

A compartmentalisation approach to understand
microbial-host interaction in irritable bowel
syndrome and short-term dietary intervention

Yi Yi Yang

Thesis presented for the degree of
Doctor of Philosophy of Imperial College London

Supervised by
Professor Gary Frost
Dr Jonathan Swann

Department of Metabolism, Digestion and Reproduction
Faculty of Medicine
Imperial College London

October 2020

Statement of originality

I certify that the content of this thesis, and the research to which it refers are the products of my own work unless otherwise stated. The work and ideas of others are fully acknowledged in accordance with the standard referencing practices and discipline.

Copyright declaration

The copyright of this thesis rests with the author unless otherwise stated. Unless otherwise indicated, its contents are licensed under a Creative Commons Attribution-Non Commercial 4.0 International Licence (CC BY-NC). Under this licence, you may copy and redistribute the material in any medium or format. You may also create and distribute modified versions of the work. This is on the condition that: you credit the author and do not use it, or any derivative works, for a commercial purpose. When reusing or sharing this work, ensure you make the licence terms clear to others by naming the licence and linking to the licence text. Where a work has been adapted, you should indicate that the work has been changed and describe those changes. Please seek permission from the copyright holder for uses of this work that are not included in this licence or permitted under UK Copyright Law.

The copyright of Figure 4.5, Figure 4.10 and Figure 4.11 in this thesis rests with Elsevier and are licensed under a Creative Commons Attribution-Non Commercial-No Derivatives 4.0 International Licence (CC BY-NC-ND). Under this licence, you may copy and distribute the figure and the accompanying legends, provided this is not done for commercial purposes, and further does not permit distribution of the figure if it is changed or edited in any way, and provided you quote this link, DOI:<https://doi.org/10.1016/j.cell.2020.08.007>, provide a link to the license, and that the licensor is not represented as endorsing the use made of the work.



Abstract

Our gut harbours a biosphere of microorganisms constantly responding to the environmental cues we provide for them. From what we eat to how we behave and when we get sick so can our gut microbiome. However, this is not a one-way communication, these microorganisms also work for us in regulating many aspects of our body's biochemical, physiological and immune functions.

The gut microbiome lives in different compartments of varying nutrient availabilities, chemical and physiological conditions. The highly fluctuant environment of the small intestine (SI) and the presence bile acids and oxygen render this compartment inhabitable by most species. However, *Streptococcus spp.* thrive in the SI because they are highly efficient at metabolising simple carbohydrates for energy. Chapter 3 showed certain strains *Streptococcus spp.* may be in competition with one another for the availability of nutrients in the SI. Inside the colon, these facultative anaerobes were replaced by a diverse community of obligate anaerobes capable for fermenting complex carbohydrates, proteins and even mucin glycans. These microbial communities are resilient and adaptable to short-term Western diet intervention although carbohydrate fermentation activity of the colonic microbiota was decreased as Chapter 3 showed.

In Chapter 4, significant alterations in the community structure and metabolic capacity of the colonic microbiota were observed in patients with irritable bowel syndrome (IBS). Lower carbohydrate fermentation and increased protein fermentation were evident in patients with constipation-predominant IBS. Diarrhoea-predominant IBS patients excreted more primary bile acids. Using multi-omics integration, a novel host-microbial co-metabolic pathway involving purine salvage was identified.

Chapter 5 showed the passage of short chain fatty acids (SCFA) from the lumen to the mucosa was interrupted by the presence of a mucus layer. A diffusion dynamic for the trans-compartmental passage of SCFA was subsequently hypothesised.

This thesis presents a compartmentalised view of the gut microbiome, demonstrating its unique adaptations to dietary change and that disturbances in its community structure and functional capacity underlie the pathophysiology of functional gastrointestinal disorders.

Acknowledgements

To my supervisors, Professor Gary Frost and Dr Jonathan Swann, thank you for your guidance throughout my PhD and the critical feedback you provided for this thesis.

A special thanks to Dr Carine Blanchard for your guidance on the *in vitro* work and other aspects of my PhD. I really benefited from our weekly discussions which always leave me with ideas to think through afterwards.

Thank you to all the current and past members of the department of Metabolism, Digestion and Reproduction and my fellow StratiGrad PhD students who had offered me help on various occasions throughout my PhD. Special thank goes to Dr Marianne Koliana for helping me with cell culture; Dr Adesola Bello for your expertise on the bile acids analysis; Dr Jerusa Brignardello for your assistance on the GC-TQ and Dr Andreas Bruckbauer for helping me with the confocal imaging work.

To my mum, dad, Tak, my brother and my husband, thank you for your love, understanding and unwavering support in the past four years which carried me through to the finishing line. This thesis is dedicated to you.

Contribution of others

Dr Purna Kashyap and his team at the Mayo Clinic, Rochester, USA designed and conducted the clinical trial for the work described in Chapter 3 and Chapter 4.

Dr Robin R. Shields-Cutler from Macalester College, Minnesota, USA prepared the 16S sequencing data for statistical analysis in Chapter 3

Dr Tonya Ward from the University of Minnesota, USA performed 16S sequencing and pre-processed the data used in Chapter 4. Figure 4.6B was analysed and produced by Dr Tonya Ward and Dr Ruben A.T. Mars from the Mayo Clinic, Rochester, USA.

Dr Yogesh Bhattarai and Dr Purna Kashyap contributed to the design and execution of the Ussing experiment described in Chapter 4. Dr Ruben A.T. Mars analysed the data and produced Figure 4.5.

Ms Sambhawa Priya from the University of Minnesota, USA performed metabolomics and microbiome data integration using Lasso regression analysis described in Chapter 4.

Dr Tal Korem from Columbia University, USA performed the data integration for bacteria structural variable region and metabolites described in Chapter 4.

Dr Purna Kashyap and his team at the Mayo Clinic, Rochester, USA designed and executed the *in vitro* and *in vivo* hypoxanthine consumption experiment described in Chapter 4.

Contents

List of Figures	11
List of Tables	13
Abbreviations	14
Chapter 1	16
Introduction	16
1.1 Microbiome in the longitudinal gut compartments	16
1.2 Microbiome in the transverse gut compartments	18
1.3 Metabolites in the longitudinal gut compartments	20
1.3.1 Short chain fatty acids	20
1.3.2 Bile acids	22
1.3.3 Metabolic activities in the small intestines	24
1.4 Diet and the microbiota	25
1.5 Irritable bowel syndrome and the microbiota	26
1.6 Aims and Objectives	28
Chapter 2	29
Methods	29
2.1 Metabolomics	29
2.1.1 Proton Nuclear Magnetic Resonance (^1H NMR) spectroscopy	30
2.1.2 ^1H NMR sample preparation	33
2.1.3 ^1H NMR experiment conditions	34
2.1.4 ^1H NMR data pre-processing	35
2.1.5 Principal components analysis	35
2.1.6 Orthogonal Projection to Latent Structures-Discriminant Analysis	35
2.1.7 ^1H NMR metabolite assignment	36
2.1.8 Liquid Chromatography-Quadrupole Time-of-Flight Mass Spectrometry for bile acids profiling	36
2.1.9 LC-MS bile acid profiling experiment	38
2.1.10 LC-MS bile acid data processing	38
2.1.11 Gas Chromatography-Triple Quadrupole mass spectrometry for SCFA analysis	39
2.1.12 GC-MS SCFA quantification experiment	40
2.1.13 GC-MS SCFA quantification data processing	40

2.2 Microbiome	41
2.2.1 16S sequencing	41
2.2.2 Shotgun metagenomics sequencing	41
2.2.3 Microbiome data analysis	41
2.2 Multi-omics integration	43
Chapter 3	44
Western diet and the small intestinal microbiota	44
3.1 Introduction	44
3.2 Aims and Objectives	45
3.2 Methods	45
3.3.1 Participant recruitment	45
3.3.2 Sample and clinical data collection	46
3.3.3 Seven-day Western diet intervention	46
3.4 Results	47
3.4.1 All participants developed symptoms during the WD intervention	47
3.4.2 No metabolic differences were detected in the duodenum between high fibre and WD diet	49
3.4.3 Stool SCFA could discriminate between high fibre and WD diet	49
3.4.4 Distinct metabolic profiles in the duodenum and the colon	51
3.4.5 No system wide metabolic differences between high fibre and WD diet	52
3.4.6 Metabolites in the duodenum are associated with the development of symptoms	52
3.4.7 Microbiome showed no variation between high fibre and WD diet	54
3.4.8 Distinct microbial communities in the duodenum and the colon.	55
3.4.9 Metabolites are associated with the microbiome in their respective compartments	57
3.5 Discussion	58
3.6 Study limitation and future work	59
3.7 Conclusion	60
Chapter 4	61
Irritable bowel syndrome and the colonic microbiota	61
4.1 Introduction	61
4.2 Aims and Objectives	62
4.3 Methods	63
4.3.1 Participant recruitment and exclusion criteria	63
4.3.2 Clinical metadata	63
4.3.3. Sample collection	63
4.4 Results	64

4.4.1 Overview of study sample collection and subject demographics	64
4.4.2 Dietary intake was not significantly different between cohorts	65
4.4.3 Luminal metabolites differ between IBS subtypes and HC	65
4.4.4 Lower stool SCFA in IBS-C is associated with the microbiome	70
4.4.5 Evidence for enhanced microbial protein fermentation in IBS-C patients	73
4.4.6 Multi-omics integration identified a novel host-microbial pathway in IBS	76
4.4.7 Increased excretion of primary bile acids in IBS-D patients	81
4.4.8 IBS symptom severity is associated with luminal metabolite levels	83
4.4.9 A rare view into the mucosa metabolome in IBS patients	85
4.4.10 The lumen microbiome is richer than the mucosa	89
4.5 Discussion	90
4.5.1 IBS-subtype specific metabolic alterations are associated with the colonic microbiota and host physiology	90
4.5.2 Comparing the metabolome and microbiome in the lumen and the mucosa	91
4.5.3 Purine salvage underlies the pathophysiology of both IBS subtypes	91
4.6 Study limitation and future work	94
4.7 Conclusion	94
Chapter 5	95
Metabolites diffusion through the mucus compartment	95
5.1 Introduction	95
5.2 Methods	96
5.2.1 Cell culture	96
5.2.2 Confocal fluorescence microscopy	96
5.2.3 Lactate dehydrogenase (LDH) cytotoxicity assay	97
5.2.4 Mucus diffusion experiment	99
5.2.5 GC-MS SCFA Quantification	100
5.3 Results	103
5.3.1 Mucus reduced the flow of SCFA	103
5.3.2 Mucus temporarily traps SCFA	104
5.3.3 Estimating SCFA uptake by the goblet cells	106
5.3.4 Diffusion through reconstituted porcine mucus	108
5.4 Discussion	110
5.5 Limitations and future work	112

Chapter 6	114
General discussion	114
6.1 Summary of key results	114
6.1.1 New learnings from exploring the SI microbiota	114
6.1.2 The microbiome in the large intestines may be more diverse but it is not always a good sign	115
6.1.3 We must not forget about the mucus compartment	117
6.2 Future work	117
6.3 Conclusions	118
References	119
Appendix	133
Publications	138

List of Figures

Chapter 1

Figure 1.1 Characteristics of the small intestines and the large intestines which influence the diversity and structure of the microbial inhabitants.

Figure 1.2 The transverse gut compartment in the colon is comprised of the lumen, the double-layered mucus and the mucosa. Oxygen concentration decreases from the mucosa to the lumen.

Figure 1.3 Majority of the SCFA in the gut are produced from the anaerobic fermentation of dietary fibre and a small proportion coming from the fermentation of amino acids.

Figure 1.4 Transcellular transport of SCFA is mediated by monocarboxylate transporters and a bicarbonate ion exchange system.

Figure 1.5 Bile acids are synthesised in the liver and modified by the microbiota in the colon.

Chapter 2

Figure 2.1 NMR theory.

Figure 2.2 NMR spectrum of 2-methylbutyrate from (A-B) one-dimensional NOSEY experiment and (C) two-dimensional J-RES experiment.

Figure 2.3 Outline of a LC-QTOF system.

Figure 2.4 Overview of a GC-TQ system.

Figure 2.5 Comparing alpha-diversity between two samples.

Chapter 3

Figure 3.1 Comparing the daily consumption of the five major nutrient groups and daily calories intake pre-intervention and during the western diet (WD) trial.

Figure 3.2. Comparing metabolites in the duodenum aspirate of high fibre versus WD diet.

Figure 3.3. Faecal short chain fatty acids and lysine are reduced following consumption of the western diet.

Figure 3.4. ^1H 600 MHz NMR spectrum of QC samples from duodenal aspirates (blue) and stool (red) showing chemical shifts (δ) regions between 0 and 4.

Figure 3.5 No system wide metabolic differences between high fibre and WD diet.

Figure 3.6 Duodenum aspirate metabolites are associated with SIBO diagnosis and post-interventional symptom development.

Figure 3.7. No difference in microbiome composition between high fibre and Western diet.

Figure 3.8 Alpha-diversity measured by the Shannon metric was significantly higher in the stool samples than the duodenum aspirates.

Figure 3.9. Longitudinal gut compartments harbour different microbial communities.

Figure 3.10. Associations between microbial species and metabolites in the duodenum and the colon.

Chapter 4

Figure 4.1. Overview of metabolomics analytical strategy.

Figure 4.2. FFQ recorded dietary intake for protein, carbohydrate, fat and fibre were compared.

Figure 4.3. No distinct clusters shown between healthy controls (HC), IBS-C and IBS-D based on unsupervised analysis of luminal (stool) metabolites.

Figure 4.4. MCCV-OPLS-DA model identified several luminal metabolite features discriminating between IBS-C patients and healthy controls (HC).

Figure 4.5. Biopsy secretory response to serotonin (5-hydroxytryptamine, 5-HT).

Figure 4.6. IBS-C microbiome is more distinct from HC.

Figure 4.7. Alteration in stool BCAA and BCFA levels indicate increased protein fermentation in IBS-C patients.

Figure 4.8. Network plot of significant positive (red edges) and negative (blue edges) associations between metabolites (blue and pink nodes) and microbial taxa (yellow node) produced from Lasso penalized regression.

Figure 4.9. Boxplots showing the relative abundance of NMR peak integrals for glutamate, aspartate, hypoxanthine and uracil in stool samples.

Figure 4.10. Hypoxanthine consumption by specific gut microbiome members as suggested by microbial gene region associations.

Figure 4.11. Evidence for hypoxanthine breakdown by the host colonic mucosa.

Figure 4.12. Elevated stool primary bile acids (BA) in IBS-D patients.

Figure 4.13. Luminal metabolites could differentiate between symptom severity in IBS-subtypes.

Figure 4.14. Sigmoid biopsy metabolites.

Figure 4.15. Bar charts showing average SCFA ratios in the lumen (stool samples) measured using 600 MHz ^1H NMR and the mucosa (sigmoid biopsy) measured using GC-TQ targeted quantification.

Figure 4.16. Network plot showing Spearman correlations between mucosal SCFA and host genes.

Figure 4.17. Comparing alpha-diversity in the lumen and the mucosa.

Chapter 5

Figure 5.1. Confocal fluorescence microscopy of a 1mm x 1mm section of membrane covered with HT29-MTX-E12 cells grown for 10 days.

Figure 5.2. Bar charts with standard deviation error bars showing the average % cell death for each treatment solution. Labels indicate the concentration for acetate in each solution.

Figure 5.3. GC-MS calibration curve for (A) acetate, (B) propionate and (C) butyrate.

Figure 5.3. Optimised GC-MS calibration curves for (A) acetate, (B) propionate and (C) butyrate where all 11 calibration points fall within the fitted linear line.

Figure 5.4. Line graphs showing the average concentration (μM) of acetate, butyrate and propionate in the apical compartments (**A-C**) and the basolateral compartments (**D-F**) over time (hr).

Figure 5.5. Line graphs comparing the average concentration (μM) of (A) acetate, (B) butyrate and (C) propionate in the apical (orange line), BL (green line) compartment and the mucus layer (yellow line) over time (hr).

Figure 5.6. Line graphs comparing the average percentage (%) of (A) acetate, (B) butyrate and (C) propionate in the remained in Transwells post treatment.

Figure 5.7. Line graphs showing the average concentration (μM) of (**A**) acetate, (**B**) butyrate and (**C**) propionate in the apical compartments over time (hr).

Figure 5.8. SCFA diffusion dynamics through the mucus layer.

List of Tables

Chapter 2

Table 2.1. Mass to charge ratio (m/z) and retention time for bile acids identified in the LC-MS profiling experiment.

Chapter 3

Table 3.1 Summary of symptom questionnaires.

Chapter 4

Table 4.1. Sample collection and cohort description.

Table 4.2 Phylogenetic level differences in stool samples between HC and IBS subgroups.

Table 4.3 Top five associations between butyrate, acetate and propionate with microbial species produced using the Lasso penalized regression model.

Chapter 5

Table 5.1. Table of treatment solutions made for the LDH cytotoxicity assay.

Abbreviations

1D one-dimensional
¹H NMR Proton Nuclear Magnetic Resonance
2D two-dimensional
2mb 2-methylbutyrate
5-HT 5-Hydroxytryptamine, Serotonin
AA Amino Acids
ABC ATP-Binding Cassette
ACN Acetonitrile
BA Bile Acids
BCAA Branched Chain Amino Acids
BCFA Branched Chain Fatty Acids
BL Basolateral
BMI Body Mass Index
CA Cholic Acid
CBP Choline-binding-proteins
CDCA Chenodeoxycholic Acid
CFU Colony Forming Units
CPMG Carr-Purcell-Meiboom-Gill
D₂O Deuterium Oxide
DCA Deoxycholic Acid
EI Electron Impact
ENS Enteric Nervous System
FDR False Discovery Rate
FFQ Food Frequency Questionnaire
FID Free Induction Decay
FXR Farnesoid X Receptor
GC-MS Gas Chromatography-Mass Spectrometry
GI Gastrointestinal
GPCR G-protein-coupled receptors
HC Healthy Control
HDAC Histone Deacetylase
Hpx Hypoxanthine
Hz Hertz
IBS Irritable Bowel Syndrome
IBS-C Constipation-predominant IBS
IBS-D Diarrhoea-predominant IBS
IEC Intestinal Epithelial Cells
J-RES J-resolved
KOH Potassium hydroxide
LC-MS Liquid Chromatography-Mass Spectrometry
LC-QTOF Liquid Chromatography-Quadrupole Time-of-Flight
LCA Lithocholic Acid
LDH Lactate dehydrogenase
m/z Mass to charge ratio

MCCV Monte Carlo Cross-Validation
MCT1/2/4 Monocarboxylate Transporter 1/2/4
MRM Multiple Reaction Monitoring
MTBE Methyl tert-butyl ether
MTBSTF *N*-tert-butyldimethylsilyl-*N*-methyltrifluoroacetamide
MUC2 Type II Mucin
NAC N-acetylcysteine
NaH₂PO₄ Sodium di-hydrogen phosphate
NaN₃ Sodium azide
NMDS Non-metric Multidimensional Scaling
NOESY Nuclear Overhauser Enhancement Spectroscopy
OPLS-DA Orthogonal Projection to Latent Structures-Discriminant Analysis
OSC Orthogona-Signal-Correction
OTU Operational Taxonomical Units
PBS Phosphate-buffered saline
PCA Principal Components Analysis
PCR Primary Chain Reaction
PEG Polyethylene glycol
pHi Intracellular pH
PNP Purine nucleoside phosphorylase
ppm Parts per million
PQN Probabilistic Quotient Normalization
PTS Phosphotransferase Systems
QC Quality Control
rf Radio frequency
RS Resistance Starch
RT Retention Time
SCFA Short Chain Fatty Acids
SI Small Intestine
SIBO Small Intestinal Bacteria Overgrowth
SMCT1 Sodium-Monocarboxylate Transporter 1
SSS Symptom Severity Score
STOCSY Statistical Total Correlation Spectroscopy
TBDMSCI tert-butyldimethylchloro-silane
TIC Total Ion Current
TMA Trimethylamine
TMAO Trimethylamine N-oxide
TQ Triple Quadrupole
TSP 3-(trimethylsilyl) propanoic acid-2,2,3,3,-d₄
TSS Total Sum Scaling
UDCA Ursodeoxycholic Acid
WD Western Diet
XDH Xanthine dehydrogenase

Chapter 1

Introduction

1.1 Microbiome in the longitudinal gut compartments

The human gastrointestinal (GI) tract is made up of a series of longitudinal tissue structures forming the central part of the digestive system. Trillions of commensal bacteria along with fungi and viruses live inside the GI tract and they are collectively known as the human gut microbiome¹. Microbial diversity and richness vary along the longitudinal compartments largely referring to the small intestines and the large intestines where differences in nutrient availabilities, chemical environments and physiological functions influence the formation of distinct microbiome niches^{2,3,4}.

The small intestine (SI) is comprised of three parts, the duodenum, the jejunum and the ileum. Digestion and absorption of proteins and monosaccharides takes place in the duodenum. The jejunum is responsible for the absorption of free fatty acids, calcium and vitamins whereas the ileum is the site for bile acids reabsorption⁵. Rapid nutrient absorption takes place in the SI which is accompanied by an influx of digestive enzymes and bile salts making it a harsh environment for the microbiota to survive in^{6,10}. The absorptive action at the intestinal mucosa can allow a small amount of oxygen to diffuse through from the blood supply at the villi². Therefore, the SI microbiome community is made up of mostly facultative anaerobes that can tolerate some levels of oxygen and the presence of bile acids⁶. The density of the SI microbiota increases from 10^3 colony forming units (CFU)/mL in the duodenum to 10^8 CFU/mL in the ileum^{2,3,4}. The SI has a thin mucus layer that does not confer a large amount of physical separation between the mucosa and bacteria in the lumen. However, there is an extensive network of immune cells in the SI that act as a chemical defence to maintain a relatively low bacterial count⁷. Paneth cells secrete antimicrobial proteins (AMPs) to the epithelial surface thereby limiting bacteria invasion⁷. Although phylogenetically less diverse, the SI microbiome is more dynamic, given the need to respond to the rapidly changing luminal conditions^{6,8}. A study showed that the temporal profile of an individual's ileum microbiome fluctuates throughout a period of nine days and even within-day variation has been observed⁹.

Indeed, the SI microbiota are highly adapted at metabolising SI nutrients as revealed by several metagenomics studies. Based on the work by Zoetendal *et al.*¹¹, the group identified many genes related to carbohydrate metabolism in the SI microbiota, much more than the faecal microbiota. In particular, enzymes involved in the pentose phosphate pathway, lactate

and propionate metabolism and sugar phosphotransferase system (PTS) were enriched. The PTS transcripts were mostly from *Streptococcus* sp. which suggest that they are the main utilisers of sugar in the SI. Interestingly, the study also identified acetate, lactate and butyrate in the ileal effluent samples. Another metagenomics study of *Streptococcus* isolates obtained from ileostomy patients showed that 20% of *Streptococcus* spp. core genome is associated with the SI metabolism of nutrients such as lipids, amino acids, carbohydrates and inorganic ions¹².

