press, Inc.

Williams, H. R. T., Cox, I. J., Walker, D. G., North, B. V., Patel, V. M., Marshall, S. E., et al. (2009). Characterization of Inflammatory Bowel Disease With Urinary Metabolic Profiling. *The American Journal Of Gastroenterology, 104*(6), 1435-1444, doi:http://www.nature.com/ajg/journal/v104/n6/suppinfo/ajg2009175s1.html.

Wishart, D. S. (2011). Advances in metabolite identification. *Bioanalysis, 3*(15), 1769-1782, doi:10.4155/bio.11.155.

Wishart, D. S., Jewison, T., Guo, A. C., Wilson, M., Knox, C., Liu, Y., et al. (2013). HMDB 3.0--The Human Metabolome Database in 2013. *Nucleic Acids Research, 41*(Database issue), D801-807, doi:10.1093/nar/gks1065.

Wishart, D. S., Knox, C., Guo, A. C., Eisner, R., Young, N., Gautam, B., et al. (2009). HMDB: a knowledgebase for the human metabolome. *Nucleic Acids Research, 37*(Database issue), D603-610, doi:10.1093/nar/gkn810.

Wojcik, R., Dada, O. O., Sadilek, M., & Dovichi, N. J. (2010). Simplified capillary electrophoresis nanospray sheath-flow interface for high efficiency and sensitive peptide analysis. *Rapid Communications in Mass Spectrometry, 24*(17), 2554-2560, doi:doi:10.1002/rcm.4672.

Wolf, S., Schmidt, S., Müller-Hannemann, M., & Neumann, S. (2010). In silico fragmentation for computer assisted identification of metabolite mass spectra. [journal article]. *BMC Bioinformatics, 11*(1), 1-12, doi:10.1186/1471-2105-11-148.

Xia, J., Broadhurst, D. I., Wilson, M., & Wishart, D. S. (2013). Translational biomarker discovery in clinical metabolomics: an introductory tutorial. [journal article]. *Metabolomics, 9*(2), 280-299, doi:10.1007/s11306-012-0482-9.

Yamamoto, H., & Sasaki, K. (2017). Metabolomics-based approach for ranking the candidate structures of unidentified peaks in capillary electrophoresis time-of-flight mass spectrometry. *ELECTROPHORESIS, 38*(7), 1053-1059, doi:10.1002/elps.201600328.

Yamashita, K., Okuyama, M., Watanabe, Y., Honma, S., Kobayashi, S., & Numazawa, M. (2007). Highly sensitive determination of estrone and estradiol in human serum by liquid chromatography–electrospray ionization tandem mass spectrometry. *Steroids, 72*(11–12), 819-827, doi:http://dx.doi.org/10.1016/j.steroids.2007.07.003.

Yano, Jessica M., Yu, K., Donaldson, Gregory P., Shastri, Gauri G., Ann, P., Ma, L., et al. (2015). Indigenous Bacteria from the Gut Microbiota Regulate Host Serotonin Biosynthesis. *Cell, 161*(2), 264-276, doi:10.1016/j.cell.2015.02.047.

Yin, P., Peter, A., Franken, H., Zhao, X., Neukamm, S. S., Rosenbaum, L., et al. (2013). Preanalytical aspects and sample quality assessment in metabolomics studies of human blood. *Clinical Chemistry, 59*(5), 833-845, doi:10.1373/clinchem.2012.199257.

Zeber-Lubecka, N., Kulecka, M., Ambrozkiewicz, F., Paziewska, A., Goryca, K., Karczmarski, J., et al. (2016). Limited prolonged effects of rifaximin treatment on irritable bowel syndrome-related differences in the fecal microbiome and metabolome. *Gut Microbes, 7*(5), 397-413, doi:10.1080/19490976.2016.1215805.

Zhang, W., Hankemeier, T., & Ramautar, R. (2017). Next-generation capillary electrophoresis–mass spectrometry approaches in metabolomics. *Current Opinion in Biotechnology, 43*, 1-7, doi:https://doi.org/10.1016/j.copbio.2016.07.002.

Zivkovic, A. M., Wiest, M. M., Nguyen, U. T., Davis, R., Watkins, S. M., & German, J. B. (2009). Effects of sample handling and storage on quantitative lipid analysis in human serum. *Metabolomics, 5*(4), 507-516, doi:10.1007/s11306-009-0174-2.

Zucchelli, M., Camilleri, M., Nixon Andreasson, A., Bresso, F., Dlugosz, A., Halfvarson, J., et al. (2011). Association of TNFSF15 polymorphism with irritable bowel syndrome. [10.1136/gut.2011.241877]. *Gut, 60*(12), 1671.

Chapter II

A Robust and High Throughput Method for Anionic Metabolite Profiling: Preventing Polyimide Aminolysis and Capillary Breakages under Alkaline Conditions in CE-MS

Thesis chapter is derived from a published peer-reviewed article, where I am the first author and my research supervisor is the corresponding author

Mai Yamamoto, Ritchie Ly, Biban Gill, Yujie Zhu, Jose Moran-Mirabal, and Philip Britz-McKibbin* *Anal. Chem.* **2016**, *88,* 10710-10719.

M.Y. conducted most of the instrumental optimization, sample analysis, data processing and interpretation on data for anionic metabolite analysis when using multiplexed separations by CE-MS. I also wrote an initial manuscript draft used for publication. R.L. and B.G. contributed on supporting information generation, including images of the aminolysis byproduct analysis and ion migration time modeling. Y.Z. provided technical assistance on microscope. J.M.M. and P.B.M. provided feedback on the manuscript draft.

Chapter II: A Robust and High Throughput Method for Anionic Metabolite Profiling: Preventing Polyimide Aminolysis and Capillary Breakages under Alkaline Conditions in CE-MS

2.1 Abstract

Capillary electrophoresis-mass spectrometry (CE-MS) represents a high efficiency microscale separation platform for untargeted profiling of polar/ionic metabolites that is ideal for volume-restricted biological specimens with minimal sample workup. Despite these advantages, the long-term stability of CE-MS remains a major obstacle hampering its widespread application in metabolomics notably for routine analysis of anionic metabolites under negative ion mode conditions. Herein, we report for the first time that commonly used ammonia containing buffers compatible with electrospray ionization (ESI)-MS can compromise the integrity of fused-silica capillaries via aminolysis of their outer polyimide coating. Unlike organic solvent swelling effects, this chemical process occurs under aqueous conditions that is dependent on ammonia concentration, buffer pH and exposure time resulting in a higher incidence of capillary fractures and current errors during extended operation. Prevention of polyimide aminolysis is achieved by using weakly alkaline ammonia containing buffers (pH < 9) in order to preserve the tensile strength of the polyimide coated fused-silica capillary. Alternatively, less nucleophilic primary/secondary amines can be used as electrolytes without polyimide degradation, whereas chemically resistant polytetrafluoroethylene coating materials offer higher pH tolerance in ammonia. In this work, multi-segment injection (MSI)-CE-MS was used as multiplexed separation platform for high throughput profiling of anionic metabolites when using optimized buffer conditions to prevent polyimide degradation. A diverse range of acidic metabolites in human urine were reliably measured by MSI-CE-MS via serial injection of seven urine samples within a single run, including organic acids, food-specific markers, microbial-derived compounds and over-the-counter drugs as their sulfate and glucuronide conjugates. This approach offers excellent throughput (< 5 min/sample) and acceptable intermediate precision (average $CV \approx 16\%$) with high separation efficiency as reflected analysis of 30 anionic metabolites following 238 repeated sample injections of human urine over 3 days while using a single non-isotope internal standard for data normalization. Careful optimization and rigorous validation of CE-MS protocols are crucial for developing a rapid, low cost and robust screening platform for metabolomics that is amenable to large-scale clinical and epidemiological studies.

2.2 Introduction

Although developed primarily as synthetic tubing materials for the fiber optic telecommunications industry,[1] the advent of open tubular fused-silica capillary columns has revolutionized separation science with the introduction of high efficiency microscale analytical techniques, including gas chromatography (GC), nano-liquid chromatography (nano-LC) and capillary electrophoresis (CE).[2] Polyimide is the outer polymer coating material of choice in the manufacture of rugged fused-silica capillaries due to its excellent thermal stability and chemical resistance that imparts flexibility with abrasive protection to prevent breakage. Despite these characteristics, polyimide coated fused-silica capillaries are susceptible to random brittleness problems during the capillary drawing process, including variations in inner surface activity and capillary dimensions that can impact separation performance.[3] Polyimide coated fused-silica capillaries are also prone to swelling after prolonged contact with certain organic solvents (*e.g.,* acetonitrile) that leads to clogging events and capillary breakage in non-aqueous CE and capillary electrochromatography (CEC).[4] Extensive heat curing prior to use has been reported to provide greater tolerance in organic solvents that maintain separation efficiency with longer capillary lifetimes.[5] Alternatively, a section of the polyimide coating at the capillary inlet/outlet in contact with organic solvent can be stripped chemically with concentrated nitric acid or charred by heating,[6] however this step can compromise mechanical stability. The latter procedure is often adopted in CE-MS methods with a coaxial sheath liquid interface in order to prevent swelling effects since a make-up flow of an electrolyte solution containing organic solvent is required for stable electrospray generation.[7] Similarly, more sensitive sheathless porous tip interface designs in CE-MS rely on an uncoated and thin walled/etched fused-silica tubing to ground the electrical circuit – a fragile section of the capillary requiring careful handling to prevent fractures or breakage.[8] As a result, the robustness of CE-MS technology remains an on-going challenge that has limited its application by a wider community of researchers involved in large-scale metabolomic profiling of complex biological fluids,[9] despite recent advances at improving concentration sensitivity, migration time reproducibility and sample throughput.[10]

Several studies have demonstrated that CE-MS offers complimentary selectivity in metabolomics in comparison to both reversed-phase LC and hydrophilic interaction chromatography (HILIC) notably for strongly ionic and poorly retained metabolites in volume-restricted or mass-limited biological samples, including bio-banked specimens.[11-14] Overall, there is growing consensus that excellent performance is achieved for untargeted profiling of metabolites and peptides by CE-MS under strongly acidic conditions using acetic acid or formic acid as a background electrolyte (BGE) with positive ion mode detection.[15-17] For instance, a recent inter-lab validation study for tryptic peptide mapping by CE-MS using a standardized operating protocol demonstrated good method robustness across 13 independent laboratories from academia and industry when using different ion source interfaces and mass analyzers.[18] Indeed, CE-MS metabolomic studies to date have primarily focused on comprehensive profiling of cationic metabolites[19-26] given the lower sensitivity and poor stability long associated with anionic metabolites under negative ion mode detection.[27] Büscher et al.[28] reported a cross-platform study for quantitative metabolomic analyses of yeast extracts that highlighted CE-MS suffers from poor reliability in comparison to LC-MS and GC-MS, which was attributed to the protocol used for anionic metabolite profiling. This problem was later linked to incidental corrosion of the stainless steel spray needle causing clogging of the capillary outlet with frequent current drops while operating CE-MS under reversed polarity with a cationic polymer-coated capillary[29] several years after publication of their original protocol[30] and subsequent applications.[31, 32] As a result, there have been sparse reports on the use of uncoated capillaries using simple alkaline buffers for anionic metabolite profiling using CE-MS under normal polarity configuration.[33] Kok et al.[34, 35] compared different background electrolyte (BGE) and sheath liquid additives for anionic metabolite profiling and concluded that triethylamine (pH 11.7) provided better sensitivity and greater metabolome coverage than conventional ammonia containing buffers. However, persistent "memory effects" when using the same instrument under positive ion detection occurs since TEA strongly adsorbs onto surfaces that contributes to ion suppression due to its high surface activity and gas phase proton affinity.[36] Recently, Gulersonmez et al.[37] developed a

sheathless CE-MS method for reliable analysis of anionic metabolites (*e.g.,* sugar phosphates) from glioblastoma cell extracts when using acetic acid as the BGE (pH 2.2) with negative ion mode detection. However, this approach is viable for separating only strongly acidic metabolites ($pK_a < 3.2$) while being limited to small sets of biological samples due to the finite lifespan of the porous-etched capillary tip. Thus, there is an urgent need for a simple, cost effective yet robust CE-MS platform amenable to large-scale metabolomic investigations involving complex biological samples, such as human urine that is composed of a diverse range of acidic/ionic compounds reflecting diet, chemical exposures, gut microflora and host metabolism.[38]

Herein we demonstrate for the first time that alkaline aqueous ammonia solutions widely used in ESI-MS lead to chemical degradation of the outer polyimide coating, which compromises the tensile strength of fused-silica capillaries during extended operation in CE-MS. Aminolysis of polyimide results in an increased incidence of capillary breakages that are prevented when using mildly basic ammonia solutions (pH < 9) at low ionic strength, ammonia-free or less nucleophilic primary or secondary amine buffers, as well as chemically resistant polytetrafluroethylene coated capillaries. A multiplexed separation platform based on multisegment injection (MSI)-CE-MS[24] was rigorously validated for high throughput anionic metabolite profiling when analyzing 238 repeated sample injections of pooled 24 h urine over 3 days under optimized buffer conditions that avoid polyimide dissolution. Accurate prediction of the electromigration behaviour of diverse classes of acidic metabolites in human urine is also demonstrated to assist in the identification of urinary metabolites in cases when authentic standards or MS/MS spectra are unavailable. This work sheds new light into longstanding concerns regarding the robustness of CE-MS technology highlighting the need for careful attention to buffer compatibility in method development.

2.3 Experimental Section

2.3.1 Chemicals and reagents. All chemicals were purchased from Sigma Aldrich (St. Louis, MO, USA) unless otherwise stated. Deuterium oxide (D, 99.9%) was purchased

from Cambridge Isotope Laboratories, Inc. (Tewksbury, MA, USA). Stock solutions for all internals standards and metabolite standards, as well as background electrolyte (BGE) and sheath liquid solutions were prepared in deionized water from a Barnstead EASYpure® II LF system (Dubuque, IA, USA). HPLC-grade methanol and acetonitrile were used to prepare sheath liquid and BGE solutions, respectively. All stock solutions for chemical standards were stored at 4°C.

2.3.2 CE-MS instrumentation. All flexible open tubular fused-silica capillaries with standard polyimide (TSP) or deep UV transparent polytetrafluoronated (TSU) outer coatings were purchased from Polymicro Technologies Inc. (Phoenix, AZ, USA) with inner diameters of 50 μm and outer diameters of 363 μm. CE-MS experiments were performed on an Agilent G7100A CE instrument (Agilent Technologies Inc., Mississauga, ON, Canada) equipped with a coaxial sheath liquid Jetstream electrospray ion source with heated nitrogen gas, which were coupled to an Agilent 6230 TOF-MS system. Separations were performed using a 110 cm polyimide coated fused-silica capillary using an applied voltage of 30 kV under normal polarity configuration at 25°C. In order to reduce total analysis times, a hydrodynamic pressure gradient was also applied during voltage application at a rate of 1 mbar min^{-1} or 0.1 kPa min^{-1} from 10 to 20 min, which was then increased to 3 mbar min^{-1} or 0.3 kPa min^{-1} from 20 min until the end of the separation. The background electrolyte (BGE) consisted of 50 mM ammonium bicarbonate, pH 8.50 that was adjusted with 1.5 M ammonium hydroxide unless otherwise stated, whereas the sheath liquid was composed of 50:50 MeOH:H$_2$O. New capillaries were conditioned by flushing at high pressure (900 mbar or 90 kPa) each morning for 15 min each in the order of MeOH, 0.1 M NaOH, deionized H$_2$O, 1 M formic acid, and deionized H$_2$O followed by BGE. In between runs, the capillary was flushed using deionized H$_2$O for 3 min followed by equilibration with BGE for 5 min. A 7-sample segment serial injection format was used for MSI-CE-MS as a multiplexed separation platform[24] to enhance sample throughput for anionic metabolite profiling under negative ion mode detection. This was performed by a programmed hydrodynamic injection sequence at 100 mbar or 10 kPa comprising seven alternating human urine and BGE plugs that is equivalent to filling about one third of total

capillary length, namely (1) 5 s sample, 40 s BGE spacer, (2) 5 s sample, 40 s BGE spacer, (3) 5 s sample, 40 s BGE spacer; (4) 5 s sample, 40 s BGE spacer; (6) 5 s sample, 40 s BGE spacer; (7) 5 s sample, 5 s BGE spacer. Between runs, the capillary was flushed at high pressure (900 mbar or 90 kPa) with deionized water for 5 min and then BGE for 10 min. The TOF-MS was operated under negative-ion mode conditions that spanned a mass range of m/z 50-1700, with an acquisition rate of 500 ms/spectrum. The ESI conditions were Vcap = 2000 V, nozzle voltage = 2000 V, nebulizer gas =10 psi, sheath gas =3.5 L/min at 195 °C, drying gas 8 L/min at 300 °C. Also, the MS voltage settings were fragmentor = 120 V, skimmer = 65V and Oct1 RF = 750 V.

2.3.3 Tensile strength of fused-silica capillaries exposed to alkaline buffer solutions.
Passive exposure studies on 2 x 5 cm length segments of polyimide-coated or polytetrafluoroethylene coated fused-silica capillaries (50 μm ID; 363 μm OD) were performed in triplicate in three separate 15 mL plastic centrifuge tubes. All six solution-exposed capillaries were then tested for changes in tensile strength by a single volunteer who manually bent the capillary to 90^0 every day for the first week and then every third day in subsequent weeks for a period of 26 days. The day upon which a capillary fractured with bending force following buffer exposure was noted and the broken edges were inspected using a Nikon ECLIPSE LV100N POL microscope equipped with objective lens and Q IMAGING RETIGA 2000R camera (Nikon Canada, Mississauga, ON, Canada). Various aqueous alkaline buffer solutions were examined in this study whose composition varied in terms of electrolyte type (*i.e.,* with/without ammonia), ionic strength, and buffer pH, which were left at room temperature. 50 mM borate and 50 mM ethylamine buffers at pH 10.0 (pH adjusted with 1M formic acid) were used as controls relative to 50 mM ammonium bicarbonate buffer at pH 8.5, 9.3 and 10 (pH adjusted with 1.5 M ammonium hydroxide), as well as 50 mM or 100 mM ammonium bicarbonate at pH 10 (pH adjusted with 1 M LiOH). In addition to evaluating the chemical stability of conventional polyimide coated capillary tubes to alkaline ammonia-based solutions, polytetrafluoroethylene coated capillary segments were also compared as a control. To assess accelerated degradation by

aminolysis during high pressure flushing of solution in CE, a 110 cm polyimide fused-silica capillary was continuously flushed for 24 h with 50 mM ammonium bicarbonate buffer at pH 8.5 or 10.0 using a pressure of 900 mbar or 90 kPa. The inlet and outlet section of each capillary end was then imaged under the light microscope to confirm changes in the morphology of polyimide coating as a function of buffer pH (pH > 9) that was also associated with reduced tensile strength, increased fragility, shorter lifespans and polyimide dissolution following prolonged exposure and/or flushing at high pressure.

2.3.4 Characterization of ammonia-induced degradation products in solution. Aliquots of the brown-coloured solutions were taken from completely solubilized polyimide from bare fused-silica capillaries following prolonged exposure (> 120 days) in 100 mM ammonium bicarbonate, pH 10. Blank-corrected UV-vis absorbance spectra were acquired from a five-fold diluted sample of polyimide solution in 50 mM borate, pH 9.0 using a Carey 50 spectrophotometer (Varian Inc., Palo Alto, CA, USA) at room temperature with high resolution scanning from 200 nm to 800 nm. ^{1}H-NMR studies were performed using a Bruker AV600 MHz spectrometer following solvent evaporation of a 0.25 mL samples using a Vacufuge Plus system (Eppendorf Inc. Mississauga, ON, Canada), reconstituted to 0.5 mL in deuterium oxide (D_2O) and transferred into a 5 mm id NMR tube. Chemical shifts and integrated proton signals for three distinct chemical environments were measured in ^{1}H NMR spectra. Chemical degradation by-products of the solubilized polyimide coating were identified by MSI-CE-MS when using a dilution trend filter in MSI-CE-MS[26] together with Molecular Feature Extractor in Mass Hunter software (Agilent Technologies Inc.). Samples were analyzed using 50 mM ammonium bicarbonate, pH 8.5 as BGE under negative ion mode (anionic metabolites), as well as 1M formic acid, pH 1.8 under positive ion mode (cationic metabolites), which confirmed the detection of six unknown species in solution following polyimide degradation (two anionic and four cationic species) defined by their *m/z:*RMT and most likely elemental formula.

2.3.5 Sample pretreatment of human urine samples. All human urine samples analyzed were pooled 24 h urine specimens derived from the Prospective Urban Rural

Epidemiological Study (PURE) as described in our previous study for urinary iodine status determination.[39] Urine samples were stored frozen at -80°C prior to their thawing at room temperature. Prior to analysis, about 50 µL of urine was centrifuged at 1,500 g for 5 min to precipitate particulate matter, then the supernatant was diluted five-fold with deionized water containing internal standard (50 µM NMS). Technical replicates of the pooled specimens ($n = 20$) were prepared in advance and kept frozen at –80 °C. Fresh aliquots were thawed three times daily and exchanged every five analysis runs in MSI-CE-MS.

2.3.6 Optimization of background electrolyte for anionic metabolite profiling. In order to avoid aminolysis of polyimide coated fused-silica capillaries to enable robust profiling of acidic metabolites in urine by MSI-CE-MS with negative ion mode detection, volatile ammonium containing buffers were limited to a pH 8.5 unlike primary or secondary amine buffers (*e.g.,* ethylamine). All BGEs tested were compared in terms of their impact on peak resolution and separation efficiency for representative urinary metabolites, as well as overall detector response based on measured signal-to-noise ratio (*S/N*) for ions. Five equimolar BGE systems (50 mM) were evaluated for MSI-CE-MS using a 7 serial plug injection format for diluted (5-fold in de-ionized water) pooled human urine, including ammonium bicarbonate (pH = 8.5), ammonium acetate (pH = 8.5), ethylamine (pH = 10.0), diethylamine (pH = 11.3) and pyrrolidine (pH = 11.5). MSI-CE-MS analyses were performed over two day using the same capillary and at least four sets of data were collected for each buffer system.

2.3.7 Intermediate precision for anionic metabolite profiling of urine. Daily mass tuning, cleaning of CE electrode (at inlet) and capillary conditioning steps were conducted by a single individual each morning using a standardized operating protocol to minimize sample carryover and systematic bias. The ion source was also rinsed with 50:50 isopropyl alcohol:H_2O and dried for 30 min prior to starting. In addition, a new polyimide coated fused-silica capillary, sheath liquid solution and BGE for CE-MS were prepared each day over three days in order to introduce random variation to the analyses. A standard metabolite mixture containing 20 organic acids with internal standard (50 µM NMS) was

then analyzed each morning to check the stability of the CE current, as well as migration times, resolution and ion responses for representative urinary metabolites (*e.g.,* uric acid, indoxyl sulfate etc.). MSI-CE-MS runs that were observed to have significantly low currents or irreversible current drops were excluded from data analysis due to instrumental bias. This problem was found to occasionally arise (3 out of 45 runs) as a result of improper alignment of capillary inlet with vial during sample injection leading to incidental injection of an air plug. Each morning following conditioning of a new fused-silica capillary, a standard mixture of anionic metabolites was analyzed followed by two consecutive equilibration runs of a five-fold diluted urine were performed by MSI-CE-MS using a 7-sample plug format for stabilization of the electroosmotic flow (EOF) with urine matrix prior to completion of about 15 runs or 105 sample injections comprising three technical replicates of urine (prepared every 5 runs) per day. A fresh BGE vial was introduced every four MSI-CE-MS runs to avoid buffer depletion and sample carry-over, whereas a fresh aliquot of diluted pooled urine sample was replaced every 8 runs to prevent evaporation losses. Between runs, the capillary was flushed at high pressure (900 mbar or 90 kPa) with deionized water for 5 min and then BGE for 10 min. A total of 45 runs were performed by MSI-CE-MS over three days, which included a standard mixture and urine samples for equilibration each morning, as well as spiked samples required for urinary metabolite identification. After rejecting runs that exhibited current drops, 34 runs acquired over 3 days equivalent to a total of 238 individual injections (34 runs x 7 injections per run = 238 total injections) of the same pooled 24 h urine sample were processed for inter-day reproducibility studies. The migration times and integrated peak areas for 30 representative acidic metabolites in urine were processed using Mass Hunter Qualitative Analysis software v B.06.00 SP (Agilent Technologies Inc.) and then normalized relative to NMS as internal standard to correct for differences in EOF and in-capillary injection volume between samples, respectively.[24] The intermediate precision of MSI-CE-MS was evaluated in terms of relative migration times (RMTs) and relative peak areas (RPA) for confirmed (*i.e.,* identified with authentic standards) and putative anionic metabolites (*i.e.,* denoted by unique *m/z*:RMT and likely molecular formula with corresponding mass error) in human

urine. All data processing and statistical analysis was performed using Igor Pro 5.0 (Wavemetrics Inc., Lake Oswego, OR, USA) and Excel (Microsoft Inc., Redmond, WA, USA).

2.3.8 Modeling of ion migration behavior to support metabolite identification. A chemically diverse set of 30 acidic metabolites consistently detected in pooled 24 h human urine samples using MSI-CE-MS (*e.g.,* mono- and divalent organic acids, zwitter-ions, glucuronide and sulfate conjugates etc.) were selected as a training set to accurately predict the relative migration time (RMT) of an ion (relative to NMS) based its absolute electrophoretic mobility (μ_o) as described by the Hubbard-Onsager equation[40] that includes the impact of hydrodynamic and dielectric friction for describing ion migration behavior in aqueous solutions at infinite dilution:

$$\mu_0 = \frac{az_0}{V^c + (\frac{bz_0{}^2}{V})} \tag{1}$$

Thus, two fundamental parameters are needed for determination of μ_o as derived from the chemical structure of a putative ion, namely its effective charge (z_o at pH 8.5) and molecular volume (*V*), whereas coefficient terms used in *eq (1)* are *a* = 0.00132, *b* = -18.8 and *c* = 0.335 as previously described.[41] In this work, z_o was determined using pK_a values listed in the Human Metabolome Database (http://www.hmdb.ca) based on the effective charge state of an acidic urinary metabolite at pH 8.5 for CE separation using the Henderson-Hasselbach equation. Also, *V* was determined *in silico* using Chem3D Pro software (Cambridge Soft, Cambridge, MA, USA), where each chemical structure was energy-minimized using an iterative molecular mechanics 2 (MM2) algorithm to determine a stable molecular conformation prior to computation of its Connolly solvent-excluded molecular volume. A correlation plot between experimentally measured RMTs for 30 urinary metabolites and predicted μ_o, demonstrated good linearity (R^2 = 0.974) and accuracy (average bias = +0.3% with range from +7.6% to -9.1%). This approach was used to confirm the assignment of putatively identified urinary metabolites based on its accurate

mass and most likely molecular formula in cases when authentic standards were not available or precursor ion signals were too low for acquisition of MS/MS spectra.

2.4 Results and Discussion

Our previous study[42] demonstrated excellent selectivity for resolving weakly acidic native estrogens and their intact sulfate and glucuronide conjugates in urine, including positional isomers when using CE-MS. However, sporadic capillary breakage and current instability events with prolonged usage of aqueous buffers containing ammonia prompted us to further investigate the underlying cause of this problem, which was unrelated to organic solvent-induced swelling or anodic corrosion of sprayer. **Figure 1(a)** depicts representative images of short sections of polyimide coated fused-silica capillaries following long-term exposure (70 days) to aqueous alkaline solutions that vary in terms of electrolyte type, buffer pH and ionic strength. It was apparent that exposure to 50 mM ammonium bicarbonate, pH 10 (or ammonium acetate) resulted in softening/warping of the outer polyimide coating with gradual dissolution of the polymer into solution. In fact, the first evidence of distortion of the polyimide coating was observed within 12 days first noticeable at the cut ends of the capillary, whereas complete dissolution/stripping of polyimide occurred within 2 days when exposed to 1 M ammonium hydroxide. This effect was not observed following exposure studies performed in equimolar "ammonia-free" aqueous alkaline buffers (50 mM borate, pH = 10) or primary/secondary amine buffer electrolytes, such as ethylamine (pH = 10) as shown in **Figure 1(a)**. Similarly, no

(a) Stability of Polyimide-coated Fused-silica Capillary in Alkaline Conditions with Ammonia

Dissolution of polyimide in buffer

pH 10 NH$_4$HCO$_3$ pH 10 Ethylamine   pH 8.5 NH$_4$HCO$_3$

pH 8.5   pH 10

Slippage of polyimide coating in ammonia-based buffers (pH > 9) with high pressure flushing
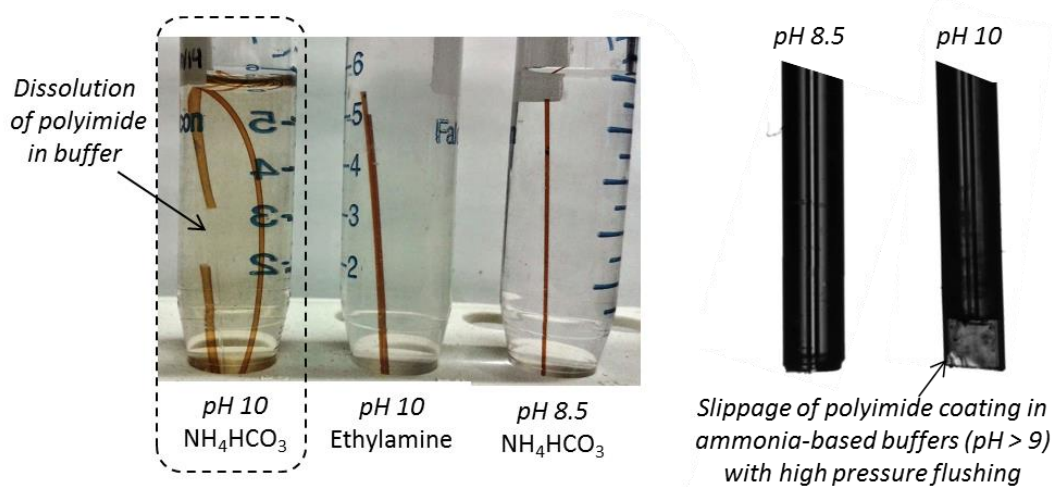
(b) Fracture Resistance for Exposed Polyimide Coated Fused-silica Capillaries

Lower tensile strength: pH > 9; Higher [NH$_4$+]   *Non-ammonium buffers*

Teflon-coated capillary

pH adjusted with LiOH

Days of Exposure

pH 8.5 50 mM NH$_4$HCO$_3$ | pH 9.3 50 mM NH$_4$HCO$_3$ | pH 10 50 mM NH$_4$HCO$_3$ | pH 10* 50 mM NH$_4$HCO$_3$ | pH 10 100 mM NH$_4$HCO$_3$ | pH 10* 100 mM NH$_4$HCO$_3$ | pH 10 100 mM Ethylamine | pH 10 100 mM Borate
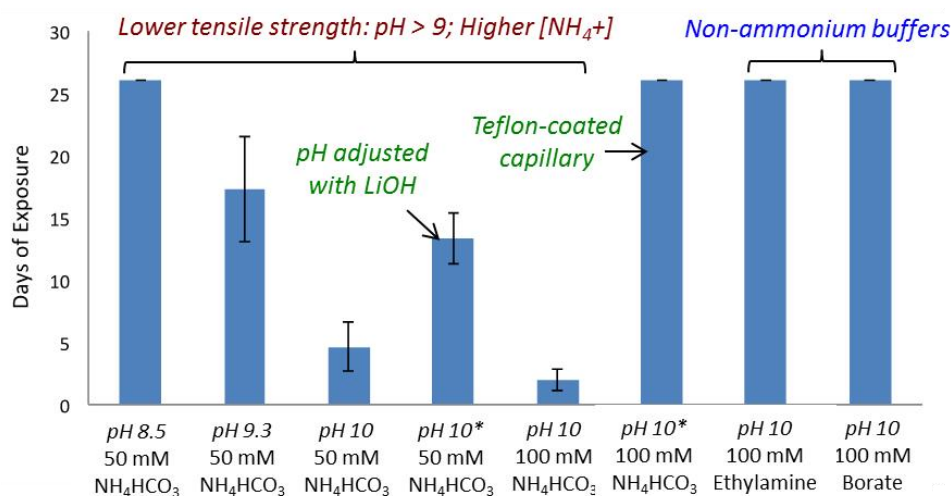
Figure 2.1 (a) Images depicting the impact of prolonged exposure (70 days) of polyimide coated fused-silica capillaries in aqueous alkaline ammonium buffers that result in softening/deformation of the outer coating and polymer dissolution. High pressure (90 kPa) flushing with 50 mM ammonium bicarbonate buffers for 24 h demonstrate elongation of the outer polymer coating beyond the fused-silica capillary tip at pH 10 that is not observed at pH 8.5 **(b)** A comparison of changes in the tensile strength and resistance to fractures with repeated bending (90°) to a series of fused-silica capillary segments (where error bars represent ±1s, *n*=6) exposed to different aqueous alkaline solution (*i.e.,* buffer type, pH, ionic strength), which confirms that higher ammonia concentrations and elevated pH conditions accelerate polyimide aminolysis shortening their average lifespan due to capillary column breakage. This irreversible chemical degradation process is prevented by using ammonium buffers at lower pH (pH = 8.5), substituting less nucleophilic or "ammonia-free" alkaline buffers (*e.g.*. borate, ethylamine) or applying more chemically-resistant polytetrafluroroethylene coated fused silica capillaries.

changes in polyimide coating morphology or polymer dissolution were apparent after prolonged exposure to 50 mM ammonium bicarbonate, pH 8.5. Since passive contact of polyimide coated fused-silica capillaries do not reflect typical use in CE, **Figure 1(a)** also depicts two microscopic images near the capillary inlets acquired following 24 h of continuous flushing at high pressure (900 mbar or 90 kPa) using 50 mM ammonium bicarbonate. This study clearly highlights that softening and elongation of the outer polyimide coating beyond the tip of the capillary is evident at pH 10, which did not occur when flushed with the same buffer at pH 8.5. This effect is similar to previous reports of acetonitrile-induced polyimide swelling of fused-silica capillaries in CEC and non-aqueous CE, where the swollen polyimide coating extending beyond the capillary inlet has been attributed to several deleterious effects, such as sample injection variability or sudden current drops during voltage application.[4-6] However, unlike organic solvent-induced swelling effects, exposure to aqueous alkaline ammonia solutions leads to chemical degradation and dissolution of the outer polyimide coating rendering the fused-silica capillary extremely fragile. **Figure 1(b)** also compares changes in tensile strength and resistance to fracture of fused-silica capillaries following application of a repeated bending force (90° angle) performed manually over a 26 day period during solution exposure. This study clearly demonstrates that early stages of polyimide degradation coincided with a higher incidence of capillary breakages and a dramatically shorter average lifespan of only 2 days when placed in 100 mM ammonium bicarbonate, pH 10. In contrast, immersion into 50 mM ammonium bicarbonate, pH 8.5 did not undergo any capillary breakage events during the entire 26 day exposure period. As expected, an intermediate pH of ammonium bicarbonate (pH = 9.3) resulted in a reduced average lifespan of about 17 days, whereas pH adjustment of 50 mM ammonium bicarbonate, pH 10 with lithium hydroxide increased average lifespan by about 8 days relative to the use of ammonium hydroxide since the latter pH adjustment additive increases the effective concentration of ammonia in solution. These observations are consistent with the much lower fraction of reactive ammonia existing at equilibrium in solution at pH 8.5 relative to pH 10 (pK_a = 9.25).

Overall, aqueous buffer solutions containing elevated ammonia concentrations at higher pH conditions accelerated polyimide degradation while increasing the susceptibility of the fused-silica capillary tubing to fracture, which was prevented by using more chemically resistant polytetrafluoroethylene coated capillaries, "ammonia-free" alkaline buffers (*e.g.,* borate) or less nucleophilic primary/secondary amines (*e.g.,* ethylamine, diethylamine, pyrrolidine) as BGE. To the best of our knowledge, only one previous technical report by NASA[43] has examined the chemical stability of polyimide when exposed to caustic solutions (pH 11-14), which indicated only modest swelling effects depending on chemical structure of the specific polyimide, as well as polymer morphology, and curing status. Aqueous degradation of Kapton-based polyimide films under strongly acidic conditions (pH < 2) have also been attributed to hydrolysis of susceptible aromatic imide and amide functional groups that ultimately reduce the tensile strength of the polymer with a corresponding decrease in elongation to failure.[44] However, our studies clearly demonstrate that polyimide coated fused-silica capillaries used in CE and nano-LC are stable in "ammonia-free" alkaline buffer solutions, but are susceptible to aminolysis at pH > 9 despite the widespread use of ammonia containing buffers as volatile electrolytes in ESI-MS. Although the exact composition of the commercial polymer coating is not known, aromatic polyimides containing a heterocylic imide linkage in their repeat unit are typically used in manufacture due to their high thermal stability.[45] **Figure S1** of the Supplemental Information section provides evidence that irreversible dissolution of the fused-silica capillary outer coating results in formation of a brown-orange solution that is consistent with an aromatic polyimide precursor as indicated by ^{1}H-NMR and UV-vis absorbance. Also, MSI-CE-MS using a dilution trend filter with molecular feature extraction[24] revealed the presence of two acidic (anionic) and four basic (cationic) ions as by-products of polyimide degradation as defined by their characteristic *m/z*:RMT under negative and positive ion mode detection, respectively. A PubChem search for these unknown synthetic compounds using their top three ranked molecular formulae revealed no known chemical structures on the database in the case for anions, and a large number of putative candidates and their isomers (> 120) for cations. Thus, a total of six highly unsaturated, aromatic,

nitrogen-rich, and water-soluble ions were identified as novel polyimide aminolysis by-products in solution. The chemical stability of polyimide coated open tubular capillaries examined in this work was derived from a single manufacturer, thus considerable differences are expected in products from other companies using different proprietary polyimide coatings and heat curing processes.

A systematic study was next performed by comparing the performance of five different alkaline BGEs that are compatible for anionic metabolite profiling of pooled 24 h urine samples when using MSI-CE-MS with conventional polyimide coated capillaries, including equimolar concentrations (50 mM) of ammonium bicarbonate (pH = 8.5), ammonium acetate (pH = 8.5), ethylamine (pH = 10.0), diethylamine (pH = 11.0) and pyrrolidine (pH = 11.3). Both caustic solutions containing diethylamine and pyrrolidine produced stable yet low currents (< 15 µA) with a fast EOF resulting in a narrow separation time window and lower peak capacity for separation. Overall, ethylamine was found to generate slightly better sensitivity for detection of anionic urinary metabolites under negative ion mode relative to ammonium acetate as BGE.[34, 35] However, ammonium bicarbonate provided far better performance in terms of improved ion signals, higher separation efficiency and shorter migration times for a diverse array of anionic metabolites in human urine without compromising the integrity of the polyimide outer coating as depicted in **Figure S2** in the Supplemental Information section. Indeed, enhanced sensitivity was acquired using ammonium bicarbonate as BGE as reflected by a 50% increase in total molecular features detected from diluted urine samples relative to ammonium acetate or ethylamine since electrolyte type can differentially impact solute ionization efficiency.[27] In our study, no electrolyte was added to the sheath liquid solution (50:50 MeOH:H$_2$O), which ultimately minimized the impact of ion suppression by bicarbonate in the BGE due to the much higher flow rate of sheath liquid used for spray formation. **Figure 2(a)** depicts a schematic of the serial injection configuration used in MSI-CE-MS that relies on repeated hydrodynamic injection of alternating plugs of urine (five-fold diluted in deionized water) and BGE prior to zonal separation of acidic urinary
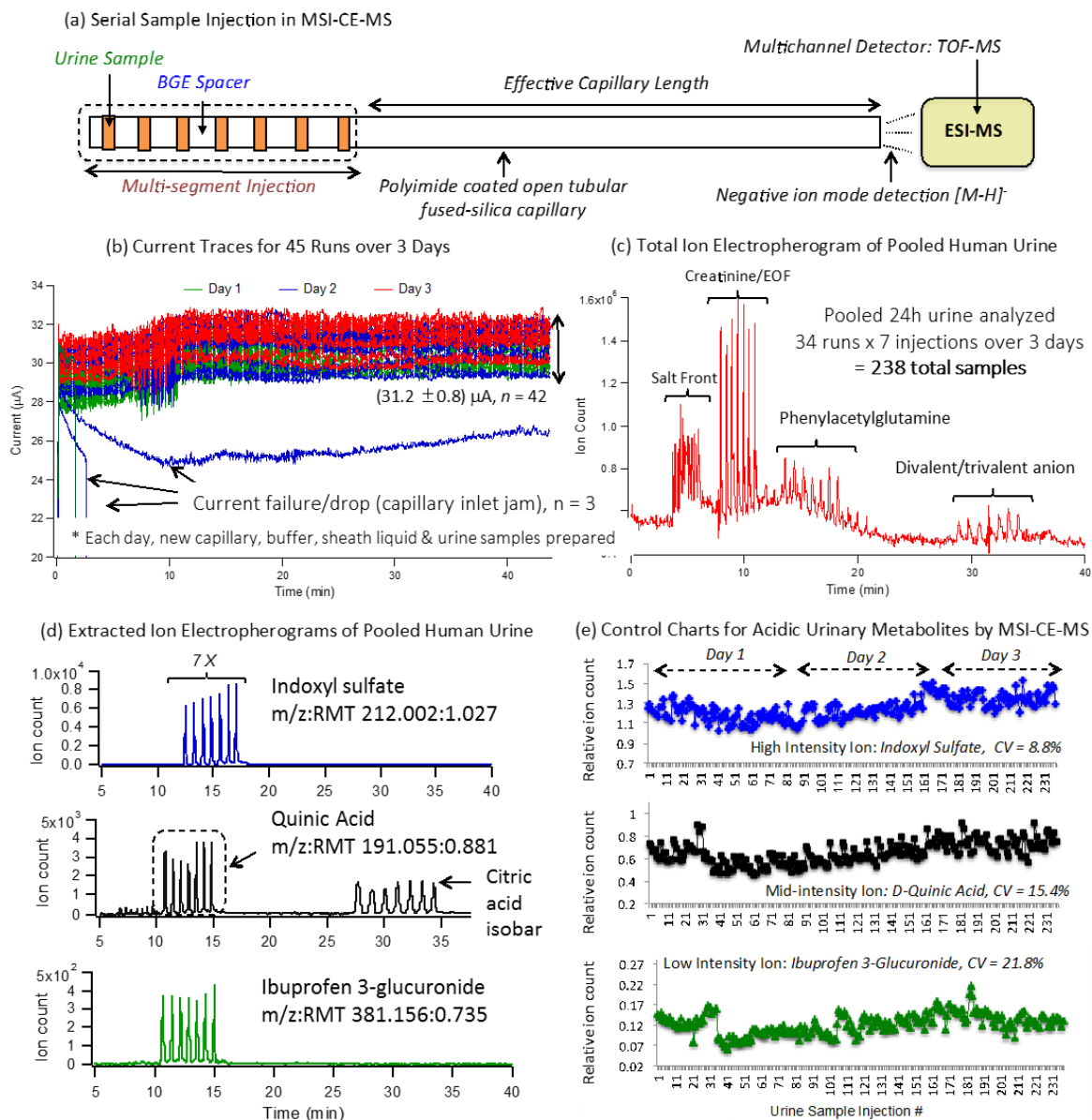
Figure 2.2 (a) Serial injection format used in MSI-CE-MS for multiplexed analysis of anionic metabolites from seven or more urine samples under alkaline conditions with negative ion mode detection. **(b)** An overlay of electrophoretic current traces from 45 runs performed over 3 days of analysis, which highlights excellent reproducibility ($CV = 2.3\%$) with exception of 3 runs that were rejected due to major current drops caused by incidental capillary inlet misalignment during sample loading. **(c)** Representative total ion electropherogram (TIE) of MSI-CE-MS when using a 7 serial injection format when analyzing pooled 24 h human urine samples, which illustrates four distinctive transitions during the separation, including a salt front of inorganic electrolytes (*e.g.,* Na$^+$), neutral compounds co-migrating with the EOF (*e.g.,* creatinine), as well as major urinary metabolites, including phenylacetylglutamine and uric acid. **(d)** Extracted ion electropherograms (EIE) for three urinary metabolites as defined by their characteristic *m/z*:RMT demonstrating excellent resolution of seven serial injections of human urine samples within a single run with isobaric/isomeric resolution. **(e)** Control charts highlighting the long-term analytical performance of MSI-CE-MS based on the integrated peak areas for ions relative to an internal standard (NMS) that differ in their natural abundance following 34 runs or 238 repeated injections of a 5-fold diluted pooled human urine sample over 3 days.

metabolites from each sample plug under steady-state electrophoretic and ionization conditions. Also, **Figure 2(b)** highlights an overlay of current traces from 45 runs performed over 3 days when using MSI-CE-MS, where each run is comprised of a multiplexed analysis of 7 (or more) samples as a way to improve sample throughput without complicated column switching programs, which increases overall productivity of MS-based metabolomics without added infrastructure costs.[24] The isocratic elution conditions of CE together with the homogenous ionization conditions provided by the sheath liquid ensures that the same metabolite entering the ion source offset in time from different samples generates a series of resolved peaks as shown in **Figure S3** of the Supplemental Information. Absolute quantification is also feasible given the high efficiency separation of urine matrix interferences when using a single non-isotope internal standard for correction of between-sample variations in injection volume.[24-26] Monitoring current traces in CE-MS provides an effective way to evaluate system stability and detect systematic error as evident by 3 runs that had a significantly lower current and/or sudden current drops, which originated from misalignment of the inlet capillary during sample loading. Overall, the CE separation current was stable and reproducible after the migration of neutral compounds with the electroosmotic flow (EOF, $\approx$ 11 min) in the majority of runs with an overall CV = 2.6% ($n = 42$) over 3 days of operation despite installing/conditioning a new capillary, as well as preparing a fresh BGE/sheath liquid solution and three technical replicates of diluted urine samples each day. Each morning started with a wipe down of platinum electrodes at inlet and ion source to reduce sample carry-over followed by analysis of a standard mixture of anionic metabolites and two equilibrating runs of pooled urine samples. As a result, 34 runs or 238 repeated sample injections of pooled 24 h human urine were used to evaluate the long-term performance of MSI-CE-MS for routine anionic metabolite profiling in complex biological samples.

Figure 2(c) depicts a representative total ion electropherogram (TIE) of the multiplexed analysis of 7 diluted human urine samples in series by MSI-CE-MS, which highlights that effective desalting of inorganic electrolytes (*e.g.,* Na$^+$ or salt front) and

neutral compounds that co-migrate with the EOF (*e.g.,* creatinine) precedes the migration of major acidic metabolites in human urine, such as phenylacetylglutamine and uric acid. **Figure 2(d)** illustrates extracted ion electropherograms (EIE) for three distinct classes of strongly ionic and weakly acidic urinary metabolites that vary in their signal intensity when using MSI-CE-MS, including indoxyl sulfate, *D*-quinic acid (resolved from the later migrating tricarboxylic acid isobar, citric acid), and ibuprofen 3-glucuronide. For instance, indoxyl sulfate is a uremic toxin associated with vascular dysfunction,[46] whereas *D*-quinic acid is a dietary phytochemical with poorly understood biological functions that is aromatized to hippuric acid by action of gut microflora.[47] High dose intake of the analgesic ibuprofen (excreted in urine as hydroxyibuprofen and its intact glucuronide conjugate) is associated with increased risk of myocardial infarction for subjects with pre-existing cardiovascular disease.[48] Effective sample throughput is increased significantly since seven samples are analyzed within the same run without loss in resolution or information content that maximizes the peak capacity of the separation. Moreover, quality control (QC) samples can be introduced within each run to assess or correct for systematic error over a large batch of runs, whereas signal pattern recognition of time-resolved signals enables unambiguous molecular feature selection for untargeted metabolite profiling that reduces false discoveries.[24-26] Each urinary metabolite is defined by its characteristic mass-to-charge ratio (*i.e.,* accurate mass) and relative migration time as a paired variable (*m/z*:RMT) together with its likely molecular formula as summarized in **Table 1** for 30 different ionic metabolites ranging from small organic acids (*e.g.,* lactic acid), microflora-derived metabolites (*e.g., p*-cresol sulfate), endogenous steroid (*e.g.,* tetrahydroaldosterone 3-glucuronide) or exogenous over-the-counter drug conjugates (*e.g.,* acetaminophen sulfate). In fact, **Figure 2(d)** highlights that later migrating metabolites (*e.g.,* indoxyl sulfate) from samples introduced at the end of the serial injection sequence in MSI-CE-MS elute with sharper peaks despite longer times for diffusion likely as a result of sample self-stacking (*i.e.,* transient isotachophoresis) when analyzing highly saline urine samples.[39] This effect was not attributed to ion suppression or enhancement effects for acidic metabolites migrating after the salt front and EOF as confirmed by monitoring changes in the ion signal

of internal mass calibrants added to the sheath liquid during the time course of the separation as shown in **Figure S4** in the Supplemental Information.

Figure 2(e) shows control charts based on integrated peak area ratios normalized to a single non-isotope internal standard (50 μM NMS) following 238 repeated sample injections (or 34 runs) by MSI-CE-MS over 3 days of analysis while applying a standardized operating protocol. Overall, controls charts indicate good intermediate precision without evidence of systematic error as reflected by coefficient of variance (*CV*) ranging from 8.8% to 21.8% depending on ion abundance. **Table 1** shows that the overall *CV* for 30 different urinary metabolites was about 16% in terms of relative peak area ratios, whereas it was about 0.9% for RMTs. In this case, the IS is primarily used to correct for variations in injection volume between samples in MSI-CE-MS, whereas improved precision is expected if co-migrating deuterated ISs are included for each metabolite in order to correct for changes in spray stability during separation.[49] In most cases, metabolite identification was confirmed after spiking authentic standards in urine when available, whereas some urinary metabolites were tentatively assigned based on their accurate mass and most likely molecular formula with a mass error < 5 ppm in most cases (with exception of low abundance metabolites), as well as characteristic electromigration behaviour in aqueous solution (*i.e.,* predicted RMTs). **Figure S5** and **Table S1** of the Supplemental Information demonstrate that the absolute mobility of anionic metabolites can be well described by two intrinsic physiochemical properties of an ion when using the Hubbard-Onsager equation,[44] namely the effective charge at pH 8.5 (z_{eff}) as derived from its pK_a and the solvent-excluded molecular volume (*V*). Overall, a good correlation ($R^2 = 0.974$) was found between the average RMT of 30 urinary metabolites and their predicted absolute mobility as derived from their chemical structure. Thus, CE offers a complimentary approach to MS for prioritizing putative isobaric candidates for biomarker discovery in metabolomics in cases when mass spectral databases are incomplete while also confirming likely isomeric compounds not readily distinguished by MS/MS.[41]

Table 2.1 Long-term stability for analysis of 30 acidic metabolites analyzed in pooled 24 h human urine over three consecutive days of analysis (34 runs or 238 sample injections) by MSI-CE-MS.

Assignment	*m/z:*RMT	Molecular formula	Mass error (*ppm*)	RMT %CV	RPA %CV
Lactic Acid	89.0244:1.154	$C_3H_6O_3$	6.0	0.7	14.2
3-Hydroxybutyric Acid	103.0401:1.026	$C_4H_8O_3$	1.3	0.1	17.8
Succinic acid	117.0193:1.950	$C_4H_6O_4$	3.1	1.8	20.6
Oxoproline	128.0353:1.020	$C_5H_7NO_3$	2.3	0.1	13.5
Glutaric Acid	131.0350:1.700	$C_5H_8O_4$	1.4	1.5	20.7
Threonic Acid	135.0299:0.993	$C_4H_8O_5$	1.7	0.1	11.6
Tartaric Acid	149.0092:2.019	$C_4H_6O_6$	0.1	1.8	28.1
4-Hydroxyphenylacetic Acid*	151.0401:0.932	$C_8H_8O_3$	0.3	0.4	14.2
2-Aminoadipic Acid*	160.0615:0.898	$C_6H_{11}NO_4$	0.8	0.6	11.7
3-Hydroxymethylglutaric Acid*	161.0455:1.602	$C_6H_{10}O_5$	0.0	1.6	22.3
Uric Acid	167.0201:0.959	$C_5H_4N_4O_3$	4.2	0.3	13.8
2-Isopropylmalic Acid*	175.0601:1.456	$C_7H_{12}O_5$	5.1	1.3	19.2
Hippuric Acid	178.0510:0.905	$C_9H_9NO_3$	1.7	0.5	10.2
Homovanillic Acid	181.0506:0.880	$C_9H_{10}O_4$	2.8	0.8	9.9
4-Pyridoxic Acid*	182.0459:0.936	$C_8H_9NO_4$	1.3	0.4	18.6
p-Cresol Sulfate	187.0071:1.068	$C_7H_8O_4S$	2.8	0.3	9.3
5-Hydroxyindoleacetic Acid	190.0510:0.880	$C_{10}H_9NO_3$	2.5	0.7	13.5
D-Quinic Acid	191.0552:0.881	$C_7H_{12}O_6$	4.7	0.7	15.4
Vanillylmandelic Acid	197.0552:0.876	$C_9H_{10}O_5$	11	0.8	16.4
Indoxyl Sulfate	212.0023:1.027	$C_8H_7NO_4S$	0.6	0.1	8.8
Pantothenic Acid	218.1034:0.816	$C_9H_{17}NO_5$	0.9	1.1	16.8

Ethyl Glucuronide	221.0746:0.828	$C_8H_{14}O_7$	1.4	1.0	24.2
2-Hydroxy Ibuprofen*	221.1183:0.799	$C_{13}H_{18}O_3$	2.8	1.2	14.2
Acetaminophen Sulfate	230.0129:0.921	$C_8H_9NO_5S$	0.3	0.4	10.3
Indolylacryloylglycine*	243.0771:0.868	$C_{13}H_{12}N_2O_3$	0.6	1.0	17.9
Phenylacetylglutamine	263.1037:0.805	$C_{13}H_{16}N_2O_4$	0.6	1.1	17.1
Acetaminophen Glucuronide*	326.0881:0.757	$C_{14}H_{17}NO_8$	0.2	1.5	22.1
Ibuprofen Glucuronide*	381.1555:0.735	$C_{19}H_{26}O_8$	0.6	1.6	21.8
Tetrahydroaldosterone 3-Glucuronide*	539.2498:0.704	$C_{27}H_{40}O_{11}$	1.9	1.8	20.2
Cortolone 3-Glucuronide*	541.2654:0.700	$C_{27}H_{42}O_{11}$	0.1	1.8	20.5

An asterisk (*) indicates putative identification of urinary metabolites based on their accurate mass and characteristic electromigration behavior using the Onsager-Hubbard equation. Other metabolites were confirmed by spiking urine with authentic chemical standards. Naphthalene monosulfonic acid (NMS) was used as a single non-deuterated internal standard for data normalization in terms of relative peak areas (RPA) and relative migration times (RMT). Mass error was calculated from the average of experimentally measured m/z (n = 18) taken from the first sample run of each day and the theoretical monoisotopic mass from molecular formula.

2.5 Conclusion

In summary, this is the first report to describe the deleterious effects of polyimide degradation when using flexible open tubular fused silica capillaries after prolonged exposure to aqueous alkaline ammonia buffers. Irreversible aminolysis of the outer polyimide coating results in polymer dissolution while severely decreasing the tensile strength of the capillary during extended operation with poor long-term stability. Our work demonstrates that MSI-CE-MS offers a robust and high throughput screening platform for characterization of the ionic metabolome,[50] such as a diverse range of acidic metabolites in human urine when operating under optimum conditions to prevent polyimide aminolysis and capillary breakages. Putative identification of metabolites in the absence of authentic standards is facilitated by accurate modeling of the electromigration behaviour *in silico* in conjunction with high resolution, accurate MS notably when MS/MS spectra cannot be

acquired, such as the case for low abundance or weakly ionizable precursor ions. Future work will explore the use of alternative polymer materials that better tolerate strongly alkaline ammonia solutions and organic solvents while overcoming the limited thermal stability of polytetrafluoroethylene coatings. Inter-laboratory validation studies of MSI-CE-MS technology under standardized operating conditions are in progress to facilitate method translation to new users and different biological samples, including round-robin studies to assess the analytical performance relative to other techniques widely used in metabolomics, including ^{1}H-NMR, GC-MS and LC-MS. MSI-CE-MS offers a rapid, low cost yet reliable method for untargeted profiling of polar/ionic metabolites in complex biological samples that maximizes the productivity of MS infrastructure without compromising separation performance.

Acknowledgements

2.6 References
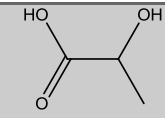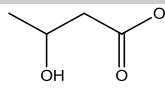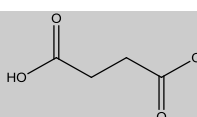
(1)   Dandeneau, R.D.; Zerenner, E.H., HRC & CC 1979, 1, 351-356.
(2)   Dandeneau, R.D.; Zerenner, E.H. LCGC North America 1990, 8, 908-912.
(3)   Griffin S, LCGC North America 2002, 20, 928-938.
(4)   Dittmann, M.M.; Rozing, G.P., Ross, G.; Adam, T.; Unger, K.K. J. Cap. Electrophoresis 1997, 5: 201–212
(5)   Baeuml, F.; Welsch, T., J. Chromatogr. A 2002, 961, 35-44.
(6)   Lindberg, P; Roeraade, J. Electrophoresis 1998, 19, 2445-2446.
(7)   In "Capillary Electrophoresis-Mass Spectrometry (CE-MS): Principles and Applications", Edited by Gerhardus de Jong, August 2016, Wiley-VCH, 368 pages, ISBN: 978-3-527-33924-2.
(8)   Lindenburg, P.W.; Haselberg, R.; Rozing, G.; Ramautar R., Chromatographia 2015, 78, 367-377.
(9)   Dunn, W.B.; Broadhurst, D.; Begley, P.; Zelena, E.; Francis-McIntyre S.; Anderson, N.; Brown, M.; Knowles, J.D.; Halsall, A.; Haselden, J.N.; Nicholls, A.W.; Wilson, I.D.; Kell, D.B.; Goodacre, R. Nat. Protoc. 2011, 6, 1060-1083.
(10)   Ramautar R., Somsen G.W.; de Jong G.J. Electrophoresis 2015 36: 212-224.
(11)   Ramautar R.; Nevedomskaya, E.; Mayboroda, O.A.; Deelder, A.M.; Wilson, I.D.; Gika, H.G.; Theodoridis, G.A.; Somsen, G.W.; de Jong, G.J. Mol Biosyst. 2011, 7, 194-199.
(12)   C. Ibáñez, C. Simó, V. García-Cañas, A. Gómez-Martínez, J.A. Ferragut, A. Cifuentes Electrophoresis, 2012, 33, 2328–2336.
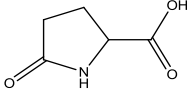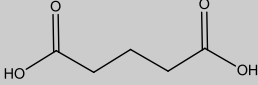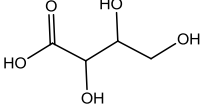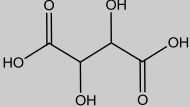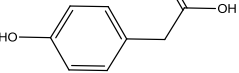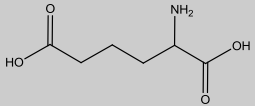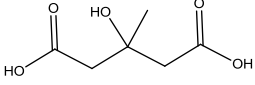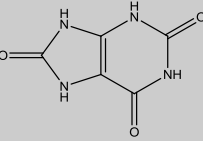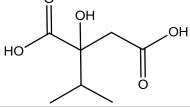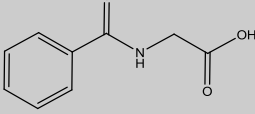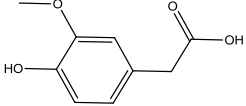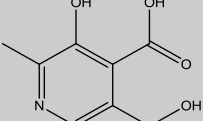
(13) Saric J.; Want, E.J.; Duthaler, U.; Lewis, M.; Keiser, J.; Shockcor, J.P.; Ross, G.A.; Nicholson, J.K.; Holmes, E.; Tavares, M.F.M. Anal. Chem., 2012, 84, 6963–6972.
(14) Kok, M.; Somsen, G.M.; de Jong, G. J. Talanta 2015, 132, 1-7.
(15) Hirayama, A.; Wakayama, M.; Soga, T. TrAC 2014, 61, 215-222.
(16) Mischak H., Vlahou A., Ioannidis J.P. Clin. Biochem. 2013, 46: 432-443.
(17) Sun, L.; Zhu, G.; Yan, X.; Dovichi, N.J. Curr. Opin. Chem. Biol. 2013, 17, 795-800.
(18) Wenz, C.; Barbas, C.; López-Gonzálvez, Á., Garcia, A.; Benavente, F.; Sanz-Nebot, V.; Blanc, T.; Freckleton, G.; Britz-McKibbin, P.; Shanmuganathan, M.; de l'Escaille, F.; Far, J.; Haselberg, R.; Huang, S.; Huhn, C.; Pattky, M.; Michels, D.; Mou, S.; Yang, F.; Neusuess, C.; Tromsdorf, N.; Baidoo, E.E.; Keasling, J.D.; Park, S.S.; J. Sep. Sci. 2015, 38, 3262-3270.
(19) Minami, Y., Kasukawa, T., Kakazu, Y., Iigo, M., Sugimoto, M., Ikeda, S., Yasui, A.; van der Horst, G.T.; Soga, T.; Ueda, H.R. Proc. Natl. Acad. Sci. USA 2009, 106, 9890-9895.
(20) Sugimoto, M.; Wong, D.T.; Hirayama, A.; Soga, T.; Tomita, M. Metabolomics 2010, 6, 78-95.
(21) Ramautar R., Busnel J.M.; Deelder A.M.; Mayboroda O.A., Anal. Chem. 2012, 84, 885-892.
(22) Alberice, J.V.; Amaral, A.F.; Armitage, E.G.; Lorente, J.A.; Algaba F.; Carrilho, E.; Márquez, M.; García, A.; Malats, N.; Barbas, C. J. Chromatogr. A 2013, 1318, 163-170.
(23) Balderas, C.; Rupérez, F.J.; Ibañez, E.; Señorans, J.; Guerrero-Fernández, J.; Casado, I.G.; Gracia-Bouthelier, R.; García, A.; Barbas, C. Electrophoresis 2013, 34, 2882-2890.
(24) Kuehnbaum, N.L.; Kormendi, A.; Britz-McKibbin, P. Anal. Chem. 2013, 85, 10664-10669.
(25) Kuehnbaum, N.L.; Gillen, J.B., Gibala M.J.; Britz-McKibbin, P. Sci. Rep. 2014, 4, 6166.
(26) Kuehnbaum, N.L.; Gillen J.B.; Kormendi ,A.; Lam, K.P.; DiBattista, A.; Gibala, M.J.; Britz-McKibbin P. Electrophoresis 2015, 36, 2226-2236.
(27) Ramautar, R.; Toraño, J.S.; Somsen, G. W.; de Jong, G.J., Electrophoresis 2010, 31. 2319-2327.
(28) Büscher, J.M.; Czernik, D.; Ewald, J. C.; Sauer, U.; Zamboni, N., Anal. Chem. 2009, 81, 2135-2143.
(29) Soga, T.; Igarashi, K.; Ito, C.; Mizobuchi, K.; Zimmermann, H.-P.; Tomita, M., Anal Chem 2009, 81, 6165-6174.
(30) Soga, T.; Ueno, Y.; Naraoka, H.; Ohashi, Y.; Tomita, M.; Nishioka, T., Anal. Chem. 2002, 74. 2233-2239.
(31) Soga, T.; Baran, R.; Suematsu, M.; Ueno, Y.; Ikeda, S.; Sakurakawa, T.; Kakazu, Y.; Ishikawa, T.; Robert, M.; Nishioka, T.; Tomita, M. J. Biol. Chem. 2006, 281, 16768-16776.
(32) A. Hirayama, A.; Kami, K.; Sugimoto, M.; Sugawara, M.; Toki, N.; Onozuka, H.; Kinoshita, T.; Saito, N.; Ochiai, A.; Tomita, M.; Esumi, H.; Soga, T. Cancer Res., 2009, 69, 4918–4925.
(33) Sato, S.; Yanagisawa, S. Metabolomics 2010, 6, 529-540.
(34) Kok, M.G.M.; de Jong, G.J.; Somsen, G.W., Electrophoresis 2011, 32, 3016-3024.
(35) Kok, M.G.M.; Ruijken, M.A.; Swann, J.; Wilson, I.; Somsen, G.; Jong, G., Anal. Bioanal. Chem. 2013, 405. 2585-2594.
(36) Rütters, H.; Möhring, T.; Rullkötter, J.; Griep-Ramingm, J.; Metzger, J.O. Rapid Commun. Mass Spectrom. 2000,14, 122-123.
(37) Gulersonmez, M.C.; Lock, S.; Hankemeier, T., Ramautar, R. Electrophoresis 2016, 37, 1007-1014.
(38) Bouatra, S; Aziat, F; Mandal, R.; Guo A.C; Wilson, M.R.; Knox, C.; Bjorndahl, T.C.;
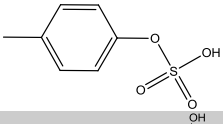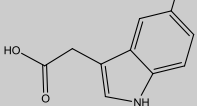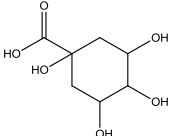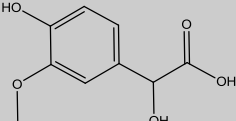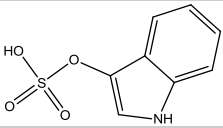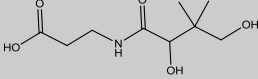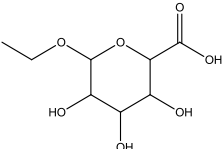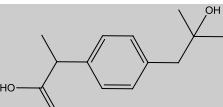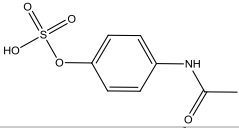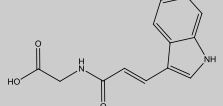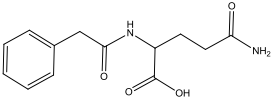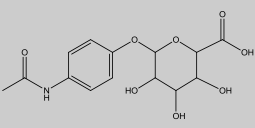
Krishnamurthy, R.; Saleem, F.; Liu, P.; Dame, Z.T.; Poelzer, J.; Huynh, J.; Yallou, F.S.; Psychogios, N.; Dong, E.; Bogumil, R.; Roehring, C.; Wishart, D.S. PLoS One 2013, 8, e73076.

(39)   Nori de Macedo, A.N.; Teo, K.; Mente, A.; McQueen, M.J., Zeidler, J.; Poirier, P.; Lear, S.A., Wielgosz, A.; Britz-McKibbin, P. Anal. Chem. 2014, 86, 10010-10015.

(40)   Hubbard, J.B.; Onsager L., J. Chem. Phys. 1977, 67, 4850-4857.

(41)   Lee, R.; Ptolemy, A.S.; Niewczas, L.; Britz-McKibbin, P. Anal. Chem. 2007, 79, 403–415.

(42)   Kuehnbaum, N.L.; Britz-McKibbin, P., Anal. Chem. 2011, 83, 8063-8068.

(43)   Croall, C.I.; Clair, T.L.S. The Mechanical Stability of Polyimide Films at High pH; NASA Technical Memorandum 102726; October 1990.

(44)   Delasi, R.; Russell, J. J. Appl. Polymer Sci. 1971, 15, 2965-2974.

(45)   Chang, W.-Y.; Chen, S.-H.; Yang, C.H.; Chuang, C.-N.; Wang, C.-K. J. Polymer Res. 2015, 22, 38.

(46)   Niwa, T. J. Ren. Nutr. 2010, 20, S2-S6.

(47)   Pero, R.W.; Lund, H.; Leanderson, T. Phytochem. Res. 2009, 23, 335-346.

(48)   Kearney, P.M.; Baigent, C.; Godwin, J.; Hal,l H.; Emberson J.R.; Patrono C. BMJ 2006, 332, 1302-1308.

(49)   Chalcraft, K.; Lee, R., Mills, C.; Britz-McKibbin, P. Anal. Chem. 2009, 81, 2506–2515.

(50)   Kuehnbaum, N. L.; Britz-McKibbin, P. Chem. Rev. 2013, 113, 2437-2468.
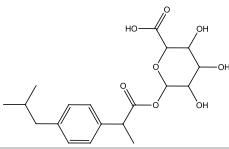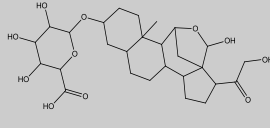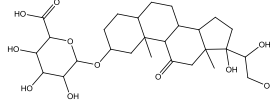
2.7 Supplemental Information

Table S2.1 *In silico* modeling of the relative migration time (RMT) of 30 different anionic metabolites in urine based on two fundamental properties of an ion using the Hubbard-Onsager equation.

ID	m/z: RMT	V (Å³)	pKₐ	z_{eff}	RMT (pred)	% error	Chemical Structure
Lactic Acid	89.024: 1.154	71.8	3.78	-1.0	1.13	-2.48	
3-Hydroxybutyric Acid	103.040: 1.026	88.9	4.44	-1.0	1.04	1.34	
Succinic acid	117.019: 1.950	88.2	4.16 5.61	-2.0	1.98	1.71	

Name	m/z: RT						Structure
Oxoproline	128.035: 1.020	98.4	3.61	-1.0	1.01	-1.47	
Glutaric Acid	131.035: 1.700	105.7	4.31 5.41	-2.0	1.81	6.30	
Threonic Acid	135.030: 0.993	102.2	3.40	-1.0	0.99	-0.04	
Tartaric Acid	149.009: 2.019	99.5	2.98 4.34	-2.0	1.86	-7.74	
Hydroxyphenyl acetic Acid*	151.040: 0.932	118.3	4.00	-1.0	0.95	1.77	
2-Aminoadipic Acid	160.062: 0.898	135.3	2.14 4.21 9.77	-1.1	0.95	5.35	
3-Hydroxymethyl-glutaric Acid	161.046: 1.602	129.8	3.68 4.44	-2.0	1.65	2.92	
Uric Acid	167.020: 0.959	90.6	7.61	-1.0	1.03	7.64	
2-Isopropylmalic Acid	175.060: 1.456	147.5	3.63 5.57	-2.0	1.57	7.48	
Hippuric Acid	178.051: 0.905	135.4	3.59	-1.0	0.91	0.74	
Homovanillic Acid	181.051: 0.880	143.7	3.74	-1.0	0.90	1.89	
4-Pyridoxic Acid	182.046: 0.936	134.0	2.55	-1.0	0.92	-2.29	

p-Cresol Sulfate	187.007: 1.068	109.6	10.36	-1.0	0.97	-9.09	
5-Hydroxy indoleacetic Acid	190.051: 0.880	142.2	10.30	-1.0	0.90	2.18	
D-Quinic Acid	191.055: 0.881	152.1	3.46	-1.0	0.88	0.20	
Vanillylmandelic Acid	197.055: 0.876	150.8	3.11	-1.0	0.89	1.02	
Indoxyl Sulfate	212.002: 1.027	121.6	8.38	-1.0	0.94	-8.40	
Pantothenic Acid	218.103: 0.816	198.3	4.35	-1.0	0.82	0.92	
Ethyl Glucuronide	221.075: 0.828	176.3	3.45	-1.0	0.85	2.49	
Hydroxy Ibuprofen	221.118: 0.799	155.1	4.63	-1.0	0.88	9.89	
Acetaminophen Sulfate	230.013: 0.921	179.5	9.46	-1.0	0.85	-8.29	
Indolyl acryloylglycine	243.077: 0.868	155.4	3.95	-1.0	0.88	1.10	
Phenylacetyl glutamine	263.104: 0.805	226.7	3.90	-1.0	0.80	-1.01	
Acetaminophen Glucuronide	326.088: 0.757	246.7	3.17	-1.0	0.78	3.16	

Ibuprofen Glucuronide	381.156: 0.735	340.8	3.41	-1.0	0.73	-1.28	
Tetrahydro aldosterone 3-Glucuronide	539.250: 0.704	481.9	3.47	-1.0	0.67	-4.20	
Cortolone 3-Glucuronide	541.265: 0.700	489.4	3.56	-1.0	0.67	-3.96	
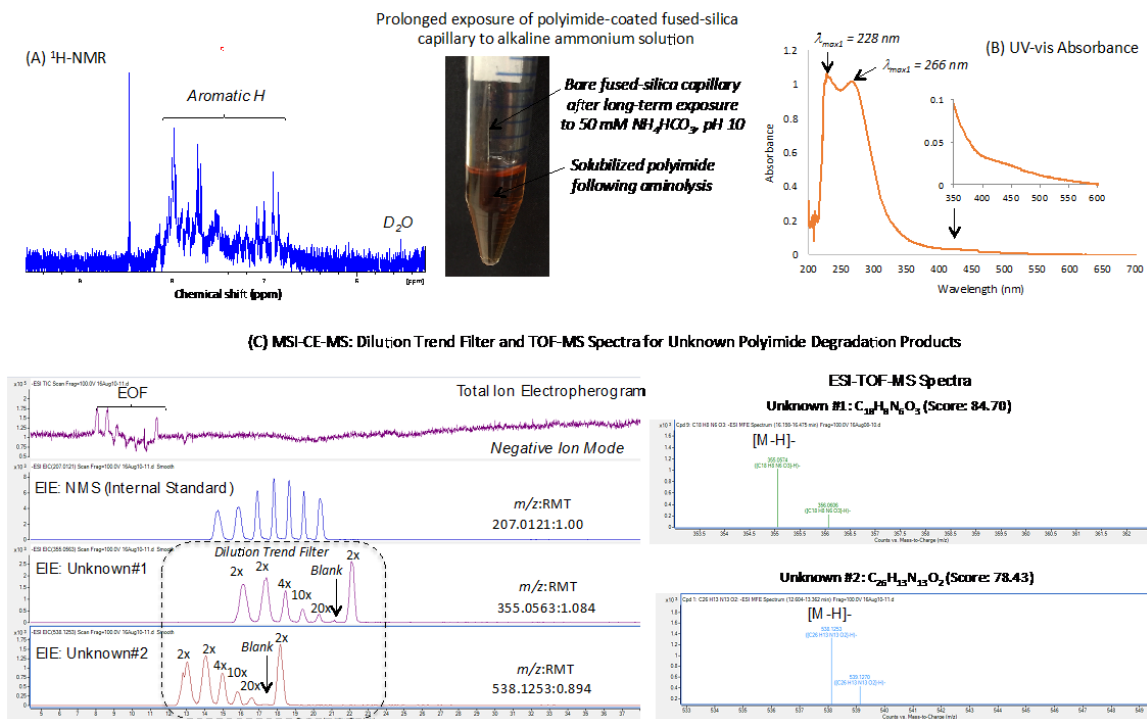
Figure S2.1 Characterization of aminolysis by-products of the outer polyimide coating of fused-silica capillaries following complete dissolution of polymer into solution after prolonged exposure to 50 mM ammonium bicarbonate, pH = 10 at room temperature. **(a)** ^{1}H-NMR spectra of the solubilized polyimide product following evaporation and reconstitution in D_2O shows a series of distinctive set of aromatic proton resonances (6.7-8.2 ppm). **(b)** Blank-corrected UV-vis absorbance spectra of 5-fold diluted orange-brown polymer solution showing two distinctive absorbance bands indicative of a highly unsaturated/aromatic chromophore with weak visible absorbance in blue spectrum. IR absorbance spectra of polyimide solution also confirmed that lack of distinctive aromatic imides stretches at 1780 cm^{-1} [$C=O$, sym.], 1720 cm^{-1} [$C=O$, asym.] and 1380 cm^{-1} [$C-N$] for the degraded polymer in solution following aminolysis (data not shown). **(c)** Chemical analysis of polymer degradation products in solution using MSI-CE-MS with a dilution trend filter under negative and ion mode conditions following use of molecular feature extractor. Two unknown anions [M-H$^-$] were identified based on their characteristic m/z:RMT together with their corresponding top-ranked molecular formula. A Pubchem search using molecular formula generated no known chemical structures in the database. A similar strategy was also performed by MSI-CE-MS under strongly acidic conditions (1 M formic acid, pH 1.8) under positive ion mode detection as described in our previous study [1], which revealed four unknown cations [MH$^+$] as defined by their m/z:RMT (4-chlorotyrosine as IS), including 116.9848:0.481, 211.0975:0797, 212.0829:0818, and 316.0939:0856. Thus, a total of six discrete unsaturated, nitrogen-rich, aromatic ions were identified as polyimide aminolysis by-products in solution.
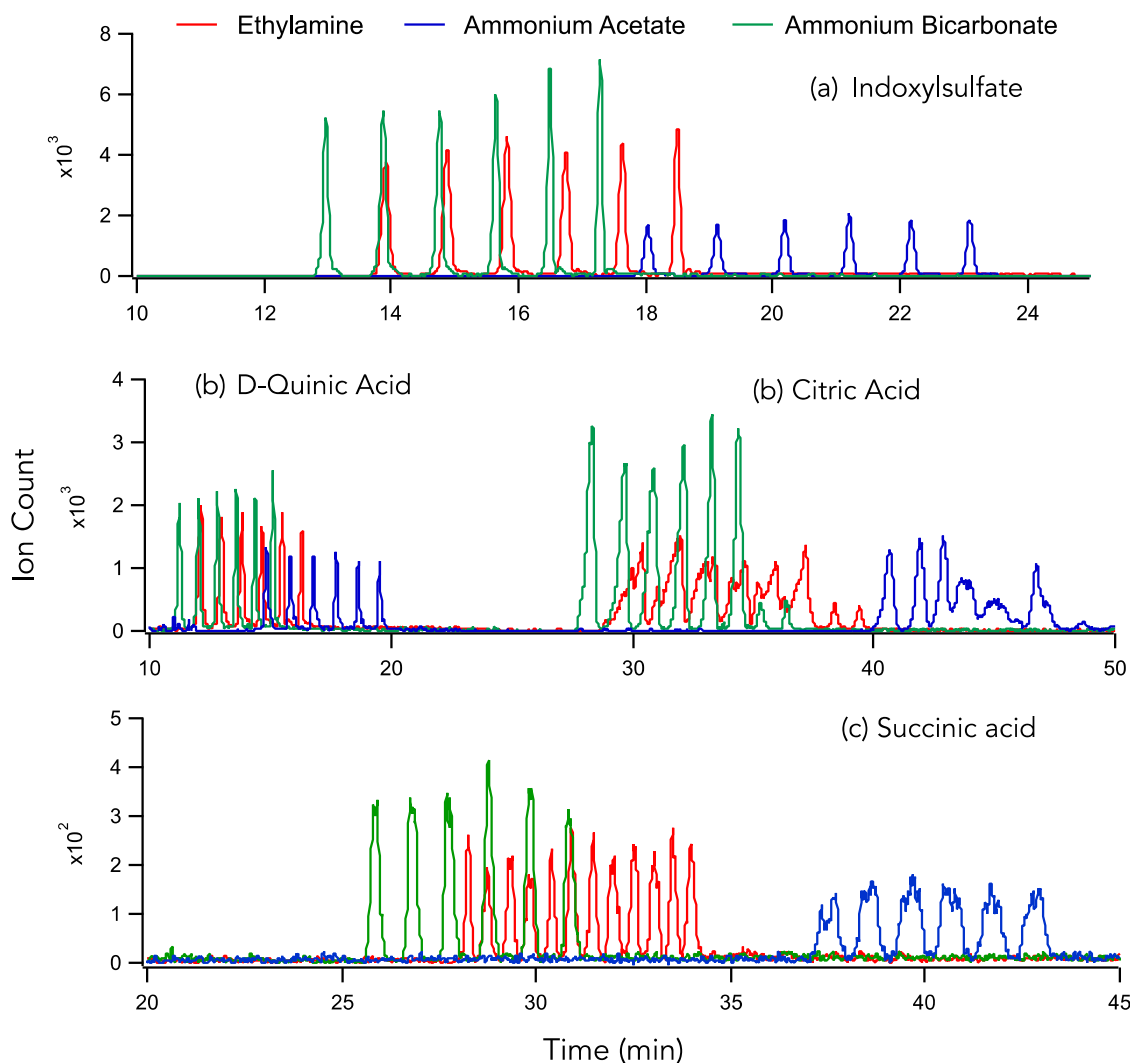
Figure S2.2 Overlay of extracted ion electropherograms comparing the separation efficiency and ionization responses for representative acidic urinary metabolites as a function of background electrolyte (BGE) composition, including 50 mM ethylamine (pH = 10), 50 mM ammonium acetate (pH = 8.5) and 50 mM ammonium bicarbonate (pH = 8.5). Overall, ammonium bicarbonate was found to provide shorter migration times, sharper peaks with greater ion responses for diverse classes of strong acids (*e.g.,* indoxylsulfate), weakly acidic organic acids (*e.g., D*-quinic acid) in human urine samples, including tricarboxylic acids without deleterious peak splitting or band dispersion, such as citric acid and succinic acid. As a result, ammonium bicarbonate was used as BGE for subsequent long-term stability studies when analyzing human urine samples by MSI-CE-MS at a pH = 8.5 that prevents polyimide degradation. Note that *D*-quinic acid and citric acid are isobaric anions [M-H-] readily resolved by CE when extracted within 10 ppm. Also, in many cases later migrating samples migrate as sharper peaks without increased band dispersion as a result of diffusion due to likely sample self-stacking (*i.e.,* transient isotachophoresis) during zonal electropherotic separation when analyzing highly saline urine samples.