In the large intestines, which comprises the caecum and colon, bacteria count increases to a staggering 10^{11} CFU/mL. This is a result of reduced luminal flow and oxygen depletion creating a favourable environment for an expansive array of microbial species^{2,3,4}. Dietary fibres that are not well absorbed from the SI enter the colon and thus constitute the primary source of nutrients for the colonic microbiota. Many species have the ability to consume a variety of complex polysaccharides for energy¹³ where *Bacteroides* spp. have the most extensive polysaccharide utilisation (carbohydrate-active enzymes [CAZmes]) genes than any other gut bacteria¹⁴. Microbial richness in the colon is highly correlated with colonic transit, which has been demonstrated in a number of studies. Using stool consistencies as a proxy for colonic transit, researchers found that short transit times were associated with increases in fast-growing species such as *Ruminococcaceae* and *Bacteroides*¹⁵. Another study found that longer colonic transit times were associated with higher microbial richness accompanied by an increase in several microbial groups including *Methanobrevibacter* and *Desulfovibrio* and a decrease in *Lachnospiraceae*¹⁶. *In vitro* evidence suggest that long transit times affect microbial composition independent of other factors¹⁷. Microbial composition also differs between different segments of the large intestines as one study showed that mucosa-associated species in the distal colon was higher in lactobacilli where *Eubacterium rectale* and *Faecalibacterium prausnitzii* dominated the ascending and descending colon²³.

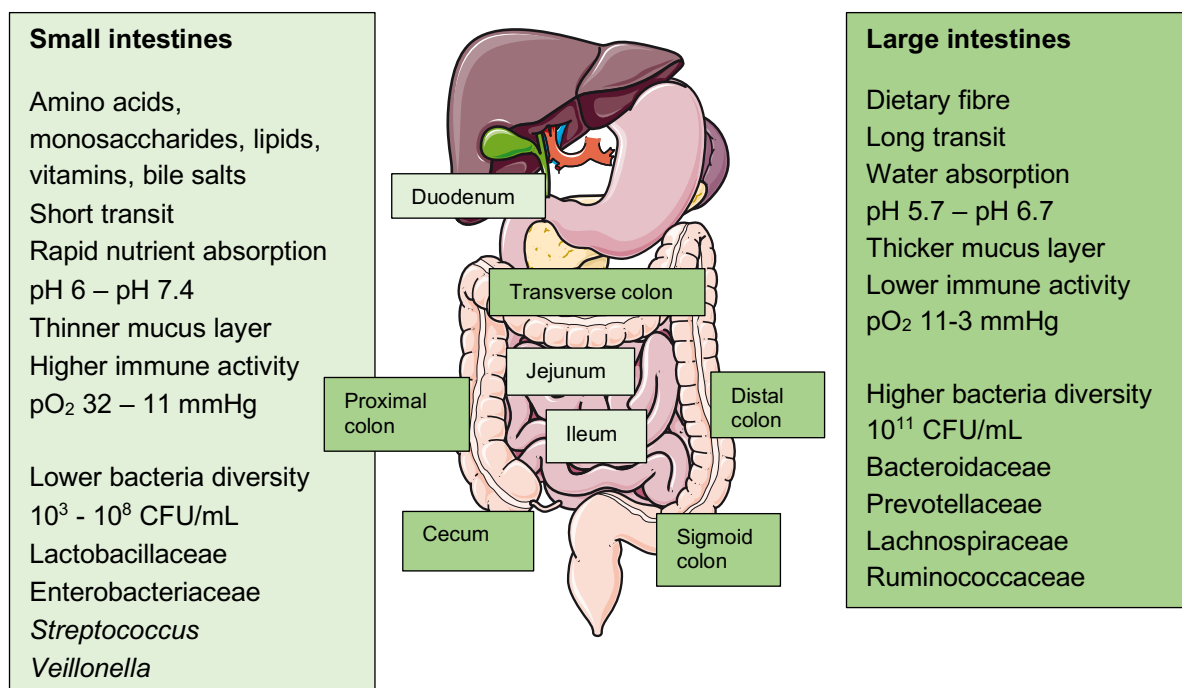


Figure 1.1 Characteristics of the small intestines and the large intestines which influence the diversity and structure of the microbial inhabitants.

1.2 Microbiome in the transverse gut compartments

Differences in microbial biogeography also exist along the transverse compartments¹⁸ (*i.e.* from the lumen to the mucosa), as evidenced by the presence of distinct microbial niches^{19,20,21}. However, inconsistencies in mucosa-associated species are found between studies largely due to the differences in sampling sites²² indicating that there could also be spatial variation within the mucosa community. These differences serve as reminders that while faecal samples, due to their accessibility, have been the primary sample type for studying the colonic microbiota, care should be taken when using them to infer functional significance of the entire colonic community.

The intestinal lumen is the primary habitat for the gut bacteria since there is a constant stream of nutrients passing through. Beneath the lumen lies the mucosa compartment where they are separated by a layer of mucus whose thickness increases from approximately 20 µm in the SI to 830 µm in the colon²⁴. The primary purpose of the mucus layer is to protect the epithelium from direct contact with the harsh chemicals and bacteria in the lumen and thus the significantly thicker mucus layer in the colon reflects its dense bacterial load. Mucus is a gel-like structure made of mucin that is excreted from the goblet cells, which then expands extracellularly into a mesh network²⁴. It was revealed that the colonic mucus is divided into two layers; a compact inner layer firmly attached to the epithelium and a loose outer layer

where most of the mucosa-associated bacteria resides^{25,26}. The importance of the mucus layer was demonstrated in a mucin-deficient mouse model where direct bacteria contact with the mucosa as a result of the absence of a mucus layer was associated with the development of colitis and inflammation-induced colorectal cancer²⁷. In addition to providing protection for the epithelium, mucus plays an important role in mediating host-microbial symbiosis where bacteria use mucin glycans as attachment sites and a rich source of carbohydrates^{28,29}. The genome of the gut commensal *Akkermansia muciniphila* is enriched with glycosyl hydroxylases which allow them to break down mucin glycans for energy³⁰. *Bacteroides spp.* are capable of consuming both dietary and host-derived complex polysaccharides by using their Sus (Starch utilisation system)-like systems¹⁴. There are many more enzyme activities associated with mucin glycan foraging and it is thought that they help confer survival advantage for these bacteria by allowing them to utilise different source of carbohydrates^{29,31}.

Another factor that drives spatial differentiation between the lumen and the mucosa is the existence of an oxygen gradient which creates a selection pressure on bacteria that can thrive in different oxygen microenvironments^{32,33}. Although the colon is generally considered an anaerobic chamber, there is an increase in oxygen concentration from the lumen to the mucosa as a result of rapid oxygen consumption by the intestinal epithelial cells³². Mucosa-associated oxygen-tolerant species were identified in the Proteobacteria and the Actinobacteria phyla³⁴. It appears that not all *Bacteroides spp.* are strictly anaerobic as *B. fragilis* can grow in nanomolar concentrations of oxygen³⁵.

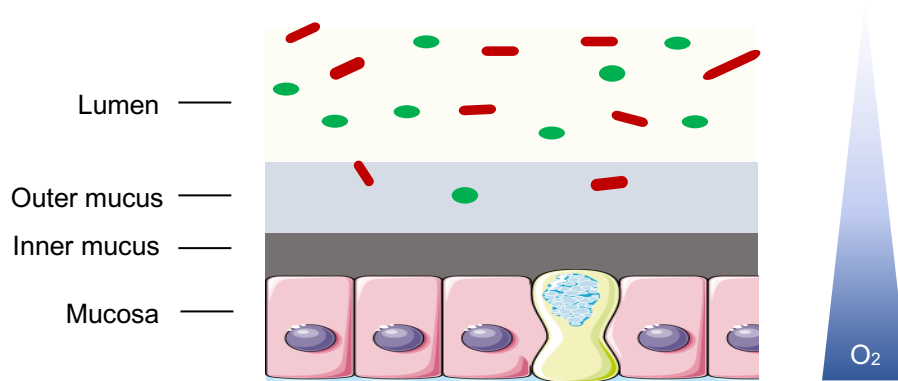


Figure 1.2 The transverse gut compartment in the colon is comprised of the lumen, the double-layered mucus and the mucosa. Oxygen concentration decreases from the mucosa to the lumen.

1.3 Metabolites in the longitudinal gut compartments

The gut microbiome is integral to human health as they extend the metabolic functionality of the host and participate in immune development and regulation³⁶. Microbial metabolites are chemical signals used by the microbiota to establish communication with the host and exert their regulatory effects. The following sections provide an overview of major metabolites in the longitudinal gut compartments and their association with the microbiota.

1.3.1 Short chain fatty acids

Much of our knowledge on the microbial metabolites comes from the study of faecal samples thus they largely represent the metabolic activities in the colon. The most prominent microbial activity in the colon is the anaerobic fermentation of dietary fibre which produces short chain fatty acids (SCFA) along with gases such as hydrogen, methane and carbon dioxide^{37,38}. Microbial cross-feeding can result in the production of SCFA via different intermediate metabolites such as lactate and succinate⁴⁰. The most abundant SCFA in the colon is acetate, butyrate and propionate where they exist in a ratio of 6:2:2 based on studying the colon content of post-mortem subjects⁴¹. Once absorbed, butyrate is rapidly oxidised by the colonocytes for energy metabolism and to a much lesser extent acetate and propionate³⁹. Around 60-70% of energy required by the epithelial cells in the distal colon are derived from bacterial fermentation products. The effects of butyrate extend beyond being the primary source of energy for colonocytes^{42,43}. Butyrate is known to inhibit histone deacetylase (HDACs) activities to exert anti-inflammatory effects⁴⁴ and induce apoptosis on cancer cells⁴⁵. In the healthy intestinal mucosa butyrate can also promote barrier function by increasing the expression of tight-junction proteins^{46,47}. All three SCFA have been found to regulate immune activities through interaction with G-protein-coupled receptors (GPCR)^{39,48}. In addition of their localised effect on the gut epithelium, SCFA can exert system wide effects modulating processes such as satiety and energy homeostasis, which has been implicated in the development of insulin resistance and obesity⁴⁹.

SCFA can also be produced from the bacterial fermentation of branched chain amino acids (BCAA)³⁹, which can also produce branched chain fatty acids (BCFA) and a series of harmful metabolites that have been negatively associated with health. This will be discussed further later on in the chapter.

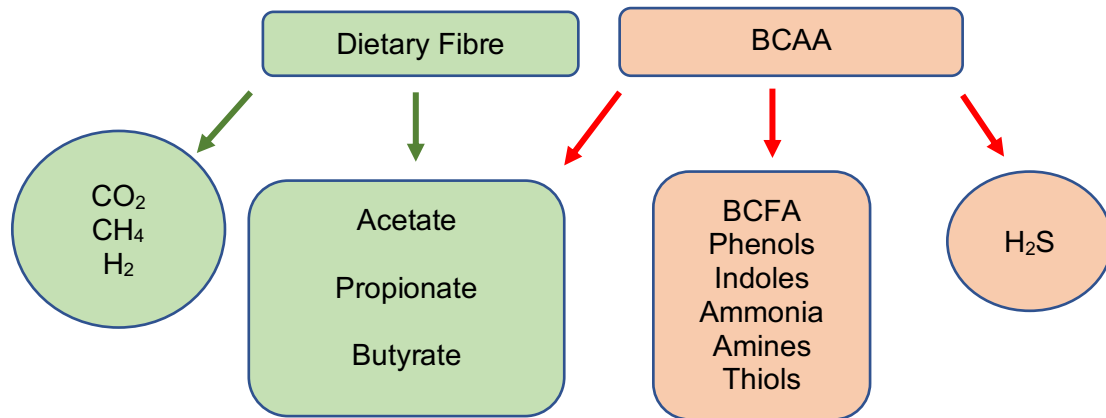


Figure 1.3 Majority of the SCFA in the gut are produced from the anaerobic fermentation of dietary fibre and a small proportion coming from the fermentation of amino acids.

Study of SCFA concentration in different colon segments of sudden death victims found that the highest concentrations of SCFA were present in the caecum with 131 mmol/kg detected, which decreased to 80 mmol/kg in the distal colon⁴¹. This indicates that SCFA are rapidly absorbed by colonic epithelium in the proximal section of the gut. A number of transporters have been identified for the uptake of SCFA into the epithelial cells. Most of them are proton exchangers where the active transport of SCFA at the apical membrane is coupled with the export of sodium and hydrogen ions out of the cell and into the lumen. The monocarboxylate transporter 1 (MCT1) is present on both the apical and the basolateral membrane whereas sodium-monocarboxylate transporter 1 (SMCT1) is exclusively present on the apical membrane and MCT4 on the basolateral membrane⁵⁰. Interestingly the expression of MCT1 was found to be much lower in the sigmoid colon than the ascending colon⁵¹ which likely indicates the epithelial surface of the ascending colon is more adapted at absorbing SCFA. A bicarbonate/SCFA exchange system was also proposed to be present on both sides of the epithelial cell although the identity of the transporter is unclear^{52,53}. It is believed that the release of HCO_3^- into the lumen is an important pH regulatory mechanism to respond to the production of SCFA⁵⁴. SCFA can also diffuse directly through the lipid membrane via the nonionic form⁵⁵.

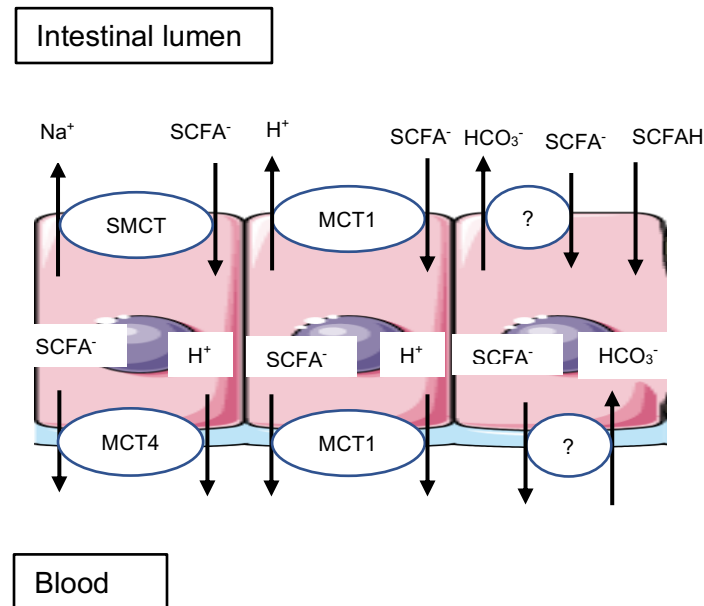


Figure 1.4 Transcellular transport of SCFA is mediated by monocarboxylate transporters and a bicarbonate ion exchange system.

1.3.2 Bile acids

Bile acids (BA) are host-microbial co-metabolites produced in the liver and modified by the microbiota in the colon. BA plays an essential role for the emulsification and absorption of dietary lipids and fat-soluble vitamins in the SI⁵⁶. After being synthesised from cholesterol in the liver by cholesterol 7 α -hydroxylases (CYP7A1/CYP27A1), primary BA namely cholic acids (CA) and chenodeoxycholic acids (CDCA) are conjugated with taurine or glycine before being excreting into the small intestines. Around 95% of the excreted BA re-enters the entero-hepatic circulation to replenish the BA pool in the liver. The remaining BA continue to flow towards the colon where they first undergo deconjugation by bacterial bile salt hydroxylases (BSH) before being converted to secondary BA, deoxycholic acids (DCA), lithocholic acids (LCA) and ursodeoxycholic (UDCA), via bacterial 7 α / β -dehydroxylases in the colon^{56,57}. BA synthesis is regulated by a negative feedback mechanism whereby BA activates the farnesoid X receptor (FXR) in the ileum and the liver which in turn inhibits enzymes involve in the synthesis of BA⁵⁸. Tauro- β -muricholic acid (T β MCA) is a potent antagonist of FXR. As authors in one study found increased levels of T β MCA accompanied by increased CYP7A1 expression in germ-free mice suggests the ability for T β MCA to reduce the inhibitory effect of FXR⁵⁹. A link between FXR inhibition by T β MCA as a result of reduced BSH activity in *Lactobacillus* and decreased obesity in mice has been described⁶⁰. BSH activity has been identified in many commensal bacteria where it is common in *Bifidobacterium* and *Lactobacillus*⁶³. BA also affect microbial community structure in the gut. DCA is a potent antimicrobial agent owing to its

hydrophobicity which can have a deterrent effect on the bacteria cell wall⁶¹. BA can also induce DNA damage on Gram-positive colonies⁶¹. *Clostridioides difficile* (*C. difficile*) is a pathogen that causes diarrhoea and uses taurocholic acid for growth⁶⁴. Interestingly, *C. difficile* spore germination was inhibited by analogues of CDCA⁶⁵. BA have also been shown to enhance absorption where DCA, glycol-DCA and tauro-DCA increased jejunum membrane permeability *in vitro*⁶².

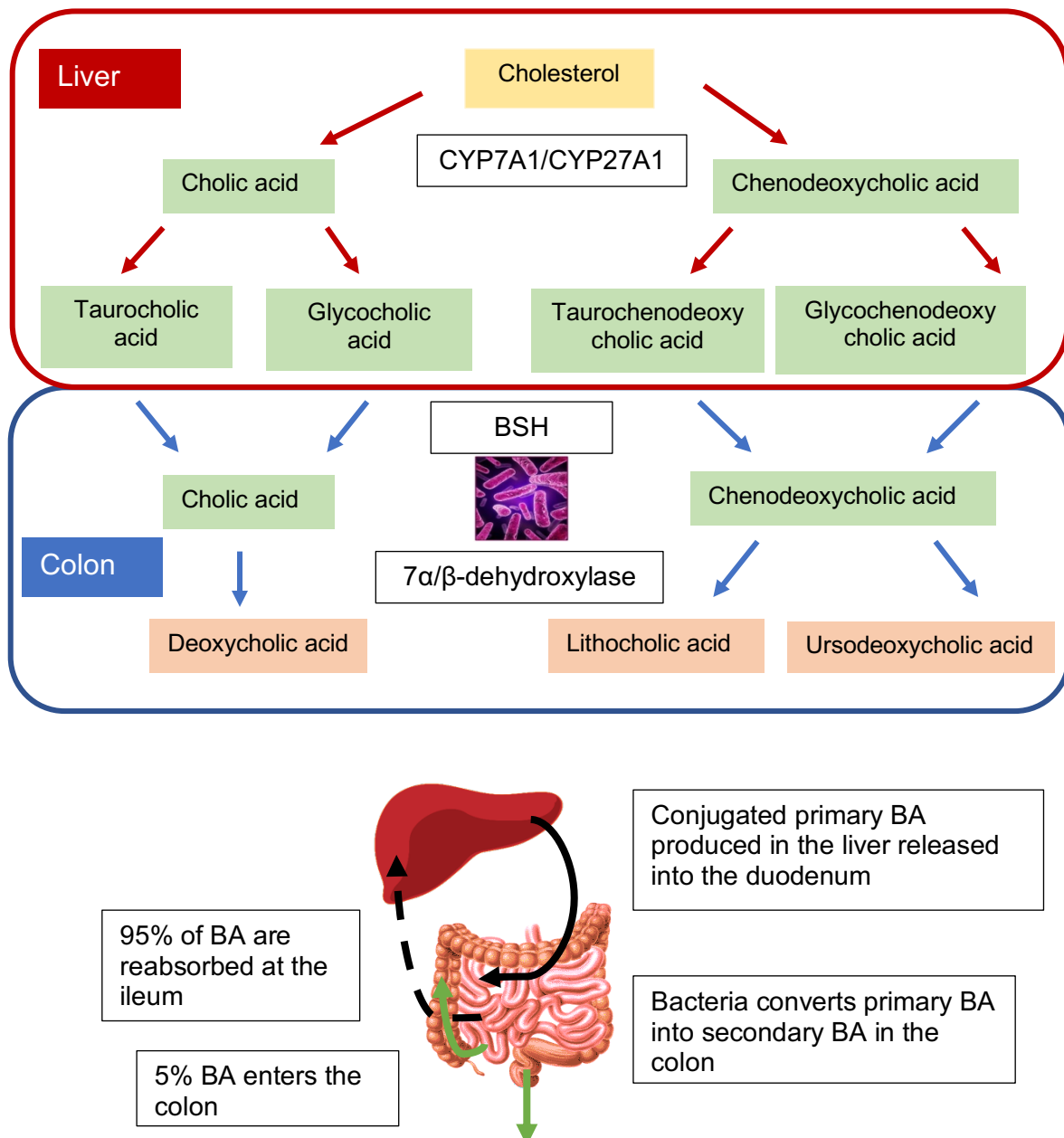


Figure 1.5 Bile acids are synthesised in the liver and modified by the microbiota in the colon.

1.3.3 Metabolic activities in the small intestines

Host digestive enzymes are responsible for breaking down macronutrients namely protein, simple carbohydrates and fats in the SI. Majority of the digested nutrients are then absorbed along the SI where some are also available to the microbiota.⁸¹ The digestibility of proteins by the host enzymes varies depending on its source where proteins from red meat are more digestible than plant-based proteins.²³⁵ They also yield different amino acids profiles which may influence the community structure and the metabolic activity of the SI microbiota. One study found alterations in stool microbiome composition of mice fed with protein diet derived from meat, dairy and plant.⁶⁷ Host-derived proteins can be digested by bacteria in the gut to release amino acids which they then incorporate into their cells as protein building blocks or catabolise further. The former conserves more energy as it relieves the need for *de novo* synthesis of amino acids⁶⁹. However, there are more interconversion steps for utilising amino acids compared to monosaccharides therefore it is clear why gut bacteria prefer monosaccharides as their primary energy source⁷⁶. The majority of our understanding of bacterial protein metabolism comes from studying the large intestines and little is known about the role played by the SI microbiota *in vivo*. It was proposed that the catabolism of lysine by the gut microbiota takes place in SI after observing a deficit between intestinal metabolism of lysine and lysine recovered from the intestinal mucosa of milk-fed piglets^{70,71}. The same reasoning has been extended to other amino acids and BCAA as reviewed by Davila *et al*⁷². A monoculture study of SI bacteria (*Escherichia. coli*, *Streptococcus spp.*) showed that amino acids were rapidly oxidised suggesting their utilisation by bacteria in the SI^{74,75}. Gut microbiota are also capable of providing amino acids to the host⁷³. Results from these studies suggest that the SI bacteria may play a role in the first pass metabolism of amino acids.

The majority of lipids from the diet are emulsified and absorbed in the SI with very little (<5%) reaching the colon⁷⁷. Given that BA are important lipid emulsifiers, microbial modulation of BA has potential to modify the absorption of lipids in the SI. Interestingly, a recent study by Martinez-Guryn *et al.*⁷⁸ found that germ-free mice were not able to absorb fat and were resistant to diet-induced obesity. In addition, germ-free mice transplanted with a 'high fat' SI microbiota had increased lipid absorption compared to those with a 'low fat' SI microbiota. Lipid absorption genes were identified to be up-regulated in the proximal SI epithelial organoids of mice with the 'high fat' SI microbiota. This study supports the notion that the SI microbiota plays an important role in the absorption of lipids.

In addition to macronutrients, micronutrients such as vitamins are also absorbed in the small intestines. The emulsifying action of BA on fat-soluble vitamins is important for their absorption. Some vitamins can be synthesised by the microbiota. It was found that all Fusobacteria and

over 90% of Bacteroidetes were predicted to be vitamin B12 producers.²³⁶ Bacteroides can also produce vitamin B6 and utilize them as cofactors for many biological reactions associated with the host immune response.²³⁶ However, majority of the body's vitamin B intakes come from food and some microbiota have shown to compete with the host for vitamin B12 (cobalamin) to use as cofactors in their own metabolic pathways⁸⁰. The SI bacteria *Veillonella* and Enterobacteriaceae produce menaquinone (vitamin K2)⁷⁹. Some microbial-derived K2 isoforms are positively associated with cognition as one study found.²³⁷

Taken together, evidence in the literature suggest that metabolism and absorption of nutrients in the SI is intimately linked to the microbiota.

1.4 Diet and the microbiota

The food that we eat has a direct impact on our gut microbiome⁸¹⁻⁸³. As described earlier, non-digestible dietary fibre is the main source of energy for the colonic microbiota. A loss of species abundance has been frequently observed when switching from a high fibre to low fibre diet as well as changes to the carbohydrate degrading capacity of the microbiota over generations⁸⁴. This is particularly concerning for Western societies, where the dietary habits of individuals have shifted to consuming foods that are low in fibre and high in simple sugars and fats⁸⁵. When analysing the gut microbiome of immigrants from non-Western countries in the US, researchers found an immediate loss of *Prevotella* strains and carbohydrate-active enzyme domains of their gut microbiota in the early stages of their settlement⁸⁶ and this effect increased with the duration of their stay. One group found that when a high sugar/high fat diet was fed to humanized gnotobiotic mice previously consuming a low sugar/high fibre diet, the microbiome structure in these mice shifted within a day and was associated with increased weight gain of the mice⁸⁷. The deleterious effects of dietary emulsifiers in promoting colitis and non-caloric artificial sweeteners in driving the development of glucose intolerance through alterations in microbiota composition have been demonstrated in mice^{88,89}.

As discussed earlier, SCFA produced by microbial carbohydrate fermentation is an important fuel for the colonocytes and can induce mucus secretion. Conversely, a lack of fibre could facilitate the degradation of the mucus layer as the microbiota feed on mucin glycans as an alternative fuel source⁹⁰. In addition to fibre, protein and fat intakes can also alter microbial composition⁹¹. A study by David *et al.*⁹² showed that short-term increases in protein and fat intake as part of an animal-based diet induce alterations in the species abundance and increase the production of faecal bile acids and branched chain fatty acids (BCFA). BCFA are a reliable marker of protein fermentation as they are produced from bacteria deamination of

branched-chain-amino-acids (BCAA)^{93,94}. Little is known about the effect of BCFA on host health, although isobutyrate can be used by the colonocytes when the supply of butyrate is limited⁹⁵. Bacterial protein fermentation has been widely associated with adverse health outcomes mainly due to the production of toxic substances such as ammonia, phenols and sulphides, which have been linked to the development of colon cancer⁶⁶. Higher concentration of ammonia was found to inhibit SCFA oxidation in rat colonic epithelial cells⁹⁷. Evidence suggest that fibre consumption may help protect the host from these toxic metabolites. When pigs fed with resistant starch (RS) and wheat bran (WB) and pigs fed with RS without WB were compared, researchers observed an increase in carbohydrate fermentation and decrease in ammonia concentration in the distal colon of the RS+WB group, which led them to suggest that consumption of RS and WB together may reduce the risk of colorectal cancer which typically occur in the distal colon.⁶⁸ In another study, high levels of p-cresol was detected in the urine of both rural South Africans and African Americans. The difference being rural South Africans also has higher concentration of SCFA in their stool but lower incidents of colorectal cancer. It was suggested that the higher fibre intake of rural South Africans protected them from the deleterious effect of protein fermentation.⁹⁹

1.5 Irritable bowel syndrome and the microbiota

Irritable bowel syndrome (IBS) is a chronic functional gastrointestinal disorder affecting approximately 11% of the global population.¹⁰⁰ According to the Rome IV criteria, patients are diagnosed based on their stool form to either constipation predominant (IBS-C), diarrhoea dominant (IBS-D) or mixed (IBS-M)¹⁰¹. Common symptoms of IBS also include abdominal pain, bloating, distention and psychological ailments such anxiety and depression. The heterogenous nature of IBS compounded by a lack of understanding of its disease aetiology make IBS difficult to diagnose and to provide an effective treatment plan. In recent years, scientists are increasingly looking to the gut microbiome for answers since many IBS symptoms triggers and pathological characteristics are associated with the gut microbiota.¹⁰²

One of the prominent features of IBS is abnormal gut motility. Evidence suggest that microbiota and gut motility are highly interrelated. The contractile response in the GI tract is mainly controlled by the enteric nervous system (ENS) which is a complex network of neurons localised along the gut.¹⁰³ Mouse enteric neurones were found to express SCFA transporter MCT2 and butyrate induced colonic transit through regulation of enteric neurons.¹⁰⁴ When mice were transplanted with faecal microbiota from IBS-D patients, increased GI transit, impaired barrier function and the development of anxiety-like behaviour was observed in these mice.¹⁰⁵ A recent study uncovered a mechanism by which gut microbial metabolite can

modulate the serotonergic secretion pathway. Serotonin (5-hydroxytryptamine, 5-HT) is an important neurotransmitter where > 90% of 5-HT in the human body is produced in the gut by enterochromaffin (EC) cells.¹⁰⁶ The work by Bhattarai Y et al. found that tryptamine, a metabolite produced from bacterial decarboxylation of dietary tryptophan accelerates colonic transit by activating the epithelial G-protein coupled receptor (GPCR) serotonin receptor 4 (5-HT₄R) and increasing anion-dependent fluid secretion in the proximal colon.¹⁰⁷ CDCA is a potent stimulator of colonic secretion and the administration of 1000mg CDCA provided in a delayed-release capsule accelerated colonic transit in IBS-C patients.^{108,143,144} Primary BA are commonly detected in the faeces of IBS-D patients^{109,145}. Interestingly, approximately 30% of IBS-D patients have bile acid malabsorption which is a condition where BA are inadequately reabsorbed in the distal small intestines leading to rapid colonic transit as a result of excess primary BA entering the colon.¹⁴⁶

As described earlier, butyrate is a key energy source for the colonocytes and can restore barrier function by increasing the expression of tight junction proteins¹¹⁰. IBS patients have less butyrate-producing gut bacteria as one study found.¹¹¹ Low-grade inflammation has been identified in some IBS patients although some argue that this may be related to their previous acute enteric infections.^{240,241} Distinct microbial phylotypes were detected in IBS subtypes and control samples where *Ruminococcus torques* was associated with IBS-D and *Ruminococcus bromii* was associated with IBS-C.¹¹² *Ruminococcus torques* is a resident mucin degrader and therefore its enrichment is suggestive of a potential disturbance to the mucus layer in IBS-D patients.²⁴³ The fermentation process of carbohydrates releases gases which can exacerbate luminal distention and bloating typically experienced by IBS-D patients. This unfavourable effect of dietary fibres provides the rationale for recommending the low-FODMAP (fermentable oligosaccharides, disaccharides, monosaccharides, and polyols) diet to some symptomatic patients.^{238,239} However, the long-term effect of reduced fermentable carbohydrate consumption needs to be evaluated given the important health benefits of SCFA.

The accumulation of anaerobic bacteria in the small intestines known as small intestinal bacteria overgrowth (SIBO) has been identified in some IBS patients prompted researchers to question whether the aetiology of IBS may have originated from the small intestines. Patients with SIBO experience similar symptoms such as visceral hypersensitivity, bloating, chronic diarrhoea, abdominal bloating and flatulence which are the characteristics of functional GI disorders.¹¹³ Very recently, it was found that dysbiosis was more apparent in the mucosa-associated microbiota than luminal microbiota in SIBO+ IBS-D patients.¹¹⁴

Evidence in the literature strongly suggest a role for the microbiota in IBS. However, many of these observations are made using animal models and often focusing on a single disease

pathway. Therefore, modelling a disease with such high level of heterogeneity requires large-scale profiling studies in human to first piece together all the possible interactions between the gut microbiome and the host before moving into more detailed mechanistic studies.

1.6 Aims and Objectives

The gut microbiota exists in unique ecological niches inside the GI tract. While their composition and functional capacity along the longitudinal and transverse gut compartments have been investigated, there is a need to explore the metabolic activities associated with these unique microbial communities in humans. Therefore, this thesis aims use of a compartmentalized approach to provide a better understanding of the different metabolic landscapes inside the GI tract and how they may contribute to the pathogenesis of IBS and the implication associated with Western diet. In Chapter 3, the effects of Western diet on the microbiome and the metabolome in the longitudinal gut compartments were explored. Chapter 4 used a multi-omics approach to understand deeper interactions between the microbiota and the host in IBS and the unique metabolic signatures in the transverse gut compartments associated with IBS subtypes. In Chapter 5, the role of the mucus compartment as a potential barrier to SCFA diffusion was investigated in an in vitro experiment. Lastly, the wider implications of the research outcomes and suggestions for future work is provided in Chapter 6.

Chapter 2

Methods

This chapter describes the metabolomics experiments and statistical analysis performed in this thesis. A general overview of microbiome data analysis is also provided.

As discussed in Chapter 1, gut microbiota is intimately connected with the metabolic and physiological functions of the host. To fully appreciate the magnitude and complexity of these interactions, a system biology approach is needed in order to generate hypotheses for further mechanistic investigation. This top-down experimental approach consists of large-scale profiling studies incorporating a variety of omics technologies that aim to capture system level interactions between the host and the microbiota.¹¹⁵

2.1 Metabolomics

Metabolomics is a form of molecular phenotyping where it produces a snapshot of the global metabolic profiles of an individual through biochemical analysis of biological fluids and tissues.¹¹⁶ Genes dictate the biological functions available to a living organism and metabolites reflect the status of these functions and how these biological systems react to environmental factors such as diet, toxins and drugs. Several metabolites essential to human health are exclusively produced by the microbiota. These small biologically active molecules produced within the gut compartment act locally by binding to epithelial receptors and triggering a series of physiological and immune responses.¹¹⁷ Some bacterial exclusive metabolites also enter the blood stream through the gut epithelium and travel to different parts of the body where they have the potential to affect other organ systems. The proportion of microbial metabolites to host-derived metabolites differ between body sites. Therefore, studying the metabolic profiles in different biofluids reveal how the microbiota and the host respond to environmental stimuli and disease on a system level.

2.1.1 Proton Nuclear Magnetic Resonance (^1H NMR) spectroscopy

^1H NMR has been used for the analysis of small molecules in biofluids since the 1980s¹¹⁸. The non-destructive and relatively straight-forward sample preparation procedure coupled with its ability to simultaneously detect multiple metabolites in a single run have made it one of the most popular techniques in metabolic profiling studies.¹¹⁹

The proton nucleus (^1H) in a neutral environment spins in random directions and this spinning motion creates a magnetic moment around the nucleus.¹²⁰ When an external magnetic field is applied, it aligns the nucleus spin in one direction. A nucleus with a half spin can be promoted from a low energy state denoted by $m = +1/2$ to a high energy state $m = -1/2$. When a radio frequency (rf) energy is applied perpendicular to the external magnetic field, it creates oscillations at the Larmour frequency which is the rate of precession of the magnetic moment of the nucleus around the external magnetic field. The nuclei resonate with the rf energy causing the a 'flip' in energy states. Once the rf energy stops, nuclei are relaxed, and they fall back to the lower energy state. The Larmour precession is recorded as a free induction decay (FID) signal which is then Fourier transformed converting the signal from a time domain (FID) to a frequency domain to produce a spectrum. Not all nuclei are promoted to the same energy level. This is caused by electrons around a nucleus that shield it from the applied field which are different depending on the environment of nucleus. The nuclei that are more shielded are promoted to a lower energy state whereas the less shielded nuclei, typically ones that are next an electron-drawing group, are more exposed to the applied magnetic field and therefore are promoted to a higher energy state. This difference in nuclei shielding are transcribed as chemical shifts (δ) or parts per million (ppm) on the NMR spectrum. Nuclei on the lower end (right-side) of the spectrum are more shielded compared to nuclei on the higher end (left-side) of the spectrum. An internal standard, 3-(trimethylsilyl) propanoic acid-2,2,3,3,- d_4 (TSP), is used as a reference signal since TSP dissolved in water produces a single resonance close to 0ppm.²⁴³ The chemical shifts for other compounds are measured as relatives to TSP given by the following formula:

$$\delta \text{ (ppm)} = \frac{\text{Frequency of signal (Hz)} - \text{Frequency of TSP (Hz)}}{\text{Frequency of TSP (Hz)}} \times 10^6$$

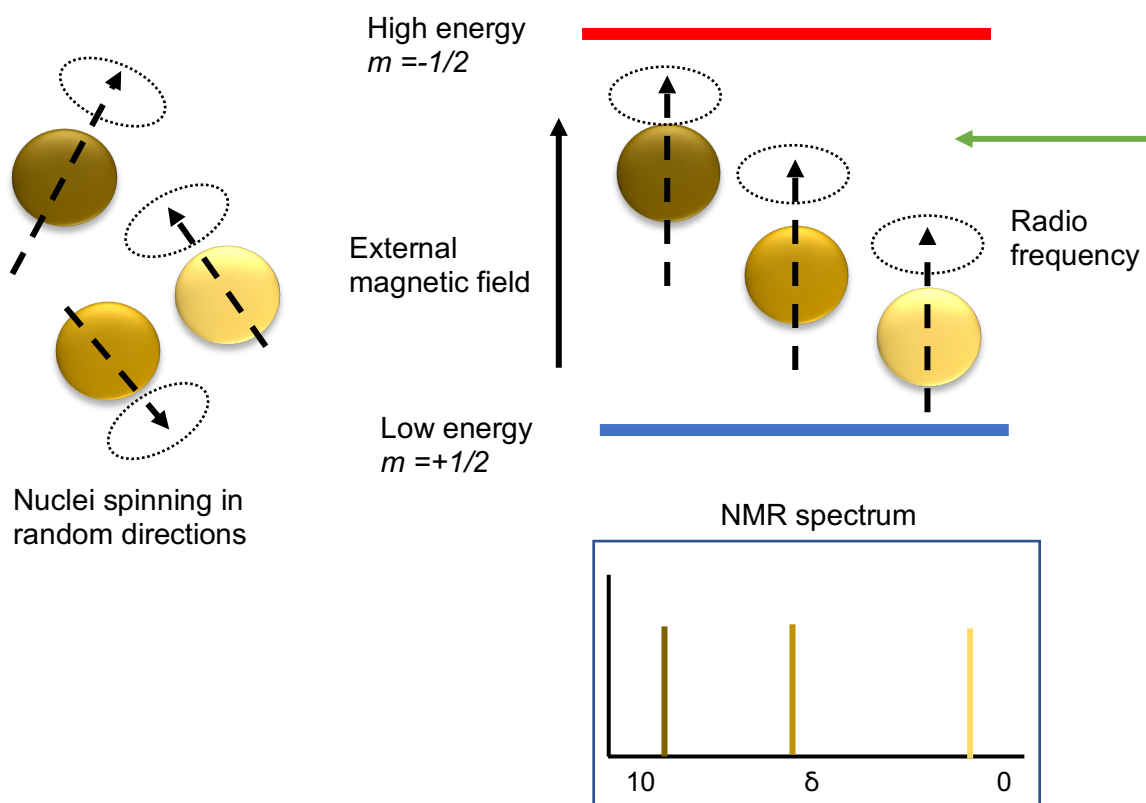
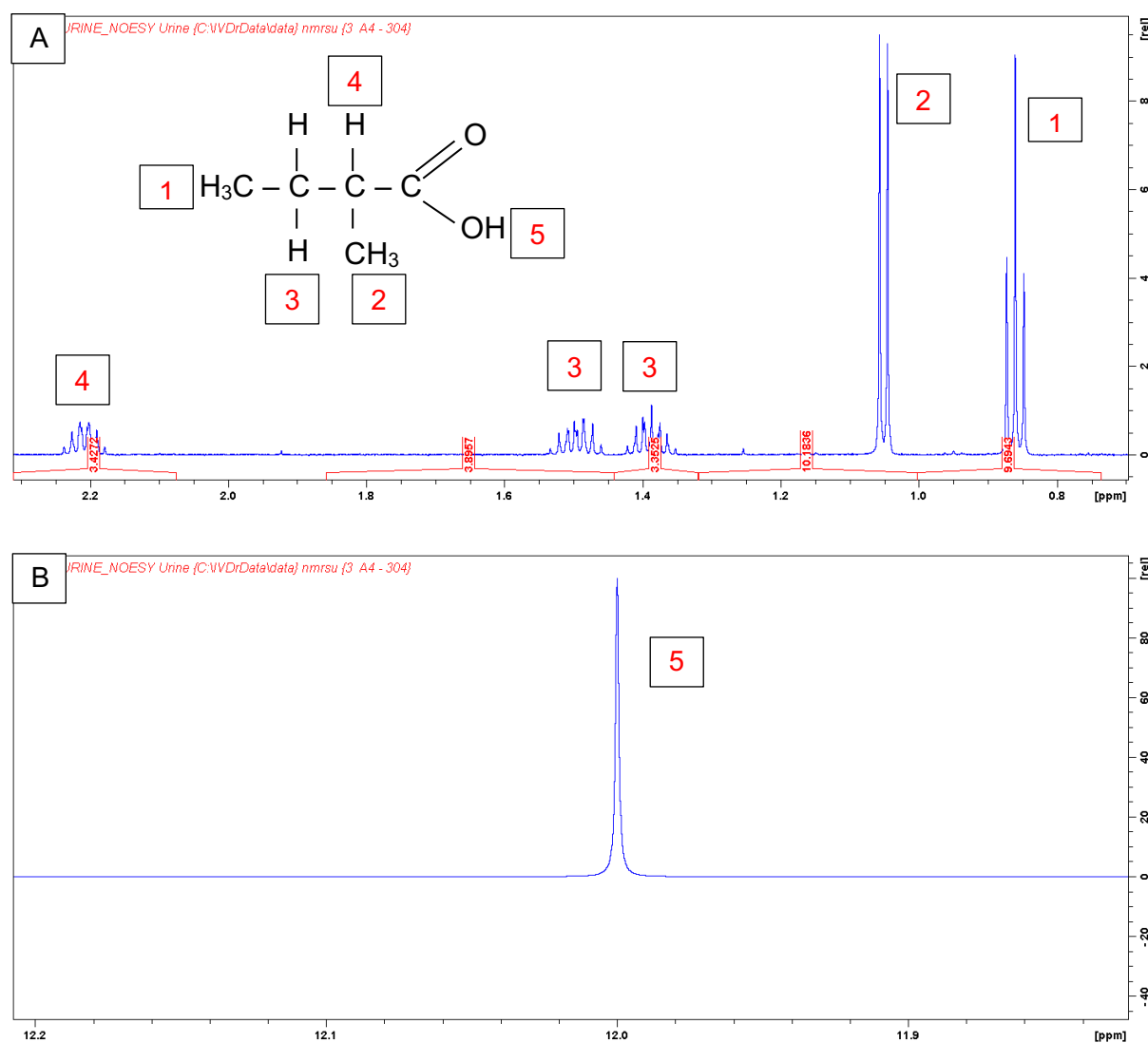


Figure 2.1 NMR theory. The spinning proton nuclei is aligned to the externally applied magnetic field and then promoted to different energy states before relaxing back down by the application of a radio frequency field. The resulting resonance is displayed in an NMR spectrum.

The environment for which the nucleus is in not only determines its chemical shift but also its splitting pattern. Splitting is caused by interactions also known as coupling occurring between neighbouring protons. The spacing between the splits is measured in Hertz (Hz) and termed the *coupling constant*, J . The $n+1$ rule is normally used for calculating the number of splits where n is the number of protons attached to the adjacent carbon atom. For example, looking at the one-dimensional nuclear overhauser enhancement spectroscopy (NOESY) spectrum for 2-methylbutyrate (2mb) (**Figure 2.2A,B**): in environment (1), the adjacent carbon has two protons attached to it, therefore the resulting signal is a triplet and in (2) the adjacent carbon only has one proton therefore the signal is a doublet. Peaks with five peaks and above are referred to as multiplets which is the case for protons (3) and (4). The proton in the carboxylic acid group (5) produces a singlet with a high chemical shift of 12ppm since it is in a strong electron-drawing environment of oxygen atoms. The spectra on display is from a solution of 2-mb standard therefore no other interfering metabolite peaks can be seen. Even with well resolved peaks, the multiplicity of some metabolite signals can be ambiguous in a biological

fluid where peaks overlap often occur. This is where a 2D J-resolved (J-RES) experiment can help (**Figure 2.2C**) as it produces a 2D spectrum with chemical shifts in one dimension and J-couplings (splitting pattern) in the other; the orthogonal display of the spectra allows easy visualisation of peak multiplicity to facilitate metabolite characterisation.



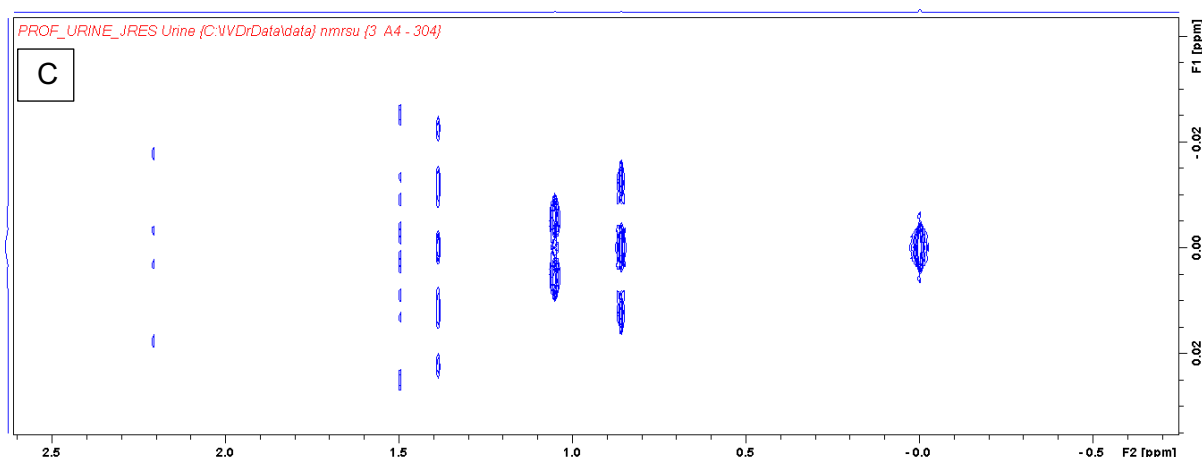


Figure 2.2 NMR spectrum of 2-methylbutyrate from (A-B) one-dimensional NOSEY experiment and (C) two-dimensional J-RES experiment.

2.1.2 ^1H NMR sample preparation

The sample preparation procedures, NMR instrument conditions and data processing workflow were adopted from previously published protocol.¹¹⁹ All samples were kept frozen at -80°C until analysed.

Serum

Serum samples were centrifuged at 4°C at 12000g for 5 min. The supernatant (90 μL) was mixed with 90 μL of plasma buffer before being transferred to a 3mm NMR tube.

The NMR buffer for the serum samples were prepared as follows: 100mL of a 0.075M NaH_2PO_4 buffer was prepared by dissolving 1.064g of NaH_2PO_4 in 80mL of D_2O . 80mg of TSP and 40mg of NaN_3 were dissolved in 4mL of D_2O . Both solutions were mixed by shaking and sonication. pH was adjusted to 7.4 by adding NaOH pellets. The final solution was then transferred to a 100mL volumetric flask to be filled to volume using D_2O .