(a) Serial Sample Injection

Urine Sample

BGE Spacer

Multichannel Detector: TOF-MS

Effective Capillary Length

ESI-MS

Multi-segment Injection

Polyimide coated bare fused-silica capillary with compatible alkaline BGE

(b) Zonal Separation *Anionic Metabolites with Characteristic Steady-state Mobilities*

+

ESI-MS

EOF

Normal polarity configuration

Negative ion mode detection [M-H]⁻

(c) Extracted Ion Electropherograms

Ibuprofen Glucuronide, m/z 381.1555

o-Cresol Sulfate, m/z 187.0071

Lactic Acid, m/z 89.0244

Figure S2.3 Schematic depicting the operation of MSI-CE-MS as a multiplexed separation method for high throughput profiling of anionic metabolites in human urine while using a compatible alkaline buffer to prevent polyimide degradation (50 mM ammonium bicarbonate, pH = 8.5) with negative ion mode detection. **(a)** Serial injection of seven or more sample segments is performed alternating with BGE as spacer plugs that overall fills approximately one third of the total capillary length; **(b)** Voltage application with zonal separation of acidic urinary metabolites from within each sample plug, where anionic metabolites and their isomers are separated based on differences in their electrophoretic mobility (*i.e.,* charge density) that migrate counter to the EOF while using a bare fused-silica capillary under normal polarity configuration; **(c)** Extracted ion electropherograms for three representative acidic metabolites as defined by their characteristic *m/z*:RMT highlighting the high peak capacity and separation efficiency of MSI-CE-MS for resolving complex and highly saline biological samples, such as human urine.
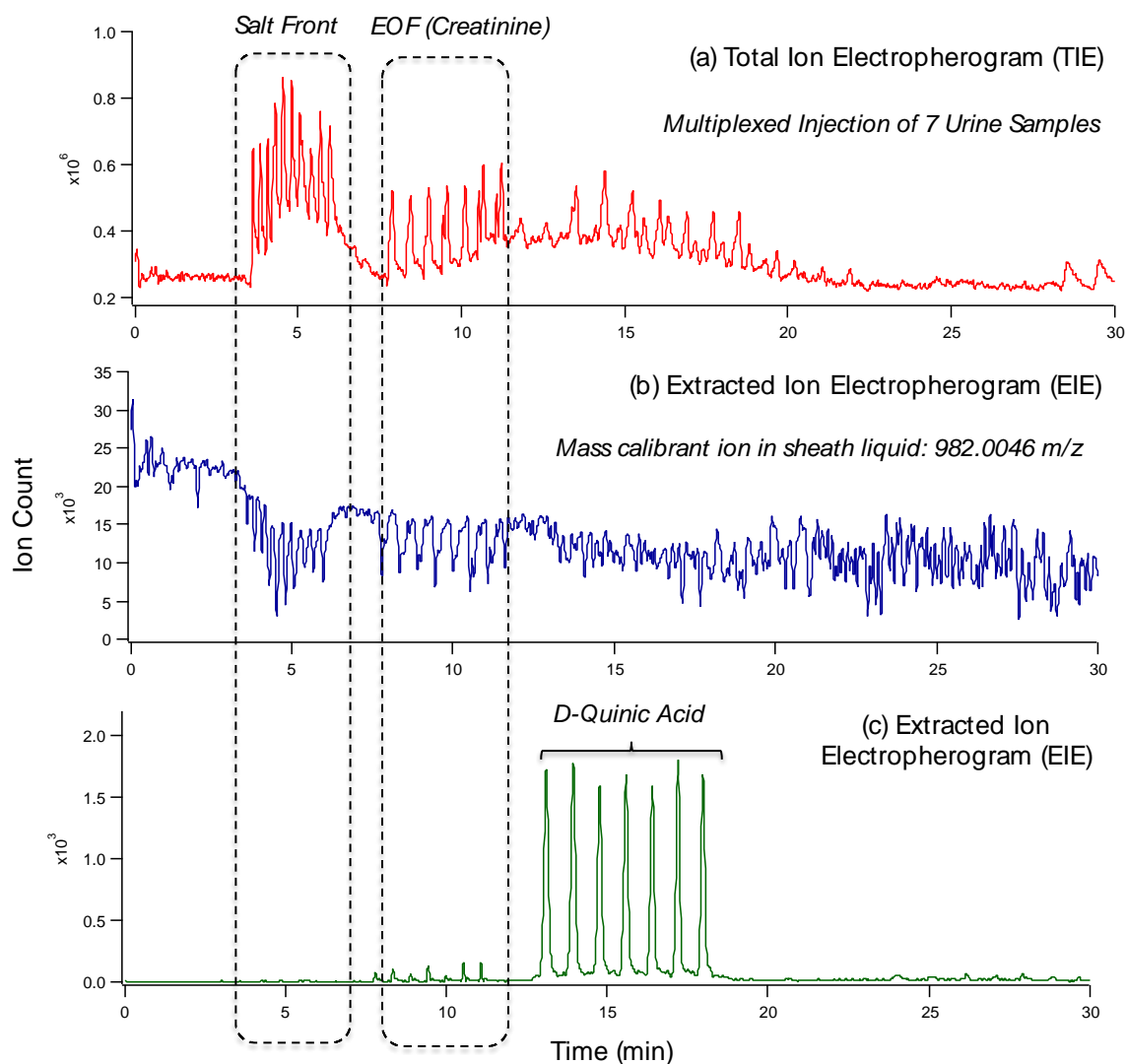
Figure S2.4 Spectral overlay for the multiplexed separation of 7 human urine samples using MSI-CE-MS when using 50 mM ammonium bicarbonate, pH 8.5 as BGE with polyimide coated fused silica capillaries, which highlights **(a)** total ion electropherogram and two early migrating zones associated with salt/electrolyte front (Na+) and EOF (co-migration of creatinine), **(b)** extracted ion electropherogram for the mass calibrant ion (982.0046 *m/z*) that is added into sheath liquid solution and continuously infused into ion source to improve mass accuracy while also enabling the monitoring of time regions susceptible to ion suppression (*i.e.,* attenuated signal) during separation, and **(c)** extracted ion electrophergram for *D*-quinic acid based on its characteristic *m/z*:RMT (191.055:0.881). In the latter case, there is no evidence of significant ion suppression in time region near migration of *D*-quinic acid and other acidic urinary metabolites due to the effective desalting and high separation efficiency of CE.

The plot shows RMT (Experimental) on the x-axis and Absolute Mobility (Predicted) on the y-axis, with the fitted equation:

$$y = 4.25\,E\text{-}5\,x + 1.17\,E\text{-}5$$
$$R^2 = 0.974$$

and the Hubbard-Onsager equation:

$$\mu_0 = \frac{a z_o}{V^c + (b z_o^2 / V)}$$
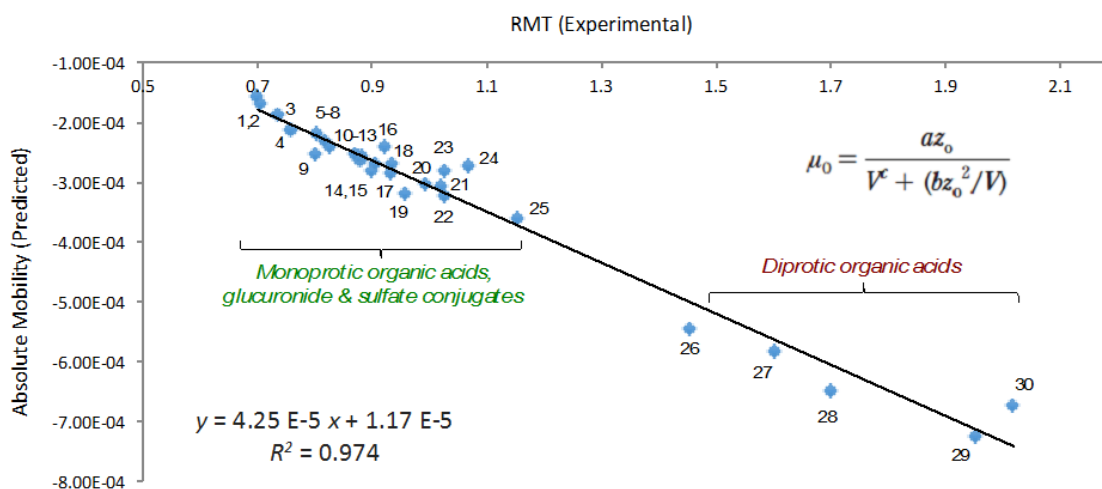
Figure S2.5 Modeling of the absolute electrophoretic mobility of weakly acidic urinary metabolites based on two intrinsic solute properties using the Hubbard-Onsager equation, namely pK_a or effective charge (z_o at pH = 8.5) and molecular volume (MV) summarized in **Table S1**. Overall, there is a good correlation in measured relative migration times (RMTs) with predicted absolute mobility for acidic metabolites (n= 30) that supports identification of putative ions based on their accurate mass or elemental formula in cases when commercial standards or MS/MS spectra are unavailable. Urinary metabolites are 1: cortolone 3-glucuronide, 2: tetrahydroaldosterone 3-glucuronide, 3: ibuprofen 3-glucuronide, 4: acetominophen glucuronide, 5: hydroxyibuprofen, 6: pantothenic acid, 7: ethylglucuronide, 8: phenylacetylglutamine, 9: indoylacroylglutamine, 10: vanillylmandelic acid, 11: *D*-quinic acid, 12: 5-hydroxyindoleacetic acid, 13: homovannilic acid, 14: 2-aminoadipic acid, 15: hippuric acid, 16: acetominophen sulfate, 17: hydroxyphenylacetic acid, 18: 4-pyridoxic acid, 19: uric acid, 20: threonic acid, 21: oxoproline, 22: 3-hydroxybutyric acid, 23: indoxylsulfate, 24: cresol sulfate, 25: lactic acid, 26: 2-isopropylmalic acid , 27: 3-hydroxymethylglutaric acid , 28: glutaric acid, 29: succinic acid, 30: tartaric acid. Experimental migration times were normalized to a single internal standard in MSI-CE-MS, namely naphthalene monosulfonic acid (NMS) in order to derive improved precision (*CV* < 1%) based on solute relative migration times (RMT).

Chapter III

Metabolomics Reveals Elevated Urinary Excretion of Collagen Degradation Products in Irritable Bowel Syndrome Patients

Thesis chapter is derived from a manuscript in preparation for submission to Metabolomics, where I am the first author and my research supervisor is the corresponding author:

Mai Yamamoto, Ines Maria Pinto Sanchez, Premysl Bercik, and Philip Britz-McKibbin
Metabolomics **2018** (submitted).

M.Y. conducted sample preparation and instrumental analysis, data processing and interpretation for all IBS and control urine samples. M.Y. also wrote an initial manuscript draft used for publication. Other co-authors provided feedback on the manuscript draft, and were involved with patient recruitment and sample collection.

3.1 Abstract

Introduction Irritable bowel syndrome (IBS), the most commonly diagnosed functional gastrointestinal (GI) disorders in developed countries, is characterized by chronic abdominal pain, and altered bowel habits.

Objectives Accurate and timely diagnosis is challenging as it relies on symptoms and an evolving set of exclusion criteria to distinguish it from other related GI disorders reflecting a complex etiology that remains poorly understood. Herein, nontargeted metabolite profiling of repeat urine specimens collected from a cohort of IBS patients (n=42) was compared to healthy controls (n=20) to gain insights into the underlying pathophysiology.

Methods An integrated data workflow for characterization of the urine metabolome with stringent quality control was developed to filter out redundant, background and spurious signals using multisegment injection-capillary electrophoresis-mass spectrometry. Complementary univariate and multivariate statistical methods were then used to rank differentially excreted urinary metabolites after normalization to osmolality.

Results Our work revealed consistently elevated urinary metabolites in repeat samples collected from IBS patients at two different time points ($q < 0.05$ after FDR adjustment), which were associated with elevated collagen degradation and intestinal mucosal turn-over likely due to low-grade inflammation, including a series of hydroxylysine metabolites (*O*-glycosylgalactosyl-hydroxylysine, *O*-galactosyl-hydroxylysine, lysine), mannopyranosyl-tryptophan, imidazole propionate, glutamine, serine, ornithine, dimethylglycine and dimethylguanosine.

Conclusion Our results from this pilot study provide intriguing new mechanistic insights into the pathology of IBS that warrant further validation as it offers a convenient way to monitor disease progression and treatment responses to novel dietary interventions while avoiding invasive blood sampling, colonoscopy and/or tissue biopsies.

3.2 Introduction

Irritable bowel syndrome (IBS) is characterized by chronically recurring abdominal pain, altered bowel habits and poor quality of life for which no anatomical abnormalities are evident following colonoscopy imaging and histopathology (Drossman 2016). Symptoms can be triggered by many factors, including diet and environmental exposures (*e.g.,* caffeinated beverages, alcohol), menstrual period, infection, immune activation and/or psychosocial stress (Enck et al. 2016). It is the most common disorder diagnosed by gastroenterologists, affecting about 11 % of the world population, however there are currently no objective biochemical tests that can accurately diagnose IBS (Lacy et al. 2016). Instead, diagnosis is based on symptoms (currently Rome IV criteria) together with exclusion of other similar conditions, such as coeliac disease and inflammatory bowel disease (Lacy 2016). Patients are diagnosed with three main sub-types of IBS, which include diarrhea-predominant (IBS-D), constipation-predominant (IBS-C) and alteration of both (mixed; IBS-M) based on frequency of bowel movements and stool consistency (Enck et al. 2016). In addition to the physical symptoms, patients often suffer from anxiety and depression, which complicates understanding of the exact pathological mechanisms of IBS (Azpiroz et al. 2007). IBS is considered to be a disorder of gut-brain communication, in which altered intestinal motility and permeability, visceral hypersensitivity, low-grade inflammation and changes in gut microbiome are implicated in its pathophysiology (Simrén et al. 2013). Dietary triggers appear to play a major role in symptom generation, and several trials suggest that diets low in fermentable oligosaccharides, disaccharides and monosaccharides and polyols (FODMAPs) (Halmos et al. 2014) may ameliorate the symptoms of IBS by reducing immune activation and modulating gut microbiota associated in gas production (McIntosh et al. 2017). However, these mechanistic insights have not been translated into effective screening tools or diagnostic tests for IBS within a clinical setting.

Metabolomics offers a hypothesis-generating approach for characterizing the molecular phenotype of an organism when using nuclear magnetic resonance (NMR) and

increasingly high efficiency separations coupled to mass spectrometry (MS) as required for comprehensive profiling of metabolites in complex biological samples. This process allows for the discovery of new biomarkers that enables pre-symptomatic diagnosis and early intervention (DiBattista et al. 2017a), while also providing new mechanistic insights into the pathophysiology of human disorders of unknown etiology (Johnson et al. 2016). Most metabolomic studies on IBS patients have applied a targeted approach to establish a link between gut microbiota and central nervous system (CNS) observed in animal models in support of a "microbiome-gut-brain" axis disorder (Eisenstein 2016; Kennedy et al. 2014). For example, endogenous spore-forming bacteria induce host serotonin production when germ-free mice were colonized with endogenous strains from normal mice (Yano et al. 2015). Serotonin is an important neurotransmitter that is implicated in mood disorders as reflected in the use of selective serotonin reuptake inhibitors as commonly used anti-depressants (Cleare et al. 2015); it also plays a key role in regulating intestinal motility by inducing relaxation and contraction of intestinal muscle tissue that is relevant to IBS (Kendig and Grider 2015). For these reasons, metabolomic studies have focused on the analysis of aberrant tryptophan/serotonin metabolism in IBS, as well as host-microbial co-metabolites, such as indoles and bile acids. For instance, Keszthelyi *et al.* (Keszthelyi et al. 2013) reported significantly lower mucosal and higher plasma concentrations of serotonin and kynurenic acid in IBS patients as compared to healthy controls when using LC-MS, which implicated disturbances in intestinal homeostasis. However, their results were derived from a small cohort of IBS patients, which has not been consistently replicated in other larger studies (Thijssen et al. 2016) that reported lower plasma 5-hydroxyindole acetic acid, a host-microbial co-metabolite of tryptophan in IBS patients. Differences in intestinal microbial community and their functions between IBS patients and healthy controls were also implicated in the analysis of volatile microbial metabolites, such as short-chain fatty acids, esters and other hydrocarbons in stool (Ahmed et al. 2013) and breath condensates (Baranska et al. 2016) when using GC-MS with head-space sample collection. Also, clear distinctions were reported among IBS-D and IBS-C sub-types from healthy controls based on bile acid profiles from fecal extracts when using LC-MS in conjunction with colonic

transit time and intestinal permeability as measured by urinary lactulose and mannitol tests, respectively (Camilleri et al. 2014). Since bile acids are largely re-absorbed at terminal ileum with unabsorbed fractions biotransformed to secondary bile acids by intestinal bacteria (Hofmann and Hagey 2008), bile acid malabsorption has been reported in up to 50% of IBS-D patients (Slattery et al. 2015).

Interestingly, the urine metabolome (Bouatra et al. 2013) of IBS patients has not been characterized to date, which offers a far more convenient biofluid in clinical investigations reflecting the dynamic interplay of host metabolism and gut microbiota activities (Obrenovich et al. 2017). In addition, urine also provides unique insights into systemic processes relevant to human health, such as biomarkers of habitual diet (Playdon et al. 2016) and lifelong environmental exposures (Rattray et al. 2018) that may not be readily detectable in blood, which also avoids the technical challenges in the collection and handling of stool specimens. In this work, we conducted nontargeted metabolic phenotyping of a cohort of IBS patients as compared to healthy controls using multisegment injection-capillary electrophoresis–mass spectrometry (MSI-CE-MS) (Kuehnbaum et al. 2013) as a high throughput platform optimal for characterization of the urine metabolome as it is comprised mostly of ionic/polar metabolites (M. Yamamoto et al. 2016). Stringent quality control (QC) was integrated when using a multiplexed metabolomics data workflow in order to identify unknown metabolites consistently elevated in repeat urine samples collected from IBS patients as compared to controls when using univariate/multivariate statistical methods in conjunction with high resolution MS and MS/MS for structural elucidation. For the first time, we report that IBS patients have higher urinary excretion of glycosylated hydroxylysine metabolites, as well as glycated tryptophan, and methylated guanosine/glycine derivatives indicative of accelerated collagen degradation and intestinal epithelial turn-over respectively, which is consistent with a low-level inflammation phenotype in IBS.

3.3 Experimental section

3.3.1 Chemicals and reagents. All chemicals were purchased from Sigma Aldrich (St. Louis, MO, USA) unless otherwise stated. Stock solutions for all internals standards and metabolite standards, as well as background electrolyte (BGE) and sheath liquid solutions were prepared in deionized water from a Barnstead EASYpure® II LF system (Dubuque, IA, USA). Ultra-grade LC-MS solvents (acetonitrile, methanol and water) purchased from Caledon Laboratories Ltd. (Georgetown, ON, Canada) were used to prepare sheath liquid solution for spray formation and BGE for CE separations. 2-(α-D-mannopyranosyl)-*L*-tryptophan standard was obtained from Toronto Research Chemicals Inc. (Toronto, ON, Canada), 5-(galactosylhydroxy)-*L*-lysine was purchased from Cayman Chemical (Ann Arbor MI, USA), and cholic acid was obtained from AVOCADO Research Chemicals Ltd. (London, UK). All stock solutions for chemical standards were stored at 4 °C.

3.3.2 Study design and urine sample collection. A total of 42 IBS patients were recruited based on Rome III Criteria, and their psychiatric comorbidity assessed by Hospital Anxiety and Depression Scale (mild to moderate anxiety and depression) (Snaith and Zigmond 1994) at the Farncombe Family Digestive Health Institute at McMaster University. Patients were initially recruited for a placebo-controlled 6-week probiotics intervention study. All patients were diagnosed with either IBS-D or IBS-M based on their predominate bowel symptoms. Patients with any concurrent organic GI pathology other than benign polyps, diverticulosis, hemorrhoids, and melanosis coli were excluded as well as the history of antibiotics use three months prior to the first sample collection. Patients who were pregnant or were breastfeeding were also excluded from the study. All participants in this study signed the informed consent prior to the participation into study-related activities. This cohort was characterized by median age of 44 years old (range from 19 to 73 years old) and 43% male (*n*=18), including four self-reported tobacco smokers. Non-IBS healthy controls (*n*=20) were also recruited based on the following inclusion criteria, age 19-70, fit and healthy, IBS negative on the IBS symptomatic questionnaire, no history of chronic digestive disorders, not pregnant or lactating, had taken no antibiotics and medication for chronic conditions, such as high blood pressure, or diabetes. Median age of this group was

28 years old (range from 19 to 43 years old) with 45% male (*n*=9) with one occasional tobacco smoker self-identified. Single-spot random urine samples were collected on two separate occasions during clinic visits, 6 weeks apart, from the majority of IBS patients, whereas only a single-spot urine sample was collected from non-IBS healthy controls. Urine samples were collected in the morning, and not necessarily the first morning void urine and were stored at -80 °C prior to analysis. Due to patient drop-out and missing samples, urine samples were collected from 42 IBS patients for the initial collection (IBS-1, *n*=42) with 34 matching IBS patients for the second collection (IBS-2, *n*=34) periods. All urine samples were prepared by thawing slowly on ice, centrifuging at 400 *g* for 5 min to sediment any particulate matter, and then a 10 μL aliquot of the supernatant was diluted 5-fold with de-ionized water containing two internal standards, namely 3-chloro-*L*-tyrosine (Cl-Tyr) and 2-naphthalenesulfonate (NMS) with a final concentration of 50 μM. A pooled quality control (QC) sample was also prepared by mixing equal aliquots of all urine samples from IBS patients (*n* = 76), which was used for monitoring technical precision and long-term system drift in ESI-MS.

3.3.3 Osmolality and creatinine measurements for urine normalization. Osmolality based on freezing point depression of urine was measured using an Advanced Micro-Osmometer 3300 (Advanced Instruments Inc., Norwood, MA, USA) in order to correct for hydration status. Measurement for all urine samples were performed within one day and a Clinitrol 290 Reference solution was used every ten urine samples to ensure the accuracy of measurements. The reference solution for QC had a mean value of 290 mOsm and a range from 288 to 292 mOsm (actual osmolality = 290 mOsm) with excellent technical precision as reflected by a coefficient of variation (CV) < 1.0 %. Urinary creatinine concentrations were also measured for all urine samples in this study using MSI-CE-MS. A seven-point calibration curve for creatinine was acquired in triplicate from 0 to 1000 μM with excellent linearity ($R^2 = 0.997$), where the integrated peak area was normalized to an internal standard (50 μM Cl-Tyr).

3.3.4 MSI-CE-MS and metabolomics data workflow. Non-targeted metabolite analysis was performed on an Agilent G7100A CE system (Agilent Technologies Inc., Mississauga, ON, Canada) equipped with a coaxial sheath liquid Jetstream electrospray ion source with heated nitrogen gas that were coupled to an Agilent 6230 TOF-MS or Agilent 6550 QTOF-MS system. Separations were performed using an uncoated fused-silica capillary (Polymicro Technologies Inc., Phoenix, AZ, USA) with an inner diameter of 50 μm, outer diameter of 360 μm and total length of 110 cm when using an applied voltage of 30 kV at 25 °C. The background electrolyte (BGE) consisted of 1.0 M formic acid with 15% *v/v* acetonitrile (pH = 1.80) for cationic metabolites and 50 mM ammonium bicarbonate (pH = 8.50) for anionic metabolites when using positive and negative ion mode detection, respectively. Fused-silica capillaries were initially conditioned by flushing at high pressure (900 mbar or 90 kPa) for 15 min each in the order of MeOH, 0.1 M NaOH, deionized water and 1.0 M formic acid. In between runs, the capillary was flushed using deionized H_2O for 3 min followed by equilibration with BGE for 15 min. Instrumental conditions for MSI-CE-MS are similar to previous studies when using a TOF-MS system with full-scan data acquisition, which was applied for screening of urinary metabolites over a mass range of *m/z* 50-1700 as described elsewhere.(Macedo et al. 2017; DiBattista et al. 2017a; DiBattista et al. 2017b) Briefly, a seven-plug serial injection sequence was used in MSI-CE-MS at 100 mbar based on alternating hydrodynamic injections of a sample (5 s) followed by background electrolyte (BGE) spacer (40 s) that was repeated six more times with different samples within a method program. An applied pressure gradient was also applied during the voltage application for anionic metabolites under negative ion mode detection, up to 72 mbar or 7.2 kPa over 38 min (2 mbar or 0.2 kPa increase every 2-3 min). The sheath liquid for the electrospray ionization (ESI) interface consisted of 60% *v/v* MeOH with 0.1% *v/v* formic acid and 50% *v/v* MeOH which were used for positive and negative ion mode detection, respectively. The TOF-MS was operated with a mass range of *m/z* 50-1700 and an acquisition rate of 500 ms/spectrum. The ESI conditions were Vcap = 2000 V, nozzle voltage = 2000 V, nebulizer gas = 10 psi, sheath gas = 3.5 L/min at 195 °C, drying gas = 8

L/min and a temperature = 300 °C for positive and negative ion modes. Furthermore, the MS voltage settings were fragmentor = 120 V, skimmer = 65V and Oct1 RF = 750 V.

An accelerated data workflow for characterization of the urine metabolome of IBS patients was applied when using multiplexed separations by MSI-CE-MS. First, a dilution trend filter (Kuehnbaum et al. 2013) of a pooled urine sample from IBS patients was used to filter out spurious and irreproducible molecular features, as well as background or redundant signals generated in ESI-MS (*e.g.,* isotopes, adducts, in-source fragments) from the original feature list generated by Mass Hunter Molecular Feature Extractor (MFE). Briefly, when a blank sample and a serially diluted QC sample were co-injected within the same run, temporal pattern recognition was used to reject molecular features that were highly variable based on triplicate injections at the same dilution (*CV* > 30%), and present in the blank sample (*i.e.,* background ions) from the data matrix (Kuehnbaum et al. 2013; Kuehnbaum et al. 2015; DiBattista et al. 2017a). Subsequently, a curated list of representative yet authentic urinary metabolites were extracted in profile mode with 10 ppm mass error allowance using Mass Hunter Workstation Software (Qualitative Analysis, version B.6.00, Agilent Technologies Inc.) and annotated based on their accurate mass for the protonated $[M+H]^+$ or deprotonated $[M-H]^-$ molecular ion (*m/z*), relative migration time (RMT) and ionization mode [(*p*) = ESI +, (*n*) = ESI -]. Peak smoothing was performed using a quadratic/cubic Savitzky-Golay algorithm (15 points) prior to peak integration. This rigorous approach to data filtering resulted in a total of 173 urinary metabolites measured consistently in a pooled urine sample, including known metabolites and unknown ions denoted by their *m/z*:RMT under positive or negative ion mode. Following peak integration, peak area and migration time data were exported to Excel (Microsoft Office, WA, USA) and relative peak area (RPA) and RMT for all metabolites were calculated relative to internal standards included in the same injection position (*i.e.,* Cl-Tyr for cationic and NMS for anionic metabolites) in order to correct for differences in on-capillary injection volume between-samples and electroosomotic flow (EOF) between-runs, respectively. Individual urine samples were then analyzed by MSI-CE-MS by injecting six urine samples and one

QC within a randomized order and injection position, however IBS patient and healthy control samples were analyzed as separate batches approximately four months apart. However, inclusion of the same QC sample in every MSI-CE-MS run served as an effective way of evaluate long-term technical precision and apply batch adjustment correction if warranted (Macedo et al. 2017). The final curated urine metabolome data matrix was comprised of 138 metabolites after rejecting ions with poor technical precision (CV > 40%), as well as ions that were detected in less than 75% of total urine samples in each group (IBS-1, IBS-2, non-IBS), whereas non-detected values were replaced by the lowest relative ion response measured for a given metabolite divided by two. Importantly, several exogenous drugs and their metabolites in urine (*e.g.,* carboxy-ibuprofen, acetaminophen sulfate, *trans*-hydroxycotinine) were excluded from the final urine metabolome data matrix given potential confounding caused by differences in prescription medications and over-the-counter drugs between patients, including between IBS and non-IBS groups.

3.3.5 Identification of unknown urinary metabolites of significance. All urinary metabolites were initially annotated based on their *m/z*:RMT within a given ionization mode, and subsequently assigned a most likely molecular formula and calculated mass error (ppm) when using high resolution TOF-MS (**Table S3.1** of Supporting Information). A single injection CE-MS format was then performed using a Q-TOF system for acquisition of MS/MS spectra as required for the identification of unknown urinary metabolites of clinical significance. Unknown metabolites in urine were identified by two complementary methods: 1) utilization of a list of known metabolites in urine (H. Yamamoto and Sasaki 2017), and 2) *de novo* structural elucidation using collisional-induced dissociation (CID) experiments for acquisition of MS/MS spectra of precursor ions at different collisional energies using a Q-TOF system under positive and negative ion mode conditions. The ESI conditions were Vcap 3500V, nozzle voltage 2000 V, nozzle voltage 2000 V, nebulizer gas = 8 psi, sheath gas = 3.5 L/min at 200 °C, drying gas 16 L/min at 200 °C for positive and negative ion mode. Both auto MS/MS and targeted MS/MS modes were used with a cycle time of 3.1 s with CID experiments performed at 0, 20 and 40V for positive ion mode, and

10, 30 and 40 V for negative ion mode. MS/MS spectra acquired by QTOF-MS for unknown urinary metabolites were then compared to experimental (deposited in public database) and *in silico* MS/MS spectra, including Human Metabolome Database (HMDB; http://hmdb.ca), MassBank (http://www.massbank.jp/), and MetFrag (http://msbi.ipb-halle.de/MetFrag/) (Ruttkies et al. 2016). Finally, MS/MS spectra of chemical standards were also acquired if available under the same instrumental conditions for unambiguous identification based on co-migration in CE and MS/MS spectral matching based on their characteristic diagnostic fragment ions and neutral losses, as well as relative peak intensities.

3.3.6 Statistical data analysis. R program (v. 3.4.3) was used for Spearman rank correlation analysis and box-whisker plot generation. False discovery rate (FDR) adjusted *p*-values (*q*-value) was also calculated also in R using package "*q*-value" (v. 2.12.0) (Storey 2003). SPSS Statistics (IBM Analytics) was used for assessment of normality (Shapiro-Wilk test; $\alpha = 0.05$), homogeneity of variance (Levine's test of equality of error variance; $\alpha = 0.05$), and age-adjusted pair-wise comparison of IBS patients and non-IBS healthy controls by ANCOVA, which is based on a linear regression model for age-adjustment (Böhning et al. 2016). A *log*-transformation of creatinine or osmolality-normalized urine metabolome data was then performed prior to univariate statistical analysis as the majority (> 70 %) of untransformed metabolite data violated normal data distribution and homogeneity of variance. Unsupervised data exploration using principal component analysis (PCA) and supervised orthogonal partial least squares-discriminate analysis (OPLS-DA) were also performed using Metaboanalyst 4.0 (Chong et al. 2018) after missing value imputation (replacing with the smallest value measured for the metabolite divided by 2) prior to a generalized *log*-transformation and autoscaling.

3.4 Results
3.4.1 IBS and control cohort and metabolomics workflow using MSI-CE-MS
Urine samples from 42 IBS patients (IBS-1) diagnosed with Rome III criteria and 20 healthy controls without a history of chronic GI disorders were included in this study.

Repeat single-spot random urine samples were also collected from 34 (of 42) IBS patients (IBS-2) during a second follow-up clinical visit. In all groups, there were more females recruited than males reflecting a higher incidence of IBS among women in the population (Meleine and Matricon 2014) with a small fraction ($< 10\%$) of self-identified tobacco smokers or heavy alcohol consumers (**Table 3.1**). Significant differences were observed between IBS-1 and non-IBS/healthy controls in terms of age ($p = 2.77$ E-3) and the Hospital Anxiety and Depression Scale for Depression (HADS-D) scores ($p = 1.30$ E-4) when using a Mann-Whitney U-test. Similarly, IBS-2 and healthy controls differed in age ($p = 3.18$ E-3) and HADS-D ($p = 8.45$ E-3), which highlights a consistent association of depression that is prevalent among all IBS sub-types (Changhyun et al. 2017). Additionally, IBS patients were characterized by a large number of prescribed and over-the-counter medications, including various low-dose antidepressants, anxiolytics, cardiovascular medications, antidiarrheals and vitamins/supplements. In contrast, the only medications taken by the non-IBS/control group were over-the-counter medications (*e.g.,* ibuprofen) and birth control pills.

Comprehensive characterization of the urine metabolome from IBS patients and controls by MSI-CE-MS resulted in a curated list of 138 unique metabolites (65 cationic, 73 anionic urinary metabolites) that were consistently detected in a majority of urine samples in this study ($> 75\%$ frequency with a CV for QCs $< 40\%$). All urinary metabolites were annotated based on their *m/z*:RMT in either positive or negative ion mode after rejection of spurious, background and redundant ions that constitute the majority ($> 90\%$) of signals in ESI-MS for complex biological samples when using nontargeted metabolomics (Mahieu and Patti 2017). In our case, a dilution trend filter in MSI-CE-MS takes advantage of signal pattern recognition to encode mass spectral information temporally within a separation to authenticate metabolites measured consistently in six independent urine samples (*i.e.,* pooled urine as QC) and a blank sample analyzed together within the same run (**Figure S3.1**).

Table 3.1. IBS and control group demographics. Age and anxiety/depression test scores are shown in mean values ± standard deviations. Medications taken by more than two patients are also listed.

	IBS-1	IBS-2	Non-IBS
Subject (n)	42	34	20
Age (mean ± 1s)*	44 ± 16	46 ± 16	30 ± 7
Sex (male:female)	18:24	16:18	9:11
IBS subtype (D:M:UD)[a]	19:17:6	17:17	NA
HADS-A	9.6 ± 3	6.7 ± 4	5.4 ± 4
HADS-D*	6.2 ± 4	4.4 ± 3	1.8 ± 2
Current smokers	4	3	1
Alcohol consumption [b] (> 10/week)	2	2	0
Medication (n)			
M1	5	4	0
M2	5	4	0
M3	4	4	0
M4	6	5	2
M5	5	4	3
M6	5	7	0
M7	9	7	0
M8	5	5	0
M9	2	1	0
M10	1	0	0

[a] *IBS subtypes: D = diarrhea predominant, M = mix of diarrhea and constipation, UD = undefined*
[b] *One drink considered as equivalent to one pint of beer*
* Significant differences (p < 0.05) in age and HADS-D scores between IBS-1/IBS-2 patients and non-IBS controls based on Mann-Whitney U test
Abbreviations: HADS-A: Hospital anxiety and depression score – anxiety; HADS-D: Hospital anxiety and depression score – depression; M1: health supplements (e.g. vitamins), M2: Low-dose antidepressant, anxyolitics, M3: Antidiarrheals, M4: Non-steroidal anti-inflammatory drugs (NSAIDs), M5 Birth control pills, M6: Proton pump inhibitors, M7: Cardiovascular medications (e.g. anticoagulants), M8: Hormone medications (e.g. thyroid), M9: Others (e.g. bronchodilators, Gravol), M10: Antibiotics

This multiplexed separation platform also provides greater data fidelity based on injection of a QC within every run as required for robust correction to instrumental drift notably when performing independent batches of runs over time (Brunius et al. 2016). As shown in **Table S3.1** in the Supplemental Information, the majority of annotated urinary metabolites were derived from amino acids, modified amino acids, organic acids, acidic sugars, purines

and pyrimidines, as well as various classes of secondary metabolites as their intact sulfate and/or glucuronide conjugates.

Confidence in chemical identities are categorized based on four levels according to reporting standards recommended by the Metabolomics Standards Initiative (Sumner et al. 2007). Compounds with identification level 1 were confidently identified with authentic chemical standards by spiking into a pooled urine sample and confirming peak co-migration based on its accurate mass (*m/z*:RMT) given potential isobaric/isomeric interferences; the majority of amino acids and organic acids were identified by this approach. In addition to spiking and confirming co-migration with standards, MS/MS spectral matching was also performed for urinary metabolites significantly associated with IBS to ensure unambiguous metabolite identification based on their diagnostic fragment ions and neutral losses measured with high mass accuracy (< 5 ppm). Other compounds were tentatively identified as confidence level 2 based on matching their MS/MS spectra with reference spectra available in public databases (*e.g.,* HMDB, MassBank) when using collisional energy of 10, 20, and 40 V. When no matching spectra were available, *in-silico* fragmentation (MetFrag) was used to increase likelihood for a correct putative assignment, which were assigned as a level 3. The remaining urinary metabolites annotated based on their *m/z*:RMT and most likely molecular formula were denoted as level 4 and listed with unknown structures, which did not have any matches with existing compounds after searching MS or MS/MS databases. We also detected a number of metabolites closely associated with intake of habitual/recent dietary intake, including *D*-quinic acid (*m/z*:RMT, 191.0552:0.889) and trigonelline (*m/z*:RMT 138.0550:0.861) that originate from plant and coffee intake (Pero et al. 2009; Rothwell et al. 2014), ethyl glucuronide (*m/z*:RMT, 221.0746:0.838), a known metabolite of ethanol (Dahl et al. 2002), and *trans*-hydroxycotinine (*m/z*:RMT, 193.0973:0.724), a long-lived nicotine metabolite of smoking or tobacco exposure (Massadeh et al. 2009). Remarkably high levels of these urinary metabolites were excreted by only a sub-set of IBS subjects especially those who were reported to drink alcohol or smoke regularly. Similarly, metabolites of over-the-counter

drugs, such as acetaminophen sulfate (*m/z*:RMT, 230.0129:0.930) and ibuprofen glucuronide (*m/z*:RMT, 381.1555:0.750), were found only in individuals who reported to take these medications in a sub-set of both IBS and control groups (**Figure S3.2** of the Supporting Information). Since these largely exogenous compounds were measured in less than 75% of total urine samples, thus they were excluded from the final metabolomics data matrix for subsequent statistical analysis based on our conservative selection criteria.

3.4.2 Urine normalization for adjustment to hydration status.

Since the absolute concentration of urinary metabolites can vary as much as 20-fold depending on the hydration status of an individual, reliable urine normalization methods are crucial in metabolomics especially when relying on single-point random urine collections (Bockenhauer and Aitkenhead 2011). In our study, creatinine (mM) and osmolality (mOsm/kg) from all urine samples showed a good linear correlation ($R^2 = 0.572$, $p = 2.20$ E-16, $n = 93$) as shown **Figure 3.1a** after exclusion of three outlier samples with high creatinine concentrations. However, there was still significant unaccounted variability in correcting for urine hydration status when relying on a single parameter (creatinine) as compared to major solutes/electrolytes that influence freezing point depression (osmolality). Nevertheless, the urine metabolome from all IBS patients and controls in this study ($n = 96$) showed similar trends after normalization by either osmolality or creatinine with lower biological variance (mean CV = 92-93%) as compared to unnormalized urine data (mean CV = 112%). Also, a 2D scores plot from PCA clearly shows tighter clusters for both normalized sets of urine data along the PC1 axis, which represents about 93% of total variance, indicating that urine normalization successfully removed some variations caused by hydration status differences between subjects. Moreover, there is evidence of good overall technical precision and stability of the method as reflected by tight clustering of QC samples ($n = 96$) within the center of unnormalized urine data from individual IBS and non-IBS subjects with a mean CV = 15% that did not require batch correction.

(A) Correlation Plot: Osmolality Vs. Creatinine

$y = 0.0345x + 0.515$
$R^2 = 0.572$
$n = 93$

[Creatinine] (mM)

Osmolality (mOsm/kg)

(B) PCA 2D Scores Plot: Impact of Urine Normalization

138 urinary metabolites, log-transformed/autoscaled

Legend:
- Crt
- Osm
- QC_raw
- raw

Pooled QC
Mean CV = 15 %

PC2 (0.7%)

Osmolality
Mean CV = 92%

Creatinine
Mean CV = 93%

Unnormalized Urine
Mean CV = 112%

PC1 (92.8%)
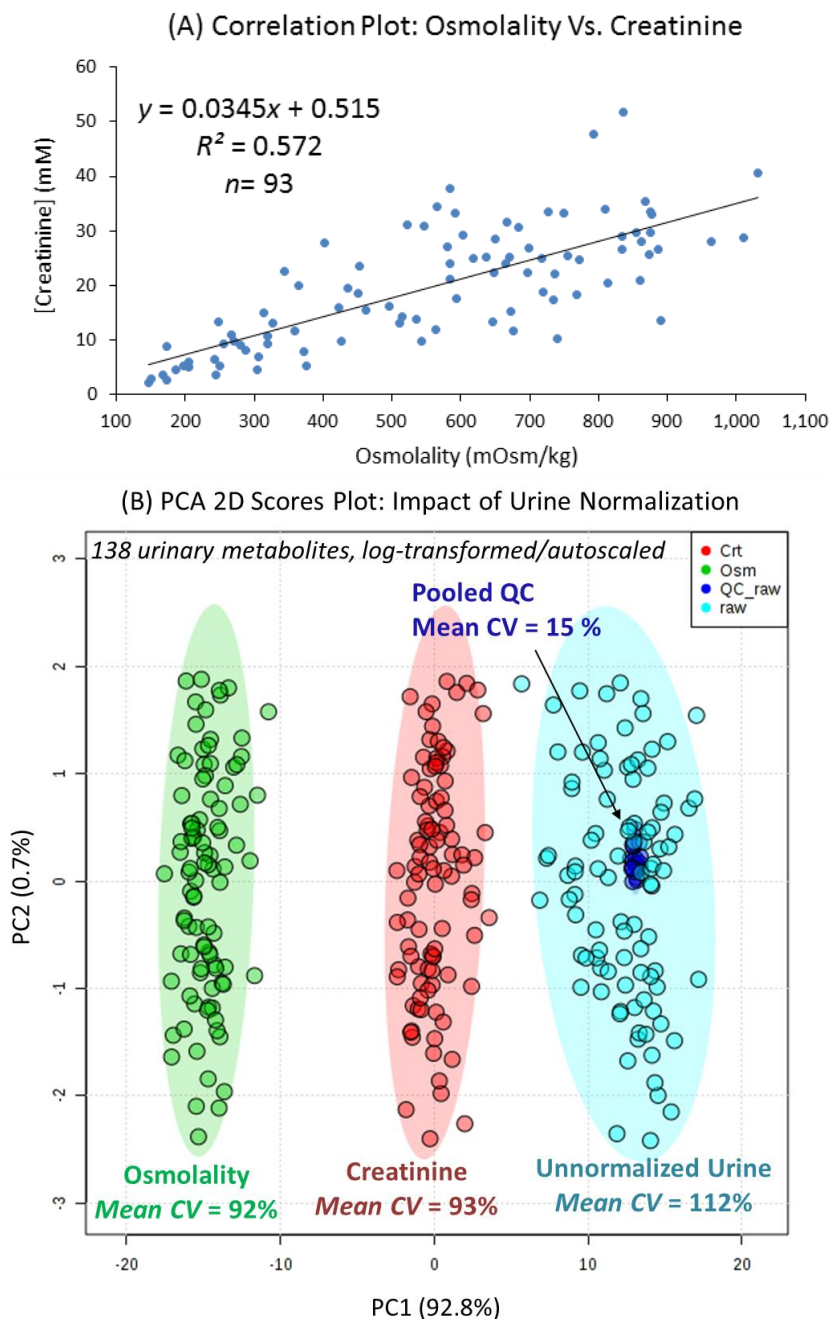
Figure 3.1 (a) A correlation plot of urinary creatinine (mM) versus urine osmolality (mOsm/kg) measured from a cohort of IBS patients and healthy controls when using linear least-squares regression. **(b)** A PCA 2D scores plot comprising 138 urinary metabolites measured consistently in the majority of urine samples following a generalized *log*-transformation and autoscaling, including unnormalized data (*n*=93) as compared to same data normalized with creatinine or osmolality. Excellent technical precision is indicated by the tight cluster of repeated analysis of pooled urine samples (QC, n=96) relative to the larger biological variance between individual urine samples.

3.4.3 Urinary metabolites differentiating between IBS and non-IBS controls

This study involved a retrospective case-control study design involving repeat urine samples collected from IBS patients and sex-matched healthy controls. Univariate and multivariate statistical analysis was performed to rank metabolites that were differentially excreted between IBS patients and non-IBS controls from urine samples collected at the first time point (IBS-1, n=42). The consistency of these findings was then confirmed by comparing the results with samples collected from a majority of IBS patients with an independent set of single-spot urine samples collected at a second time point (IBS-2, n=34). Additionally, correlation between top-ranked urinary metabolites and age was investigated to account for potential confounding between groups (Mann-Whitney U-test, $p < 0.05$) while also applying a Benjamini-Hochberg false discovery rate (FDR) adjustment to correct for multiple hypothesis testing. **Figure 3.2a, b** shows a supervised multivariate statistical model using OPLS-DA based on 138 osmolality-normalized urinary metabolites measured in IBS patients at first time point (IBS-1, n=42) and healthy controls (n=20) following a generalized *log*-transformation and autoscaling. The model performance was modest overall following a leave-one-out-at-a-time cross-validation ($R^2 = 0.389$: $p < 0.001$, $Q^2 = 0.252$: $p = 0.011$), which is indicative of large unexplained biological variance since IBS disease status reflected only 6.3% of total variance. Nevertheless, a S-plot of this OPLS-DA model shows that a subset of urinary metabolites (p[1] > 1.5) were significantly associated with IBS, including lysine, trimethyllysine, deoxycytidine, and two unknown cationic metabolites [M+H]$^+$ with m/z:RMT of and 325.1601:0.720 and 367.1605:0.720, which were subsequently identified as *O*-galactosyl-5-hydroxylysine, (Gal-OHLys, **Figure 3.2c-e**) and *C*-mannopyranosyl-tryptophan (Man-Trp, **Figure S3.3** of the Supporting Information), respectively. Both metabolites undergo CID to generate characteristic fragment ions in their MS/MS spectra based on their amino acid core structure (*i.e.,* OH-Lys and Trp) with neutral losses corresponding to a monosaccharide (m/z 162).
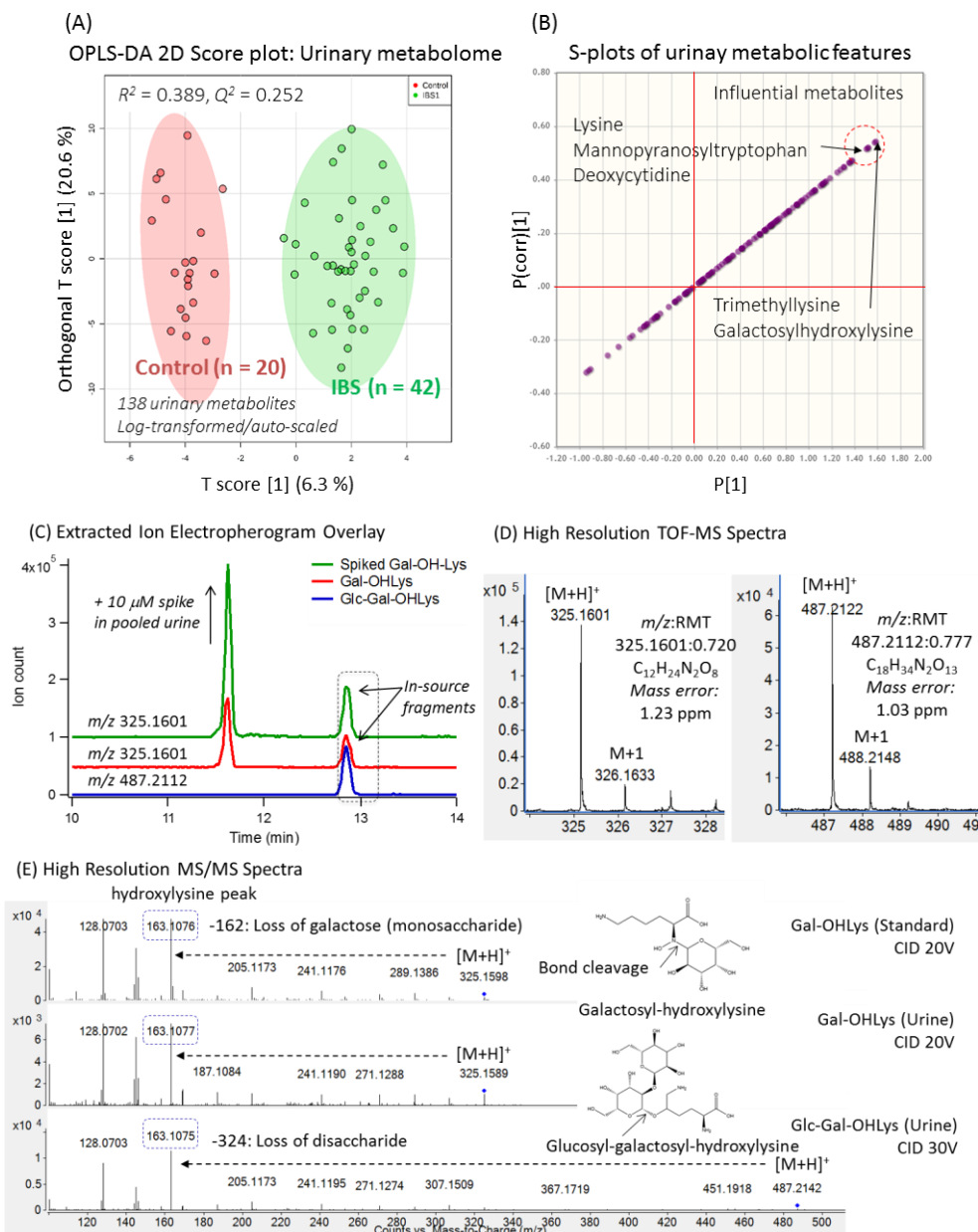
Figure 3.2 (a) OPLS-DA 2D score plot of 138 urinary metabolites measured in diagnosed IBS patients (IBS-1, n=42) and healthy/non-IBS controls (*n*=20). A generalized *log*-transformation and autoscaling of osmolality-corrected urine metabolome data was applied prior to statistical analysis. **(b)** S-plot for ranking of urinary metabolites associated with IBS status. Data points representing the top-ranked metabolites with the highest coefficients (p[1] > 1.5) are circled with red dotted line, which were excreted at higher concentrations in IBS patients as compared to controls. **(c)** An extracted ion electropherogram overlay of a single injection of a pooled urine sample and the same sample spiked with an authentic chemical standard containing 10 μM *O*-galactosyl-hydroxylysine. **(d)** High resolution MS spectrum showing the protonated molecular ions [M+H]$^+$ and isotope patterns for two unknown metabolites that were used to generate most likely molecular formulae with low mass error (< 2 ppm). **(e)** High resolution MS/MS spectra of an authentic standard as compared to two unknown glycated hydroxylysine metabolites in pooled urine. Arrows indicate proposed bond cleavage sites for each precursor ion. Abbreviations include Gal-OHLys: *O*-galactosyl-hydroxylysine; Glc-Gal-OHLys: *O*-glucosyl-galactosyl-hydroxylysine.

The exact stereochemistry of the amino acid glycans were tentatively deduced based on a review of urine metabolites reported in the literature, as well as known biochemical pathways involving OH-Lys and other glycated amino acids used in collagen biosynthesis (Krane et al. 1977; Bateman et al. 1984; Cunningham et al. 1967).

Age-adjusted comparisons between IBS and healthy controls was then performed using ANCOVA with osmolality-normalized urine metabolome data (*log*-transformed) in order to identify top-ranked urinary metabolites (FDR, $q < 0.05$) associated with IBS as summarized in **Table 3.2**. In this case, ten metabolites satisfying both age and FDR adjustments ($q < 0.05$) were identified as consistently elevated in urine samples of IBS patients at both collection time points (IBS-1, IBS-2) when compared to non-IBS controls (**Table 3.2; Figure S3.6** of the Supporting Information). These metabolites were consistent with most of the lead candidate biomarkers also identified in the OPLS-DA model (**Figure 3.2a, b**), which can be categorized into three major metabolite classes, including primary amino acids and their derivatives, glycated amino acids, and methylated nucleosides. Putative identification of imidazole propionate (**Figure S3.4** of the Supporting Information) and confirmatory identification of *N,N*-dimethylguanosine (**Figure S3.5** of the Supporting Information) were made based on accurate mass and most likely molecular formula, MS/MS spectral matching and spiking with authentic chemical standards. As expected, a strong correlation was observed among glycosylated amino acids and modified nucleosides based on Spearman rank correlation test (**Figure 3.3**), which is indicative of their common underlying biochemical pathways.

For instance, a urinary metabolite was identified as one the top-ranked urinary metabolites elevated in IBS patients at both time points collected, which was subsequently identified as Glu-Gal-OH-Lys (*m/z*:RMT, 487.2112:0.777) based on its diagnostic MS/MS spectrum as shown in **Figure 3.2e** comprising a product ion for OH-Lys (*m/z* 163.1075) and a neutral loss corresponding to a disaccharide (*m/z* 324).

Table 3.2. Top-ranked osmolality-normalized urinary metabolites associated with IBS status when applying ANCOVA pair-wise comparison of IBS-1 and IBS-2 patients with non-IBS/healthy controls. All *p*-values were adjusted using age as a covariate while applying a Benjamini-Hochberg FDR adjustment.

m/z:RMT	ID	*p*-value1	*q*-value1	FC 1[a]	*p*-value2	*q*-value2	FC 2[a]
367.1500: 1.092	*C*-Mannopyranosyl-tryptophan	8.90E-05	1.04 E-03	1.84	0.0108	> 0.05	1.48
147.1128: 0.515	Lysine[c]	2.83E-04	2.27 E-03	2.04	2.82 E-03	0.0151	1.72
141.0659: 0.625	Imidazole propionate[b]	3.54E-04	2.70E-03	2.21	2.21 E-04	2.12 E-03	2.18
487.2117: 0.777	*O*-Glucosyl galactosyl-hydroxylysine[c]	5.92E-04	4.16 E-03	1.68	1.55 E-03	8.57 E-03	1.84
325.1605 : 0.720	*O*-Galactosyl-hydroxylysine[c]	9.25E-04	6.27 E-03	1.66	8.84 E-04	5.95 E-03	1.64
147.0764: 0.886	Glutamine[c]	1.24E-03	8.35E-03	1.32	3.06 E-04	2.58 E-03	1.40
133.0608: 0.857	Ornithine	1.62E-03	0.0110	1.23	1.12 E-03	6.51 E-03	1.23
106.0499: 0.805	Serine	0.0121	> 0.05	1.30	6.93 E-04	0.0427	1.50
104.0706: 0.909	*N*-*N*-Dimethyl glycine[c]	0.0184	> 0.05	1.98	1.00 E-05	3.09 E-04	2.09
312.1297: 1.102	*N*,*N*-Dimethyl guanosine	0.0257	> 0.05	1.53	2.07 E-03	0.0111	1.49

[a] *Fold-change (FC) calculated by median untransformed RPAs of IBS/non-IBS*

[b] *Putative identification by high resolution MS and MS/MS, whereas other metabolites confirmed by spiking with chemical standard together with MS/MS spectral matching.*

[c] *These urinary metabolites were also found to significant when using creatinine normalization.*

In fact, Glu-Gal-OH-Lys also undergoes significant in-source fragmentation during ionization as shown in extracted ion electropherograms (EIE) in **Figure 3.2c** with a glycan loss (*m/z* 162) to generate an isobaric ion to Gal-OH-Lys. Note however that intact Glu-Gal-OH-Lys and its pseudo- MS/MS product ions (Gal-OH-Lys) co-migrate with each other with longer migration times than Gal-OH-Lys with a higher RMT in CE. As expected, both Glc-Gal-OH-Lys and Gal-OH-Lys are strongly co-linear with a Spearman rho > 0.70, whereas they cluster together with Lys and Man-Trp in the heat map shown in **Figure 3.3**. Overall, there were no urinary metabolites significantly associated with anxiety and depression status based on HADS-A and HADS-D scores (rho < 0.4). Although, a chemical standard for Glc-Gal-OH-Lys was not available, the combined evidence of distinctive migration times, MS/MS fragmentation pattern, and in-source fragmentation patterns were
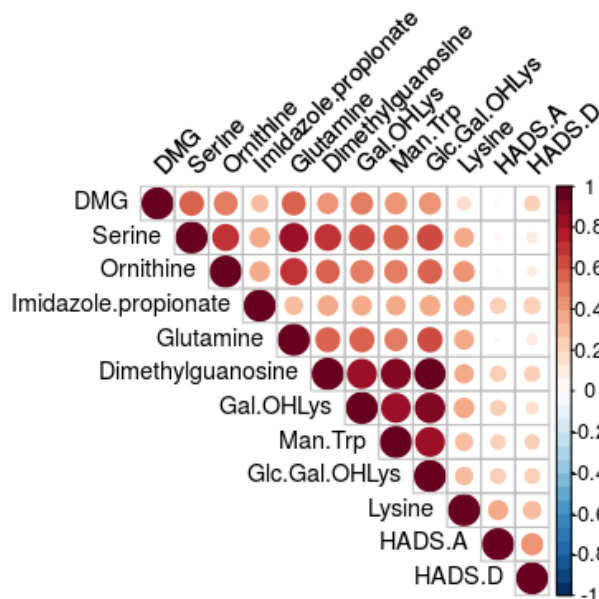
Figure 3.3 A correlation heat map for the lead candidate urinary biomarkers of IBS as compared to healthy controls, as well as anxiety and depression questionnaire scores when using a Spearman correlation rank analysis, where colour gradient indicates Spearman rho correlation coefficients. All urine metabolite responses were osmolality normalized and untransformed. Abbreviations include, HADS.A: Hospital Anxiety and Depression Scale–Anxiety; HADS.D: Hospital Anxiety and Depression Scale–Depression; DMG: *N,N*-Dimethylglycine; Man.Trp: C-Mannopyranosyl-tryptophan, GalOHLys: *O*-Galactosyl-hydroxylysine; GlcGalOHLys: *O*-Glucosylgalactosyl-hydroxylysine.

sufficient to conclude that the latter migrating peak was indeed Glc-Gal-OH-Lys. Additionally, when creatinine normalized data was also examined when using ANCOVA for age-adjustment, higher excretion of both glycosylated hydroxylysine adducts, as well as lysine, imidazole propionate, dimethylglycine and glutamine were also found to be significant ($p < 0.05$) in at least one time period irrespective of method of urine normalization (**Table S3.2**).

Despite a previously studied association between gut microbial metabolism and IBS as seen in increased amount of volatile compounds in stool and breath of IBS patients (Ahmed et al. 2013; Baranska et al. 2016), as well as tryptophan metabolism in relation to serotonin (Quigley 2016), significant differences in excretion levels were observed mostly among products of human metabolism in our study. For example, metabolites involved with

gut microbiota activity, such as indoxyl sulfate and *p*-cresol sulfate that are also known as uremic toxins (Edamatsu et al. 2018), and short-chain fatty acids (*e.g.,* lactate and butyrate) were also measured in this study; however, they did not show a statistically significant difference between healthy controls and IBS patients (**Figure S3.7**) with the exception of urinary indole acetic acid but only with creatinine normalization (**Table S3.2**).

3.5 Discussion

Comprehensive analysis of urinary metabolites of IBS patients and sex-matched non-IBS controls were characterized and statistically compared when using a high throughput metabolomics platform based on MSI-CE-MS that takes advantage of stringent QC to minimize false discoveries. In this study, spurious/background signals were effectively rejected by using a dilution trend filter of a pooled urine sample to authenticate reproducible yet frequently measured sample-derived metabolites in an untargeted manner prior to individual sample analysis, which greatly streamlined subsequent annotation of the urine metabolome. This combination of untargeted and targeted approaches resulted in 138 authentic urinary metabolites measured in majority (> 75%) of urine samples with good technical precision (CV < 30% for QC samples) as summarized in **Table S3.1**. For normalization of urine metabolite levels to account for different hydration status, creatinine is most commonly used in clinical investigations, but suffers major limitations relevant to our study since its concentration is highly dependent on sex, age and habitual diet, such as protein intake (Carrieri et al. 2000). Alternatively, urine normalization to osmolality was found to be a more robust approach to correct for hydration status when relying on single-spot/random urine specimen collections as it is dependent on several major solutes/electrolytes in urine instead of a creatinine alone (Chadha et al. 2001). In addition, osmolality measurements based on freezing point depression are known to be highly reproducible with a CV < 1 % (Bockenhauer and Aitkenhead 2011), which was also achieved in our study.

Using osmolality as a normalization parameter, we identified a panel of urinary metabolites that were significantly elevated in the urine of IBS patients as compared to

healthy controls, which in most cases remained significant with a follow-up independent repeat urine collection for a majority of the same IBS patients (IBS-2). However, the chemical structures of the top-three lead candidate biomarkers differentially excreted between two groups were initially unknown. Since there can be hundreds of potential isobaric or isomeric candidates that match an accurate mass or a molecular formula of an unknown metabolite, structural elucidation remains one of the major bottlenecks in MS-based metabolomics for biomarker discovery (Bowen and Northen 2010). Typically, information derived from unknown molecular features are searched against entries of authentic compounds, which is challenging given the far greater number of expected metabolites than well curated MS/MS spectra available in public databases (Peironcely et al. 2011). Additionally, investigators rarely undertake the challenging and time-consuming process of *de novo* structural elucidation when interpreting high resolution MS/MS spectra (Viant et al. 2017). However, successful identification of unknown metabolites is necessary as it provides valuable biochemical insights regarding mechanisms, as well as valuable insights into the potential clinical significance of putative biomarkers. In our work, we identified several unknown urinary metabolites based on high resolution MS, MS/MS fragmentation, as well as a comparison of their electromigration behavior with that of authentic standards as illustrated for OH-Lys, Glc-Gal-OH-Lys and Man-Trp in **Figure 3.2c-e** and **Figure S3.3** of the Supporting Information.

IBS is a functional bowel disorder characterized by altered GI motility and heightened pain perception (*i.e.* visceral hypersensitivity), which are frequently associated with psychological conditions, such as depression. This close association between neurological pathway and GI symptoms makes IBS a classic example of disorder of the "brain-gut axis" (Dinan and Cryan 2017). Additionally, despite the absence of overt inflammation and/or intestinal epithelial tissue damage, low-grade inflammation and immune activation has long been implicated in the etiology of IBS based on epidemiological studies reporting past episodes of viral and bacterial infection as the most significant environmental risk factor for developing IBS (Rodríguez and Ruigómez 1999;

Spiller et al. 2000; Gwee et al. 2003; Marshall et al. 2007). In fact, opportunistic infection by *Helicobacter pylori* caused visceral hypersensitivity and persistent abnormal feeding behavior in conjunction with higher inflammatory markers (*e.g.* TNF-α) even after complete eradication of the infection in mice (Bercik et al. 2009). Moreover, a full-thickness biopsy of the jejunum from patients with severe IBS showed evidence of low-grade infiltration of lymphocytes in the myenteric plexus, a clear sign of immune response, in 9 out of 10 patients (Törnblom et al. 2002).

In this pilot study, two glycosylated hydroxylysine metabolites, a glycated tryptophan, a modified nucleoside (dimethylguanosine), primary amino acids (lysine, serine, ornithine, and glutamine) and other amino acid catabolites (dimethylglycine and imidazole propionate) were found significantly elevated in urine samples of IBS patients as compared to controls even after adjusting for age and multiple hypothesis testing (FDR). The majority of these urinary metabolites were also consistently elevated in osmolality corrected repeat urine specimens collected from the same IBS patients, as well as with creatinine normalization. The presence of glycosylated OH-Lys metabolites in human urine was first reported in literature back in 1967 and later recognized as the index for collagen degradation and turn-over especially for type I collagen in the context of bone tissue formation (Cunningham et al. 1967; Krane et al. 1977; Bateman et al. 1984). In addition to type I collagen, glycosylation of OH-Lys are found in other types of collagen, such as type III and IV, which are part of intestinal stroma and basement membrane of colon (William 1978; Hilska et al. 1998). Collagen is one of the most abundant proteins in our body, which plays a critical role in elasticity and tensile strength of tissue, including the intestinal epithelium (Nimni 1983). Altered collagen production and function has been associated with disorders such as biliary cirrhosis (Savolainen et al. 1983) and colonic diverticulosis (Wess et al. 1995), highlighting its importance in structural integrity of organs associated with digestive system. To the best of our knowledge, increased excretion of collagen degradation biomarkers has not previously been reported in the context of IBS, but the implicated reduction of tissue elasticity via accelerated collagen degradation may represent

an underlying mechanism of altered function of GI motility in IBS (**Figure 3.4**). We also found higher excretion of a modified nucleoside, *N'N*-dimethylguanosine, which has been previously reported as a biomarker of higher cell turnover as observed in various types of cancers (Zheng et al. 2005). Furthermore, high excretion of imidazole propionate may also support the presence of low-grade immune activation implicated among IBS patients since it is a catabolite of histidine metabolism that has recently been shown to be produced by gut microbiota (Koh et al. 2018), which may also regulate histamine that is involved in local immune responses and regulation of normal GI function (Deiteren et al. 2015). Similarly, Man-Trp has been found in immune-related proteins such as cytokine receptor type I family proteins (Hamming et al. 2012), and functions as an important component for protein folding and protein-protein interactions (Ihara et al. 2015). Despite being widely marketed as a nutritional supplement due to its primary roles in methionine and folic acid biosynthesis (Friesen et al. 2007), recent evidence demonstrates that dimethylglycine metabolism can generate reactive oxygen species in liver mitochondria (Mailloux et al. 2016), which may be responsible for its putative immunostimulating functions and use as an immunoadjuvant in animals (Graber et al. 1981; Reap and Lawson 1990). In fact, dietary choline is largely excreted as dimethylglycine in urine whose metabolism is involved in a complex host/gut microbiota molecular cross-talk (Dumas et al. 2006).

Although this is the first comprehensive urinary metabolomics study conducted on IBS patients in comparison to non-IBS controls, Goo *et al*. (Goo et al. 2012) investigated urinary proteome of female IBS patients with different symptom subgroups in comparison to healthy control women. Their LC-MS/MS based method identified significantly increased urinary glycoproteins associated with mucosal integrity (*e.g*. MUC1, TFF3) in 49 IBS patients compared to 24 healthy controls regardless of IBS subtype. Moreover, urinary excretion of proteins associated with proteolysis (*e.g*. SERPIN B4, γ-glutamylhydrolase) was specifically higher in IBS-D patients, which may explain the higher urinary excretion of amino acids (*i.e.* Ser, Orn, Lys, and Gln) observed in our study.
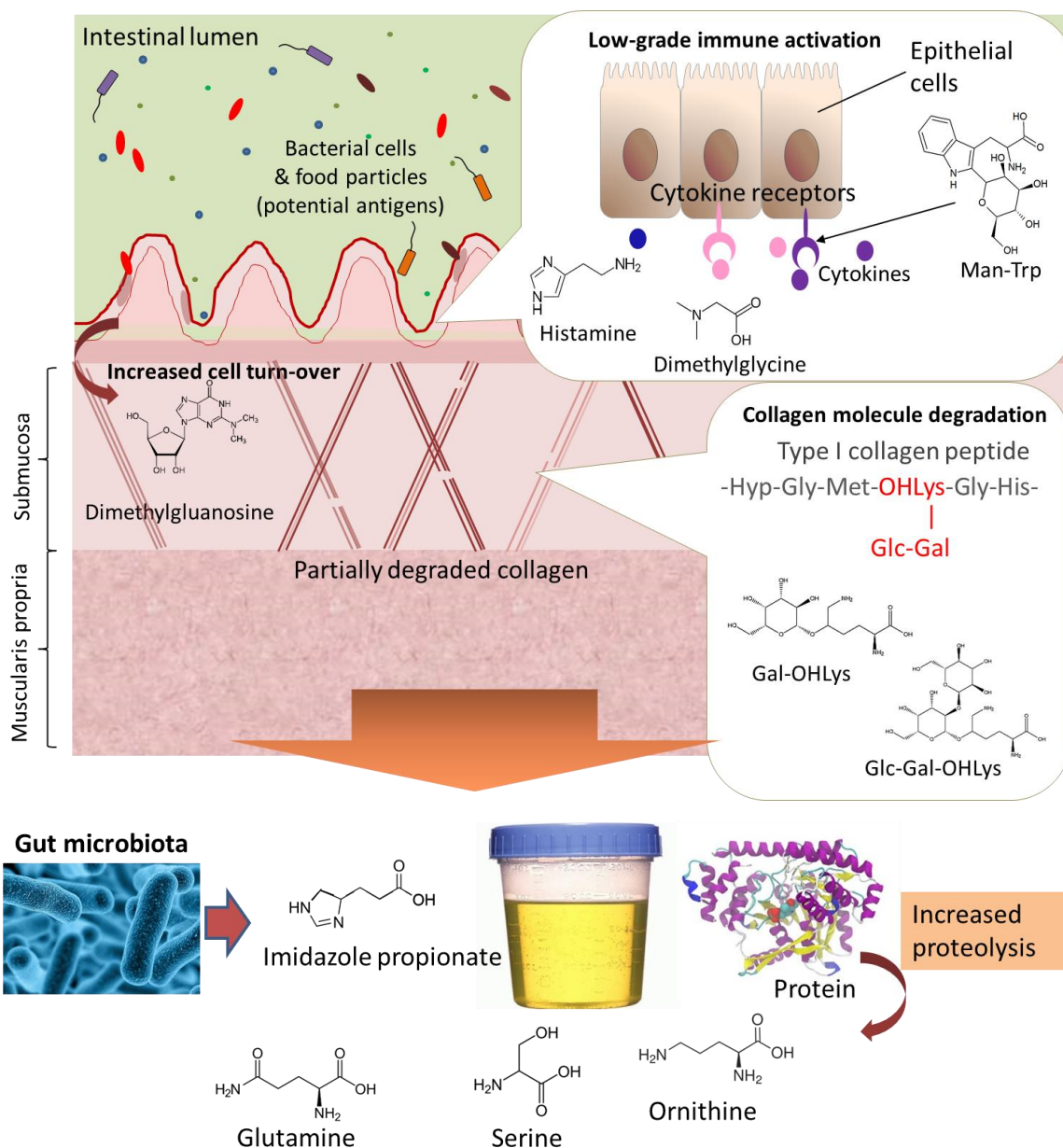
Figure 3.4 A proposed scheme illustrating the putative pathological mechanisms of IBS based on lead candidate biomarkers identified in urine from this pilot study. IBS is characterized by low-grade immune activation, potentially caused by food- or bacteria-derived antigens, and increased degradation of collagen within the intestinal mucosa (Glc-Gal-OHLys, Gal-OHLys, OH-Lys), which leads to higher cell turn-over (dimethylguanosine) and excretion of metabolites related to inflammatory responses/immunomodulation (imidazole propionate, Man-Trp, dimethylglycine). The higher excretion of primary and modified amino acids (Gln, Ser, Orn) may be the result of their greater urinary clearance due to increased proteolysis as a result of disturbed homeostasis in related protein expression. Abbreviations: Man-Trp: *C*-mannopyranosyl-tryptophan; Gal-OHLys: *O*-galactosyl-hydroxylysine; Glc-Gal-OHLys: *O*-glucosylgalactosyl-hydroxylysine.