Duodenal aspirates, urine and faecal samples

Aspirate and urine samples were centrifuged at 4°C at 12,000 g for 5 min. The supernatant of aspirate (162 μL) was mixed with 18 μL of NMR buffer and transferred to a 3 mm NMR tube. The supernatant of urine (540 μL) was mixed with 60 μL of NMR buffer and transferred to a 5mm NMR tube. Faecal samples (~100mg) were pre-weighed, randomised, and emptied into a screw-cap tube containing 50mg of 1.0mm Zirconia beads to which 400 μL of ACN:H $_2$ O (1:3) was added. The tube was placed in a Biospec bead beater for 30 seconds. The homogenised sample was then centrifuged for 20 minutes at 16,000g. The supernatant was carefully

transferred into spin filter tubes and centrifuged for 30 minutes at 16,000g. The filtered faecal water (80µL) was aliquoted into a 96 well plate, 10µL were used for quality control. Filtered faecal water in the 96 well plate was dried down under nitrogen flow before reconstituting with 540µL of D₂O and 60µL of the NMR buffer. The plate was then put on an Eppendorf MixMate plate shaker at 1300 rpm for 5 min. The reconstituted faecal water and buffer mixture was transferred to 5mm NMR experiment tubes.

The following reagents and consumables were purchased from Sigma-Aldrich. Acetonitrile (ACN); LC-MS grade water; 3-(Trimethylsilyl) propionic-2,2,3,3-d₄ acid sodium salt (TSP); Sodium azide (NaN₃); Potassium dihydrogen phosphate (KH₂PO₄); Potassium Hydroxide (KOH); Deuterium Oxide (D₂O). Consumables purchased from elsewhere are: Spin-X 0.22µm Centrifuge Tube Filter (COSTAR®); 96 well microplates – 2mL capacity (COSTAR®); 1.0mm Zirconia beads (BioSpec)

NMR buffer for urine, duodenal aspirates and faecal samples – 100mL of a 1.5M KH₂PO₄ buffer was prepared by dissolving 20.4g of KH₂PO₄ in 80mL of D₂O. 100mg of TSP and 13mg of NaN₃ were dissolved in 6 to 10mL of D₂O. Both solutions were mixed by shaking and sonication. pH was adjusted to 7.4 by adding KOH pellets. The final solution was then transferred to a 100mL volumetric flask to be filled to volume using D₂O.

2.1.3 ¹H NMR experiment conditions

Metabolic profiles were measured on a Bruker 600 MHz spectrometer (Bruker Biospin, Rheinstetten, Germany) set at a constant temperature of 300K for faecal, urine and duodenal aspirates samples and 310K for serum samples. A standard one-dimensional NOESY experiment with water suppression was performed for each sample. In a blood plasma or serum spectrum, there are regions of broad peaks due to the presence of large macromolecules such as lipoproteins. Metabolite peaks in these regions can be resolved using Carr – Purcell – Meiboom – Gill (CPMG) spin – echo sequence to edit peak intensity based on NMR relaxation times to selectively highlight the signals from small molecule metabolites. Therefore, an additional CPMG experiment was performed for each serum sample. For the experiments performed in Chapter 3, a total of 128 scans were acquired per faecal sample and 64 scans per serum, duodenal aspirate and urine sample. For the experiments performed in Chapter 4, a total of 32 scans were acquired per faecal sample and 64 scans per serum sample. Automatic phasing, baseline correction and spectral calibration to TSP (0 ppm) was performed in Topspin 3.1 (Bruker Biospin, Rheinstetten, Germany).

2.1.4 ^1H NMR data pre-processing

Raw data from the NMR experiment were imported into MATLAB (Version 8.3.0.532 R2014a, Mathworks Inc, Natick, MA, USA) for pre-processing. A series of in-house developed scripts were used for the following executions. First, peak corresponds to water and TSP were removed due to their biological irrelevance. During a sequence run, subtle alterations in chemical shifts can occur due to pH variation and other sample-dependant or technical factors and therefore spectral alignment is needed to ensure chemical shifts are comparable across all samples.^{246,247} Spectra were aligned using the recursive segment-wise alignment algorithm (RSPA) where a reference spectrum was first selected, one that is most similar to all other spectra and then alignment was performed in segments and recursively. To account for the difference in sample concentrations, a probabilistic quotient normalization (PQN) was applied to the samples.¹²¹ PQN works by first selecting a reference spectrum where its peak intensity is the median of all samples. Then a quotient or dilution factor for each peak in a sample is calculated by dividing the sample peak intensity by the intensity of the corresponding reference peak. The median quotient which is unique to each sample is used to normalise all peak intensities for that particular sample spectrum. The pre-processed spectral signal and chemical shifts can then be exported for multivariate analysis.

2.1.5 Principal components analysis

Principal components analysis (PCA) is an unsupervised dimensionality-reduction technique widely used in the analysis of high dimensional metabolomics datasets. An NMR experiment of biofluids typically generates tens of thousands of spectral variables which is much larger than the number of observations, therefore it is unfeasible to apply traditional linear regression methods.^{122,123} PCA is useful in this case as it tries to explain the variation of a high dimensional dataset by summarising them into principal components. In principal component (PC) 1, a line is fitted through the data points whereby those that are nearest to the line display the largest variance and thus PC1 captures the largest proportion of variation of a multidimensional dataset. The projection of the next PC is perpendicular to the first PC and therefore accounts for the second largest variance of the dataset. In every PC, a score is calculated for each row of observation which is a linear combination of variables for that particular row and the direction vector of the PC projection. The direction vector is also known as loadings whereby variables contribute the most to the direction of the PC projection is given a higher loading score. Plotting the PC scores for each observation whilst highlighting their class membership allows easy visualisation of group clustering as well as spotting anomalies.

2.1.6 Orthogonal Projection to Latent Structures-Discriminant Analysis (OPLS-DA)

Orthogonal projections to latent structures analysis-discriminant analysis (OPLS-DA) is a supervised dimensionality reduction technique widely used for identifying class discriminating features in metabolomics data.¹²⁴ In OPLS, variations in X that are uncorrelated with Y is removed from the predictive component using Orthogonal-Signal-Correction (OSC). The OPLS model uses a Monte Carlo cross-validation (MCCV) procedure based on a previously published method by Posma *et al.*,¹²⁵ which randomly splits the data into five parts and builds a model using four parts of the data while leaving one part out to assess how well the model can predict the outcomes/class of the data that was left out. The MCCV continues for 10 rounds which generates a total of 50 iterations, i.e. the number of partitions of the data. For each model a score of its predictability (Q^2Y) and the proportion of variation in Y explained by the model (R^2Y) are calculated. To ensure that the prediction does not happen at random a 100 rounds permutation test was carried out where 100 models were built using randomly permuted outcome labels. A correlation line is then drawn between the R^2Y and Q^2Y of models built using the permuted outcomes labels and one with the correct outcome labels and the model is robust if there is no correlation between the two.

To determine variable importance, regression coefficient (β) for outcome Y was recalculated 25 times for each model by bootstrap resampling. The mean model β and variance of the bootstrap β s can then be used to calculate p values for each spectral variable. The p values are corrected for multiple hypothesis testing using the Benjamini-Hochberg method for calculating False Discovery Rate.¹²⁶ An FDR cut-off of 0.05 was used to identify class discriminating variables.¹²⁵

2.1.7 ^1H NMR metabolite assignment

To improve confidence of metabolite assignment, the correlation between peaks arising from the same metabolite can be assessed using the Statistical Total Correlation Spectroscopy (STOCSY).¹²⁷ Peaks belonging to the same metabolites/compound should be highly correlated, although overlapping peaks from other molecules may adversely affect this correlation. First, a spectral signal of interest is selected, STOCSY then compute a correlation matrix based on the intensities of all the spectral signal in the matrix in relation to the signal of interest. The correlation matrix is then projected onto a 1D spectra where highly correlated signals can be identified together. An in-house peak integration script was applied to calculate peak integrals for assigned metabolite features.

2.1.8 Liquid chromatography-Quadrupole Time-of-Flight Mass Spectrometry for bile acids profiling

Liquid chromatography-Quadrupole Time-of-Flight Mass Spectrometry (LC-QTOF) is a highly sensitive and selective mass spectrometry technique ideal for measuring metabolites that have similar nominal masses such as bile acids (BA).^{128,129}

LC-QTOF has a reverse liquid chromatography phase where solvent of increasing hydrophobicity is used to separate molecules based on their affinity with the liquid phase (the solvent mixture) and the hydrophobic stationary phase of the LC column. Therefore, in this experiment, polar BA elude first due to their higher affinity with the more polar liquid phase followed by more hydrophobic BA. This difference in elution order is displayed as retention time on the chromatogram. After exiting the chromatography column, BA in solution arrive at the high voltage capillary where they are ionized by electrospray ionization operated at negative mode which removes a proton from the bile acids molecule. The negatively charged molecules are sprayed out of the capillary cone into droplets. A nebulizing gas is applied to help evaporate the solvent around the molecule. After solvent has been evaporated, the negatively charge BA are now in a gaseous state and ready for entering the vacuumed QTOF analyser. First part of the QTOF consists of a triple quadrupole system. Each quadrupole is made of four parallel cylinder rods. Ions are guided through the quadrupoles by a radio frequency. Since this is a profiling experiment, BA molecules do not need to be fragmented therefore no collision energy is applied to the quadrupoles. After leaving the quadrupoles, ions are re-accelerated to the required energy and ion optics are used to focus their path into the TOF analyser. Upon entering the TOF, a pulse electric field is applied which forces the ions down an orthogonal trajectory to their original direction. The ions are now in a field-free drift space where TOF separation can occur. Different energy is applied to the ion mirror and the detector to allow the ions to travel at their own speed in the two orthogonal paths. The time-to-digital converter (TDC) is used to record arrival time of the ions. The TOF analyser can differentiate ions by the time of their arrival at the TDC and their mass to charge ratio therefore offers a much higher mass resolution than a conventional quadrupole detector.

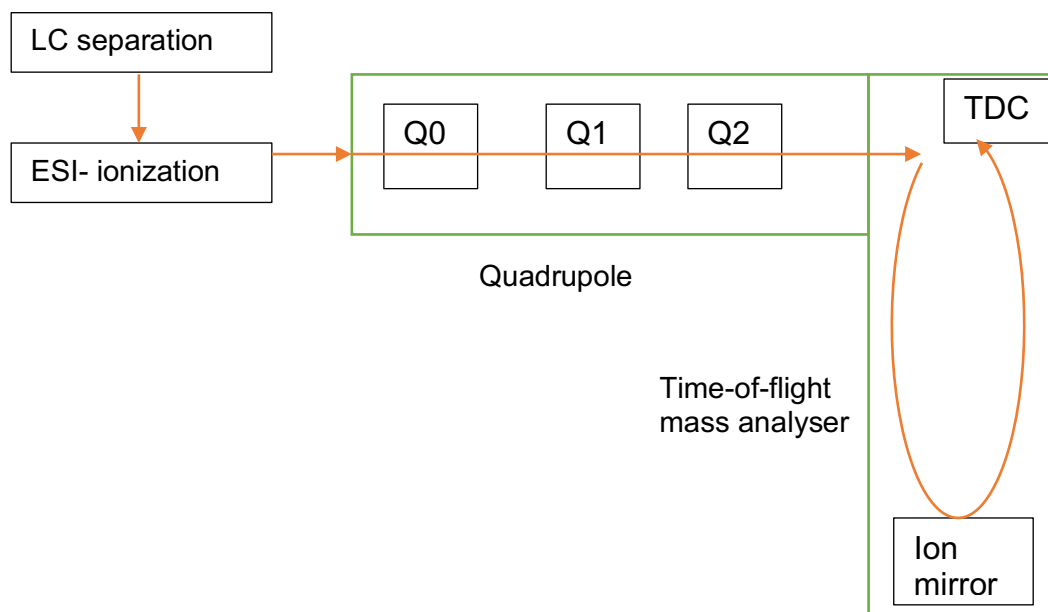


Figure 2.3 Outline of a LC-QTOF system.

2.1.9 LC-MS bile acid profiling experiment

This protocol was adapted from Sarafian *et al.*¹³⁰ Separate aliquots of the faecal water in 25% ACN extracted for ^1H NMR profiling were transferred into 96-well plates (Costar® 3960). Samples were analysed by ACQUITY ultra-performance liquid-chromatography (UPLC) (Waters Ltd., UK) coupled to a Xevo G2-S Q-ToF mass spectrometer (Waters Ltd., UK) equipped with an electrospray ionization source operating in negative mode. Mass spectrometry parameters were as follows: capillary voltage was set at 1.5 kV, cone voltage at 60 V, source temperature at 150 °C, desolvation temperature at 600 °C, desolvation gas flow at 1000 L/h, and cone gas flow at 150 L/h. A reversed-phase column ACQUITY BEH C8 column (1.7 μm , 100 mm $\times$ 2.1 mm) was used at an operating temperature of 60 °C. The aqueous part of the mobile phase includes 1 mM ammonia acetate in ultra-pure water, pH 4.15. The organic mobile phase consists of isopropanol and acetonitrile in a 1:1 ratio. Chromatographic gradient is outlined in figure S-3 Sarafian *et al.* A standard mix of 58 bile acids was kindly provided by Dr Alexandros Pechlivanis and used to make spiked quality control samples for identification purpose. The same bile acids are listed in the supplementary figure of Sarafian *et al.*¹³⁰

2.1.10 LC-MS bile acid data processing

Data acquired from the LC-MS experiment was imported into MassLynx (Waters Ltd.) and processed using method files containing information on mass-to-charge ratio and retention time of the bile acids listed in **Table 2.1**. Manual inspection on each processed sample file was carried out to ensure bile acid peaks have been correctly integrated. The extracted peak integral was then normalized to the total ion current (TIC) of each sample.

Bile acid	m/z	Retention time (minutes)
Cholic acid	407.29	8.80
Chenodeoxycholic acid	391.29	10.46
Deoxycholic acid	391.29	10.57
Lithocholic acid	375.29	10.92
Ursodeoxycholic acid	391.29	8.60
Glycocholic acid	464.32	5.98
Glychenodeoxycholic acid	448.32	7.85
Taurocholic acid	514.30	5.45
Taurochenodeoxycholic acid	498.31	7.40

Table 2.1. Mass to charge ratio (m/z) and retention time for bile acids identified in the LC-MS profiling experiment.

2.1.11 Gas chromatography-triple quadrupole mass spectrometry for SCFA analysis

Gas chromatography-mass spectrometry is used for the separation and quantification of volatile compounds such as SCFA.^{131,132} GC-TQ is comprised of an oven containing the GC column and a triple quadrupole (TQ) mass analyser. The same separation principle of LC applies to the GC, but the molecules are eluted from the column by a carrier gas instead of liquid solvent. A temperature programme is used to separate the molecules based on the volatility. The more volatile the molecule the easier it is to enter gas phase and thus the quicker it leaves the column. After exiting the column, molecules are ionized by electron impact (EI) ionization where electrons emitting from a high voltage filament bombard the molecule and such impact knocks off an electron from the molecule thus it now carries a positive charge. Inside the triple quad system, Q1 and Q3 are supplied with a direct current and a radio frequency and the collision cell Q2 is controlled by radio frequency only. SCFA are quantified using multiple reaction monitoring where precursor ions are selected in Q1 and a collision energy is applied to fragment the precursor ion in Q2 and finally in Q3 a targeted product ion is selected. Since ten SCFA were analysed, there were ten unique pairs of precursor and

product ions. Only ions of the same m/z are detected by the electron multiplier at any one time therefore increases the selectivity of the compound.

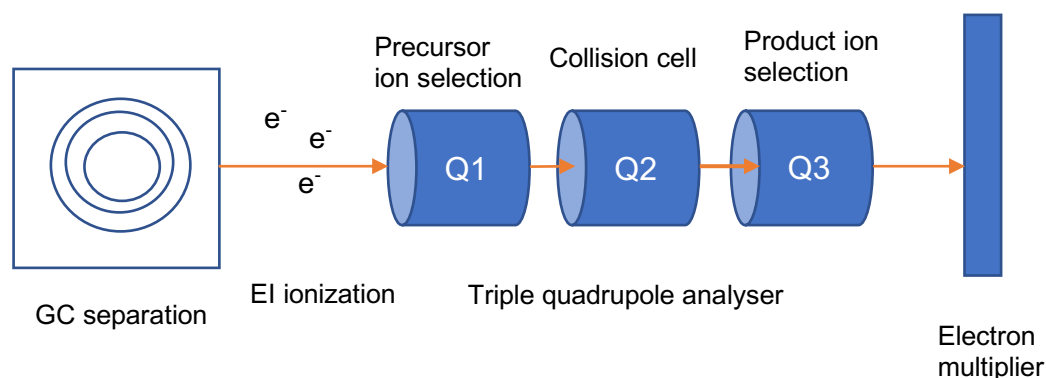


Figure 2.4 Overview of a GC-TQ system.

2.1.12 GC-MS SCFA quantification experiment

The experimental procedures were adopted from an in-house method (not yet published) and a previously published protocol by García-Villalba et al.¹³³ Sigmoid colon biopsies were stored at -80 °C until analysed. Biopsies were weighed and then transferred to a screw-cap tube followed by the addition of 5 x 1.0mm Zirconia beads and 100 µL of LC-MS grade water. The samples were homogenized using a Biospec bead-beater for two cycles at 30 seconds per cycle. SCFA in the homogenized tissue were then extracted using methyl tert-butyl ether (MTBE) (Sigma) and derivatized with *N*-tert-butyldimethylsilyl-*N*-methyltrifluoroacetamide with 1 % tert-butyldimethylchloro-silane (MTBSTF + 1% TBDMSCI). An eleven-point calibration curve was prepared with SCFA standards and a pooled biopsy QC sample was prepared in the same manner. The derivatized samples, calibration curve and QC were analysed on a 7890B GC coupled to a 7000D Triple-Quadrupole mass spectrometer (Agilent Technologies, Palo Alto, CA, USA), using Chemstation software (Hewlett–Packard, Palo Alto, CA, USA). The GC was equipped with an automatic liquid sampler (MPS2) (Gerstel, Mülheim, Germany) fitted with a high polarity, polyethylene glycol (PEG), fused silica capillary column DB-WAXetr (30 m, 0.25 mm id, 0.25 µm film thickness) was used for chromatographic separation. Helium was used as the carrier gas at 1 ml/min and injection was made in splitless mode with an injection volume of 1 µL and an injector temperature of 250 °C. A glass liner with a glass wool plug at the lower end of the liner was used to avoid the contamination of the GC column with non-volatile fecal material. For every ten faecal samples injected, a solvent wash of MTBE was

injected to check for carry-over. The column temperature gradient was as follows; start temp 90 °C, increasing to 150 °C at 15 °C/min, 170°C at 5 °C/min, 250 °C at 20 °C/min, finishing at 250 °C for 2 min (total time 14 min). Solvent delay was 3.5 min. The MS detector was operated in electron impact ionization mode (electron energy 70 eV), scanning the 30–250 m/z range. The temperature settings were ion source 230 °C, quadrupole 150 °C, and interface 280 °C. The SCFA quantified were acetic acid, butyric acid, propionic acid, isobutyric acids, caproic acid, lactic acid, isovaleric acid, 2-methylbutyric acid, valeric acid and 2-hydroxybutyric acid. All standards were purchased from Sigma-Aldrich, UK.

2.1.13 GC-MS SCFA quantification data processing

Data files were imported into MassHunter Workstation Software Quantitation Analysis for QQQ version B.07.01 (Agilent Technologies Ltd.). A data processing method consists of MRM, retention times and standard curve concentrations was constructed to allow the software to automatically calculate the concentration (µM) for individual SCFA in each sample.

2.2 Microbiome

2.2.1 16S sequencing

16S rRNA gene sequencing is a type of targeted amplicon analysis where a marker gene in this case 16S ribosomal RNA gene is selected for amplification using polymerase chain reaction (PCR) technologies. 16S ribosomal RNA gene encodes the component of the small 30S ribosomal subunit found in all prokaryotes. Within 16S gene there are highly conserved regions which are the same across all bacteria, therefore ideal for designing primers that target a wide range of bacterial groups. While the hyper variable regions in 16S that have changed with evolution can be used to distinguish between bacteria from different taxonomy groups. Bacteria groups that showed more than 97% similarities in their 16S gene sequence variants are clustered together and the number of sequence reads in a cluster is measured which is recorded as Operational Taxonomical Units (OTU).²⁴⁴

2.2.2 Shotgun metagenomics sequencing

In large scale microbiome studies, other than community composition, researchers are also interested in the functional capacity of these microbial communities which can be achieved by shotgun metagenomics sequencing. Instead of targeting a specific gene like in 16S rRNA

gene sequencing, in metagenomics all DNA in a given sample is independently sequenced. The resulting sequences are then mapped to genomic loci which include regions that give taxonomical information such as the 16S gene and also regions that encode a particular biological function. While the in-depth information provided by shotgun metagenomics is useful for identifying novel microbial species and gain insights into gene function, it can be expensive to perform for a large-scale study.¹³⁴

2.2.3 Microbiome sequencing and alignment

DNA extraction and sequencing were performed at the University of Minnesota Genomics Center (UMGC). DNA was extracted from stool and biopsy sections using the QIAGEN PowerSoil kit (QIAGEN, Germantown, MD, USA) and was quantified using a NanoDrop-8000 UV-Vis Spectrophotometer (Thermo Scientific, Wilmington, DE, USA) and PicoGreen assays. Shotgun metagenomic sequencing library preparation for stool samples was completed using a modified NexteraXT protocol followed by sequencing on a HiSeq 2500 (Rapid Mode) with 100 bp single-end reads (1x100) or on a NextSeq with 150 bp single-end reads (1x150). To minimize sequencing batch effects samples from the three cohorts were assigned randomly to two sequencing runs. Shotgun reads were trimmed to a maximum of 100 bp prior to alignment. Shotgun sequences were aligned to the RefSeq representative prokaryotic genome collection (release 86) at 97% identity with BURST using default settings.²⁴⁸ The generated alignment table was filtered by dropping samples with low depth (< 10,000 reads per sample). Functional profiling of the shotgun sequencing data was completed using the KEGG Orthology group annotations for RefSeq-derived genes from direct alignment. KEGG Orthology profiles were also predicted from reference genomes and the predicted profiles were augmented to improve the estimates of low-abundance genes using SHOGUN (<https://github.com/knights-lab/SHOGUN>).

Biopsy samples were sequenced via amplification of the V4 region of the 16S ribosomal RNA gene,²⁴⁹ followed by paired-end 2x250 bp sequencing on an Illumina MiSeq. Adapters were trimmed and low-quality reads (< 25 Q-score) were dropped using Shi7 (Al-Ghalith et al., 2018). Amplicon reads were stitched also using Shi7.²⁵⁰ Amplicon sequences were aligned to the 16S rRNA genes from the same bacterial genomes in the shotgun sequencing approach using BURST,²⁴⁸ with the same setting as above.

2.2.4 Microbiome data analysis

Data analysis were carried out in R version 4.0.2 (R Foundation for Statistical Computing, Vienna, Austria). OTU abundance was normalised using Total Sum Scaling (TSS) where abundance data for each species was converted into a value between zero and one to reflect their proportion in a sample. Species which were present in less than 10% of the total samples and with a relative abundance of less than 0.00005 (0.005%) were removed by using the *taxa.filter* function in the *metamicrobiomeR* (version 1.1) package.¹³⁵ The dedicated microbiome analysis package *phyloseq* (version 1.32.0) was used for alpha-diversity calculation and for collapsing OTU table to different phylogenetic levels.¹³⁶ The *vegan* package (version 2.5-6)¹³⁷ was used to calculate beta-diversity. A list of significantly different OTUs between cohorts with FDR correction ($q < 0.05$) between cohorts was calculated using the *simper_pretty.R* and *R_krusk.R* scripts developed by Andrew J. Steinberger (<https://doi.org/10.5281/zenodo.1458243>).¹³⁸ *simper_pretty.R* is an upgraded version of the *SIMPER* function in *phyloseq* which identifies all OTUs that are most variable in all pairwise comparison groups at once. The *R_krusk.R* script takes the output file of *simper_pretty.R* and performs the non-parametric Kruskal-Wallis rank-sum test on each OTU in the output file.