Among the key amino acids found in our study, Gln is known to participate in the glutamate-glutamine flux in the brain, which plays an essential role to replenish a major excitatory neurotransmitter within the CNS, glutamate (Zhou et al. 2010). Indeed, reduced hippocampal glutamate and glutamine levels were found in IBS patients by magnetic resonance spectroscopy when studying 15 IBS patients and 15 age-matched healthy controls (Niddam et al. 2011). Therefore, higher excretion of Gln in our study may be an indication of disturbed glutamate-glutamine flux homeostasis in CNS of IBS patients. Taken together, our study indicated the involvement of immune-activation, structural/functional integrity of GI tissue, higher cell turnover, altered proteolysis and neurotransmitter homeostasis IBS pathology based on untargeted urinary metabolome analysis. The increased permeability of GI epithelia and associated immune activation are also found in IBD, which is characterized by overt inflammation and tissue damage and considered as more severe condition than IBS. IBD is classically defined as immune-mediated disorder, whereas IBS does not involve immunoactivation by definition (Drossman 2016; Geremia et al. 2014). Thus, these two disorders are considered as separate entities in clinical practice despite their overlapping symptoms. However, as previously suggested by Qin, IBS and IBD may share the presence of mucosal degradation with subsequent damage to GI wall occurring at different rates (Qin 2017), which would explain the ten-fold higher risk among newly diagnosed IBS patients to be diagnosed with IBD within an average of three years as compared healthy subjects (L. A. García Rodríguez 2000), and consistently reported presence of IBS-like symptoms in adult (Forough et al. 2006; Abdalla et al. 2017) and pediatric (Diederen et al. 2016) patients with IBD.

Overall, the major limitations in this study were significant co-morbidity of IBS patients with other illnesses, including depression with a far greater prescribed drug history as compared to healthy controls. Also, this study lacked dietary control prior to urine sample collection, such as a wash-out period with standardized meals prior to sampling, which are known to influence the urine metabolome (Walsh et al. 2006; Want et al. 2010; Nicholson et al. 2002). For instance, we have identified several urinary metabolites derived

from over-counter analgesics (*e.g.* acetaminophen sulfate) and tobacco exposure (*e.g.,* *trans*-hydroxycotinine), but there were other medications, such as thyroid hormones, bronchodilators, anticoagulants and antibiotics, that we did not directly measure. To account for the heterogeneity of the IBS group, it was ideal to have larger control group; however, our control group was about half the size of IBS group and it did not include repeat urine specimens collected. Urinary excretion of amino acids are expected to depend not only on differences in genetics and gut microbiota co-metabolism, but also on habitual dietary patterns (Fowler et al. 1957; Walsh et al. 2006). Since a majority of IBS patients reported that meals or specific food triggered their symptoms in surveys on perceived food intolerance (Monsbakken et al. 2005; Simren et al. 2001; Hayes et al. 2014), it is likely that IBS patients, especially those who have been experiencing the disorder for many years, tend to follow restrictive diets to avoid exacerbation of their symptoms, and thus may have influenced urinary metabolic profiles notably for amino acids.

In summary, we report several novel metabolic signatures from human urine that offers new insights into an unrecognized mechanism associated with the pathophysiology of IBS. Notably, this is the first study to report high excretion of glycosylated amino acids in urine of IBS patients. Our findings support the mucosal layer degradation hypothesis underlying IBS symptoms and the presence of chronic low-grade inflammation and immune activation, which may contribute to impaired structural integrity and easier access of luminal contents to GI walls (*i.e.,* leaky gut) compromising GI motility due to accelerated collagen turn-over. Once validated within a larger cohort of IBS patients with different severity of symptom replicated across multiple centres, a panel of urinary metabolites identified in this study may replace the need for invasive colonoscopy and serve as a convenient way to monitor progression and treatment responses to therapy. Future work will also investigate the potential of using urine profiling for IBS screening in order to identify high-risk groups susceptible in developing IBS before severe GI symptoms manifest, including differential diagnosis of IBS from other GI disorders, such as Crohn's disease and ulcerative colitis. Additionally, repeat urine specimens from longitudinal

studies involving treatment naïve patients recently diagnosed with IBS would help reduce potential confounding and increase study power together with metabolomic studies of complementary biospecimens (*e.g.,* plasma and stool) to better delineate the underlying biochemical mechanisms that distinguish IBS from major IBD subtypes.

Acknowledgements

Ethical approval: All procedures performed in studies involving human participants were in accordance with the ethical standards of the Hamilton Health Sciences and St. Joseph's Health Care Research Ethics Boards. Informed consent was collected from every participant in this study prior to any personal information or sample collection.

3.6 References

Abdalla, M. I., Sandler, R. S., Kappelman, M. D., Martin, C. F., Chen, W., Anton, K., et al. (2017). Prevalence and Impact of Inflammatory Bowel Disease–Irritable Bowel Syndrome on Patient-reported Outcomes in CCFA Partners *Inflammatory Bowel Diseases, 23*(2), 325-331.

Ahmed, I., Greenwood, R., Costello, B. d. L., Ratcliffe, N. M., & Probert, C. S. (2013). An Investigation of Fecal Volatile Organic Metabolites in Irritable Bowel Syndrome. *PLoS ONE, 8*(3), e58204, doi:10.1371/journal.pone.0058204.

Azpiroz, F., Bouin, M., Camilleri, M., Mayer, E. A., Poitras, P., Serra, J., et al. (2007). Mechanisms of hypersensitivity in IBS and functional disorders. *Neurogastroenterology & Motility, 19*(1 Suppl), 62-88, doi:10.1111/j.1365-2982.2006.00875.x.

Baranska, A., Mujagic, Z., Smolinska, A., Dallinga, J. W., Jonkers, D. M. A. E., Tigchelaar, E. F., et al. (2016). Volatile organic compounds in breath as markers for irritable bowel syndrome: a metabolomic approach. *Alimentary Pharmacology & Therapeutics, 44*(1), 45-56, doi:doi:10.1111/apt.13654.

Bateman, J. F., Mascara, T., Chan, D., & Cole, W. G. (1984). Abnormal type I collagen metabolism by cultured fibroblasts in lethal perinatal osteogenesis imperfecta. *Biochemical Journal, 217*(1), 103-115.

Bercik, P., Verdú, E. F., Foster, J. A., Lu, J., Scharringa, A., Kean, I., et al. (2009). Role of gut-brain axis in persistent abnormal feeding behavior in mice following eradication of Helicobacter pylori infection. [10.1152/ajpregu.90752.2008]. *American Journal of Physiology - Regulatory, Integrative and Comparative Physiology, 296*(3), R587-R594.

Bockenhauer, D., & Aitkenhead, H. (2011). The kidney speaks: interpreting urinary sodium and osmolality. *Archives of disease in childhood - Education & practice edition, 96*(6), 223-227, doi:10.1136/archdischild-2011-300115.

Böhning, D., Böhning, W., Guha, N., Cowan, D. A., Sönksen, P. H., & Holt, R. I. G. (2016). Statistical methodology for age-adjustment of the GH-2000 score detecting growth hormone misuse. *BMC Medical Research Methodology, 16*(1), 147, doi:10.1186/s12874-016-0246-8.

Bouatra, S., Aziat, F., Mandal, R., Guo, A. C., Wilson, M. R., Knox, C., et al. (2013). The human urine metabolome. *PLoS One, 8*(9), e73076, doi:10.1371/journal.pone.0073076.

Bowen, B. P., & Northen, T. R. (2010). Dealing with the Unknown: Metabolomics and Metabolite Atlases. *Journal of the American Society for Mass Spectrometry, 21*(9), 1471-1476, doi:http://dx.doi.org/10.1016/j.jasms.2010.04.003.

Brunius, C., Shi, L., & Landberg, R. (2016). Large-scale untargeted LC-MS metabolomics data correction using between-batch feature alignment and cluster-based within-batch signal intensity drift correction. *Metabolomics, 12*(11), 173, doi:10.1007/s11306-016-1124-4.

Camilleri, M., Shin, A., Busciglio, I., Carlson, P., Acosta, A., Bharucha, A. E., et al. (2014). Validating Biomarkers of Treatable Mechanisms in Irritable Bowel Syndrome. *Neurogastroenterology and motility : the official journal of the European Gastrointestinal Motility Society, 26*(12), 1677-1685, doi:10.1111/nmo.12421.

Carrieri, M., Trevisan, A., & Bartolucci, G. B. (2000). Adjustment to concentration-dilution of spot urine samples: correlation between specific gravity and creatinine. *International Archives of Occupational and Environmental Health, 74*(1), 63-67, doi:10.1007/s004200000190.

Chadha, V., Garg, U., & Alon, U. S. (2001). Measurement of urinary concentration: a critical appraisal of methodologies. *Pediatric Nephrology, 16*(4), 374-382, doi:10.1007/s004670000551.

Changhyun, L., lt, sup, gt, lt, sup, et al. (2017). The Increased Level of Depression and Anxiety in Irritable Bowel Syndrome Patients Compared with Healthy Controls: Systematic Review and Meta-analysis. *Journal of Neurogastroenterology and Motility, 23*(3), 349-362.

Chong, J., Soufan, O., Li, C., Caraus, I., Li, S., Bourque, G., et al. (2018). MetaboAnalyst 4.0: towards more transparent and integrative metabolomics analysis. *Nucleic Acids Research, 46*(W1), W486-W494, doi:10.1093/nar/gky310.

Cleare, A., Pariante, C. M., Young, A. H., Anderson, I. M., Christmas, D., Cowen, P. J., et al. (2015). Evidence-based guidelines for treating depressive disorders with antidepressants: A revision of the 2008 British Association for Psychopharmacology guidelines. *J Psychopharmacol, 29*(5), 459-525, doi:10.1177/0269881115581093.

Cunningham, L. W., Ford, J. D., & Segrest, J. P. (1967). The Isolation of Identical Hydroxylysyl Glycosides from Hydrolysates of Soluble Collagen and from Human Urine. *Journal of Biological Chemistry, 242*(10), 2570-2571.

Dahl, H., Stephanson, N., Beck, O., & Helander, A. (2002). Comparison of Urinary Excretion Characteristics of Ethanol and Ethyl Glucuronide. *Journal of Analytical Toxicology, 26*(4), 201-204, doi:10.1093/jat/26.4.201.

Deiteren, A., De Man, J. G., Pelckmans, P. A., & De Winter, B. Y. (2015). Histamine H4 receptors in the gastrointestinal tract. *British Journal of Pharmacology, 172*(5), 1165-1178, doi:doi:10.1111/bph.12989.

DiBattista, A., McIntosh, N., Lamoureux, M., Al-Dirbashi, O. Y., Chakraborty, P., & Britz-McKibbin, P. (2017a). Temporal Signal Pattern Recognition in Mass Spectrometry: A Method for Rapid Identification and Accurate Quantification of Biomarkers for Inborn Errors of Metabolism with Quality Assurance. *Analytical Chemistry, 89*(15), 8112-8121, doi:10.1021/acs.analchem.7b01727.

DiBattista, A., Rampersaud, D., Lee, H., Kim, M., & Britz-McKibbin, P. (2017b). High Throughput Screening Method for Systematic Surveillance of Drugs of Abuse by Multisegment Injection–Capillary Electrophoresis–Mass Spectrometry. *Analytical Chemistry, 89*(21), 11853-11861, doi:10.1021/acs.analchem.7b03590.

Diederen, K., Hoekman, D. R., Hummel, T. Z., Meij, T. G., Koot, B. G. P., Tabbers, M. M., et al. (2016). The prevalence of irritable bowel syndrome-type symptoms in paediatric inflammatory bowel disease, and the relationship with biochemical markers of disease activity. *Alimentary Pharmacology & Therapeutics, 44*(2), 181-188, doi:doi:10.1111/apt.13636.

Dinan, T. G., & Cryan, J. F. (2017). Brain–gut–microbiota axis — mood, metabolism and behaviour. *Nature Reviews Gastroenterology & Hepatology, 14*, 69, doi:10.1038/nrgastro.2016.200.

Drossman, D. A. (2016). Functional Gastrointestinal Disorders: History, Pathophysiology, Clinical Features, and Rome IV. *Gastroenterology, 150*(6), 1262-1279.e1262, doi:https://doi.org/10.1053/j.gastro.2016.02.032.

Dumas, M.-E., Barton, R. H., Toye, A., Cloarec, O., Blancher, C., Rothwell, A., et al. (2006). Metabolic Profiling Reveals a Contribution of Gut Microbiota to Fatty Liver Phenotype in Insulin-Resistant Mice. *Proceedings of the National Academy of Sciences of the United States of America, 103*(33), 12511-12516.

Edamatsu, T., Fujieda, A., & Itoh, Y. (2018). Phenyl sulfate, indoxyl sulfate and p-cresyl sulfate decrease glutathione level to render cells vulnerable to oxidative stress in renal tubular cells. *PLOS ONE, 13*(2), e0193342, doi:10.1371/journal.pone.0193342.

Eisenstein, M. (2016). Microbiome: Bacterial broadband. *Nature, 533*(7603), S104-S106, doi:10.1038/533S104a.

Enck, P., Aziz, Q., Barbara, G., Farmer, A. D., Fukudo, S., Mayer, E. A., et al. (2016). Irritable bowel syndrome. [Primer]. *Nature Reviews Disease Primers, 2*, 16014, doi:10.1038/nrdp.2016.14.

Forough, F., K, M. J., Brock, E., & Jan, I. E. (2006). Functional gastrointestinal disorders and mood disorders in patients with inactive inflammatory bowel disease: Prevalence and impact on health. *Inflammatory Bowel Diseases, 12*(1), 38-46, doi:doi:10.1097/01.MIB.0000195391.49762.89.

Fowler, D. I., Norton, P. M., Cheung, M. W., & Pratt, E. L. (1957). Observations on the urinary amino acid excretion in man: the influence of age and diet. *Archives of Biochemistry and Biophysics, 68*(2), 452-466, doi:https://doi.org/10.1016/0003-9861(57)90376-4.

Friesen, R. W., Novak, E. M., Hasman, D., & Innis, S. M. (2007). Relationship of dimethylglycine, choline, and betaine with oxoproline in plasma of pregnant women and their newborn infants. *The Journal of Nutrition, 137*(12), 2641-2646, doi:10.1093/jn/137.12.2641.

Geremia, A., Biancheri, P., Allan, P., Corazza, G. R., & Di Sabatino, A. (2014). Innate and adaptive immunity in inflammatory bowel disease. *Autoimmunity Reviews, 13*(1), 3-10, doi:https://doi.org/10.1016/j.autrev.2013.06.004.

Goo, Y. A., Cain, K., Jarrett, M., Smith, L., Voss, J., Tolentino, E., et al. (2012). Urinary Proteome Analysis of Irritable Bowel Syndrome (IBS) Symptom Subgroups. *Journal of Proteome Research, 11*(12), 5650-5662, doi:10.1021/pr3004437.

Graber, C. D., Goust, J. M., Glassman, A. D., Kendall, R., & Loadholt, C. B. (1981). Immunomodulating Properties of Dimethylglycine in Humans. *The Journal of Infectious Diseases, 143*(1), 101-105, doi:10.1093/infdis/143.1.101.

Gwee, K.-A., Collins, S. M., Read, N. W., Rajnakova, A., Deng, Y., Graham, J. C., et al. (2003). Increased rectal mucosal expression of interleukin 1β in recently acquired post-infectious irritable bowel syndrome. *Gut, 52*(4), 523-526, doi:10.1136/gut.52.4.523.

Halmos, E. P., Power, V. A., Shepherd, S. J., Gibson, P. R., & Muir, J. G. (2014). A Diet Low in FODMAPs Reduces Symptoms of Irritable Bowel Syndrome. *Gastroenterology, 146*(1), 67-75.e65, doi:10.1053/j.gastro.2013.09.046.

Hamming, O. J., Kang, L., Svensson, A., Karlsen, J. L., Rahbek-Nielsen, H., Paludan, S. R., et al. (2012). Crystal structure of interleukin-21 receptor (IL-21R) bound to IL-21 reveals that sugar chain interacting with WSXWS motif is integral part of IL-21R. *Journal of Biological Chemistry, 287*(12), 9454-9460, doi:10.1074/jbc.M111.311084.

Hayes, P. A., Fraher, M. H., & Quigley, E. M. M. (2014). Irritable Bowel Syndrome: The Role of Food in Pathogenesis and Management. *Gastroenterology & Hepatology, 10*(3), 164-174.

Hilska, M., Collan, Y., Peltonen, J., Gullichsen, R., Paajanen, H., & Laato, M. (1998). The distribution of collagen types I, III, and IV in normal and malignant colorectal mucosa. *European Journal of Surgery, 164*(6), 457-464, doi:10.1080/110241598750004274.

Hofmann, A. F., & Hagey, L. R. (2008). Bile Acids: Chemistry, Pathochemistry, Biology, Pathobiology, and Therapeutics. [journal article]. *Cellular and Molecular Life Sciences, 65*(16), 2461-2483, doi:10.1007/s00018-008-7568-6.

Ihara, Y., Inai, Y., Ikezaki, M., Matsui, I.-S. L., Manabe, S., & Ito, Y. (2015). C-Mannosylation: Modification on Tryptophan in Cellular Proteins. In N. Taniguchi, T. Endo, W. G. Hart, H. P. Seeberger, & C.-H. Wong (Eds.), *Glycoscience: Biology and Medicine* (pp. 1091-1099). Tokyo: Springer Japan.

Johnson, C. H., Ivanisevic, J., & Siuzdak, G. (2016). Metabolomics: beyond biomarkers and towards mechanisms. *Nature Review Molecular Cell Biology, 17*(7), 451-459, doi:10.1038/nrm.2016.25.

Kendig, D. M., & Grider, J. R. (2015). Serotonin and Colonic Motility. *Neurogastroenterology and motility : the official journal of the European Gastrointestinal Motility Society, 27*(7), 899-905, doi:10.1111/nmo.12617.

Kennedy, P. J., Cryan, J. F., Dinan, T. G., & Clarke, G. (2014). Irritable bowel syndrome: A microbiome-gut-brain axis disorder? *World Journal of Gastroenterology : WJG, 20*(39), 14105-14125, doi:10.3748/wjg.v20.i39.14105.

Keszthelyi, D., Troost, F. J., Jonkers, D. M., Kruimel, J. W., Leue, C., & Masclee, A. A. M. (2013). Decreased levels of kynurenic acid in the intestinal mucosa of IBS patients: Relation to serotonin and psychological state. *Journal of Psychosomatic Research, 74*(6), 501-504, doi:http://dx.doi.org/10.1016/j.jpsychores.2013.01.008.

Krane, S. M., Kantrowitz, F. G., Byrne, M., Pinnell, S. R., & Singer, F. R. (1977). Urinary excretion of hydroxylysine and its glycosides as an index of collagen degradation. *Journal of Clinical Investigation, 59*(5), 819-827.

Kuehnbaum, N. L., Gillen, J. B., Kormendi, A., Lam, K. P., DiBattista, A., Gibala, M. J., et al. (2015). Multiplexed separations for biomarker discovery in metabolomics: Elucidating adaptive responses to exercise training. *ELECTROPHORESIS, 36*(18), 2226-2236, doi:doi:10.1002/elps.201400604.

Kuehnbaum, N. L., Kormendi, A., & Britz-McKibbin, P. (2013). Multisegment Injection-Capillary Electrophoresis-Mass Spectrometry: A High-Throughput Platform for Metabolomics with High Data Fidelity. *Analytical Chemistry, 85*(22), 10664-10669, doi:10.1021/ac403171u.

L. A. García Rodríguez, A. R., M.-A. Wallander, S. Johansson, L. Olbe (2000). Detection of Colorectal Tumor and Inflammatory Bowel Disease during Follow-up of Patients with Initial Diagnosis of Irritable Bowel Syndrome. *Scandinavian Journal of Gastroenterology, 35*(3), 306-311, doi:doi:10.1080/003655200750024191.

Lacy, B. E. (2016). Perspective: An easier diagnosis. *Nature, 533*(7603), S107-S107, doi:10.1038/533S107a.

Lacy, B. E., Mearin, F., Chang, L., Chey, W. D., Lembo, A. J., Simren, M., et al. (2016). Bowel Disorders. *Gastroenterology, 150*(6), 1393-1407.e1395, doi:10.1053/j.gastro.2016.02.031.

Macedo, A. N., Mathiaparanam, S., Brick, L., Keenan, K., Gonska, T., Pedder, L., et al. (2017). The Sweat Metabolome of Screen-Positive Cystic Fibrosis Infants:

Revealing Mechanisms beyond Impaired Chloride Transport. *ACS Central Science, 3*(8), 904-913, doi:10.1021/acscentsci.7b00299.

Mahieu, N. G., & Patti, G. J. (2017). Systems-Level Annotation of a Metabolomics Data Set Reduces 25 000 Features to Fewer than 1000 Unique Metabolites. *Analytical Chemistry, 89*(19), 10397-10406, doi:10.1021/acs.analchem.7b02380.

Mailloux, R. J., Young, A., Chalker, J., Gardiner, D., O'Brien, M., Slade, L., et al. (2016). Choline and dimethylglycine produce superoxide/hydrogen peroxide from the electron transport chain in liver mitochondria. *FEBS Letters, 590*(23), 4318-4328, doi:doi:10.1002/1873-3468.12461.

Marshall, J. K., Thabane, M., Borgaonkar, M. R., & James, C. (2007). Postinfectious Irritable Bowel Syndrome After a Food-Borne Outbreak of Acute Gastroenteritis Attributed to a Viral Pathogen. *Clinical Gastroenterology and Hepatology, 5*(4), 457-460, doi:10.1016/j.cgh.2006.11.025.

Massadeh, A. M., Gharaibeh, A. A., & Omari, K. W. (2009). A Single-Step Extraction Method for the Determination of Nicotine and Cotinine in Jordanian Smokers' Blood and Urine Samples by RP-HPLC and GC-MS. *Journal of Chromatographic Science, 47*(2), 170-177, doi:10.1093/chromsci/47.2.170.

McIntosh, K., Reed, D. E., Schneider, T., Dang, F., Keshteli, A. H., De Palma, G., et al. (2017). FODMAPs alter symptoms and the metabolome of patients with IBS: a randomised controlled trial. [10.1136/gutjnl-2015-311339]. *Gut, 66*(7), 1241.

Meleine, M., & Matricon, J. (2014). Gender-related differences in irritable bowel syndrome: Potential mechanisms of sex hormones. *World Journal of Gastroenterology : WJG, 20*(22), 6725-6743, doi:10.3748/wjg.v20.i22.6725.

Monsbakken, K. W., Vandvik, P. O., & Farup, P. G. (2005). Perceived food intolerance in subjects with irritable bowel syndrome – etiology, prevalence and consequences. *European Journal Of Clinical Nutrition, 60*, 667, doi:10.1038/sj.ejcn.1602367.

Nicholson, J. K., Connelly, J., Lindon, J. C., & Holmes, E. (2002). Metabonomics: a platform for studying drug toxicity and gene function. *Nature Reviews Drug Discovery, 1*, 153, doi:10.1038/nrd728.

Niddam, D. M., Tsai, S.-Y., Lu, C.-L., Ko, C.-W., & Hsieh, J.-C. (2011). Reduced Hippocampal Glutamate–Glutamine Levels in Irritable Bowel Syndrome: Preliminary Findings Using Magnetic Resonance Spectroscopy. [Functional GI Disorders]. *The American Journal Of Gastroenterology, 106*, 1503, doi:10.1038/ajg.2011.120.

Nimni, M. E. (1983). Collagen: structure, function, and metabolism in normal and fibrotic tissues. *Semin Arthritis Rheum, 13*(1), 1-86.

Obrenovich, M. E., Tima, M. A., Polinkovsky, A., Zhang, R., Emancipator, S. N., & Donskey, C. J. (2017). A Targeted Metabolomics Analysis Identifies Intestinal Microbiota-Derived Urinary Biomarkers of Colonization Resistance in Antibiotic-Treated Mice. [10.1128/AAC.00477-17]. *Antimicrobial Agents and Chemotherapy*.

Peironcely, J. E., Reijmers, T., Coulier, L., Bender, A., & Hankemeier, T. (2011). Understanding and Classifying Metabolite Space and Metabolite-Likeness. *PLoS ONE, 6*(12), e28966, doi:10.1371/journal.pone.0028966.

Pero, R. W., Lund, H., & Leanderson, T. (2009). Antioxidant metabolism induced by quinic acid. increased urinary excretion of tryptophan and nicotinamide. *Phytotherapy Research, 23*(3), 335-346, doi:10.1002/ptr.2628.

Playdon, M. C., Sampson, J. N., Cross, A. J., Sinha, R., Guertin, K. A., Moy, K. A., et al. (2016). Comparing metabolite profiles of habitual diet in serum and urine. *The American Journal of Clinical Nutrition, 104*(3), 776-789, doi:10.3945/ajcn.116.135301.

Qin, X. (2017). Damage of the Mucus Layer: The Possible Shared Critical Common Cause for Both Inflammatory Bowel Disease (IBD) and Irritable Bowel Syndrome (IBS). *Inflammatory Bowel Diseases, 23*(2), E11-E12, doi:10.1097/mib.0000000000001010.

Quigley, E. M. M. (2016). Editorial: serotonin and irritable bowel syndrome – reconciling pharmacological effects with basic biology. *Alimentary Pharmacology & Therapeutics, 43*(5), 644-646, doi:doi:10.1111/apt.13501.

Rattray, N. J. W., Deziel, N. C., Wallach, J. D., Khan, S. A., Vasiliou, V., Ioannidis, J. P. A., et al. (2018). Beyond genomics: understanding exposotypes through metabolomics. *Human Genomics, 12*, 4, doi:10.1186/s40246-018-0134-x.

Reap, E. A., & Lawson, J. W. (1990). Stimulation of the immune response by dimethylglycine, a nontoxic metabolite. *The Journal of laboratory and clinical medicine, 115*(4), 481-486.

Rodríguez, L. A. G., & Ruigómez, A. (1999). Increased risk of irritable bowel syndrome after bacterial gastroenteritis: cohort study. [10.1136/bmj.318.7183.565]. *BMJ, 318*(7183), 565.

Rothwell, J. A., Fillâtre, Y., Martin, J.-F., Lyan, B., Pujos-Guillot, E., Fezeu, L., et al. (2014). New Biomarkers of Coffee Consumption Identified by the Non-Targeted Metabolomic Profiling of Cohort Study Subjects. *PLoS One, 9*(4), e93474, doi:10.1371/journal.pone.0093474.

Ruttkies, C., Schymanski, E. L., Wolf, S., Hollender, J., & Neumann, S. (2016). MetFrag relaunched: incorporating strategies beyond in silico fragmentation. *Journal of Cheminformatics, 8*(1), 3, doi:10.1186/s13321-016-0115-9.

Savolainen, E. R., Miettinen, T. A., Pikkarainen, P., Salaspuro, M. P., & Kivirikko, K. I. (1983). Enzymes of collagen synthesis and type III procollagen aminopropetide in the evaluation of D-penicillamine and medroxyprogesterone treatments of primary biliary cirrhosis. [10.1136/gut.24.2.136]. *Gut, 24*(2), 136.

Simrén, M., Barbara, G., Flint, H. J., Spiegel, B. M. R., Spiller, R. C., Vanner, S., et al. (2013). Intestinal microbiota in functional bowel disorders: a Rome foundation report. *Gut, 62*(1), 159-176, doi:10.1136/gutjnl-2012-302167.

Simren, M., Mansson, A., Langkilde, A. M., Svedlund, J., Abrahamsson, H., Bengtsson, U., et al. (2001). Food-related gastrointestinal symptoms in the irritable bowel syndrome. *Digestion, 63*(2), 108-115, doi:10.1159/000051878.

Slattery, S. A., Niaz, O., Aziz, Q., Ford, A. C., & Farmer, A. D. (2015). Systematic review with meta-analysis: the prevalence of bile acid malabsorption in the irritable bowel syndrome with diarrhoea. *Alimentary Pharmacology & Therapeutics, 42*(1), 3-11, doi:doi:10.1111/apt.13227.

Snaith, P., & Zigmond, T. (1994). *The hospital anxiety and depression scale with the irritability-depression-anxiety scale and the Leeds situational anxiety scale manual.* : GL Assessment Ltd.

Spiller, R. C., Jenkins, D., Thornley, J. P., Hebden, J. M., Wright, T., Skinner, M., et al. (2000). Increased rectal mucosal enteroendocrine cells, T lymphocytes, and increased gut permeability following acuteCampylobacter enteritis and in post-dysenteric irritable bowel syndrome. *Gut, 47*(6), 804-811, doi:10.1136/gut.47.6.804.

Storey, J. D. (2003). The positive false discovery rate: a Bayesian interpretation and the q-value. *Ann. Statist., 31*(6), 2013-2035, doi:10.1214/aos/1074290335.

Sumner, L. W., Amberg, A., Barrett, D., Beale, M. H., Beger, R., Daykin, C. A., et al. (2007). Proposed minimum reporting standards for chemical analysis. [journal article]. *Metabolomics, 3*(3), 211-221, doi:10.1007/s11306-007-0082-2.

Thijssen, A. Y., Mujagic, Z., Jonkers, D. M. A. E., Ludidi, S., Keszthelyi, D., Hesselink, M. A., et al. (2016). Alterations in serotonin metabolism in the irritable bowel syndrome. *Alimentary Pharmacology & Therapeutics, 43*(2), 272-282, doi:doi:10.1111/apt.13459.

Törnblom, H., Lindberg, G., Nyberg, B., & Veress, B. (2002). Full-thickness biopsy of the jejunum reveals inflammation and enteric neuropathy in irritable bowel syndrome. *Gastroenterology, 123*(6), 1972-1979, doi:http://dx.doi.org/10.1053/gast.2002.37059.

Viant, M. R., Kurland, I. J., Jones, M. R., & Dunn, W. B. (2017). How close are we to complete annotation of metabolomes? *Current Opinion in Chemical Biology, 36*, 64-69, doi:https://doi.org/10.1016/j.cbpa.2017.01.001.

Walsh, M. C., Brennan, L., Malthouse, J. P. G., Roche, H. M., & Gibney, M. J. (2006). Effect of acute dietary standardization on the urinary, plasma, and salivary metabolomic profiles of healthy humans. *The American Journal of Clinical Nutrition, 84*(3), 531-539, doi:10.1093/ajcn/84.3.531.

Want, E. J., Wilson, I. D., Gika, H., Theodoridis, G., Plumb, R. S., Shockcor, J., et al. (2010). Global metabolic profiling procedures for urine using UPLC-MS. [10.1038/nprot.2010.50]. *Nature Protocols, 5*(6), 1005-1018.

Wess, L., Eastwood, M. A., Wess, T. J., Busuttil, A., & Miller, A. (1995). Cross linking of collagen is increased in colonic diverticulosis. [10.1136/gut.37.1.91]. *Gut, 37*(1), 91.

William, T. B. (1978). Structure and Biosynthesis of Connective Tissue Proteoglycans. In H. Martin (Ed.), *THE GLYCOCONJUGATES* (1 ed., Vol. 2): Academic Press, Inc.

Yamamoto, H., & Sasaki, K. (2017). Metabolomics-based approach for ranking the candidate structures of unidentified peaks in capillary electrophoresis time-of-flight mass spectrometry. *ELECTROPHORESIS, 38*(7), 1053-1059, doi:10.1002/elps.201600328.

Yamamoto, M., Ly, R., Gill, B., Zhu, Y., Moran-Mirabal, J., & Britz-McKibbin, P. (2016). Robust and High-Throughput Method for Anionic Metabolite Profiling: Preventing Polyimide Aminolysis and Capillary Breakages under Alkaline

Conditions in Capillary Electrophoresis-Mass Spectrometry. *Analytical Chemistry*, doi:10.1021/acs.analchem.6b03269.

Yano, Jessica M., Yu, K., Donaldson, Gregory P., Shastri, Gauri G., Ann, P., Ma, L., et al. (2015). Indigenous Bacteria from the Gut Microbiota Regulate Host Serotonin Biosynthesis. *Cell, 161*(2), 264-276, doi:10.1016/j.cell.2015.02.047.

Zheng, Y.-F., Yang, J., Zhao, X.-J., Feng, B., Kong, H.-W., Chen, Y.-J., et al. (2005). Urinary nucleosides as biological markers for patients with colorectal cancer. *World Journal of Gastroenterology : WJG, 11*(25), 3871-3876, doi:10.3748/wjg.v11.i25.3871.

Zhou, Q., Souba, W. W., Croce, C. M., & Verne, G. N. (2010). MicroRNA-29a Regulates Intestinal Membrane Permeability in Patients with Irritable Bowel Syndrome. *Gut, 59*(6), 775-784, doi:10.1136/gut.2009.181834.

3.7 Supplemental Information

Table S3.1 Summary of 143 urinary metabolites detected in IBS patients that are annotated based on their accurate mass (m/z), relative migration time (RMT), ionization mode (p = ESI+, n = ESI-), most likely molecular formula, compound name, confidence level for identification, and metabolite classification.

m/z:RMT:mode	Formula	ID	Confidence level [e]	Classification
73.0295:1.841: n	$C_3H_6O_2$	Propionic acid [b]	1	Carboxylic acid
87.0088:1.631: n	$C_3H_4O_3$	Pyruvic acid	1	Alpha-keto acid
87.0452:1.627: n	$C_4H_8O_2$	Butyric acid	1	Fatty acid
89.0244:1.141: n	$C_3H_6O_3$	Lactic acid	1	Hydroxy carboxylic acid
115.0037:1.982: n	$C_4H_4O_4$	Fumaric acid	1	Dicarboxylic acid
115.0401:1.042: n	$C_5H_8O_3$	Alpha-Ketoisovaleric acid	3	Short-chain keto acid
117.0193:1.845: n	$C_4H_6O_4$	Methylmalonic acid	1	Dicarboxylic acid
128.0353:1.019: n	$C_5H_7NO_3$	L-5-Oxoproline	1	Amino acid
130.051:0.966: n	$C_5H_9NO_3$	Propionylglycine	3	Amino acid derivative
131.035:1.628: n	$C_5H_8O_4$	Ethylmalonic acid	3	Fatty acid
132.0302:1.025: n	$C_4H_7NO_4$	Aspartic acid	1	Amino acid
133.0142:1.867: n	$C_4H_6O_5$	Malic acid	1	Beta-hydroxy acid
135.0299:0.994: n	$C_4H_8O_5$	Threonic acid	3	Sugar acid
141.0193:0.979: n	$C_6H_6O_4$	Sumiki's acid	3	Furoic acid
144.0299:1.001: n	$C_5H_7NO_4$	2-Hydroxy-5-oxoproline	3	Amino acid derivative
144.0666:0.915: n	$C_6H_{11}NO_3$	N-Butyrylglycine	3	Amino acid derivative
144.0666:0.930: n	$C_6H_{11}NO_3$	Unknown [a]	4	Unknown
145.0506:1.461: n	$C_6H_{10}O_4$	2-Methylglutaric acid	3	Methyl fatty acid
151.0401:0.937: n	$C_8H_8O_3$	3-Hydroxyphenylacetic acid	1	Hydroxy-benzenoid
156.0657:0.915: n	$C_7H_{11}NO_3$	Tiglylglycine	3	Amino acid derivative
160.0615:0.905: n	$C_6H_{11}NO_4$	Alpha-Aminoadipic acid	1	Alpha-amino acid
166.0146:1.558: n	$C_7H_5NO_4$	Quinolinic acid	3	Pyridinecarboxylic acids
167.0201:0.963: n	$C_5H_4N_4O_3$	Uric acid	1	Xanthine
172.9912:1.143: n	$C_6H_6O_4S$	Phenyl sulfate	2	Phenylsulfate
173.0092:2.124: n	$C_6H_6O_6$	Cis-Aconitic acid	3	Tricarboxylic acid
174.0408:1.477: n	$C_6H_9NO_5$	N-Acetylaspartic acid	3	Amino acid derivative
174.0561:0.903: n	$C_{10}H_9NO_2$	Indoleacetic acid	1	Indole-3-acetic acid
174.0561:0.925: n	$C_{10}H_9NO_2$	Unknown [a]	4	Unknown
175.0601:1.411: n	$C_7H_{12}O_5$	2-Isopropylmalic acid	3	Hydroxy fatty acid
178.051:0.912: n	$C_9H_9NO_3$	Hippuric acid	1	Hippuric acid
181.0355:0.932: n	$C_6H_6N_4O_3$	Methyluric acid	1	Xanthine
181.0506:0.889: n	$C_9H_{10}O_4$	3-(3-Hydroxyphenyl)-3-hydroxypropanoic acid (HPHPA)	3	Phenylpropanoic acids
182.0459:0.941: n	$C_8H_9NO_4$	4-Pyridoxic acid	1	Pyridinecarboxylic acid
184.0977:0.870: n	$C_9H_{15}NO_3$	Unknown [a]	4	Unknown
187.0071:1.062: n	$C_7H_8O_4S$	p-Cresol sulfate	1	Phenylsulfates
188.0353:0.923: n	$C_{10}H_7NO_3$	Kynurenic acid	1	Quinoline carboxylic acid
190.051:0.888: n	$C_{10}H_9NO_3$	5-Hydroxyindole acetic acid	1	Indole
191.0552:0.889: n	$C_7H_{12}O_6$	Quinic acid	1	Sugar acid
193.03574:0.875: n	$C_6H_{10}O_7$	Glucuronic acid	1	Glucuronic acid
195.0524: 0.877: n	$C_7H_8N_4O_3$	1,3 (or 1,7) Dimethyl uric acid [f]	2	Xanthine

197.0455:0.885: *n*	$C_9H_{10}O_5$	*Fatty acid* [a]	4	Fatty acid
201.112:0.821: *n*	$C_{10}H_{18}O_4$	2-Ethylsuberic acid	3	Fatty acid
204.0666:0.866: *n*	$C_{11}H_{11}NO_3$	Indole-3-lactic acid	1	Indole
211.0613:0.851: *n*	$C_{11}H_8N_4O$	*Unknown*[a]	4	*Unknown*
212.0023:1.025: *n*	$C_8H_7NO_4S$	Indoxyl sulfate	1	Arylsulfate
218.0495:0.854: *n*	$C_8H_{13}NO_4S$	*Unknown* [a]	4	*Unknown*
218.1034:0.827: *n*	$C_9H_{17}NO_5$	Pantothenic acid	4	Beta-amino acid
221.0746:0.838: *n*	$C_8H_{14}O_7$	Ethyl glucuronide	2	O-Glucuronide
225.0639:0.851: *n*	$C_8H_{10}N_4O_4$	5-acetylamino-6-formylamino-3-methyluracil	3	Hydroxypyrimidine
227.9968:0.980: *n*	$C_8H_7NO_5S$	*Unknown* [a]	4	*Unknown*
230.0129:0.930: *n*	$C_8H_9NO_5S$	Acetaminophen sulfate [d]	1	Phenylsulfate
235.0975:1.183: *n*	$C_{13}H_{16}O_4$	Carboxy-ibuprofen [d]	2	Phenylpropanoic acids
241.1193:0.815: *n*	$C_{11}H_{18}N_2O_4$	*Unknown* [a]	4	*Unknown*
243.0771:0.751: *n*	$C_{13}H_{12}N_2O_3$	Indolylacryloylglycine	3	Amino acid derivative
260.023:0.922: *n*	$C_9H_{11}NO_6S$	*Unknown* [a]	4	*Unknown*
263.1037:0.816: *n*	$C_{13}H_{16}N_2O_4$	Phenylacetylglutamine	1	Amino acid derivative
269.1500:0.794: *n*	$C_9H_{11}NO_6S$	*Unknown* [a]	4	*Unknown*
287.0227:0.893: *n*	$C_{11}H_{12}O_7S$	*Unknown* [a]	4	*Unknown*
290.0882:0.784: *n*	$C_{11}H_{17}NO_8$	2,3-Dehydro-2-deoxy-N-acetylneuraminic acid	3	Sialic acid
302.114:0.800: *n*	$C_{15}H_{17}N_3O_4$	Indoleacetyl glutamine	2	Amino acid derivative
308.0987:0.778: *n*	$C_{11}H_{19}NO_9$	N-Acetylneuraminic acid	1	N-Acylneuraminic acid
324.0936:0.786: *n*	$C_{11}H_{19}NO_{10}$	N-Glycolylneuraminic acid	2	Neuraminic acid
326.0881:0.769: *n*	$C_{14}H_{17}NO_8$	Acetaminophen glucuronide [d]	1	Phenolic glycoside
331.1757:0.766: *n*	$C_{15}H_{22}N_7O_2$	*Unknown* [a]	4	*Unknown*
350.1093:0.775: *n*	$C_{13}H_{21}NO_{10}$	N-Acetyl-9-O-acetylneuraminic acid	3	N-Acylneuraminic acid
381.1555:0.750: *n*	$C_{19}H_{26}O_8$	Ibuprofen glucuronide [d]	2	O-Glucuronide
412.8833:1.645: *n*	$C_{10}N_5O_8S_3$	*Unknown* [a]	4	*Unknown*
448.3068:0.730: *n*	$C_{26}H_{43}NO_5$	Glycoursodeoxycholic acid	3	Cholic acid
464.3018:0.722: *n*	$C_{26}H_{43}NO_6$	Glycocholic acid	3	Cholic acid
465.2483:0.734: *n*	$C_{25}H_{38}O_8$	Dihydroxytestosterone-glucuronide	2	Steroid glucuronide
473.1452:0.743: *n*	$C_{24}H_{26}O_{10}$	Enterolactone glucuronide		O-Glucuronide
481.2439:0.728: *n*	$C_{25}H_{38}O_9$	*Unknown* [a]	4	*Unknown*
525.2688:0.718: *n*	$C_{25}H_{40}N_3O_9$	*Unknown* [a]	4	*Unknown*
539.2493:0.718: *n*	$C_{27}H_{40}O_{11}$	Tetrahydroaldosterone-3-glucuronide	3	Steroid glucuronide
541.2649:0.714: *n*	$C_{27}H_{42}O_{11}$	Cortolone-3-glucuronide	3	Steroidal glycoside
543.2811: 0.710: *n*	C27H42O11	Cortol-3-glucuronide	3	Steroidal glycoside
632.2044:0.700: *n*	$C_{23}H_{39}NO_{19}$	3'-Sialyllactose	2	N-Acylneuraminic acid
673.2309: 0.695: *n*	$C_{25}H_{42}N_2O_{19}$	3-Sialyl-N-acetyllactosamine	3	N-Acylneuraminic acid
76.0393:0.642: *p*	$C_2H_5NO_2$	Glycine	1	Amino acid
76.0757:0.484: *p*	C_3H_9NO	Trimethylamine N-oxide	2	Amino oxide
90.055:0.551: *p*	$C_3H_7NO_2$	Alanine	1	Amino acid
104.0706:0.588: *p*	$C_4H_9NO_2$	GABA	1	Amino acid derivative
104.0706:0.909: *p*	$C_4H_9NO_2$	Dimethylglycine (DMG)	1	Alpha-amino acid
106.0499:0.805: *p*	$C_3H_7NO_3$	Serine	1	Amino acid
118.0611:0.629: *p*	$C_5H_{11}NO_2$	Valine	1	Amino acid
118.0863:0.945: *p*	$C_5H_{11}NO_2$	Betaine	1	Amino acid
120.0652:0.855: *p*	$C_4H_9NO_3$	Threonine	1	Amino acid
122.0270: 0.957: *p*	$C_3H_7NO_2S$	Cysteine	1	Amino acid
126.1026:0.615: *p*	$C_6H_{11}N_3$	Methylhistamine	3	2-arylethylamine

129.0659:0.672: *p*	$C_5H_8N_2O_2$	Dihydrothymine	3	Ureide
131.1179:0.643: *p*	$C_6H_{14}N_2O$	N-Acetylputrescine	3	Carboximidic acid
132.0768:0.682: *p*	$C_4H_9N_3O_2$	Creatine	1	Amino acid derivative
133.0608:0.857: *p*	$C_5H_{12}N_2O_2$	Ornithine	2	Amino acid
137.0457:1.131: *p*	$C_5H_4N_4O$	Hypoxanthine	1	Hypoxanthine
138.0550:0.861: *p*	$C_7H_7NO_2$	Trigonelline	1	Alkaloid
139.0503:0.638: *p*	$C_6H_6N_2O_2$	Urocanic acid	2	Amino acid derivative
141.0659: 0.625: *p*	$C_6H_8N_2O_2$	Iimidazole propionate	2	Imidazoyl carboxylic acid
144.1019:0.954: *p*	$C_7H_{13}NO_2$	Proline betaine	2	Amino acid derivative
146.0924:0.633: *p*	$C_5H_{11}N_3O_2$	*Unknown [a]*	4	*Unknown*
147.0764:0.886: *p*	$C_5H_{10}N_2O_3$	Glutamine	1	Amino acid
147.1128:0.515: *p*	$C_6H_{14}N_2O_2$	Lysine	1	Amino acid
148.0604:0.903: *p*	$C_5H_9NO_4$	Glutamic acid	2	Amino acid
149.0554: 0.887: *p*	$C_4H_8N_2O_4$	*Unknown [a]*	4	*Unknown*
150.0583: 0.867: *p*	$C_5H_{11}NO_2S$	Methionine	1	Amino acid
154.0499: 0.989: *p*	$C_7H_7NO_3$	3-Hydroxyanthranilic acid	3	Hydroxy benzoic acid
156.0767:0.554: *p*	$C_6H_9N_3O_2$	Histidine	1	Amino acid
162.1125: 0650: *p*	$C_7H_{15}NO_3$	Carnitine	1	Carnitine
164.0748:0664: *p*	$C_6H_{13}NO_2S$	S-Propyl-L-cysteine	3	Amino acid derivative
166.0723:0.639: *p*	$C_6H_7N_5O$	Methylguanine	3	6-oxopurine
170.0924:0.557: *p*	$C_7H_{11}N_3O_2$	1-Methylhistidine	1	Amino acid derivative
176.0658:0.813: *p*	$C_6H_9NO_5$	N-Acety-L-aspartic acid	3	Amino acid derivative
182.0809:0.942: *p*	$C_9H_{11}NO_3$	Tyrosine	1	Amino acid
189.1598:0.535: *p*	$C_9H_{20}N_2O_2$	N6,N6,N6-Trimethyl-L-lysine	1	Amino acid derivative
190.1191:0.926: *p*	$C_7H_{15}N_3O_3$	Homocitrulline	2	Amino acid
191.0661:1.02: *p*	$C_6H_{10}N_2O_5$	L-Beta-Aspartyl-L-glycine	3	Amino acid derivative
192.1131:0.775: *p*	$C_{10}H_{13}N_3O$	N'-Nitrosoanabasine	3	Alkaloid
193.0973:0.724: *p*	$C_{10}H_{12}N_2O_2$	Hydroxycotinine [d]	1	pyrrolidinylpyridine
195.0764:0.873: *p*	$C_9H_{10}N_2O_3$	4-Aminohippuric acid	3	Hippuric acid derivative
201.1597:0.685: *p*	$C_9H_{10}N_2O_3$	Pyridylacetylglycine	3	Amino acid derivative
204.1230:0.698: *p*	$C_9H_{17}NO_4$	Acetylcarnitine	1	Carnitine derivative
205.0972:0.953: *p*	$C_{11}H_{12}N_2O_2$	Tryptophan	1	Amino acid
209.0921:0.843: *p*	$C_{10}H_{12}N_2O_3$	Kynurenine [b]	2	Alkyl-phenylketone
217.1294:0.795: *p*	$C_8H_{16}N_4O_3$	Acetyl-arginine	3	Amino acid derivative
222.0796:0.769: *p*	$C_8H_{15}NO_4S$	5-(delta-carboxybutyl)homocysteine	3	Amino acid derivative
223.0747:0.801: *p*	$C_7H_{14}N_2O_4S$	Cystathionine	1	Amino acid derivative
226.0830:0.753: *p*	$C_9H_{13}N_3O_4$	Deoxycytidine	3	Pyrimidine 2'-deoxyribonucleoside
232.1543:0.738: *p*	$C_{11}H_{21}NO_4$	Butyrylcarnitine	2	Carnitine derivative
238.0916:1.12: *p*	$C_8H_{15}NO_7$	Xylosyl-serine	3	Amino acid derivative
241.0311:0.921: *p*	$C_6H_{12}N_2O_4S_2$	Cystine	1	Amino acid derivative
258.1084:0.811: *p*	$C_{10}H_{15}N_3O_5$	Methylcytidine	3	Pyrimidine nucleoside
258.1084:0.820: *p*	$C_8H_{13}N_6O_4$	*Unknown [a]*	4	*Unknown*
259.0918:0.831: *p*	$C_{10}H_{14}N_2O_6$	*Ribothymidine or imidazoleacetic acid ribose [f]*	3	Ribonucleoside
269.1238:0.879: *p*	$C_{11}H_{16}N_4O_4$	Acetylcarnosine	1	Peptide
276.1442:0.815: *p*	$C_{12}H_{21}NO_6$	Glutarylcarnitine	1	Carnitine derivative
282.1197:0.798: *p*	$C_{11}H_{15}N_5O_4$	1-Methyladenosine	1	Purine nucleoside
286.2013:0.817: *p*	$C_{15}H_{27}NO_4$	2-Octenoylcarnitine	1	Carnitine derivative
298.1146:1.103: *p*	$C_{11}H_{15}N_5O_5$	1-Methylguanosine	2	Purine nucleoside
302.2326:0.842: *p*	$C_{19}H_{29}N_2O$	*Unknown [a]*	4	*Unknown*
312.1297:1.100: *p*	$C_{12}H_{17}N_5O_5$	N2,N2-Dimethylguanosine	1	Purine nucleoside

325.1605: 0.720: *p*	$C_{12}H_{24}N_2O_8$	Galactosylhydroxylysine	1	Glycosyl-amino acid
367.1500: 1.092: *p*	$C_{17}H_{22}N_2O_7$	2-(α-D-Mannopyranosyl)-L-tryptophan	1	Amino acid derivative
487.2117:0.777: *p*	$C_{18}H_{34}N_2O_{13}$	Glucosylgalactosyl hydroxylysine	2	Glycosyl-amino acid

[a] *Unknown compounds are assigned with most probable formula based on their accurate mass.*

[b] *Low signal-to-noise ratio and not included in the final matrix*

[c] *Migration time differed from the chemical standard for originally assigned identity (Pantothenic acid).*

[d] *Not included in the final data matrix for statistical analysis because they were present in below 75% of samples in each group.*

[e] *Confidence levels indicate 1) Confidently identified with authentic chemical standards, 2) Putatively identified based on MS/MS fragmentation, accurate mass, and migration time 3) Putatively assigned 4) Unknown*

[f] *MS/MS spectral information was indecisive to choose one from these candidates*

Table S3.2. Significant urinary metabolites found by pair-wise comparison of ANCOVA between IBS and non-IBS/healthy controls, including between a second repeat urine samples collected for IBS patients. Creatinine (mM) was used instead of osmolality (mOsm/kg) to normalize data.

m/z: RMT	ID	*p*-value1	*q*-value2	FC-1[a]	*p*-value2	*q*-value2	FC-2[a]
174.0561: 0.903	Indoleacetic acid	2.25 E-04	5.60 E-03	0.62	2.65 E-03	0.0248	0.73
147.1128: 0.515	Lysine [c]	1.26 E-03	0.0187	2.18	> 0.05	> 0.05	1.44
132.0302: 1.025	Aspartate	7.45 E-03	> 0.05	0.66	0.0265	> 0.05	0.68
487.2117: 0.777	*O*-Glucosyl galactosyl-hydroxylysine [c]	9.98 E -03	> 0.05	1.30	6.53 E-03	0.0422	1.57
141.0659: 0.625	Imidazole propionate[b]	0.0127	> 0.05	1.18	0.0103	> 0.05	1.51
181.0355:0.932	Methyluric acid [b]	0.0145	> 0.05	0.72	0.0144	> 0.05	0.79
325.1605: 0.720	*O*-Galactosyl-hydroxylysine [c]	0.0267	> 0.05	1.24	9.78 E -03	> 0.05	1.42
104.0706: 0.909	*N-N*-Dimethyl glycine[c]	> 0.05	> 0.05	1.82	1.05 E-03	0.0162	1.61
147.0764: 0.886	Glutamine [c]	> 0.05	> 0.05	1.10	0.0179	> 0.05	1.52
312.1297: 1.102	*N,N*-Dimethyl guanosine [c]	> 0.05	> 0.05	1.15	0.0251	> 0.05	1.37
367.1500: 1.092	*C*-Mannopyranosyl-tryptophan [c]	> 0.05	> 0.05	1.27	> 0.05	> 0.05	1.54
106.0499: 0.805	Serine [c]	> 0.05	> 0.05	1.09	> 0.05	> 0.05	1.22
133.0608: 0.857	Ornithine [c]	> 0.05	> 0.05	1.15	> 0.05	> 0.05	1.32

[a] *Fold-change (FC) calculated by median untransformed RPAs of IBS/non-IBS.*

[b] *Putative metabolite identification based on MS/MS spectra without chemical standard.*

[c] *Significant in osmolality normalized data.*
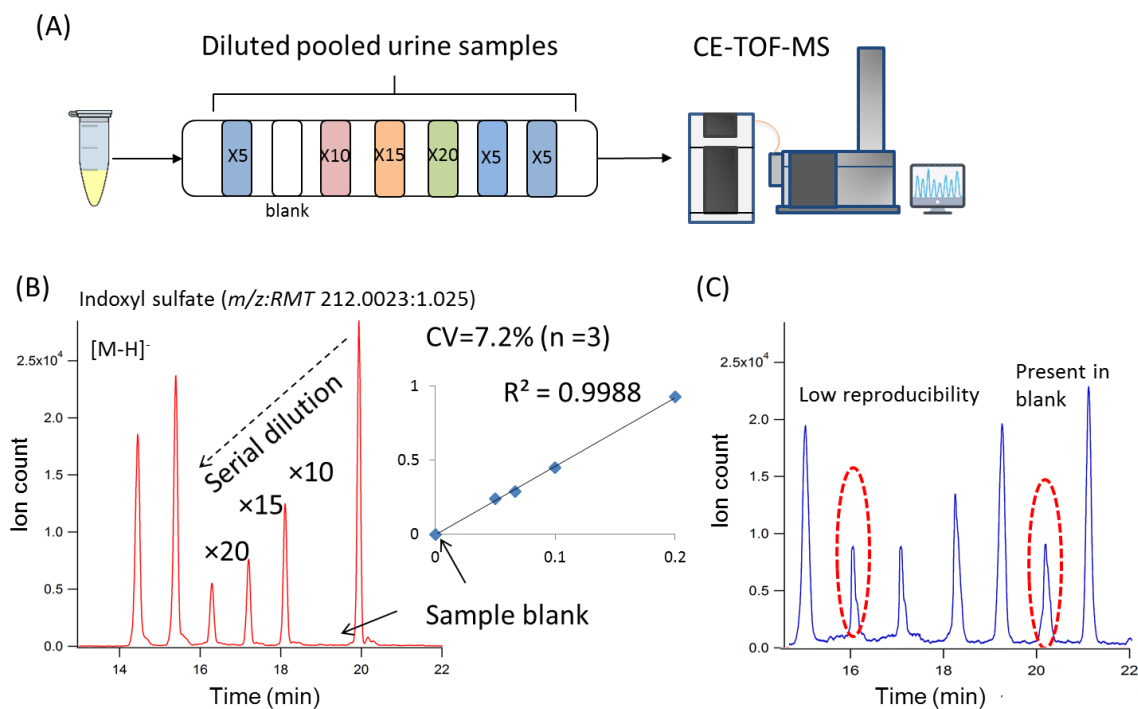
Figure S3.1 (a) Multiplexed separations using a 7-sample serial injection sequence by MSI-CE-MS based on a dilution trend filter of a pooled urine QC ($n = 76$) analyzed within a single run. The least diluted urine sample was injected in triplicate, whereas a blank sample containing only internal standards in deionized water was injected within the same run to generate a distinctive temporal signal pattern in ESI-MS reflecting the serial injection order/dilution. **(b)** Representative extracted ion electropherogram (EIE) of an authentic urinary metabolite, indoxyl sulfate based on its characteristic *m/z*:RMT (212.0023:1.025 ESI-) which fully satisfies a serial dilution trend filter based on a reproducible signal (CV = 7.2%, *n*=3), no signal measured in the blank, and a linear decrease in relative ion response upon dilution ($R^2 = 0.999$). **(c)** Representative EIE of a background molecular feature (*m/z*:RMT 328.0464: 1.105 ESI-) which does not show the dilution trend while having poor precision (CV = 43%) with a signal detected in the blank position.
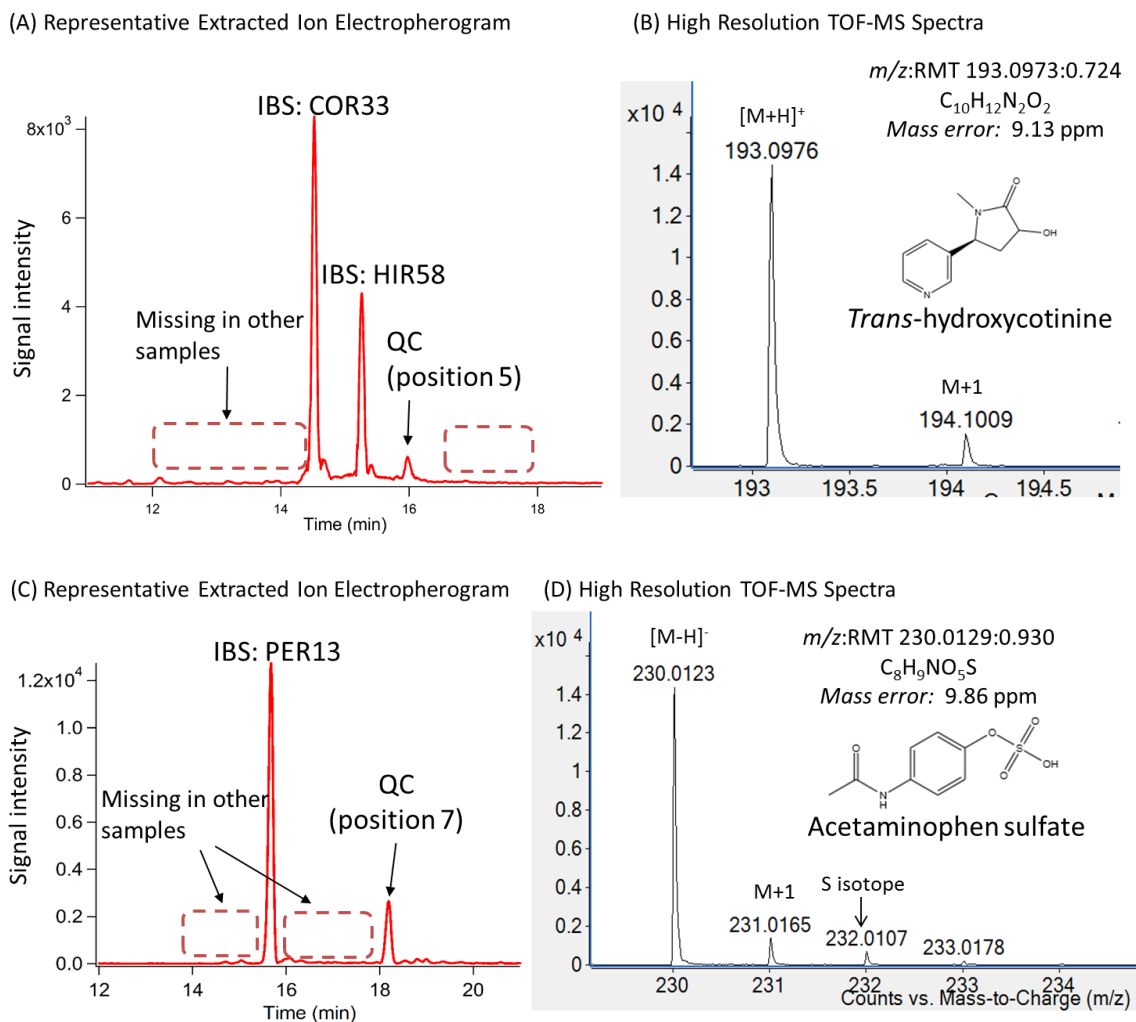
Figure S3.2 Examples of infrequently detected urinary metabolites among IBS patients (specific code assigned to each sample is shown) that were removed from the final metabolomics data matrix. Representative extracted ion electropherogram (EIE) and high resolution TOF-MS spectra of over-the-counter drug metabolites detected in a sub-set of urine samples from IBS patients, including **(a, b)** *trans*-hydroxycotinine (*m/z:*RMT 193.0973:0.724) that is a long-lived metabolite of nicotine (*e.g.,* tobacco exposure; detected in only two of the six urine samples analyzed in this run together with pooled QC sample) and **(c, d)** acetaminophen sulfate (*m/z:*RMT 230.0129:0.930) that is a secondary metabolite of acetaminophen and a commonly used analgesic (*i.e.,* detected in only one of the six urine samples analyzed in this run together with pooled QC sample).
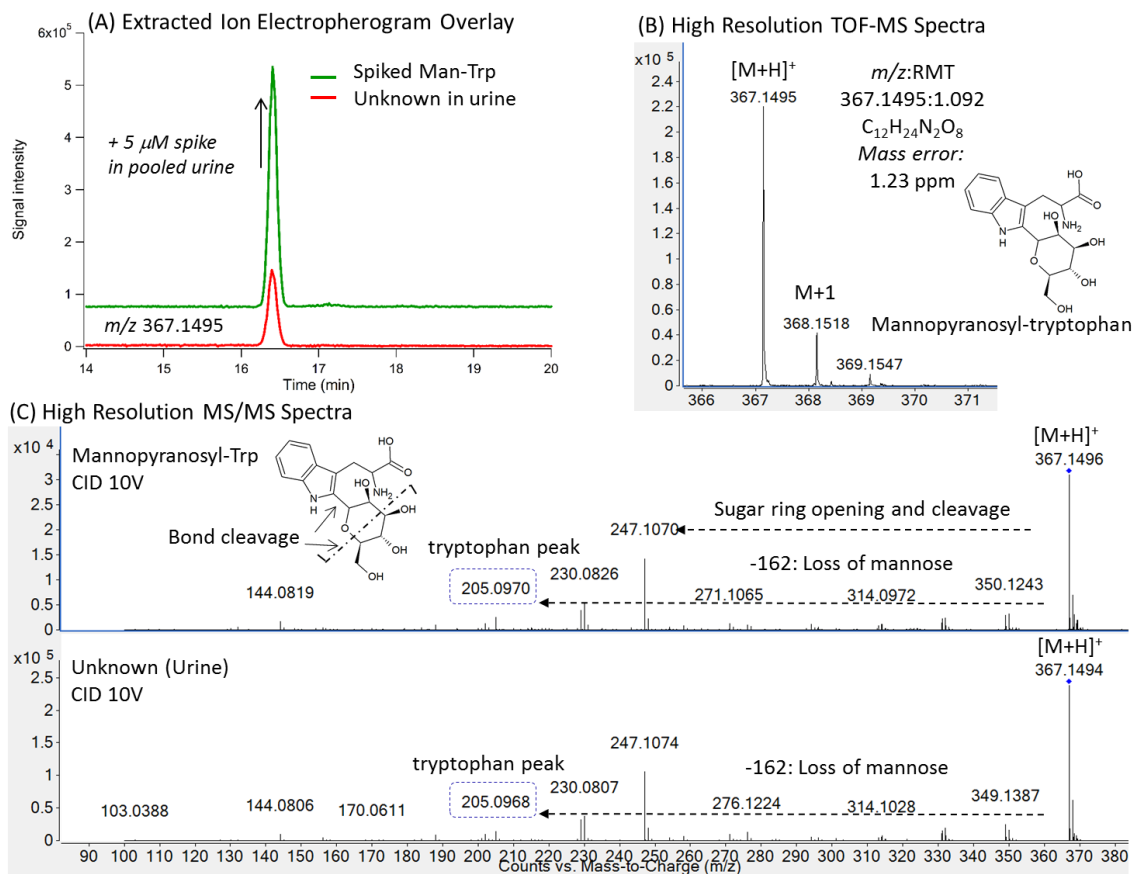
Figure S3.3 Structural elucidation (level 1) of an unknown cation excreted at significantly elevated levels in IBS patients as compared to healthy controls as annotated by its *m/z*:RMT 367.1605:0.720. **(a)** Extracted ion electropherogram overlay of a single injection of a pooled urine sample and the same sample spiked with an authentic chemical standard of *C*-mannopyranosyl-tryptophan confirming its co-migration. **(b)** High resolution TOF-MS spectrum depicts the accurate mass of its protonated molecular ion (MH⁺) and an isotope pattern that confirms detection of a singly charged ion as required to determine its most likely molecular formula with low mass error (< 2 ppm) **(c)** Collisional-induced dissociation experiments (at 10V) used to compare MS/MS spectra of the authentic standard with unknown cation present in pooled urine sample. Arrows are pointing at the proposed bond cleavage sites of the precursor ion as reflected by matching fragment ions and neutral losses generated for standard as compared to unknown ion in urine.

Figure S3.4 Putative identification (level 2) of an unknown cation excreted at significantly elevated levels in IBS patients as compared to healthy controls as annotated by its *m/z*:RMT 141.0657:0.734. **(a)** Extracted ion electropherogram overlay of a single injection of a pooled urine sample and the same sample spiked with an authentic chemical standard of methylimidazole acetate that did not confirm its co-migration, which suggested its likely positional isomer, namely imidazole propionate (standard not available) due to its similar MS/MS spectra. **(b)** High resolution TOF-MS spectrum depicts the accurate mass of its protonated molecular ion (MH$^+$) and an isotope pattern that confirms detection of a singly charged ion as required to determine its most likely molecular formula with low mass error (< 2 ppm) **(c)** Collisional-induced dissociation experiments (at 20V) used to compare MS/MS spectra of the authentic standard with unknown cation present in pooled urine sample highlighting similar fragmentation patterns despite their different migration times. Arrows are pointing at the proposed bond cleavage sites of the precursor ion as reflected by matching fragment ions and neutral losses generated for standard as compared to unknown ion in urine.

147

(A) Extracted Ion Electropherogram Overlay

(B) High Resolution TOF-MS Spectra

Spiked N'N-dimethylguanosine

Unknown in urine

+ 10 µM spike in pooled urine

m/z 312.1297

m/z 312.1297

N'N-Dimethylguanosine

[M+H]+ 312.1297

m/z:RMT 312.1297:1.100
$C_{12}H_{17}N_5O_5$
Mass error: 6.08 ppm

M+1 313.1334

(C) High Resolution MS/MS Spectra

N'N-Dimethylguanosine CID 10V

180.0882

Bond cleavage

84.9606   115.0389

Loss of ribose

[M+H]+ 312.1297

Unknown (urine) CID 10V

180.0881

91.0534

Loss of ribose

[M+H]+ 312.1297

Counts vs. Mass-to-Charge (m/z)

Figure S3.5 Structural elucidation (level 1) of an unknown cation excreted at significantly elevated levels in IBS patients as compared to healthy controls as annotated by its *m/z*:RMT 312.1297:1.100. **(a)** Extracted ion electropherogram overlay of a single injection of a pooled urine sample and the same sample spiked with an authentic chemical standard of *C*-mannopyranosyl-tryptophan confirming its co-migration. **(b)** High resolution TOF-MS spectrum depicts the accurate mass of its protonated molecular ion (MH+) and an isotope pattern that confirms detection of a singly charged ion as required to determine its most likely molecular formula with low mass error (< 2 ppm) **(c)** Collisional-induced dissociation experiments (at 10V) used to compare MS/MS spectra of the authentic standard with unknown cation present in pooled urine sample. Arrows are pointing at the proposed bond cleavage sites of the precursor ion as reflected by matching fragment ions and neutral losses generated for standard as compared to unknown ion in urine.

Figure S3.6. Box-whisker plots of significant features found in urine samples collected from IBS patients and non-IBS control. Y-axis represents log-transformed relative peak area, and *q*-value > 0.05 is considered significant after FDR adjustment.
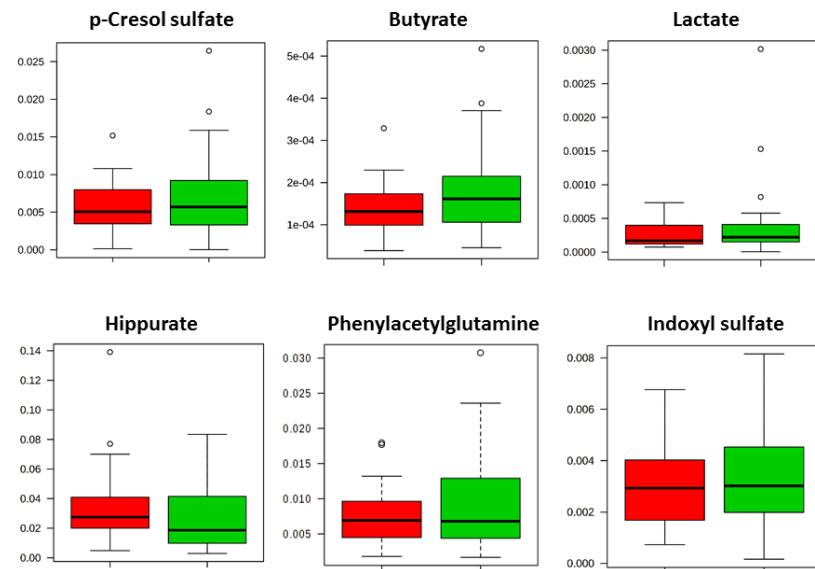
Figure S3.7. Boxplots of commonly known microbial metabolites found in human urine. None of these metabolites show statistically significant difference based on a Mann-Whitney U-test.

Chapter IV

Metabolomics and Gut Microbiome Reveal New Insights into Pediatric Inflammatory Bowel Disease: Urinary Biomarkers for Differential Diagnosis of Crohn's Disease and Ulcerative Colitis

Thesis chapter is derived from a manuscript in preparation for submission to Nature Metabolism, where I am the first author and my research supervisor is the corresponding author:

Mai Yamamoto, Lara Hart, Nikhil Pai, Jasmine Chong, Philip Britz-McKibbin *Nature Metabolism* (to be submitted).

M.Y. conducted all of the instrumental optimization, sample analysis, data processing and interpretation on data for urinary and fecal metabolomics, and conducted part of microbiome metagenomics data analysis under the supervision of P.B.M. M.Y. also wrote an initial manuscript draft used for publication. J.C. contributed on additional microbiome data analysis, and all other co-authors provided feedback on metabolomics data interpretation and manuscript draft.

Chapter IV: Metabolomics and Gut Microbiome Reveal New Insights into Pediatric Inflammatory Bowel Disease: Urinary Biomarkers for Differential Diagnosis of Crohn's Disease and Ulcerative Colitis

4.1 Abstract

Inflammatory bowel disease (IBD) is a chronic autoimmune disorder of the gastrointestinal tract that is characterized by lifelong episodes of rectal bleeding, weight loss, and abdominal pain. The incidence of IBD is increasing worldwide, particularly among children. However, differential diagnosis of Crohn's disease (CD) and ulcerative colitis (UC) is challenging when relying on invasive endoscopic imaging and mucosal biopsies that are also prone to indeterminate results. Herein, an integrative approach for urinary and fecal metabolic phenotyping using multisegment injection-capillary electrophoresis-mass spectrometry (MSI-CE-MS) together with functional microbiome analysis was performed to elucidate the underlying mechanisms distinguishing major pediatric IBD sub-types. This pilot study was comprised mainly of newly diagnosed and sex-matched children with CD (n=19) and UC (n=11) at early stages of disease progression with a mean age of 13 y. Non-targeted metabolite profiling of urine using MSI-CE-MS revealed distinct differences ($p < 0.05$) in tryptophan/serotonin metabolism and several human-microbial co-metabolites, whereas CD patients had higher excretion of phenylacetylglutamine, phenylsulfate, indoxylsulfate, hydroxyindole sulfate, and indole 3-acetic acid glucuronide as compared to UC cases. Conversely, affected UC children had higher excretion of urinary kynurenine, hypoxanthine, serine and threonine when normalized to either osmolality or creatinine. Excellent discrimination of CD from UC patients was realized based on the ratio in urinary serine/indoxylsulfate (AUC = 0.974; p = 2.73 E-5) with stability studies demonstrating reliable measurements for over 2 days upon refrigeration. Similarly, metabolic phenotyping of matching stool extracts demonstrated that CD patients had lower elimination of tryptophan, choline and lactate, but higher trimethyllysine, cholic acid and butyric acid as compared to UC patients when normalized to dried stool weight. Although there were no overall differences in gut microbiota alpha diversity, 16S rRNA gene profiling results indicated that UC patients are characterized by mucus depletion and colonization by opportunistic pathogens (*e.g., Klebsiella*, *Streptococcus*), whereas CD patients have enriched bacterial activities in tryptophan and phenylalanine metabolic pathways corresponding to elevated concentrations in indole and phenol derived metabolites measured in urine. Overall, a panel of urinary metabolites offers the potential for non-invasive screening and accurate diagnosis of pediatric IBD not feasible by non-specific inflammatory biomarkers (*e.g.,* serum C-reactive protein, fecal calprotectin), which may also be applied to monitor for therapeutic responses to novel dietary, biologic and/or immunomodulating interventions.

4.2 Introduction

Inflammatory bowel disease (IBD), including Crohn's disease (CD) and ulcerative colitis (UC), is a serious chronic gastrointestinal (GI) condition that carries significant burden of disease in pediatrics. About 20-30% of IBD patients experience their first symptoms before the age of 18 years.[1] The remitting and relapsing nature of their GI inflammation results in episodes of debilitating abdominal pain, bloody stools, fatigue, malabsorption and weight loss.[2] Children who develop IBD are uniquely affected by its unpredictable course and atypical symptoms, which impairs normal growth and development, including a higher risk for depression.[3] The early onset of IBD significantly increases the long-term complications of colitis, such as fibrosis, stricturing and the need for intestinal resections. As a result, early diagnosis and treatment is critical to manage inflammation while promoting physical growth and psychosocial well-being in severe pediatric IBD cases using biologic therapies and immunomodulators to prolong remission of active disease symptoms without corticosteroid treatment.[4] However, delineating CD from UC is diagnostically challenging yet critical for optimal treatment decisions on individual patients. Although pediatric CD is more prevalent than UC, invasive tests are needed for reliable diagnosis based on disease location and extent of inflammation due to overlapping clinical presentations.[5] In CD, inflammation appears anywhere along the digestive tract but most often in ileal and colonic regions of the GI tract with transmural inflammation.[6] For UC, inflammation is typically confined to the mucosal and sub-mucosal layers of the colon, starting from the rectum and progressing proximally.[6] While anatomic differences help differentiate CD from UC, there are no validated serological or stool biochemical markers (*i.e.,* biomarkers) that can accurately distinguish between these two major disease subtypes.[7] Thus, repeat endoscopic imaging and histopathological assessment of IBD activity and mucosal healing are often necessary for reliable diagnosis while evaluating treatment responses to therapy despite their invasiveness. These procedures are also costly, may contribute to delays in diagnosis due to tissue biopsy handling and do not always lead to conclusive results notably for early onset IBD, including indeterminate colitis highlighting the complex disease spectrum in affected children.[8-10] These diagnostic dilemmas reflect the poorly understood etiology of pediatric IBD who often

exhibit a more aggressive or complicated disease courses than adults[11] that is mediated by a complex interplay of genetic, immunological, and environmental factors.