Microbiome alpha-diversity was measured by Shannon evenness index and Chao1 richness score.^{139,140} Sample A has more species varieties, so it has a higher Chao1 score. Sample B has a higher Shannon score because the species abundance is more evenly distributed.

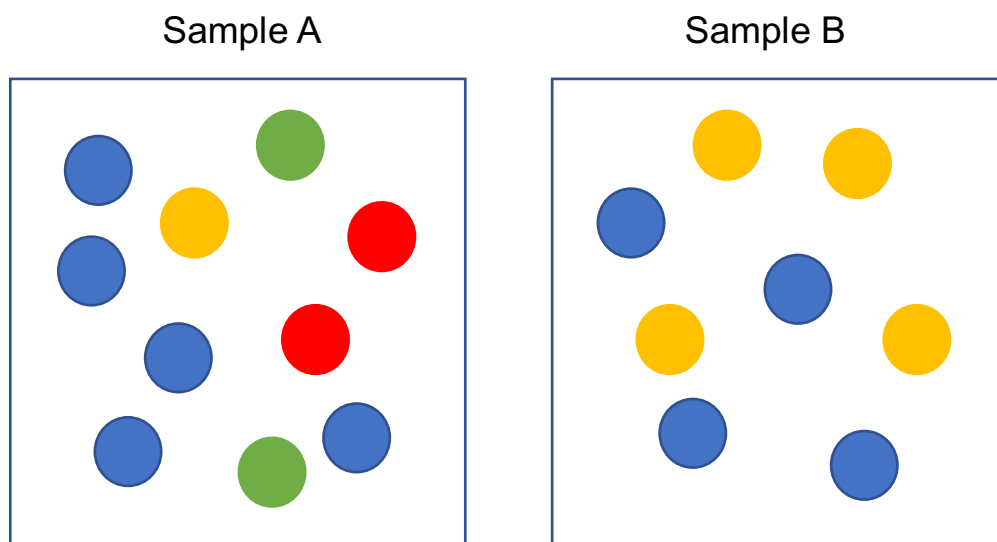


Figure 2.5 Comparing alpha-diversity between two samples. Sample A is represented by a richer community whereas sample B has a more balanced community.

Bray-Curtis dissimilarity was developed by J. Roger Bray and John T. Curtis in 1957.¹⁴¹ It measures the differences in community abundance between different sites. The dissimilarity matrix was presented in a 2-dimensional space using Non-metric Multidimensional Scaling (NMDS). NMDS ranks objects based on their distance from each other so it mitigates the problem associated with using absolute abundance such as sensitivity to normalization.

2.3 Multi-omics integration

Metabolomics data were correlated with the microbiome data using the *corr.test* function in the *psych* (version 2.0.7) package in R version 4.0.2 (R Foundation for Statistical Computing, Vienna, Austria). Spearman correlation was used. For each correlation, correlation coefficient (r) and FDR corrected P value corrected were calculated. Outputs of the correlation analysis were visualised using Cytoscape 3.8.0.¹⁴²

Chapter 3

Western diet and the small intestinal microbiota

3.1 Introduction

The rapid industrialization of the Western world has had a profound impact on the dietary habits of its population. Human societies in these industrialised parts of the world are increasingly consuming foods that have been ultra-processed to enhance their flavour and texture which are high in simple sugar, fats and low in fibre. These foods are the main constituents of a Western diet (WD).^{85,147} These diets have been implicated as one of the leading causes of obesity, type 2 diabetes and even cancer are the biggest public health issues faced by the Western world.^{85,147} Microbial dysbiosis induced by Western diet can have serious implication on host health as discussed in Chapter 1.^{148,151} While much of our understanding of WD-induced effects on the human gut microbiome came from studying stool samples, not much is known about the effect on the small intestinal (SI) microbiome and their metabolic activities. The SI harbours a distinct microbial community that is highly adapted to metabolise carbohydrates; therefore, they may play an important role in modulating nutrient availability and absorption in the SI. A study by Angelaskis *et al.* found the duodenal microbiota of obese individuals exhibit alterations in fatty acid and sucrose breakdown pathways.¹⁴⁹ A recent study comparing rats fed with standard diet and high fat/high sugar diets found that SI microbiome community composition varied more than the fecal microbiome where the SI phyla *Bacterioides* and *Allobaculum* were negatively correlated with weight gain.¹⁵²

Plant cell walls consist mainly of fibres that are not well absorbed thus they flow down to the colon to be broken down by the fibre fermenting gut bacteria. However, modern food processes such as milling wheat grains into flour and extracting starch from seed crops destroy the plant cell wall and rendering all nutrients acellular.¹⁵⁵ When foods made with these ingredients are ingested, nutrients that used to be contained within the plant cell walls are now becoming available in the SI which can drive the growth of potentially pathogenic bacteria.^{156,157} Furthermore, research have shown that there is a limit to how much sugar can be absorbed in the human SI meaning that more can be used by the gut bacteria driving the expansion in growth.¹⁵⁸ One study in mice receiving WD observed a bloom of bacteria in the distal colon capable of metabolising simple sugars that had not been seen in this region previously.¹⁵⁹ Given that diet can aggravate the symptoms of functional GI disorders typically

seen in SIBO patients this further highlights the importance of understanding the interaction between diet and the SI microbiome.¹⁶⁰

3.2 Aims and Objectives

Studies investigate the impact of WD on the SI microbiome have mostly been done on animal models and data from human intervention are lacking.

This chapter aims to characterise the microbial and metabolic alterations induced by Western diet in both the human small and large intestine as well as its impact on the systemic host metabolic system. The comparison between the SI and the colonic microbiome and metabolome offered a compartmental view of host-microbial interactions in the gut. Associations between metabolic variations and GI symptoms were also performed.

3.3 Methods

This was a pilot prospective single-centre dietary intervention study in healthy volunteers consuming a high-fibre baseline diet. The study was approved by the Mayo Clinic Institutional Review Board and was registered with clinicaltrials.gov under NCT03266536. The Mayo clinical team was responsible for recruiting the participants, conducting clinical assessments and collecting dietary information from them. 16S rRNA gene sequencing and the associated data processing were performed by researchers at the University of Minnesota. In addition, pre and post-intervention samples were tested for small intestinal bacterial overgrowth (SIBO) at the Mayo Clinic laboratory. SIBO is diagnosed based on the presence of $\geq 1 \times 10^5$ bacteria [colony-forming units (cfu)] per ml of duodenum aspirates.¹⁶⁰

3.3.1 Participant recruitment

All subjects were asked to record their typical diet for the past 1-3 months by completing a food-frequency questionnaire (FFQ).¹⁶¹ Healthy adults above 18 years of age consuming a typical baseline diet that consists of: fibre intake ≥ 11 g/1000 calories/day; $< 10\%$ daily calories from added sugar; ≥ 5 servings of fruits and vegetables/day; and $\leq 13\%$ daily calories from saturated fat based on FFQ were recruited. However, participants were excluded if they had any of the following: a known diagnosis of inflammatory bowel disease, microscopic colitis, celiac disease or other inflammatory conditions, presence of abdominal symptoms based on baseline questionnaire, oral antibiotic or probiotic use within the past 4 weeks, pregnancy or plans to become pregnant within the study time frame, or any other disease(s), condition(s), or habit(s) that would interfere with completion of study.

3.3.2 Sample and clinical data collection

At the pre and post-intervention visit, participants underwent esophagogastroduodenoscopy (EGD) with conscious sedation. Participants were in a fasting state for at least 8 hours prior to the procedure. Duodenal aspirates were obtained via a standard suction catheter passed through the suction port with minimal inflation in the stomach; an aliquot was submitted to a clinical laboratory for testing for small intestinal bacterial overgrowth (SIBO) and the remaining sample was stored for microbiome and metabolomics analysis at -80°C . At both visits, participants completed a symptom and demographic questionnaire and provided a stool and a urine sample. A plasma sample was also taken. The following symptoms were assessed using questionnaires: stool frequency, straining, incomplete evacuation, hard/lumpy stools, abdominal pain associated with bowel movements, diarrhoea/loose-watery stools, bloating, nausea/vomiting, heartburn, fatigue, and appetite. Answers were recorded in a binary (yes/no) fashion.

3.3.3 Seven-day Western diet intervention

All participants consumed a seven-day standardized diet with typical United States macronutrient calorie distribution: 50% from carbohydrates, 35% from fats, and 15% from protein. The diet was low in fibre ($< 10\text{ g}/1000\text{ calories/day}$) and high in simple sugar ($\geq 50\%$ of daily carbohydrates). Dietary instructions and food choices were discussed at the pre-intervention visit by a licensed dietitian. A summary of the dietary consumption presented as grams per nutrient group (protein, fat, carbohydrates, fibre and sugar) and total calories was calculated by the dietitian based on participants' pre-intervention FFQ and their food choices from the seven-day WD.

3.4 Results

3.4.1 All participants developed symptoms during the Western diet intervention.

The study recruited 16 participants in total of which 7 were male and 9 were female. Their average age was 28 years-old (+/- 5.4 standard deviation (sd)) with an average BMI of 24.7 (+/- 4.7 sd). All participants developed at least one GI symptoms during the WD intervention where fatigues, feeling of hunger and bloating were among the most common symptoms as shown in **Table 3.1**.

Symptoms	Pre-intervention		Post-intervention	
	Yes	No	Yes	No
Duodenum aspirates SIBO positive	8 (50%)	8 (50%)	6 (37.5%)	10 (62.5%)
≤2 bowel movements/week	0	16 (100%)	4 (25%)	12 (75%)
Strain/incomplete bowel evacuation >25%	0	16 (100%)	8 (50%)	8 (50%)
Lumpy stool >25%	0	16 (100%)	7 (43.75%)	9 (56.25%)
Abdominal pain/discomfort relieved by defecation	0	16 (100%)	4 (25%)	12 (75%)
≤3 bowel movements/week	0	16 (100%)	0	16 (100%)
Loose stool	0	16 (100%)	1 (6.25%)	15 (93.75%)
Bloating	1 (6.25%)	15 (93.75%)	8 (50%)	8 (50%)
Postprandial upper abdominal pain >1/meal	1 (6.25%)	15 (93.75%)	4 (25%)	12 (75%)
Postprandial abdominal bloating	1 (6.25%)	15 (93.75%)	7 (43.75%)	9 (56.25%)
Nausea/vomiting >1/meal	1 (6.25%)	15 (93.75%)	1 (6.25%)	15 (93.75%)
Heartburn >1/week	1 (6.25%)	15 (93.75%)	2 (12.5%)	14 (87.5%)
Post-intervention fatigue/decreased energy	0	16 (100%)	9 (56.25%)	7 (43.75%)
Post-intervention hunger	0	16 (100%)	9 (56.25%)	7 (43.75%)

Table 3.1 Summary of symptom questionnaires. Number and percentage of participants who experienced symptoms pre and post-intervention.

A break-down of the main dietary components and calories consumed on a daily basis before and during the intervention is presented in **Figure 3.1**. Participants on the whole consumed more protein, fat, carbohydrates during the trial and therefore contributed to an overall increase in daily calories intakes.

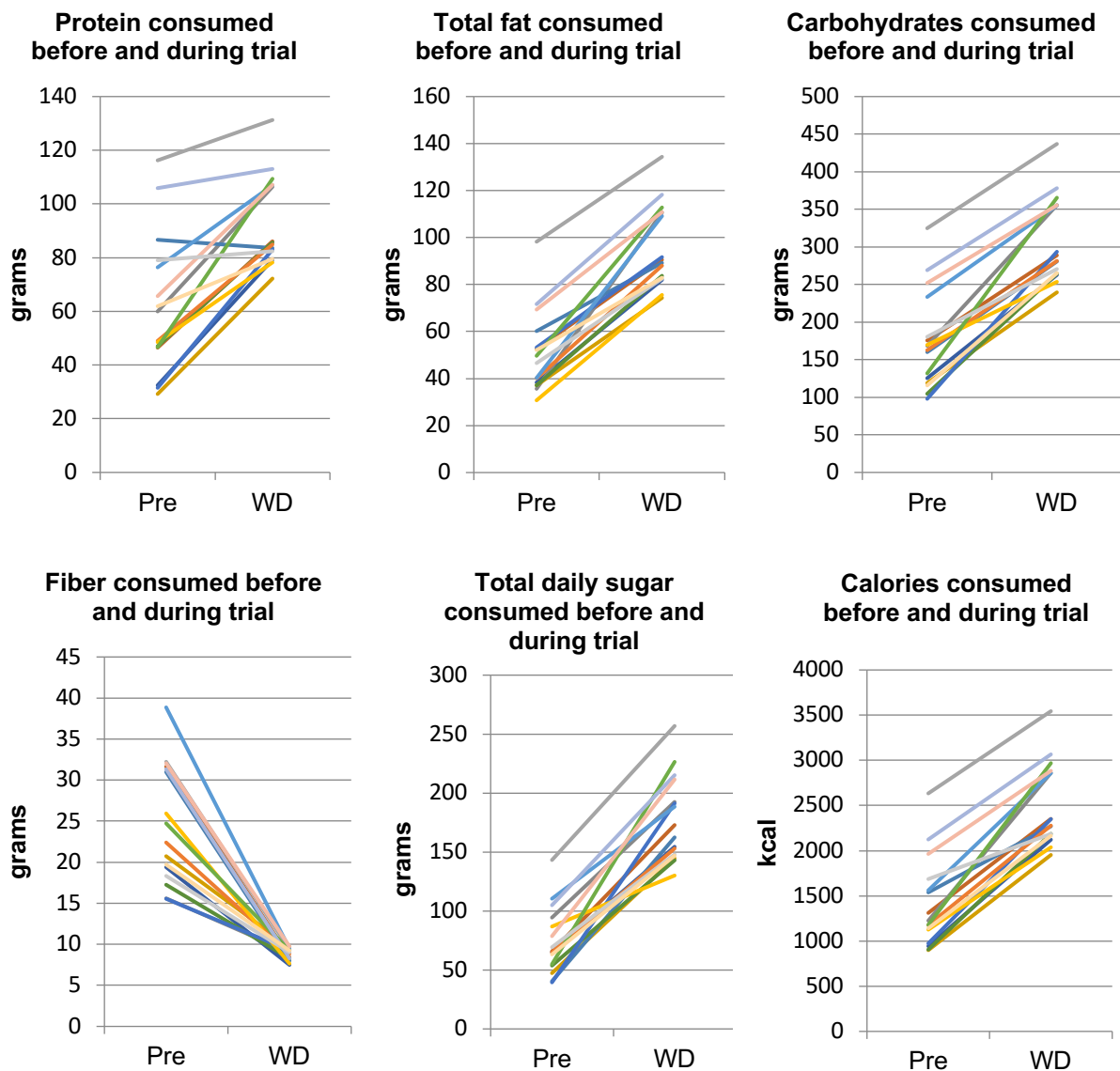


Figure 3.1 Comparing the daily consumption of the five major nutrient groups and daily calories intake pre-intervention and during the western diet (WD) trial based on participant's FFQ. Each coloured line represents a single participant.

3.4.2 No metabolic differences were detected in the duodenum aspirate sample between high fibre and WD diet

^1H NMR global metabolic profiles obtained for duodenal aspirates were compared using PCA. As shown in **Figure 3.2A, B**, no clear group clustering can be seen for the two diet. Two outliers were detected as they are outside the ellipse showing 95% confidence limit for Hotelling's T^2 where T^2 is the distance from the centre of the plane to the projection of the observation onto the plane. However, even after removing the outliers from the PCA model (**Figure 3.2B**), there were still no clear group separation.

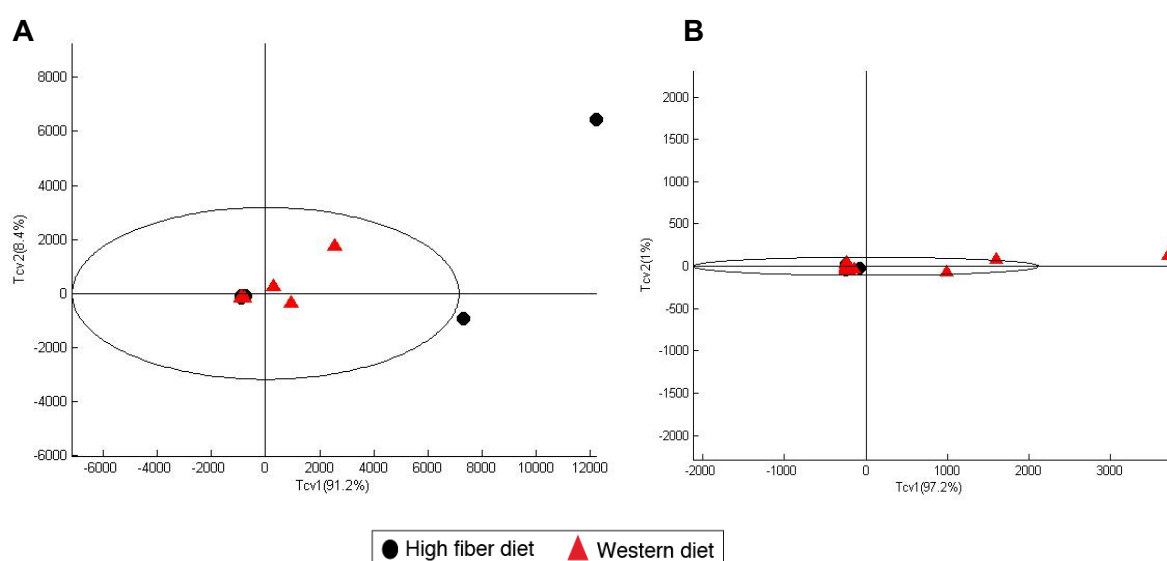


Figure 3.2. Comparing metabolites in the duodenum aspirate of high fibre versus WD diet. A) Cross-validated scores plot from the PCA model built using 600 MHz ^1H NMR spectral variables showing Hotelling T^2 scores in the first and the second principal component and 95% confidence ellipse. B) Same as A) but with the two high fibre diet outlier samples removed.

3.4.3 Stool SCFA could discriminate between high fibre and WD diet

In contrast to the metabolic profiles in the duodenal aspirates, clearer group separation could be seen in the PCA scores plot of stool metabolites where SCFA were observed to change (**Figure 3.3A-B**). When peak integrals were extracted for paired-wise analysis, SCFA and lysine were found to be significantly reduced in WD stool samples compared to pre-diet samples. This result was not surprising as there was less fibre available for the bacteria in the large intestines to ferment as part of the WD diet.

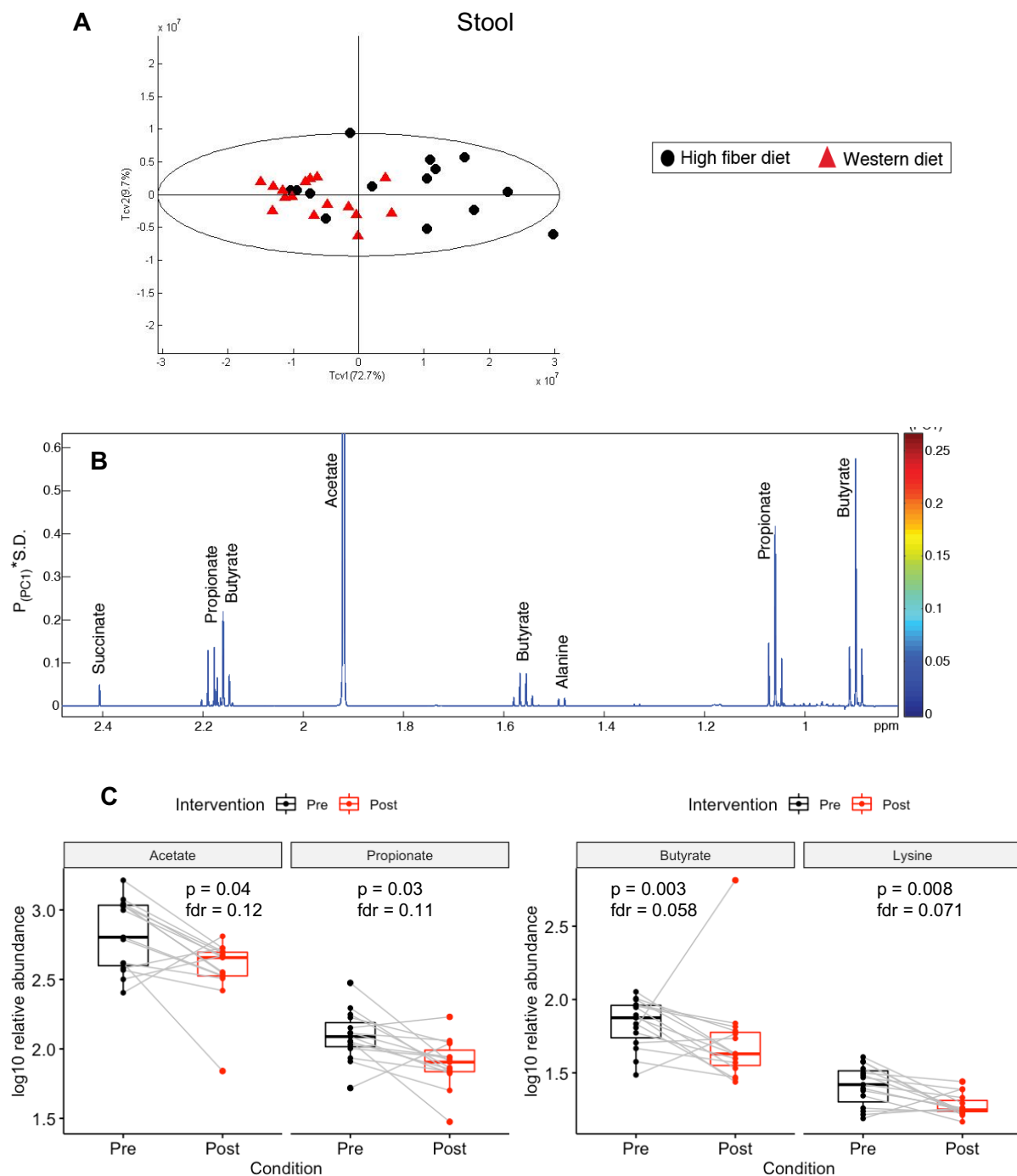


Figure 3.3. Faecal short chain fatty acids and lysine are reduced following consumption of the western diet. A) Cross-validated scores plot from the PCA model built using 600 MHz ^1H NMR spectral variables. B) PCA loading plot for PC1 showing metabolites that contribute to separation in the scores plot. Positive peaks indicate metabolites positively correlated with PC1 (higher in higher fibre samples) and the peaks are coloured by their PC1 loading. C) Boxplots showing log₁₀ relative abundance of metabolite peak integrals for acetate, propionate, butyrate and lysine. (paired-Wilcoxon test, *P* value and FDR (Benjamini-Hochberg) corrected *P* value displayed.)

3.4.4 Distinct metabolic profiles in the duodenum and the colon

The duodenum lumen has a distinct metabolic profile to the colonic lumen, reflected by the faecal samples, as shown in **Figure 3.4**. Here, sections of the NMR spectrum of QC samples from duodenal aspirates and stool were compared. The duodenal lumen (blue spectrum) was dominated by amino acids, including branched-chain amino acids (BCAA) and glucose whereas the colonic lumen (red spectrum) was dominated by the presence of SCFA which dwarfs the signals of amino acids and glucose. The stool metabolic profile reflects the higher abundance of fibre fermenting bacteria inside the colonic lumen. Branched-chain fatty acids (BCFA) such as 2-methylbutyrate (2mb) can also be detected in the stool indicating the presence of bacteria protein fermentation.

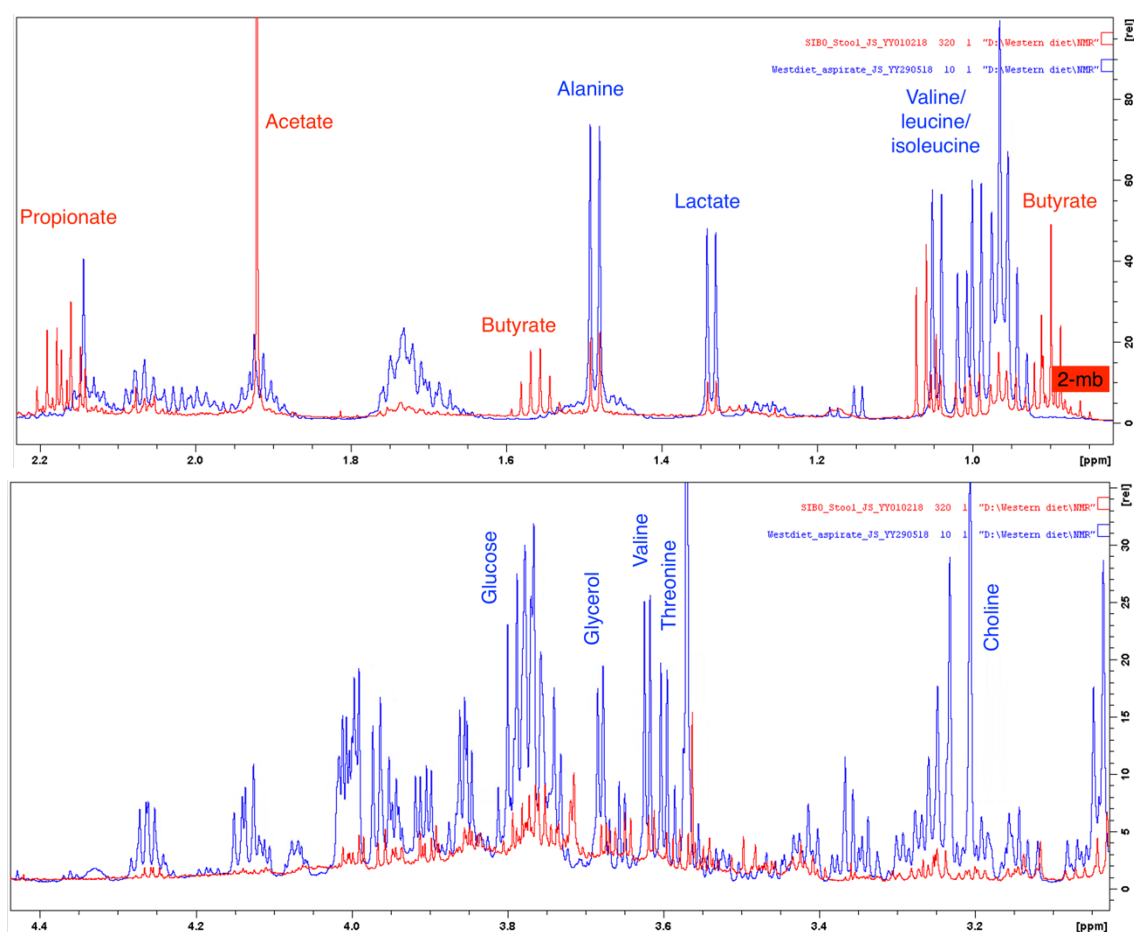


Figure 3.4. ¹H 600 MHz NMR spectrum of QC samples from duodenal aspirates (blue) and stool (red) showing chemical shifts (δ) regions between 0 and 4.

3.4.5 No system wide metabolic differences between high fibre and WD diet

No diet-associated metabolic variation was seen in the plasma (**Figure 3.5A**) and urine profiles (**Figure 3.5B**). This was indicated by the PCA scores plots shown in Figure 3.6A-B. Pairwise OPLS-DA models comparing the urine ($Q^2Y = -0.26$) and plasma ($Q^2Y = -0.92$) profiles were not significant further confirming a lack of biochemical alterations following the dietary switch.

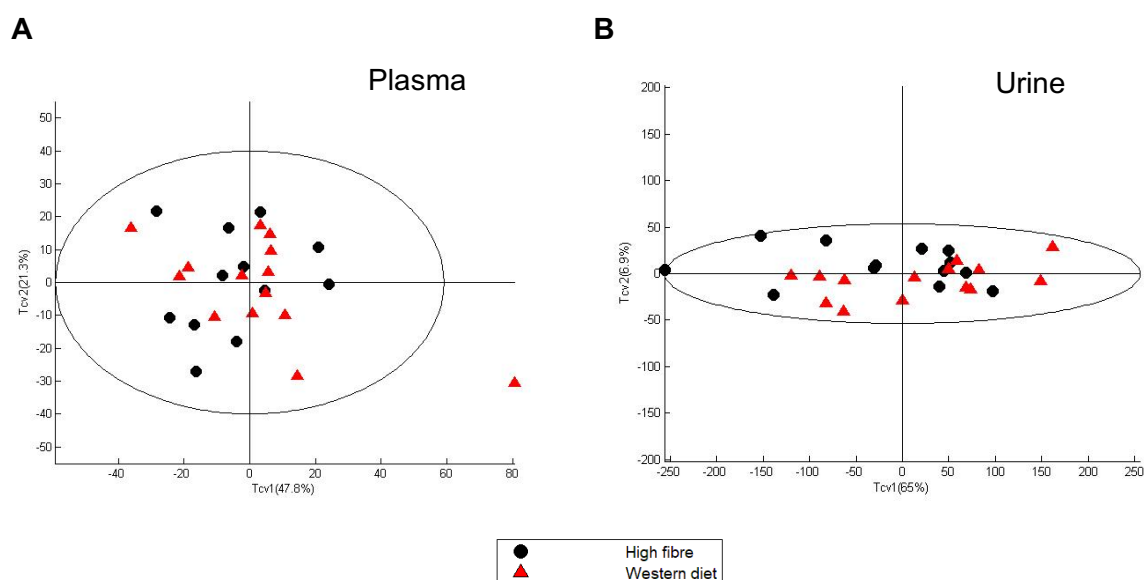


Figure 3.5 No system wide metabolic differences between high fibre and WD diet.

(A-B) Cross-validated scores plot from the PCA model built using 600 MHz ^1H NMR spectral variables of (A) plasma samples and (B) urine samples.

3.4.6 Metabolites in the duodenum are associated with the development of symptoms

For the diagnosis for SIBO, 8 out of 16 participants had a positive diagnosis pre-intervention and 6 had a positive diagnosis post-intervention (**Table 3.1**). To understand the associations between SIBO and metabolic in the duodenum and the faeces, the relative abundance of metabolites identified in **Figure 3.5** were compared in participants with SIBO+ and SIBO- diagnosis using the Wilcoxon Ranks Sum Tests. SIBO+ participants had higher amounts of choline in their duodenum compared to those with a negative SIBO diagnosis (**Figure 3.6A**). Since nutrient absorption takes places in the SI, higher duodenal choline content may reflect a reduction in absorption and a subsequent reduction in choline entering the host metabolic system. To investigate this, the peak integral for choline was extracted from the plasma NMR spectra. It was found that even though no significance was reached, on average participants with a positive SIBO diagnosis did have a trend for lower choline in their plasma samples, as

shown in **Figure 3.6B**. Choline was also lower in participants that reported feeling fatigue post-intervention (**Figure 3.6C**) although no difference found for plasma choline (**Figure 3.6D**). Since choline is found in dairy and protein-rich foods, protein consumption associated with these two symptoms was also analysed and no significance shown. (**Figure 3.6E**).

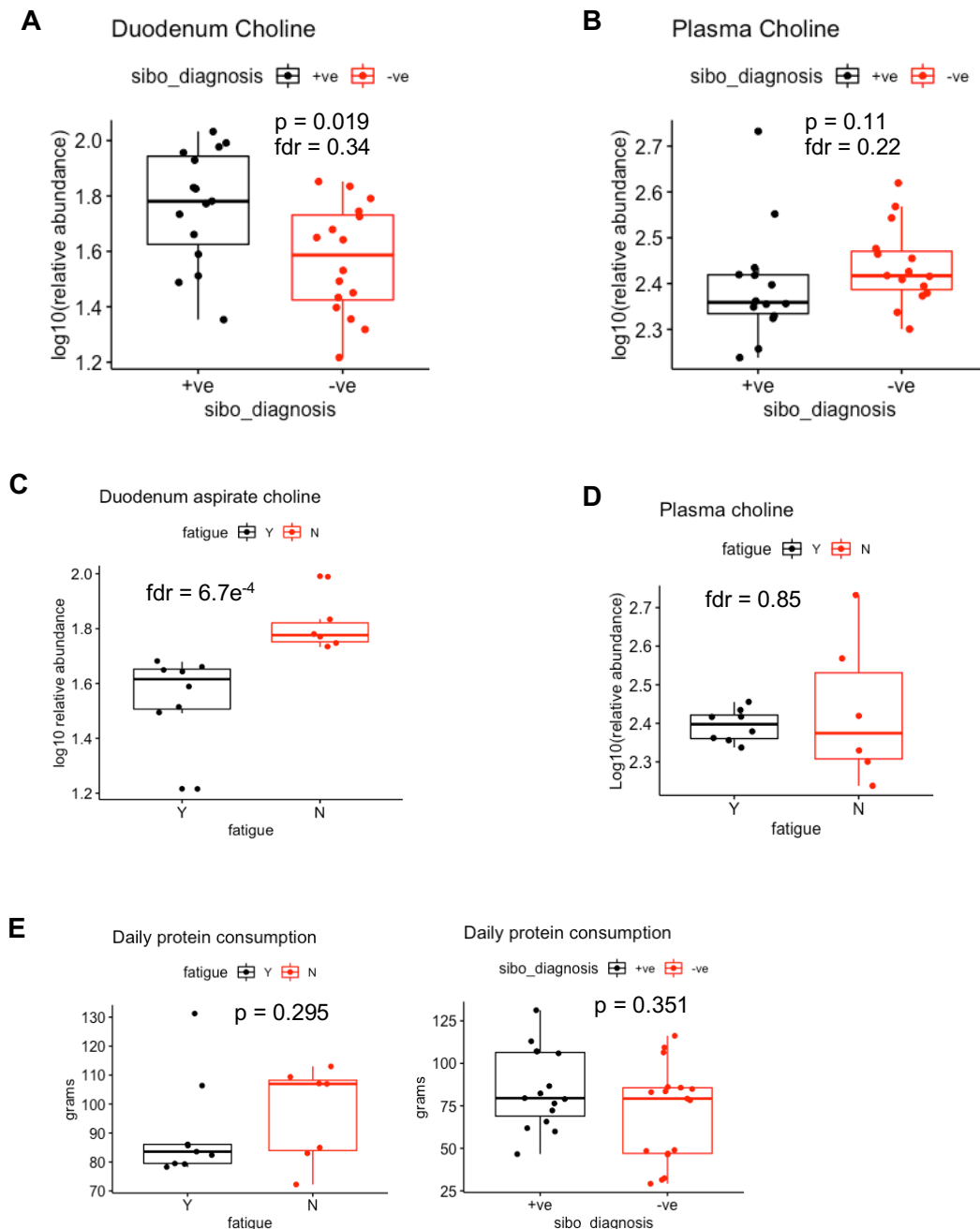


Figure 3.6 Duodenum aspirate metabolites are associated with SIBO diagnosis and post-interventional symptom development.

(A-B) Boxplots showing log10 relative abundance of choline in (A) duodenum aspirates and (B) plasma comparing between participants with a positive and a negative SIBO

diagnosis. Both pre and post intervention samples were included. (Wilcoxon Rank Sum test, *P* value and fdr displayed.)

(C-D) Boxplots showing log₁₀ relative abundance of choline in (A) duodenum aspirates and (B) plasma comparing between participants with either a Yes or a No answer to a question on fatigue. Both pre and post intervention samples were included. (Wilcoxon Rank Sum test, fdr displayed.)

(E) Boxplots showing grams of daily protein consumption between participants with either a Yes or a No answer to a question on fatigue. Both pre and post intervention samples were included. (Wilcoxon Rank Sum test, fdr displayed.)

Boxplots shows the median and the interquartile range (IQR) of the dataset. The whiskers are drawn between 5 to 95 percentiles of the dataset.

3.4.7 Microbiome showed no variation between high fibre and WD diet

Neither the duodenal aspirate microbiota (**Figure 3.7A**) nor the stool microbiota (**Figure 3.7B**) showed any dissimilarities before and after the WD intervention as shown in the Bray-Curtis β -diversity ordination plot which is a measure of between sample variation. However, considerable between-subject microbiome variation in the duodenum and the colon was observed, which was greater than the variation induced by the diet.

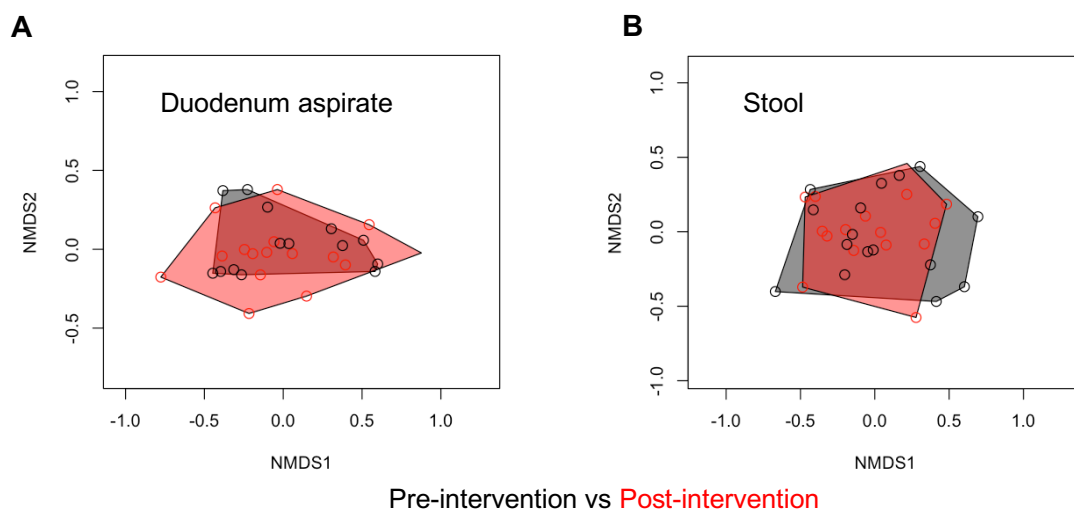


Figure 3.7. No difference in microbiome composition between high fibre and Western diet.

- (A) Bray-Curtis β -diversity ordination plot of duodenum aspirates 16S sequencing data comparing between pre-intervention samples (black polygon and circles) and post-intervention samples (red polygon and circles).
- (B) Same as (A) but for stool samples.

3.4.8 Distinct microbial communities in the duodenum and the colon

Microbial community composition differs between the duodenum and the colon. First, microbial diversity measured by the Shannon evenness index was significantly higher in the colon than the duodenum (**Figure 3.8**) which indicates that microbial diversity in the duodenum was not as evenly distributed as the colon.

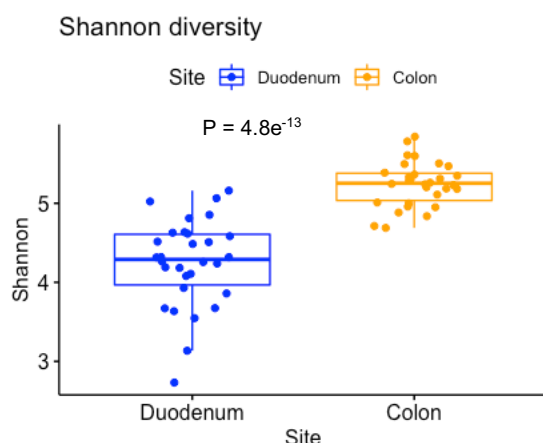


Figure 3.8 Alpha-diversity measured by the Shannon metric was significantly higher in the stool samples than the duodenum aspirates. All pre and post-intervention samples were included. (Wilcoxon Rank Sum test, fdr displayed.)

Next, microbial phylotype in duodenum aspirates and stool was analysed in more detail. At the phylum level (**Figure 3.9A-B**), Firmicutes was the most dominant phylum in both samples with a relative abundance over taking the other four top phyla combined. Bacteroides was the second most abundant phylum in the stool samples whereas both Bacteroides and Proteobacteria were equally abundant in the duodenum aspirates. The divergence at genus level was more apparent between the two locations (**Figure 3.9C-D**). The duodenum lumen was dominated by *Streptococcus* whereas the proportion of the top four genera were more evenly distributed in the stool. At the species level, it was not surprising to find the top three species in the duodenum all belonged to *Streptococcus spp* (**Figure 3.9E**). In the stool samples, difference in species abundance between pre and post interventions started to emerge where there was higher abundance of *Ruminococcus bromii* (*R.bromii*) and *Prevotella copri* (*P. copri*) in the pre-intervention samples (**Figure 3.9F**). Both species are key degrader of resistance starch in the human colon.¹⁶² However, there were no statistically significant differences in alpha-diversity between pre and post-intervention samples nor were there differences in phylogenetic abundances in either sample types.

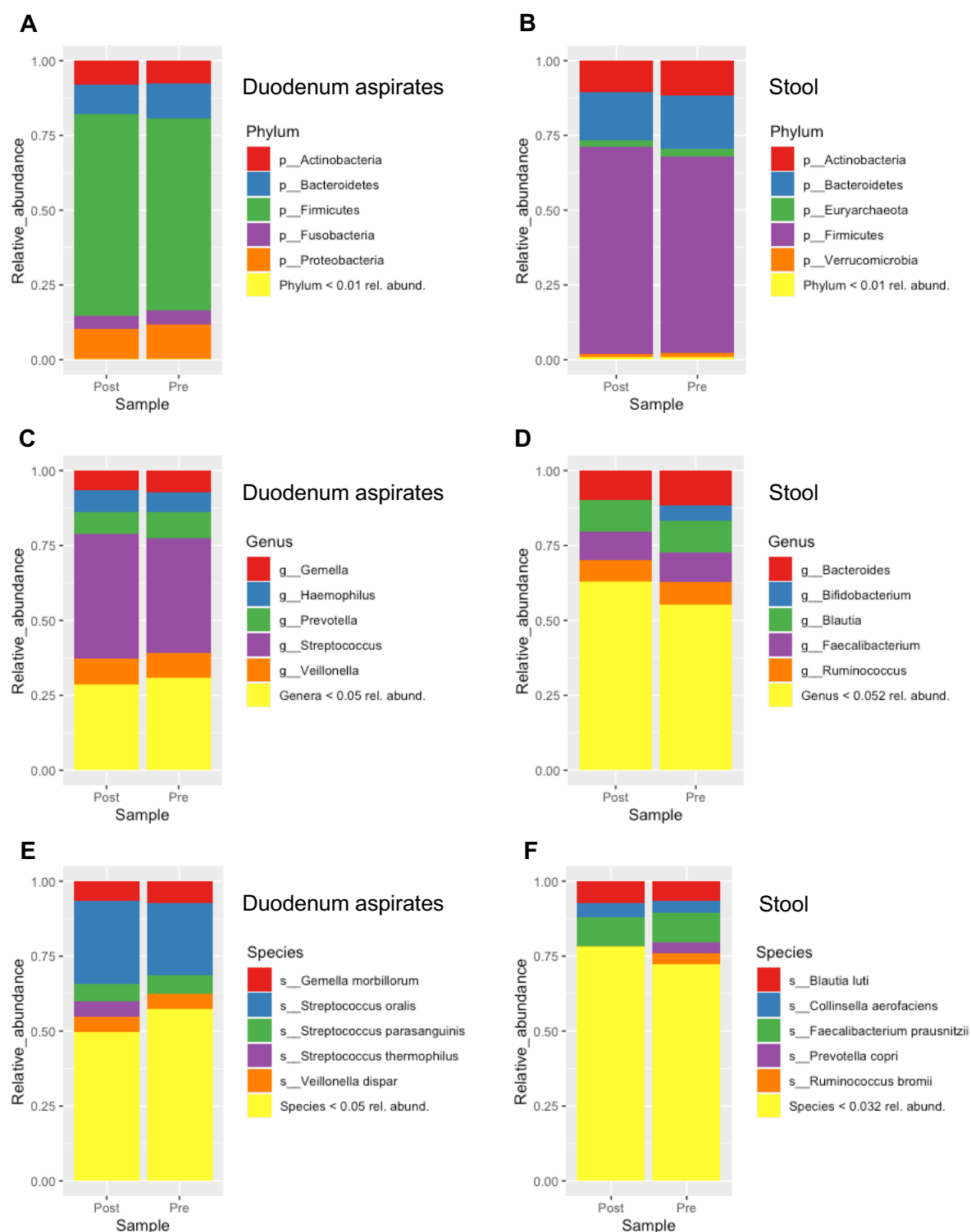


Figure 3.9. Longitudinal gut compartments harbour different microbial communities.

(A-B) Stacked bar charts displaying the relative abundance of the top five most abundant phyla in (A) duodenum aspirates and (B) stool samples comparing between post (WD) and pre intervention (high fibre).

(C-D) Same as (A – B) but showing the top five most abundant genera.

(E-F) Same as (A – B) but showing the top five most abundant species.

3.4.9 Metabolites are associated with the microbiome in their respective compartments

Correlation analyses were performed to identify statistical associations between microbial species abundance and metabolites in the duodenum and the colon. In the duodenum, out of the top five most abundant species, *S. parasanguinis* was the only *Streptococcus* spp. to positively associated with lactate, aspartate, glucose and choline (**Figure 3.10A**). *S. parasanguinis* may be the species driving the metabolism of these nutrients in the duodenum, particularly choline, as it was elevated in the duodenum of SIBO+ patients (**section 3.4.6**). Interestingly, choline-binding-proteins (CBP) are commonly found in the cell wall of a number of *Streptococcus* spp and they are involved in the interaction with host cells.¹⁶³ CBP are also responsible for mediating virulent activities of *S. pneumoniae* although the same mechanism has yet to be found in other *Streptococcus* spp.¹⁶⁴ The metagenomics study by Van denbogert *et al.* identified gene regions in *S. parasanguinis* predicted to encode ATP-Binding cassette (ABC) transporters for multiple sugars which may explain its strong association with glucose.¹² The negative association between *S. parasanguinis* and *S. oralis* was also interesting, as it seems to suggest that there was potential competition for survival between these two species. In the colon, *Blautia luti* was positively associated with a number of amino acids and BCAA. *P. copri* was positively associated with lysine and *R. bromii* was negatively associated with succinate (**Figure 3.10B**).