Metabolomics offers a systemic approach for characterizing the molecular phenotype of an organism since metabolites represent real-world end-products of gene expression and protein activity, as well as bioactive molecules derived from habitual diet and lifelong exposures.[12] To date, metabolomics studies have been largely focused on understanding the pathophysiology of IBD by comparing metabolic differences between IBD patients and healthy controls when analyzing serum/plasma, urine or stool specimens.[13] Previous studies have reported significant changes in microbial derived metabolites reflecting underlying dysbiosis of commensal gut microbiota communities, such as secondary bile acids, butyric acid, and/or trimethylamine in fecal water extracts when using nuclear magnetic resonance (NMR) and liquid-chromatography-mass spectrometry (LC-MS).[14-16] As expected, given the clinical manifestations of IBD, such as fatigue and weight loss, down-regulated amino acid metabolism and organic acids from the citric acid cycle reflect impairments in energy homeostasis.[17-19] In contrast to clear differences between healthy controls and patients with IBD, differentiation between CD and UC based on metabolic profiles have proven far more challenging often with contradictory results. Several studies reported significantly lower levels of human-microbial co-metabolites, including *p*-cresol sulfate and hippuric acid in urine samples from CD patients as compared to UC and healthy controls.[20, 21] In contrast, no significant differences were reported in an independent cohort of CD and UC patients in another study despite using the same sample workup protocol and analytical platform (NMR) for metabolomic studies.[22] Similarly in stool samples, significantly reduced levels of short-chain fatty acids (*e.g.,* butyric acid, propionic acid) have been found in adult CD as compared to UC patients and healthy controls,[14] however these same stool-derived metabolites failed to differentiate CD from UC in a recent study that focused on treatment naïve pediatric IBD patients.[23] Unlike urine specimens, stool is a highly heterogeneous mixture of bacteria, host cells, protein, and undigested food components (*i.e.,* fiber, unabsorbed nutrients, waste products). Additionally, there are currently no standardized methods developed for routine stool collection and sample workup as reflected in the wide

variety of extraction methods used for stool analysis, including solvent types, homogenization and filtration methods, as well as data normalization as discussed in a recent review of stool sample preparation for metabolomics.[24] Delays to storage of fecal specimens and the lack of standardized collection and processing methods likely contribute to the many discrepancies reported among stool metabolome studies to date, which are strongly dependent on the intrinsic chemical stability of metabolites and the activity of bacteria in stool.

Adult IBD patients sometimes require intestinal resections for treatment or management of their disease, which induces drastic changes in the composition of the gut microbiome.[25] On the other hand, pediatric populations are disproportionately characterized by patients who are relatively treatment naïve, which may contribute to less confounding caused by long-term pharmaceutical or surgical interventions; this may be another important reason why the handful of pediatric studies are not consistent with different cohorts of primarily adult IBD patients. Additionally, the importance of sample quality especially those biospecimens with high biological activity, such as stool, is rarely investigated in recently published studies.[26, 27] Lastly, integration of microbiome metagenomic data, especially functional metagenomics investigations, is still in its infancy despite the prominent focus placed on the dynamic interplay of host-microbial metabolism. In this context, our work focuses on the characterization of the metabolic phenotype of a cohort of pediatric IBD patients diagnosed early in their disease progression, largely before the onset of medications or surgeries. Thus, the study design and patient cohort is more likely to reveal authentic biological differences in the pathophysiology between CD and UC at earlier stages of development with fewer confounding variables. Additionally, optimal collection and sample workup of matching urine and stool specimens from each patient provides complementary insights into aberrant metabolism while identifying putative biomarker candidates that enable differential diagnosis of major IBD sub-types in affected children with high accuracy. Optimization of a stool extraction procedure together with implementation of a validated multiplexed separation platform and data workflow for biomarker discovery with stringent quality control (QC) is also outlined.[28-30] Additionally, several unknown metabolites of clinical significance in both urine and stool extracts were identified for the first time when

using multisegment injection-capillary electrophoresis-mass spectrometry (MSI-CE-MS) in conjunction with high resolution MS/MS and *in silico* fragmentation modeling.[31] Our metabolomic studies are further corroborated by the identification of differentially abundant bacterial taxa and their predicted functional potential using 16S rRNA gene analyses. In addition, we have conducted stability studies of lead biomarker candidates identified in urine and stool specimens stored under different conditions to validate our findings while showcasing their potential for classifying UC and CD pediatric patients non-invasively.

4.3 Experimental section

4.3.1 Chemicals and reagents. All chemicals were purchased from Sigma Aldrich (St. Louis, MO, USA) unless otherwise stated. For preparation of buffer and sheath liquid, ultra-grade LC-MS solvents (water, methanol and acetonitrile) obtained from Caledon Laboratories Ltd. (Georgetown, ON, Canada) were used. Stock solutions for calibrants, sheath liquid and buffer solutions were all prepared in deionized water (Barnstead EASY-pure II LF ultrapure water system (Dubuque, IA, USA).

4.3.2 Pediatric IBD study cohort. This study was approved by the Hamilton Integrated Research Ethics Board (#15-365) and parental consent was obtained for all the participants. The study enrolled children from 5 to 18 years old who had been diagnosed with IBD by endoscopy, histology and radiography at McMaster Children's Hospital. Patients were included if they were admitted to hospital to be initiated on exclusive enteral nutrition therapy, or intravenous corticosteroids for induction of remission of CD or UC. Patients were excluded if they were younger than 5 years, received antibiotic therapy, or did not require admission to hospital. None of these patients have undergone resection surgery.

4.3.3 Urine and fecal sample collection, storage and workup procedure. All urine and stool samples included in this study were collected prior to induction therapy (*i.e.* corticosteroid or exclusive enteral nutrition) at McMaster Children's Hospital. Single-spot urine samples were collected in the morning and did not necessarily represent the first

morning urine void. Following collection, 1 mM of sodium azide was added to all urine samples as an antimicrobial preservative and then samples were stored in a fridge before being transferred to a freezer at -80 °C. All thawed urine samples (on ice) were prepared by a simple dilution (from 5 to 10-fold) using 25 μL of urine in ultra-grade LC-MS water containing the internal standards, 3-chloro-*L*-tyrosine (Cl-Tyr) and sodium 2-naphthalenesulfonate (NMS) to the final internal standards concentration of 10 μM each, which was followed by mixing using a vortex for 30 s. Also, stool specimens were aliquoted into two tubes at the clinic, one for metabolomics analysis and a second for microbiome analysis, into 1.5 mL centrifuge tubes and then placed in a freezer at -80°C at the earliest time possible. For metabolomics analysis, frozen stool samples were transferred to 15 mL Falcon tubes, tube caps were replaced with two layers of Kimwipes, which were secured with rubber bands, and samples were lyophilized (Labconco FreeZone Freeze Dry System; MO, USA) over approximately two days. Freeze-dried stool samples (*i.e.,* powder) were then weighed out (15-20 mg) accurately on an electronic balance prior to performing a modified Bligh Dyer extraction.[32] After ensuring that all samples are completely dry, freeze dried stool was mixed with pre-chilled methanol, deionized water and chloroform at 4:3.6:4 ratio with a total volume of 424 μL. Methanol was added to increase solubility of polar but relatively hydrophobic metabolites such as bile acids, and to precipitate proteins upon centrifugation.[33] Also, 4-fluoro-*L*-phenylalanine (10 μM F-Phe as final concentration) was used as a recovery standard to evaluate extraction efficiency and included in the deionized water for all stool samples at this step. This mixture was vortexed for 10 min and centrifuged for 20 min at 450 *g*. Subsequently, the upper aqueous layer was transferred to a separate clean tube, and then the process was repeated the second time to maximize recovery of metabolites. Finally, the two stool extracts were then combined and stored at -80 °C until analysis. On the day of analysis, extracts were slowly thawed on ice and mixed with deionized water containing internal standards (10 μM), namely 3-chloro-*L*-tyrosine (Cl-Tyr) and naphthalene monosulfonic acid (NMS) for data normalization in positive and negative ion mode detection, respectively resulting in an overall dilution of 5- and 10-fold for about 15 mg of lyophilized stool. Prior to individual sample analysis, a subset (*n* = 24) of lyophilized stool samples were homogenized and separately extracted as six technical

replicates for evaluating the reproducibility of the stool extraction protocol. A pooled quality control (QC) sample was also prepared from a sub-set of urine and stool extract samples (n = 30 for urine, n = 24 for stool) for the purpose of monitoring instrumental signal drift throughout the analysis.

4.3.4 Urinary osmolality and creatinine measurements. Osmolality was measured using Advanced Micro-Osmometer 3300 (Fisher Scientific Company). Measurement for all IBD urine samples were done in one day and Clinitrol 290 reference solution was measured intermittently every ten urine samples to ensure the accuracy of measurement. Reference solution read from 288 to 292 mOsm (actual osmolality: 290 mOsm) throughout the analysis (n = 8, average reading = 290 mOsm). All metabolite responses from single-spot urine samples were normalized to osmolality in order to correct for between-subject differences in hydration status. Urinary creatinine concentrations were also measured for all urine samples in this study using MSI-CE-MS. A seven-point calibration curve for creatinine was acquired in triplicate from 200 to 3000 µM with excellent linearity (R^2 = 0.999), where the integrated peak area was normalized to an internal standard (10 µM Cl-Tyr).

4.3.5 Urinary and stool metabolome stability studies. Five random single-spot urine samples were collected from healthy volunteers ranging from 25 to 30 years old, including two males and three females. Each sample was placed on ice and a pooled sample was prepared within 1 hr upon initial urine collection. Aliquots of this pooled urine sample were stored at either room temperature (~ 22 °C) or in a fridge (4 °C) for 6, 12, 24, 36 and 48 h, which were performed in triplicate. After assigned storage duration, 1 mM sodium azide was then added and samples were stored at -80 °C prior to analysis. Six urine aliquots of a pooled urine sample were also prepared as a control and immediately transferred to a freezer at -80 °C after addition of 1 mM sodium azide. An additional three urine aliquots were transferred to the freezer without sodium azide as a negative control. The same dilution and analysis protocol as described for IBD urine samples were performed. A representative stool sample was also collected from healthy six year old twin brothers upon receiving parental consent. Each stool sample was homogenized with sterile spatula and transferred into 18

tubes for different conditions in duplicates: control (-80 °C freezer), freezer (-20 °C) for 48 h, fridge (4 °C) for 2, 4, 8 and 48 h, and room temperature (~ 22 °C) for 2, 4, and 8 h of storage. Initial sample processing was completed within 30 min. After assigned storage duration, they were transferred to a freezer at -80 °C followed by lyophilisation, extraction and analysis. The same extraction protocol as freeze-dried IBD stool samples was also applied.

4.3.6 High throughput metabolite screening of urine and stool extract by MSI-CE-MS

MSI-CE-MS experiments were performed as previously described[34] on an Agilent G7100A CE system (Agilent Technologies Inc., Mississauga, ON, Canada) equipped with a coaxial sheath liquid Jetstream electrospray ion source with heated nitrogen gas to an Agilent 6550 iFunnel Q-TOF-MS system. Separations were performed using an uncoated fused silica capillary (Polymicro Technologies, AZ, USA) with an inner diameter of 50 μm, outer diameter of 360 μm, and total length of 110 cm using an applied voltage of 30 kV at 25 °C. Each sample was analyzed by MSI-CE-MS under two conditions based on a background electrolyte (BGE) comprised of 1.0 M formic acid with 15 % *v/v* acetonitrile (pH = 1.80) and 50 mM ammonium bicarbonate (pH = 8.50) for positive and negative ion mode detection, respectively. The sheath liquid composition for electrospray formation when using the coaxial sheath liquid interface in CE-MS was comprised of 60% *v/v* methanol with 0.1% *v/v* formic acid for analysis of cationic metabolites under positive ion mode detection, whereas 50% *v/v* MeOH was used as the sheath liquid for anionic metabolites under negative ion mode detection. Prior to sample injection, the capillary was conditioned with BGE for 15 min to ensure adequate equilibration. A seven serial sample injection format was used for multiplexed separations by MSI-CE-MS, which utilized an alternating hydrodynamic injection sequence of 5 s (at 100 mbar) for each sample followed by a 40 s (at 100 mbar) of BGE that served as spacer plug between each pair of diluted urine or stool extract sample. A pressure gradient was also applied during voltage application when analyzing anionic metabolites under negative ion mode detection to reduce total analysis times within 45 min, which comprised by up to 72 mbar or 7.2 kPa pressure application over 38 min (2 mbar or 0.2 kPa increase every 2-3 min).

Temporal signal pattern recognition using MSI-CE-MS was applied in this work to ensure correct sample assignment to each electropherogram peak as described elsewhere.[28] Briefly, all samples from a longitudinal study of the therapeutic effects of exclusive enteral nutrition for pediatric UC and CD patients were fully randomized and analyzed. Briefly, three samples (urine or fecal extract) were injected in duplicate and sample pairs were then diluted using a characteristic pattern within the run (*i.e.,* 1:1, 1:2 or 2:1) to facilitate identification of sample position especially when a metabolite is not consistently detected in all samples analyzed within the same run. Additionally, a QC (*i.e.,* pooled urine or fecal extract) was included in every run, which is essential for evaluating system stability over time resulting from non-biological experimental variation. In this case, QC was injected randomly at a position 1, 3, 5, or 7 in all runs performed by MSI-CE-MS. Q-TOF-MS was operated using full-scan data acquisition (TOF-MS mode) when performing nontargeted metabolite screening under positive and negative-ion modes over a mass range of *m/z* 50-1700 with an acquisition rate of 500 ms/spectrum. The ESI conditions were Vcap=3500 V, nozzle voltage=2000 V, nebulizer gas=8 psi, sheath gas=3.5 L/min at 200 °C and drying gas=16 L/min at 200 °C for both ionization modes. Furthermore, the MS voltage settings were fragmentor=120V, skimmer=65V and Oct1 RF=750 V. The Q-TOF-MS was used for collisional-induced dissociation (CID) experiments for metabolite identification when using both auto MS/MS and targeted MS/MS modes over a mass range of *m/z* 50-1700 with a cycle time of 3.1 s in conjunction with collisional voltages set at 10, 20 or 30V and 40 V for a precursor ion. A combination of deposited MS/MS spectral databases (*e.g.,* HMDB),[35] *in-silico* fragmentation (*e.g.,* MetFrag),[31] and manual annotation was used for MS/MS spectra interpretation.

4.3.7 Metabolomics data processing and statistical analysis. The initial raw data included all data from a longitudinal study of treatment effects (corticosteroid or exclusive enteral nutrition) with samples collected different time points. For this study, only the baseline (pre-treatment) data were selected and processed. Raw data (.d format) was processed using Mass

Hunter Workstation Software (Qualitative Analysis, version B.6.00, Agilent Technologies, 2012). Initial feature detection and identification was performed using Mass Hunter Molecular Feature Extractor, Molecular Formula Generator tools and an in-house compound database. Molecular features were extracted using a 10 ppm mass window and ions were annotated by their accurate mass (m/z), relative migration time (RMT) as compared to an internal standard, and ionization mode used for detection. Peak smoothing was performed using a quadratic/cubic Savitzky-Golay function (15 points) prior to peak integration. Peak areas and migration times for all molecular features and internal standards were transferred to Excel (Microsoft Office) and saved as .csv file. A list of authentic stool metabolites were initially curated using dilution trend filter as previously described.[34] Briefly, when a blank sample and a serially diluted QC sample were co-injected within the same run, temporal pattern recognition was used to reject molecular features that were highly variable based on triplicate injections at the same dilution ($CV > 30\%$), and present in the blank sample (*i.e.,* background ions) from the data matrix. Subsequently, compounds in the list were further filtered by rejecting compounds with high variability during sample analysis (CV in QC > 40 %) and frequently missing (present in < 75 % of total samples). Subsequently, R program (v. 3.5.1) was used for pre-processing of the data matrix, which included calculation of relative peak area (RPA) and RMT, coefficient of variation (CV) from QC samples in every run, removal of compounds with high variance (CV > 40 %) from QCs, and evaluation of frequently missing compounds. Subsequently, values for each metabolite was divided by osmolality (urine) or dried weight (stool) to account for different hydration status and water content, respectively. Code for functions used for these process is freely available at https://github.com/maiyama/DataProcess-Tools under the name "initial_process.R". Most univariate and multivariate data analysis was performed using Metaboanalyst 4.0,[36] including Mann-Whitney U-test, FDR adjustment for multiple hypothesis testing, principle component analysis (PCA), orthogonal partial least-squares-discriminant analysis (OPLS-DA) and receiver operating characteristic (ROC) curves. In all cases, missing values were replaced with half of the lowest detected value, whereas metabolomic data sets were (generalized) log transformed and autoscaled when performing multivariate data analysis. Also, data normality assessment and effect size calculations were performed using the

Statistical Package for the Social Science (SPSS, version 21). Normality assumption was violated for majority of metabolites (88% for stool and 73% for urine) based on Shapiro-Wilk test ($\alpha = 0.05$). As a result, non-parametric univariate test (Mann-Whitney U-test) was also performed on untransformed metabolomics data.

4.3.8 Microbiome DNA isolation, illumina sequencing and data analysis. Genomic DNA was extracted as described in Stearns *et al.* with some modifications.[37] Samples were transferred to screw cap tubes containing 2.8 mm ceramic beads, 0.1 mm glass beads, GES and sodium phosphate buffer as described. Samples were bead beat and centrifuged as described and the supernatant was further processed using the MagMAX Express 96-Deep Well Magnetic Particle Processor from Applied Biosystems with the Multi-Sample kit (Life Technologies#4413022). Purified DNA was used to amplify the v3 region of the 16S rRNA gene by PCR. 50 ng of DNA was used as template with 1U of Taq, 1x buffer, 1.5 mM MgCl2, 0.4 mg/mL BSA, 0.2 mM dNTPs, and 5 pmoles each of 341F (CCTACGGGAGGCAGCAG) and 518R (ATTACCGCGGCTGCTGG) Illumina adapted primers, as described in Bartram *et al.*[38] The reaction was carried out at 94 °C for 5 min, 25 cycles of 94 °C for 30 s, 50 °C for 30 s and 72 °C for 30 s, with a final extension of 72 °C for 10 min. Resulting PCR products were visualized on a 1.5% agarose gel. Positive amplicons were normalized using the SequalPrep normalization kit (ThermoFisher#A1051001) and sequenced on the Illumina MiSeq platform at the McMaster Genomics Facility. Resulting sequences were run through the sl1p pipeline as described in Whelan *et al.*[39] Initially, a total of 5234 taxa were present from all samples analyzed. The OTUs were annotated using Greengenes database. 16S rRNA data was filtered to remove OTUs present in less than 10 % of samples prior to rarefaction, resulting in 392 OTUs. The library size ranged from minimum 4940 to maximum 173466. Alpha diversity was calculated in phyloseq (v. 1.22.0)[40] based on Chao1, Shannon and Simpson diversity measures at genus level after rarefying reads to minimum library size. Taxonomic abundance was visualized using fantaxtic (v. 0.1.0). Inferred metagenomics and predicted functional analysis were imputed using PICRUSt v.1.1.3[41]. Briefly, the OTU table was used as the input file after rarefaction to 4000 reads per sample. Copy-number normalization was also

performed using raw counts. Using MicrobiomeAnalyst,[42] A linear discriminate analysis effect size (LEfSe)[43] analysis was performed on the PICRUSt output containing the relative abundance of KEGG orthologous groups (KO) per sample with a threshold linear discriminant analysis (LDA) score of 2.0. KEGG pathway enrichment analysis was then performed on the LEfSe identified differential KO features.

4.4 Results

4.4.1 Inflammatory markers show no difference between CD and UC

Most pediatric IBD patients in this sex and age-balanced cohort were newly diagnosed with an average age of about 12-13 years for CD and UC sub-groups. However, three and four patients within the CD and UC sub-group of patients, respectively were diagnosed within the past 3 years. Also, only a small sub-set of these patients were receiving maintenance medications at the time of sample collection, including immunomodulators, biologics and/or anti-inflammatory drugs, whereas corticosteroids were not prescribed prior to this study. Thus, pediatric IBD patients were largely treatment naïve without an extensive history of long-term prescribed medications or prior surgery. Importantly, standard inflammatory biochemical measures for IBD did not show any significant differences between CD and UC patients sub-groups in our study (Mann-Whitney U-test, $p > 0.05$); however, most values were still well above the recommended threshold to be considered inactive (CRP ≤ 1.0 mg/L, calprotectin ≤ 250 µg/g) clearly indicating an active inflammatory state. IBD classification and differential disease diagnosis was determined following colonoscopy imaging together with colonic mucosal tissue biopsies collected for histopathology. Inflammation was clearly visible in all cases, further confirming the active disease state. All samples included in this study were collected at a single hospital site, and a subset of patients in each group did not provide samples prior to induction therapy, resulting in slightly unbalanced sample sizes as shown in **Table 4.1**. Overall, there were 15 (out of 19 total) matching urine and stool samples in the CD group and 8 (out of 11 total) paired samples collected from the UC group. As expected, compliance for specimen collection was greater for urine as it was more easily collected as compared to stool specimens.

Table 4.1 Summary of pediatric patient cohort who participated in this study with clinical measurements based on mean values and errors as ± 1s. All biochemical measurements were derived from serum samples except for fecal calprotectin, whereas disease location was based on colonoscopy imaging.

Criteria	CD ($n = 19$)	UC ($n = 11$)
Age; mean ± sd	13 ± 2	12 ± 3
Sex; male:female	9:8	6:5
New diagnosis (n)	13	6
CRP (mg/L)[a]	40 ± 40	36 ± 65
Fecal calprotectin (µg/g)[a]	3360 ± 2230	2420 ± 1570
Hemoglobin (g/L)	107 ± 16	111 ± 17
ESR (mm/hr)	37 ± 21	35 ± 26
Albumin (g/L)	27 ± 4	32 ± 4
Disease location (n)		
Ileocolonic	11	NA
Ileocolonic + UGI	2	NA
Colonic	3	11
Colonic + UGI	3	NA
IBD medications at sampling		
Biologic[b]	1	0
Immunomodulator[c]	1	0
5-ASA[d]	0	3
Biologic + Immunomodulator[e]	0	1
Urine samples available	19	8
Stool samples available	15	11

Abbreviations include, CD: Crohn's disease; ESR: Erythrocyte sedimentation rate; UC: Ulcerative colitis; PCDAI: Pediatric Crohn's disease activity index; PUCAI: Pediatric ulcerative colitis activity index; UGI: Upper Gastrointestinal tract
[a] There were no significant differences ($p > 0.05$, Mann-Whitney U test) measured between inflammatory markers between CD and UC patients, including serum CRP and fecal calprotectin.
[b] Infliximab; [c] Methotrexate;; [d] Aminosalicylic Acid (ASA) or Mesalamine; [e] Azathioprine + Adalimumab

4.4.2 Sample workup protocol, metabolomics data workflow and quality control

In order to ensure reproducible and quantitative metabolite extractions from freeze-dried stool samples, a micoscaled extraction protocol was developed prior to IBD sample analysis using a representative pooled stool from a subset of patients ($n = 24$). Most of stool samples collected from children with IBD were extremely loose with frequent inclusion of blood as expected based on their active inflammatory disease status. Water content in stool samples was highly variable, which necessitated a lyophilisation process to remove water and more efficiently extract metabolites from a fine powdered sample, which also enabled normalization of measured metabolite responses to dried weight (about 15-20 mg) of each stool sample. Stool extracts were prepared using a modified Bligh Dyer extraction method,[32, 44] which utilized a mixture of methanol, water and chloroform to preferentially extract a wide range of polar/ionic metabolites within the upper methanol/water layer that are ideally suited for analysis using MSI-CE-MS (**Figure 4.1a**). In contrast, lipids are partitioned within the bottom chloroform layer, whereas denatured proteins reside at the bi-phasic solvent interface.[24] Prior to individual stool sample analysis, the reproducibility of the stool extraction protocol was evaluated using MSI-CE-MS, which offers a high throughput metabolomics platform by allowing for serial injection of six technical replicate stool extracts and a blank extract within a single run. Absence of sample carry-over effects or background contamination was ensured based on the absence of a signal at the blank position, and reproducibility of extraction process was evaluated by the mean metabolite response ratio measured from six technical replicates of a pooled stool extract analyzed. Excellent reproducibility and recovery (average > 85 %) was found for representative cationic and anionic metabolites after normalization to internal standards and dried stool weight as shown in **Table S4.1** of the Supporting Information as reflected by a CV of 3.7% and 9.1% for succinic acid and tryptophan, respectively.

Additionally, excellent long-term precision for the analysis of all stool extracts over several days when using MSI-CE-MS was confirmed (overall CV = 14%, $n = 214$) based on inspection of the control chart for the recovery standard (F-Phe) as shown in **Figure 4c** with few outliers exceeding confidence intervals ($\pm$ 2s).
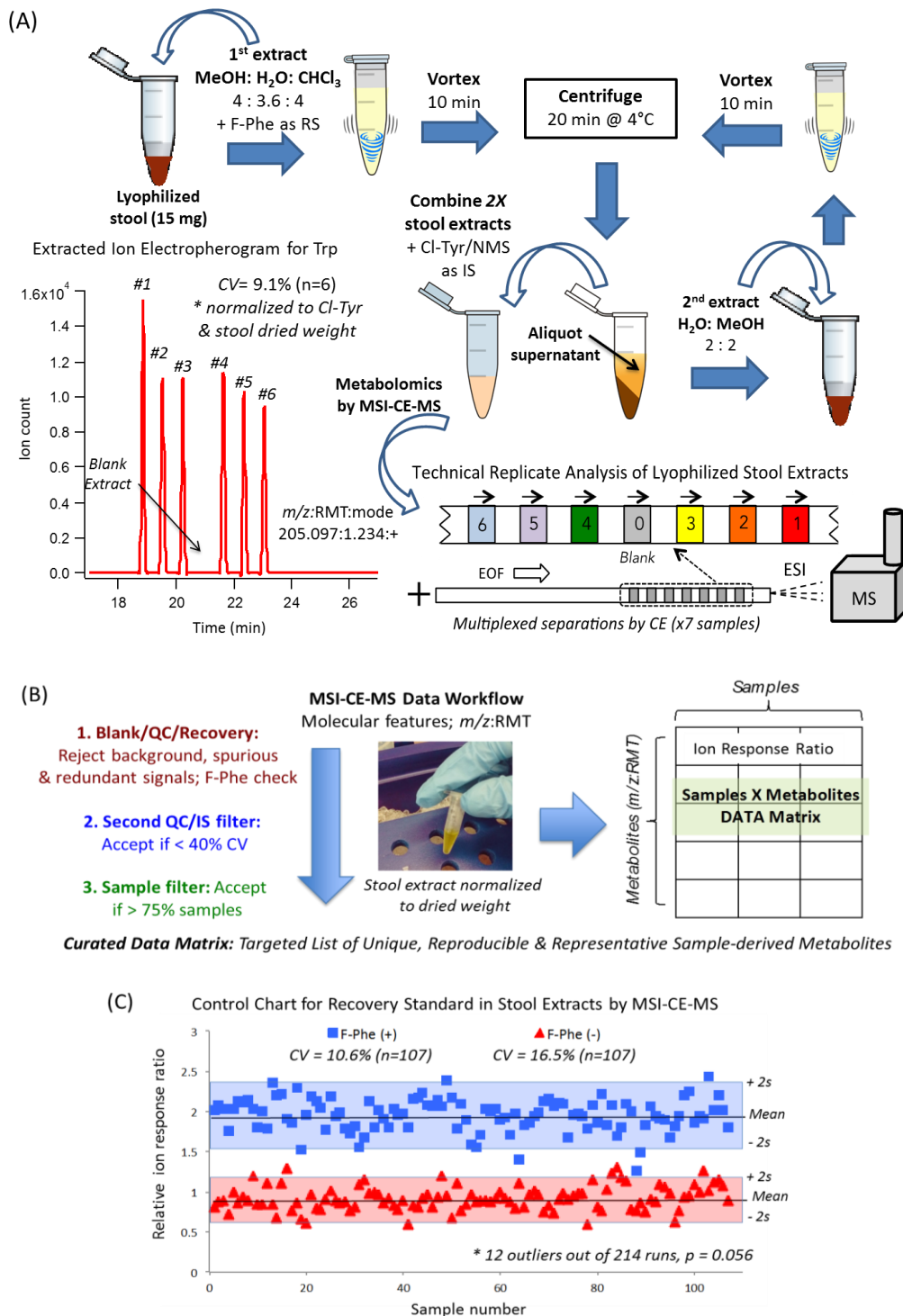
Figure 4.1. (a) Stool extraction protocol and schematic of MSI-CE-MS for high throughput metabolite profiling of stool extracts with stringent QC. **(b)** Overview of the metabolomics data workflow for selection of reproducible yet authentic metabolites detected in the majority of stool extracts with high data fidelity. **(c)** Control chart for recovery standard (10 µM F-Phe) included in every sample to monitor long-term method precision across the entire analysis of stool samples when analyzed by MSI-CE-MS under positive and negative ion mode detection.

Subsequently, IBD stool extract samples were analyzed using MSI-CE-MS with temporal signal pattern recognition[28] to ensure reliable sample assignment to each signal while maintaining high sample throughput. This multiplexed separation platform effectively filters out spurious signals, background ions and redundant peaks initially when using dilution trend filter on a pooled QC sample, which is then followed by a second QC filter to ensure adequate technical precision was achieved for each metabolite (CV < 40%) that was also detected in a majority of samples (> 75%) to avoid biases in subsequent statistical analysis (**Figure 4.1b**). In this case, MSI-CE-MS offers a high data fidelity metabolomics platform for biomarker discovery since a QC is injected randomly within the serial sample injection sequence for every run as required for robust correction of long-term monitoring of signal drift and batch correction in ESI-MS.[45] This rigorous filtering process ensures that only authentic, reproducible yet representative metabolites from urine or stool extracts are included in the final metabolomics data matrix, including known and unknown metabolites annotated by their characteristic *m/z*:RMT. Finally, all measured ion responses for a metabolite (measured in duplicate for each sample) are then normalized to an internal standard (Cl-Tyr or NMS) and further normalized to dried weight (stool) or osmolality (urine) when using univariate or multivariate statistical methods with appropriate data transformation.

4.4.3 Characterization of the urine and stool metabolome of pediatric IBD patients

A total of 104 fecal metabolites (66 cations, 38 anions) and 131 urinary metabolites (66 cations, 65 anions) were initially detected from stool ($n = 27$) and urine ($n = 26$) specimens collected from pediatric IBD patients prior to exclusion of infrequently (< 75 %) detected metabolites in samples. As a result, the final metabolomics data matrix used in this study was comprised of a total of 72 stool and 122 urinary metabolites in order to minimize missing data values. In this case, the majority of excluded compounds in urine were putatively identified as exogenous drugs and their metabolites based on their accurate mass and diagnostic MS/MS spectra. For example, mesalamine (*m/z:*RMT*, 154.0500:0.887; cation) and salicyluric acid (*m/z:*RMT, 194.0458:0.921; anion) were detected only in a subset of IBD patients who were prescribed 5-aminosalicylic acid-based drugs. However, compounds

identified as propofol glucuronide (*m/z:*RMT 353.1598:0.766; anion) and hydroxypropofol glucuronide (*m/z:*RMT, 369.1545:0.760; anion) were detected in the urine and stool of most patients as propofol was used as a sedative prior to colonoscopy and subsequent urine and stool sample collection. The propofol metabolites were detected only in the baseline urine samples with residual amounts still detected after 2-3 days as shown in **Figure S4.1** of the Supporting Information. Thus, compounds identified as drug metabolites were also excluded from the final data matrix as they are a source of confounding while not being consistently detected in the majority of pediatric IBD patients. Overall, there were 50 polar/ionic metabolites commonly measured in both urine and stool samples, and the major class of metabolite measured when using MSI-CE-MS was amino acids and their derivatives followed by organic acids, amines/amino sugars and purines/pyrimidines (**Figure S4.2**). Additionally, a majority of steroid metabolites, including bile acids, were detected only in stool. The most abundant metabolite in urine besides creatinine was hippuric acid (*m/z:*RMT; 178.0510:0.914), whereas amino acids, especially leucine and isoleucine were the most abundant metabolites measured in stool extracts. In general, more chemically diverse array of metabolites were detected in stool extracts, whereas secondary metabolites excreted as their intact glucuronide and sulfate conjugates were more abundant in urine. Metabolites were annotated based on their *m/z:*RMT and most likely molecular formula with low mass error (< 5 ppm) from high resolution MS, whereas collisional-induced dissociation (CID) experiments were performed for unknown metabolites associated with CD or UC in order to acquire MS/MS spectra for database matching (*e.g.,* HMDB)[35] with *in silico* fragmentation modeling (e.g., MetFrag).[31] Additionally, spiking experiments with authentic chemical standards were also performed to confirm their co-migration with 67 metabolites in urine and 56 metabolites from stool extracts identified with a known chemical structure (level 1), whereas a small sub-set of unknown metabolites remained if standards and matching MS/MS spectra were unavailable, and precursor ion signal was too low for CID experiments.

As compared to the urine metabolome, there was a much greater extent of biological (between-subject) variance for stool metabolites even after exclusion of infrequently

detected (< 75% samples) or highly variable metabolites (CV of QC > 40%), such as polyamines, amino acids, organic acids and drug metabolites. However, a complete list of all detected metabolites in urine and stool extracts are summarized in the Supplemental Information (**Table S4.6** and **Table S4.7**). This major difference in measured biological variance between urine and stool metabolites were also apparent when using principal component analysis (PCA) as shown in **Figure 4.2 a, b**. In this case, technical precision was effectively monitored throughout the analysis by including QC samples in every single run, which resulted in an average CV of 20% and 21% based on 122 urinary and 72 stool metabolites, respectively. In contrast, the average biological variance of osmolality normalized urine samples was 63%, whereas average variance in dried weight normalized stool extracts was 135%, reflecting much greater heterogeneity of stool samples. In addition to inherent between-subject variability in stool samples, delays in storage and chemical stability of metabolites may also have attributed this large variance as not all stool specimens were immediately stored frozen after sample collection at the clinic. All urine and stool samples were transferred to a refrigerator soon after sample collection; however, this storage temperature may not have been effective to ensure the stability of stool metabolites. Additionally, morphological variability of stool samples was much greater than urine samples due to variable severity of diarrhea and bleeding among pediatric IBD patients, which was not the case for urine.

A wide range of polar/ionic metabolites were detected in urine samples, which are ideal for analysis by MSI-CE-MS when using an acidic buffer/positive ion mode and alkaline buffer/negative ion mode for cations and anions, respectively. However, the absolute concentration of urinary metabolites is highly dependent on hydration status when relying on single-spot urine sample collection,[46] which necessitated correction using osmolality or creatinine normalization. In our study, creatinine was not a suitable parameter for urine normalization given that the cohort comprised both male and female IBD patients, including young children and teenagers with different muscle mass and dietary patterns.[47]
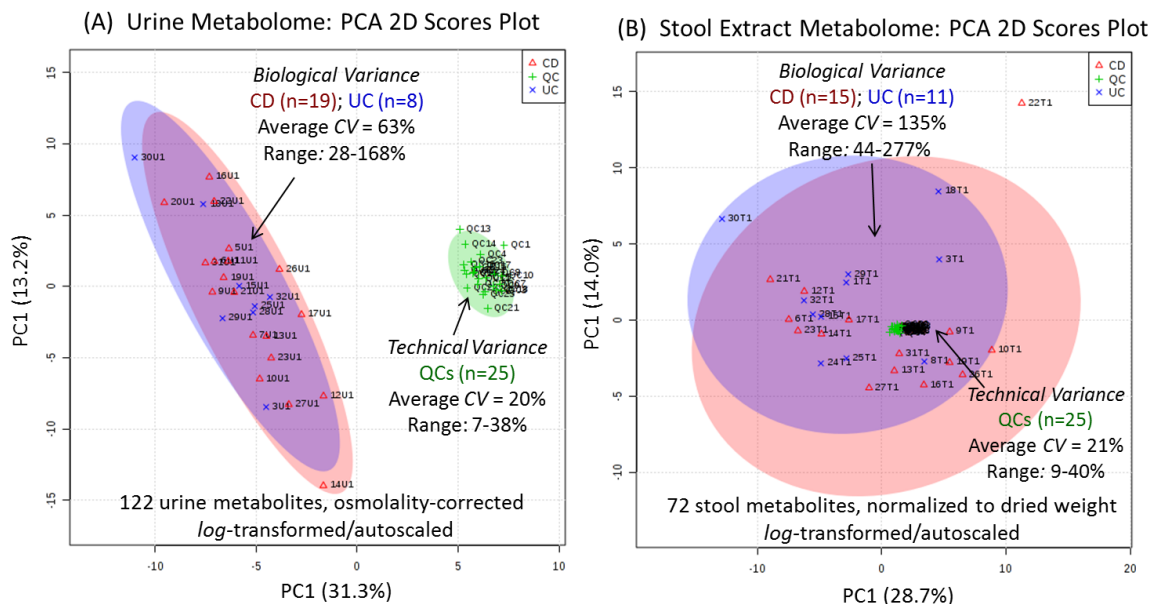
Figure 4.2. (a) A PCA 2D scores plot of 122 urine metabolites from 19 CD and 8 UC patients and (**b**) A PCA 2D scores plot of 72 stool metabolites from 15 CD and 11 UC patients with corresponding pooled QC samples analyzed in every MSI-CE-MS. Overall technical variance in urine and stool analysis was equivalent (CV = 20%), but between-subject biological variance in stool extracts was more than two-fold larger as that of the urine metabolome. Data was normalized to osmolality (urine) or total dried weight (stool) followed by a generalized *log*-transformation with autoscaling.

Nevertheless, urinary creatinine and osmolality showed good overall correlation as demonstrated by a linear regression model ($R^2 = 0.610$, $p = 1.52$ E-03, $n = 27$). Additionally, similar statistical outcomes were realized when comparing urinary metabolites that differentiate between CD and UC regardless of normalization parameter used as shown in **Figure S4.3** of the Supporting Information.

4.4.4 Biomarker candidates in stool and urine that differentiate CD from UC

A comparison of the stool and urine metabolome between CD and UC patients was next performed using nonparametric univariate analysis based on Mann-Whitney U-test due to skewed data distribution observed in majority of metabolites (93% in stool, 70% in urine) based on a Shapiro-Wilk normality test ($p < 0.05$). Results of univariate test are summarized in **Table 4.2** (stool) and **Table 4.3** (urine) with unadjusted *p*-values (< 0.10), median fold-change (FC) and effect sizes.

Table 4.2. Top-ranked biomarker candidates from stool extracts normalized to dried mass identified by MSI-CE-MS that differentiate pediatric CD (*n*=15) from UC (*n*=11) patients.

m/z:RMT:polarity	Chemical ID	*p*-value	Median FC	Effect size
405.2646:0.753:n	7-Ketodeoxycholate	0.0173	26	0.47
189.1598:0.612:p	Trimethyllysine	0.0203	2.6	0.45
87.0452:1.054:n	Butyric acid	0.0402	3.1	0.41
407.2803:0.748:n	Cholic acid	0.0423	14	0.40
89.0244:1.140:n	Lactic acid	0.0464	0.50	0.40
205.0972:0.927:p	Tryptophan	0.0472	0.50	0.39
391.2865:0.755:n	Deoxycholic acid	0.0501	4.3	0.39
106.0499:0.844:p	Serine	0.0554	0.30	0.38
104.1069:0.560:p	Choline	0.0613	0.50	0.37
120.0655:0.887:p	Threonine	0.0619	0.40	0.37
204.123:0.792:p	Acetylcarnitine	0.0862	0.60	0.34

* Statistical significance based on Mann-Whitney U-test, p < 0.1 with a median FC > 2 or < 0.5, where a 50% missing rate was applied to allow inclusion of more differentially excreted stool metabolites.

Table 4.3. Top-ranked biomarker candidates from osmolality normalized urine identified by MSI-CE-MS that differentiate pediatric CD (*n*=19) from UC (*n*=8) patients.

m/z:RMT:polarity	Chemical ID	*p*-value	FC	Effect size
212.0023:1.025:n	Indoxylsulfate	1.18 E-03	2.55	0.59
345.1553:0.770:n	Unknown branched oxodecanoic acid	1.18 E-03	2.68	0.59
106.0499:0.868:p	Serine	1.33 E-02	0.70	0.47
227.9968:0.979:n	5-Hydroxyindole sulfate	1.89 E-03	2.76	0.60
120.0652:0.905:p	Threonine	6.21 E-03	0.46	0.53
209.0921:0.887:p	Kynurenine	0.0112	0.53	0.48
263.1037:0.826:n	Phenylacetylglutamine	0.0112	2.34	0.48
172.9912:1.135:n	Phenolsulfate	0.0217	2.67	0.44
137.0457:1.039:p	Hypoxanthine	0.0253	0.69	0.43
222.0796:0.849:p	5-(delta-carboxybutyl) homocysteine	0.0253	2.76	0.43
350.0880:0.788:n	Indole-3-acetate glucuronide	0.0293	2.18	0.42
104.0706:0.940:p	Dimethylglycine	0.0463	0.66	0.39

* Statistical significance determined by a Mann-Whitney U-test, p < 0.05

Despite the large biological variability observed in the stool metabolome, clear trends in bile acids, short-chain fatty acids and certain amino acids excretion patterns were observed between CD and UC. For instance, pediatric UC patients were characterized by a significantly higher excretion of tryptophan and lactic acid, however lactic acid was not detected in nearly half of stool samples from CD patients (**Figure 4.3a, Table 4.2**). Conversely, cholic acid and trimethyllysine was more abundant in stool extracts from CD patients as compared to UC patients ($p < 0.05$). Although only marginally significant ($p = 0.05$-0.10) with lower effect sizes (< 0.04), higher excretion of serine, choline and acetylcarnitine were also noted in stool extracts for UC. Since frequently missing compounds may provide important information about differential metabolism if any metabolites are completely missing from one group, the initially removed 32 metabolites from stool were reviewed for a such trend. All metabolites excluded were measured in both groups, however, an uneven rate of detection was found in majority of bile acids, including 7-ketodeoxycholate, deoxycholate, taurodeoxycholate and taurocholate (average 69% for CD and 38% for UC). When 50 % missing rate cut-off was applied instead of 75% to allow inclusion of these differentially excreted metabolites, 14 metabolites were included back into the data matrix resulting in a total of 86 stool metabolites, which revealed significantly higher excretion of 7-ketodeoxycholic acid, as well as butyric acid in CD patients. When discriminating characteristics of these metabolites were evaluated using a receiver operating characteristic (ROC) curve, trimethyllysine was most significant with an $AUC = 0.773$ and $p = 0.026$ (81% sensitivity and 67% specificity), whereas a ratio of guanine/choline showed slightly better performance with an $AUC = 0.842$ and $p = 0.0031$ (81% sensitivity and 73% specificity) as shown in **Figure 4.3b**.

In the urine metabolome analyses, samples from CD patients were characterized by significantly higher excretion of indole metabolites, including indoxylsulfate, phenylsulfate, phenylacetylglutamine, indole-3-acetic acid glucuronide, and an unknown compound later identified as 5-hydroxy-6-indolyl-*O*-sulfate (**Figure 4.4a**).
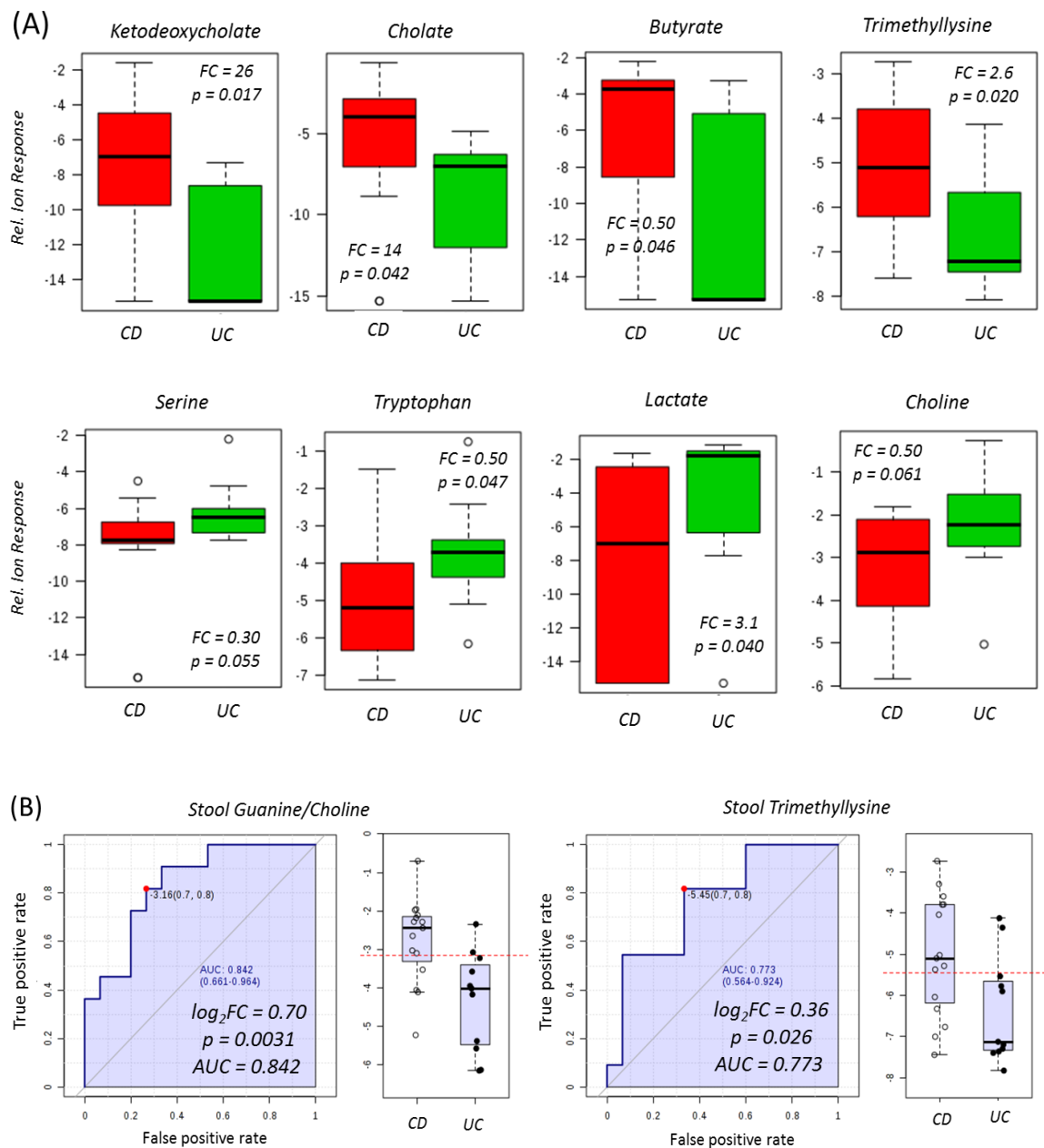
Figure 4.3. (a) Box-whisker plots of metabolites that were differentially excreted in stool extracts ($p < 0.05$) collected from pediatric UC (n=11) and CD (n=15) patients, where the y-axis of boxplots show generalized *log*-transformed relative ion responses. **(b)** ROC curves for top-ranked ratiometric and single stool metabolites differentially excreted in UC compared to CD by median fold-change (FC), Mann-Whitney U-test ($p < 0.05$) and area under the curve (*AUC*). Data were normalized to dried weight for stool and no transformation was performed prior to statistical analysis. Missing values were inputed with the lowest value found for the metabolite across all samples divided by 2 with metabolites detected > 50% of stool samples analyzed.
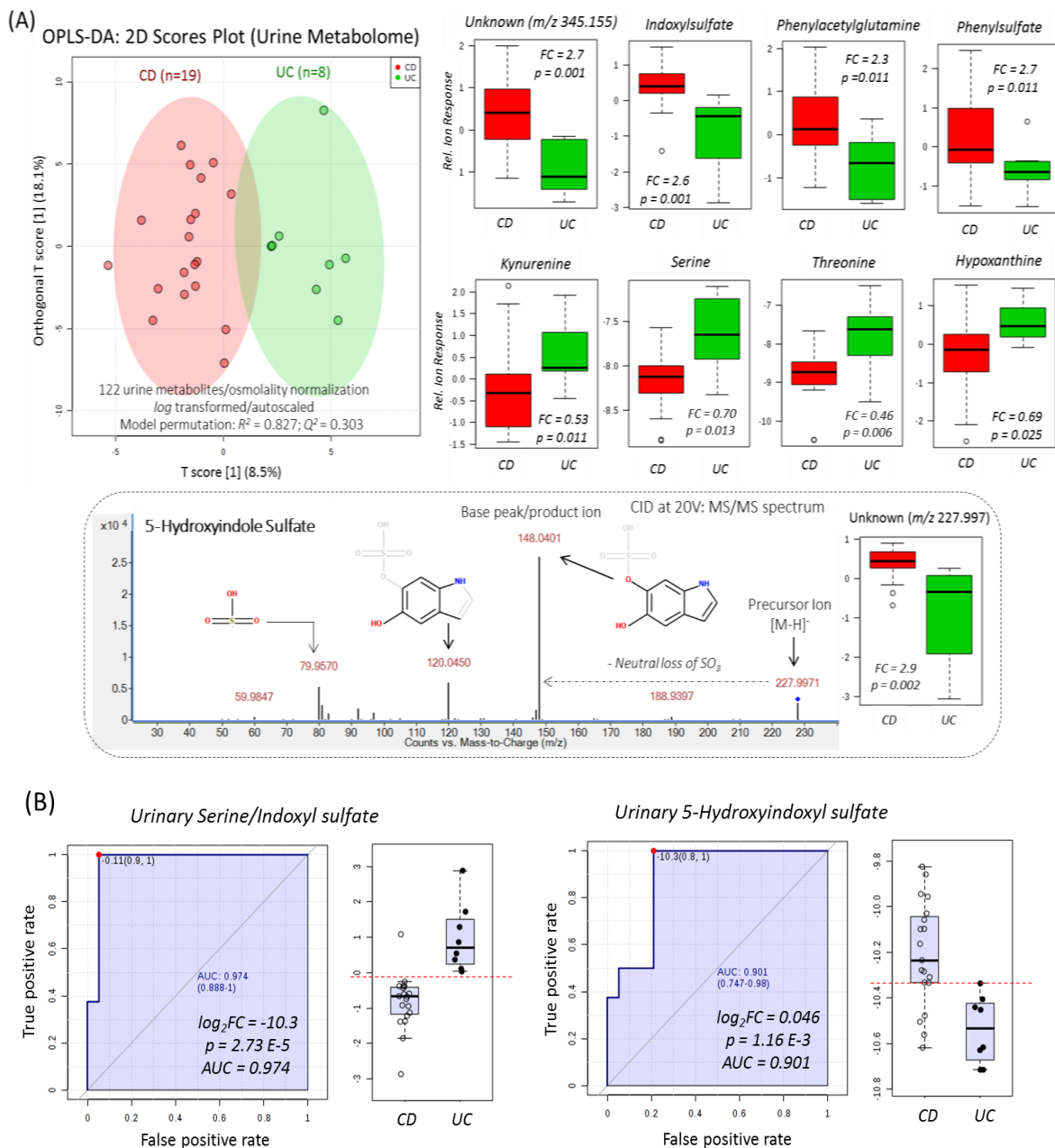
Figure 4.4 (a) Overall biological variance that discriminate between UC and CD patients when projected as 2D score plot from OPLS-DA based on 122 urinary metabolites consistently detected in majority of urine samples following cross-validation ($R^2 = 0.827$, $Q^2 = 0.303$) together with boxplots of top-ranked biomarker candidates in urine (Mann-Whitney U-test, $p < 0.05$). Unknown compound with m/z 345.155 is assigned with molecular formula of $C_{16}H_{26}O_8$ and tentatively identified as branched oxodecanoic acid. Urine metabolomics data was normalized to osmolality with generalized *log*-transformation used only for plot generation. Missing values were imputed with the lowest value found for the compound across all samples divided by 2. Y-axis of boxplots show log-transformed relative ion response. **(b)** ROC curves of differentially excreted urinary metabolites ROC curves indicate the area under the curve (AUC).

The MS/MS fragmentation spectrum of this oxidized indole metabolite clearly showed a product ion and neutral loss for sulfate resulting in formation of dihydroxyindole as the base peak (*m/z* 148.0401). Additionally, significantly higher levels of unknown anion (*m/z:*RMT, 345.1550:0.770) and cation (*m/z:*RMT, 222.0796:0.849) were found in CD patients. There were no database matches for these compounds in HMDB, but the MS/MS fragmentation spectra are indicative of a branched oxodecanoic acid and carboxybutyl-homocysteine based on *in-silico* fragmentation predictions using MetFrag as shown in **Figure S4.4** of the Supporting Information. Urine samples from UC patients on the other hand, contained much higher levels of amino acids, such as threonine, serine, kynurenine, hypoxanthine and dimethylglycine. These metabolites contributed to the overall separation of urinary metabolic profiles of UC and CD patients when using orthogonal partial least square-discriminant analysis (OPLS-DA) with good model performance following cross validation ($R^2 = 0.827$, $Q^2 = 0.303$) as shown in **Figure 4a**. Moreover, these same urinary metabolites remained significant when using a univariate Mann-Whitney U-test (without data transformation) after adjustment for multiple hypothesis testing using Bejamini-Hochberg FDR adjustment ($q < 0.05$) with moderate effect sizes of 0.50 to 0.60 (**Table 4.3**). Ratiometric analysis also uncovered a correlative relationship between kynurenine and indole derived metabolites, which are all derived from the catabolism of tryptophan as summarized in **Table S4.2** of the Supporting Information. In this case, ROC curves of top-ranked single and ratio metric urinary metabolites showed excellent performance as candidate biomarkers for differential diagnosis of CD and UC as compared to stool. Specifically, the ratio of serine to indoxylsulfate achieved an $AUC = 0.967$ and $p = 2.73$ E-5 as reflected by a sensitivity of 100% and a specificity of 95% (**Figure 4.4b**). Similarly, osmolality normalized urinary 5-hydroxyindole sulfate also showed very good discriminating performance with an $AUC = 0.901$ and $p = 1.16$ E-3 with a sensitivity of 100%, but with a lower specificity of 79%.

4.4.5 Chemical stability of lead urinary and fecal biomarkers with delayed storage

To investigate the effect of different storage conditions and other pre-analytical factors that may have affected our results, the impacts of delayed storage and storage temperature were

studied systematically with representative urine and fecal samples collected from two healthy volunteers. Both subjects who provided stool samples were 6-year-old twin boys. Despite the same diet, bile acids and butyrate had an opposite trend of changes between the twins, while amino acids (serine, tryptophan, trimethyllysine) and choline generally increased with longer storage duration at room temperature in both subjects, which may indicate differences in microbial metabolism of butyrate and bile acids. Amino acids and butyric acid were affected by storage conditions by a larger extent than bile acids (**Figure 4.5**). Overall, greater changes in stool metabolite levels were observed at room temperature as expected from the higher microbial activities at warmer temperatures; however, storage of stool samples under refrigeration did not prevent alteration in metabolite levels after as short as 2 hr as seen in ketodeoxycholate and trimethyllysine. Furthermore, freeze-thaw cycle by storing samples in a freezer for 48 h and thawing at room temperature affected the stability of the majority of stool metabolites to a greater extent than a 48 h storage at room temperature or under refrigeration. In contrast to the variable stability observed among stool metabolites, important urinary indole and phenolic metabolites were remarkably stable under different storage condition, however amino acids were found to significantly decrease at room temperature within 6 h (**Figure 4.6**). The unknown branched fatty acid (*m/z* 345.1550) with high excretion in CD group was also stable with only a minor decrease at room temperature after 36 h. The decreasing trends observed in hypoxanthine, kynurenine, serine and threonine may indicate oxidation or biotic degradation of these compounds at room temperature, which was effectively prevented when urine samples were stored at 4 $^{\circ}$C. Furthermore, the addition of sodium azide as a preservative did not perturb metabolite responses analyzed based on the comparison with control values. Longer duration in storage under refrigeration even up to 48 h, did not show appreciable changes in responses confirming that these urinary metabolites are likely robust biomarkers for differentiation of CD and UC within a clinical setting since they are largely insensitive to delays to storage and storage conditions.
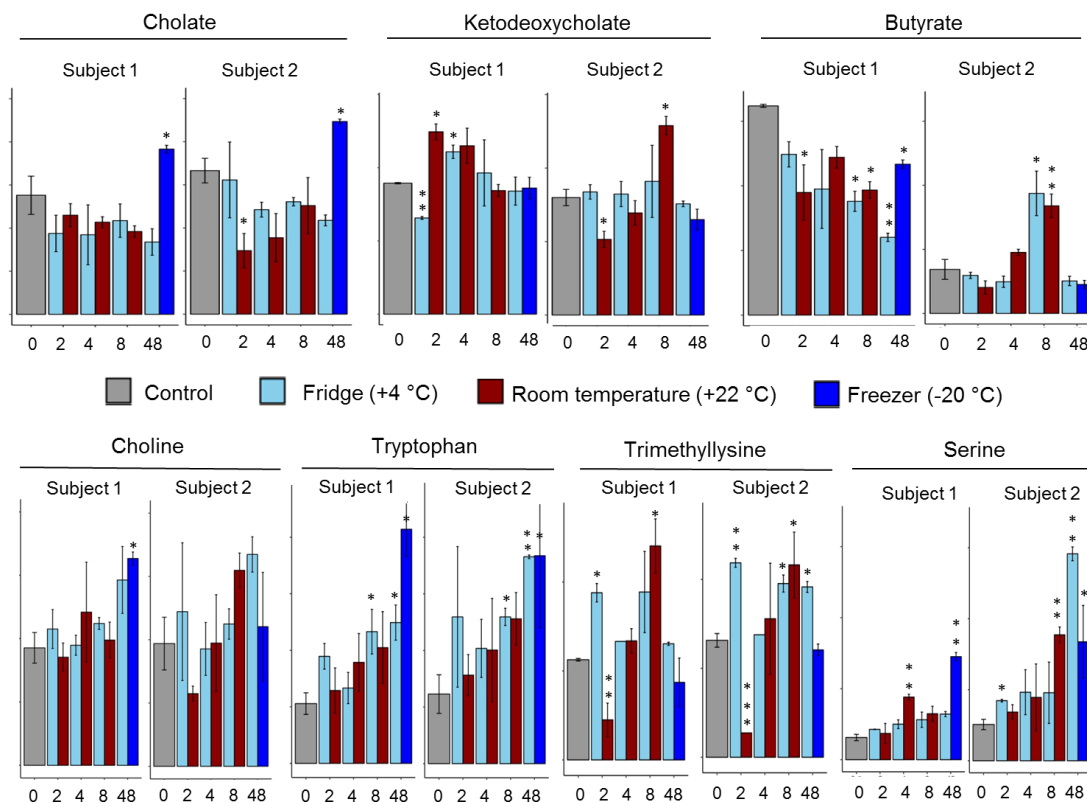
Figure 4.5. Representative plots illustrating the chemical stability of select metabolites in stool samples stored under different conditions, where *, ** and *** indicate differences between control and each condition at $p < 0.05$, $p < 0.01$, and $p < 0.001$, respectively. A two-tailed homoscedastic student's t-test was used for analysis, where y-axis represents relative peak area, and x-axis is storage duration over time.

4.4.6 Different taxonomic abundance and predicted functions between CD and UC

In order to investigate differentially abundant bacterial taxa that could explain the observation in metabolome, 16S rRNA gene sequencing was applied to stool samples. Alpha diversity at genus level was slightly higher in CD samples when using Chao1 diversity measure but the difference did not attain statistical significance ($p = 0.328$; Mann-Whitney U test; **Figure 4.7a**). In Shannon and Simpson diversity measures, p-values were 0.669 and 0.973, respectively, further indicating that there was no differences in bacterial diversity in stool samples of UC and CD patients. Wide diversity in relative abundance at genus level was observed in both CD and UC groups with notable abundance in phylum *Firmicutes*, *Bacteroidetes* and *Proteobacteria* (**Figure 4.7b**).
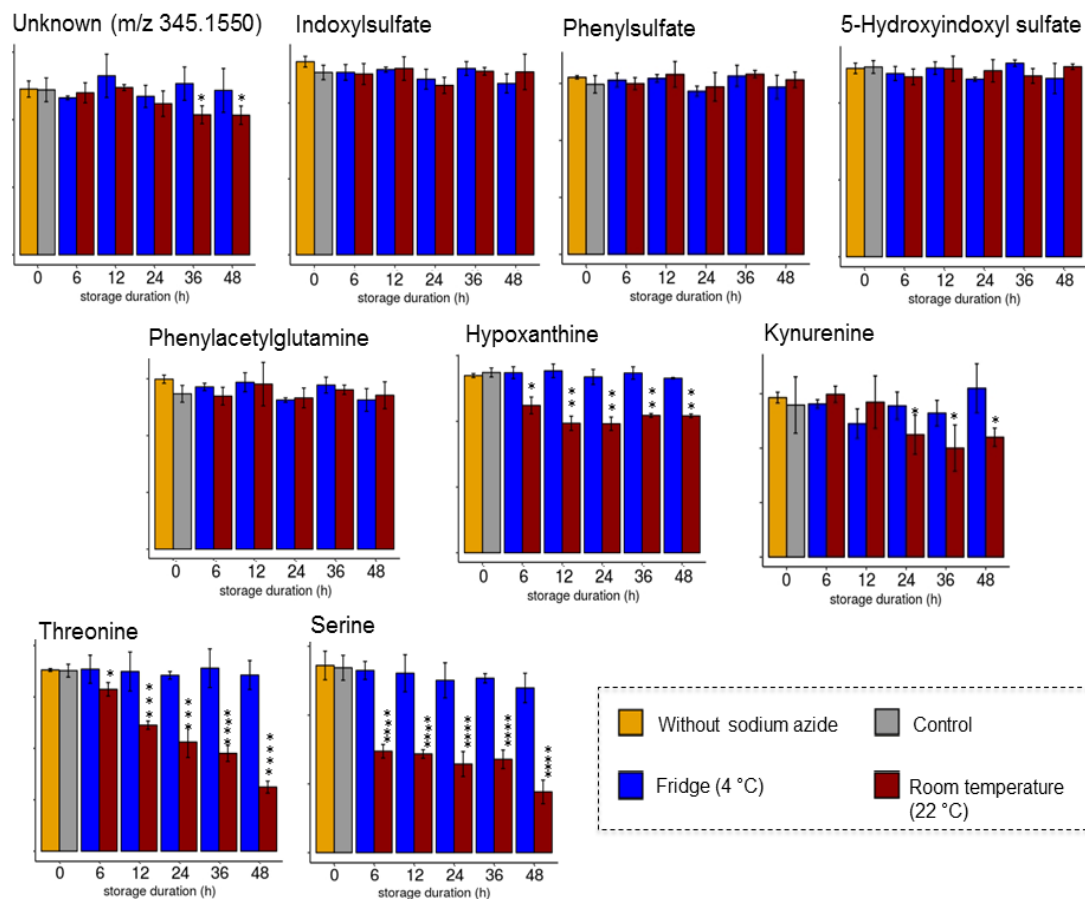
Figure 4.6. Representative plots illustrating the chemical stability of select urinary metabolites, where *, **, *** and **** indicate differences between control and each condition at $p < 0.05$, $p < 0.01$, $p < 0.001$, and $p < 0.0001$, respectively. A two-tailed homoscedastic student's t-test was used for analysis, where y-axis represents relative peak area, and x-axis is storage duration over time.

Relative abundance of each taxa may vary, but approximately 75 – 85 % of total fecal microbiome was comprised of 25 genera across the samples. Based on a simple visual observation, some samples contained strikingly high relative content of *Proteobacteria*, especially genus *Escherichia*. Additionally, this phylum was represented higher in UC than CD samples while phylum *Firmicutes* showed the opposite trend. This initial observation was further confirmed with linear discriminant analysis (LDA) effect size (LEfSe) based on LDA score and statistical significance (p < 0.05) with non- parametric factorial Kruskal-Wallis sum-rank test. Twenty-seven OTU were identified to be differentially abundant
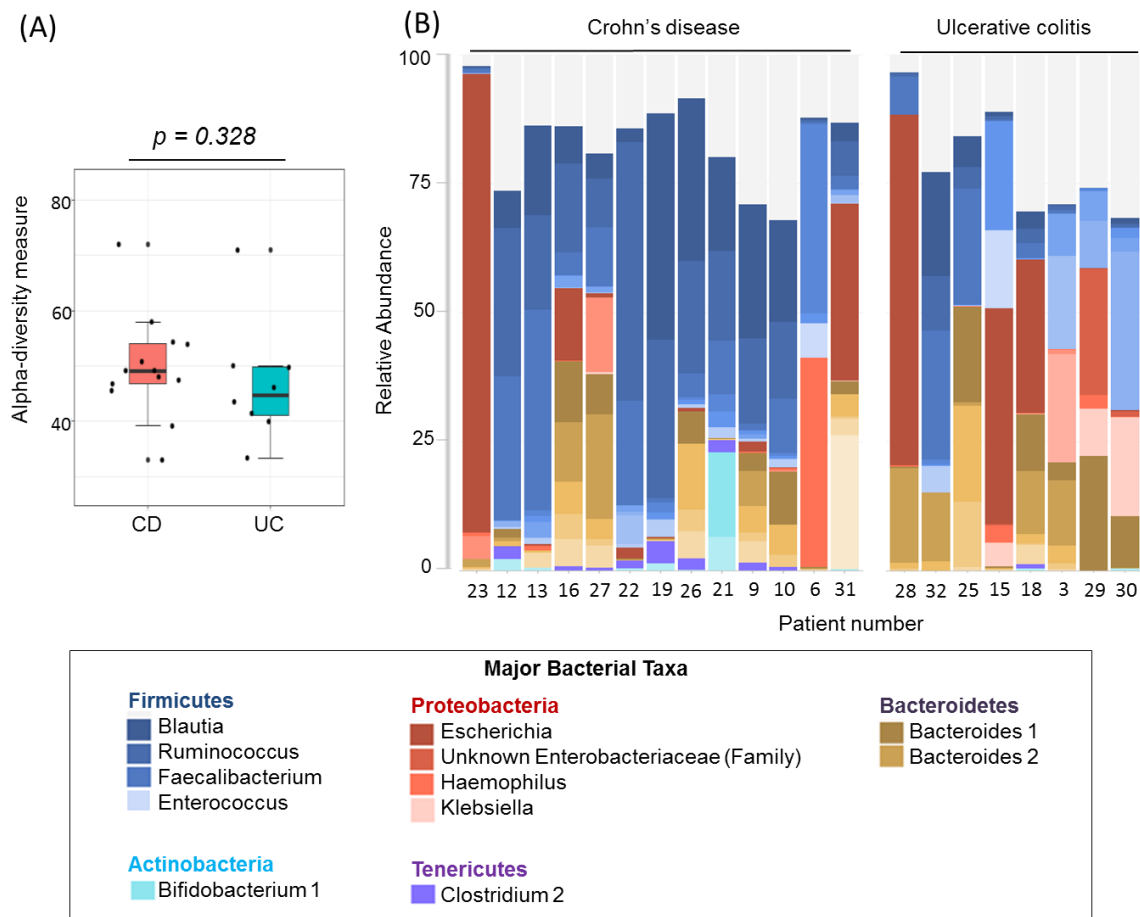
Figure 4.7. (a) Alpha diversity measure using Chao1 at genus level. Statistical significance based on Mann-Whitney U-test. (**b**) 25 most abundant genus overall across samples are shown in a relative abundance bar plot. Each phylum is differentially colour-coded and major bacterial taxa at genus level are highlighted in the box below.

between CD and UC. Fifteen of them were higher in UC samples and 12 others had greater abundances in CD patients (**Figure 4.8**). Taxonomic analysis of the differential abundance identified higher abundance of genera in phylum *Firmicutes* in CD, and higher abundance of genera in phylum *Proteobacteria* in UC (**Table S4.3**). Notably, abundant genera of bacteria in UC that belong to this phylum are generally known as pathogens, namely, *Klebsiella, Veillonella, Haemophilus* and *Campylobacter*.
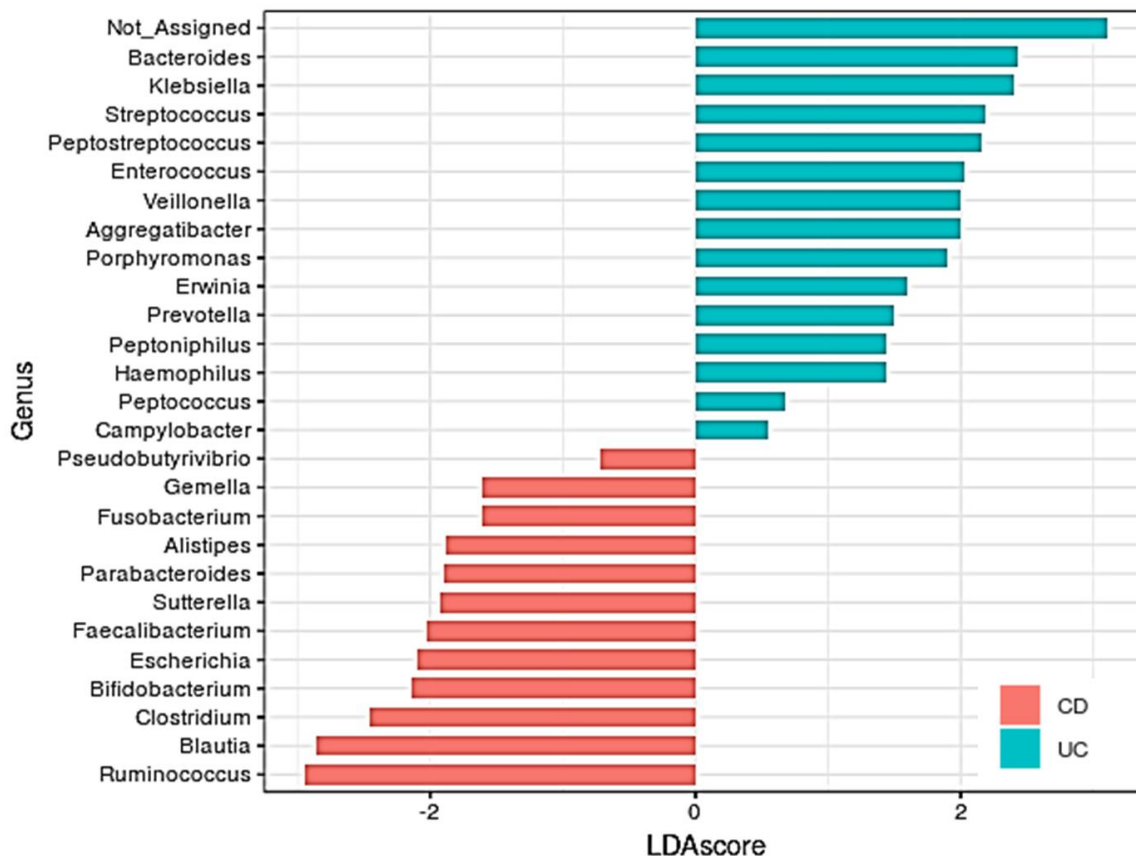
Figure 4.8. Differentially abundant taxa at genus level was identified by linear discriminant analysis effect size (LEfSe). Rarefaction of counts to 4000 reads per sample and copy-number normalization was performed prior to the statistical analysis. Taxonomic annotation was based on Greengenes database.

Although analysis was not performed down to the species level, the abundance of these genera in UC indicated higher pathogenic potential of microbial community in these patients' GI tract. PICRUSt was used to predict functions of gut microbiota based on 16S rRNA collected from stool samples. There were 98 metabolism associated functional features differentially abundant in CD and UC samples (LEfSe LDA score > 2.0). 93 features out of 98 were more abundant in CD patients while only 5 were more abundant in UC patients (**Table S4.4**). KEGG pathway enrichment analysis highlighted significant prevalence of predicted functions in amino acid metabolic pathways, particularly for histidine, phenylalanine, tyrosine and tryptophan, enriched in CD patients (**Table S4.5**).

4.5   Discussion

Untargeted metabolite profiling of matching urine and stool specimens from pediatric CD and UC patients was performed using MSI-CE-MS when using a rigorous data workflow with stringent QC (**Figure 4.1**). A microextraction procedure was first developed for sample workup of stool extracts to ensure good recovery and reproducibility. Fecal water is frequently used for stool metabolomic studies based on a simple mixing procedure with buffer followed by filtration when using murine fecal specimens, which are typically analyzed by NMR.[48-52] However, human fecal specimens especially from active IBD patients poses significant challenges for sample workup because of the highly variable water content and inclusion of mucus and blood. Therefore, lyophilisation followed by use of a Bligh-Dyer extraction procedure using a mixture of organic solvents with deionized water was developed in order to normalize metabolite levels to total dry weight. By using chloroform to remove hydrophobic constituents and methanol to precipitate proteins, this extraction method also served as an effective method for sample cleanup that is compatible with MSI-CE-MS. Overall, excellent reproducibility (CV < 10%) with good recovery (> 85%) for a wide range of polar/ionic metabolites was realized when pooling together two separate extracts from a freeze-dried stool specimen as shown in **Figure 4.1a** and **Table S4.1** of the Supporting Information. Excellent long-term reproducibility of the stool extraction protocol and MSI-CE-MS methodology was further demonstrated by monitoring control charts for the recovery standard, F-Phe that was added to all fecal specimens analyzed in this study as shown in **Figure 4.1c**. Additionally, the metabolomics data workflow developed when using MSI-CE-MS for stool extracts effectively excluded spurious, background and redundant signals to reduce false discoveries while ensuring good technical variance. Some amino acids such as leucine, isoleucine, and arginine were only measured in stool extracts and not in urine samples, which reflected recent diet/protein intake and nutrient malabsorption prevalent among IBD patients. Additionally, the majority of bile acids were only detected in stool samples, as they are generally re-absorbed and recycled in the liver without being transferred to the kidney.[53] Conversely, highly polar/hydrophilic organic acids, indole and glycosylated metabolites were far more abundant in urine, but mostly undetected in stool. Overall, the biological variance of metabolites measured in stool

extracts was far greater than urine metabolome even after normalization to dried weight as illustrated in **Figure 4.2**. This is a result of the greater pre-analytical variance in stool collection that is more sensitive to delays to storage and storage temperature due to high metabolic activities of bacteria present in stool (≈ 30% of dried mass), which was apparent in our stool metabolome stability studies using fresh stool samples collected from healthy male twins (**Figure 4.5**). Despite genetic similarities, bile acids and butyric acid showed distinctive changes over time that likely reflect underlying differences in gut microbial communities/activities present as highlighted in a previous metabolomics study of discordant discordant twin pairs.[54] As reported in another study investigating fecal metabolome integrity,[55] thawing stool samples after freezing had a large impact on measured metabolite concentrations. As a result, stool metabolome data in our work was innately more variable with only modest effects sizes (0.34 to 0.47) realized for the top candidate biomarkers differentiating major pediatric IBD subtypes (**Table 4.2**). Nevertheless, all three bile acids (ketodeoxycholic acid, cholic acid and deoxycholic acid) were measured to be consistently elevated (median FC > 4; $p < 0.05$) in stool extracts collected from CD as compared to UC patients. In this case, primary (cholic acid) and notably secondary (deoxycholic acid, ketodeoxycholic acid) bile acids were elevated in stool specimens from CD affected children, which play essential roles in fat absorption in the intestinal lumen, but have also been reported to promote oxidative stress and tissue damage with greater risk for colorectal and other GI tract cancers at higher concentrations.[56]

Urine samples, on the other hand, showed remarkable stability for the majority of relevant metabolites associated with IBD status when stored refrigerated even up to 48 h (**Figure 4.6**). This assures the reliability of our findings in this work, where urine samples were immediately stored in a refrigerator after the addition of azide as an antimicrobial agent. However, prolonged storage at room temperature storage prior to azide addition/refrigeration can alter the concentration of certain urinary metabolites over time (> 6 h) likely due to microbial uptake of key nutrients excreted in urine, such as amino acids. Therefore, prompt azide addition and/or cold temperature storage of urine specimens without excessive delays (< 3 h) to freezing is optimum for robust metabolomics studies. Overall,

the urine metabolome was found to be a more reliable specimen for the discovery of putative biomarkers that differentiate pediatric CD from UC patients following normalization to osmolality (or creatinine) to correct for differences in hydration status when relying on random single-spot urine collection. We have chosen to use osmolality, which is a measure of total solutes in urine that impact freezing point depression, instead of creatinine because of a heterogeneous cohort in terms of sex, age, and muscle mass, all of which have been shown to influence urinary creatinine excretion besides hydration.[47, 57] Nonetheless, there was a good correlation between urinary creatinine and osmolality values measured independently in our study with similar trends observed for top-ranked metabolite that differentiated between CD and UC (**Figure S4.3** of the Supporting Information).

When comparing metabolic profiles of UC and CD patients, there has been little agreement among published results[19, 20, 22] due to the heterogeneity of IBD, and effects of medications and surgeries on host and microbial metabolism. Therefore, we focused our study on newly diagnosed pediatric IBD, with the majority patients who were treatment-naïve, and all patients were in active disease state at the point of recruitment. We found higher excretion of amino acids, particularly serine and threonine in urine and stool of UC compared to CD patients, indicating altered amino acid metabolism and subsequent development of autoimmunity as seen in other chronic pathology, such as pediatric type 1 diabetes.[58] This is a consistent finding with another study with larger cohort of pediatric IBD patients,[23] as well as colorectal cancer patients whose metabolism in glycine, serine and threonine pathway was enriched compared to other amino acid metabolic pathways when colonic mucosa of patients were studied by LC-MS.[59] More specifically in IBD, over expression of serine/threonine specific protein kinase has been implicated, which plays an essential role in cell proliferation, adhesion, as well as inflammation.[60, 61] High excretion of serine and threonine in UC patients may indicate over production of these amino acids, which complements the higher activity of enzymes that take these amino acids as specific substrates. In addition to serine and threonine, UC patients were characterized by higher excretion of metabolites in purine metabolism and key metabolite for DNA synthesis, namely hypoxanthine in urine and choline in stool. On the other hand, guanine, a nucleic

acid in purine pathway, was found more enriched in stool of CD patients, which resulted in modest performance as a biomarker when ratio of fecal guanine / choline was considered. Purine metabolism has been reported as an essential initiator of purine nucleotides synthesis through the purine salvage pathway for cell proliferation and re-establishment of intestinal barrier function, which is known to be disrupted in IBD.[62, 63] The purine salvage pathway is initiated through hypoxanthine-guanine phosphoribosyltransferase (HGPRT), which produces inosine monophosphate from hypoxanthine and guanosine monophosphate from guanine as a substrate.[64] High excretion of HGPRT substrates in both groups, as well as choline, which has been shown to stabilize DNA structure,[65] may indicate different mode of disruption in intestinal barrier function present in UC and CD.

Besides the ability for recovering mucosal integrity, intestinal mucus plays a critical protective role against pathogens. Depletion of mucus is more frequently observed in treatment naïve UC patients than CD patients,[6] and altered mucus properties have been suggested in human and animal models of colitis.[66] This alteration might be attributed to bacterial mucin-desulphating enzymes found in certain bacterial strains such as *Bacteroides spp.*[67] and *Prevotella spp.*,[68] both of which have been found enriched in feces of UC patients.[69] The presence of *Bacteroidetes* spp. was previously found not only in inflamed mucosal sites but also in non-inflamed regions, indicating the wide-spread presence of this group of bacteria in the colonic mucosa of UC patients. Higher abundance of these organisms are also found in our study along with an array of pathogens such as *Klebsiella* and *Campylobacter*, most of which belong to phylum *Proteobacteria,* class *Gammaproteobacteria*. Bacteria in this class are largely facultative anaerobes, which are not expected to be abundant in a strictly anaerobic environment such as human intestine.[70] Given this, abundance of these bacteria may suggest a disturbance of normal flora caused by an oxidative environment. The dominance of *Proteobacteria* found in the majority of UC samples in our study (**Figure 4.7**) is consistent with Papa *et al*. who investigated stool microbiome of newly diagnosed pediatric IBD patients.[71] *Campylobacter* in particular has been associated with relapse in UC,[72] and infection with this bacteria has been found to induce short- and long-term colitis in healthy populations.[73] Other notable genus of bacteria

enriched in UC patients were *Enterococcus* and *Veillonella*, both of which are known for their lactate producing abilities, which corresponds to the marked increase in lactate in fecal samples of UC patients in our study and multiple others.[74-76] *Enterococcus* is known as a highly stress-resistant opportunistic pathogen, whose colonization in the intestine seems to depend on lactate dehydrogenase, an enzyme capable of converting between pyruvate and lactate.[77] Association between disease activity and lactate concentration found in these studies may have resulted from such bacterial enzymes and associated pathogenic activities in colonizing epithelial surface that lacks protective mucous layer and ability to re-establish mucosal barrier.

Samples from CD patients were characterized by remarkable increase in bacterial metabolites such as indole and phenolic compounds in urine, and trimethyllysine and bile acids in fecal samples. Defects in primary bile acid absorption is expected in CD patients due to the disease involvement in ileum, where 95% of secreted bile acids are reabsorbed and sent back to the liver.[78] Approximately half of primary bile acids left in the intestine are normally deconjugated and/or hydroxylated by certain bacteria for their energy source.[79] In addition to cholate, other bile acids were also elevated in feces of CD patients in our study (**Table 4.2**), indicating the overall increase of unabsorbed bile acids compared to UC patients who lack disease involvement in their ileum. Trimethyllysine is a precursor of carnitine, which is essential for transport and utilization of fatty acid in mitochondria as an energy source.[80] Trimethyllysine had been traditionally considered to originate from degradation of protein that contains trimethylated lysine residues, but it was recently found abundant in a number of vegetables.[81] Therefore, high excretion of this compound by CD patient may indicate relatively higher consumption of vegetables containing this compound or higher degradation of proteins for producing carnitine to account for high energy demand posed by constant immune activation.

Urinary indole and phenolic sulfates that were abundant in urine of CD patients have been known as uremic solutes associated with chronic kidney disease[82] but have never previously been reported in the context of IBD. Indole and its metabolites in serum and fecal

samples, on the other hand, have been gaining attention as aryl hydrocarbon receptor (AHR) agonists, capable of inducing expression of genes involved in immune responses.[83] AHR is best known as a ligand-dependent transcription factor for mediating effects from exogenous toxins such as dioxin. Recent investigations on AHR ligands identified multiple endogenous compounds in tryptophan metabolism, namely, kynurenine and indole metabolites as human AHR agonists.[84-86] In the context of gastrointestinal disorder, contradicting benefits of AHR activation have been reported in literature. Indole-3-propionic acid[87] and indole-3-acetic acid[86] were found as AHR ligands that promoted intestinal homeostasis through activation of AHR and consequential regulation of IL-10 and IL-22 in vitro. Paradoxically, proinflammatory gene expression induced by significant up-regulation of AHR and overexpression of IL-22 was found in IBD patients,[88, 89] indicating that lack of homeostasis in the extent of AHR activation may play a crucial role in autoimmunity of IBD. In alignment with these findings, significant increase of kynurenine and consequent depletion of precursor tryptophan in serum has been found in IBD patients in disease activity-dependent manner, further indicating the involvement of AHR activation in the pathogenesis of IBD.[90] Furthermore, tryptophan is the sole precursor of serotonin, a neurotransmitter that plays a critical role in the regulation of mood and cognition, as well as gastrointestinal motility.[91] Indeed, most of serotonin in our body is produced in enterocytes[92] and there are numerous different serotonin receptors found in the intestinal wall.[93] There was moderately higher excretion of tryptophan in the stool of UC patients observed in this study, which may indicate defect in utilization of this amino acid in addition to increased transformation to kynurenine, both of which would lead to reduced production of serotonin. Increased activity of an enzyme responsible for the oxidation of tryptophan to kynurenine, indoleamine 2, 3-dioxygenase (IDO), has also been found in the inflamed mucosa of IBD patients.[90] High urinary excretion of kynurenine by UC patients observed in our study perhaps indicates IDO-mediated dysregulation of immune response more likely taking place in UC than CD patients. Conversely, significantly higher urinary excretion of indole metabolites in CD patients indicates increased conversion of tryptophan to indole through bacterial enzyme, tryptophanase, which consequently leaves less tryptophan available for kynurenine, as well as serotonin production.
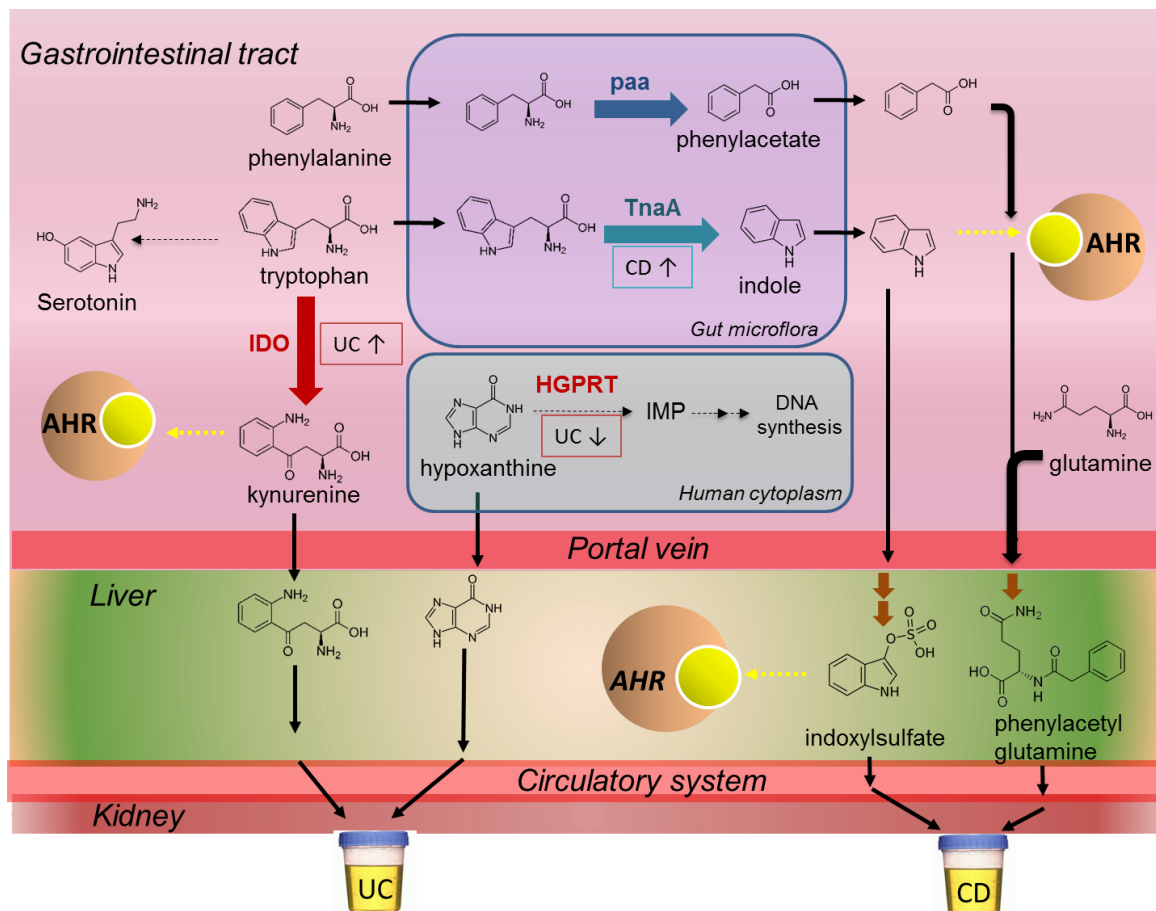
Figure 4.9. Proposed metabolic pathways associated with lead biomarker candidates identified in urine. Thick arrows represent enzymatic conversions and dotted arrow represent binding to AHR. AHR: Aryl hydrocarbon receptor; IDO: indoleamine 2, 3-dioxygenase; TnaA: Tryptophanase, paa: phenylacetic acid gene cluster;[94] HGPRT: hypoxanthine-guanine phosphoribosyltransferase.

Diverse strains of bacteria found in phylum *Proteobacteria* especially *Escherichia coli* carry tryptophanase encoding gene, tnaA.[95] Considering the high abundance of *Proteobacteria* in UC group and genus *Escherichia* in CD group in this study, both groups seem to harbour bacteria capable of indole production in their intestine (**Table S4.3**), provided that bacteria in fecal samples reflect bacterial community in the intestinal tract. The puzzling selective increase in indole and other microbial metabolites only in CD patients in our study may be explained by the difference in colonic transit time between UC and CD patients. In a study correlating urinary metabolites with colonic transit time of healthy individuals, urinary excretion of phenylacetylglutamine and indole compounds were highly

correlated with longer colonic transit time, which allowed bacterial transformation of amino acids into these compounds.[96] Aggressive diarrhea resulting from inflammation and increased secretion in colon is commonly seen in IBD especially in UC, which likely results in shorter colonic transit time.[97] Since colon and rectum harbour the largest number of bacteria in our gastrointestinal tract,[98] fast transit of luminal content through this region in UC may result in overall reduction of microbial metabolites as seen in our study. This assumption was supported by the significantly different overall bacterial functions predicted from 16S rRNA genes in fecal samples between UC and CD patients, and abundance of functions in amino acid metabolism especially for tryptophan and phenylalanine predicted from samples of CD patients (**Table S 4.4, 4.5**).

Our study is the first to report significant increase in urinary excretion of indoxylsulfate, phenol sulfate, other indole compounds and putatively identified long-chain fatty acid (*m/z* 345.1550) in CD patients as opposed to UC patients in pediatric population. Long-chain fatty acids are best known as energy storing molecules, but their metabolism has been recognized in association with modulation of inflammatory response.[99] Although exact identification is yet to be clarified, significantly higher excretion of this unknown fatty acid may reflect lipid metabolism-mediated inflammation that is preferentially taking place in CD patients. On the other hand, elevated excretion of the indole and phenolic compounds are expected to result in impaired cellular protective mechanisms against oxidative stress, induced by reduction of glutathione as seen in a recent in-vitro study.[100] Additionally, indoxylsulfate was identified as an extremely potent ligand of AHR in human hepatocyte, capable of altering immune response and drug metabolisms in the liver.[101] Taken together, these microbial compounds that are cpable of inducing oxidative stress in circulatory system, liver and kidneys may contribute to the extraintestinal manifestations frequently observed in pediatric IBD patients particularly in CD (**Figure 4.9**).[102, 103]

We reported differentially excreted urinary and fecal metabolites by CD and UC patients in pediatric population with microbial taxonomic and functional abundance analysis for the first time. Majority of findings in microbial and fecal metabolites were consistent with two of previous reports with a focus on pediatric, treatment-naïve IBD patients.[23, 104] We also presented novel perspectives on pathological mechanisms behind IBD that are

distinct between CD and UC. In short, presence of opportunistic pathogens, concurrent intestinal mucus depletion, impaired mucosal barrier function, and overexpression of enzyme capable of producing AHR ligands may play a central role in immune activation in UC patients while increased production of potentially toxic bacterial metabolites in the liver, kidney and circulatory system may contribute particularly to extra-intestinal symptoms of CD patients. Additionally, defects in cell proliferation to re-gain mucosal barrier function was implicated in both groups with involvement of different substrates for the enzyme, HGPRT, in purine pathway.

A major limitation in our study was the small and unbalanced sample size that prevented us from adjusting for confounding factors, especially the disease location of CD patients, which included mostly ileocolonic involvement but also upper gastrointestinal tract and isolated colonic cases. Additionally, lack of controls prevented us from determining abnormality of our findings in reference to 'healthy' levels. Finally, there was no dietary control or survey taken for this study, which hindered us from comparing levels of available precursors for metabolites identified in our study between UC and CD group. Nonetheless, our findings in urinary metabolome was supported by metabolite stability study and comparison of commonly used normalization parameters, osmolality and creatinine. Moreover, we showcased the effectiveness of urinary ratio metric features that expose correlational relationships between metabolites and cancels out any discrepancies that could be caused by different normalization factors. To confirm our results, follow-up study including pediatric CD, UC and chronic constipation as a control is on its way. Once validated, urinary metabolomics based markers hold potential for the least invasive and effective tool to differentiate UC and CD without colonoscopy or histological measurement. These markers may also be applied as prognostic and predictive markers to monitor the disease progress and effectively assign the suitable treatment based on the unique metabolic status of each patient.

Acknowledgements

4.6   References

(1)   Heyman, M. B.; Kirschner, B. S.; Gold, B. D.; Ferry, G.; Baldassano, R.; Cohen, S. A.; Winter, H. S.; Fain, P.; King, C.; Smith, T.; El-Serag, H. B., Children with early-onset inflammatory bowel disease (IBD): Analysis of a pediatric IBD consortium registry. *The Journal of Pediatrics* **2005,** *146* (1), 35-40.

(2)   Bernstein, C. N.; Fried, M.; Krabshuis, J. H.; Cohen, H.; Eliakim, R.; Fedail, S.; Gearry, R.; Goh, K. L.; Hamid, S.; Khan, A. G.; LeMair, A. W.; Malfertheiner; Ouyang, Q.; Rey, J. F.; Sood, A.; Steinwurz, F.; Thomsen, O. O.; Thomson, A.; Watermeyer, G., World Gastroenterology Organization Practice Guidelines for the Diagnosis and Management of IBD in 2010. *Inflamm. Bowel Dis.* **2010,** *16* (1), 112-124.

(3)   Keethy, D.; Mrakotsky, C.; Szigethy, E., Pediatric IBD and Depression: Treatment Implications. *Current opinion in pediatrics* **2014,** *26* (5), 561-567.

(4)   Punati, J.; Markowitz, J.; Lerer, T.; Hyams, J.; Kugathasan, S.; Griffiths, A.; Otley, A.; Rosh, J.; Pfefferkorn, M.; Mack, D.; Evans, J.; Bousvaros, A.; Moyer, S. M.; Wyllie, R.; Oliva-Hemker, M.; Mezoff, A.; LeLeiko, N.; Keljo, D.; Crandall, W., Effect of early immunomodulator use in moderate to severe pediatric Crohn disease. *Inflamm. Bowel Dis.* **2008,** *14* (7), 949-954.

(5)   Guariso, G.; Gasparetto, M., Treating children with inflammatory bowel disease: Current and new perspectives. *World Journal of Gastroenterology* **2017,** *23* (30), 5469-5485.

(6)   Russo, P., The Pathology of Chronic Inflammatory Bowel Disease In *Pediatric Inflammatory Bowel Disease* 2nd ed.; Mamula, P.; Markowitz, J.; Baldassano, R., Eds. Springer: 2013; pp 223-240.

(7)   Cuffari, C., Diagnostic Considerations in Pediatric Inflammatory Bowel Disease Management. *Gastroenterology & Hepatology* **2009,** *5* (11), 775-783.

(8)   Robert, M. E.; Skacel, M.; Ullman, T.; Bernstein, C. N.; Easley, K.; Goldblum, J. R., Patterns of Colonic Involvement at Initial Presentation in Ulcerative ColitisA Retrospective Study of 46 Newly Diagnosed Cases. *American Journal of Clinical Pathology* **2004,** *122* (1), 94-99.

(9)   Escher, J. C.; Kate, F. t.; Lichtenbelt, K.; Schornagel, I.; Büller, H.; Derkx, B.; Taminiau, J., Value of Rectosigmoidoscopy with Biopsies for Diagnosis of Inflammatory Bowel Disease in Children. *Inflamm. Bowel Dis.* **2002,** *8* (1), 16-22.

(10)   Glickman, J. N.; Bousvaros, A.; Farraye, F. A.; Zholudev, A.; Friedman, S.; Wang, H. H.; Leichtner, A. M.; Odze, R. D., Pediatric Patients With Untreated Ulcerative Colitis May

Present Initially With Unusual Morphologic Findings. *The American Journal of Surgical Pathology* **2004,** *28* (2), 190-197.

(11)   Haller, C. A.; Markowitz, J., IBD in children: Lessons for adults. *Current Gastroenterology Reports* **2007,** *9* (6), 528-532.

(12)   Warth, B.; Spangler, S.; Fang, M.; Johnson, C. H.; Forsberg, E. M.; Granados, A.; Martin, R. L.; Domingo-Almenara, X.; Huan, T.; Rinehart, D.; Montenegro-Burke, J. R.; Hilmers, B.; Aisporna, A.; Hoang, L. T.; Uritboonthai, W.; Benton, H. P.; Richardson, S. D.; Williams, A. J.; Siuzdak, G., Exposome-Scale Investigations Guided by Global Metabolomics, Pathway Analysis, and Cognitive Computing. *Anal. Chem.* **2017,** *89* (21), 11505-11513.

(13)   De Preter, V., Metabolomics in the Clinical Diagnosis of Inflammatory Bowel Disease. *Digestive Diseases* **2015,** *33(suppl 1)* (Suppl. 1), 2-10.

(14)   Marchesi, J. R.; Holmes, E.; Khan, F.; Kochhar, S.; Scanlan, P.; Shanahan, F.; Wilson, I. D.; Wang, Y., Rapid and Noninvasive Metabonomic Characterization of Inflammatory Bowel Disease. *J. Proteome Res.* **2007,** *6* (2), 546-551.

(15)   Thibault, R.; De Coppet, P.; Daly, K.; Bourreille, A.; Cuff, M.; Bonnet, C.; Mosnier, J. F.; Galmiche, J. P.; Shirazi–Beechey, S.; Segain, J. P., Down-Regulation of the Monocarboxylate Transporter 1 Is Involved in Butyrate Deficiency During Intestinal Inflammation. *Gastroenterology* **2007,** *133* (6), 1916-1927.

(16)   Duboc, H.; Rajca, S.; Rainteau, D.; Benarous, D.; Maubert, M.-A.; Quervain, E.; Thomas, G.; Barbu, V.; Humbert, L.; Despras, G.; Bridonneau, C.; Dumetz, F.; Grill, J.-P.; Masliah, J.; Beaugerie, L.; Cosnes, J.; Chazouillères, O.; Poupon, R.; Wolf, C.; Mallet, J.-M.; Langella, P.; Trugnan, G.; Sokol, H.; Seksik, P., Connecting dysbiosis, bile-acid dysmetabolism and gut inflammation in inflammatory bowel diseases. *Gut* **2013,** *62* (4), 531-539.

(17)   Scoville, E. A.; Allaman, M. M.; Brown, C. T.; Motley, A. K.; Horst, S. N.; Williams, C. S.; Koyama, T.; Zhao, Z.; Adams, D. W.; Beaulieu, D. B.; Schwartz, D. A.; Wilson, K. T.; Coburn, L. A., Alterations in lipid, amino acid, and energy metabolism distinguish Crohn's disease from ulcerative colitis and control subjects by serum metabolomic profiling. *Metabolomics* **2017,** *14* (1), 17.

(18)   Lu, K.; Knutson, C. G.; Wishnok, J. S.; Fox, J. G.; Tannenbaum, S. R., Serum Metabolomics in a Helicobacter hepaticus Mouse Model of Inflammatory Bowel Disease Reveal Important Changes in the Microbiome, Serum Peptides, and Intermediary Metabolism. *J. Proteome Res.* **2012,** *11* (10), 4916-4926.

(19)   Balasubramanian, K.; Kumar, S.; Singh, R. R.; Sharma, U.; Ahuja, V.; Makharia, G. K.; Jagannathan, N. R., Metabolism of the colonic mucosa in patients with inflammatory bowel diseases: an in vitro proton magnetic resonance spectroscopy study. *Magnetic Resonance Imaging* **2009,** *27* (1), 79-86.

(20)   Williams, H. R. T.; Cox, I. J.; Walker, D. G.; North, B. V.; Patel, V. M.; Marshall, S. E.; Jewell, D. P.; Ghosh, S.; Thomas, H. J. W.; Teare, J. P.; Jakobovits, S.; Zeki, S.; Welsh, K. I.; Taylor-Robinson, S. D.; Orchard, T. R., Characterization of Inflammatory Bowel Disease With Urinary Metabolic Profiling. *Am. J. Gastroenterol.* **2009,** *104* (6), 1435-1444.

(21)   Schicho, R.; Shaykhutdinov, R.; Ngo, J.; Nazyrova, A.; Schneider, C.; Panaccione, R.; Kaplan, G. G.; Vogel, H. J.; Storr, M., Quantitative Metabolomic Profiling of Serum, Plasma, and Urine by 1H NMR Spectroscopy Discriminates between Patients with Inflammatory Bowel Disease and Healthy Individuals. *J. Proteome Res.* **2012,** *11* (6), 3344-3357.

(22)   Stephens, N. S.; Siffledeen, J.; Su, X.; Murdoch, T. B.; Fedorak, R. N.; Slupsky, C. M., Urinary NMR metabolomic profiles discriminate inflammatory bowel disease from healthy. *J. Crohns Colitis* **2013,** *7* (2), e42-e48.

(23)   Kolho, K.-L.; Pessia, A.; Jaakkola, T.; de Vos, W. M.; Velagapudi, V., Faecal and Serum Metabolomics in Paediatric Inflammatory Bowel Disease. *J. Crohns Colitis* **2017,** *11* (3), 321-334.

(24)   Deda, O.; Gika, H. G.; Wilson, I. D.; Theodoridis, G. A., An overview of fecal sample preparation for global metabolic profiling. *J. Pharm. Biomed. Anal.* **2015,** *113*, 137-150.

(25)   De Cruz, P.; Kang, S.; Wagner, J.; Buckley, M.; Sim, W. H.; Prideaux, L.; Lockett, T.; McSweeney, C.; Morrison, M.; Kirkwood, C. D.; Kamm, M. A., Association between specific mucosa-associated microbiota in Crohn's disease at the time of resection and subsequent disease recurrence: A pilot study. *J. Gastroenterol. Hepatol.* **2015,** *30* (2), 268-278.

(26)   Loftfield, E.; Vogtmann, E.; Sampson, J. N.; Moore, S. C.; Nelson, H.; Knight, R.; Chia, N.; Sinha, R., Comparison of Collection Methods for Fecal Samples for Discovery Metabolomics in Epidemiologic Studies. *Cancer Epidemiology Biomarkers & Prevention* **2016,** *25* (11), 1483-1490.

(27)   Matysik, S.; Le Roy, C. I.; Liebisch, G.; Claus, S. P., Metabolomics of fecal samples: A practical consideration. *Trends in Food Science & Technology* **2016,** *57*, 244-255.

(28)   DiBattista, A.; McIntosh, N.; Lamoureux, M.; Al-Dirbashi, O. Y.; Chakraborty, P.; Britz-McKibbin, P., Temporal Signal Pattern Recognition in Mass Spectrometry: A Method for Rapid Identification and Accurate Quantification of Biomarkers for Inborn Errors of Metabolism with Quality Assurance. *Anal. Chem.* **2017,** *89* (15), 8112-8121.

(29)   DiBattista, A.; Rampersaud, D.; Lee, H.; Kim, M.; Britz-McKibbin, P., High Throughput Screening Method for Systematic Surveillance of Drugs of Abuse by Multisegment Injection–Capillary Electrophoresis–Mass Spectrometry. *Anal. Chem.* **2017,** *89* (21), 11853-11861.

(30)   Macedo, A. N.; Mathiaparanam, S.; Brick, L.; Keenan, K.; Gonska, T.; Pedder, L.; Hill, S.; Britz-McKibbin, P., The Sweat Metabolome of Screen-Positive Cystic Fibrosis Infants: Revealing Mechanisms beyond Impaired Chloride Transport. *ACS Central Science* **2017,** *3* (8), 904-913.

(31)   Ruttkies, C.; Schymanski, E. L.; Wolf, S.; Hollender, J.; Neumann, S., MetFrag relaunched: incorporating strategies beyond in silico fragmentation. *J. Cheminform.* **2016,** *8* (1), 3.

(32)   Bligh, E. G.; Dyer, W. J., A RAPID METHOD OF TOTAL LIPID EXTRACTION AND PURIFICATION. *Canadian Journal of Biochemistry and Physiology* **1959,** *37* (8), 911-917.

(33)   Gao, X.; Pujos-Guillot, E.; Sébédio, J.-L., Development of a Quantitative Metabolomic Approach to Study Clinical Human Fecal Water Metabolome Based on Trimethylsilylation Derivatization and GC/MS Analysis. *Anal. Chem.* **2010,** *82* (15), 6447-6456.

(34)   Kuehnbaum, N. L.; Kormendi, A.; Britz-McKibbin, P., Multisegment Injection-Capillary Electrophoresis-Mass Spectrometry: A High-Throughput Platform for Metabolomics with High Data Fidelity. *Anal. Chem.* **2013,** *85* (22), 10664-10669.

(35)   Wishart, D. S.; Feunang, Y. D.; Marcu, A.; Guo, A. C.; Liang, K.; Vazquez-Fresno, R.; Sajed, T.; Johnson, D.; Li, C.; Karu, N.; Sayeeda, Z.; Lo, E.; Assempour, N.; Berjanskii, M.; Singhal, S.; Arndt, D.; Liang, Y.; Badran, H.; Grant, J.; Serra-Cayuela, A.; Liu, Y.; Mandal, R.; Neveu, V.; Pon, A.; Knox, C.; Wilson, M.; Manach, C.; Scalbert, A., HMDB

4.0: the human metabolome database for 2018. *Nucleic Acids Res.* **2018,** *46* (D1), D608-d617.

(36)   Chong, J.; Soufan, O.; Li, C.; Caraus, I.; Li, S.; Bourque, G.; Wishart, D. S.; Xia, J., MetaboAnalyst 4.0: towards more transparent and integrative metabolomics analysis. *Nucleic Acids Res.* **2018,** *46* (W1), W486-W494.

(37)   Stearns, J. C.; Davidson, C. J.; McKeon, S.; Whelan, F. J.; Fontes, M. E.; Schryvers, A. B.; Bowdish, D. M. E.; Kellner, J. D.; Surette, M. G., Culture and molecular-based profiles show shifts in bacterial communities of the upper respiratory tract that occur with age. *The Isme Journal* **2015,** *9*, 1246.

(38)   Bartram, A. K.; Lynch, M. D. J.; Stearns, J. C.; Moreno-Hagelsieb, G.; Neufeld, J. D., Generation of Multimillion-Sequence 16S rRNA Gene Libraries from Complex Microbial Communities by Assembling Paired-End Illumina Reads. *Applied and Environmental Microbiology* **2011,** *77* (11), 3846.

(39)   Whelan, F. J.; Surette, M. G., A comprehensive evaluation of the sl1p pipeline for 16S rRNA gene sequencing analysis. *Microbiome* **2017,** *5* (1), 100.

(40)   McMurdie, P. J.; Holmes, S., Phyloseq: An R Package for reproducible interactive analysis and graphics of microbiome census data. *PLoS One* **2013,** *8*.

(41)   Langille, M. G. I.; Zaneveld, J.; Caporaso, J. G.; McDonald, D.; Knights, D.; Reyes, J. A.; Clemente, J. C.; Burkepile, D. E.; Vega Thurber, R. L.; Knight, R.; Beiko, R. G.; Huttenhower, C., Predictive functional profiling of microbial communities using 16S rRNA marker gene sequences. *Nat. Biotechnol.* **2013,** *31*, 814.

(42)   Dhariwal, A.; Chong, J.; Habib, S.; King, I. L.; Agellon, L. B.; Xia, J., MicrobiomeAnalyst: a web-based tool for comprehensive statistical, visual and meta-analysis of microbiome data. *Nucleic Acids Res.* **2017,** *45* (W1), W180-W188.

(43)   Segata, N.; Izard, J.; Waldron, L.; Gevers, D.; Miropolsky, L.; Garrett, W. S.; Huttenhower, C., Metagenomic biomarker discovery and explanation. *Genome Biology* **2011,** *12* (6), R60-R60.

(44)   Lin, C. Y.; Wu, H.; Tjeerdema, R. S.; Viant, M. R., Evaluation of metabolite extraction strategies from tissue samples using NMR metabolomics. *Metabolomics* **2007,** *3* (1), 55-67.

(45)   Brunius, C.; Shi, L.; Landberg, R., Large-scale untargeted LC-MS metabolomics data correction using between-batch feature alignment and cluster-based within-batch signal intensity drift correction. *Metabolomics* **2016,** *12* (11), 173.

(46)   Warrack, B. M.; Hnatyshyn, S.; Ott, K. H.; Reily, M. D.; Sanders, M.; Zhang, H.; Drexler, D. M., Normalization strategies for metabonomic analysis of urine samples. *J Chromatogr B Analyt Technol Biomed Life Sci* **2009,** *877* (5-6), 547-52.

(47)   Alessio, L.; Berlin, A.; Dell'Orto, A.; Toffoletto, F.; Ghezzi, I., Reliability of urinary creatinine as a parameter used to adjust values of urinary biological indicators. *International Archives of Occupational and Environmental Health* **1985,** *55* (2), 99-106.

(48)   Zhao, Y.; Wu, J.; Li, J. V.; Zhou, N. Y.; Tang, H.; Wang, Y., Gut microbiota composition modifies fecal metabolic profiles in mice. *J. Proteome Res.* **2013,** *12* (6), 2987-99.

(49)   Li, J. V.; Saric, J.; Wang, Y.; Keiser, J.; Utzinger, J.; Holmes, E., Chemometric analysis of biofluids from mice experimentally infected with Schistosoma mansoni. *Parasit Vectors* **2011,** *4*, 179.

(50)   Wu, J.; An, Y.; Yao, J.; Wang, Y.; Tang, H., An optimised sample preparation method for NMR-based faecal metabonomic analysis. *Analyst* **2010,** *135* (5), 1023-30.

(51)   Romick-Rosendale, L. E.; Goodpaster, A. M.; Hanwright, P. J.; Patel, N. B.; Wheeler, E. T.; Chona, D. L.; Kennedy, M. A., NMR-based metabonomics analysis of mouse urine and

fecal extracts following oral treatment with the broad-spectrum antibiotic enrofloxacin (Baytril). *Magn. Reson. Chem.* **2009,** *47 Suppl 1*, S36-46.

(52)   Calvani, R.; Brasili, E.; Pratico, G.; Capuani, G.; Tomassini, A.; Marini, F.; Sciubba, F.; Finamore, A.; Roselli, M.; Marzetti, E.; Miccheli, A., Fecal and urinary NMR-based metabolomics unveil an aging signature in mice. *Exp Gerontol* **2014,** *49*, 5-11.

(53)   Hofmann, A. F.; Hagey, L. R., Bile Acids: Chemistry, Pathochemistry, Biology, Pathobiology, and Therapeutics. *Cellular and Molecular Life Sciences* **2008,** *65* (16), 2461-2483.

(54)   Bondia-Pons, I.; Maukonen, J.; Mattila, I.; Rissanen, A.; Saarela, M.; Kaprio, J.; Hakkarainen, A.; Lundbom, J.; Lundbom, N.; Hyötyläinen, T.; Pietiläinen, K. H.; Orešič, M., Metabolome and fecal microbiota in monozygotic twin pairs discordant for weight: a Big Mac challenge. *The FASEB Journal* **2014,** *28* (9), 4169-4179.

(55)   Gratton, J.; Phetcharaburanin, J.; Mullish, B. H.; Williams, H. R. T.; Thursz, M.; Nicholson, J. K.; Holmes, E.; Marchesi, J. R.; Li, J. V., Optimized Sample Handling Strategy for Metabolic Profiling of Human Feces. *Anal. Chem.* **2016,** *88* (9), 4661-4668.

(56)   Ajouz, H.; Mukherji, D.; Shamseddine, A., Secondary bile acids: an underrecognized cause of colon cancer. *World Journal of Surgical Oncology* **2014,** *12* (1), 164.

(57)   Walsh, M. C.; Brennan, L.; Malthouse, J. P. G.; Roche, H. M.; Gibney, M. J., Effect of acute dietary standardization on the urinary, plasma, and salivary metabolomic profiles of healthy humans. *The American Journal of Clinical Nutrition* **2006,** *84* (3), 531-539.

(58)   Orešič, M.; Simell, S.; Sysi-Aho, M.; Näntö-Salonen, K.; Seppänen-Laakso, T.; Parikka, V.; Katajamaa, M.; Hekkala, A.; Mattila, I.; Keskinen, P.; Yetukuri, L.; Reinikainen, A.; Lähde, J.; Suortti, T.; Hakalax, J.; Simell, T.; Hyöty, H.; Veijola, R.; Ilonen, J.; Lahesmaa, R.; Knip, M.; Simell, O., Dysregulation of lipid and amino acid metabolism precedes islet autoimmunity in children who later progress to type 1 diabetes. *The Journal of Experimental Medicine* **2008,** *205* (13), 2975.

(59)   Brown, D. G.; Rao, S.; Weir, T. L.; O'Malia, J.; Bazan, M.; Brown, R. J.; Ryan, E. P., Metabolomics and metabolic pathway networks from human colorectal cancers, adjacent mucosa, and stool. *Cancer & Metabolism* **2016,** *4*, 11.

(60)   Coskun, M.; Olsen, J.; Seidelin, J. B.; Nielsen, O. H., MAP kinases in inflammatory bowel disease. *Clin. Chim. Acta* **2011,** *412* (7), 513-520.

(61)   Monteleone, G.; Pallone, F.; MacDonald, T. T., Smad7 in TGF-β-mediated negative regulation of gut inflammation. *Trends in Immunology* **2004,** *25* (10), 513-517.

(62)   De Cruz, P.; Kamm, M. A.; Prideaux, L.; Allen, P. B.; Moore, G., Mucosal healing in Crohn's disease: A systematic review. *Inflamm. Bowel Dis.* **2012**, n/a-n/a.

(63)   Dorofeyev, A. E.; Vasilenko, I. V.; Rassokhina, O. A.; Kondratiuk, R. B., Mucosal Barrier in Ulcerative Colitis and Crohn’s Disease. *Gastroenterology Research and Practice* **2013,** *2013*, 9.

(64)   Lee, J. S.; Colgan, S. P., Purine Metabolism and Barrier Formation in Intestinal Epithelial Cells. *The FASEB Journal* **2017,** *31* (1_supplement), 976.1-976.1.

(65)   Nakano, M.; Tateishi-Karimata, H.; Tanaka, S.; Sugimoto, N., Choline Ion Interactions with DNA Atoms Explain Unique Stabilization of A–T Base Pairs in DNA Duplexes: A Microscopic View. *The Journal of Physical Chemistry B* **2014,** *118* (2), 379-389.

(66)   Johansson, M. E. V.; Gustafsson, J. K.; Holmén-Larsson, J.; Jabbar, K. S.; Xia, L.; Xu, H.; Ghishan, F. K.; Carvalho, F. A.; Gewirtz, A. T.; Sjövall, H.; Hansson, G. C., Bacteria penetrate the normally impenetrable inner colon mucus layer in both murine colitis models and patients with ulcerative colitis. *Gut* **2014,** *63* (2), 281.

(67)   Breeling, J. L.; Onderdonk, A. B.; Cisneros, R. L.; Kasper, D. L., Bacteroides vulgatus outer membrane antigens associated with carrageenan-induced colitis in guinea pigs. *Infect. Immun.* **1988,** *56* (7), 1754-1759.

(68)   Wright, D. P.; Rosendale, D. I.; Roberton, A. M., Prevotella enzymes involved in mucin oligosaccharide degradation and evidence for a small operon of genes expressed during growth on mucin. *FEMS Microbiol. Lett.* **2000,** *190* (1), 73-79.

(69)   Lucke, K.; Miehlke, S.; Jacobs, E.; Schuppler, M., Prevalence of Bacteroides and Prevotella spp. in ulcerative colitis. *Journal of Medical Microbiology* **2006,** *55* (5), 617-624.

(70)   Eckburg, P. B.; Bik, E. M.; Bernstein, C. N.; Purdom, E.; Dethlefsen, L.; Sargent, M., Diversity of the human intestinal microbial flora. *Science.* **2005,** *308*.

(71)   Papa, E.; Docktor, M.; Smillie, C.; Weber, S.; Preheim, S. P.; Gevers, D.; Giannoukos, G.; Ciulla, D.; Tabbaa, D.; Ingram, J.; Schauer, D. B.; Ward, D. V.; Korzenik, J. R.; Xavier, R. J.; Bousvaros, A.; Alm, E. J., Non-Invasive Mapping of the Gastrointestinal Microbiota Identifies Children with Inflammatory Bowel Disease. *PLOS ONE* **2012,** *7* (6), e39242.

(72)   Campieri, M.; Gionchetti, P., Bacteria as the cause of ulcerative colitis. *Gut* **2001,** *48* (1), 132.

(73)   Gradel, K. O.; Nielsen, H. L.; Schønheyder, H. C.; Ejlertsen, T.; Kristensen, B.; Nielsen, H., Increased Short- and Long-Term Risk of Inflammatory Bowel Disease After Salmonella or Campylobacter Gastroenteritis. *Gastroenterology* **2009,** *137* (2), 495-501.

(74)   Le Gall, G.; Noor, S. O.; Ridgway, K.; Scovell, L.; Jamieson, C.; Johnson, I. T.; Colquhoun, I. J.; Kemsley, E. K.; Narbad, A., Metabolomics of Fecal Extracts Detects Altered Metabolic Activity of Gut Microbiota in Ulcerative Colitis and Irritable Bowel Syndrome. *J. Proteome Res.* **2011,** *10* (9), 4208-4218.

(75)   Vernia, P.; Caprilli, R.; Latella, G.; Barbetti, F.; Magliocca, F. M.; Cittadini, M., Fecal Lactate and Ulcerative Colitis. *Gastroenterology* **1988,** *95* (6), 1564-1568.

(76)   Bjerrum, J. T.; Wang, Y.; Hao, F.; Coskun, M.; Ludwig, C.; Günther, U.; Nielsen, O. H., Metabonomics of human fecal extracts characterize ulcerative colitis, Crohn's disease and healthy individuals. *Metabolomics* **2015,** *11* (1), 122-133.

(77)   Rana, N. F.; Sauvageot, N.; Laplace, J.-M.; Bao, Y.; Nes, I.; Rincé, A.; Posteraro, B.; Sanguinetti, M.; Hartke, A., Redox Balance via Lactate Dehydrogenase is Important for Multiple Stress Resistance and Virulence in Enterococcus faecalis. *Infect. Immun.* **2013**.

(78)   Cipriani, S.; Mencarelli, A.; Chini, M. G.; Distrutti, E.; Renga, B.; Bifulco, G.; Baldelli, F.; Donini, A.; Fiorucci, S., The Bile Acid Receptor GPBAR-1 (TGR5) Modulates Integrity of Intestinal Barrier and Immune Response to Experimental Colitis. *PLoS ONE* **2011,** *6* (10), e25637.

(79)   Martínez-Augustin, O.; de Medina, F. S., Intestinal bile acid physiology and pathophysiology. *World J. Gastroenterol.* **2008,** *14* (37), 5630-5640.

(80)   Bremer, J., Carnitine--metabolism and functions. *Physiol. Rev.* **1983,** *63* (4), 1420-1480.

(81)   Servillo, L.; Giovane, A.; Cautela, D.; Castaldo, D.; Balestrieri, M. L., Where Does Nε-Trimethyllysine for the Carnitine Biosynthesis in Mammals Come from? *PLOS ONE* **2014,** *9* (1), e84589.

(82)   Leong, S.; Sirich, T., Indoxyl Sulfate—Review of Toxicity and Therapeutic Strategies. *Toxins* **2016,** *8* (12), 358.

(83)   Schmidt, J. V.; Bradfield, C. A., AH RECEPTOR SIGNALING PATHWAYS. *Annu. Rev. Cell Dev. Biol.* **1996,** *12* (1), 55-89.

(84)   Opitz, C. A.; Litzenburger, U. M.; Sahm, F.; Ott, M.; Tritschler, I.; Trump, S.; Schumacher, T.; Jestaedt, L.; Schrenk, D.; Weller, M.; Jugold, M.; Guillemin, G. J.; Miller, C. L.; Lutz, C.; Radlwimmer, B.; Lehmann, I.; von Deimling, A.; Wick, W.; Platten, M., An endogenous tumour-promoting ligand of the human aryl hydrocarbon receptor. *Nature* **2011,** *478*, 197.

(85)   Hubbard, T. D.; Murray, I. A.; Bisson, W. H.; Lahoti, T. S.; Gowda, K.; Amin, S. G.; Patterson, A. D.; Perdew, G. H., Adaptation of the human aryl hydrocarbon receptor to sense microbiota-derived indoles. *Sci. Rep.* **2015,** *5*, 12689.

(86)   Lamas, B.; Richard, M. L.; Leducq, V.; Pham, H.-P.; Michel, M.-L.; Da Costa, G.; Bridonneau, C.; Jegou, S.; Hoffmann, T. W.; Natividad, J. M.; Brot, L.; Taleb, S.; Couturier-Maillard, A.; Nion-Larmurier, I.; Merabtene, F.; Seksik, P.; Bourrier, A.; Cosnes, J.; Ryffel, B.; Beaugerie, L.; Launay, J.-M.; Langella, P.; Xavier, R. J.; Sokol, H., CARD9 impacts colitis by altering gut microbiota metabolism of tryptophan into aryl hydrocarbon receptor ligands. *Nature Medicine* **2016,** *22*, 598.

(87)   Alexeev, E. E.; Lanis, J. M.; Kao, D. J.; Campbell, E. L.; Kelly, C. J.; Battista, K. D.; Gerich, M. E.; Jenkins, B. R.; Walk, S. T.; Kominsky, D. J.; Colgan, S. P., Microbiota-Derived Indole Metabolites Promote Human and Murine Intestinal Homeostasis through Regulation of Interleukin-10 Receptor. *The American Journal of Pathology* **2018,** *188* (5), 1183-1194.

(88)   Arsenescu, R.; Arsenescu, V.; Zhong, J.; Nasser, M.; Melinte, R.; Dingle, C. R. W.; Swanson, H.; de Villiers, W. J., Role of the xenobiotic receptor in inflammatory bowel disease. *Inflamm. Bowel Dis.* **2011,** *17* (5), 1149-1162.

(89)   Brand, S.; Beigel, F.; Olszak, T.; Zitzmann, K.; Eichhorst, S. T.; Otte, J.-M.; Diepolder, H.; Marquardt, A.; Jagla, W.; Popp, A.; Leclair, S.; Herrmann, K.; Seiderer, J.; Ochsenkühn, T.; Göke, B.; Auernhammer, C. J.; Dambacher, J., IL-22 is increased in active Crohn's disease and promotes proinflammatory gene expression and intestinal epithelial cell migration. *American Journal of Physiology-Gastrointestinal and Liver Physiology* **2006,** *290* (4), G827-G838.

(90)   Nikolaus, S.; Schulte, B.; Al-Massad, N.; Thieme, F.; Schulte, D. M.; Bethge, J.; Rehman, A.; Tran, F.; Aden, K.; Häsler, R.; Moll, N.; Schütze, G.; Schwarz, M. J.; Waetzig, G. H.; Rosenstiel, P.; Krawczak, M.; Szymczak, S.; Schreiber, S., Increased Tryptophan Metabolism Is Associated With Activity of Inflammatory Bowel Diseases. *Gastroenterology* **2017,** *153* (6), 1504-1516.e2.

(91)   Jenkins, T. A.; Nguyen, J. C. D.; Polglaze, K. E.; Bertrand, P. P., Influence of Tryptophan and Serotonin on Mood and Cognition with a Possible Role of the Gut-Brain Axis. *Nutrients* **2016,** *8* (1), 56.

(92)   Racké, K.; Schwörer, H., Regulation of serotonin release from the intestinal mucosa. *Pharmacol. Res.* **1991,** *23* (1), 13-25.

(93)   Bornstein, J. C., Serotonin in the Gut: What Does It Do? *Frontiers in Neuroscience* **2012,** *6*, 16.

(94)   Olivera, E. R.; Miñambres, B.; García, B.; Muñiz, C.; Moreno, M. A.; Ferrández, A.; Díaz, E.; García, J. L.; Luengo, J. M., Molecular characterization of the phenylacetic acid catabolic pathway in Pseudomonas putida U: The phenylacetyl-CoA catabolon. *PNAS* **1998,** *95* (11), 6419.

(95)   DeMoss, R. D.; Moser, K., Tryptophanase in Diverse Bacterial Species. *J. Bacteriol.* **1969,** *98* (1), 167.

(96)   Roager, H. M.; B. S. Hansen, L.; I. Bahl, M.; Frandsen, H.; Carvalho, V.; Gøbel, R.; Dalgaard, M.; Plichta, D.; H. Sparholt, M.; Vestergaard, H.; Hansen, T.; Sicheritz-Ponten,

T.; Nielsen, H.; Pedersen, O.; Lauritzen, L.; Kristensen, M.; Gupta, R.; Licht, T., Colonic transit time is related to bacterial metabolism and mucosal turnover in the gut. *Nat. Microbiol.* **2016,** *1*, 16093.

(97)   Reddy, S. N.; Bazzocchi, G.; Chan, S.; Akashi, K.; Villanueva-Meyer, J.; Yanni, G.; Mena, I.; Snape, W. J., Jr., Colonic motility and transit in health and ulcerative colitis. *Gastroenterology* **1991,** *101* (5), 1289-1297.

(98)   Tannock, G. W., *Normal Microbiota: An Introduction to Microbes Inhabiting the Human Body*. Chapman & Hall London, UK, 1995.

(99)   Varga, T.; Czimmerer, Z.; Nagy, L., PPARs are a unique set of fatty acid regulated transcription factors controlling both lipid metabolism and inflammation. *Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease* **2011,** *1812* (8), 1007-1022.

(100)   Edamatsu, T.; Fujieda, A.; Itoh, Y., Phenyl sulfate, indoxyl sulfate and p-cresyl sulfate decrease glutathione level to render cells vulnerable to oxidative stress in renal tubular cells. *PLOS ONE* **2018,** *13* (2), e0193342.

(101)   Schroeder, J. C.; DiNatale, B. C.; Murray, I. A.; Flaveny, C. A.; Liu, Q.; Laurenzana, E. M.; Lin, J. M.; Strom, S. C.; Omiecinski, C. J.; Amin, S.; Perdew, G. H., The Uremic Toxin 3-Indoxyl Sulfate Is a Potent Endogenous Agonist for the Human Aryl Hydrocarbon Receptor. *Biochemistry* **2010,** *49* (2), 393-400.

(102)   Rabizadeh, S.; Oliva-Hemker, M., Extraintestinal Manifestations of Pediatric Inflammatory Bowel Disease In *Pediatric Inflammatory Bowel Disease* 2nd ed.; Mamula, P.; Markowitz, J. E.; Baldassano, R., Eds. Springer: 2013.

(103)   Vavricka, S. R.; Schoepfer, A.; Scharl, M.; Lakatos, P. L.; Navarini, A.; Rogler, G., Extraintestinal Manifestations of Inflammatory Bowel Disease. *Inflamm. Bowel Dis.* **2015,** *21* (8), 1982-1992.

(104)   Gevers, D.; Kugathasan, S.; Denson, Lee A.; Vázquez-Baeza, Y.; Van Treuren, W.; Ren, B.; Schwager, E.; Knights, D.; Song, Se J.; Yassour, M.; Morgan, Xochitl C.; Kostic, Aleksandar D.; Luo, C.; González, A.; McDonald, D.; Haberman, Y.; Walters, T.; Baker, S.; Rosh, J.; Stephens, M.; Heyman, M.; Markowitz, J.; Baldassano, R.; Griffiths, A.; Sylvester, F.; Mack, D.; Kim, S.; Crandall, W.; Hyams, J.; Huttenhower, C.; Knight, R.; Xavier, Ramnik J., The Treatment-Naive Microbiome in New-Onset Crohn's Disease. *Cell Host & Microbe* **2014,** *15* (3), 382-392.

4.7 Supporting Information

Table S4.1. Technical precision (*n*=6) for extraction protocol from lyophilized stool calculated for representative cationic and anionic metabolites based on pooling together two separate extracts.

Chemical ID	Ion mode	%CV (n = 6)
Alanine	Positive	4.6
Carnitine	Positive	4.2
Choline	Positive	4.3
Glutamate	Positive	5.0
Glutamine	Positive	5.0
Glycine	Positive	4.5
Isoleucine	Positive	4.9
Leucine	Positive	4.9
Tryptophan	Positive	9.1
Butyrate	Negative	8.8
Cholate	Negative	7.2
Propionate	Negative	5.2
Succinate	Negative	3.7
Valerate	Negative	5.6
3-Hydroxybutyrate	Negative	4.6
3-Hydroxyphenylacetate	Negative	7.0

Table S4.2. Top-ranked ratiometric biomarkers identified by MSI-CE-MS that differentiate pediatric CD from UC in osmolality normalized urine.

Ratiometric Biomarkers	*p*-value	FC	Effect size	*q*-value
Serine/Indoxylsulfate	1.71 E-05	0.18	0.73	1.86 E-03
Threonine/Phenylacetylglutamine	2.70 E-05	0.24	0.72	1.86 E-03
Serine/Hydroxyindole sulfate	4.05 E-05	0.19	0.71	1.86 E-03
Kynurenine/Indoxylsulfate	1.23 E-04	0.08	0.67	3.90 E-03
Serine/Phenylacetylglutamine	1.69 E-04	0.37	0.66	3.90 E-03
Threonine/Hydroxyindole sulfate	1.69 E-04	0.10	0.66	3.90 E-03
Threonine/Indoxylsulfate	2.32 E-04	0.09	0.65	4.58 E-03
Tryptophan /Indoxylsulfate	7.16 E-04	0.32	0.61	9.88 E-03
Kynurenine/Hydroxyindole sulfate	9.23 E-04	0.11	0.60	0.0106
Tryptophan/Hydroxyindole sulfate	1.50 E-03	0.28	0.58	> 0.05

* Statistical significance calculated by a Mann-Whitney U-test, p < 0.05, whereas q-value is based on Benjamini-Hochberg FDR; Abbreviation: FC = median fold-change

Table S4.3. Differentially enriched bacterial taxa identified by LEfSe.

IBD subtype	Genus	Phylum	LDA score	*p*-value
CD	*Ruminococcus*	*Firmicutes*	2.95	0.014
	Blautia	*Firmicutes*	2.86	0.036
	Clostridium	*Firmicutes*	2.46	0.030
	Bifidobacterium	*Actinobacteria*	2.14	> 0.05
	Escherichia	*Proteobacteria*	2.10	> 0.05
	Faecalibacterium	*Firmicutes*	2.03	> 0.05
	Sutterella	*Proteobacteria*	1.93	> 0.05
	Parabacteroides	*Bacteroidetes*	1.90	0.0041
	Alistipes	*Bacteroidetes*	1.88	> 0.05
	Fusobacterium	*Fusobacteria*	1.61	> 0.05
	Gamella	*Firmicutes*	1.61	> 0.05
	Pseudobutyvibrio	*Firmicutes*	0.72	0.008
UC	Unidentified taxa	NA	3.11	0.011
	Bacteroides	*Bacteroidetes*	2.43	> 0.05
	Klebsiella	*Proteobacteria*	2.40	> 0.05
	Streprococcus	*Firmicutes*	2.18	> 0.05
	Peptostreptococcus	*Firmicutes*	2.16	> 0.05
	Enterococcus	*Firmicutes*	2.03	> 0.05
	Veillonella	*Firmicutes*	2.00	> 0.05
	Aggregatibacter	*Proteobacteria*	2.00	> 0.05
	Porphyromonas	*Bacteroidetes*	1.90	> 0.05
	Erwinia	*Proteobacteria*	1.60	> 0.05
	Prevotella	*Bacteroidetes*	1.54	> 0.05
	Peptoniphilus	*Firmicutes*	1.50	> 0.05
	Haemophilus	*Proteobacteria*	1.50	> 0.05
	Peptococcus	*Firmicutes*	0.67	0.020
	Campylobacter	*Proteobacteria*	0.55	0.042

Abbreviation: LDA = Linear discriminate analysis; LEfSe = Linear discriminant analysis effect size

Table S4.4. Predicted metagenomics functions in each group differentiated by LEfSe analysis.

CD Differential functional group (KEGG level 3)	LDA score	UC Differential functional group (KEGG level 3)	LDA score
Peptidoglycan biosynthesis and degradation	2.99	Glutathione metabolism	2.11
Pentose phosphate pathway	2.97	Pyruvate metabolism	2.08
Amino sugar and nucleotide sugar metabolism	2.97	Uniquinone and other terpenoid-quinone biosynthesis	2.06
Carbon fixation pathways in prokaryotes	2.87	Pentose and glucuronate interconversions	2.02
Pyruvate metabolism	2.80	Riboflavin metabolism	2.01
Valine, leucine and isoleucine biosynthesis	2.78		
Alanine, aspartate and glutamate metabolism	2.75		
Arginine biosynthesis	2.69		
Phenylalanine, tyrosine and tryptophan biosynthesis	2.68		
Cysteine and methionine metabolism	2.68		
Starch and sucrose metabolism	2.67		

Abbreviation: LEfSe = Linear discriminant analysis effect size

Table S4.5 Pathway enriched in identified functions from CD stool samples. Number of hits in the KEGG database for the pathway and associated *p*-values (Fisher's exact test) are shown.

CD Differential functional group	Hits	*p*-value
Biosynthesis of amino acids	30	2.11 E-09****
Histidine metabolism	9	9.2 E-06***
Phenylalanine, tyrosine and tryptophan biosynthesis	9	0.0013
Glycerolipid metabolism	7	0.0096
C5-Branched dibasic acid metabolism	3	0.0098
Nicotinate and nicotinamide metabolism	5	0.017
Amino sugar and nucleotide sugar metabolism	7	0.018
Valine, leucine and isoleucine biosynthesis	3	0.024
Tetracycline biosynthesis	2	0.024
2-Oxocarboxylic acid metabolism	6	0.032
Fructose and mannose metabolism	5	0.037
Carbon fixation pathways in prokaryotes	6	0.040

**** False discovery rate corrected p-value (FDR) < 0.0001, ***FDR < 0.001

Table S4.6. Summary of 104 stool metabolites measured in samples of pediatric IBD patients with *m/z*, RMT, ionization mode (p = ESI positive ion mode, n = ESI negative ion mode), molecular formula, chemical ID or tentative ID, confidence level of the ID and chemical classification based on HMDB. Column "common" indicates the common metabolites found in both urine and stool metabolite (1 = commonly found, 0 = not found in urine).

*m.z:*RMT:mode	Chemical ID	Molecular Formula	Metabolite Class	Confirmed level	Common
87.0452:1.054:n	Butyric acid [a]	C4H8O2	organic acid	1	1
88.0404:1.005:n	Sarcosine	C3H7NO2	amino acid	3	0
89.0244:1.1400:n	Lactic acid	C3H6O3	organic acid	1	1
101.0608:0.99:n	Ethylmethylacetic acid [a]	C5H10O2	organic acid	3	0
117.0193:1.844:n	Succinic acid	C4H6O4	organic acid	1	1
117.0555:0.989:n	2-Methyl-3-hydroxybutyric or 3-Hydroxyisovaleric acid	C5H10O3	organic acid	3	0
133.0142:1.862:n	Malic acid	C4H6O5	organic acid	1	0
145.0514:1.456:n	2-Methylglutaric acid [a]	C6H10O4	organic acid	2	1
146.0459:0.956:n	Glutamic acid	C5H9NO4	amino acid	1	0
147.0299:1.618:n	2-Hydroxyglutaric acid	C5H8O5	organic acid	2	0
149.0608:0.943:n	Hydrocinnamic acid/phenylpropanoic acid [a]	C9H10O2	organic acid	3	0
167.0201:0.962:n	Uric acid	C5H4N4O3	organic acid	1	1
178.051:0.911:n	Hippuric acid [a]	C9H9NO3	organic acid	1	1
182.0459:0.942:n	4-Pyridoxic acid	C8H9NO4	organic acid	1	1
187.0071:1.059:n	*p*-Cresol sulfate	C7H8O4S	phenylsulfate	1	1
188.0558:1.354:n	*N*-Acetylglutamate	C7H11NO5	amino acid	3	1
218.1034:0.832:n	Pantothenic acid	C9H17NO5	organonitrogen	1	1
308.0987:0.788:n	*N*-Acetylneuraminic acid	C11H19NO9	amino sugar	1	1
391.2865:0.755:n	Deoxycholic acid [a]	C24H40O4	steroid	1	0
405.2646:0.753:n	7-Ketodeoxycholic acid [a]	C24H38O5	steroid	3	0
407.2803:0.748:n	Cholic acid	C24H40O5	steroid	1	1
448.3068:0.739:n	Deoxycholic acid glycine conj.	C26H43NO5	steroid	1	0
464.3018:0.736:n	Glycocholic acid	C26H43NO6	steroid	1	0
471.2422:0.993:n	Chenodeoxycholic acid 3-sulfate [a]	C24H40O7S	steroid	2	0
498.2895:0.741:n	Taurodeoxycholic acid [a]	C26H45NO6S	steroid	1	0
514.2844:0.738:n	Taurocholic acid	C26H45NO7S	steroid	1	0
76.0393:0.704:p	Glycine	C2H5NO2	amino acid	1	1
89.1079:0.402:p	Putrescine	C4H12N2	amine	1	0
90.055:0.747:p	Alanine	C3H7NO2	amino acid	1	1
103.1236:0.424:p	Cadaverine	C5H14N2	amine	1	0
104.0706:0.664:p	Gamma-aminobutyrate	C4H9NO2	amino acid	1	1
104.1069:0.56:p	Choline	C5H14NO	organonitrogen	1	1
106.0499:0.844:p	Serine	C3H7NO3	amino acid	1	1

112.0875:0.410:p	Histamine	C5H9N3	amine	3	0
114.0662:0.646:p	Creatinine	C4H7N3O	amino acid	1	1
116.0706:0.586:p	Proline	C5H9NO2	amino acid	1	0
116.0706:0.907:p	4-Amino-2-methylenebutanoic acid	C5H9NO2	amino acid	3	0
118.0862:0.839:p	Betaine	C5H11NO2	amino acid	1	0
118.0862:0.690:p	Valine	C5H11NO2	amino acid	1	0
120.0652:0.887:p	Threonine	C4H9NO3	amino acid	1	1
124.0393:0.86:p	Picolinic acid [a]	C6H5NO2	amino acid	3	0
131.1179:0.745:p	Acetylputriscine	C6H14N2O	amine	1	1
132.0768:0.768:p	Creatine	C4H9N3O2	amino acid	1	1
132.101:0.862:p	Isoleucine	C6H13NO2	amino acid	1	0
132.101:0.852:p	Leucine	C6H13NO2	amino acid	1	0
133.0969:0.586:p	Ornithine	C5H12N2O2	amino acid	1	1
134.0445:0.969:p	Aspartate	C4H7NO4	amino acid	1	0
137.0457:1.063:p	Hypoxanthine	C5H4N4O	purine	1	1
138.055:0.892:p	Trigonelline [a]	C7H7NO2	organic acid	1	1
141.0659:0.698:p	Imidazole propionate	C6H8N2O2	organic acid	2	1
145.1329:0.733:p	Acetylcadaverine [a]	C7H16N2O	amine	2	0
146.1175:0.671:p	Acetylcholine/4-Trimethylammoniobutanoic acid	C7H16NO2	organonitrogen	3	0
147.0764:0.905:p	Glutamine	C5H10N2O3	amino acid	1	1
147.1128:0.566:p	Lysine	C6H14N2O2	amino acid	1	1
148.0598:0.918:p	Glutamate	C5H9NO4	amino acid	1	0
150.0583:0.897:p	Methionine	C5H11NO2S	amino acid	1	1
152.0567:0.734:p	Guanine	C5H5N5O	purine	1	0
152.0567:1.152:p	Unknown	$C_7H_7N_2O_2$	NA	4	0
156.0767:0.606:p	Histidine	C6H9N3O2	amino acid	1	1
160.1331:0.699:p	Aminooctanoic acid	C8H17NO2	NA	4	0
162.1125:0.742:p	Carnitine	C7H15NO3	organonitrogen	1	1
166.0723:0.716:p	Methylguanine	C6H7N5O	purine	1	1
166.0863:0.921:p	Phenylalanine	C9H11NO2	amino acid	1	1
170.0924:0.643:p	3-Methylhistidine	C7H11N3O2	amino acid	1	1
175.119:0.586:p	Arginine	C6H14N4O2	amino acid	1	0
176.0658:0.862:p	Acetyl-aspartic acid	C6H9NO5	amino acid	3	1
176.1029:0.938:p	Citrulline	C6H13N3O3	amino acid	1	0
182.0809:0.954:p	Tyrosine	C9H11NO3	amino acid	1	1
189.1234:0.837:p	Alpha-Acetyllysine/N_6-Acetyllysine	C8H16N2O3	amino acid	3	0
189.1598:0.612:p	Trimethyllysine	C9H20N2O2	amino acid	1	1
191.0668:0.998:p	Aspartyl-glycine [a]	C6H10N2O5	amino acid	3	1
198.0851:0.846:p	Unknown	C8H11N3O3	NA	4	0

204.123:0.792:p	Acetyl-carnitine	C9H17NO4	organonitrogen	1	1
205.0972:0.927:p	Tryptophan	C11H12N2O2	amino acid	1	1
231.1704:0.888:p	*Unknown*	C14H20N3	NA	4	0
241.0311:0.932:p	Cystine	C6H12N2O4S2	amino acid	1	1
243.1095:0.606:p	*Unknown*	C10H14N2O5	NA	4	0
247.0373:0.588:p	*Unknown*	C9H6N6OS	NA	4	0
268.104:0.857:p	Deoxyguanosine	C10H13N5O4	purine	3	0
276.1559:0.637:p	*Unknown*	C14H19N4O2	NA	4	0
284.0989:1.151:p	Guanosine	C10H13N5O5	purine	2	0
287.2447:0.643:p	N_1,N_{12}-Diacetylspearmine	C14H30N4O2	polyamine	2	0
295.1288:0.936:p	Phenylalanyl-Glutamine	C14H19N3O4	amino acid	3	0
298.097:0.873:p	*Unknown*	C7H15N5O8	NA	4	0
298.097:0.617:p	Methylthioadenosine	C11H15N5O3S	purine	2	1
308.5167:0.735:p	*Unknown* [a]	C21H41N	NA	4	0
235.1178:0.994:n	*Unknown* [b]	C6H16N6O4	NA	4	0
632.2044:0.713:n	Sialyllactose [b]	C23H39NO19	amino sugar	1	1
673.2309:0.716:n	Sialyl-*N*-acetyllactosamine [b]	C25H42N2O19	amino sugar	1	1
230.0127:0.932:n	Paracetamol sulfate [b]	C8H9NO5S	sulfar conjugate	1	1
353.1597:0.764:n	Propofol glucuronide [b]	C18H26O7	drug metabolite	2	1
369.1545:0.757:n	Hydroxypropofol glucuronide [b]	C18H26O8	drug metabolite	2	1
108.0456:1.041:n	*Unknown* [b]	C6H7NO	NA	4	0
175.0612:0.881:n	2-Isopropylmalic acid [b]	C7H12O5	organic acid	3	0
273.0802:0.789:n	*Unknown* [b]	C8H14N6O3S	NA	4	0
375.2898:0.763:n	Lithocholic acid [b]	C24H40O3	steroid	2	0
593.3345:0.722:n	Stercobilin [b]	C33H46N4O6	steroid	3	0
598.3032:0.716:n	Urobilinogen [b]	C33H42N4O6	steroid	3	0
154.0499:0.865:n	Mesalamine [b]	C7H7NO3	aminobenzoic acid	2	1
160.0975:0.714:p	Acetylvaline [b]	C7H13NO3	amino acid	3	0
228.0979:0.844:p	Deoxycytidine [b]	C9H13N3O4	purine	3	0
291.1305:0.812:p	Arginosuccinate [b]	C10H18N4O6	amino acid	3	1
407.2385:1.015:p	*Unknown* [b]	C21H32N3O5	NA	4	0
471.2185:0.839:p	*Unknown* [b]	C20H26N10O4	NA	4	0

[a] *Compound that were removed with 75% missing rate cut-off but are added with 50 % cut-off value.*
[b] *Detection rate was lower than 50% for these metabolites.*

Table S4.7. Summary of 131 urinary metabolites detected in urine samples of pediatric IBD patients with *m/z*, RMT, ionization mode (p = ESI positive ion mode, n = ESI negative ion mode), molecular formula, chemical ID or tentative ID, confidence level of the ID and chemical classification based on HMDB. Column "common" indicates the common metabolites found in both urine and stool metabolite (1 = commonly found, 0 = not found in urine).