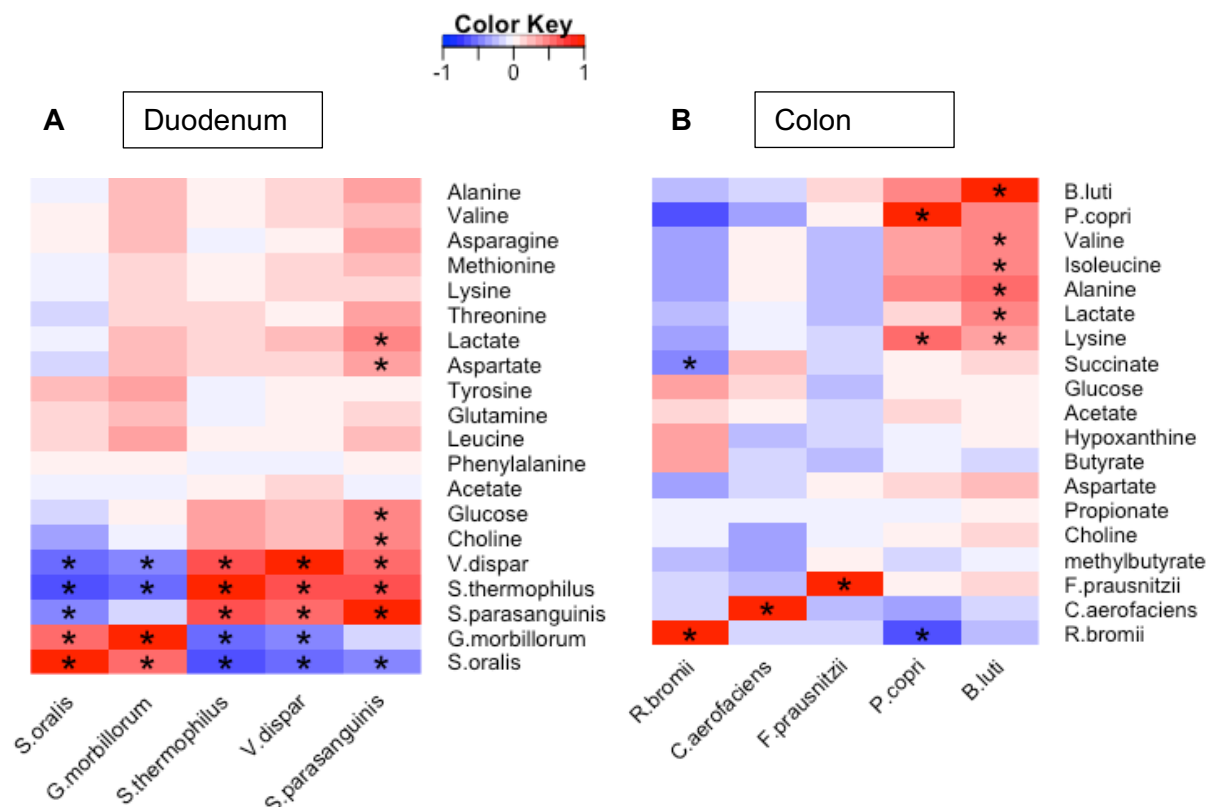


Figure 3.10. Associations between microbial species and metabolites in the duodenum and the colon. Heatmap showing the statistical associations between the stool and duodenal metabolites and microbial species abundance in the (A) duodenal and the (B) colon. Colour intensity is proportional to the correlation coefficients (Spearman correlation). *Significant correlations ($fdr < 0.05$).

3.5 Discussion

This chapter was set out to explore the effect of Western diet on the metabolome and the microbiome of the small and the large intestines in human. All participants who were consuming a high fibre baseline diet developed one or more GI symptoms during the seven-day Western diet trial. No metabolic alterations were noted in the duodenum following consumption of the WD, but significant reductions were seen in the faecal SCFA. Given that these bacteria metabolites arise from the breakdown of fibre, this observation was expected transitioning from a high-fibre diet to one lacking these components. However, no change was observed in the microbial composition of the faeces including no changes in SCFA-producing species. This finding was in agreement with a previous short-term dietary intervention study in human where no difference in alpha-diversity was observed between low fibre and high fibre diets even though SCFA production was increased.⁹² In the same study, investigators found that participants who were consuming an animal-based diet with twice the amount of

protein and fat than the WD study and a lack of fibre, did cause alterations in the community structures of the microbiota. Although after 3 days on the baseline diet, the participant's microbiomes did return to pre-intervention levels⁹². Collectively, these results suggest that the commensal microbiota is both resilient and adaptable to short term changes in fibre and carbohydrate intakes. Discrepancies between human and mouse studies also warrant attention. A study by Turnbaugh *et al.* observed an immediate shift in microbiome structure when humanized mice, i.e. germ-free mice transplanted with human faecal microbiota were fed a WD diet after having consumed a low-fat, fibre-rich diet for one month.¹⁵⁹ This apparent disconnect highlights the challenges in translating results from animal studies to infer human physiology. One of the limitations commonly associated with humanized mouse model is the use of a single faecal donor which fails to account for the inter-individual variations in microbial composition, thus potentially leads to biased results.¹⁶⁷

It is unclear whether the development of SIBO is associated with WD as SIBO+ve participants were identified at both pre and post-intervention. Interestingly, choline in the duodenum was elevated in participants diagnosed with SIBO irrespective of diet. It is possible that increased choline in the gut could fuel the abnormal growth of anaerobic bacteria in the duodenum and affect its availability in the plasma. Consistently, a mouse study by Romano *et al.* showed that mice colonised with choline-consuming microbes had a reduction of choline in their circulation.¹⁶⁵ Choline is an important nutrient in the human body required for the synthesis of the neurotransmitter, acetylcholine and the membrane components, phosphatidylcholine, and therefore is essential for maintaining brain function and cellular structure. Verifying the presence of choline-consuming species in SIBO patients should be explored further.¹⁶⁶ Participants who reported feeling fatigue during the WD trial also had decreased levels of choline in their duodenum. Surprisingly, increased simple sugar intakes did not result in an increase of glucose in either the duodenum or the plasma post-intervention. This points to a possible usage by *Streptococcus.spp* since it is highly adapted at metabolising sugars.

3.6 Study limitation and future work

In this study, baseline food intakes were estimated from patient reported questionnaires (FFQ) and then summarised into the main nutrient groups, carbohydrates, proteins, fat and sugar. However, research have shown that microbiome exhibit high variability associated with food choices but not with conventional nutrients.¹⁷² Adding to this complication, it was also observed that similar foods produced different effects on different people's microbiome. This personalised response to foods can affect the interpretation of a dietary intervention study. Recently, a new tool of accessing individuals' dietary response based on their metabotype

(metabolic phenotype) was developed.¹⁷³ Such tool in combination with the microbiome profile could provide more accurate assessment of the metabolic activities of individual's microbiome in response to food.

3.7 Conclusion

A key finding from this pilot study was that metabolic output rather than community composition of the microbiota was influenced by short-term dietary change. A longer study lasting one month in a larger cohort may be required to establish the long-term effect of WD on the SI microbiome. It would also be interesting to investigate whether switching from a baseline WD to a week of high fibre diet also cause a change limited to the functional capacity of the microbiome. Results from this study could provide basis for a longer dietary intervention to attempt at restoring healthy fibre-degrading species back into the gut of long term WD-consuming participants.

Chapter 4

Irritable bowel syndrome and the colonic microbiota

4.1 Introduction

This chapter details the work of a longitudinal multi-omics study to identify microbial mediators that drive subtype-specific phenotypes in IBS. Specifically, results focus on the metabolic disturbances in the gut lumen and the mucosa and their associations with the microbiome, host transcriptomes and patient's symptom severity.

IBS is a globally prevalent functional GI disorder.¹⁷⁴ In the UK, IBS accounts for 40%-60% of referrals in the GI clinic.¹⁷⁵ Patients with IBS typically suffer from recurrent abdominal pain and discomfort and are characterised by the frequencies of their bowel movements to either constipation dominant (IBS-C), diarrhoea dominant (IBS-D) or mixed (IBS-M).¹⁰¹ In addition to altered bowel motility, IBS pathogenesis also includes visceral hypersensitivity, changes in intestinal secretion and permeability.¹⁷⁶ The origin of IBS disease manifestation is multifactorial where diet, infection and psychological factors all contribute to triggering IBS symptoms.¹⁷⁶ In recent years, the microbiome has emerged as an important player in the pathogenesis of IBS, where transit time changes associated with IBS-C and IBS-D were replicated in gnotobiotic mice following transplantation of faecal microbiota from patients.¹⁰⁴ More recently, mice humanized with microbiota from IBS patients with increased permeability also demonstrated increased *in vivo* permeability.¹⁷⁷

While rodent studies have been instrumental in establishing mechanistic links between IBS pathogenesis and the microbiome, they do not fully recapitulate the human physiology.¹⁸⁰ There is also a lack of stratification of IBS subtypes since transit time is a well-known factor for microbiome variability in the gut which is the main reason for the discrepancies seen in a large number of microbiome studies.^{178,179} In addition, IBS symptom severity varies between individuals and fluctuates over time thus longitudinal sampling would enable the establishment of a temporal profile of patient's clinical and molecular characteristics. The heterogeneous nature of IBS requires deep molecular phenotyping studies to identify perturbed disease pathways and capture system level interactions between the microbiome and the host. Furthermore, studies comparing the microbiome communities in the mucosa and the lumen

of IBS patients are limited to taxonomy profiling with a lack of understanding of the metabolic alterations within these two unique compartments.

4.2 Aims and Objectives

This longitudinal multi-omics study in human was led by the Mayo Clinic to elucidate microbial-mediated mechanism in IBS. The study collected extensive metadata as well as longitudinal host and microbial samples from defined subsets of IBS patients. **Figure 4.1** provides an overview of the metabolomics experiments performed for this study which aimed to explore the metabolome in the transverse compartments of the gut and in the systemic circulation. Therefore, results in this chapter are presented in a compartmental-specific manner focusing on the lumen and the mucosa metabolome and their associations with the microbiome and transcriptome in the context of IBS physiology. Insights generated from this multi-omics study is a joint effort with other institutions with an overall goal aimed at generating a wealth of hypotheses to guide future IBS research. Given the nature of this collaborative project some of the analysis in this chapter were performed by collaborators where they are appropriately acknowledged in the relevant sections.

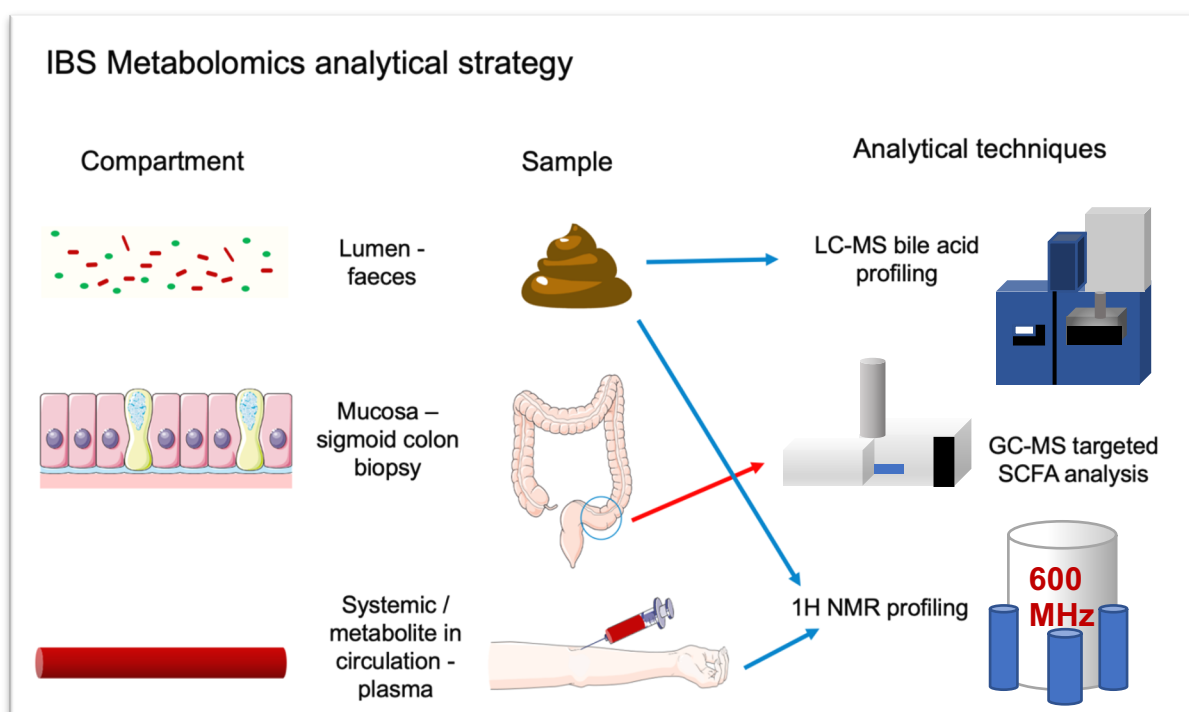


Figure 4.1. Overview of metabolomics analytical strategy.

4.3 Methods

4.3.1 Participant recruitment and exclusion criteria

This study was approved by the Mayo Clinic Institutional Review Board and was registered with clinicaltrials.gov under NCT02249169. Mayo Clinic (Rochester, USA) recruited participants for the trial. Healthy subjects and IBS-C and IBS-D patients between 18-65 years old who responded to an advert for recruitment were invited to undergo screening to assess eligibility. Participants were given the option of undergoing two flexible sigmoidoscopies. All IBS-C and IBS-D subjects fulfilled the Rome III criteria. Recruitment of healthy subjects was matched with IBS subjects for age, sex and BMI. Volunteers with prior history of abdominal surgeries (except appendectomy and cholecystectomy), diagnosis of inflammatory bowel disease, microscopic colitis, celiac disease, or other inflammatory conditions, antibiotic use within the past 4 weeks, bleeding risk or taking medication that increases bleeding risk (only for those who chose to undergo the flexible sigmoidoscopies), bowel prep for colonoscopy in the past week, pregnancy, plan to become pregnant during study, being a vulnerable adult, and age below 18 or over 65 were excluded. In addition, people were excluded with other diseases, conditions, or habits that would interfere with study completion, increased risks with flexible sigmoidoscopies (if chosen), or that in judgment of the investigator would potentially interfere with compliance to study or adversely affect outcomes.

4.3.2 Clinical metadata

Participants underwent a medical examination to measure their height, weight and BMI and their medical history was taken by the study physician during their first visit. They were also given training by a dietitian on how to fill out the Food Frequency Questionnaires (FFQ) and the 24-hour dietary recall questionnaire. Additional questionnaires at the first visit were Rome III criteria for IBS diagnosis, IBS symptom severity (also completed monthly for IBS participants), microbiome health, bowel disease questionnaire (BDQ-6), Hospital Anxiety and Depression, IBS Quality of Life, and a seven-day Bowel Diary (also completed monthly for all participants).

4.3.3 Sample collection

Baseline stool samples were acquired via home collection kits at the earliest convenience after the initial visit and then monthly for six months. Sample tubes were returned on frozen gel packs overnight via courier service or dropped off at the clinical core facility of Mayo Clinic Centre for Cell Signalling where samples were stored at -80°C. Serum samples were collected

at the initial visit only and stored at -80 °C upon further processing. Biopsies were obtained through flexible sigmoidoscopy from the sigmoid colon 20-30 cm from the anal verge essentially as described previously. Up to two tap water enemas were given to cleanse colon for each procedure. All endoscopic procedures were performed by a single endoscopist and up to twelve colonic biopsies were collected using a large-capacity (2.8 mm) biopsy forceps without pin. Depending on downstream processing the biopsies were placed in RNAlater stabilisation solution (Life Technologies), directly frozen in liquid nitrogen, or placed in glucose Krebs solution on ice (composition in mM: 11.5 D-glucose, 120.3 NaCl, 15.5 NaHCO₃, 5.9 KCl, 1.2 NaH₂PO₄, 2.5 CaCl₂·2H₂O, and 1.2 MgCl₂; pH 7.3– and immediately transported to the laboratory for analysis.

4.4 Results

4.4.1 Overview of study sample collection and subject demographics

A total of 75 participants provided stool samples for at least one time point, and 42 participants agreed to undergo flexible sigmoidoscopy to have their biopsy taken and 29 of them allowed the physician to obtain a second biopsy at the second visit. Study subjects provided 24-hr dietary recall and symptom severity at each visit. A total of 12 subjects provided the additional stool sample at the time of flare (6 each for IBS-C and IBS-D) which was defined as a substantial worsening of symptoms. Detailed information on number of samples collected at each time point and the participants demographics are show in **Table 4.1**.

A

	Study Timeline (month)							
Sample type / clinical data	0	1	2	3	4	5	6	Flare
Stool	76	72	69	64	62	63	56	12
Biopsy	42						29	
Serum	78							

B

	Healthy controls	IBS-C	IBS-D
Mean age (+/- sd)	34.5 (+/- 11.0)	41.6 (+/- 13.6)	37.4 (+/- 12.3)
Mean BMI (+/- sd)	26.8 (+/- 6.1)	27.7 (+/- 8.2)	28.2 (+/- 6.8)
Female (n)	19	21	19
Male (n)	5	1	10

Table 4.1. Sample collection and cohort description. A) Number of samples collected at each time point. B) Study subject demographics.

4.4.2 Dietary intake was not significantly different between cohorts

Since diet is a major determining factor for microbial composition and metabolic output in the gut, dietary intakes between cohorts were compared. Data on the specific food groups from the FFQ such as fibre, protein, carbohydrates and fat were compared using Kruskal-Wallis test. No significant difference was found between cohorts therefore diet should not be a confounding factor for subsequent metabolomics and microbiome analyses.

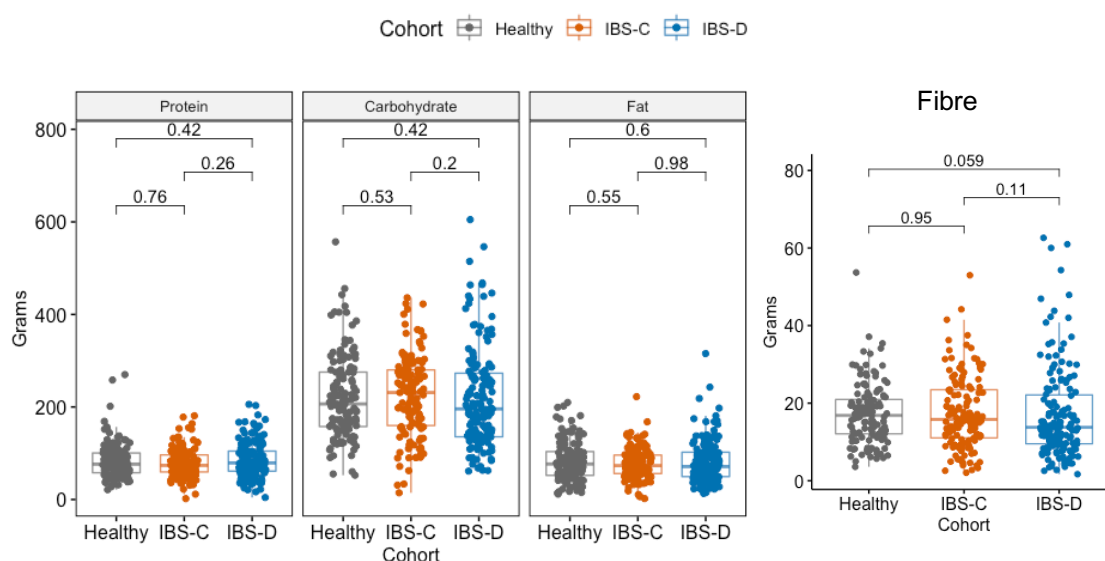


Figure 4.2. FFQ recorded dietary intake for protein, carbohydrate, fat and fibre were compared. (Kruskal-Wallis test, p value displayed.)

4.4.3 Luminal metabolites differ between IBS subtypes and healthy controls

Global metabolic profiling of stool samples and serum samples was performed using 600 MHz ^1H NMR spectroscopy. Initial inspection of the stool profiles measured by NMR using principle components analysis (PCA) is shown in **Figure 4.3A**. No clustering or separation could be seen between the study groups, although a few of the IBS-C samples were observed to have a different profile to the other IBS-C samples. Additional analyses were performed measuring bile acids (BA) in the stool samples using liquid chromatography-quadrupole-time-of-flight mass spectrometry. Similarly, the PCA scores plot on this data (**Figure 4.3B**) showed no clustering of sample cohorts, but a number of IBS-D samples fell outside the main IBS-D cluster. Both methods were robust with minimum systematic variation as indicated by the small

clustering of pooled quality control samples (QCs) taking a central position in relation to all samples. No cohort clustering was detected in the serum samples (Appendix Figure 3).