*m.z:*RMT:mode	Chemical ID	formula	class	Confirmed level	common
87.0452:1.607:n	Butyric acid	C4H8O2	organic acid	1	1
89.0244:1.131:n	Lactic acid	C3H6O3	organic acid	1	1
117.0193:1.826:n	Succinic acid	C4H6O4	organic acid	1	1
128.0353:1.102:n	Oxo-proline	C5H7NO3	amino acid	1	0
131.035:1.609:n	Glutaric acid	C5H8O4	organic acid	2	0
132.0302:1.025:n	Aspartic acid	C4H7NO4	amino acid	1	0
135.0299:0.995:n	Threonic acid	C4H8O5	organic acid	1	0
145.0506:1.473:n	2-Methylglutaric acid	C6H10O4	organic acid	2	1
156.0657:0.917:n	Tiglylglycine	C7H11NO3	organonitrogen	3	0
157.051:1.429:n	Isopropylmaleate	C7H10O4	organic acid	2	0
159.1027:0.861:n	7-Hydroxyoctanoic acid	C8H16O3	organic acid	3	0
160.0615:0.911:n	Aminoadipic acid	C6H11NO4	amino acid	1	0
166.0146:1.014:n	Quinolinic acid	C7H5NO4	organic acid	3	0
167.0201:0.968:n	Uric acid	C5H4N4O3	organic acid	1	1
172.9912:1.135:n	Phenyl sulfate	C6H6O4S	phenylsulfate	2	0
178.051:0.914:n	Hippuric acid	C9H9NO3	organic acid	1	1
181.0506:0.903:n	3-(3-Hydroxyphenyl)-3-hydroxypropanoic acid (HPHPA)	C9H10O4	organic acid	3	0
182.0459:0.948:n	4-Pyridoxic acid	C8H9NO4	organic acid	1	1
184.0977:0.876:n	2-Hepteneoylglycine	C9H15NO3	amino acid	3	0
187.0071:1.059:n	*p*-Cresol sulfate	C7H8O4S	phenylsulfate	1	1
188.0353:0.926:n	Kynurenic acid	C10H7NO3	Quinoline	1	0
188.0558:1.35:n	acetylglutamate	C7H11NO5	amino acid	3	1
191.0552:0.895:n	Quinic acid	C7H12O6	organic acid	1	0
193.0357:0.882:n	Glucuronic acid	C6H10O7	organic acid	1	0
195.0524:0.888:n	dimethyluric acid/gluconate	C7H8N4O3	purine	3	0
197.0455:0.9:n	*Unknown*	C5H6N6O3	NA	4	0
201.1129:1.218:n	Sebacic acid	C10H18O4	organic acid	3	0
212.0023:1.025:n	Indoxyl sulfate	C8H7NO4S	phenylsulfate	1	0
218.1034:0.836:n	Pantothenic acid	C9H17NO5	organonitrogen	1	1
222.9916:0.973:n	*Unknown*	C9H4O7	NA	4	0
225.0629:0.86:n	5-acetylamino-6-formylamino-3-methyluracil	C8H10N4O4	NA	3	0
227.9968:0.979:n	5-Hydroxy-6-indolyl-O-sulfate	C8H7NO5S	phenylsulfate	2	0
241.1193:0.824:n	*Unknown*	C7H14N8O2	NA	4	0

243.0771:0.845:n	Indolylacryloylglycine	C13H12N2O3	amino acid	3	0
263.1037:0.826:n	Phenylacetyl glutamine	C13H16N2O4	amino acid	1	0
269.15:0.805:n	*Unknown*	C8H22N4O6	NA	4	0
283.0823:0.809:n	*p*-Cresol-glucuronide	C13H16O7	phenolic glycoside	3	0
287.0227:0.912:n	5'-(3',4'-Dihydroxyphenyl)-gamma-valerolactone sulfate	C11H12O7S	phenylsulfate	3	0
290.0882:0.796:n	2,3-Dehydro-2-deoxy-N-acetylneuraminic acid	C11H17NO8	amino sugar	1	0
302.114:0.812:n	Indoleacetyl glutamine	C15H17N3O4	amino acid	2	0
308.0987:0.791:n	N-Acetylneuraminic acid	C11H19NO9	amino sugar	1	1
331.1757:0.777:n	Neomenthol-glucuronide	C16H28O7	O-glucuronide	4	0
336.0725:0.794:n	Indole-3-carboxylic acid glucuronide	C15H15NO8	O-glucuronide	3	0
338.0881:0.791:n	6-Hydroxy-5-methoxyindole glucuronide	C15H17NO8	O-glucuronide	3	0
345.1553:0.770:n	*Unknown*	C16H26O8	NA	3	0
347.0853:0.826:n	*Unknown*	C23H12N2O2	NA	4	0
350.088:0.788:n	Indole-3-acetic-acid-O-glucuronide	C16H17NO8	O-glucuronide	3	0
352.0868:0.863:n	4-Hydroxybenzyl isothiocyanate 4"-acetylrhamnoside	C16H19NO6S	phenolic glycoside	3	0
407.2803:0.750:n	Cholic acid	C24H40O5	steroid	1	1
464.3018:0.737:n	Glycocholic acid	C26H43NO6	steroid	1	1
481.2439:0.742:n	11-beta-Hydroxyandrosterone-3-glucuronide	C25H38O9	steroid	3	0
525.2688:0.733:n	*Unknown*	C21H36N9O7	NA	4	0
539.2493:0.733:n	Tetrahydrocortisone-glucuronide	C27H40O11	steroid	3	0
541.2649:0.729:n	Cortolone-glucuronide	C27H42O11	steroid	2	0
543.2811:0.725:n	Cortol-3-glucuronide	C27H44O11	steroid	2	0
632.2044:0.717:n	Sialyllactose	C23H39NO19	amino sugar	1	1
673.2309:0.713:n	Sialyl-*N*-acetyllactosamine	C25H42N2O19	amino sugar	1	1
62.06:0.577:p	Ethanolamine	C2H7NO	amine	1	0
76.0393:0.743:p	Glycine	C2H5NO2	amino acid	1	1
76.0757:0.594:p	Trimethylamine-*N*-oxide	C3H9NO	organonitrogen	1	0
90.055:0.795:p	Alanine	C3H7NO2	amino acid	1	1
104.0706:0.94:p	Dimethylglycine	C4H9NO2	amino acid	1	0
104.0706:0.702:p	Gamma-aminobutyric acid (GABA)	C4H9NO2	amino acid	1	1
104.1069:0.618:p	Choline	C5H14NO	organonitrogen	1	1
106.0499:0.868:p	Serine	C3H7NO3	amino acid	1	1
114.0662:0.656:p	Creatinine	C4H7N3O	amino acid	1	1
118.0611:0.737:p	*Unknown*	C3H7N3O2	NA	4	0
120.0652:0.905:p	Threonine	C4H9NO3	amino acid	1	1
129.0659:0.774:p	Dihydrothymine	C5H8N2O2	purine	3	0

131.1179:0.752:p	Acetylputriscine	C6H14N2O	amine	1	1
132.0768:0.782:p	Creatine	C4H9N3O2	amino acid	1	1
133.0969:0.628:p	Ornithine	C5H12N2O2	amino acid	1	1
137.0457:1.039:p	Hypoxanthine	C5H4N4O	purine	1	1
138.055:0.909:p	Trigonelline	C7H7NO2	organic acid	1	1
141.0659:0.734:p	imidazole propionate	C6H8N2O2	organic acid	2	1
146.0924:0.742:p	4-Guanidinobutanoate	C5H11N3O2	organic acid	3	0
147.0764:0.926:p	Glutamine	C5H10N2O3	amino acid	1	1
147.1128:0.631:p	Lysine	C6H14N2O2	amino acid	1	1
150.0583:0.915:p	Methionine	C5H11NO2S	amino acid	1	1
156.0767:0.667:p	Histidine	C6H9N3O2	amino acid	1	1
162.1125:0.757:p	Carnitine	C7H15NO3	organonitrogen	1	1
163.1077:0.654:p	5-hydroxylysine	C6H14N2O3	amino acid	1	0
164.0748:0.769:p	Propyl-*S*-cysteine	C6H13NO2S	amino acid	3	0
166.0723:0.744:p	Methylguanine	C6H7N5O	purine	1	1
166.0863:0.94:p	Phenylalanine	C9H11NO2	amino acid	1	1
170.0924:0.681:p	3-Methylhistidine	C7H11N3O2	amino acid	1	1
176.0658:0.876:p	Acetyl-aspartic acid	C6H9NO5	amino acid	3	1
182.0809:0.964:p	Tyrosine	C9H11NO3	amino acid	1	1
189.1598:0.651:p	Trimethyllysine	C9H20N2O2	amino acid	1	1
190.1191:0.954:p	Homocitrulline	C7H15N3O3	amino acid	1	0
191.0661:1.007:p	Aspartyl-glycine	C6H10N2O5	amino acid	3	1
195.0764:0.895:p	Aminohippuric acid	C9H10N2O3	organic acid	3	0
204.123:0.796:p	Acetyl-carnitine	C9H17NO4	organonitrogen	1	1
205.0972:0.938:p	Tryptophan	C11H12N2O2	amino acid	1	1
209.0921:0.887:p	Kynurenine	C10H12N2O3	amino acid	1	0
217.1294:0.869:p	Acetyl-arginine	C8H16N4O3	amino acid	3	0
222.0796:0.849:p	5-(delta-carboxybutyl) Homocysteine	C8H15NO4S	amino acid	2	0
223.0747:0.866:p	Cystathionine	C7H14N2O4S	amino acid	1	0
232.1543:0.829:p	Butyryl carnitine	C11H21NO4	organonitrogen	2	0
238.0916:1.064:p	Xylosylserine	C8H15NO7	amino sugar	2	0
241.0311:0.946:p	Cystine	C6H12N2O4S2	amino acid	1	1
243.0981:0.91:p	Thymidine	C10H14N2O5	pyrimidine	2	0
244.1543:0.849:p	Tiglylcarnitine	C12H21NO4	organonitrogen	3	0
258.1084:0.862:p	Methylcytidine	C10H15N3O5	pyrimidine	3	0
259.0918:0.893:p	Ribothymidine	C10H14N2O6	pyrimidine	2	0
269.1238:0.926:p	Acetylcarnosine	C11H16N4O4	amino acid	1	0
276.1442:0.883:p	Glutaryl-carnitine	C12H21NO6	organonitrogen	3	0
282.1197:0.872:p	Methyl adenosine	C11H15N5O4	purine	1	0

286.2013:0.886:p	Fumaric acid, 2-dimethylaminoethyl heptyl ester	C18H25N2O	NA	3	0
290.1598:0.897:p	Methylglutarylcarnitine	C13H23NO6	organonitrogen	3	0
291.1305:0.827:p	Arginosuccinate	C10H18N4O6	amino acid	3	1
298.097:0.656:p	Methylthioadenosine	C11H15N5O3S	purine	2	1
298.1146:1.058:p	1- or 2- or 3'-O-Methylguanosine	C11H15N5O5	purine	2	0
304.1755:0.908:p	Pimelyl carnitine	C14H25NO6	organonitrogen	3	0
304.2109:0.92:p	Hydroxyoctanoyl carnitne	C15H29NO5	organonitrogen	3	0
312.1297:1.039:p	Dimethyl guanosine	C12H17N5O5	purine	1	0
325.165:0.777:p	Galactosyl-hydroxylysine	C12H24N2O8	amino acid	1	0
367.15:1.065:p	Mannopyranosyl-Trptophan	C17H22N2O7	amino sugar	1	0
399.1451:0.638:p	S-Adenosylmethionine	C15H23N6O5S	amino acid	3	1
487.2117:0.854:p	Glucosylgalactosyl-hydroxylysine	C18H34N2O13	amino acid	1	0
194.0458:0.921:n	Salicyluric acid [a]	C9H9NO4	organic acid	1	0
204.0666:0.876:n	Indole lactic acid [a]	C11H11NO3	organic acid	1	0
230.0127:0.933:n	Paracetamol sulfate [a]	C8H9NO5S	sulfar conjugate	1	1
263.629:0.953:n [M-2H]$^{2-}$	Unknown bile acid glycine-sulfate conjugate adduct [a]	C26H43NO8S	steroid	3	0
319.14:0.782:n	Octanoylglucuronide	C14H24O8	O-glucuronide	3	0
353.1597:0.766:n	Propofol glucuronide [a]	C18H26O7	drug metabolite	2	1
359.1857:0.587:n	Prednisolone [a]	C21H28O5	drug metabolite	2	0
369.1545:0.760:n	Hydroxypropofol glucuronide [a]	C18H26O8	drug metabolite	2	1
154.0499:0.887:p	Mesalamine [a]	C7H7NO3	aminobenzoic acid	2	1
262.1028:1.079:p	Aspartyl-glutamine [a]	C9H15N3O6	amino acid	2	0
288.217:0.941:p	Octanoyl-carnitine [a]	C15H29NO4	organonitrogen	1	0

[a] *Compounds that were missing in more than 50% of total samples in this study.*

(A) Extracted ion electropherogram of anion (*m/z:RMT* 353.1597:0.766)

(B) High resolution MS spectra

Propofol-glucuronide

[M+H]⁻ 353.1599 $C_{18}H_{26}O_7$ Mass error: 9.06 ppm

C13 isotope

(C) MS/MS spectra (CID at 20 V)

Deprotonated propofol

Bond cleavage

Loss of glucuronic acid

[M-H]⁻ 353.1598
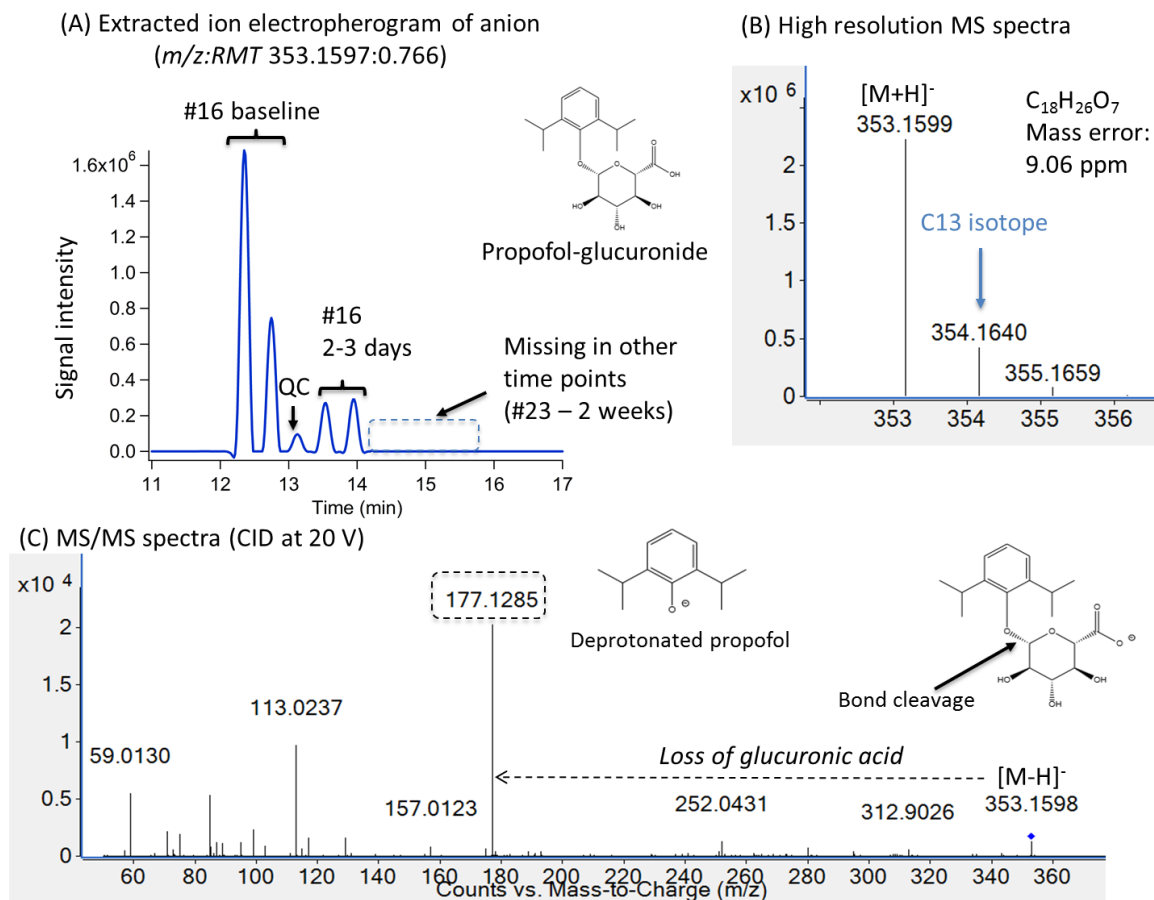
Figure S4.1. Structural elucidation of urinary propofol glucuronide (*m/z:*RMT, 353.1597:0.766), an exogenous drug metabolite excluded from the final metabolomics data matrix. **(a)** Extracted ion electropherogram of a representative run when using a seven serial sample injection format in MSI-CE-MS. Large peaks only appeared in baseline samples and residual metabolites were still detected 2-3 days after but not at latter time points. **(b)** High resolution MS spectrum for the deprotonated molecular ion [M-H⁻] and isotopic pattern with most likely molecular formula and calculated mass error (< 10 ppm). **(c)** MS/MS spectrum with collisional induced dissociation at 20 V showing a characteristic neutral loss of glucuronic acid (*m/z* 176.0313) and the base peak for the product ion corresponding the original drug molecule that was used as an anesthetic agent for the colonscopy procedure of pediatric IBD patients, namely propofol (*m/z* 177.1285).
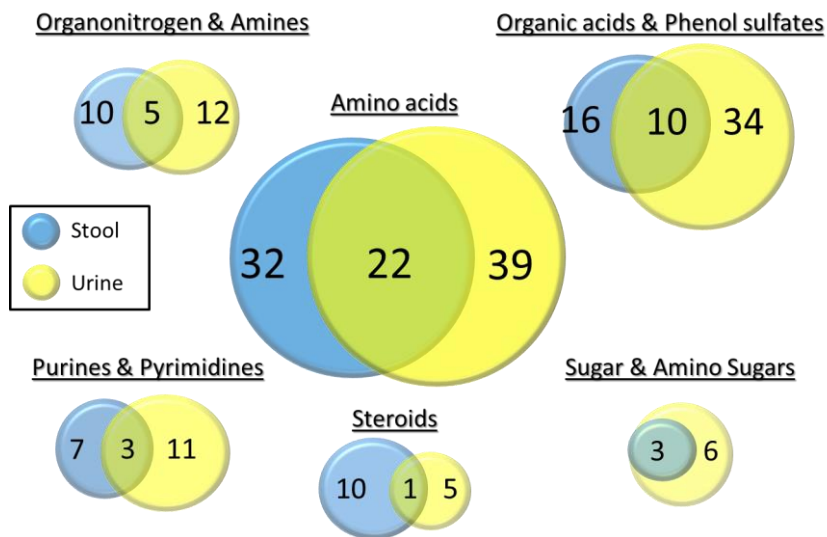
Figure S4.2. Number of metabolites detected in urine and stool samples of IBD patients categorized based on chemical class. Metabolites detected in both bio-specimens are shown in the overlapped region. Unidentified metabolic features are excluded from this number.
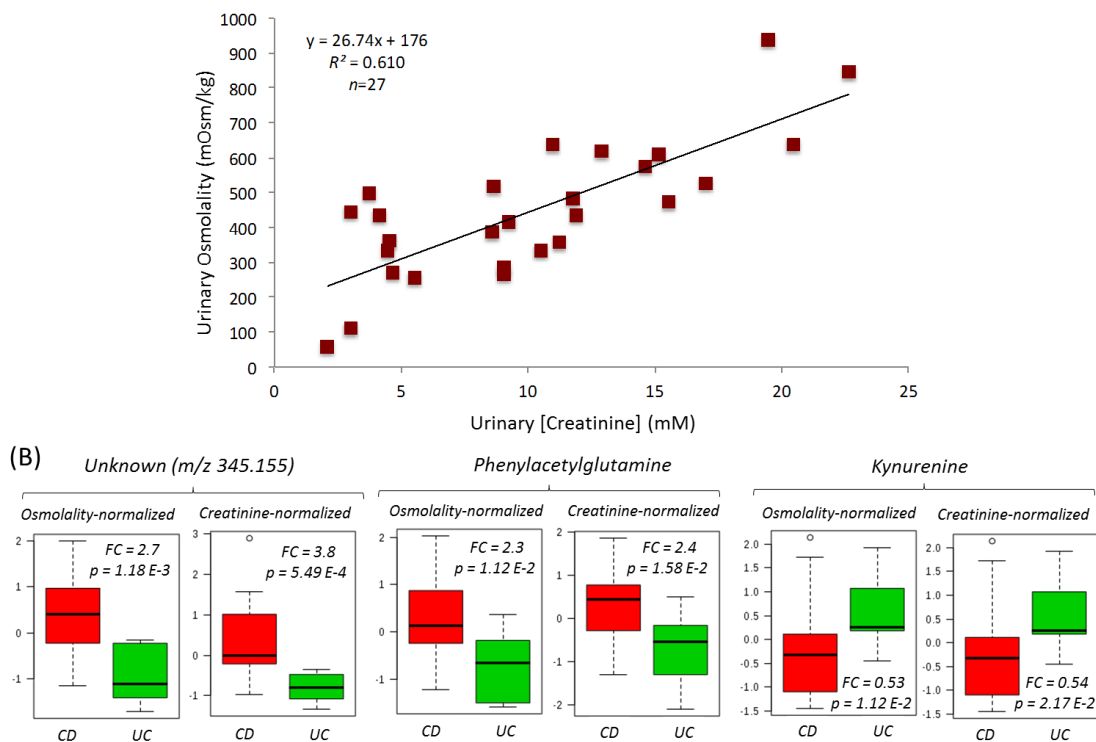
Figure S4.3. (a) Scatter plot of urine osmolality (mOsm/kg) and creatinine (mM) measured from 27 urine samples of IBD patients. Linear regression equation and R^2 values corresponding to the linearity are shown. **(b)** Boxplots of differentially excreted urinary metabolites between UC and CD patients normalized by osmolality or creatinine. Mann-Whitney U-test, $p < 0.05$. FC: median fold change of CD/UC.
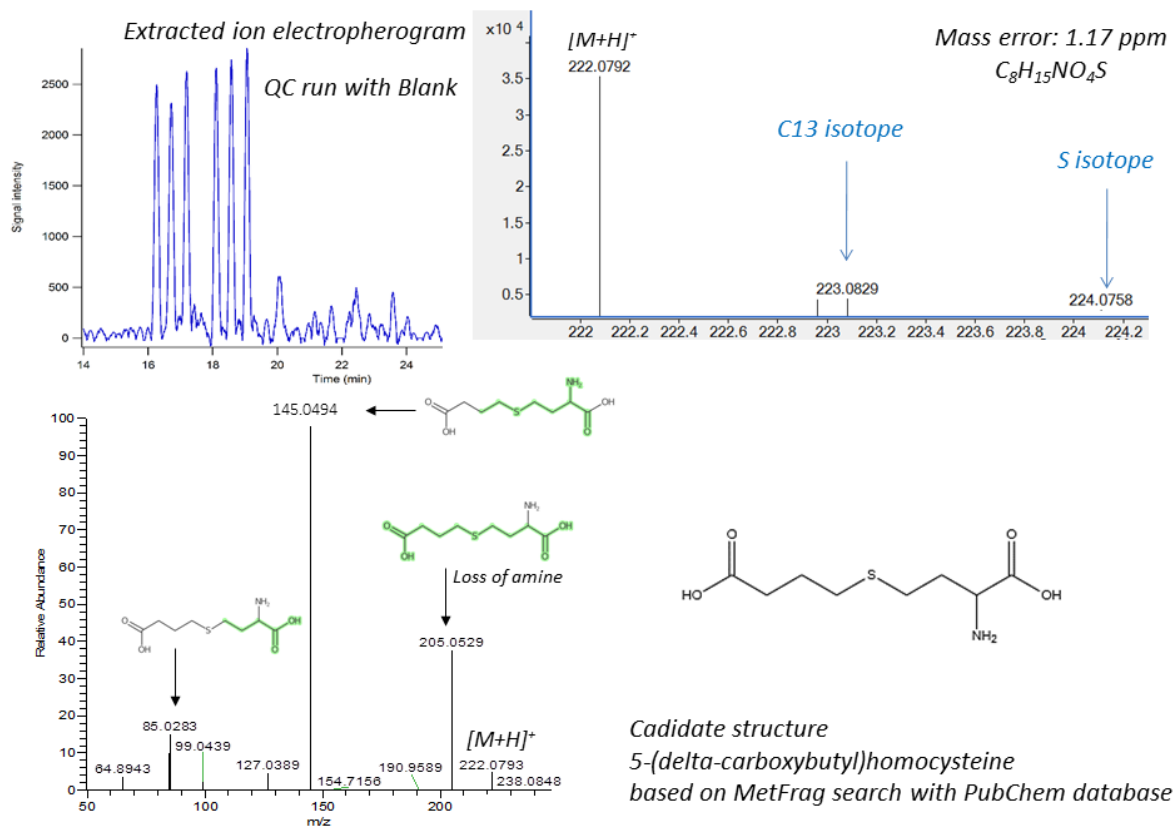
Figure S4.4. Putative identification of urinary 5-(delta-carboxybutyl)homocysteine based on high resolution MS and MS/MS as shown for **(a)** an extracted ion electropherogram for the unknown anion (*m/z:*RMT, 222.0796:0.849) that is measured consistently in a pooled QC sample run by MSI-CE-MS without sample carry-over effects in blank, **(b)** a high resolution TOF-MS spectrum showing its protonated molecular ion [M+H$^+$], charge state and isotope pattern required for determining its mostly likely molecular formula and **(c)** annotation of its MS/MS spectrum following collisional-induced dissociation at 20V of the precursor ion to confirm its likely chemical structure based on diagnostic fragment ions and neutral losses.

Chapter V

Elucidating Treatment Responses of Pediatric Crohn's Disease Patients Following Exclusive Enteral Nutrition Therapy: A Longitudinal Metabolomics Study

Mai Yamamoto, Lara Hart, Nikhil Pai, and Philip Britz-McKibbin

M.Y. prepared all samples and conducted the urinary and stool metabolite analysis using CE-MS and analyzed the metabolomics data and microbiome data, as well as wrote an initial manuscript draft for publication. Other authors designed the study, obtained ethics approval, recruited patients, collected samples, and provided feedback on the manuscript draft.

Chapter V: Elucidating Treatment Responses of Pediatric Crohn's Disease Patients Following Exclusive Enteral Nutrition Therapy: A Longitudinal Metabolomics Study

5.1 Abstract

Exclusive enteral nutrition (EEN) therapy has become first-line therapy for treatment of Crohn's disease (CD) in pediatric inflammatory bowel disease due to its efficacy in inducing mucosal healing and clinical remission. Importantly, EEN promotes linear growth in pediatric CD patients while avoiding the adverse effects of corticosteroid therapy for children. However, the underlying molecular mechanisms of EEN therapy remains poorly understood. Herein, dynamic metabolomic studies were performed to better decipher biochemical changes following EEN therapy when using multisegment injection-capillary electrophoresis-mass spectrometry (MSI-CE-MS) in conjunction with stool metagenomics analysis. A cohort of largely treatment naïve pediatric CD patients (mean age=13 y, n=16) were recruited to complete a 8-week EEN intervention with repeat urine and stool specimens collected as a function treatment time course. A robust metabolomics data workflow that takes advantage of multiplexed separations and stringent quality control (QC) was developed to characterize the metabolic phenotype of urine and stool extracts in order to identify metabolic trajectories corresponding to biomarkers of treatment responses to EEN therapy from baseline. Unknown metabolites with potential clinical application were also identified when using high resolution, accurate MS and MS/MS experiments. Complementary univariate, multivariate and correlation statistical methods were used to select top-ranked metabolites associated with inflammatory markers (*i.e.,* C-reactive protein and fecal calprotectin) and/or displayed time-dependent treatment responses coinciding with the onset of remission from osmolality normalized urine and stool extracts normalized to dried weight. Several nutrients were found to undergo significant increases in excretion following initiation of EEN therapy, including pantothenic acid, pyridoxic acid, and trigonelline. Our work also revealed that EEN treatment significantly lowered excretion of several biomarker candidates associated with immune activation and oxidative stress, as well as alteration in gut microbial activity implicated in reduced inflammation. For instance, urinary sialic acids and several glycosylated amino acids derived from glycoprotein turn-over were attenuated with EEN therapy, whereas *p*-cresol sulfate and its glucuronide conjugate, which are produced by specific gut microbiota, increased during the therapy. Additionally, fecal microbiome analysis showed significant decreases in opportunistic pathogens, *Haemophilus* and *Veillonella* following EEN therapy. Overall, our findings indicate improved nutrition absorption and gut microbial modulation as key initiators of symptom improvement following initiation of EEN with subsequent recovery in immune function and cellular protective mechanisms. For the first time, we report global metabolic trajectories as a result of EEN therapy, which may provide an effective and convenient tool for predicting treatment responses at the early stage of therapy induction, as well as therapeutic monitoring for risk assessment of disease recurrence.

5.2 Introduction

The incidence of Crohn's disease (CD) in children is rapidly increasing around the world,[1] especially in Western countries, including Canada.[2] Early onset of CD poses increased risk of impaired linear growth, delayed puberty and poor bone health, in addition to the recurrent clinical symptoms, such as abdominal pain, diarrhea and weight loss.[3] The pathogenesis of CD is considered to involve a complex interplay of genetic predisposition, environmental factors, innate immunity and intestinal bacterial communities (*i.e.,* microbiome) which overlaps with other inflammatory bowel diseases (IBD), such as ulcerative colitis (UC).[4] However, based on the alarming increase in incidence rates among migrant populations and the rapid course of urbanization in developing nations,[1] environmental changes are considered to play a key role in the pathogenetisis of CD. Among these changes, population-wide dietary changes have been implicated. Short-term dietary modification is known to play in the treatment of CD.[5] Exclusive enteral nutrition (EEN) therapy is first line therapy for CD, and while its adoption has been greater in pediatric medicine, its efficacy is considered equivalent in adults[6,7] and children[8] globally. EEN is an important therapeutic consideration due to its ability to induce clinical remission as well in traditional first-line treatments, immunosuppression with corticosteroids (CS), but without the multiple deleterious effects of therapy. During EEN, patients take liquid formula containing complete nutrients orally or through nasogastric tubing to ensure better adherence for 6-8 weeks to supply all (or most) of their caloric needs. A number of clinical studies have shown that EEN induces remission based on a reduction in symptom disease severity and inflammatory markers, including C-reactive protein (CRP) and fecal calprotectin (FCP), at equivalent or better rates than CS therapy.[9-11] These benefits are paired with the reduction in treatment-related risks associated with CS therapy: namely, long-term growth retardation and loss of bone mineral density especially for children or adolescents. Additionally, some studies also showed a significantly greater rate of mucosal healing in patients treated with EEN, which was correlated well with a prolonged duration of remission.[12-15] Although treatment efficacy has been consistently reported in several

pediatric CD treatment studies, the exact mechanisms of action of EEN remain unclear, including its effectiveness in the treatment of UC. Some proposed mechanisms include complete avoidance of potential yet unknown allergens in the diet, anti-inflammatory components in the formula and/or modulation of commensal gut microbiota.[16] The efficacy based on exclusion of dietary antigen is compelling based on the correlative associations between increased meat consumption and incidence of CD,[17,18] and high remission inducing rates (86%) of exclusion diets that lacked dairy, animal fats, emulsifiers and most processed food.[19] However, no single dietary antigen has been identified to have an exacerbating effect on CD symptoms, and the specific metabolic changes that connect dietary changes with symptom improvement have yet to be identified. Similarly, since EEN formula typically consists of polysaccharide, polymeric or elemental forms of protein, vitamins and other essential micronutrients, there has not been any single anti-inflammatory macronutrient identified in typical EEN formula.[16,20] EEN induced modulation of gut microbiome is considered one mechanism of action. It is recognized that short-term dietary changes can induce profound effects on the intestinal microbiome.[21] Thus, exclusive nutritional therapy may result in beneficial changes in the intestinal microbiome. This may have therapeutic value, considering that differences have been reported in gut microbiota between healthy individuals and CD patients.[22,23]

In this context, metabolomics holds great potential in providing new insights into the mechanisms of EEN therapy for inducing remission in CD patients. Metabolomics involves the comprehensive analysis of metabolites in complex biological samples, which are not only by-products of energy metabolism or building blocks of biopolymer synthesis, but also include bioactive xenobiotics from diet and environmental exposures that modulate host immune function, such as vaccines.[24] As a result, metabolomics offers a hypothesis-generating approach for characterizing the metabolic phenotype of an organism that is closely associated with clinical outcomes when using nuclear magnetic resonance (NMR) and high efficiency separation platforms coupled to high resolution mass spectrometry (MS). Particularly, nontargeted metabolite profiling aims to resolve, identify and quantify the widest range of metabolites without *a priori* knowledge for biomarker discovery, which

can be integrated with other -omics approaches to decipher the underlying mechanisms of chronic human diseases.[25] Longitudinal metabolomics studies, which identify dynamic metabolic trajectories associated with treatment responses or exposures over time, has been increasingly applied for understanding the underlying mechanism(s) of pharmacological agents[26] or environmental contaminants[27] in recent years. However, only a few reports have been reported in the context of therapeutic monitoring for IBD patients, which used targeted metabolomic methods with longitudinal sampling. For instance, Nikolaus *et al.*[28] focused on serum tryptophan levels of adult CD ($n = 31$) and UC ($n = 15$) patients at baseline and following therapy with biologic (*i.e.* infliximab) over 24 weeks when using liquid chromatography (LC) with UV absorbance. In this work, a significant increase in serum tryptophan levels was measured for responders to therapy as evaluated by disease score, whereas therapy non-responders showed decreases in circulating tryptophan levels.

In the context of EEN therapy for IBD, existing studies have taken a combined approach of gut microbiome metagenomics and metabolomics approach, which specifically targeted microbial metabolites. Gerasimidis *et al.*[29] reported significant decrease in short chain fatty acid (SCFA), butyrate, in fecal samples of 15 children with CD during 8-week EEN therapy, when using gas chromatography. Butyric acid plays a critical role in immune homeostasis in the intestine by promoting differentiation of regulatory T-cells, and its decrease has been linked to CD pathology.[30] Concurrently, they also identified decrease in presumably beneficial bacteria, *Faecalibacterium prausnitzii*, which is known as predominant butyrate producer in the gut,[31] and its reduced level has consistently reported in CD as compared to healthy controls.[32,33] This paradoxical decrease in "healthy" bacteria and butyric acid during EEN has also been reported in patients with juvenile idiopathic arthritis,[34] another autoimmune disorder prevalent among children that rely on similar therapeutic interventions as IBD.[35] As the authors suggested, the lack of butyric acid sources, namely fermentable carbohydrates, in EEN formula was likely the reason for these observations, which still did not explain the molecular mechanisms related to EEN treatment efficacy. Despite the clear link between changes in habitual diet and CD

prevalence, comprehensive examination of metabolic phenotype changes that take place during a therapeutic dietary intervention, including EEN, in CD patients have not been reported to date. For exploratory metabolomic studies involving pilot clinical investigations of modest cohort sizes, it is ideal to have repeat and complementary biospecimens to monitor dynamic changes in metabolic phenotypes that are linked to symptom improvement of patients. In addition to stool samples, human urine has been shown to contain polar/ionic metabolic end-products reflecting recent dietary intake and gut microbiota activity, such as trigonelline from coffee and plant intake [36,37] as well as products of human host-microbial co-metabolism, including *p*-cresol sulfate.[38] Indeed, the complex interactions involving gut microbiota, host immunity and diet warrants an untargeted metabolomics investigation using multiple biospecimens beyond the targeted analysis of fecal short-chain fatty acids (SCFA).

Herein, we report for the first time a longitudinal metabolomics study involving matching urine and stool specimens collected from a cohort of pediatric CD patients following EEN therapy. Using an optimized fecal extraction protocol and a rigorous data workflow for biomarker discovery based on multisegment injection-capillary electrophoresis-mass spectrometry (MSI-CE-MS) with quality control (QC), metabolic trajectories associated with EEN treatment intervention were characterized in osmolality normalized urine and stool normalized to dried weight. In this case, our work revealed distinctive metabolite signatures as a function of EEN time course reflecting improved nutrient absorption, modulation in immune functions and changes to intestinal microbiome. Additionally, a comparison of top-ranked treatment responsive metabolites between CD responders to EEN (*n*=11) and CS (*n*=3), as well as an EEN non-responder (*n*=1) confirmed that the same metabolic trajectories coincide with clinical remission regardless of therapeutic intervention. Our results were further validated by performing metabolite stability studies as related to specimen collection conditions and delays to storage, which demonstrated excellent stability for the top-ranked urinary biomarker candidates implicating their practical use to monitor treatment responses of CD patients in a routine

clinical settings, including identifying non-responders who did not undergo remission following EEN.

5.3 Experimental section

5.3.1 Subjects and induction therapy. This study was approved by the Hamilton Integrated Research Ethics Board (HIREB# 15-365) and parental consent was obtained for all the participants. The study enrolled children 5 − 18 years old who have been diagnosed with IBD by endoscopy, histology and radiography. Patients were included if they were admitted to hospital to be initiated on exclusive enteral nutrition therapy, or intravenous corticosteroids for induction of remission of CD or UC. Patients were excluded if they were younger than 5 years old, received antibiotic therapy, or did not require admission to hospital. None of these patients have undergone resection surgery. Each patient was assigned a number in the order of enrollment. Polymeric EEN formula (PEPTAMEN®, Nestle Health Science) was administered to patients through nasogastric tubing, except for one patient who orally took the formula, for the period of 8 weeks. Patients who received corticosteroid (prednisolone) therapy followed the Division of Pediatric Gastroenterology and Nutrition protocol for CS induction therapy, with 2 weeks of high dose IV/PO glucocorticoid (maximum 40mg/day) followed by a 6 week wean (approximately decreasing 5mg/day per week).

5.3.2 Selection of subjects and samples for statistical analysis. A total of 11 UC patients and 20 CD patients were recruited for this study; however, numbers of patients were unequally distributed in each treatment arm (*i.e.* only 1 UC patient and 16 CD patients for EEN) due to more widely studied efficacy of EEN toward CD[8,14,39] and paucity of literature in its effect for UC patients,[40] which influenced preference of patients' and physicians' treatment choices. Additionally, missing sampling points were prevalent in this study due to challenges associated with sample collection from children and adolescents with unpredictable bowel movements, particularly after leaving hospital, which resulted in more inconsistency of longitudinal stool sampling compared to urine sample collection. Missing

clinical visits further attributed some missing sampling points. Largely missing and slightly different sampling points pose problems in statistical analysis when comparing metabolic profiles of each time point. Therefore, it was necessary to carefully select subjects to be included in this study. This study focused only on CD patients due to smaller sample size in UC group and larger heterogeneity resulted in UC patients under CS therapy, which failed to induce remission for 4 patients. For longitudinal analysis of metabolome and microbiome during EEN therapy, the following inclusion criteria were applied to select subjects, namely 1) subjects who provided a urine or stool specimen at 2 weeks or longer after the initiation of the therapy, in addition to baseline, 2) subjects who responded to the therapy based on pediatric Crohn's disease activity score (PCDAI). Based on these criteria, data acquired from repeat urine samples from 11 patients and stool specimens from 8 patients with CD during EEN therapy were included in this study and multivariate (PCA) and univariate analysis (correlation analysis and paired Wilcoxon test) were applied. Although subject #19 provided samples up to 1-week time point, this subject was included because of significant decline in FCP from baseline (3148 µg/g) to 1-week point (201 µg/g), which was indicative of a positive treatment response equivalent to later time points of other patients. For paired Wilcoxon rank-sum test, available samples collected at the most advanced time point was used as samples to be compared with baseline because inflammatory markers declined with time, which resulted in the lowest value and state of remission in later time points.

Data for the three patients who were prescribed corticosteroid treatment and one non-responder to EEN therapy were used to compare levels of metabolites that were identified as significant in the paired Wilcoxon test. This group of three patients in CS arm was characterized with mean age of 13 year-old, 33% male ($n=1$), and 100% treatment naïve at the time of diagnosis. The non-responder to EEN therapy was 15 year-old female who was previously diagnosed and prescribed for EEN therapy. Samples from subject #6 were removed from most of data analysis beyond exploratory unsupervised multivariate data analysis to avoid introducing extra variability potentially caused by colonic CD, as CD with isolated colonic involvement have shown different clinical and histological features,[41,42] as

well as different genetic characteristics,[43] from CD that affects other regions, which was also implicated in this study.

5.3.3 Urine and fecal sample collection, storage and preparation. Baseline urine and stool samples were collected prior to or within a few hours of the start of induction therapy (*i.e.* CS or EEN) at McMaster Children's Hospital. Subsequently, samples were collected at patient's home after 2-3 days, 1 week, 2 weeks, 4 weeks, 6 weeks and 8 weeks from the start of the therapy. Urine samples were collected in the mornings and not necessarily the first morning urine. All samples were transferred with cooler packs during transportation. Median time between sample production and collection at the clinic was 4.5 h for urine (range 1 – 48 h) and 4 h for stool (range 1 – 48 h). As soon as samples were brought to a clinic, 6 mM of sodium azide solution (final concentration of 1 mM) was added to each urine sample as antimicrobial preservative and stored in a fridge before being transferred to a freezer at -80 °C. Stool samples were aliquoted into two tubes: one for metabolomics analysis and the other for microbiome analysis, and placed in a freezer at - 80°C at the earliest time possible after sample collection. Chemical reagent information and methods used for stool sample extraction, urine sample preparation, osmolality measurement, and instrumental configuration of MSI-CE-MS are described in *Chapter IV*. Briefly, all stool samples were lyophilized and stool metabolites were extracted from approximately 15 mg of dried stool using a modified Bligh and Dyer method with chilled chloroform, water and methanol in 4:3.6:4 ratio,[44,45] which included 4-fluoro-*L*-phenylalanine (F-Phe) as recovery standard, followed by repeated vortex and centrifugation. Methanol and water layer was stored at -80 °C until analysis. Urine samples were centrifuged at 400 *g* for 5 min to remove particulates, and then the supernatant was withdrawn and diluted with deionized water with IS to 5- and 10-fold dilution prior to analysis of anionic and cationic metabolites.

5.3.4 Urinary and stool metabolome stability study. Stability of selected metabolites in stool and urine were evaluated using samples freshly collected from healthy volunteers as described in the previous chapter for differential IBD characterization at baseline. Briefly, five single-spot mid-stream urine samples were collected from healthy volunteers (25–30

years old), including two males and three females. Each sample was placed on ice and a pooled sample was prepared within one hour following sample collection. Aliquots of this pooled sample were stored at either room temperature (~ 22 °C) or in a fridge (4 °C) for 6, 12, 24, 36 and 48 h in triplicates. After assigned storage duration, sodium azide was added to the final concentration of 1 mM and samples were stored at -80 °C until analysis. Six replicates were prepared as control and immediately transferred to a freezer at -80 °C after preparation of the pooled sample and addition of sodium azide. Additional three aliquots were transferred to the freezer without sodium azide. The same dilution and analysis protocol as IBD samples were applied. Stool samples were collected from a healthy six year-old twin brothers upon receiving a parental consent. Each sample was homogenized with sterile spatula and transferred into 18 tubes for different conditions in duplicates: control (-80 °C freezer), freezer (-20 °C) for 48 h, fridge (4 °C) for 2, 4, 8 and 48 h, and room temperature (~ 22 °C) for 2, 4, and 8 h of storage. Initial sample processing was completed within 30 min. After assigned storage duration, they were transferred to a freezer at -80 °C and stored until extraction and analysis.

5.3.5 Metabolomics data workflow and quality control. To account for the instrumental variations in migration time due to changes in electroosmotic flow (EOF) and integrated peak areas due to differences in on-column injection volumes when using MSI-CE-MS, two internal standards (3-chloro-*L*-tyrosine; Cl-Tyr for positive and 2-naphthalene monosulfonic acid; NMS for negative ion mode) were included for data normalization in every stool extract and diluted urine sample analyzed with a fixed concentration of 10 µM. Additionally, a quality control (QC) sample, comprised of a pooled urine or stool extract sample from the pediatric IBD cohort, was included as a QC sample in every single run within a random position in order to evaluate overall technical variance and correct for potential long-term signal drift in ESI-MS.[46-48] A seven-plug serial injection format was used in MSI-CE-MS for the analysis of stool extracts and urine samples as described previously.[46] Additionally, a QC-based dilution trend filter was applied when using MSI-CE-MS to authenticate reproducible yet representative sample-derived metabolites annotated based on their accurate mass and relative migration time (*m/z*:RMT) as a paired

variable.[49,50] In this case, metabolites from pooled urine and stool extract samples were included in final data matrix provided they had adequate precision (RSD < 40%) with no measurable signal in the blank following rigorous rejection of spurious, redundant and irreproducible signals that constitute the majority of ions detected in ESI-MS.[51] Additionally, metabolites that did not appear in more than 75% of total samples were removed from the final metabolomics data matrix to focus on metabolites that were consistently and reliably detected in the majority of CD patients with missing values substituted by the lowest measured relative ion response measured in study divided by two.

5.3.6 Metabolomics data processing and unknown compound identification. The same data processing scheme used in *Chapter IV* applies in this study. Briefly, after raw data processing using Mass Hunter Workstation Software (Qualitative Analysis, version B.6.00, Agilent Technologies, 2012), migration time and electropherogram peak area data were transferred to Excel spreadsheet (Microsoft Office) and saved as .csv file. Relative migration time (RMT) and relative peak area (RPA) were calculated for each compound in reference to internal standards (Cl-Tyr and NMS) in statistical software R environment (v. 3.5.1). Application of 40% CV cut-off filter based on QC data, as well as 75% cut-off filter based on detection rate in all samples were also applied in R. Finally, normalization of stool and urine metabolome was achieved using dried stool weight (mg) and osmolality (mOsm/kg). Osmolality was specifically selected over creatinine for hydration status normalization from single-spot urine samples collected due to acute changes in diet from baseline following initiation of EEN in this study.[52] Data of over 70% of urinary and stool metabolites were not normally distributed based on Shapiro-Wilk test. Therefore, non-parametric statistical measures were applied by default to untransformed data, except for multivariate analysis (PCA), in which data was generalized *log*-transformed and autoscaled. Unknown metabolite identification was based on formula generated from accurate monoisotopic mass, isotopic ratio, ionization state, and collisional-induced dissociation (CID) experiments in order to generate a product ion/fragmentation spectrum (MS/MS) that was matched with deposited online database spectra (HMDB),[53] *in-silico*

fragmentation, or by manual annotation when standard references were not available for direct comparison.

5.3.7 Statistical analysis and data presentation for metabolomics data. Most of data analysis and plot generation was completed in R environment (v. 3.5.1). Specifically, univariate (paired Wilcoxon rank sum test), principal component analysis by factoextra (v. 1.0.5), p-value adjustment for multiple hypothesis testing by p.adjust function (FDR) were performed after removal of metabolic features that were missing in > 75% of all samples. Missing values were replaced with half of the smallest detected value for the compound as a surrogate for a detection limit prior to statistical analysis. Metabolomics data sets applied a generalized *log* transformation and autoscaling prior to multivariate statistical analysis. Effect size calculation was performed using the Statistical Package for the Social Science (SPSS, version 21).

5.3.8 Microbiome data collection and analysis. Methods used for microbiome raw data collection and raw data processing are described in the method section of *Chapter IV*. Statistical analysis for alpha diversity based on Chao 1 at genus level and visualization of taxonomic composition were conducted using R statistical software (v.3.5.1) with phyloseq (v. 1.22.0)[54] and fantaxtic (v. 0.1.0), respectively. Using MicrobiomeAnalyst,[55] differentially enrichment analysis was conducted based on package DESeq[56] after filtering OTU whose counts were lower than 10% prevalence in samples and 10% variance in inter-quantile range, which resulted in reduction of 2304 OTUs to 477 OTUs.

5.4 Results

5.4.1 Clinical measures and sample classification.

Collecting a complete and balanced set of biological samples from pediatric IBD patients at multiple time points during therapeutic interventions in both arms (CS and EEN) was challenging due to difficulties in specimen collection adherence when young children are included, and resistance among family in choosing a therapeutic intervention whose efficacy has not been well-studied.[40] This resulted in only one UC patient in the EEN arm

as compared with 16 CD patients. Also, missing urine and stool samples were prevalent over seven collection time points during 8 weeks of EEN therapy, which necessitated exclusion of these patients for longitudinal study purposes. Therefore, inclusion criteria were set as CD patients who responded the EEN therapy and provided baseline samples, as well as samples collected at 2-week or later time point during the therapy. Eleven patients in EEN arm who met these inclusion criteria were included in this time-resolved metabolomics study. Matching urine and stool samples were collected from most of these patients, and multivariate and univariate statistical analysis was applied (**Table S5.1, S5.2**), which provided complementary information on metabolic phenotype changes coinciding with therapeutic responses to EEN. Seven out of 11 were diagnosed for the first time and nine out of 11 were prescribed for EEN therapy for the first time. Additionally, clinical information of three patients in CS arm and one non-responder to EEN therapy is summarized in **Table S5.3**. All patients listed in **Table 5.1** responded to the therapy, and most of them achieved clinical remission, which was defined as PCDAI score ≤ 10. Serum CRP and stool FCP also decreased toward the values associated with remission (*i.e.,* CRP: 1.0 mg/L, FCP: 250 μg/g) (**Figure S5.2**). Statistically significant reduction in FCP was not achieved due to large between-subject variability. A total of 27 stool samples and 35 urine samples were collected from patients but at variable and inconsistent time points (**Table S5.1**). In order to maximize samples that can be included in the data analysis, we classified each samples into three inflammatory states based on an aggregate score derived from clinically used thresholds for CRP, FCP and disease score as an alternative to categorizing samples based on time point of collection (**Table 5.2**).

5.4.2 Exploratory analysis of urinary and stool metabolome and microbiome.

Metabolomic analysis of 39 stool extracts and 49 urine samples collected during clinical visits by pediatric CD patients were included in this study initially resulting in 130 and 163 metabolites detected in stool and urine samples, respectively. A rigorous data workflow in metabolomics when using MSI-CE-MS was applied to remove spurious, background and redundant signals.

Table 5.1. Pediatric CD patients demographics and clinical measurements shown in mean and standard deviation. All measurements were taken from serum except for fecal calprotectin.

	Number of patients	11	
	Age	13 ± 2	
	Sex; male : female	8 : 3	
	First time diagnosed (n)	7	
	First time EEN (n)	9	
	Disease location (n)		
	Ileocolonic	7	
	Ileocolonic + UGI	0	
	Colonic	2	
	Colonic + UGI	2	
	Maintenance medication (n)		
	Biologic[a]	5	
	Immunomodulator[b]	7	
	Supplementation (n)		
	Folic acid	5	
	Folic acid + iron supplement	2	
		Baseline	**Post 2-6 weeks**	**p-value[c]**
	CRP (mg/L)	46 ± 51	4 ± 4	0.024
	Fecal calprotectin (µg/g)	2364 ± 1680	973± 1165	> 0.05
	Hemoglobin (g/L)	110 ± 15	116 ± 8	> 0.05
	ESR (mm/hr)	34 ± 32	21 ± 10	> 0.05
	Albumin (g/L)	27 ± 4	34 ± 6	0.044
	Quality of Life score	28 ± 7	34 ± 7	> 0.05
	PCDAI	10 ± 8	6 ± 4	> 0.05

Abbreviations: ESR, Erythrocyte sedimentation rate; UGI, Upper Gastrointestinal tract; PCDAI, Pediatric Crohn's disease activity index
[a]Infliximab; [b]Methotrexate (one patients took a combination of infliximab and methotrexate)
[c]Based on Mann-Whitney U-test (α = 0.05)

Table 5.2 Classification criteria of inflammatory state for each sample based on disease score (PCDAI) and inflammatory markers.

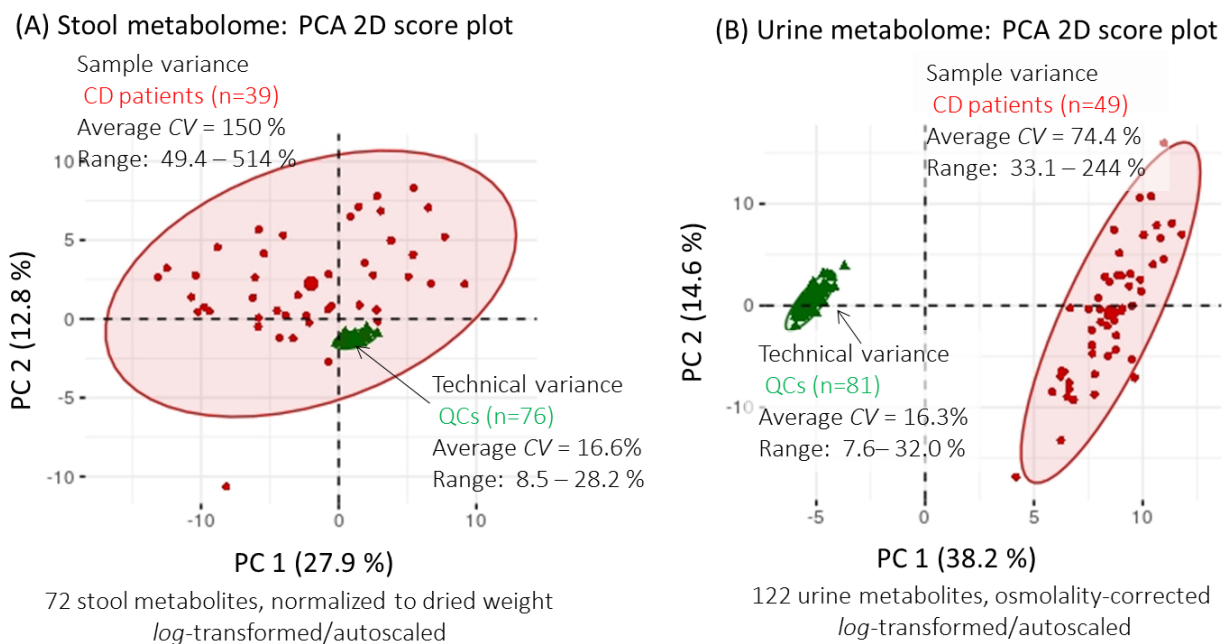| Classification	Criteria
Flaring	FCP > 250 or CRP > 1.0
Healing	FCP or CRP or disease score decreased from baseline
Remission	FCP ≤ 250 or CRP ≤ 1.0

FCP: Fecal calprotectin; CRP: C-reactive protein

Briefly, authentic yet reproducible metabolites from stool extracts or urine samples were first identified when using a dilution trend filter[46,57] and annotated based on their characteristic *m/z*:RMT under positive or negative ion mode detection. Additionally, metabolites that had poor technical precision in pooled QC samples analyzed in every run by MSI-CE-MS were excluded as unreliable signals (RSD > 40%). Subsequently, metabolites detected in less than 75% of samples were also removed in order to focus on representative metabolites that are measured consistently in most specimens, which is important to reduce data over fitting, skewed data and bias that can lead to false discoveries. Finally, the metabolomics data matrix was carefully inspected for co-migration of redundant molecular features, including in-source fragments, isotopes and adducts generated from the same metabolite. After this rigorous processing, a curated list of 72 stool metabolites and 122 urinary metabolites were reliably measured from 27 repeat fecal extracts from 8 CD patients, and 35 repeat urine samples from 11 CD patients in EEN arm, which were included in exploratory multivariate data analysis using principal component analysis (PCA) (**Figure 5.2**), as well as correlation analysis to inflammatory markers.

As shown in the PCA 2D scores plot, good long-term technical variance based on repeat analysis of pooled QC specimens (*n*=76) in every run was achieved for all metabolites measured in stool extracts (mean CV = 16.6%) and urine samples (mean CV = 16.3%) when using MSI-CE-MS (**Figure 5.1**). In contrast, a much greater extent of biological (between-subject) variance was reflected in the stool metabolome (mean CV = 150 %) as compared to urine metabolome (mean CV = 74%), which were both much larger than technical variance. Similarly, when inflammatory state classification code is used to visualize overall data trends using PCA, time-resolved stool and notably urine metabolome changes from patients categorized in "remission" tend to cluster tightly together reflecting an analogous metabolic phenotype for CD during EEN treatment, whereas samples at baseline and in intermediate transition (*i.e.,* "healing phase") showed a much wider distribution (**Figure 5.2a, b**).

(A) Stool metabolome: PCA 2D score plot

Sample variance
CD patients (n=39)
Average CV = 150 %
Range: 49.4 – 514 %

Technical variance
QCs (n=76)
Average CV = 16.6%
Range: 8.5 – 28.2 %

PC 2 (12.8 %)

PC 1 (27.9 %)

72 stool metabolites, normalized to dried weight
log-transformed/autoscaled

(B) Urine metabolome: PCA 2D score plot

Sample variance
CD patients (n=49)
Average CV = 74.4 %
Range: 33.1 – 244 %

Technical variance
QCs (n=81)
Average CV = 16.3%
Range: 7.6– 32.0 %

PC 2 (14.6 %)

PC 1 (38.2 %)

122 urine metabolites, osmolality-corrected
log-transformed/autoscaled

Figure 5.1. (a) PCA plots of 72 stool metabolites from 39 samples and **(b)** 122 urinary metabolites from 49 samples of pediatric CD patients with QC, which represents technical variance. Data were normalized with dried weight (stool) or osmolality (urine), log-transformed and autoscaled. CV: coefficient of variance.

Loading variables on PC2 of the urinary metabolome PCA plot were associated with increased urinary excretion of pyridoxic acid, aspartyl glutamine (putative assignment; **Figure S5.1a**), pantothenic acid and octanoyl glucuronide (putative assignment; **Figure S5.1b**), whereas PC1 was largely influenced by mannopyranosyl-tryptophan (Man-Trp) and methylcytidine, which resulted in general clustering of samples in the "flaring" category in the lower right quadrant (**Figure 5.2b**).

As expected, a larger heterogeneity was observed among measured changes in the stool metabolome following EEN therapy, but a general trend of moving across PC3 when transitioning from baseline to remission state was observed. Between-subject variance was greater among stool metabolites compared to that of urine, which resulted in clusters of samples from the same patient along PC1 and PC2 instead of clusters that represent the inflammatory state.
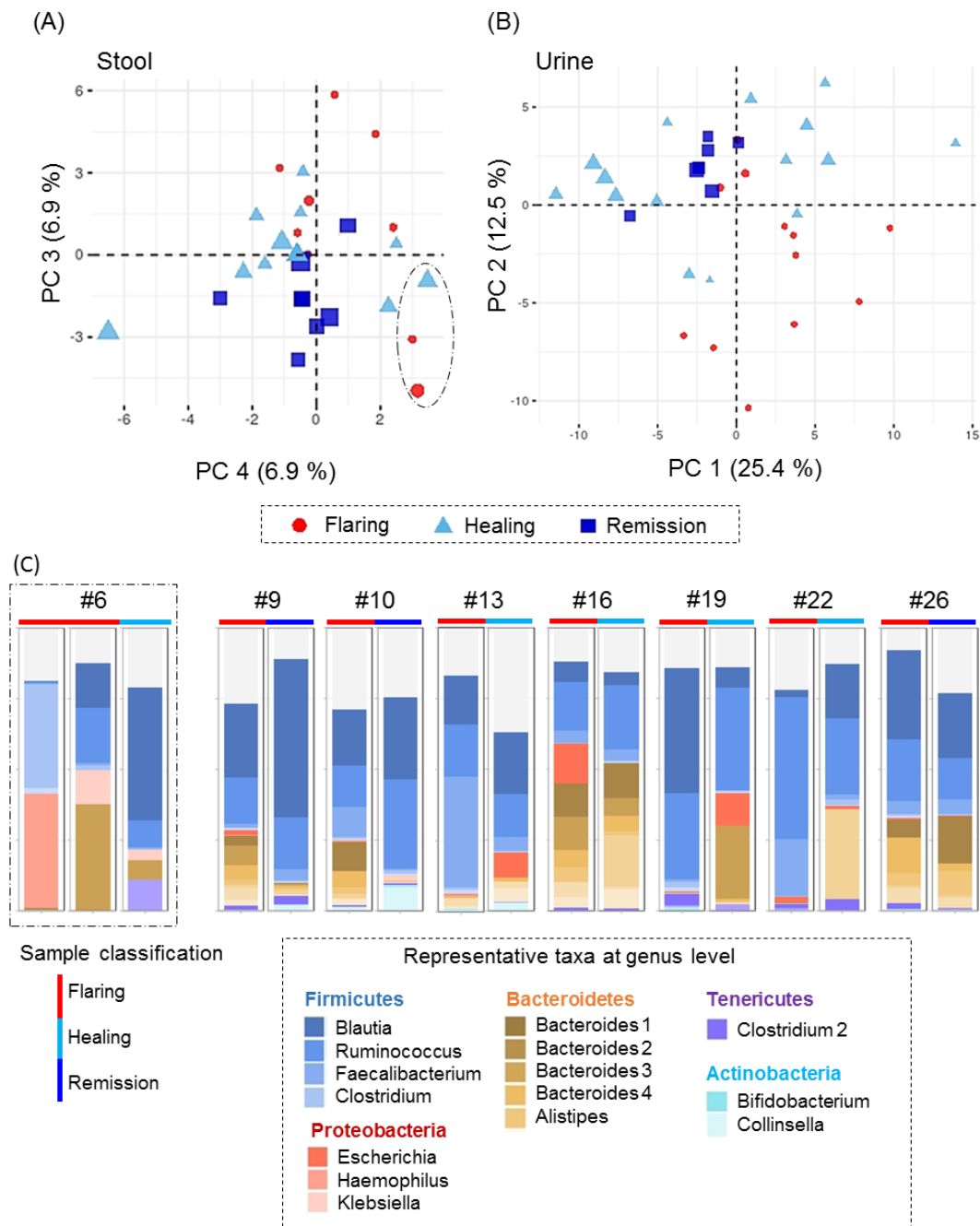
Figure 5.2. Principal component analysis (PCA) of dynamic changes in (A) 72 stool metabolites measured from 8 CD patients and (B) 122 urinary metabolites of 11 CD patients following EEN therapy. Principal components 3 and 4 were selected in the plot of stool metabolome for better between-subject clustering. Size of data points reflect time point of collection (*i.e.* small size = early time point) and the colour is based on the inflammatory state criteria on **Table 5.2**. Three stool samples from patients #6 is circled with dotted line on stool metabolome PCA plot to highlight the deviation in metabolic profiles compared to other samples. (C) Relative abundance plot of 20 most abundant genera identified from samples at baseline and closest to the end of the treatment. Representative genera are highlighted in a dotted box.

Thus, PC3 and 4 were chosen for better visualization of the separation between "flaring" and "remission" phases. PC3 generally represented a metabolic trajectory of decreasing diacetylspermine (putative assignment; **Figure S5.1c**), acetylaspartate and 3-methylhistidine toward the lower half of the plot where samples categorized as "remission" clustered. These trends in stool metabolome were not applicable to subject #6, who was diagnosed as colonic CD. This patient's samples took the opposite trajectory of other samples along PC 3. Stool microbiome composition of this patient was also found distinct from others and characterized by significant abundance of genus *Haemophilus* and *Clostridium*, while all other samples showed abundance of *Blautia* and *Ruminococcus* (**Figure 5.2c**). These two genera, *Blautia* and *Ruminococcus,* in this patient's samples subsequently increased at 2-week and 4-week following initiation of EEN therapy. Alpha diversity measure of fecal microbiome also showed a general decreasing trend for the majority of pediatric CD patients, while it increased for three patients, including subject #6 (**Figure S5.3**). To avoid introducing additional confounding factor in this pilot study, samples from this patient were removed from subsequent data analysis.

5.4.3 Correlation analysis of stool and urinary metabolites with inflammatory markers.

Spearman correlation analysis was next applied to fecal and urinary metabolites to investigate the potential association with classic inflammatory markers used to assess IBD (*i.e.,* serum CRP and stool FCP) at baseline and over the course of EEN therapy. Fecal metabolites that showed strong positive correlations with decreasing inflammation following EEN therapy (Spearman rho > 0.50 and $p < 0.05$; hence metabolite responses decreased over time) were mostly primary amino acids (*e.g.,* threonine and tryptophan) and methylated amino acid analogs (*e.g.,* trimethyllysine, 3-methylhistidine) (**Figure 5.3a**; **Table S5.4**). All of these metabolites also showed moderate to strong positive correlations with each other (*i.e.,* collinear).
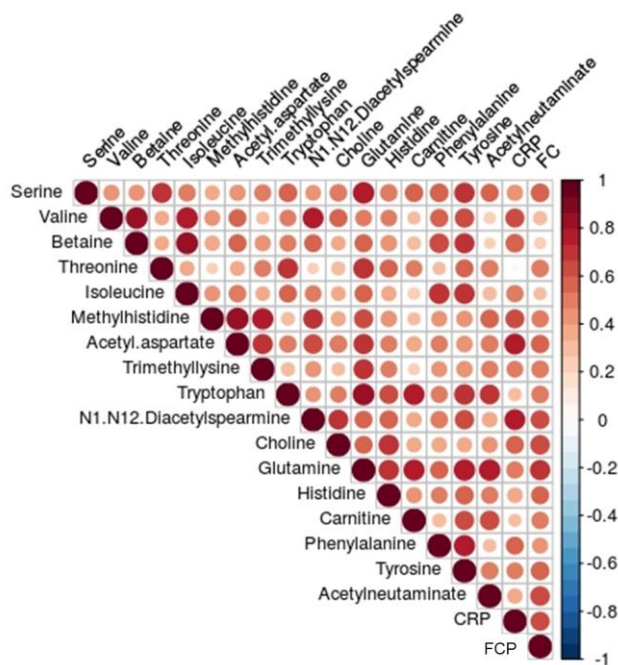
Figure 5.3 (a) Heatmap of a Spearman correlation analysis between fecal metabolites and inflammatory markers (serum CRF and stool FCP) with a cut-off value of rho ≥ 0.5 or rho $\leq$ - 0.5 and $p < 0.05$. Abbreviations include CRF: C-reactive protein and FCP: Fecal calprotectin.
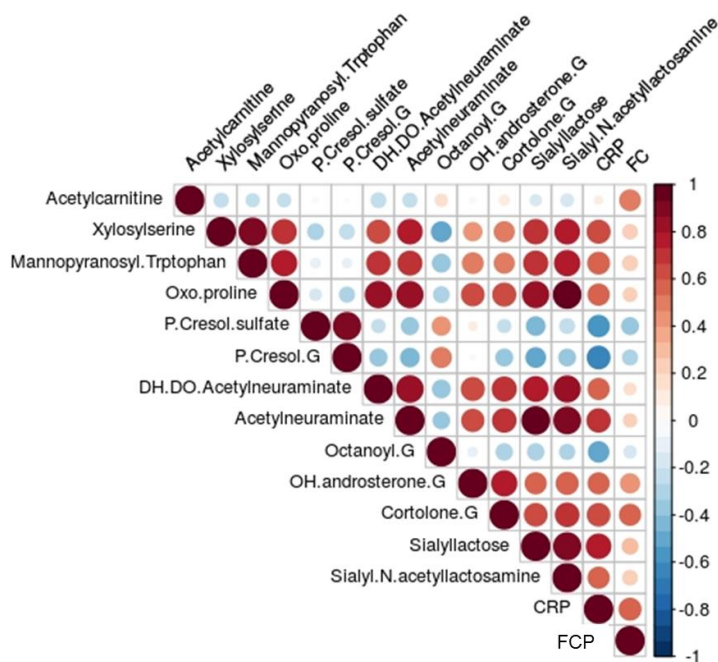
Figure 5.3b. Heat map of a Spearman rank-order correlation analysis between urinary metabolites and inflammatory markers with a cut-off value of rho ≥ 0.5 and rho $\leq$ - 0.5 and $p < 0.05$. Abbreviations include G: Glucuronide, OH: Hydroxyl, DO: Deoxy, CRP: C-reactive protein and FCP: Fecal calprotectin.

Other fecal derived metabolites that did not belong among amino acids were *N1-N2*-diacetylspermine (putative assignment; polyamine) and *N*-acetylneuraminic acid (amino sugar). Noteworthy, *N1-N2*-diacetylspermine was correlated with both CRP and FCP with correlation coefficient of 0.74 ($p < 0.001$) and 0.60 ($p < 0.01$), respectively (supporting MS/MS spectra in **Figure S5.1c**).

In contrast to metabolites from stool extracts, some primary and modified amino acids in urine showed negative correlations, in other words, had an increasing trend over the course of EEN therapy (**Figure 5.3b**; **Table S5.5**). These compounds were bacterial co-metabolites derived from aromatic amino acids, such as *p*-cresol sulfate and its intact glucuronide conjugate (supporting MS/MS spectra **Figure S5.1d**), as well as a compound putatively assigned as octanoyl glucuronide (**Figure S5.1b**). Unlike fecal metabolites, most urinary metabolites with strong correlations with CRP and FCP were not co-linear with each other except for a sub-set of metabolites that showed nearly perfect correlation (Spearman rho ≈ 1.0), which is indicative of a common underlying metabolic pathway. Notably, glycosylated amino acids, mannopyranosyl-tryptophan (Man-Trp) and a compound putatively identified as xylosyl-serine (Xyl-Ser, supporting MS/MS spectra in **Figure S5.1e**) showed a strong positive correlation with each other, as well as with CRP. Additionally, four sialic acids, namely *N*-acetylneuraminic acid, 2,3-dehydrodeoxy-*N*-acetylneuraminic acid, 3'-sialyllactose and sialyl-*N*-acetyllactosamine also showed strong positive correlations to CRP (Spearman rho from 0.54 to 0.73). Noteworthy, *N*-acetylneuraminic acid and 3'-sialyllactose also showed strong correlations in Pearson product moment correlation analysis as well, which considers corresponding changes in raw/untransformed data, instead of changes in ranked values as in Spearman correlation (**Figure S5.4**). Spearman correlation was used by default due to largely skewed data structure observed for majority ($> 70\,\%$) of metabolites in urine and stool samples. Among all urinary metabolites with significant correlation with inflammatory markers, the only class of metabolite that showed strong correlation with both CRP and FCP were tentatively ascribed as steroid conjugates, including hydroxyandrosterone glucuronide and cortolone glucuronide (supporting MS/MS spectra in **Figure S5.1f, g**).

5.4.4 Dynamic metabolic trajectories and microbial genera changes following EEN therapy.

Time-resolved changes in metabolite levels between baseline and at 2 weeks or after the start of the therapy were next analyzed using a heteroscedastic paired-Wilcoxon rank-sum test. While correlation analysis showed overall trends in metabolites regardless of individual trajectory in healing process, a paired non-parametric univariate test takes advantage of the repeat measures study design of stool and urine samples collected from the same patients and helps to identify common metabolic signatures undergoing dynamic changes during EEN therapy. A few fecal metabolites, such as guanine, choline, 3-methylhistidine and a compound putatively identified as imidazole propionate (refer to high resolution MS and MS/MS spectra in **Figure S5.1h**) showed significant decreases in response following therapy; although these stool metabolites were statistically significant ($p < 0.05$), they did not satisfy a false discovery (FDR, Benjamini-Hochberg) adjustment for multiple hypothesis testing ($q < 0.05$) (**Table S5.7**). Small sample size and the rather high biological variability of stool extracts likely contributed to this outcome despite acceptable long-term technical variance (**Figure 5.1**). Additionally, the statistical significance of diacetylspermine, which was correlated positively with CRP and FCP, was influenced by one subject whose value increased, while the trend for all of other patients dropped, highlighting large between-subject variations in treatment responses in some cases (**Figure 5.4**). The average FCP in response of this patient marginally decreased from baseline (3393 µg/g) to 4-weeks (3273 µg/g) when stool specimens were collected, but this value was still the highest among all other EEN responders at the same time point and well-above the critical threshold of inflammation (≤ 250 µg/g)[58] for this patient to be considered as in remission. Interestingly, fecal butyric acid was not found to be significantly changed following EEN therapy. Also, microbial differential abundance analysis of 14 fecal specimens from seven pediatric CD patients showed a significant reduction in genera *Veillonella* and *Haemophilus* between the two time points following EEN therapy (**Figure 5.5**). Other genera, such as *Porphyromonas* and *Lactococcus* showed increase and decrease

between these two time points respectively, but their changes over time were not significant after multiple hypothesis testing correction (FDR, Benjamini-Hochberg $q < 0.05$).

As expected from PCA exploratory data analysis, osmolality normalized urinary metabolites displayed much more distinctive changes in their trajectories from baseline over a 2-week time course when compared to changes in stool metabolites. For instance, dimethlyglycine (a major catabolite of choline) displayed an increasing trajectory over time, which was found however to decrease in stool extracts. Other urinary metabolites that significantly increased during EEN intervention were octanoyl glucuronide (OG), a putatively identified dipeptide, aspartyl-glutamine (Asp-Gln) and p-cresol as its sulfate and glucuronide conjugate (refer to high resolution MS and MS/MS spectra in **Figure S5.1**), as well as two vitamins likely derived from intake of the EEN formula, namely pantothenic acid (vitamin B5) and pyridoxic acid (vitamin B6). Highest median fold-changes (FC) were observed in octanoyl glucuronide (13.7), as well as Xyl-Ser and trigonelline (0.26) when the ratio of 2-week time point over baseline was considered (**Figure 5.4**; **Table S5.6**). All urinary metabolites highlighted in **Figure 5.4** were also significant after multiple hypothesis testing correction based on FDR ($q < 0.05$).

Common classes identified for urinary metabolites that underwent a decrease in response over time from baseline were amino acid derivatives, Xyl-Ser, trimethyllysine, oxo-proline, as well as pyrimidine/purine metabolites, including methylcytidine and uric acid. Additionally, trigonelline, and a steroid conjugate putatively identified as cortolone glucuronide also decreased significantly during EEN therapy. Indeed, a majority of urinary metabolites with significant changes identified in this Wilcoxon test were consistent with significant compounds identified previously with the Spearman correlation analysis with CRP and FCP.
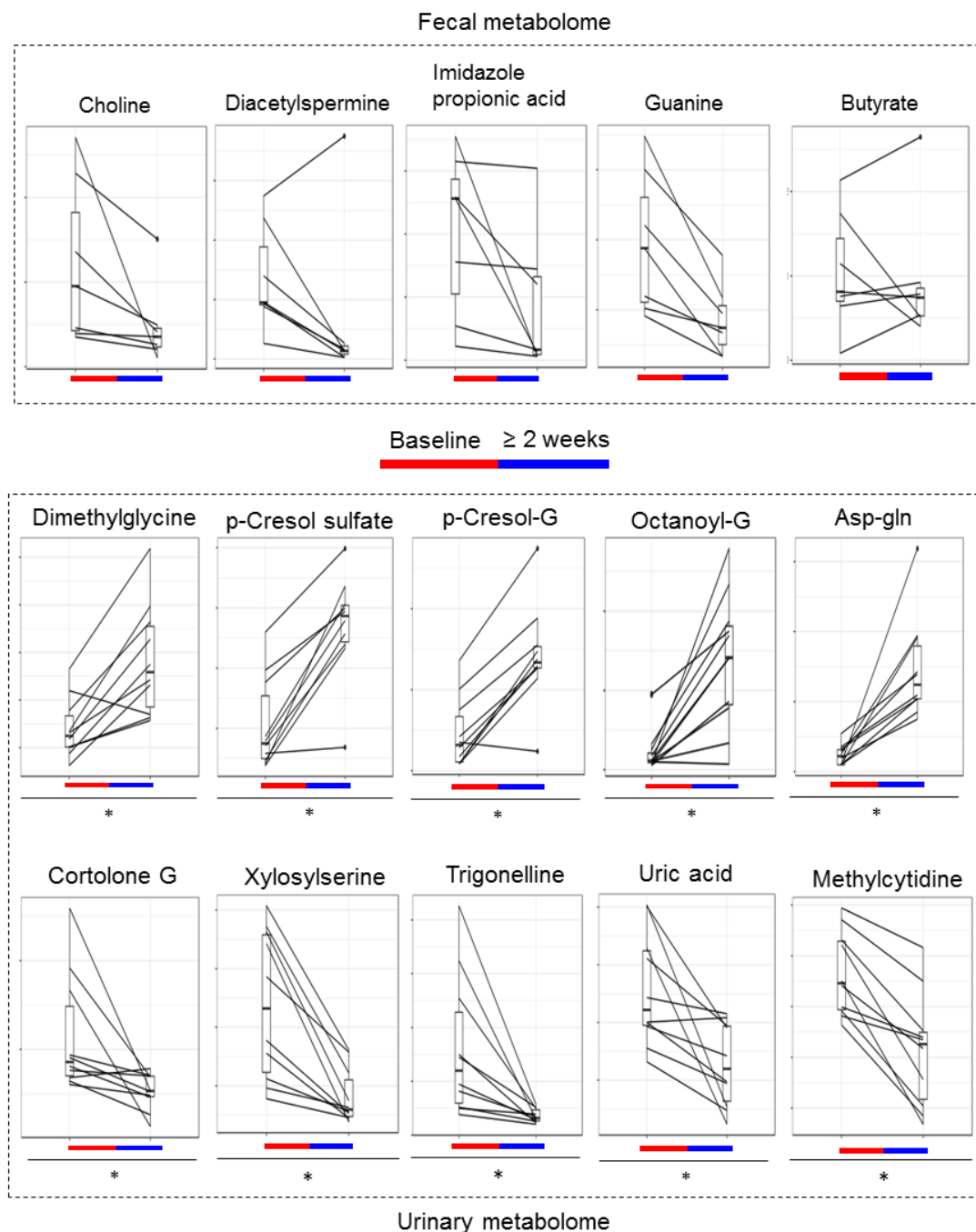
Figure 5.4. Metabolic trajectories for a cohort of pediatric CD patients based on a paired box plots of dynamic changes in the response of fecal and urinary metabolites from baseline and after two weeks following initiation of ENN therapy based on paired Wilcoxon rank-sum test. All metabolite were found to undergo significant changes in stool or urine over time ($p < 0.05$) with the exception of fecal butyrate, whereas * indicates statistical significance following a FDR-adjustment, $q < 0.05$. Abbreviations include, Asp-Gln: Aspartyl-glutamine dipeptide, and G: glucuronide.
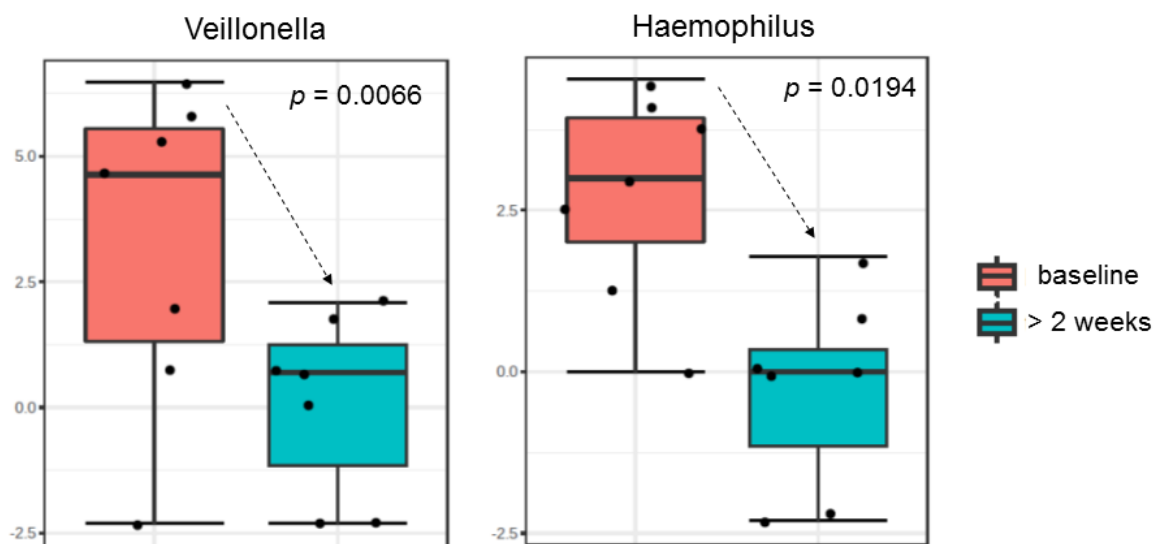
Figure 5.5. Differentially enriched microbial genera from matching fecal specimens collected at baseline and after two weeks following EEN therapy on a cohort of pediatric CD patients (except for subject #19, 1-week). Y-axis shows *log*-transformed counts and FDR adjusted *p*-values ($q < 0.05$).

5.4.5 Stability of metabolites as putative biomarkers of treatment responses to ENN.

Based on records taken by parents and nurses who received samples during clinical visits, a sub-set of urine and stool specimens were brought to a clinic over 24 h after sample collection at home. This pre-analytical source of variation can be a significant concern for contributing to potential bias caused by metabolite instability especially in a sample with high metabolic activity, such as stool. As shown in our previous studies in *Chapter IV*, stool metabolites tend to be highly unstable due to bacterial activities even 2 h after collection while being stored at refrigerated at +4 ^{o}C. Levels of key stool metabolites identified in this study also changed considerably as a function of a delay time to storage following an increasing trend when not immediately stored frozen at -80 ^{o}C (**Figure S5.5**). However, choline and imidazole propionate were relatively unaffected by the duration and temperature of the storage, whereas guanine and diacetylspermine increased over time with an increasing delay to storage. Butyrate, on the other hand, showed divergent trends between the twin boys examined, which may be due to different microbial community harboured in their intestine. Additionally, initial levels of diacetylspermine were strikingly

different between them, which showed an increasing trend as delay in sample storage was extended. Considering the significant between-subject variance despite their the similarities in diet and genes, the pre-analytical variance caused by delays to stool storage was likely the major factor associated with the small effect sizes measured for top-ranked stool metabolites following EEN therapy of pediatric CD patients.

Compared to the most significant metabolites identified in stool, urinary metabolites showed remarkable stability irrespective of storage conditions or time delay to storage with the exception for *p*-cresol glucuronide, *N*-acetylneuraminic acid and methylcytidine when stored at room temperature (**Figure S5.6**). While stool metabolites generally increased, significant changes in labile urinary metabolite levels were found to display a decreasing trend in average responses over time, which may be due to chemical or microbial degradation of these compounds. Conversely, none of the key urinary metabolites showed significant changes when samples were stored in a fridge at 4 °C even up to 48 h. This result assures the robustness of our urinary metabolome data acquired from pediatric CD patients when using 1.0 mM azide as an antimicrobial agent/preservative, which were stored in a home fridge for average of 4.5 h with a maximum of 48 h based on our records obtained at clinical visits.

5.4.6 Confirming metabolic trajectories in CS responders and EEN non-responder

We next confirmed whether biomarker candidates reflecting positive treatment responses to EEN identified in urine and stool extracts would follow the same trend in CS treatment arm (*n*=3, patient #12, 21, 27) and an EEN non-responder (*n*=1, patient #17) when comparing metabolic trajectories of representative CD patients considered as optimum EEN responders who achieved clinical remission (patient#9, 22, 26, *n*=3) as shown in **Figure 5.6**. Only urinary metabolites were investigated in this case because of their greater chemical stability and lack of pre-analytical storage artifacts, as well as larger number of samples collected from CD patients, and greater effect sizes from non-parametric statistical analysis.
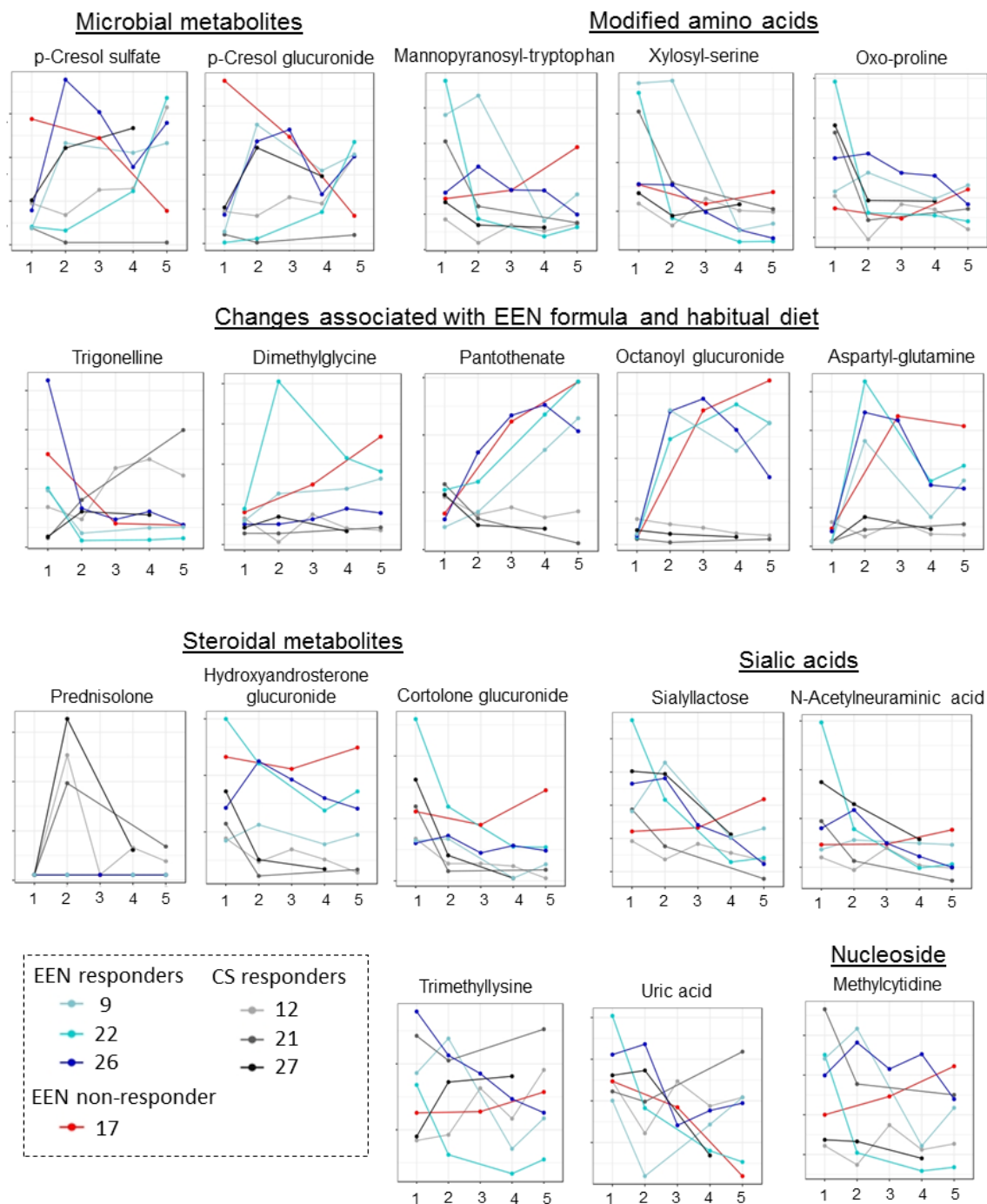
Figure 5.6. Line plots of urinary metabolite trajectories for pediatric CD patients that showed distinct trends among three representative EEN responders (n=3), an EEN non-responder (n=1) and three CS responders (n=3). Time point 5 includes week 6 and 8 when week 4 sample is not available. Y-axis corresponds to relative peak area.

Overall, similar metabolic trajectories were observed when comparing EEN and CS responders in terms of dynamic changes in p-cresol sulfate and its glucuronide conjugate, as well as glycosylated amino acids, sialic acids, trimethylysine, methylcytidine and uric acid. Some metabolites associated with vitamin intake from EEN formula, such as pantothenic acid, and octanoyl glucuronide showed rapid increases within a week from initiation of therapy only in the EEN arm confirming dietary adherence, while there was no notable changes in the CS arm. As expected, there was complete absence of prednisolone, the corticosteroid prescribed for CS arm, detected in urine samples of EEN-treated patients. Moreover, the sharp increase in prednisolone excretion corresponded to a rapid decrease in compounds putatively identified as cortisol metabolites in urine of CS-treated patients within a few days from the intravenous administration of prednisolone to affected pediatric CD patients. These putatively identified cortisol metabolites also decreased in EEN-responders, but with more gradual trend compared to the CS-arm. Conversely, these metabolites showed an increasing trajectory over time in EEN non-responders, along with other metabolites, such as sialic acids and glycosylated amino acids, which decreased in urine of treatment responders. An opposite trend was also noted in urinary p-cresol conjugates, which decreased in treatment responders. Trimethyllysine and uric acid, on the other hand, showed largely mixed trend or no change among CS responders and EEN non-responder.

5.5 Discussion

5.5.1 Impact of localized inflammation and changes in metabolites and microbiome.

The main objective of this study was to identify metabolic trajectories and gut microbial changes associated with treatment responses to EEN therapy that lead to clinical remission. A cohort of pediatric CD patients were recruited in this pilot study who were heterogeneous in terms of the affected GI location, while majority of them were newly diagnosed and treatment naïve for EEN. Several clinical studies to date have shown that EEN is most effective when used for the first time than after repeated EEN treatments.[59] Indeed, all patients who were prescribed EEN for the first time in our study responded well, whereas

one out of two patients who had previously been treated with EEN did not achieve clinical remission. In terms of disease location, EEN has been shown to be effective regardless of affected area; however, one earlier study[60] showed reduced efficacy for colonic CD and a more recent study showed no clinical effect for perianal CD.[61] Interestingly, one patient whose fecal metabolic and microbial profile was distinct from other patients had a colonic CD phenotype (**Figure 5.1**). Although this patient responded to the therapy based on disease score (PCDAI ≤ 10) with a decrease in inflammatory markers achieved by week 4, distinct fecal microbial composition and corresponding fecal metabolic profile may indicate that differences in affected location may create an environment that favour certain bacterial strains than others. The striking dominance of *Haemophilus* in the baseline sample of the colonic CD patient also suggests that inflammation was likely triggered by infection-mediated immune reaction that favours the growth of this organism. *Haemophilus* belongs to a class *Gammaproteobacteria*, which is composed of facultative anaerobes and thus, not expected to be abundant in strictly anaerobic environment, such as the colon.[62,63] Conversely, the abundance of this class is indicative of an oxidizing microenvironment within the colon. In fact, this genus has been found in treatment naïve CD patients at much higher abundance than healthy children, indicating that CD patients are exposed to higher oxidative stress in the intestine.[22] Indeed, the abundance of *Haemophilus* decreased significantly between baseline and at or after 2 weeks into the therapy for all EEN therapy responders in our study (**Figure 5.5**). Another genus that significantly decreased following EEN therapy was *Veillonella*, which is known to adhere to mucosal layer and potentially exacerbate symptoms of extra-intestinal manifestation of IBD in the liver and gall bladder.[64] This genus has also been directly associated with concentrations of the fecal polyamine, cadaverine, which shares a biochemical pathway with N(1),N(2)-diacetylspermine.[65,66] In our case, N(1)-N(2)-diacetylspermine, which decreased in fecal extract specimens following EEN therapy as compared to baseline (median FC = 0.14 , $p > 0.05$, effect size = 0.43), has been shown to contribute to oxidative stress on cellular metabolism when it is upregulated, which promotes virulence of certain pathogenic strains of bacteria, such as *Streptococcus pneumonia* and *Helicobacter pylori*.[67] Similarly,

increased expression of N(1)-N(2)-diacetylspermine has been also reported to impact colonic bacterial biofilms that increase risk for colon cancer.[68] Microbial involvement in the production of this compound was also implicated in our chemical stability study, where significant increase was measured after 8 h of room temperature storage (**Figure S5.5**).

There were several other metabolites associated with gut microbiota activity that were found to undergo significant changes following EEN therapy (**Table S5.6, 5.7**), including imidazole propionate in stool extracts (median FC = 0.47 , p = 0.0156, effect size = 0.63), and both *p*-cresol sulfate (median FC = 3.8 , p = 1.90E-03, effect size = 0.63) and *p*-cresol glucuronide conjugates (median FC = 3.2 , p = 3.90E-03, effect size = 0.60) excreted in urine. Bacterial involvement in the production of imidazole propionate has been known for almost a century,[69] and its presence in urine and stool has been associated with various GI disorders, such as coeliac disease[70] and more recently, colorectal cancer.[71] The exact biochemical mechanism is still incompletely understood, but as an imidazole derivative likely derived from the metabolism of histamine, a mediator of local immune response, it may indicate EEN-induced lowering of immune activation. In contrast to N(1)-N(2)-diacetylspermine, this compound was relatively unaffected by delays to storage and storage conditions, which may indicate the involvement of largely anaerobic bacteria in its metabolism. Metabolic trajectories for the *p*-cresol sulfate and *p*-cresol glucuronide (which are co-linear, rho ≈ 1.0) increased in urine during EEN therapy in contrast to changes measured in stool gut microbiota derived metabolites. In this case, *p*-cresol is produced via bacterial proteolysis of tyrosine under an anaerobic environment and only certain bacteria, especially *Clostridia*, are known to be involved in its primary metabolism,[72-74] followed by its absorption in the colon, secondary metabolism as its sulfate and glucuronide conjugates in the liver, and subsequent excretion in urine.[75] Increased excretion of *p*-cresol sulfate in urine has been shown to be associated with various pathological conditions, such as insulin resistance in chronic kidney disease,[76] autism spectral disorder in pediatric population,[77] and *Clostridium difficile* infection.[78] Paradoxically, significantly lower urinary excretion of *p*-cresol sulfate has been reported in CD patients as compared to healthy controls,[79] and reduction of commensal *Clostridia* has been replicated in treated CD patients in

remission.[80] Although largely known as deleterious bacterial species because of the widely known pathogenicity of *C. difficile, Clostridia spp.* represent a diverse class of commensal bacteria involved in intestinal homeostasis.[81] Therefore, the increasing trajectories for urinary *p*-cresol conjugates following EEN therapy may indicate modulation of these commensal strains capable of *p*-cresol production that is associated with lower inflammation and improved GI function.

One of the most investigated class of microbial metabolites in the context of GI diseases is short-chain fatty acids (SCFAs), especially butyric acid, which is a major energy source of colonocytes and associated with intestinal homeostasis.[82] Two previous studies on metabolite and microbiome changes of pediatric CD patients following EEN therapy specifically targeted fecal SCFAs and investigated microbial changes associated with SCFA levels; however, a study conducted by Tjellstrom *et al.*[61] reported an increase in fecal butyric acid, while Gerasimidis *et al.*[29] reported the opposite trend. As authors of these studies suggested, fecal SCFAs levels are highly dependent on dietary fibre intake, in addition to the presence of intestinal bacteria that are capable of butyrate production via fermentation of carbohydrates. Additionally, bacterial production of butyric acid may not be effective in mitigating the intestinal inflammation if the butyrate receptor is defective due to genetic factors, which has been reported in subset of CD patients.[83,84] Fecal butyrate levels in our study also showed inconsistent changes with EEN therapy ($p > 0.05$) since a modest reduction observed in two patients was obscured by a marginal increase or no net change in others (**Figure 5.4**). There is clearly a complex interplay between intestinal bacteria, habitual diet, and genetics when considering the beneficial role of bacterial metabolites in intestinal homeostasis. Similarly, the large heterogeneity and modest sample size in our pilot study may have resulted in the lack of measurable changes in bacterial genera reported in other studies, such as increase in probiotic strain, *Bifidobacterium spp*, and reduction of *Bacterioides/Prevotella.*

5.5.2 Metabolite changes associated with EEN formula and habitual diet.

Metabolic profiles, especially that of urine, tends to be strongly influenced by changes in habitual diet[52] and thus, one of challenges in this study was to distinguish metabolites/nutrients directly associated with intake of EEN formula from downstream changes in host/gut microbiota metabolism as a result of EEN intervention that coincide with clinical remission. As expected, initiation of EEN therapy constitutes a drastic change in dietary patterns and oral feeding from baseline as it is comprised of a liquid formula delivered via nasogastric tubing that contains complete nutrition to meet the energy demands of CD patients over 8 weeks. As a result, fecal and urine metabolic profiles were found to reflect an increase in specific formula constituents, as well as a decrease in metabolites of common components derived from habitual diet. For example, a significant increase in urinary excretion of pantothenic acid (median FC = 3.3 , p = 5.92 E-03, effect size = 0.58), which is included in the formula as vitamin B5, and pyridoxic acid (median FC = 4.3 , p = 5.92 E-03, effect size = 0.58), a metabolite of pyridoxine or vitamin B6 (**Table S5.8**) were measured following EEN therapy as compared to baseline. Additionally, a large increase was found in urinary excretion of octanoyl glucuronide (median FC = 13.2, p = 3.90 E-03, effect size = 0.60) and dimethylglycine (median FC = 2.1, p = 3.90 E-03, effect size = 0.60), which are metabolites of medium-chain triglycerides[85] and choline,[86] respectively that are also constituents of the EEN formula. An increase in a dipeptide putatively identified as aspartyl-glutamine in urine was also observed, which may be derived from hydrolysis of whey protein in the formula. Furthermore, an absence of changes in these metabolites in patients in the CS treatment arm (*n*=3) further confirmed that these urinary metabolites are likely dietary biomarkers confirming adherence to EEN therapy by CD patients (**Figure 5.6**). Indeed, CD patients undergoing CS therapy were clearly evident based on the excretion of high levels of prednisone in urine as compared to baseline, which was not the case for patients in the EEN treatment arm.

Conversely, several other metabolites in this study serve as biomarkers of recent food intake or habitual diet, which decreased following EEN therapy in urine and stool extract

samples as compared to baseline. Notably, excretion of 3-methylhistidine[87] (median FC = 0.26, $p > 0.0469$, effects size = 0.53, **Table S5.7**) has been associated with protein intake, while trigonelline (median FC = 0.26, $p = 1.95$ E-03, effects size = 0.63) is an alkaloid associated with plant polyphenols, hence generally correlated with vegetable and caffeinated beverage intake.[36] Reduction of 3-methylhistidine was also observed in stool in our study, which is likely reflects changes in muscle protein turn-over due to whey supplementation as compared to habitual diets for CD patients who likely had a lower average daily protein intake.[87] There was also a reduction in guanine in stool extracts following EEN therapy. This is also likely a result of whey protein supplementation in the absence of animal protein intake following initiation of EEN, which was reflected by a decreased urinary excretion of uric acid that is a metabolic end product in the purine pathway.[88] Trimethyllysine is another metabolite recently shown to be abundant in various vegetables, while it had traditionally been considered to originate from breakdown of proteins, such as myosin and histone.[89] Trimethyllysine also serves as a precursor for *de novo* carnitine biosynthesis, an essential metabolite for the fatty acid transport and utilization in mitochondria,[90] which explains the reduced need for this energetically expensive pathway. CD patients in our study were supplemented with carnitine in the EEN formula (**Table S5.8**), which may have resulted in reduced need for breakdown of protein to release trimethyllysine for *de novo* biosynthesis of carnitine.

5.5.3 Metabolic trajectories associated with intestinal mucosal healing outcomes.

A major clinical goal for treatment of CD patients is mucosal healing associated with reduction in inflammatory response during remission, which are frequently measured by FCP and serum CRP. Indeed, there was a strong positive correlation (rho > 0.5, $p < 0.01$) between fecal amino acids (*e.g.,* tyrosine, glutamine, histidine) and these inflammatory markers, which decreased over the course of EEN therapy reflecting improved nutrient absorption within the colon as a result of mucosa healing. High excretion of fecal amino acids has been observed previously in CD patients in comparison to healthy controls, which is a hallmark nutrient malabsorption.[30] Additionally, better amino acid nutrition

supplemented with micronutrients and vitamins in the EEN formula was likely to have modulated a number of downstream metabolic pathways in affected CD patients, such as enhancement in intracellular glutathione production. Glutathione is a major thiol-based antioxidant that regulates redox cellular environment and protein activity via reversible glutathionylation, which is involved in various essential functions, such as xenobiotic detoxification and immune regulation.[91,92] Oxo-proline (pyroglutamic acid) is a catabolite of the glutathione cycle that can be recycled to regenerate glutamic acid via 5-oxoprolinase. In this case, glutathione depletion due to oxidative stress in the GI tract is associated with an accumulation of oxo-proline that is excreted in urine, which has been reported in malnourished adults[93] and children.[94] Conversely, reduced urinary excretion for oxo-proline following EEN therapy in this study suggests replenished precursor amino acids for glutathione biosynthesis (notably cysteine) and improved intracellular antioxidant capacity has been implicated as a key mechanism in the pathogenesis of IBD as indicated in antioxidant therapy used for its treatment.[95] Although not strongly correlated with CRP or FCP as inflammatory markers, modified nucleosides, including methylcytidine (median FC = 0.73, p = 5.92 E-03, effects size = 0.58) in urine and putatively identified methylthioadenosine (median FC = 0.48, p = 0.046, effects size = 0.50; supporting high resolution MS and MS/MS in **Figure S5.1j**) in stool decreased following EEN therapy. Modified nucleosides are primary degradation products of tRNA and indicative of high cell-turnover,[96] in this case, it is likely an indication of epithelial tissue damage in the intestine due to chronic inflammation. In fact, tissue methylation has been shown to play vital roles in protecting against mucosal inflammation as reflected by higher methylcytidine levels, which can be augmented by supplementation with folic acid as applied in a colitis murine model.[97] As a result, a reduction in metabolites associated with cellular methylation activity (*e.g.,* methylcytidine, choline, trimethyllysine) in urine or stool from pediatric CD patients further indicates healing of intestinal mucosa with a corresponding improvement in gut function for better nutrient absorption. In addition to moderate correlation with serum CRP (rho = 0.55, p = 4.25 E-03), oxo-proline was strongly correlated (rho > 0.8, p < 0.001) with sialic acids and two glycosylated amino acids, namely Man-Trp and a

compound putatively identified as Xyl-Ser, indicating that they form part of a common metabolic pathway or biological function. Indeed, sialic acids, including 2,3-dehydro-2-deoxy-*N*-acetylneuraminic acid, *N*-acetylneuraminate, 3'-sialyllactose and sialyl-*N*-acetyllactosamine, are ubiquitous terminal glycans on cellular surfaces and in glycoproteins, which are essential for signalling, recognition and immune functions.[98] Additionally, *N*-acetylneuraminic acid also serves as a terminal glycan in mucin, which is a component of mucus that protects epithelial layer of various organs, including intestine, from toxins and pathogens.[99] The strong positive correlation between serum CRP and urinary sialic acids, especially sialyllactose and sialyl-*N*-acetyllactosamine, has been previously reported in various inflammatory conditions, such as low-grade inflammation in healthy individuals,[100] rheumatoid arthritis[101] and CD.[102] Since acute phase glycoproteins, including CRP, are highly sialylated,[103] increased excretion of sialic acid metabolites in urine were likely derived from glycoproteins related to immune responses. In addition to notably changes in urinary sialic acid following EEN therapy, we also observed reduction in fecal *N*-acetylneuraminate, which may have originated from attenuation in the breakdown of intestinal mucosa during the healing process.

Although urinary excretion of glycosylated amino acids has not been previously reported in the context of CD, earlier studies have demonstrated that glycosylated tryptophan excreted in human urine was identified as Man-Trp from other possible configurations when using NMR.[104] Our *de-novo* characterization of high resolution MS/MS spectra (*Chapter III,* **Figure S3.2**) showed characteristic neutral loss of *m/z* 162, which corresponds to a hexose, as well as resulting peak of tryptophan (*m/z* 205.097). C-mannosylation is a unique, conserved glycosylation that occurs at the first tryptophan in the consensus amino acids sequence,[105] which has been found in proteins involved in immune functions, such as the thrombospondin type-I receptor super family and the cytokine receptor type-I family proteins.[106] Moreover, C-mannosylated peptides have been suggested to modulate production of tumor necrosis factor (TNF)-α, which is an acute phase cell signaling protein involved in systemic inflammation.[107] Similarly, xylosyl-serine

has been detected in human urine and associated with regulation of immune functions.[108] For instance, *O*-xylosyltransferase, an enzyme responsible for transfer of xylose to serine residue in a conserved amino acid sequence,[109] is also responsible for initiation of glycosaminoglycan (GAG) biosynthesis,[110,111] which modulate immune responses by interacting with potential pathogens in a similar way as sialic acids.[112] Taken together, immune functions and protective mechanism against oxidative stress seem to be the unifying theme behind the strong correlation among oxo-proline, sialic acids, and glycosylated amino acids excreted in urine. Thus, a significant reduction in the excretion of these urinary metabolites reflect lower oxidative stress and immune activation following EEN therapy, which coincides with positive clinical responses for pediatric CD patients due to reduced inflammation that allows for GI mucosal healing.

When changes of these metabolites identified to be involved in healing processes were compared between three representative EEN responders (*n=3*) and CS responders (*n=3*), generally the same metabolic trajectories were observed except for trimethyllysine, uric acid and putatively identified cortisol metabolites (**Figure 5.6**). Noteworthy, the rapid decrease in cortisol metabolites seen only in CS-treated patients between baseline and a few days after corresponds well to the excretion of excess prednisolone in urine (**Figure S5.1f,g,i**), which had been intravenously administered to these patients before the baseline sample collection. Cortisol is an essential hormone released from the adrenal cortex in response to adrenocorticotropic hormone, which is produced upon physiological or emotional stress.[113] The opposite trend between prednisolone and cortisol metabolite excretion has previously been reported,[114] and thus an expected metabolic signature measured in the small number of pediatric CD patients recruited for the CS arm. Interestingly, a more gradual but steady reduction in urinary cortisol metabolites were also measured in peditric CD patients following EEN therapy, especially for those patients with high level at baseline. This diet-induced suppression in stress hormone responses highlights the effectiveness of EEN therapy in reducing psychosocial and physical stress due to

chronic inflammation that is analogous to corticosteroid therapy but without the potential adverse effects on growth and development with long-term usage.

Nevertheless, we had one patient in the EEN treatment arm who did not respond to therapy based on unchanged PCDAI scores (15), despite a relatively low FCP of 276 µg/mg after 4 weeks. Additionally, the serum CRP of this patient increased from 3.3 to 3.8 mg/L between time point 3 (1-week) to 5 (4-week), which corresponded to increases in urinary metabolites (*e.g.* oxo-proline, Man-Trp) that correlated strongly with serum CRP in other patients (**Figure 5.6**). Based on metabolic trajectories that are specific to changes in habitual diet due to EEN formula intake, such as pantothenic acid and octanoyl glucuronide, adherence to the therapy does not seem to be the reason of her clinical non-responsiveness. These metabolic trajectory plots also highlight a lack of dynamic changes in metabolites from baseline to 1-week following EEN therapy for this patient, whereas most of the positive treatment responders showed rapid changes in urinary metabolite profiles by this early time point of the intervention. The underlying reason for this lack of treatment response is unclear from this work, however certain urinary metabolites associated with attenuation of immune response activation (*e.g.,* Man-Trp, xylosyl-serine) and oxidative stress (*e.g.,* oxo-proline) remained the same or increased during EEN therapy. For example, the greatest difference in metabolite trajectories between responders and the non-responder to EEN therapy was evident in time resolved profiles for urinary *p*-cresol sulfate and glucuronide conjugates suggesting that alterations in commensal gut microbiota was not successfully induced for this patient. This hypothesis could not be investigated in this study due to lack of stool samples from this non-responder.

5.5.4 Proposed mechanism of EEN efficacy and biomarkers for treatment response

We reported the involvement of biochemical pathways in cell cycling, protection against oxidative stress, and immune functions during EEN therapy in treatment-responsive CD patients for the first time through untargeted metabolomics using MSI-CE-MS. Additionally, modulation of intestinal microbiome corresponded to the response to EEN therapy, which was highlighted by significant reduction of opportunistic pathogens and

potential increase of bacterial strains responsible for production of *p*-cresol conjugates, which increased significantly in 90 % of EEN responders. Based on all observations in this study, we propose that mechanisms of EEN efficacy involves reduced immune activation through gut microbiota modulation and restoration of mucosal integrity and subsequent improvement of nutrient absorption through intestine (**Figure 5.7**). Since habitual diet may contain potential toxin and allergen that trigger immune response, effect of eliminating habitual diet is likely a part of mechanisms in EEN efficacy as well. Improved nutrient absorption and balanced nutrient in the formula supported and maintained protective cellular mechanisms against oxidative stress and reduced damages that used to be caused by toxic compounds produced by intestinal pathogens or components in habitual diet. We also highlighted the wide between-subject variability of stool metabolites and their instability in delayed storage situations at room and refrigerator (+4 °C) temperature. On the other hand, most of key urinary metabolites highlighted in this study were stable for at least 48 hours at room temperature, which is ideal to be used as a biomarker in clinical settings in which samples can be left unprocessed for extended amount of time.

Study limitations in this work included small sample sizes especially for stool, incomplete time points for sample collection, and lack of dietary information to infer potential triggers of immune response in diet at baseline. The small sample size and inherent variability in metabolite levels likely to have resulted in the relatively small changes in fecal metabolites during the therapy. Larger sample size would also help detecting changes in microbiome that corresponds to changes in urinary *p*-cresol conjugates, whose level best reflected treatment response to EEN. For these metabolites to be useful as biomarkers of treatment response, further studies are required with larger number of treatment responders and non-responders. Nevertheless, our study reported EEN-induced modulation of immune function at the equivalent extent as corticosteroid therapy that is free of pharmacological suppression of immune response and potential side effects. Currently, disease severity questionnaire and inflammatory markers, such as CRP and FCP, are primarily used to assess mucosal healing, predict relapse and response to a therapy.[115]
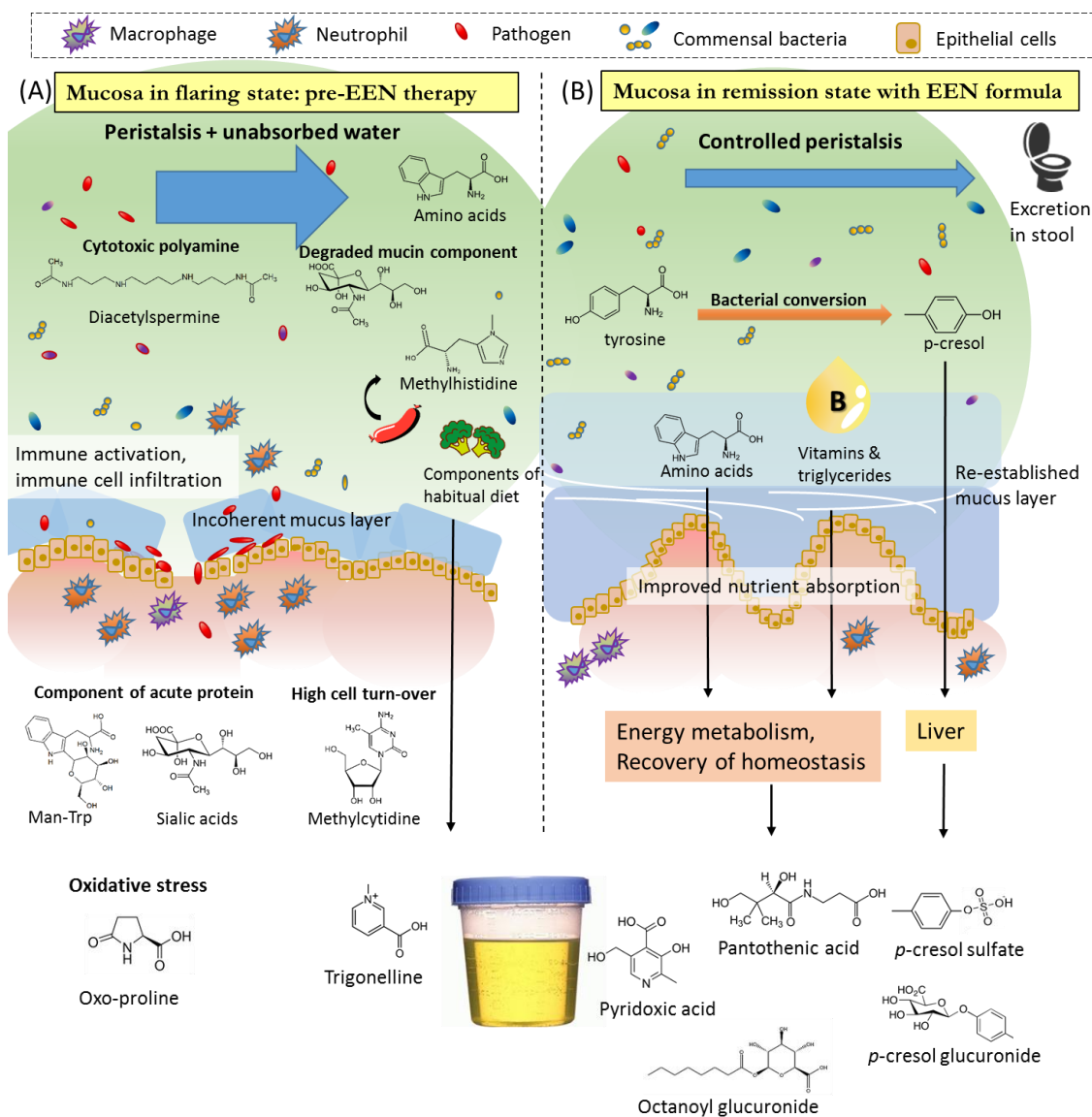
Figure 5.7 Schematic of GI and metabolic phenotypes of pediatric CD patients **(a)** before and **(b)** during EEN therapy. Active disease state is characterized by immune activation with microbial pathogen and potential antigen from food, resulting in mucosal degradation, neutrophil infiltration and pathogenic bacterial infiltration. Malabsorbed water and higher peristaltic movement causes diarrhea, which flushes nutrient components (*e.g.* amino acids), component of degraded mucus and other luminal contents into stool. Immune activation coincides with higher urinary excretion of components in immune-related proteins (*e.g.* Man-Trp, sialic acids), modified nucleosides from increased cell turn-over, and oxidative stress (oxo-proline). **(b)** During EEN therapy, reduction in pathogens and dietary antigen results in decreased immune activation and accelerated healing of mucosa, which supports proper absorption of nutrients in the formula (*e.g.* vitamins, amino acids, triglycerides) and associated recovery in energy homeostasis. Reduced water content in the intestinal lumen leads to controlled peristalsis and longer colonic transit time, leading to increased bacterial proteolysis and absorption of the end product (*i.e.* p-cresol), which is conjugated with sulfate and glucuronide and excreted in urine. Unused component and metabolic end products of the formula are also excreted in urine (*e.g.* octanoyl glucuronide, pyridoxic acid, pantothenic acid). During EEN therapy, excretion of metabolites derived from habitual diet significantly decreases (*e.g.* trigonelline).

Although FCP is generally effective in evaluating mucosal healing, it may not reflect underlying real-time biochemical changes that corresponds to symptoms as shown in the non-responder with relatively low FCP value in this study. We demonstrated that urinary excretion of *p*-cresol sulfate, glycosylated amino acids (Man-Trp, Xyl-Ser) and sialyllactose had opposite trajectories during the EEN therapy between responders and a non-responder. This clear trend and reliability of these metabolites under delayed storage of up to 48 h make them ideal candidates for biomarker panel to monitor and predict treatment responses to aid clinical decision making that results in suitable assignment of therapy within a short frame of time. Additionally, some urinary metabolites that showed significant increase as a result of formula intake (*e.g.,* octanoyl glucuronide) can serve as a marker of compliance, which is essential for the successful nutritional therapy. Metabolomics-based biomarkers hold promising potential for the development of optimal therapies for individual patients, which is urgently needed for complex multifactorial disorders, such as CD.

5.6 References

1. Gasparetto M, Guariso G. Highlights in ibd epidemiology and its natural history in the paediatric age. *Gastroenterology Research and Practice* 2013;**2013**:12.
2. Bernstein CN, Wajda A, Svenson LW*, et al.* The epidemiology of inflammatory bowel disease in canada: A population-based study. *Am J Gastroenterol* 2006;**101**:1559.
3. Ezri J, Marques-Vidal P, Nydegger A. Impact of disease and treatments on growth and puberty of pediatric patients with inflammatory bowel disease. *Digestion* 2012;**85**:308-19.
4. Ahmed I, Roy B, Khan S, Septer S, Umar S. Microbiome, metabolome and inflammatory bowel disease. *Microorganisms* 2016;**4**:20.
5. Lee D, Albenberg L, Compher C*, et al.* Diet in the pathogenesis and treatment of inflammatory bowel diseases. *Gastroenterology* 2015;**148**:1087-106.
6. Ueno F, Matsui T, Matsumoto T*, et al.* Evidence-based clinical practice guidelines for crohn's disease, integrated with formal consensus of experts in japan. *J Gastroenterol* 2013;**48**:31-72.
7. Lee J, Allen R, Ashley S*, et al.* British dietetic association evidence-based guidelines for the dietary management of crohn's disease in adults. *J Hum Nutr Diet* 2014;**27**:207-18.

8. Ruemmele FM, Veres G, Kolho KL*, et al.* Consensus guidelines of ecco/espghan on the medical management of pediatric crohn's disease. *J Crohns Colitis* 2014;**8**:1179-207.

9. Lafferty L, Tuohy M, Carey A*, et al.* Outcomes of exclusive enteral nutrition in paediatric crohn's disease. *Eur J Clin Nutr* 2016;**71**:185.

10. Soo J, Malik BA, Turner JM*, et al.* Use of exclusive enteral nutrition is just as effective as corticosteroids in newly diagnosed pediatric crohn's disease. *Dig Dis Sci* 2013;**58**:3584-91.

11. Connors J, Basseri S, Grant A*, et al.* Exclusive enteral nutrition therapy in paediatric crohn's disease results in long-term avoidance of corticosteroids: Results of a propensity-score matched cohort analysis. *J Crohns Colitis* 2017;**11**:1063-70.

12. Grover Z, Muir R, Lewindon P. Exclusive enteral nutrition induces early clinical, mucosal and transmural remission in paediatric crohn's disease. *Journal of Gastroenterology* 2014;**49**:638-45.

13. Grover Z, Lewindon P. Two-year outcomes after exclusive enteral nutrition induction are superior to corticosteroids in pediatric crohn's disease treated early with thiopurines. *Dig Dis Sci* 2015;**60**:3069-74.

14. Borrelli O, Cordischi L, Cirulli M*, et al.* Polymeric diet alone versus corticosteroids in the treatment of active pediatric crohn's disease: A randomized controlled open-label trial. *Clinical Gastroenterology and Hepatology* 2006;**4**:744-53.

15. Rubio A, Pigneur B, Garnier-Lengliné H*, et al.* The efficacy of exclusive nutritional therapy in paediatric crohn's disease, comparing fractionated oral vs. Continuous enteral feeding. *Aliment Pharmacol Ther* 2011;**33**:1332-9.

16. Otley A, Day AS, Zachos M. Nutritional management of inflammatory bowel disease. In: Mamula P, Markowitz JE, Baldassano RN, editors. *Pediatric inflammatory bowel disease*. 2nd edn.: Springer; 2013: 295-312.

17. Hou JK, Abraham B, El-Serag H. Dietary intake and risk of developing inflammatory bowel disease: A systematic review of the literature. *Am J Gastroenterol* 2011;**106**:563.

18. Jantchou P, Morois S, Clavel-Chapelon F, Boutron-Ruault M-C, Carbonnel F. Animal protein intake and risk of inflammatory bowel disease: The e3n prospective study. *Am J Gastroenterol* 2010;**105**:2195.

19. Sigall-Boneh R, Pfeffer-Gik T, Segal I*, et al.* Partial enteral nutrition with a crohn's disease exclusion diet is effective for induction of remission in children and young adults with crohn's disease. *Inflamm Bowel Dis* 2014;**20**:1353-60.

20. Ashton JJ, Gavin J, Beattie RM. Exclusive enteral nutrition in crohn's disease: Evidence and practicalities. *Clinical Nutrition* 2018.

21. David L, Maurice CF, Carmody RN*, et al.* Diet rapidly and reproducibly alters the human gut microbiome. *Nature* 2014;**505**.

22. Gevers D, Kugathasan S, Denson Lee A*, et al.* The treatment-naive microbiome in new-onset crohn's disease. *Cell Host & Microbe* 2014;**15**:382-92.

23. Morgan XC, Tickle TL, Sokol H*, et al.* Dysfunction of the intestinal microbiome in inflammatory bowel disease and treatment. *Genome Biology* 2012;**13**:R79.

24. Li S, Sullivan NL, Rouphael N*, et al.* Metabolic phenotypes of response to vaccination in humans. *Cell* 2017;**169**:862-77.e17.

25. Johnson CH, Ivanisevic J, Siuzdak G. Metabolomics: Beyond biomarkers and towards mechanisms. *Nature Review Molecular Cell Biology* 2016;**17**:451-9.

26. Keun HC. Metabonomic modeling of drug toxicity. *Pharmacology & Therapeutics* 2006;**109**:92-106.

27. Viant MR, Bundy JG, Pincetich CA, de Ropp JS, Tjeerdema RS. Nmr-derived developmental metabolic trajectories: An approach for visualizing the toxic actions of trichloroethylene during embryogenesis. *Metabolomics* 2005;**1**:149-58.

28. Nikolaus S, Schulte B, Al-Massad N*, et al.* Increased tryptophan metabolism is associated with activity of inflammatory bowel diseases. *Gastroenterology* 2017;**153**:1504-16.e2.

29. Gerasimidis K, Bertz M, Hanske L*, et al.* Decline in presumptively protective gut bacterial species and metabolites are paradoxically associated with disease improvement in pediatric crohn's disease during enteral nutrition *Inflamm Bowel Dis* 2014;**20**:861-71.

30. Marchesi JR, Holmes E, Khan F*, et al.* Rapid and noninvasive metabonomic characterization of inflammatory bowel disease. *J Proteome Res* 2007;**6**:546-51.

31. Miquel S, Martín R, Rossi O*, et al.* Faecalibacterium prausnitzii and human intestinal health. *Curr Opin Microbiol* 2013;**16**:255-61.

32. Sokol H, Pigneur B, Watterlot L*, et al.* Faecalibacterium prausnitzii is an anti-inflammatory commensal bacterium identified by gut microbiota analysis of crohn disease patients. *PNAS* 2008;**105**:16731.

33. Cao Y, Shen J, Ran ZH. Association between faecalibacterium prausnitzii reduction and inflammatory bowel disease: A meta-analysis and systematic review of the literature. *Gastroenterology Research and Practice* 2014;**2014**:7.

34. Berntson L, Agback P, Dicksved J. Changes in fecal microbiota and metabolomics in a child with juvenile idiopathic arthritis (jia) responding to two treatment periods with exclusive enteral nutrition (een). *Clinical Rheumatology* 2016;**35**:1501-6.

35. Toussi SS, Pan N, Walters HM, Walsh TJ. Infections in children and adolescents with juvenile idiopathic arthritis and inflammatory bowel disease treated with tumor necrosis factor–α inhibitors: Systematic review of the literature. *Clinical Infectious Diseases* 2013;**57**:1318-30.

36. Rothwell JA, Fillâtre Y, Martin J-F*, et al.* New biomarkers of coffee consumption identified by the non-targeted metabolomic profiling of cohort study subjects. *PLoS One* 2014;**9**:e93474.

37. Pero RW, Lund H, Leanderson T. Antioxidant metabolism induced by quinic acid. Increased urinary excretion of tryptophan and nicotinamide. *Phytotherapy Research* 2009;**23**:335-46.

38. Wijeyesekera A, Clarke PA, Bictash M*, et al.* Quantitative uplc-ms/ms analysis of the gut microbial co-metabolites phenylacetylglutamine, 4-cresyl sulphate and hippurate in human urine: Intermap study. *Analytical Methods* 2012;**4**:65-72.

39. Shaikhkhalil AK, Crandall W. Enteral nutrition for pediatric crohn's disease: An underutilized therapy. *Nutrition in Clinical Practice* 2018;**33**:493-509.

40. Seidman EG. Nutritional management of inflammatory bowel disease. *Gastroenterol Clin North Am* 1989;**18**:129-55.

41. Andrews HA, Lewis P, Allan RN. Prognosis after surgery for colonic crohn's disease. *BJS* 1989;**76**:1184-90.

42. Soucy G, Wang HH, Farraye FA*, et al.* Clinical and pathological analysis of colonic crohn's disease, including a subgroup with ulcerative colitis-like features. *Modern Pathology* 2011;**25**:295.

43. Hancock L, Beckly J, Geremia A*, et al.* Clinical and molecular characteristics of isolated colonic crohn's disease. *Inflamm Bowel Dis* 2008;**14**:1667-77.

44. Bligh EG, Dyer WJ. A rapid method of total lipid extraction and purification. *Canadian Journal of Biochemistry and Physiology* 1959;**37**:911-7.

45. Lin CY, Wu H, Tjeerdema RS, Viant MR. Evaluation of metabolite extraction strategies from tissue samples using nmr metabolomics. *Metabolomics* 2007;**3**:55-67.

46. Kuehnbaum NL, Kormendi A, Britz-McKibbin P. Multisegment injection-capillary electrophoresis-mass spectrometry: A high-throughput platform for metabolomics with high data fidelity. *Anal Chem* 2013;**85**:10664-9.

47. Macedo AN, Mathiaparanam S, Brick L*, et al.* The sweat metabolome of screen-positive cystic fibrosis infants: Revealing mechanisms beyond impaired chloride transport. *ACS Central Science* 2017;**3**:904-13.

48. DiBattista A, Rampersaud D, Lee H, Kim M, Britz-McKibbin P. High throughput screening method for systematic surveillance of drugs of abuse by multisegment injection–capillary electrophoresis–mass spectrometry. *Anal Chem* 2017;**89**:11853-61.

49. Yamamoto M, Ly R, Gill B*, et al.* Robust and high-throughput method for anionic metabolite profiling: Preventing polyimide aminolysis and capillary breakages under alkaline conditions in capillary electrophoresis-mass spectrometry. *Anal Chem* 2016;**88**:10710-9.

50. DiBattista A, McIntosh N, Lamoureux M*, et al.* Temporal signal pattern recognition in mass spectrometry: A method for rapid identification and accurate quantification of biomarkers for inborn errors of metabolism with quality assurance. *Anal Chem* 2017;**89**:8112-21.

51. Mahieu NG, Patti GJ. Systems-level annotation of a metabolomics data set reduces 25 000 features to fewer than 1000 unique metabolites. *Anal Chem* 2017;**89**:10397-406.

52. Walsh MC, Brennan L, Malthouse JPG, Roche HM, Gibney MJ. Effect of acute dietary standardization on the urinary, plasma, and salivary metabolomic profiles of healthy humans. *The American Journal of Clinical Nutrition* 2006;**84**:531-9.

53. Wishart DS, Feunang YD, Marcu A*, et al.* Hmdb 4.0: The human metabolome database for 2018. *Nucleic Acids Res* 2018;**46**:D608-d17.

54. McMurdie PJ, Holmes S. Phyloseq: An r package for reproducible interactive analysis and graphics of microbiome census data. *PLoS One* 2013;**8**.

55. Dhariwal A, Chong J, Habib S*, et al.* Microbiomeanalyst: A web-based tool for comprehensive statistical, visual and meta-analysis of microbiome data. *Nucleic Acids Res* 2017;**45**:W180-W8.

56. Anders S, Huber W. Differential expression analysis for sequence count data. *Genome Biology* 2010;**11**:R106.

57. Kuehnbaum NL, Gillen JB, Kormendi A*, et al.* Multiplexed separations for biomarker discovery in metabolomics: Elucidating adaptive responses to exercise training. *ELECTROPHORESIS* 2015;**36**:2226-36.

58. D'Haens G, Ferrante M, Vermeire S*, et al.* Fecal calprotectin is a surrogate marker for endoscopic lesions in inflammatory bowel disease. *Inflamm Bowel Dis* 2012;**18**:2218-24.

59. Frivolt K, Schwerd T, Werkstetter KJ*, et al.* Repeated exclusive enteral nutrition in the treatment of paediatric crohn's disease: Predictors of efficacy and outcome. *Aliment Pharmacol Ther* 2014;**39**:1398-407.

60. Afzal NA, Davies S, Paintin M*, et al.* Colonic crohn's disease in children does not respond well to treatment with enteral nutrition if the ileum is not involved. *Dig Dis Sci* 2005;**50**:1471-5.

61. Tjellström B, Högberg L, Stenhammar L*, et al.* Effect of exclusive enteral nutrition on gut microflora function in children with crohn's disease. *Scand J Gastroenterol* 2012;**47**:1454-9.

62. Eckburg PB, Bik EM, Bernstein CN, *et al.* Diversity of the human intestinal microbial flora. *Science* 2005;**308**.
63. Hidenori H, Mitsuo S, Yoshimi B. Fecal microbial diversity in a strict vegetarian as determined by molecular analysis and cultivation. *Microbiology and Immunology* 2002;**46**:819-31.
64. Quraishi MN, Sergeant M, Kay G, *et al.* The gut-adherent microbiota of psc–ibd is distinct to that of ibd. *Gut* 2016.
65. Santoru ML, Piras C, Murgia A, *et al.* Cross sectional evaluation of the gut-microbiome metabolome axis in an italian cohort of ibd patients. *Sci Rep* 2017;**7**:9523.
66. Santoru ML, Piras C, Murgia A, *et al.* Author correction: Cross sectional evaluation of the gut-microbiome metabolome axis in an italian cohort of ibd patients. *Sci Rep* 2018;**8**:4993.
67. Louis P, Hold GL, Flint HJ. The gut microbiota, bacterial metabolites and colorectal cancer. *Nat Rev Microbiol* 2014;**12**:661.
68. Johnson Caroline H, Dejea Christine M, Edler D, *et al.* Metabolism links bacterial biofilms and colon carcinogenesis. *Cell Metabolism* 2015;**21**:891-7.
69. Koessler KK, Hanke MT, Sheppard MS. Production of histamine, tyramine, bronchospastic and arteriospastic substances in blood broth by pure cultures of microorganisms. *The Journal of Infectious Diseases* 1928;**43**:363-77.
70. van der Heiden C, Wadman SK, de Bree PK, Wauters EA. Increased urinary imidazolepropionic acid, n-acetylhistamine and other imidazole compounds in patients with intestinal disorders. *Clin Chim Acta* 1972;**39**:201-14.
71. Brown DG, Rao S, Weir TL, *et al.* Metabolomics and metabolic pathway networks from human colorectal cancers, adjacent mucosa, and stool. *Cancer & Metabolism* 2016;**4**:11.
72. Nicholson JK, Holmes E, Kinross J, *et al.* Host-gut microbiota metabolic interactions. *Science* 2012;**336**:1262-7.
73. Bone E, Tamm A, Hill M. The production of urinary phenols by gut bacteria and their possible role in the causation of large bowel cancer. *Am J Clin Nutr* 1976;**29**:1448-54.
74. Smith EA, Macfarlane GT. Enumeration of human colonic bacteria producing phenolic and indolic compounds: Effects of ph, carbohydrate availability and retention time on dissimilatory aromatic amino acid metabolism. *J Appl Bacteriol* 1996;**81**:288-302.
75. De Preter VVTHGSJDVLRPVK. Effects of lactobacillus casei shirota, bifidobacterium breve, and oligofructose-enriched inulin on colonic nitrogen-protein metabolism in healthy humans. *Am J Physiol Gastrointest Liver Physiol* 2007;**292**:G358-G68.
76. Koppe L, Pillon NJ, Vella RE, *et al.* P-cresyl sulfate promotes insulin resistance associated with ckd. *Journal of the American Society of Nephrology* 2013;**24**:88.
77. Altieri L, Neri C, Sacco R, *et al.* Urinary p-cresol is elevated in small children with severe autism spectrum disorder. *Biomarkers* 2011;**16**:252-60.
78. Selmer T, Andrei PI. P-hydroxyphenylacetate decarboxylase from clostridium difficile. *Eur J Biochem* 2001;**268**:1363-72.
79. Williams HRT, Cox IJ, Walker DG, *et al.* Characterization of inflammatory bowel disease with urinary metabolic profiling. *Am J Gastroenterol* 2009;**104**:1435-44.
80. Sokol H, Pigneur B, Watterlot L, *et al.* Faecalibacterium prausnitzii is an anti-inflammatory commensal bacterium identified by gut microbiota analysis of crohn disease patients. *Proceedings of the National Academy of Sciences of the United States of America* 2008;**105**:16731-6.
81. Lopetuso LR, Scaldaferri F, Petito V, Gasbarrini A. Commensal clostridia: Leading players in the maintenance of gut homeostasis. *Gut Pathogens* 2013;**5**:23.

82. Canani RB, Costanzo MD, Leone L*, et al.* Potential beneficial effects of butyrate in intestinal and extraintestinal diseases. *World J Gastroenterol* 2011;**17**:1519-28.

83. Thibault R, Blachier F, Darcy-Vrillon B*, et al.* Butyrate utilization by the colonic mucosa in inflammatory bowel diseases: A transport deficiency. *Inflamm Bowel Dis* 2010;**16**:684-95.

84. Kovarik JJ, Tillinger W, Hofer J*, et al.* Impaired anti-inflammatory efficacy of n-butyrate in patients with ibd. *European Journal of Clinical Investigation* 2011;**41**:291-8.

85. Kuhara T, Matsumoto I, Ohno M, Ohura T. Identification and quantification of octanoyl glucuronide in the urine of children who ingested medium-chain triglycerides. *Biomedical & Environmental Mass Spectrometry* 1986;**13**:595-8.

86. Moolenaar SH, Poggi-Bach J, Engelke UFH*, et al.* Defect in dimethylglycine dehydrogenase, a new inborn error of metabolism: Nmr spectroscopy study. *Clin Chem* 1999;**45**:459.

87. Dragsted LO. Biomarkers of meat intake and the application of nutrigenomics. *Meat Science* 2010;**84**:301-7.

88. Maiuolo J, Oppedisano F, Gratteri S, Muscoli C, Mollace V. Regulation of uric acid metabolism and excretion. *International Journal of Cardiology* 2016;**213**:8-14.

89. Servillo L, Giovane A, Cautela D, Castaldo D, Balestrieri ML. Where does nε-trimethyllysine for the carnitine biosynthesis in mammals come from? *PLOS ONE* 2014;**9**:e84589.

90. Steiber A, Kerner J, Hoppel CL. Carnitine: A nutritional, biosynthetic, and functional perspective. *Molecular Aspects of Medicine* 2004;**25**:455-73.

91. Franco R, Schoneveld OJ, Pappa A, Panayiotidis MI. The central role of glutathione in the pathophysiology of human diseases. *Arch Physiol Biochem* 2007;**113**:234-58.

92. Forman HJ, Zhang H, Rinna A. Glutathione: Overview of its protective roles, measurement, and biosynthesis. *Molecular Aspects of Medicine* 2009;**30**:1-12.

93. PERSAUD C, McDERMOTT J, BENOIST B, JACKSON AA. The excretion of 5-oxoproline in urine, as an index of glycine status, during normal pregnancy. *BJOG: An International Journal of Obstetrics & Gynaecology* 1989;**96**:440-4.

94. Lenton C, Ali Z, Persaud C, Jackson AA. Infants in trinidad excrete more 5-l-oxoproline (l-pyroglutamic acid) in urine than infants in england: An environmental not ethnic difference. *British Journal of Nutrition* 2007;**80**:51-5.

95. Moura FA, de Andrade KQ, dos Santos JCF, Araújo ORP, Goulart MOF. Antioxidant therapy for treatment of inflammatory bowel disease: Does it work? *Redox Biology* 2015;**6**:617-39.

96. Orešič M, Simell S, Sysi-Aho M*, et al.* Dysregulation of lipid and amino acid metabolism precedes islet autoimmunity in children who later progress to type 1 diabetes. *The Journal of Experimental Medicine* 2008;**205**:2975.

97. Kominsky DJ, Keely S, MacManus CF*, et al.* An endogenously anti-inflammatory role for methylation in mucosal inflammation identified through metabolite profiling. *The Journal of Immunology* 2011.

98. Sillanaukee, Pönniö, Jääskeläinen. Occurrence of sialic acids in healthy humans and different disorders. *European Journal of Clinical Investigation* 1999;**29**:413-25.

99. Dhanisha SS, Guruvayoorappan C, Drishya S, Abeesh P. Mucins: Structural diversity, biosynthesis, its role in pathogenesis and as possible therapeutic targets. *Critical Reviews in Oncology / Hematology* 2018;**122**:98-122.

100. Pietzner M, Kaul A, Henning A-K, *et al.* Comprehensive metabolic profiling of chronic low-grade inflammation among generally healthy individuals. *BMC Medicine* 2017;**15**:210.

101. Maury CP, Teppo AM, Wegelius O. Relationship between urinary sialylated saccharides, serum amyloid a protein, and c-reactive protein in rheumatoid arthritis and systemic lupus erythematosus. *Annals of the Rheumatic Diseases* 1982;**41**:268-71.

102. Baba R, Yashiro K, Nagasako K, Obata H. Significance of serum sialic acid in patients with crohn's disease. *Gastroenterologia Japonica* 1992;**27**:604-10.

103. Crook MA, Tutt P, Simpson H, Pickup JC. Serum sialic acid and acute phase proteins in type 1 and type 2 diabetes mellitus. *Clin Chim Acta* 1993;**219**:131-8.

104. Gutsche B, Grun C, Scheutzow D, Herderich M. Tryptophan glycoconjugates in food and human urine. *Biochem J* 1999;**343**:11-9.

105. Aleksandra F, Jan H. Protein c-mannosylation: Facts and questions. *Acta Biochim Pol* 2000;**47**:781-9.

106. Hamming OJ, Kang L, Svensson A, *et al.* Crystal structure of interleukin-21 receptor (il-21r) bound to il-21 reveals that sugar chain interacting with wsxws motif is integral part of il-21r. *J Biol Chem* 2012;**287**:9454-60.

107. Muroi E, Manabe S, Ikezaki M, *et al.* C-mannosylated peptides derived from the thrombospondin type 1 repeat enhance lipopolysaccharide-induced signaling in macrophage-like raw264.7 cells. *Glycobiology* 2007;**17**:1015-28.

108. TOMINAGA F, OKA K, YOSHIDA H. The isolation and identification of o-xylosyl-serine and s-methylcysteine sulfoxide from human urine. *J Biochem* 1965;**57**:717-20.

109. Takeuchi H, Fern, xe, *et al.* Rumi functions as both a protein o-glucosyltransferase and a protein o-xylosyltransferase. *Proceedings of the National Academy of Sciences of the United States of America* 2011;**108**:16600-5.

110. Bourdon MA, Krusius T, Campbell S, Schwartz NB, Ruoslahti E. Identification and synthesis of a recognition signal for the attachment of glycosaminoglycans to proteins. *Proceedings of the National Academy of Sciences of the United States of America* 1987;**84**:3194-8.

111. Otsuka Y, Sato T. Saccharide primers comprising xylosyl-serine primed phosphorylated oligosaccharides act as intermediates in glycosaminoglycan biosynthesis. *ACS Omega* 2017;**2**:3110-22.

112. Aquino RS, Park PW. Glycosaminoglycans and infection. *Frontiers in bioscience (Landmark edition)* 2016;**21**:1260-77.

113. Kornel L, Saito Z. Studies on steroid conjugates—viii: Isolation and characterization of glucuronide-conjugated metabolites of cortisol in human urine. *Journal of Steroid Biochemistry* 1975;**6**:1267-84.

114. Gray CH, Green MA, Holness NJ, Lunnon JB. Urinary metabolic products of prednisone and prednisolone. *J Endocrinol* 1956;**14**:146-54.

115. Lewis JD. The utility of biomarkers in the diagnosis and therapy of inflammatory bowel disease. *Gastroenterology* 2011;**140**:1817-26.e2.

5.7. Supplemental Information

Table S5.1 Time points of samples available from each pediatric CD patient included in the study.

EEN arm		
Patient #	Stool	Urine
6	1, 4, 5	1, 5
9	1, 2, 4, 6	1, 2, 4, 6
10	1, 4	1, 2, 4
13	1, 2, 3, 6	1, 3, 6
16	1, 2, 4	1, 2, 4
19	1, 3	1, 3
22	1, 2, 4, 6	1, 2, 4, 6
26	1, 3, 4, 5, 6	1, 2, 3, 4, 5, 6
5	NA	1, 2, 4
7	NA	1, 5
11	NA	1, 2, 4
Total sample number	27	35
Corticosteroid arm		
12	1, 2, 3, 4, 5	1, 2, 3, 4, 5
21	1, 2, 7	1, 2, 7
27	1, 2, 4, 7	1, 2, 4

* Time points represent; 1= Day 0, 2= Day 2-3, 3 = Day 7, 4 = Week 2, 5 = Week 4, 6 = Week 6, 7 = Week 8

Table S5.2 Number of subjects and their samples included in statistical analyses.

Number of subjects	Biospecimen	Number of samples	Statistical method
11	urine	27	PCA
8	stool	35	PCA
10	urine	32	Spearman's correlation
7	stool	24	Spearman's correlation
10	urine	20	Paired Wilcoxon test
7	stool	14	Paired Wilcoxon test
7	stool	14	Microbial differential abundance analysis

Only subjects who responded to the EEN therapy and provided baseline (pre-treatment) sample, as well as sample collected at 2-week or later time point were included.
Samples of one colonic CD patient were included only in the exploratory data analysis, such as PCA.
The cohort of patients was characterized by mean age of 13 year-old, 73 % male, 64 % treatment naïve, and 82 % of patients were prescribed with EEN for the first time. Inflammatory state was confirmed by average CRP of 46 mg/L and FCP of 2364 μg/g.

Table S5.3 Patients demographics and clinical information of CD patients who were prescribed for corticosteroid therapy and who did not respond to EEN therapy. Clinical measures are shown in mean and standard deviation. All measurements were taken from serum except for FCP.

Number of patients	3	
Age	12.5 ± 1	
Sex; male : female	1:2	
First time diagnosed (n)	3	
Disease location (n)		
Ileocolonic	2	
Colonic + UGI	2	
Maintenance medication (n)		
Biologic[a]	1	
Immunomodulator[b]	1	
5-ASA[c]	1	
Supplements (n)		
Folic acid	1	
	baseline	**Post 2-6 weeks**
CRP (mg/L)	34.4 ± 22	NA
FCP (µg/g)	3309.3 ± 1723	1142.0 ± 184
Quality of Life score	33.5 ± 3	38.6 ± 3
PCDAI	0	0
Patient #17 (non-responder to EEN)		
Age	15	
Sex	Female	
Disease location	ileocolonic	
Maintenance medication	Immunomodulator[b]	
Supplements	Iron, calcium, vitamin D	
Final PCDAI	15	

Abbreviations: CRP: C-reactive protein; ESR: Erythrocyte sedimentation rate; FCP: fecal calprotectin; UGI: Upper Gastrointestinal tract; NA: Data not available
[a]Infliximab; [b]Methotrexate; [c]Pentasa

Table S5.4 Spearman rank correlation between fecal metabolites and inflammatory markers.

Chemical ID	rho (CRP)	*p*-value	rho (FCP)	*p*-value
Serine	0.44	0.067	0.59	0.007
Valine	0.62	0.006	0.33	0.170
Betaine	0.60	0.009	0.24	0.329
Threonine	0.06	0.810	0.52	0.021
Isoleucine	0.51	0.031	0.28	0.254
Methylhistidine	0.61	0.007	0.51	0.026
Acetyl.aspartate	0.76	0.000	0.58	0.010
Trimethyllysine	0.47	0.050	0.53	0.021
Tryptophan	0.29	0.236	0.53	0.021
N1.N12.Diacetylspermine	0.74	0.000	0.60	0.006
Choline	0.55	0.018	0.66	0.002
Glutamine	0.50	0.033	0.70	0.001
Histidine	0.40	0.103	0.55	0.014
Carnitine	0.27	0.273	0.51	0.025
Phenylalanine	0.55	0.019	0.42	0.074
Tyrosine	0.48	0.045	0.59	0.008
N-Acetylneuraminic acid	0.39	0.113	0.66	0.002

Table S5.5. Spearman rank correlation between urinary metabolites and inflammatory markers.

Chemical ID	rho (CRP)	*p*-value	rho (FCP)	*p*-value
Acetylcarnitine	0.08	> 0.05	0.52	2.11E-02
Xylosyl-serine	0.65	4.47E-04	0.25	> 0.05
Mannopyranosyl-Trptophan	0.58	2.23E-03	0.26	> 0.05
oxo-proline	0.55	4.25E-03	0.22	> 0.05
p-Cresol sulfate	-0.56	3.45E-03	-0.35	> 0.05
p-Cresol glucuronide	-0.63	6.70E-04	-0.27	> 0.05
Dehydro-deoxy-N-acetylneuraminic acid	0.55	4.65E-03	0.17	> 0.05
N-Acetylneuraminic acid	0.67	2.64E-04	0.21	> 0.05
Octanoyl glucuronide	-0.51	9.36E-03	-0.18	> 0.05
Hydroxyandrosterone glucuronide	0.54	4.97E-03	0.45	5.37E-02
Cortolone glucuronide	0.63	7.78E-04	0.55	1.53E-02
3'-Sialyllactose	0.74	2.74E-05	0.30	> 0.05
Sialyl-N-acetyllactosamine	0.55	4.54E-03	0.26	> 0.05

Table S5.6. Heteroscedastic paired Wilcoxon rank sum test results applied to urinary metabolites between baseline and measurement taken after two weeks or longer from the start of the therapy. FDR-adjusted *p*-value > 0.05 is considered as significant.

m/z	Chemical ID	*p*-value	FDR	FC	Effect size
138.0550	Trigonelline	1.95E-03	0.040	0.26	0.63
238.0916	Xylosyl-serine[a]	1.95E-03	0.040	0.26	0.63
262.1028	Aspartyl-glutamine[a]	1.90E-03	0.040	4.1	0.63
167.0201	Uric acid	1.90E-03	0.040	0.64	0.63
187.0071	p-Cresol sulfate	1.90E-03	0.040	3.8	0.63
104.0706	Dimethylglycine	3.90E-03	0.040	2.1	0.60
189.1598	Trimethyllysine	3.90E-03	0.040	0.62	0.60
128.0353	Oxo-proline	3.90E-03	0.040	0.80	0.60
283.0823	p-Cresol glucuronide[a]	3.90E-03	0.040	3.2	0.60
319.1400	Octanoyl glucuronide[a]	3.90E-03	0.040	13.2	0.60
541.2649	Cortolone glucuronide[a]	3.90E-03	0.040	0.69	0.60
258.1084	Methylcytidine[b]	5.92E-03	0.045	0.73	0.58
182.0459	Pyridoxate	5.92E-03	0.045	4.3	0.58
218.1034	Pantothenic acid	5.92E-03	0.045	3.3	0.58
632.2044	3'-Sialyllactose	9.81E-03	0.055	0.63	0.56
673.2309	Sialylacetyllactosamine	9.81E-03	0.055	0.76	0.56
367.1500	Mannopyranosyl-tryptophan	0.0195	0.083	0.57	0.51
290.0882	2,3-Dehydrodeoxy-N-acetylneuraminic acid	0.0273	0.093	0.71	0.49
308.0987	N-Acetylneuraminic acid	0.0488	0.130	0.79	0.44

FC: Median fold change (> two weeks / baseline); FDR: False discovery rate
[a] Putative assignment based on tandem MS experiments.
[b] Total methylcytidine

Table S5.7. Heteroscedastic paired Wilcoxon rank sum test results of stool metabolites when samples between baseline and after 2 weeks or more from the start of the therapy were compared.

m/z	Chemical ID	*p*-value	FDR	FC	Effect size
141.0659	Imidazole propionate[a]	0.0156	> 0.05	0.47	0.63
152.0567	Guanine	0.0156	> 0.05	0.50	0.63
104.1069	Choline	0.0156	> 0.05	0.40	0.62
170.0924	Methylhistidine	0.0469	> 0.05	0.26	0.53
298.0970	Methylthioadenosine[a]	0.0469	> 0.05	0.48	0.50
308.0987	N-Acetylneuraminic acid	0.0469	> 0.05	0.48	0.59

*FC: Median fold change (> 2 weeks / baseline)*
[a] Putative assignment based on tandem MS experiments.

Table S5.8 Constituents of formula based on information provided by the manufacturer.

Formula components	Detected metabolites
Water	ND
Maltodextrin	ND
Enzymatically Hydrolyzed Whey Protein (from Milk)	Amino acids, Peptide
Medium Chain Triglycerides (from Coconut and/or Palm Kernel Oil)	Octanoyl glucuronide
less than 2% of Cornstarch	ND
Soybean Oil	ND
Soy Lecithin	ND
Magnesium Chloride	ND
Sodium Ascorbate	ND
Sodium Phosphate	ND
Calcium Phosphate	ND
Guar Gum	ND
Calcium Citrate	ND
Choline Chloride	Choline, Dimethylglycine
Potassium Chloride	ND
Salt	ND
Sodium Citrate	ND
Taurine	ND
L-Carnitine	Carnitine, Trimethyllysine
Magnesium Oxide	ND
Alpha-Tocopheryl Acetate	ND
Zinc Sulfate	ND
Ferrous Sulfate	ND
Niacinamide	ND
Calcium Pantothenate	Pantothenic acid
Vitamin A Palmitate	ND
Potassium Citrate	ND
Manganese Sulfate	ND
Pyridoxine Hydrochloride	Pyridoxic acid
Vitamin D3	ND
Copper Sulfate	ND
Thiamine Mononitrate	ND
Riboflavin	ND
Beta Carotene	ND
Folic Acid	ND
Biotin	ND
Citric Acid	ND
Potassium Iodide	ND
Chromium Chloride	ND
Sodium Selenate	ND
Sodium Molybdate	ND
Phytonadione	ND
Vitamin B12.	ND

ND: Metabolite/nutrient not detected when using MSI-CE-MS

Figure S5.1 Unknown compound identification process, including formula generation based on accurate mass, reproducible signal on electrophoregram and MS/MS spectra annotation. Proposed chemical identities are; **(a)** Aspartyl-glutamine in urine, **(b)** octanoyl glucuronide in urine, **(c)** N1, N2-diacetylspermine in stool, **(d)** p-Cresol glucuronide in urine, **(e)** Xylosyl-serine in ruine, **(f)** Hydroxyandrosterone glucuronide in urine, **(g)** Cortolone glucuronide in urine, **(h)** Imidazole propionate in urine, **(i)** Prednisolone in urine, **(j)** Methylthioadenosine in stool. Imidazole propionate identification process included spike experiment with an authentic chemical standard of another candidate, methyl-imidazole acetate (isomer), which did not comigrate with the unknown peak. Based on the frequent detection of imidazole propionate in human urine based on HMDB entry, the tentative assignment of this identification was made. MetFrag and HMDB were used for all compounds as reference databases when interpreting MS/MS spectra.
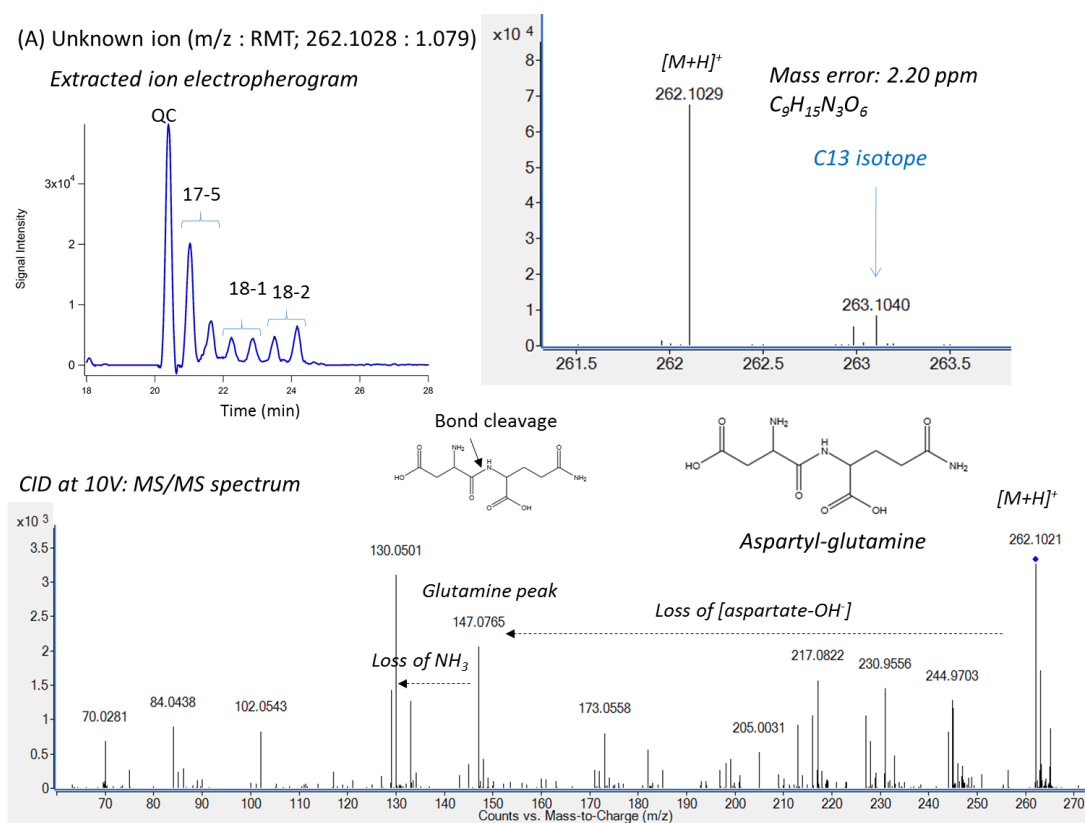
(B) Unknown (m/z : RMT; 319.14 : 0.782)

Extracted ion electropherogram

High resolution MS spectrum
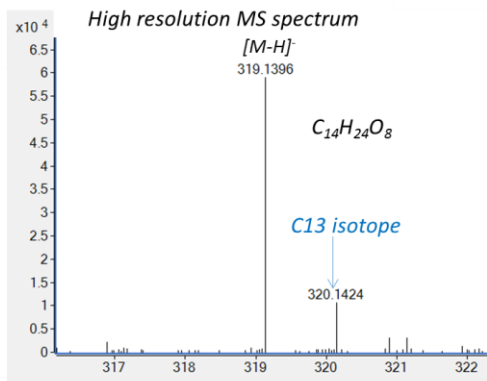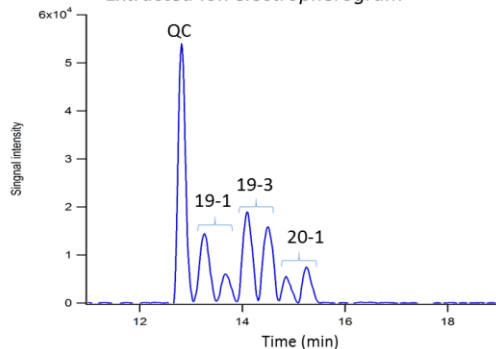
CID at 10V: MS/MS spectrum
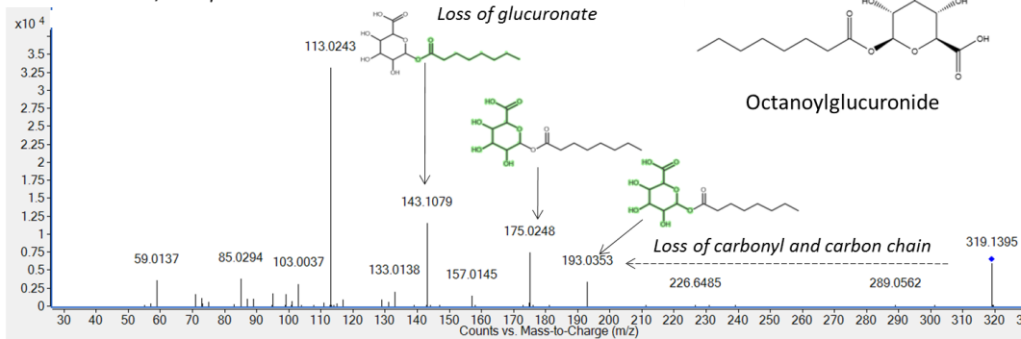
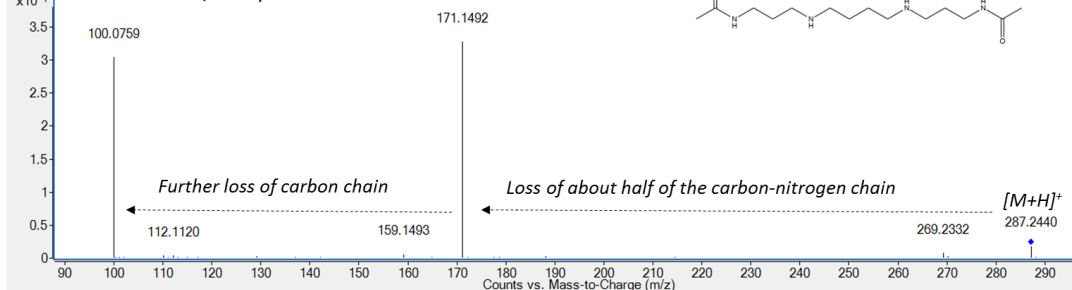(C) Unknown ion (m/z : RMT; 287.2447 : 0.643) in stool

Extracted ion electropherogram
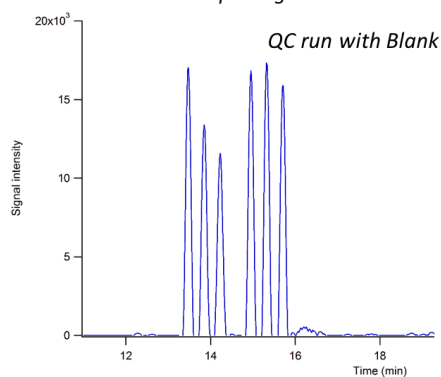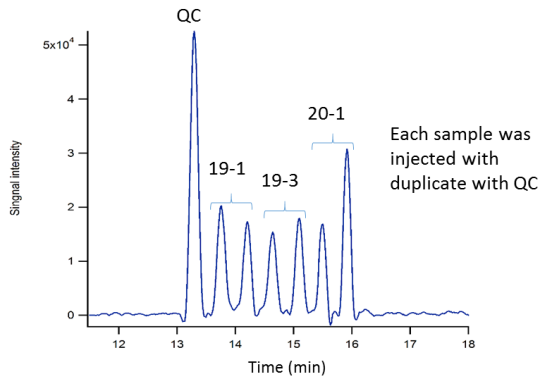
High resolution MS spectrum
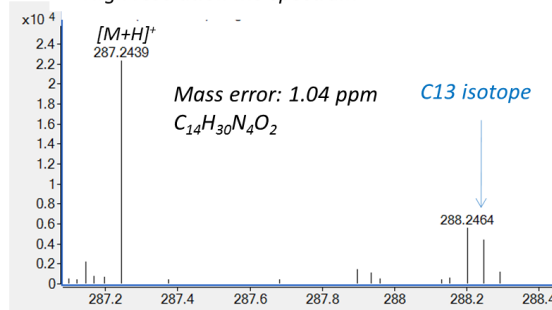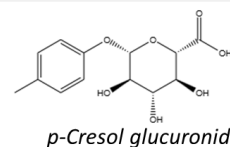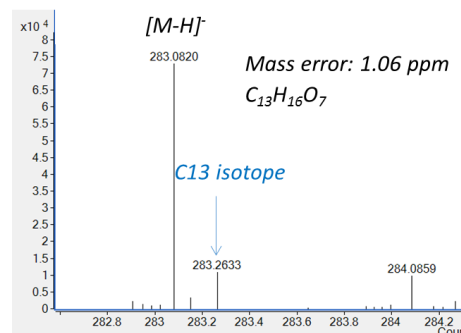
CID at 20V: MS/MS spectrum
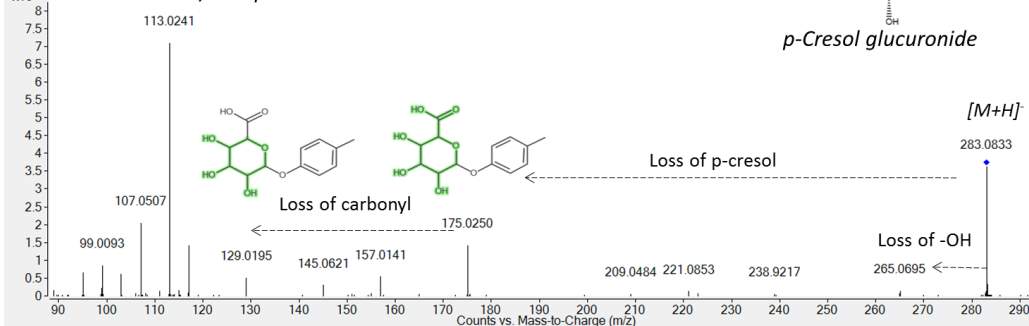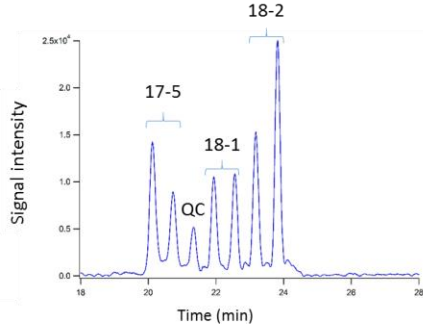
(D) Unknown (m/z : RMT; 283.082 : 0.809))

Extracted ion electropherogram

High resolution MS spectrum

$[M-H]^-$

Mass error: 1.06 ppm
$C_{13}H_{16}O_7$

C13 isotope

p-Cresol glucuronide

CID at 10V: MS/MS spectrum

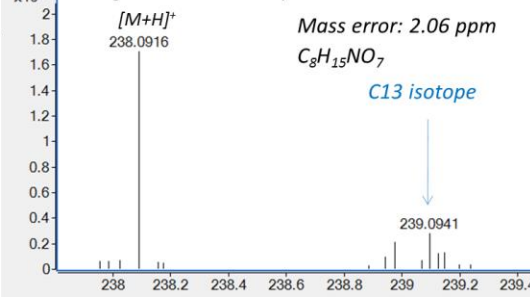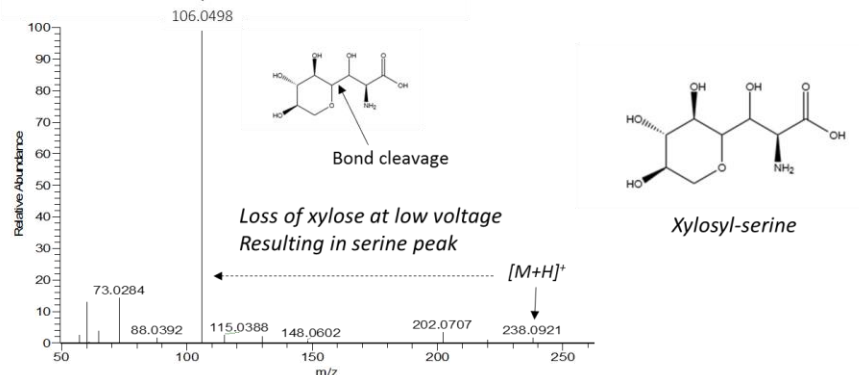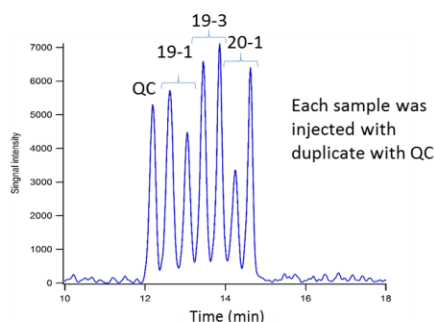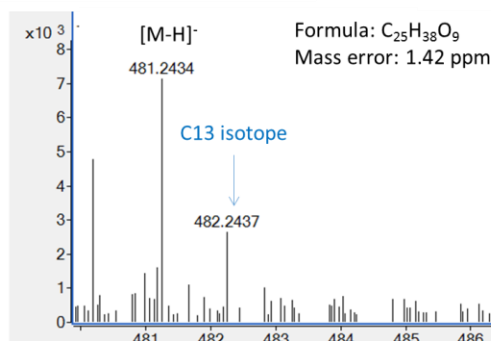$[M+H]^-$

Loss of p-cresol

Loss of carbonyl

Loss of -OH

(E) Unknown ion(m/z : RMT; 238.0916 : 1.064)

Extracted ion electropherogram

High resolution MS spectrum

$[M+H]^+$

Mass error: 2.06 ppm
$C_8H_{15}NO_7$

C13 isotope

CID at 10V: MS/MS spectrum

Bond cleavage

*Loss of xylose at low voltage
Resulting in serine peak*

Xylosyl-serine

$[M+H]^+$

262

(F) Unknown (m/z : RMT; 481.244 : 0.742)
Extracted ion electropherogram

High resolution MS spectrum

Formula: $C_{25}H_{38}O_9$
Mass error: 1.42 ppm

Each sample was injected with duplicate with QC

CID at 40V: MS/MS spectrum

11-beta-hydroxyandrosterone-3-glucuronide

Fragmentation of glucuronate

(G) Unknown (m/z : RMT; 541.264 : 0.729)
Extracted ion electropherogram

High resolution MS spectrum

$C_{27}H_{42}O_{11}$
Mass error: 5.55 ppm

CID at 40V: MS/MS spectrum

Cortolone-3-glucuronide

Loss of cortolone + H_2O

Glucuronic acid

Loss of [cortolone- H_2O]

263
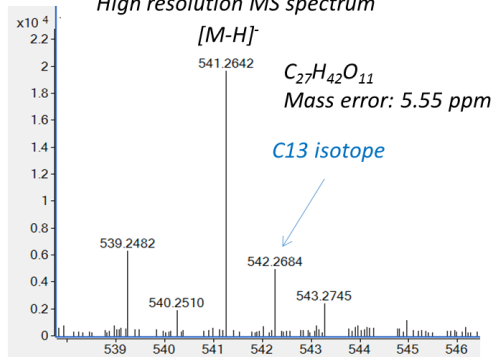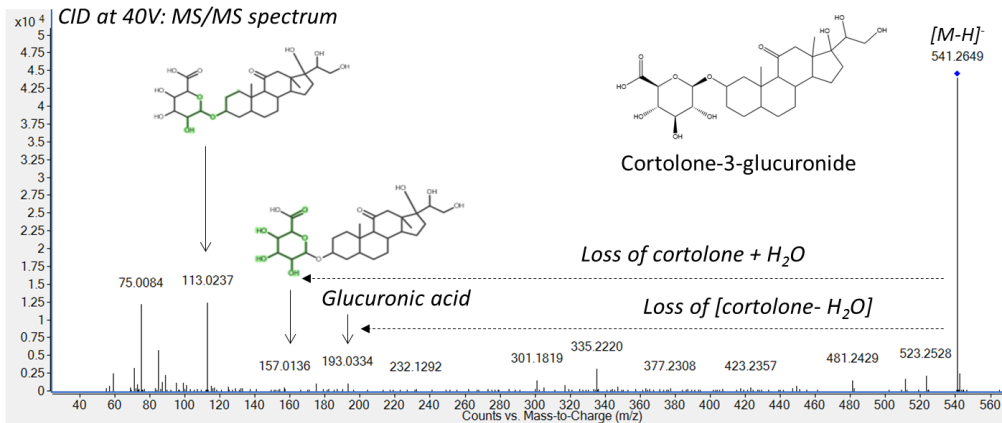
(H) Unknown (m/z : RMT; 141.0657 : 0.734) in stool

Extracted Ion Electropherogram Overlay
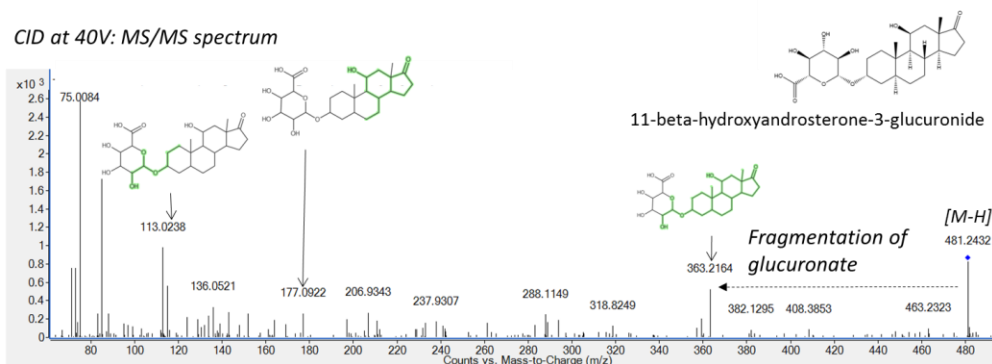
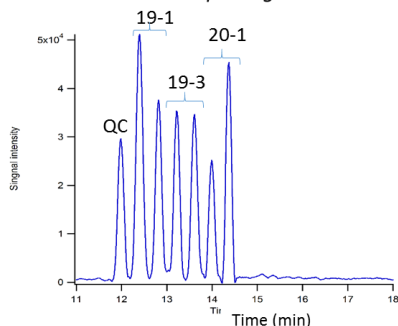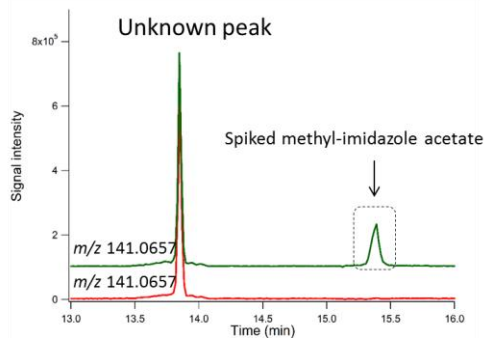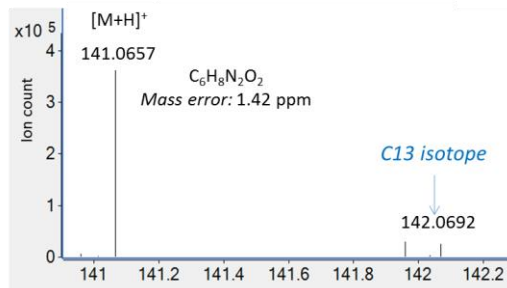High Resolution MS Spectrum

CID at 20V: MS/MS spectrum
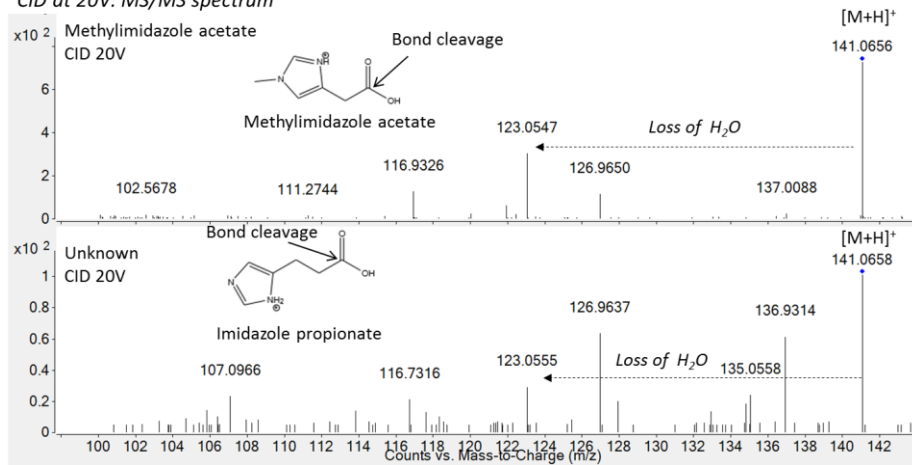
(I) Unknown (m/z : RMT; 359.1857 : 0.595)
Extracted ion electropherogram

QC run with Blank
in the middle
Migrate with EOF

High Resolution MS Spectrum

[M-H]⁻
359.1855

Formula: $C_{21}H_{28}O_5$
Mass error: 9.36 ppm

C13 isotope

360.1894

361.2010

CID at 20V: MS/MS spectrum

Prednisolone

[M-H]⁻
359.1855

125.0605

217.0863

Loss of dimethyl
and H_2O

89.0241

310.1204

59.0131

199.0767 Loss of three carbon rings

343.1541

149.0602 179.0546

Counts vs. Mass-to-Charge (m/z)
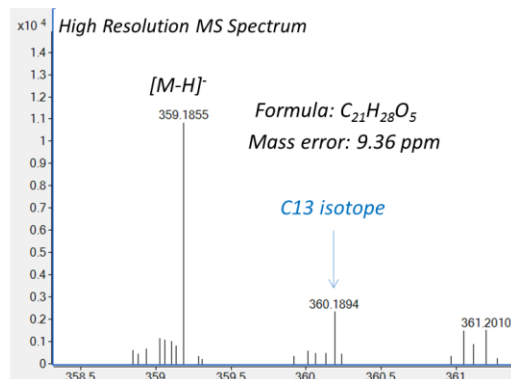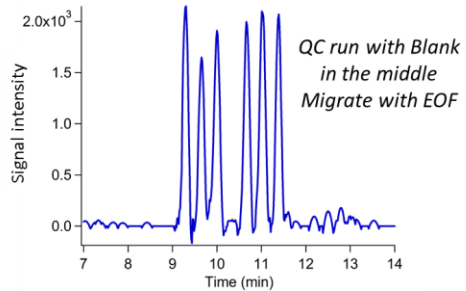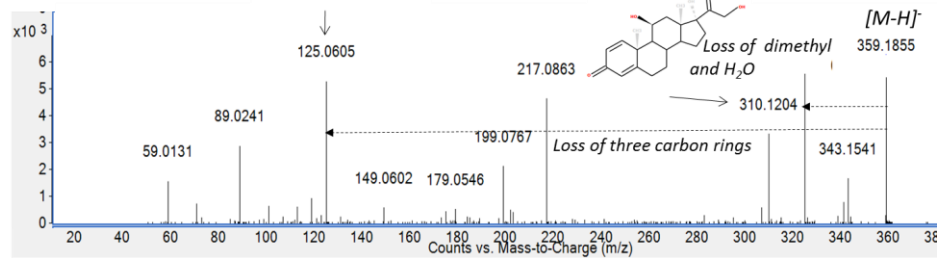
(J) Unknown (m/z : RMT; 298.0970 : 0.0.617)

Extracted ion electropherogram

QC run with Blank in the middle

High Resolution MS Spectrum

$[M-H]^+$ 298.0969

Formula: $C_{11}H_{15}N_5O_3S$

Mass error: 0.671 ppm

C13 isotope

299.0999

S+2 isotope

300.0949

CID at 20V: MS/MS spectrum

Adenine peak
136.0616

$[M-H]^+$
298.0963

Loss of thio-ribose

163.0414

Methylthioadenosine

Counts vs. Mass-to-Charge (m/z)
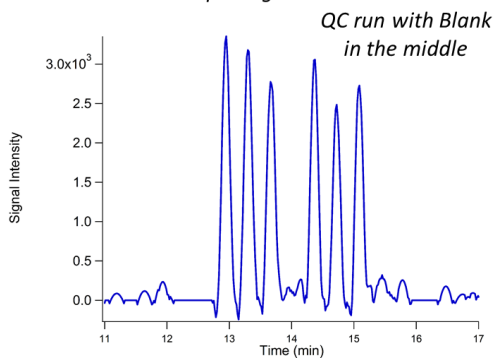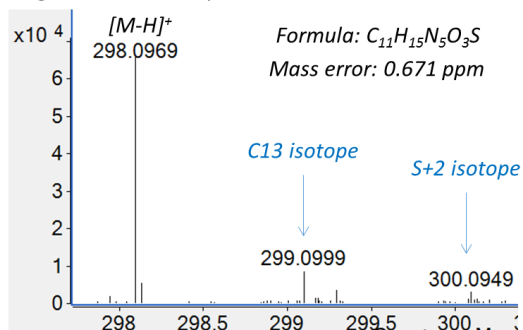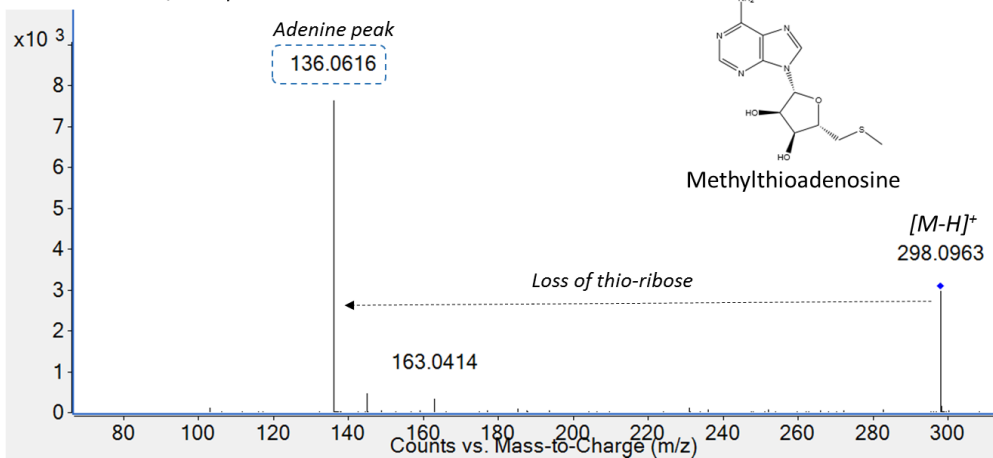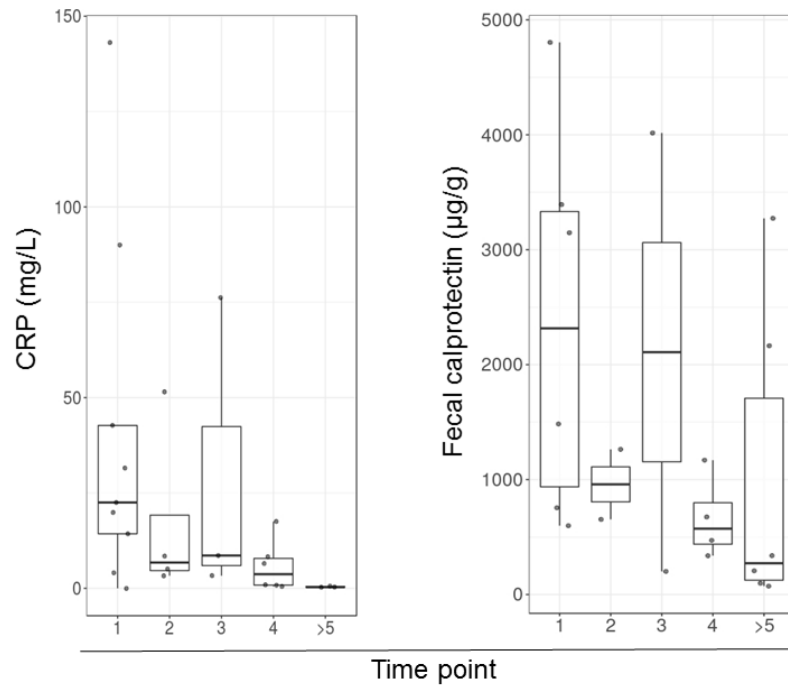
Figure S5.2 C-reactive protein and fecal calprotectin measured from 11 CD patients during EEN induction therapy.
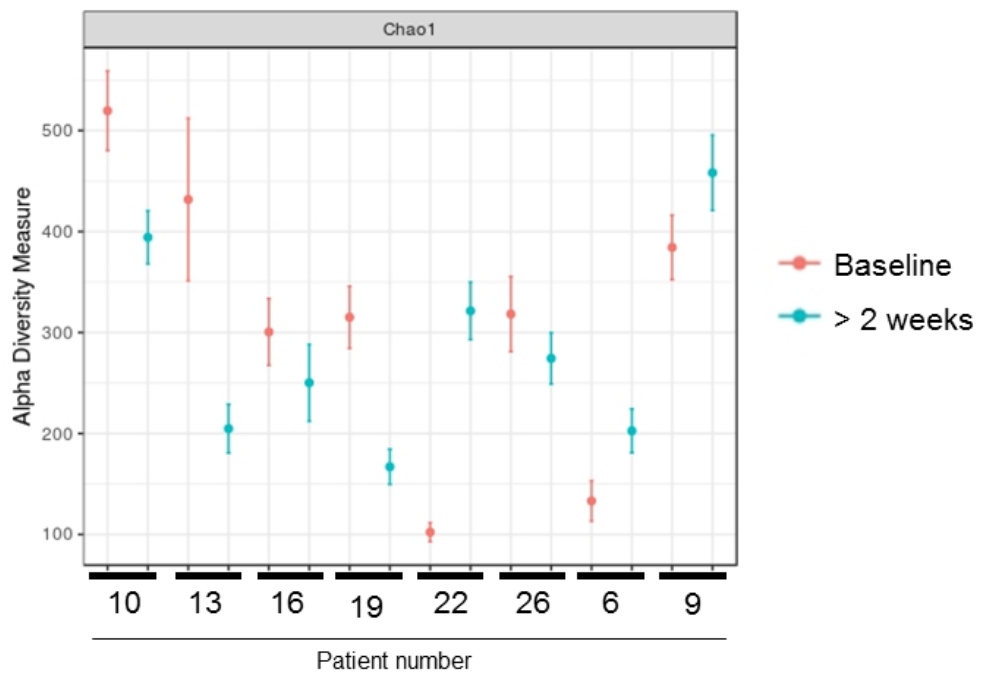
Figure S5.3 Chao1 alpha diversity measures of fecal microbiota of CD patients at baseline and after two weeks from the start of the therapy.
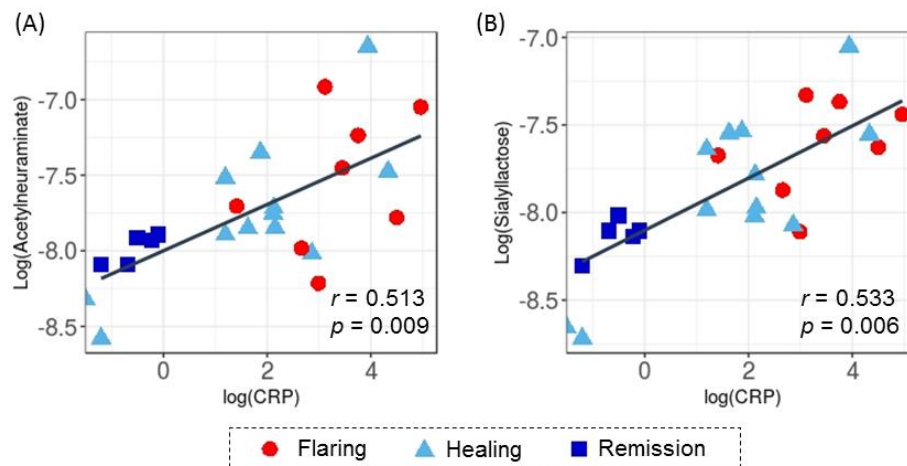
Figure S5.4 Scatter plots of **(a)** N-acetylneuraminate and CRP, and **(b)** 3'-sialyllactose and CRP with each sample point colour-coded with inflammatory state assignment. Correlation coefficients and p-values are from Pearson correlation. Both metabolites and CRP values were log-transformed.
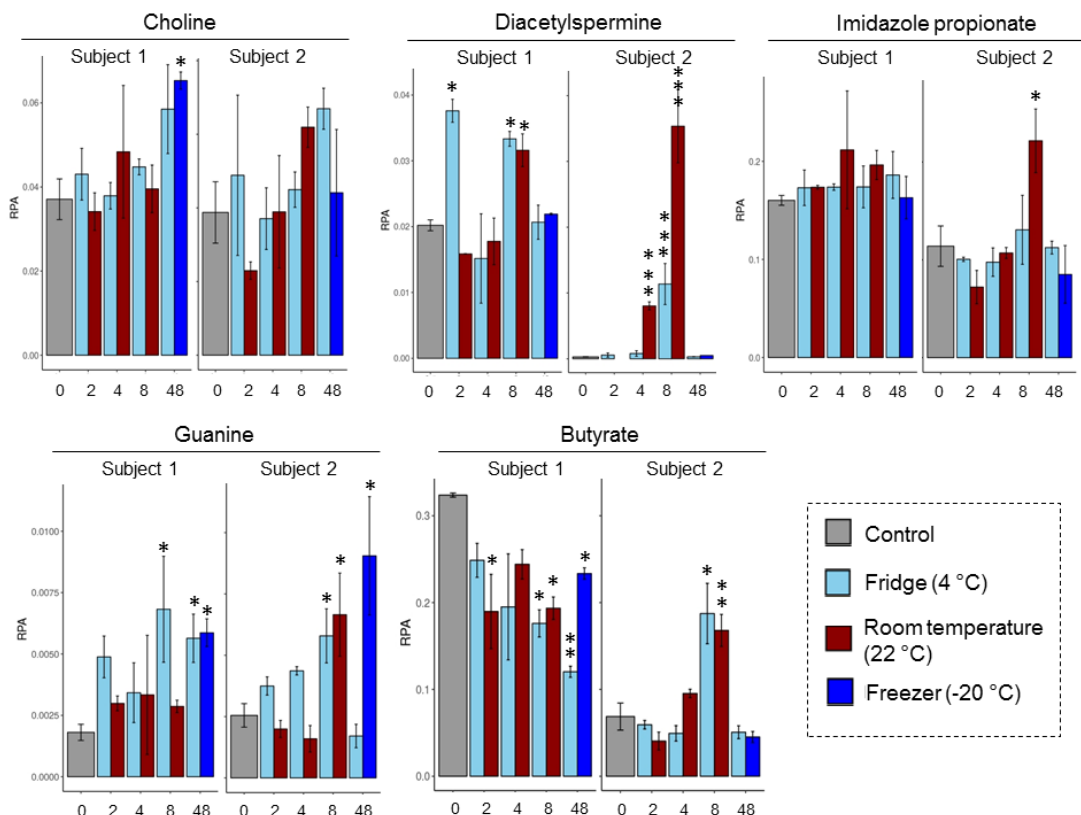
Figure S5.5 Representative plots for selected metabolites in stool samples stored at different conditions. Each x-axis shows a duration of storage in hours and y-axis shows relative peak area. *, ** and *** indicate differences between control and each condition at $p < 0.05$, $p < 0.01$, and $p < 0.001$ respectively with two-tailed homoscedastic t-test.
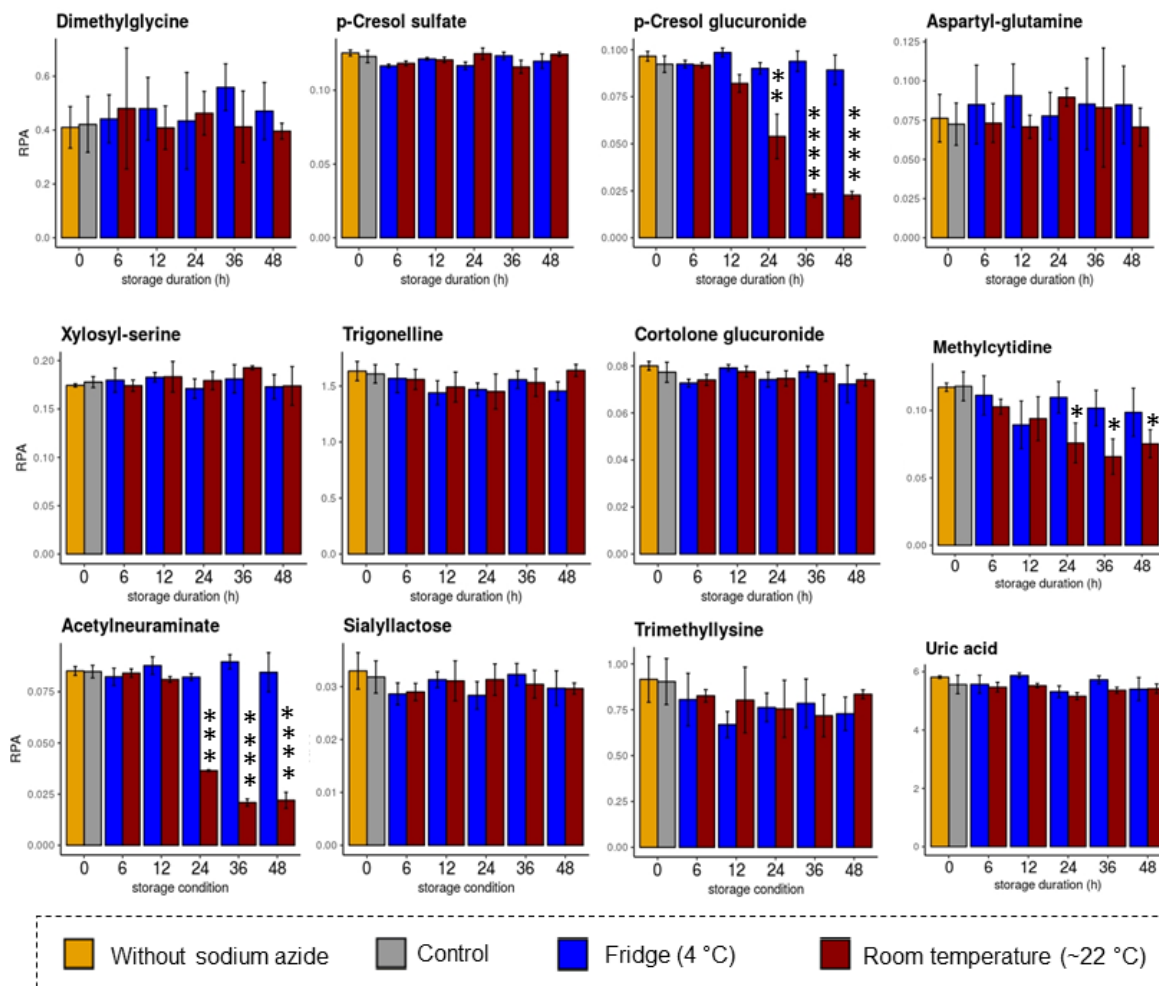
Figure S5.6 Representative plots of relative signal of urinary metabolite levels when samples were stored at different conditions. *, **, *** and **** indicate differences between control and each condition at p < 0.05, p < 0.01, p < 0.001, and p < 0.0001 respectively. Two-tailed homoscedastic t-test was used. Y-axis represents relative peak area, and x-axis shows storage duration in hour. Stability profile of Man-Trp was nearly identical to that of xylosyl-serine.

Chapter VI

Future Directions in Metabolomics of Chronic Digestive Disorders and Impacts on Decision Making in the Clinic

Chapter VI: Future Directions in Metabolomics of Chronic Digestive Disorders and Impacts on Decision Making in the Clinic

6.1 Overview of major thesis contributions

The work presented in this thesis contributed four major research innovations relevant to new advances towards better understanding the etiology of irritable bowel syndrome (IBS) and inflammatory bowel disease (IBD) that also have promising clinical applications for improved screening, diagnosis and/or treatment monitoring, including: (1) development of robust and high throughput method for analyzing anionic metabolites in human urine by multisegment injection-capillary electrophoresis-mass spectrometry (MSI-CE-MS), which included the identification of aminolysis as a major cause of current instability and capillary breakages when using ammonia-based buffer systems. This technical advancement is needed to enable reliable analysis of acidic metabolites that may serve as new biomarkers for differentiation of IBS and IBD using urine as a convenient specimen for clinical testing; (2) a comprehensive characterization of the urine metabolome of IBS patients, which involved unambiguous identification of several novel urinary metabolites implicated in accelerated collagen degradation and immune-activation processes in IBS as compared to healthy controls; this work represented the first report of aberrant hydroxylysine metabolism and collagen degradation in IBS that is likely associated with underlying gut dysfunction, including poor motility; (3) untargeted characterization of the metabolome from matching stool extracts and urine samples from a cohort of pediatric IBD patients together with corresponding microbiome functional metagenomics analysis that provided new insights into the underlying pathophysiology of Crohn's disease (CD) as compared to ulcerative colitis (UC). Noteworthy, this work outlined the discovery and structural elucidation of a single and ratiometric metabolites that allow for differential diagnosis of major IBD subtypes not feasible by conventional inflammatory biomarkers based on non-invasive analysis of urine samples; and (4) a longitudinal study of dynamic metabolic trajectories from matching urine and stool specimens urinary of pediatric CD patients following exclusive enteral nutrition (EEN) therapy, which also incorporated microbiome metagenomics analysis to better understand the mechanisms of EEN action. Both IBS and

IBD are chronic digestive disorders with unknown etiology. Differentiating these disorders from each other, as well as from other chronic gastrointestinal (GI) conditions, such as coeliac disease and chronic constipation, is challenging and time-consuming due to lack of reliable biomarkers that can accurately diagnose these conditions in a non-invasive manner. As a result, current clinical practice relies on standardized questionnaires, blood and stool tests and various invasive examinations, including invasive colon endoscopy and tissue histopathology. Improved understanding of the pathogenesis of IBS and IBD will lead to development of specific biomarkers for improved screening and early detection of high risk populations, including differential diagnosis, prognosis, and treatment monitoring, which is crucial for improving the quality of life for patients suffering from these debilitating chronic disorders. In this context, untargeted metabolomics offers a hypothesis generating approach to identify new biomarkers and mechanistic insights that are often overlooked in targeted analysis. Despite their high global prevalence, comprehensive metabolic phenotyping of IBS and IBD have not been explored thoroughly to date. Additionally, longitudinal analysis of stool and urinary metabolome of pediatric CD patients during EEN therapy holds potential to guiding clinical decision making by identifying a subset of patients who do not respond to standard therapies and thus require alternative interventions.

Chapter I of this thesis provided comprehensive review of metabolomics as related to IBS and IBD, including their prevalence and clinical history, current diagnostic procedures and treatment strategies, and the roles of metabolomics when integrated with other -omics especially microbial metagenomics. The review also included current metabolomics research in IBS and IBD, and highlighted major knowledge gaps and diagnostic dilemmas to be resolved in future research of these chronic digestive disorders. *Chapter II* of this thesis introduced a robust and high throughput method for analysing anionic metabolites in urine when using multiplexed separations based on multisegment injection-capillary electrophoresis-mass spectrometry (MSI-CE-MS) when using a polyimide coated fused-silica capillary in ammonia-based alkaline buffers.[1] Through extended exposure of the capillary to different buffer compositions with variable amount of ammonium hydroxide and pH, aminolysis of the polyimide outer coating of the capillary

in CE was reported for the first time as a major cause of frequently observed current instability and capillary breakage in CE-MS metabolomic studies to date. Following the use of optimum buffer pH conditions that prevents aminolysis, separation performance of MSI-CE-MS was evaluated when analyzing various acidic metabolites in pooled 24 h urine samples. The optimized method with ammonium bicarbonate buffer was validated by analyzing freshly thawed pooled urine sample each day for three consecutive days, which resulted in 238 sample injections with good long-term precision and robustness for a wide range of acidic urinary metabolites of clinical significance, including drug metabolites (*i.e.* acetaminophen sulfate), vitamins (*i.e.* pantothenate), and host-microbial co-metabolites (*i.e.* indoxyl sulfate and *p*-cresol sulfate) with coefficient of variation (CV) ranging from 9.3 to 28% over consecutive three days of analysis using different capillaries and buffer conditions.

Chapter III described an untargeted characterization of the ionic metabolome from repeat urine samples collected form IBS patients as compared to healthy controls when using a metabolomics data workflow MSI-CE-MS with stringent quality control to reduce false discoveries. This study was the first work to comprehensively analyze a diverse range of urinary metabolites from IBS patients, including several unknown metabolites associated with IBS status as compared to healthy controls reported here for the first time. This pilot study identified significantly higher excretion of several biomarker candidates following osmolality normalization that were associated with IBS patients after adjustments for age and false discovery rate (FDR), including mannopyranosyl-tryptophan (Man-Trp), glucosylgalactosyl-hydroxylysine (Glc-Gal-OHLys), galactosyl-hydroxylysine (Gal-OHLys), dimethylguaonosine, and a few primary (*i.e.* serine, lysine, ornithine, glutamine) and modified amino acid metabolites (*i.e.* imidazole propionate, dimethylglycine). High excretion of these urinary metabolites provided evidence of low-grade inflammation in the pathophysiology of IBS as characterized by increased immune-activation, collagen degradation and proteolysis. Clinically, IBS is defined as a functional gut disorder with comorbid anxiety and/or depression, but without morphological/structural changes to GI tract that can be measured by conventional tests, including inflammation

273

based on serum C-reactive protein or fecal calprotectin.[2] However, animal studies,[3, 4] and observations in IBD cases[5-7] consistently suggested the presence of low-grade immune activation and potential mucosal damage as an underlying etiology of IBS symptoms. Additionally, IBS patients have been identified to have a 10-fold higher risk to be later diagnosed with IBD when compared to subjects without IBS.[8] Taken together, the boundary between IBS and inflammatory disease seems vague for at least subset of IBS patients, and our results in this thesis support this notion. Further validation is required to confirm our findings but once validated, urinary biomarkers of immune-activation and collagenous tissue damage will offer a convenient and non-invasive method to more accurately screen or diagnose IBS, and potentially assess disease severity and optimize treatment responses to novel dietary interventions to relieve underlying symptoms.

Chapter IV introduced comprehensive characterization of the metabolome of lyophilized stool extracts when using MSI-CE-MS with temporal pattern recognition, which was also combined with urinary metabolome and stool microbiome metagenomics data collected from a cohort of pediatric IBD patients. Both UC and CD poses equally significant negative impacts on affected children in terms of poor nutrition, delayed onset of puberty, growth impairment and overall quality of life; however, treatment decisions differ significantly between UC and CD,[9] which makes an initial accurate diagnosis essential for successful therapeutic interventions. Currently, physicians primarily depend on endoscopic and histological examinations to differentiate these two major IBD subtypes, which can still result in inconclusive diagnoses (*i.e.* IBD-U; unclassified) especially in pediatric IBD populations.[10, 11] Therefore, better mechanistic insights and diagnostic tests for differentiating CD and UC in affected children is urgently needed in a clinical setting. In this context, this pilot study contributed by identifying a ratiometric urinary biomarker (*i.e.* serine/indoxyl sulfate) that had excellent specificity (95%) and sensitivity (100%) for distinguishing CD from UC in largely treatment naïve and recently diagnosed IBD patients. Additionally, metagenomic studies revealed a higher prevalence of opportunistic pathogens in the stool of pediatric UC patients, whereas phenolic and indole metabolites involved in gut microflora metabolism, were significantly elevated in the urine of CD patients,

indicating different bacterial communities and functions as contributing factors in the etiology of UC and CD. This study was also unique by including investigation of urinary and stool metabolite stability as a result of delayed storage under different conditions to simulate actual sample collection challenges in a clinical setting while confirming potential biomarker robustness for routine clinical testing. There was a large biological variability in measured metabolite concentrations between-subjects in stool as expected from the high metabolic capacity of bacterial content and heterogeneity of stool specimens even after normalization to total dried mass following lyophilization. In contrast, lower biological variance and better chemical stability was found in most of clinically significant urinary metabolites when samples were stored in a refrigerator for up to 48 h following the addition of sodium azide as an antimicrobial preservative. In summary, this study proposed novel pathological mechanisms of UC and CD, which involves mucosal degradation and concurrent increase in bacterial pathogens in the colon of pediatric UC patients, and higher microbial activity for production of phenols and indoles that are likely associated with oxidative stress in affected children with CD.

Chapter V described a time-resolved longitudinal metabolomic study of changes in urinary and stool derived metabolite trajectories that were associated with clinical remission of the majority of pediatric CD patients following EEN therapy. EEN therapy is an effective induction therapy for CD, which has shown equivalent or superior rate in inducing clinical remission and mucosal healing as compared to corticosteroid without adverse side effects related to growth and development in children.[12, 13] Thus, EEN therapy is currently considered as a first-line therapy for pediatric CD in Europe, North America, Japan and Britain;[14-16] however, underlying mechanisms related to the efficacy of EEN remains poorly understood. This pilot study included nontargeted analysis of polar/ionic metabolites in matching urine and stool samples of largely treatment naïve pediatric IBD patients over a period of 8-week EEN therapy, which identified distinctive metabolic trajectories associated with symptom improvement and mucosal healing. For instance, a positive correlation was found between primary amino acids and other nutritional

components of the EEN formula and two classic inflammatory biomarkers (*i.e.* serum C-reactive protein, CRP and fecal calprotectin, FCP) indicating that improved nutrient absorption via intestinal epithelia coincided with lower tissue/systemic inflammation. This effect was likely mediated by a number of homeostatic metabolisms, including energy metabolism and cellular protection against oxidative stress as reflected in time-dependent reduction in trimethyllysine and oxo-proline excretion in urine. Additionally, reduced immune-activation was correlated with lower excretion of glycosylated amino acids (*i.e.* Man-Trp and unknown putatively identified as xylosyl-serine) and sialic acid metabolites in urine. These urinary metabolites serve as functionally important components of glycoproteins mediating biomolecular interactions in the immune system, including CRP.[17-22] An unexpected finding of this study was the paradoxical increase in urinary excretion of *p*-cresol conjugates (*i.e.* sulfate and glucuronide) during the therapy, which have been reported as microbially derived metabolite and uremic toxin in chronic kidney disease,[23] as well as metabolites with strong positive correlation with colonic transit times.[24] Symptom improvement shortly after initiation of EEN therapy includes a reduction in diarrhea, which increases colonic transit time, and thus, significant increase in *p*-cresol conjugate formation was likely a reflection of symptomatic improvement. In terms of stool microbiome, there was a significant reduction in opportunistic pathogen, *Haemophilus* and *Veillonella* following intervention, whereas overall alpha-diversity did not change. Interestingly, when metabolic trajectories of key urinary metabolites were compared with that of one EEN non-responder and three corticosteroid responders, similar trends were observed among positive responders regardless of the treatment, whereas the one non-responder showed opposite trends among the most significant treatment responsive biomarker candidates. Taken together, dynamic urinary and stool metabolomics confirmed improved nutrition absorption and gut microbial modulation as key initiators of symptom improvement following EEN therapy, which subsequently reduced immune activation and inflammation. After validation with larger number of patients, lead urinary biomarker candidates identified in this study for the first time (*e.g.,* Man-Trp) can be applied to improve clinical decision making in terms of predicting treatment prognosis of IBD patients

as required for continuing therapeutic interventions (or switching to a different therapy for non-responders) at an early stage of induction therapy as an alternative to more invasive colonoscopy and tissue histopathology.

6.2 Validation Study on Collagen Degradation behind IBS Pathology

The novel finding of high collagen degradation products excretion in urine by IBS patients has shed new light onto the unknown pathological mechanism of IBS. To validate this finding, follow-up study should follow multiple steps involving the analysis of collagen content in intestinal mucosa and enzyme expression responsible for collagen degradation using an animal model of IBS. The most direct validation of the link between IBS symptoms and potential collagen degradation in intestinal tissue is an investigation of full-thickness biopsy at multiple sites along the small intestine and colon of IBS patients, and measurement of collagen content in those samples in comparison to samples collected from healthy controls. However, tissue biopsy at multiple sites, especially from healthy individuals, would pose challenges in ethical approval processes due to the invasiveness of the procedure. In this context, animal models play an essential role, and rat and mouse IBS models have been utilized in multiple studies, such as the investigation of post-infectious IBS[4] and visceral hypersensitivity.[25] Quantification of collagen is commonly performed histologically using picrosirus red, a staining agent that binds to a collagen molecule and appears red in light microscopy.[26, 27] This technique can be used to measure collagen molecules actually present in intestinal tissue of rat IBS model and normal rat as a negative control.

Additionally, enzyme expression associated with the collagen degradation should be measured to gain further insights into key biochemical pathways involved in the loss of tissue integrity and subsequent functional loss. Degradation of collagen is generally tightly controlled in a healthy state because of the ubiquitous presence of this protein and its crucial roles in keeping the integrity of extracellular matrix. Therefore, loss of this control has been implicated in various disorders, such as pulmonary fibrosis,[28] peridontal disease,[29] and diabetes.[30] There are different types of proteases involved in the degradation process,

however, the major types belong to the matrix metalloproteinase (MMP) family, which is a family of Ca^{2+} and Zn^{2+} dependent proteins.[31] Among over 20 members of MMP,[32] MMP-1 preferentially targets type III collagen, and MMP-13 and -8 selectively targets type II and type I collagen, respectively, while majority of other MMPs target different proteins in extracellular matrix.[33] This variety in substrate specificity makes it necessary to target specific MMPs when measuring their expression levels using enzyme-linked immunosorbent assay (ELISA) as each antibody specifically bind to its substrate. Although, there has not been a study on MMPs associated with IBS, expression of MMP-13 and -9 were found up-regulated in association with anastomotic leakage, one of complications following a colorectal surgery.[34, 35] Beside cases resulting from inexperience of surgeons, there are always some cases of post-operational leakage following a colorectal surgery, presumably as a result of certain defects in intestinal wound healing associated with collagen metabolism as indicated from several animal-based studies.[36, 37] When comparing 10 subjects with anastomotic leakage and 14 subjects who healed without the complication, Stumpf *et al.*[34] measured significantly lower amount of collagen in colonic wall of patients with the complication. These patients also showed presence of MMP-13 in their colonic tissue, while it was missing in more than half of subjects without the complication. Additionally, Shogan *et al.*[35] recently reported the association between accelerated collagen degradation and *Enterococcus faecalis*-induced activation of MMP-9 as a contributing factor in delayed wound healing and resulting anastomotic leakage. Based on these reports, MMP-13 and -9 are the most interesting targets to start with for the validation study with IBS animal model to investigate the extent of enzyme expression involved in collagen degradation.

Finally, urinary excretion of glycosylated hydroxylysine should be measured in addition to the quantification of collagen and collagenase expression in intestinal tissue to validate the connection between the aberrant collagen metabolism and increased excretion of these key metabolites found in our pilot study. Once validated with the animal model, investigation of the enzyme expression in colonic biopsy of actual IBS patients should follow as it holds a potential as a therapeutic target to prevent the loss of intestinal function.

6.3 Follow-up Study on Urinary Metabolome of CD and UC Patients with Controls

One of major limitations of the study in *Chapter IV*, which involved a comparison of the urine and stool metabolomes of CD and UC patients was a lack of control group. This limited us from evaluating the concentrations of metabolites as "high" or "low" in comparison to what would be observed in healthy age/sex-match population (*e.g.,* reference ranges), as well as other related but different GI disorders. In order to investigate the reproducibility of our findings and overall specificity of lead biomarker candidates, follow-up urine samples will be collected from pediatric IBD patients and pediatric patients with chronic constipation as clinically relevant controls, including independent cohorts of both UC and CD patients from multiple centres to increase study power and ensure replication of findings. Since stability of stool metabolome using current collection method with a simple sterile collection cup was found to be unreliable when samples are kept at ambient or refrigerated temperatures for 2 or more hours, this follow-up study will primarily focus on urine samples, which is easier to collect and stable under refrigeration for up to 48 h as described in *Chapter IV* and *V*. Chronic constipation is commonly seen among children, with prevalence ranging from 0.7 % - 29.6 % worldwide.[38, 39] Chronic constipation can be categorized into three different types; 1) "functional" constipation, in which patients perceive constipation despite the normal stool frequency, 2) disorders in defecation due to dysfunction in pelvic floor or anal sphincter, and 3) slow transit constipation due to alterations in neurons and certain types of cells in GI tract that are responsible for intestinal motility.[40] Regardless of these sub-types, chronic constipation is not an immune-mediated disorder,[41] and thus it can serve as an excellent negative control for IBD. Moreover, chronic constipation can be cured by elimination diet,[42] pharmaceutical agent (*e.g.* prucalopride)[43] and prebiotics/probiotics,[44] whereas IBD can only be managed but not completely cured. In addition, the vast majority of IBD cases involve diarrhea, whereas constipation has been reported in very rare occasions.[45] Therefore, urinary metabolites proposed to be associated with immune activation in this thesis (*i.e.* Man-Trp, xylosyl serine, sialic acids) are expected to be higher in both UC and CD patients, while compounds associated with

colonic transit time (*i.e. p*-cresol sulfate and glucuronide) are likely elevated in chronically constipated children as it has already been reported in studies involving autistic children with chronic constipation.[46] Since chronic constipation is distinctly different from IBD, differences in urinary metabolite observed between UC and CD can be compared with levels measured in constipated children to gauge the specificity of lead urinary biomarkers (*e.g.* indoxyl sulfate, kynurenine, serine) in association with specific symptoms and morphological manifestation of UC and CD.

6.4 Improved Stool Sample Collection and Storage for Metabolomics

Stool metabolome can provide the most direct information from metabolism in GI tract, including digestion of food, host-microbial interactions and associated immune responses. Stool is highly abundant in live microorganisms with associated high metabolic capacity to alter composition of stool metabolites.[47] Indeed, bacterial diversity and total counts have been reported to change significantly when human stool samples were stored at either room temperature or at 4 °C for 8 and 24 h, respectively,[48] which implies associated changes in measured metabolite levels that increases overall biological variance and likelihood for bias and false discoveries. Therefore, it was not surprising to see the high susceptibility of stool metabolites to delayed storage in metabolomic studies presented in *Chapter IV* and *V*, and elsewhere in the literature.[49-51] Unlike genetic sequencing of DNA or RNA, untargeted metabolomics studies typically involve analysis of hundreds of metabolites that have chemically diverse structures, properties and stabilities that require careful method optimization during sample collection/storage and sample workup. Therefore, it is challenging to optimize a storage method for global untargeted metabolomics notably for stool metabolites. Currently there is no standardized method for stool sample collection in metabolomics, but there are two studies evaluating the effects of various storage conditions and collection tools on metabolite integrity. Gratton *et al.*[49] evaluated room temperature, 4 °C (*i.e.* refrigerator) and -20 °C (*i.e.* freezer) storage for the duration of 1, 5, 10 and 24 h when analyzing crude stool extract and fecal water collected from five healthy adults using NMR. They also investigated the effect of freeze-thaw cycles by storing samples at 4 °C or

-20 °C followed by 5 h of thawing period at room temperature to mimic the transportation of samples between homes to a clinical laboratory. Based on multivariate statistical analysis (*e.g.* PCA, OPLS-DA), they reported relatively well-preserved metabolome integrity when samples were stored in a refrigerator as compared to room temperature and interestingly, freezer storage. Storing stool samples in a freezer has been recommended for microbiome analysis,[52] but freezing was found to significantly impact metabolite composition in this study. Moreover, metabolites in samples stored in a freezer for 24 h and subsequently thawed at room temperature showed greater deviation from control samples, which was immediately stored at -80 °C after being processed, whereas samples stored in a refrigerator for 24 h and stored at room temperature for the same duration of time clustered together tightly with control samples when using PCA. The same observation was found based on results outlined in this thesis, as well as another study pertaining to fecal fermentation.[53] Based on these findings, the recommended sample storage condition for stool specimens is in a refrigerator or cold pack immediately after sample collection with transfer of samples to a clinic on ice within 24 h. However, as encountered in the IBD study in this thesis, obtaining samples from children and their families collected at home within 24 h before a clinical visit may not be realistic for all samples. In order to mitigate this issue, different sample collection procedures are needed to halt post-defecation microbial activities while preserving stool metabolome integrity. Loftfield *et al.*[51] examined the performance of three sample collection methods, including addition of 95% *v/v* ethanol, as well as commercial fecal occult blood test (FOBT) kit, and fecal immunochemical test (FIT) tubes when comprehensively analyzing stool extracts collected from 18 healthy adults when using LC-MS. Both FOBT and FIT are routinely used in clinical settings to test for inclusion of blood in stool, which is an indication of colorectal cancer, IBD and potentially other immune-mediated GI disorders.[54] FOBT kit includes sample collection cards, on which a stool sample is smeared and attached flap is closed on top of the sample to minimize contamination. FIT tube is equipped with a stick to be used for scraping the surface of stool samples, which is subsequently mixed with a solution containing antibody that binds to hemoglobin in the tube. A total of 859 molecular features ranging from lipids, amino acids,

nucleosides and xenobiotics were detected from control stool samples, which were collected without any preservatives and immediately processed and stored at -80 °C. When three collection methods were compared to the control after 4 days of delayed storage at room temperature, the highest number of metabolites (508) were consistently detected when using 95% *v/v* ethanol as a preservative, while lower number of metabolites (435 and 220) were detected when using FOBT and FIT, respectively. However, the authors recommended to select collection methods depending on the class of metabolites of interest as FOBT showed high detectability of amino acids and carbohydrates, but lower detectability of vitamins, lipids, nucleosides, peptides and xenobiotics. Overall, FIT tubes detected fewer classes metabolites than other methods compared in this study. As a result, addition of absolute ethanol at the point of sample collection at home offers a simple, inexpensive, and convenient option for future metabolomic studies involving stool specimens. When stool samples contain highly variable amount of water due to diarrhea as observed in IBD studies in this thesis, lyophilisation process and subsequent normalization to total dried weight is necessary to accurately compare stool metabolite levels across samples. Alternatively, anaerobic stool collection kit may also be considered as an effective collection tool to preserve the microbiome and resulting metabolome composition as close as how they were at the point of defecation. The GI tract is primarily anaerobic, which is a natural and ideal environment for many of strictly anaerobic commensal bacteria to thrive.[55] Defecation and consequential exposure to ambient oxygen levels are expected to significantly alter the stool microbial composition post-defecation and resulting microbial metabolism. To mitigate this problem during sample collection while avoiding dilution of stool samples with organic solvents that may be incompatible with lyophilization, media-free anaerobic specimen collection tube equipped with a built-in oxygen-elimination system has been effectively utilized for collection of key anaerobic species involved in gut homeostasis,[56] and evaluation of antibiotic dosage.[57] As a more cost-effective option, a gas-impermeable (anaerobic) bag[58] is also an option in combination with conventional sterile collection cup, which has been applied for investigations of colon cancer prevalence and associated gut microflora,[59] effect of probiotics in preserving microflora after antibiotics

use,[60] and colonic bile acid composition changes associated with fecal microbiota transportation.[61] In summary, future work involving stool collection and metabolome analysis will include evaluation of absolute ethanol to preserve stool metabolism, as well as the implementation of an anaerobic sample collection kit, for more reliable and reproducible analysis of stool metabolome as required for new mechanistic insights into chronic digestive disorders.

6.5 Concluding Remarks

Overall, this thesis has contributed novel insights into IBS and IBD pathology and mechanism behind the efficacy of EEN therapy through untargeted urinary and stool metabolomics when using rigorous quality control with MSI-CE-MS in combination with stool microbiome metagenomics analysis for IBD studies. This thesis included the first comprehensive characterization of polar/ionic urinary metabolome of IBS patients when compared to healthy controls, as well as the first untargeted longitudinal analysis of urinary and stool metabolome of pediatric IBD patients during EEN therapy. These pilot studies consistently highlighted urinary signatures of immune-activation in both IBS and IBD patients. Additionally, physiological changes, such as collagenous tissue structure and colonic transit time, were also reflected onto levels of urinary metabolites. These key urinary metabolites identified in this thesis comprise initial step toward simple, reliable and more accurate diagnosis and personalized therapeutic intervention for increasing cases of IBS and IBD worldwide. Future studies will improve stool standardized collection procedures, expansion of lipid and fatty acid profiling by non-aqueous MSI-CE-MS methods, as well as metabolomic studies of fecal transplantation on pediatric IBD patients who are not responsive to conventional dietary or pharmacological therapies.

6.6 References

(1)   Yamamoto, M.; Ly, R.; Gill, B.; Zhu, Y.; Moran-Mirabal, J.; Britz-McKibbin, P., Robust and High-Throughput Method for Anionic Metabolite Profiling: Preventing Polyimide Aminolysis and Capillary Breakages under Alkaline Conditions in Capillary Electrophoresis-Mass Spectrometry. *Anal. Chem.* **2016**.

(2)   Lacy, B. E.; Mearin, F.; Chang, L.; Chey, W. D.; Lembo, A. J.; Simren, M.; Spiller, R., Bowel Disorders. *Gastroenterology* **2016,** *150* (6), 1393-1407.e5.

(3)   Bercik, P.; Verdu, E. F.; Collins, S. M., Is Irritable Bowel Syndrome a Low-Grade Inflammatory Bowel Disease? *Gastroenterol. Clin. North Am.* **2005,** *34* (2), 235-245.

(4)   Bercik, P.; Verdu, E. F.; Foster, J. A.; Macri, J.; Potter, M.; Huang, X.; Malinowski, P.; Jackson, W.; Blennerhassett, P.; Neufeld, K. A.; Lu, J.; Khan, W. I.; Corthesy–Theulaz, I.; Cherbut, C.; Bergonzelli, G. E.; Collins, S. M., Chronic Gastrointestinal Inflammation Induces Anxiety-Like Behavior and Alters Central Nervous System Biochemistry in Mice. *Gastroenterology* **2010,** *139* (6), 2102-2112.e1.

(5)   Spiller, R.; Lam, C., The shifting interface between IBS and IBD. *Curr. Opin. Pharmacol.* **2011,** *11* (6), 586-592.

(6)   Keohane, J.; O'Mahony, C.; O'Mahony, L.; O'Mahony, S.; Quigley, E. M.; Shanahan, F., Irritable Bowel Syndrome–Type Symptoms in Patients With Inflammatory Bowel Disease: A Real Association or Reflection of Occult Inflammation? *Am. J. Gastroenterol.* **2010,** *105*, 1789.

(7)   Forough, F.; K, M. J.; Brock, E.; Jan, I. E., Functional gastrointestinal disorders and mood disorders in patients with inactive inflammatory bowel disease: Prevalence and impact on health. *Inflamm. Bowel Dis.* **2006,** *12* (1), 38-46.

(8)   Porter, C. K.; Cash, B. D.; Pimentel, M.; Akinseye, A.; Riddle, M. S., Risk of inflammatory bowel disease following a diagnosis of irritable bowel syndrome. *BMC Gastroenterology* **2012,** *12* (1), 55.

(9)   Pithadia, A. B.; Jain, S., Treatment of inflammatory bowel disease (IBD). *Pharmacological Reports* **2011,** *63* (3), 629-642.

(10)   Levine, A.; de Bie, C. I.; Turner, D.; Cucchiara, S.; Sladek, M.; Murphy, M. S.; Escher, J. C., Atypical disease phenotypes in pediatric ulcerative colitis: 5-year analyses of the EUROKIDS Registry. *Inflamm. Bowel Dis.* **2012,** *19* (2), 370-377.

(11)   Gasparetto, M.; Guariso, G., Highlights in IBD Epidemiology and Its Natural History in the Paediatric Age. *Gastroenterology Research and Practice* **2013,** *2013*, 12.

(12)   Lafferty, L.; Tuohy, M.; Carey, A.; Sugrue, S.; Hurley, M.; Hussey, S., Outcomes of exclusive enteral nutrition in paediatric Crohn's disease. *Eur. J. Clin. Nutr.* **2016,** *71*, 185.

(13)   Soo, J.; Malik, B. A.; Turner, J. M.; Persad, R.; Wine, E.; Siminoski, K.; Huynh, H. Q., Use of Exclusive Enteral Nutrition Is Just as Effective as Corticosteroids in Newly Diagnosed Pediatric Crohn's Disease. *Dig. Dis. Sci.* **2013,** *58* (12), 3584-3591.

(14)   Lee, J.; Allen, R.; Ashley, S.; Becker, S.; Cummins, P.; Gbadamosi, A.; Gooding, O.; Huston, J.; Le Couteur, J.; O'Sullivan, D.; Wilson, S.; Lomer, M. C., British Dietetic Association evidence-based guidelines for the dietary management of Crohn's disease in adults. *J Hum Nutr Diet* **2014,** *27* (3), 207-18.

(15)   Ueno, F.; Matsui, T.; Matsumoto, T.; Matsuoka, K.; Watanabe, M.; Hibi, T., Evidence-based clinical practice guidelines for Crohn's disease, integrated with formal consensus of experts in Japan. *J Gastroenterol* **2013,** *48* (1), 31-72.

(16)   Ruemmele, F. M.; Veres, G.; Kolho, K. L.; Griffiths, A.; Levine, A.; Escher, J. C.; Amil Dias, J.; Barabino, A.; Braegger, C. P.; Bronsky, J.; Buderus, S.; Martín-de-Carpi, J.; De Ridder, L.; Fagerberg, U. L.; Hugot, J. P.; Kierkus, J.; Kolacek, S.; Koletzko, S.; Lionetti, P.; Miele, E.; Navas López, V. M.; Paerregaard, A.; Russell, R. K.; Serban, D. E.; Shaoul, R.; Van Rheenen, P.; Veereman, G.; Weiss, B.; Wilson, D.; Dignass, A.; Eliakim, A.; Winter, H.; Turner, D., Consensus guidelines of ECCO/ESPGHAN on the medical management of pediatric Crohn's disease. *J. Crohns Colitis* **2014,** *8* (10), 1179-1207.

(17)   Ihara, Y.; Manabe, S.; Kanda, M.; Kawano, H.; Nakayama, T.; Sekine, I.; Kondo, T.; Ito, Y., Increased expression of protein C-mannosylation in the aortic vessels of diabetic Zucker rats. *Glycobiology* **2005,** *15* (4), 383-392.

(18)   Ihara, Y.; Inai, Y.; Ikezaki, M.; Matsui, I.-S. L.; Manabe, S.; Ito, Y., C-Mannosylation: Modification on Tryptophan in Cellular Proteins. In *Glycoscience: Biology and Medicine*, Taniguchi, N.; Endo, T.; Hart, W. G.; Seeberger, H. P.; Wong, C.-H., Eds. Springer Japan: Tokyo, 2015; pp 1091-1099.

(19)   Takeuchi, H.; Fern; xe; ndez-Valdivia, R. C.; Caswell, D. S.; Nita-Lazar, A.; Rana, N. A.; Garner, T. P.; Weldeghiorghis, T. K.; Macnaughtan, M. A.; Jafar-Nejad, H.; Haltiwanger, R. S., Rumi functions as both a protein O-glucosyltransferase and a protein O-xylosyltransferase. *Proceedings of the National Academy of Sciences of the United States of America* **2011,** *108* (40), 16600-16605.

(20)   Otsuka, Y.; Sato, T., Saccharide Primers Comprising Xylosyl-Serine Primed Phosphorylated Oligosaccharides Act as Intermediates in Glycosaminoglycan Biosynthesis. *ACS Omega* **2017,** *2* (7), 3110-3122.

(21)   Maury, C. P.; Teppo, A. M.; Wegelius, O., Relationship between urinary sialylated saccharides, serum amyloid A protein, and C-reactive protein in rheumatoid arthritis and systemic lupus erythematosus. *Annals of the Rheumatic Diseases* **1982,** *41* (3), 268-271.

(22)   Pietzner, M.; Kaul, A.; Henning, A.-K.; Kastenmüller, G.; Artati, A.; Lerch, M. M.; Adamski, J.; Nauck, M.; Friedrich, N., Comprehensive metabolic profiling of chronic low-grade inflammation among generally healthy individuals. *BMC Medicine* **2017,** *15* (1), 210.

(23)   Koppe, L.; Pillon, N. J.; Vella, R. E.; Croze, M. L.; Pelletier, C. C.; Chambert, S.; Massy, Z.; Glorieux, G.; Vanholder, R.; Dugenet, Y.; Soula, H. A.; Fouque, D.; Soulage, C. O., p-Cresyl Sulfate Promotes Insulin Resistance Associated with CKD. *Journal of the American Society of Nephrology* **2013,** *24* (1), 88.

(24)   Roager, H. M.; B. S. Hansen, L.; I. Bahl, M.; Frandsen, H.; Carvalho, V.; Gøbel, R.; Dalgaard, M.; Plichta, D.; H. Sparholt, M.; Vestergaard, H.; Hansen, T.; Sicheritz-Ponten, T.; Nielsen, H.; Pedersen, O.; Lauritzen, L.; Kristensen, M.; Gupta, R.; Licht, T., Colonic transit time is related to bacterial metabolism and mucosal turnover in the gut. *Nat. Microbiol.* **2016,** *1*, 16093.

(25)   La, J. H.; Kim, T. W.; Sung, T. S.; Kim, H. J.; Kim, J. Y.; Yang, I. S., Role of mucosal mast cells in visceral hypersensitivity in a rat model of irritable bowel syndrome. *J Vet Sci* **2004,** *5* (4), 319-24.

(26)   Vogel, B.; Siebert, H.; Hofmann, U.; Frantz, S., Determination of collagen content within picrosirius red stained paraffin-embedded tissue sections using fluorescence microscopy. *MethodsX* **2015,** *2*, 124-134.

(27)   Nadkarni, S. K.; Pierce, M. C.; Park, B. H.; de Boer, J. F.; Whittaker, P.; Bouma, B. E.; Bressner, J. E.; Halpern, E.; Houser, S. L.; Tearney, G. J., Measurement of Collagen and Smooth Muscle Cell Content in Atherosclerotic Plaques Using Polarization-Sensitive Optical Coherence Tomography. *Journal of the American College of Cardiology* **2007,** *49* (13), 1474.

(28)   Jenkins, R. G.; Simpson, J. K.; Saini, G.; Bentley, J. H.; Russell, A.-M.; Braybrooke, R.; Molyneaux, P. L.; McKeever, T. M.; Wells, A. U.; Flynn, A.; Hubbard, R. B.; Leeming, D. J.; Marshall, R. P.; Karsdal, M. A.; Lukey, P. T.; Maher, T. M., Longitudinal change in collagen degradation biomarkers in idiopathic pulmonary fibrosis: an analysis from the prospective, multicentre PROFILE study. *The Lancet Respiratory Medicine* **2015,** *3* (6), 462-472.

(29)   Ram, V. S.; Parthiban; Sudhakar, U.; Mithradas, N.; Prabhakar, R., Bonebiomarkers in periodontal disease: a review article. *Journal of clinical and diagnostic research : JCDR* **2015,** *9* (1), ZE07-ZE10.

(30)   Chen, Y.-J.; Chan, D.-C.; Lan, K.-C.; Wang, C.-C.; Chen, C.-M.; Chao, S.-C.; Tsai, K.-S.; Yang, R.-S.; Liu, S.-H., PPARγ is involved in the hyperglycemia-induced inflammatory responses and collagen degradation in human chondrocytes and diabetic mouse cartilages. *Journal of Orthopaedic Research* **2015,** *33* (3), 373-381.

(31)   Jablonska-Trypuc, A.; Matejczyk, M.; Rosochacki, S., Matrix metalloproteinases (MMPs), the main extracellular matrix (ECM) enzymes in collagen degradation, as a target for anticancer drugs. *J Enzyme Inhib Med Chem* **2016,** *31* (sup1), 177-183.

(32)   Page-McCaw, A.; Ewald, A. J.; Werb, Z., Matrix metalloproteinases and the regulation of tissue remodelling. *Nature Reviews Molecular Cell Biology* **2007,** *8*, 221.

(33)   Lu, P.; Takai, K.; Weaver, V. M.; Werb, Z., Extracellular matrix degradation and remodeling in development and disease. *Cold Spring Harbor perspectives in biology* **2011,** *3* (12), 10.1101/cshperspect.a005058 a005058.

(34)   Stumpf, M.; Cao, W.; Klinge, U.; Klosterhalfen, B.; Kasperk, R.; Schumpelick, V., Collagen distribution and expression of matrix metalloproteinases 1 and 13 in patients with anastomotic leakage after large-bowel surgery. *Langenbeck's Archives of Surgery* **2002,** *386* (7), 502-506.

(35)   Shogan, B. D.; Belogortseva, N.; Luong, P. M.; Zaborin, A.; Lax, S.; Bethel, C.; Ward, M.; Muldoon, J. P.; Singer, M.; An, G.; Umanskiy, K.; Konda, V.; Shakhsheer, B.; Luo, J.; Klabbers, R.; Hancock, L. E.; Gilbert, J.; Zaborina, O.; Alverdy, J. C., Collagen degradation and MMP9 activation by Enterococcus faecalis contribute to intestinal anastomotic leak. *Sci. Transl. Med.* **2015,** *7* (286), 286ra68-286ra68.

(36)   Hesp, F. L. E. M.; Hendriks, T.; Lubbers, E.-J. C.; deBoer, H. H. M., Wound healing in the intestinal wall. *Diseases of the Colon & Rectum* **1984,** *27* (2), 99-104.

(37)   Jiborn, H.; Ahonen, J.; Zederfeldt, B., Healing of experimental colonic anastomoses: The effect of suture technic on collagen concentration in the colonic wall. *The American Journal of Surgery* **1978,** *135* (3), 333-340.

(38)   Mugie, S. M.; Benninga, M. A.; Di Lorenzo, C., Epidemiology of constipation in children and adults: A systematic review. *Best Practice & Research Clinical Gastroenterology* **2011,** *25* (1), 3-18.

(39)   Benninga, M. A.; Voskuijl, W. P.; Taminiau, J. A. J. M., Childhood Constipation: Is There New Light in The Tunnel? *Journal of Pediatric Gastroenterology and Nutrition* **2004,** *39* (5), 448-464.

(40)   Lembo, A.; Camilleri, M., Chronic Constipation. *N. Engl. J. Med.* **2003,** *349* (14), 1360-1368.

(41)   Camilleri, M.; Ford, A. C.; Mawe, G. M.; Dinning, P. G.; Rao, S. S.; Chey, W. D.; Simrén, M.; Lembo, A.; Young-Fadok, T. M.; Chang, L., Chronic constipation. *Nat. Rev. Dis. Primers* **2017,** *3*, 17095.

(42)   Carroccio, A.; Scalici, C.; Maresi, E.; Prima, L. D.; Cavataio, F.; Noto, D.; Porcasi, R.; Averna, M. R.; Iacono, G., Chronic constipation and food intolerance: A model of proctitis causing constipation. *Scand. J. Gastroenterol.* **2005,** *40* (1), 33-42.

(43)   Camilleri, M.; Van Outryve, M. J.; Beyens, G.; Kerstens, R.; Robinson, P.; Vandeplassche, L., Clinical trial: the efficacy of open-label prucalopride treatment in patients with chronic constipation – follow-up of patients from the pivotal studies. *Aliment. Pharmacol. Ther.* **2010,** *32* (9), 1113-1123.

(44)   Ford, A. C.; Quigley, E. M. M.; Lacy, B. E.; Lembo, A. J.; Saito, Y. A.; Schiller, L. R.; Soffer, E. E.; Spiegel, B. M. R.; Moayyedi, P., Efficacy of Prebiotics, Probiotics, and Synbiotics in Irritable Bowel Syndrome and Chronic Idiopathic Constipation: Systematic Review and Meta-analysis. *Am. J. Gastroenterol.* **2014,** *109*, 1547.

(45)   D'Aleo, C. M.; Sabbi, T.; Gismondi, P.; Scattoni, R.; Ferrara, P., Childhood Crohn's disease presenting as chronic constipation. *Pediatr Med Chir* **2009,** *31* (4), 168-71.

(46)   Gabriele, S.; Sacco, R.; Altieri, L.; Neri, C.; Urbani, A.; Bravaccio, C.; Riccio, M. P.; Iovene, M. R.; Bombace, F.; De Magistris, L.; Persico, A. M., Slow intestinal transit contributes to elevate urinary p-cresol level in Italian autistic children. *Autism Research* **2016,** *9* (7), 752-759.

(47)   Sender, R.; Fuchs, S.; Milo, R., Revised Estimates for the Number of Human and Bacteria Cells in the Body. *PLOS Biology* **2016,** *14* (8), e1002533.

(48)   Ott, S. J.; Musfeldt, M.; Timmis, K. N.; Hampe, J.; Wenderoth, D. F.; Schreiber, S., In vitro alterations of intestinal bacterial microbiota in fecal samples during storage. *Diagnostic Microbiology and Infectious Disease* **2004,** *50* (4), 237-245.

(49)   Gratton, J.; Phetcharaburanin, J.; Mullish, B. H.; Williams, H. R. T.; Thursz, M.; Nicholson, J. K.; Holmes, E.; Marchesi, J. R.; Li, J. V., Optimized Sample Handling Strategy for Metabolic Profiling of Human Feces. *Anal. Chem.* **2016,** *88* (9), 4661-4668.

(50)   Matysik, S.; Le Roy, C. I.; Liebisch, G.; Claus, S. P., Metabolomics of fecal samples: A practical consideration. *Trends in Food Science & Technology* **2016,** *57*, 244-255.

(51)   Loftfield, E.; Vogtmann, E.; Sampson, J. N.; Moore, S. C.; Nelson, H.; Knight, R.; Chia, N.; Sinha, R., Comparison of Collection Methods for Fecal Samples for Discovery Metabolomics in Epidemiologic Studies. *Cancer Epidemiology Biomarkers & Prevention* **2016,** *25* (11), 1483-1490.

(52)   Cardona, S.; Eck, A.; Cassellas, M.; Gallart, M.; Alastrue, C.; Dore, J.; Azpiroz, F.; Roca, J.; Guarner, F.; Manichanh, C., Storage conditions of intestinal microbiota matter in metagenomic analysis. *BMC Microbiology* **2012,** *12*, 158-158.

(53)   Bosch, G.; Wrigglesworth, D. J.; Cone, J. W.; Pellikaan, W. F.; Hendriks, W. H., Effects of preservation conditions of canine feces on in vitro gas production kinetics and fermentation end products1. *Journal of Animal Science* **2013,** *91* (1), 259-267.

(54)   Sanford, K. W.; McPherson, R. A., Fecal Occult Blood Testing. *Clinics in Laboratory Medicine* **2009,** *29* (3), 523-541.

(55)   Lozupone, C. A.; Stombaugh, J. I.; Gordon, J. I.; Jansson, J. K.; Knight, R., Diversity, stability and resilience of the human gut microbiota. *Nature* **2012,** *489*, 220.

(56)   Guilhot, E.; Khelaifia, S.; Scola, B. L.; Raoult, D.; Dubourg, G., Methods for culturing anaerobes from human specimen. *Future Microbiology* **2018,** *13* (3), 369-381.

(57)   Walker, C.; Preshaw, P. M.; Novak, J.; Hefti, A. F.; Bradshaw, M.; Powala, C., Long-term treatment with sub-antimicrobial dose doxycycline has no antibacterial effect on intestinal flora. *Journal of Clinical Periodontology* **2005,** *32* (11), 1163-1169.

(58)   Rosenblatt, J. E.; Stewart, P. R., Anaerobic bag culture method. *J. Clin. Microbiol.* **1975,** *1* (6), 527.

(59)   O'Keefe, S. J. D.; Chung, D.; Mahmoud, N.; Sepulveda, A. R.; Manafe, M.; Arch, J.; Adada, H.; van der Merwe, T., Why Do African Americans Get More Colon Cancer than Native Africans? *J. Nutr.* **2007,** *137* (1), 175S-182S.

(60)   Madden, J. A. J.; Plummer, S. F.; Tang, J.; Garaiova, I.; Plummer, N. T.; Herbison, M.; Hunter, J. O.; Shimada, T.; Cheng, L.; Shirakawa, T., Effect of probiotics on preventing disruption of the intestinal microflora following antibiotic therapy: A double-blind,

placebo-controlled pilot study. *International Immunopharmacology* **2005,** *5* (6), 1091-1097.

(61)   Weingarden, A. R.; Dosa, P. I.; DeWinter, E.; Steer, C. J.; Shaughnessy, M. K.; Johnson, J. R.; Khoruts, A.; Sadowsky, M. J., Changes in Colonic Bile Acid Composition following Fecal Microbiota Transplantation Are Sufficient to Control Clostridium difficile Germination and Growth. *PLOS ONE* **2016,** *11* (1), e0147210.