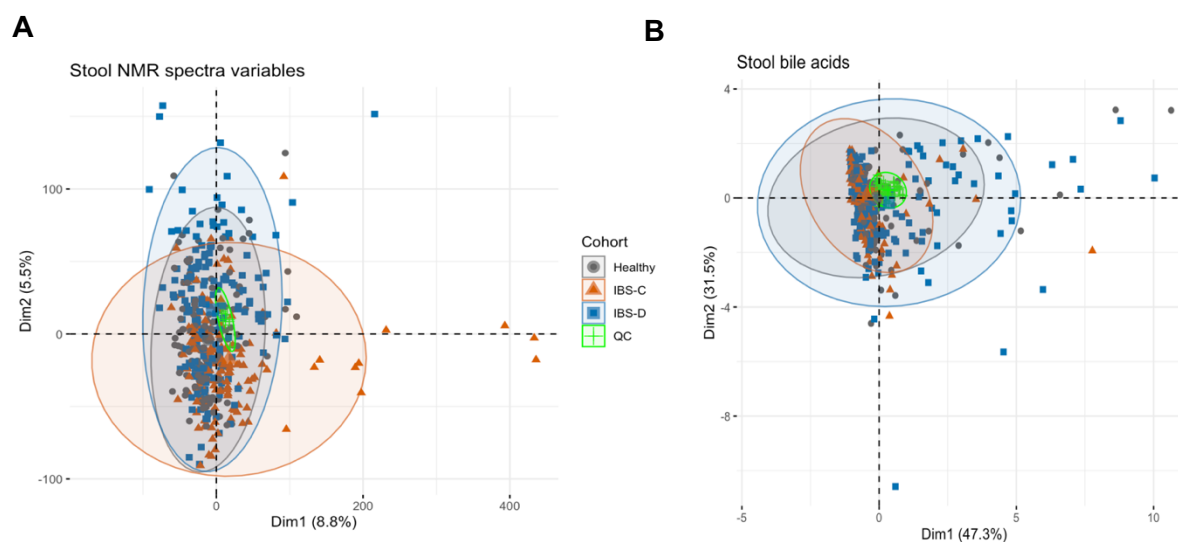


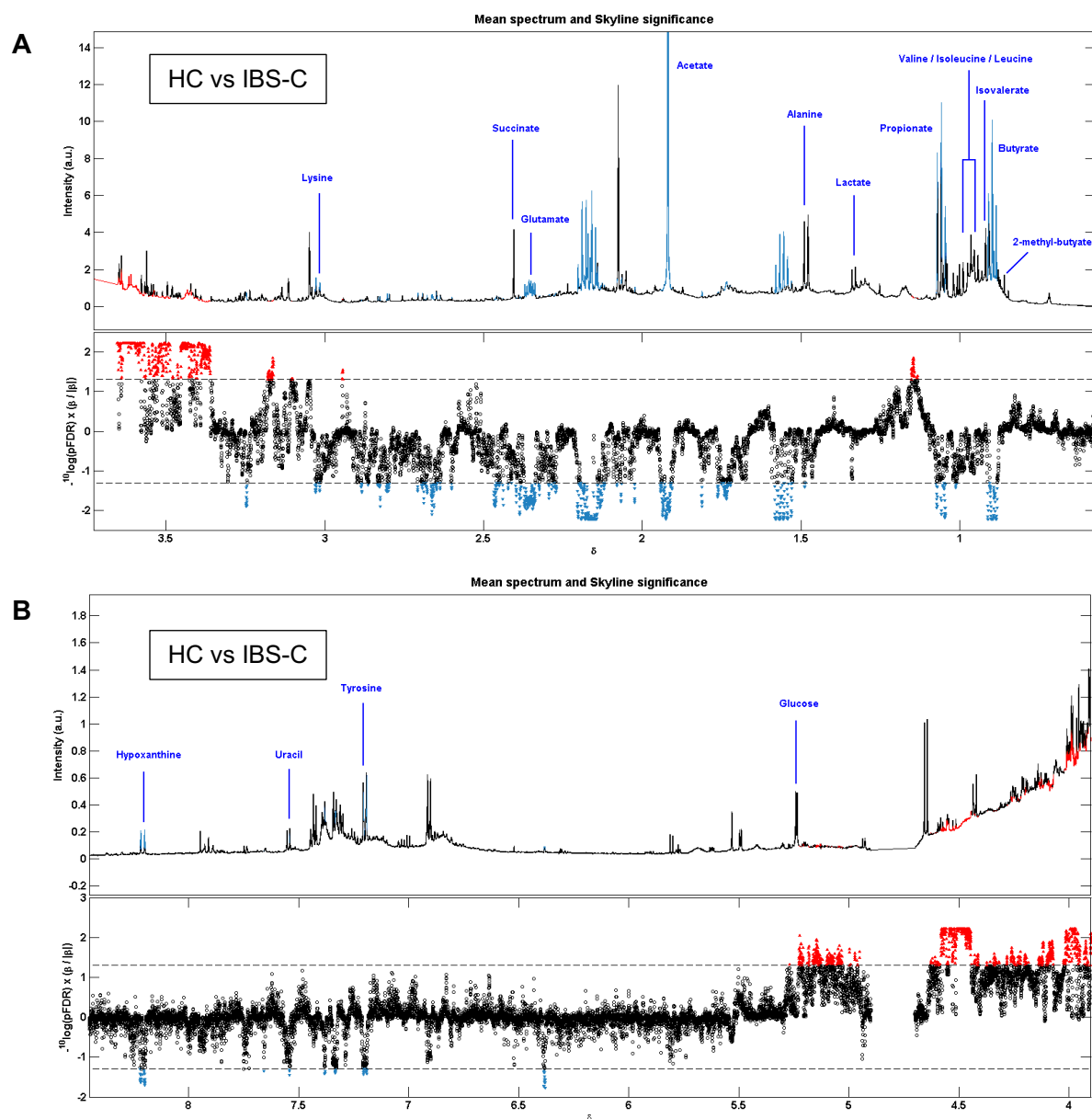
Figure 4.3. No distinct clusters shown between healthy controls (HC), IBS-C and IBS-D based on unsupervised analysis of luminal (stool) metabolites.

- (A) Principle component analysis (PCA) of 30611 spectral variables obtained from the 600 MHz ^1H NMR experiment showing no distinct clusters between healthy control (H), IBS-C (C) and IBS-D (D) as well as pooled quality control (QC) samples.
- (B) Same as (A) but showing PCA was built with the signal intensities of 12 identifiable bile acids (BA) obtained from the untargeted LC-QToF BA profiling experiment. No distinct clusters between H, C, D and QC.

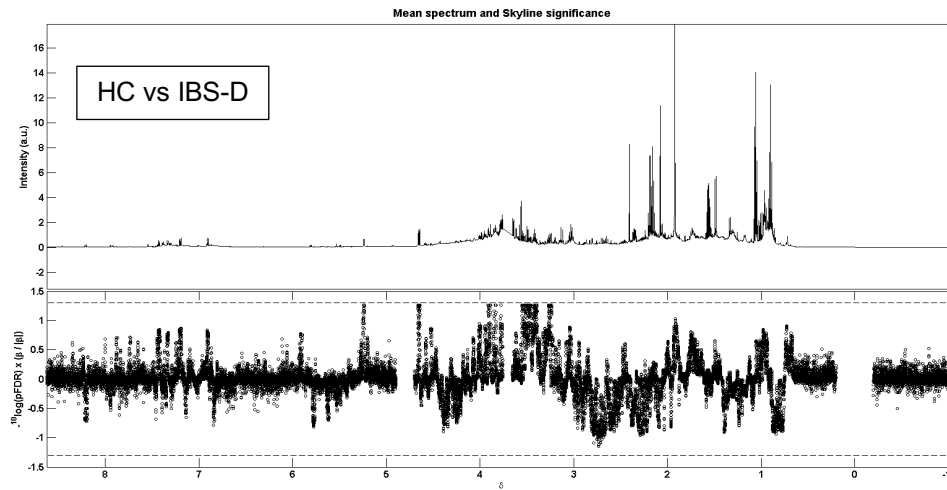
Further analysis of the NMR data on the faeces using MCCV-OPLS-DA analysis identified several metabolic features ($q < 0.05$) discriminating between IBS-C patients and healthy controls (HC) as shown in **Figure 4.4A-B**. Faecal short chain fatty acids (SCFA) namely acetate, butyrate and propionate were significantly lower in IBS-C patients compared to healthy controls. A number of amino acids (AA) were also identified to be reduced. In contrast, no metabolic features were found to differ in the faecal metabolic profiles between IBS-D and HC, as shown in **Figure 4.4C**. This finding confirms with that from a meta-analysis of SCFA in IBS patients where faecal propionate and butyrate were significantly lower in IBS-C patients²⁰⁵.

The OPLS models were assessed for their predictability (Q^2Y) and how well they explain the variation between cohorts (R^2Y). The model that successfully identified discriminating metabolites between HC and IBS-C had relatively high diagnostic values at $R^2Y=0.48$ and $Q^2Y=0.35$ (Figure 4.4D) compared to the model for HC vs IBS-D ($R^2Y=0.18$, $Q^2Y=0.098$)

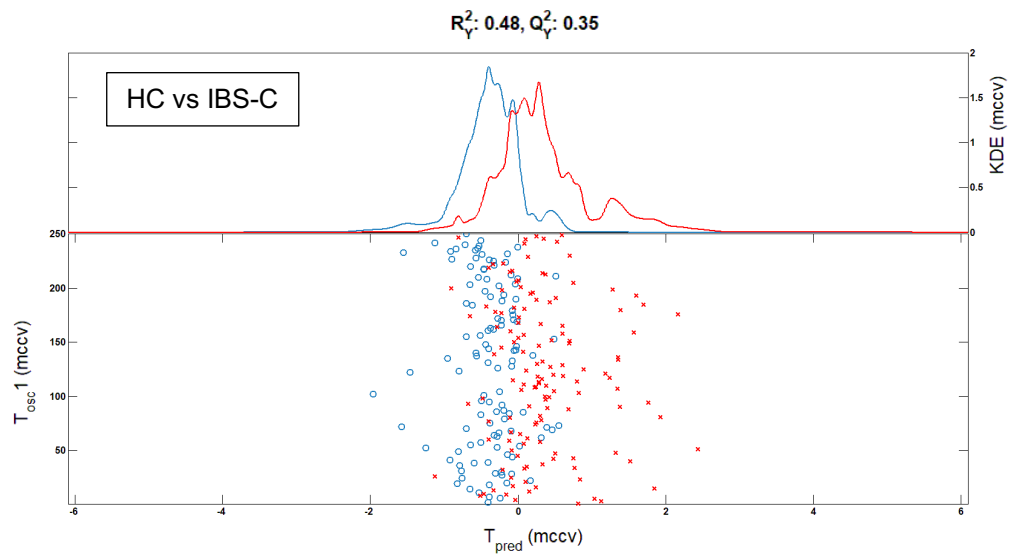
(Figure 4.4E). A total of 10 cross-validation rounds and 25 bootstraps in each model was used to calculate the mean prediction (T_{pred}) of each sample. In Figure 4.4D, a negative (T_{pred}) or blue circle indicates the faecal metabolic profile of the sample resembles more IBS-C than HC. A permutation test was also carried out to assess model robustness. Here, models were built on wrongly assigned (randomly permuted) data and their predictive ability was compared to the model generated from the correctly labelled data. This demonstrated that the predictive ability of the 'correct' model was superior to that generated on 'junk' data (**Figure 4.5F**).



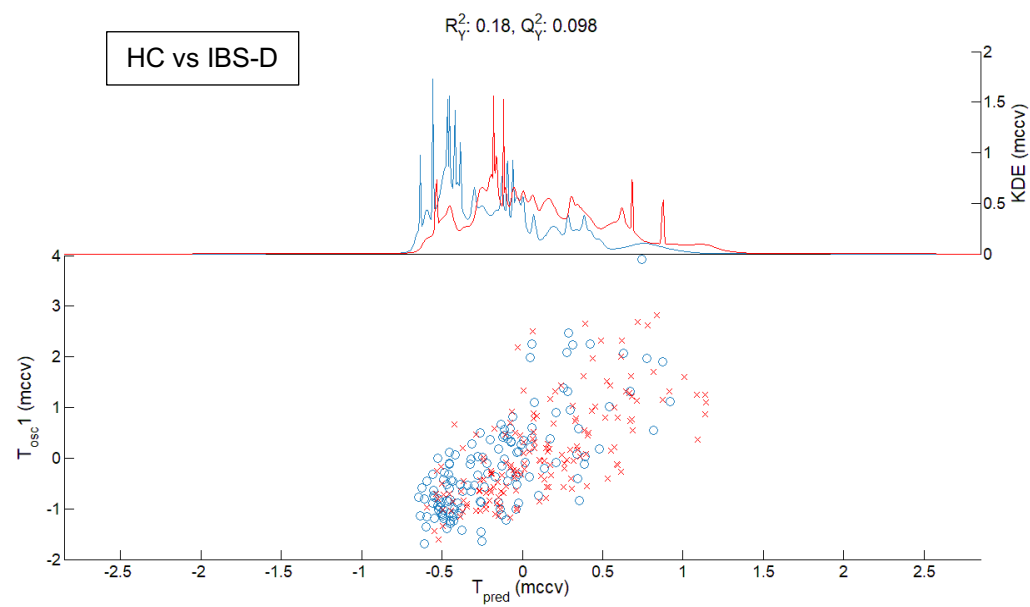
C



D



E



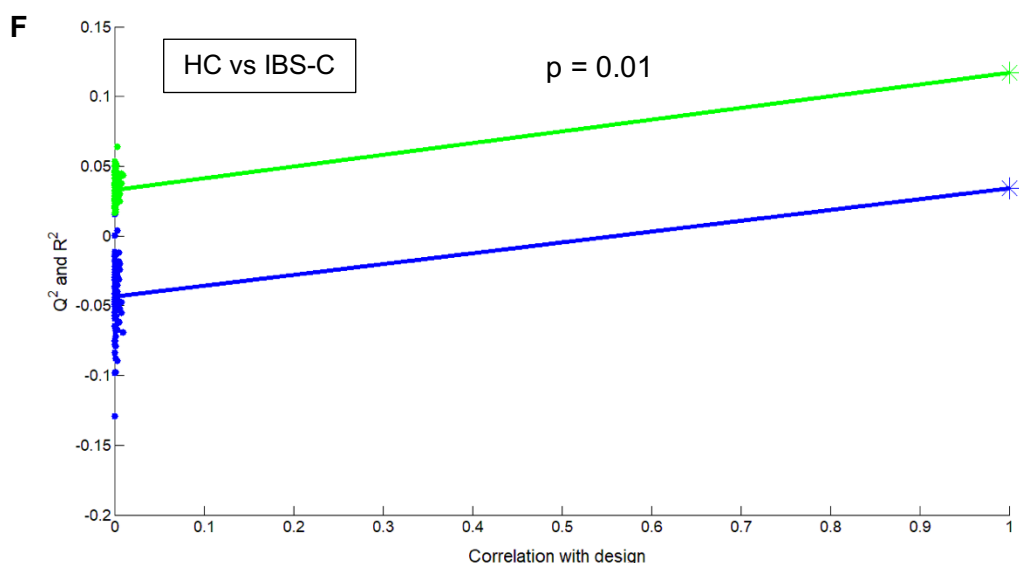


Figure 4.4. MCCV-OPLS-DA model identified several luminal metabolite features discriminating between IBS-C patients and healthy controls (HC)

- (A – B) An OPLS-DA model was constructed using 30611 spectra variables obtained from the 600 MHz ^1H NMR experiment of stool samples showing chemical shifts region (δ) between (A) 0 and 3.8 ppm and (B) regions between 3.8 and 8.5 ppm. Top panel display the averaged spectra whereas bottom panel shows the Manhattan plot of all corresponding spectral variables. Blue-coloured variables passing the lower dotted line shown in the Manhattan plot represent metabolite features that the model identifies to have a significantly lower relative abundance in IBS-C compare to HC ($q < 0.05$). Variables in red indicate higher relative abundance. (Number of samples used were, HC=142, IBS-C=143.)
- (C) Same as (A) and (B) but showing comparison between HC and IBS-D for the entire spectra region (0 to 8.5 ppm). No significant metabolite features were identified by the model. (Number of samples used were, HC=142, IBS-C=180.)
- (D – E) Scores plot of the MCCV-OPLS-DA model showing mean prediction (T_{pred}) of each sample (x-axis) with kernel density estimate (KDE) for (D) HC vs IBS-C and (E) HC vs IBS-D. The KDE was obtained by calculating the mean and standard deviation for each sample in the test set in MCCV which shows the spread of the prediction. A blue circle or a negative (T_{pred}) indicates the sample was more IBS-C like or IBS-D like in the case of HC vs IBS-D. OPLS model characteristics were $R^2Y = 0.48$, $Q^2Y = 0.35$ for HC vs IBS-C and $R^2Y = 0.18$, $Q^2Y = 0.098$ for HC vs IBS-D.
- (F) Plot of permutation test for the OPLS-DA model of HC vs IBS-C. It shows the predicted R^2Y (green dots) and Q^2Y (blue dots) scores of each sample (total 285) after class labels were randomly permuted as having no correlation with the predicted scores of a sample with a correct class label (dots on the far right of the plot). 100 permutation tests resulted in a P value of 0.01.

The physiological significance of lower SCFA in IBS-C patients was investigated in an Ussing chamber experiment designed and executed by Dr Yogesh Bhattari and Dr Purna Kashyap from the Mayo Clinic. This experiment was inspired by their previous work where they showed SCFA can modulate the serotonergic secretory pathway in host GI tissue.¹⁸¹ Results from the Ussing chamber experiment showed the secretory response of colonic biopsies from IBS-C patients to serotonin (5-hydroxytryptamine, 5-HT) was significantly lower than HC (**Figure 4.5**). This suggests that impaired secretory response in IBS-C patients may be a consequence of lower luminal SCFA concentration.

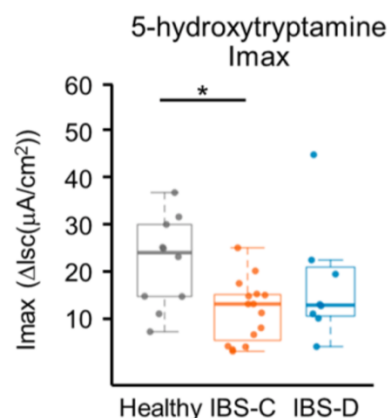


Figure 4.5. Biopsy secretory response to serotonin (5-hydroxytryptamine, 5-HT). Maximum differences in short circuit current ($I_{max}\Delta I_{sc}$) following application of increasing concentrations of serotonin (5-HT) basolaterally in colonic biopsies from time point 1. (ANOVA Tukey HSD, * $p < 0.05$)

4.4.4 Lower stool SCFA in IBS-C is associated with the microbiome

The microbial profiles of the stool samples were acquired using metagenomics and 16S rRNA gene sequencing performed by Dr Tonya Ward from the University of Minnesota.

The faecal microbial diversity of IBS-C patients was greater than that of HC, as indicated by a higher Chao1 richness score and these species were more evenly distributed as indicated by a higher Shannon diversity score (**Figure 4.6A**). IBS-D patients on the other hand have a similar microbiome richness to HC but a lower Shannon diversity score suggesting that one or more species dominated their stool microbiome pool. Between-sample differences were compared in all three cohorts by calculating a Bray-Curtis dissimilarity (BCD)-based irregularity (BCDI) score (performed by Dr Tonya Ward). As shown in **Figure 4.6B**, when considering a sample to be irregular beyond the 90th percentile of the HC distribution, more IBS-C samples are irregular than IBS-D (31.7% for IBS-C, 14.1% for IBS-D, equality-of-

proportions test p -value <0.001). Based on previous studies, it is believed that the longer transit time, being a characteristic of IBS-C patients, promoted the growth of a more diverse microbiome community^{15,178,179}. Furthermore, there is evidence supporting that distal colonic transit is a determinant of SCFA concentration.¹⁸²

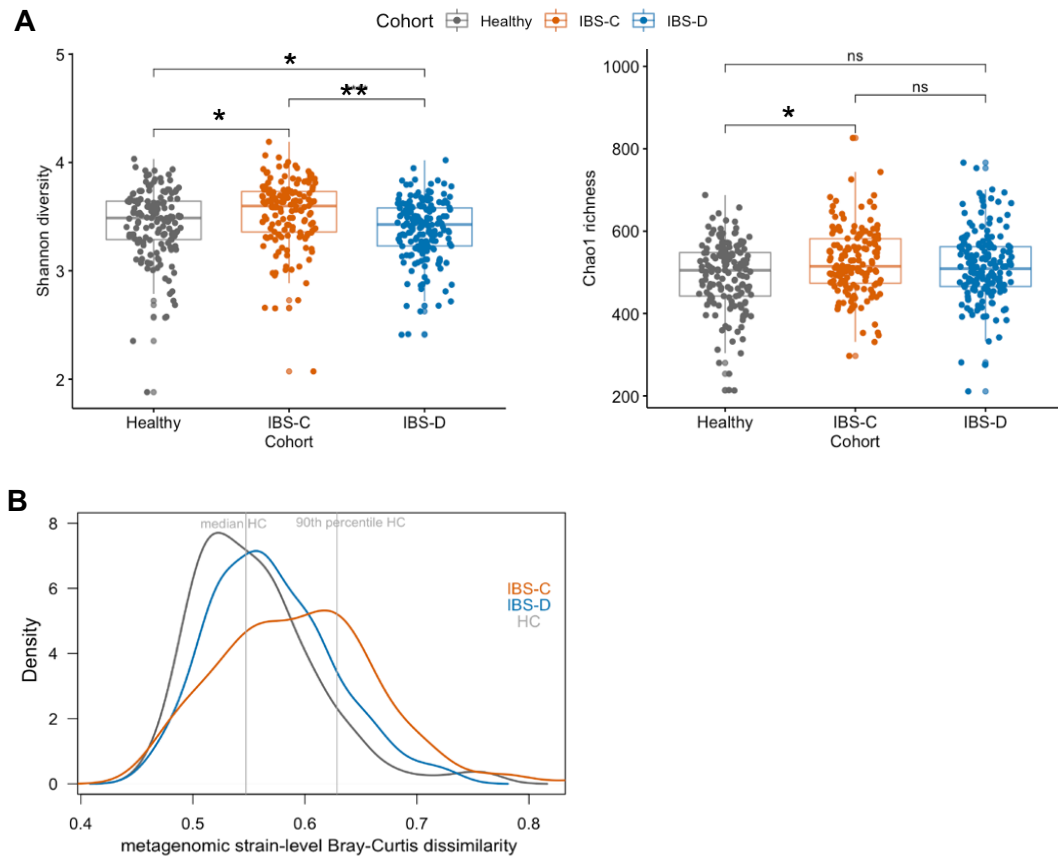


Figure 4.6. IBS-C microbiome is more distinct from HC.

- (A) Box plots showing α -diversity metrics, Shannon diversity and Chao1 richness. (Kruskal – Wallis test, * $p < 0.05$, ** $p < 0.005$, FDR corrected)
- (B) Bray-Curtis dissimilarity (BCD)-based irregularity (BCDI) showing distribution of the three groups. (HC vs IBS-C $p < 0.011$)

Further interrogation of the microbiome data found phylogenetic level differences between HC and IBS-subtypes (**Table 4.2**). More species were lower in IBS-C patients than in IBS-D patients and the majority of them are prominent dietary fibre degraders in the gut. A powerful machine learning method developed by Ms Sambhawa Priya from the University of Minnesota identified significant associations between microbial species and metabolites. Special interest was paid to the associations with SCFA since they were significantly reduced in IBS-C patients. The top five most significant microbial associations with acetate, butyrate and propionate are displayed in **Table 4.3**. Four of them came from butyrate suggesting that a reduction in prominent butyrate producers in the colon such as *R. intestinalis* and *R. faecis* contributed to the decrease in butyrate in the stool. *Ruthenibacterium lactatiformans* (*R. lactatiformans*) was negatively associated with butyrate and acetate. The genus *Ruthenibacterium* was also found to be significantly elevated in IBS-C patients. Interestingly, the genome of *R. lactatiformans* encodes a repertoire of glycan-degrading enzymes which suggests a potential for *R. lactatiformans* to degrade mucus.¹⁸³ The loss of butyrate producers may be a consequence of transit time as this further depletes carbohydrates in the colon which are the main source of fuel for these saccharolytic commensals. However, it is worth noting that SCFA are readily absorbed and metabolised by the epithelial cells therefore alterations in their absorptive functionality and gut permeability cells can affect luminal SCFA concentration in addition to lower abundance of butyrate-producing species.^{179,182}

Cohort	Phylum	Family	Genus	Species
Higher in IBS-C	p_Actinobacteria	f_Coriobacteriaceae	g_Collinsella	s__[Ruminococcus]_torques s__Alistipes_shahii s__Collinsella_aerofaciens
	p_Euryarchaeota	f_Eggerthellaceae f_Enterobacteriaceae f_Methanobacteriaceae f_Streptococcaceae	g_Escherichia g_Methanobrevibacter g_Ruthenibacterium g_Streptococcus	s__Escherichia_coli s__Methanobrevibacter_smithii
Lower in IBS-C		f_Prevotellaceae	g_Clostridium g_Dorea g_Prevotella g_Roseburia	s__[Eubacterium]_rectale s__Bacteroides_coprocola s__Bacteroides_ovatus s__Bacteroides_vulgatus s__Dorea_longicatena s__Prevotella_copri s__Roseburia_faecis s__Roseburia_intestinalis
Higher in IBS-D		f_Lachnospiraceae f_Veillonellaceae f_Streptococcaceae	g_Blautia g_Anaerostipes g_Streptococcus	s__[Ruminococcus]_gnavus s__[Ruminococcus]_torques s__Anaerostipes_hadrus s__Bacteroides_thetaiotaomicron
Lower in IBS-D	p_Actinobacteria	f_Rikenellaceae f_Bifidobacteriaceae f_Clostridiaceae f_Ruminococcaceae f_Oscillospiraceae	g_Ruminiclostridium g_Alistipes g_Bifidobacterium g_Clostridium g_Faecalibacterium	s__Alistipes_putredinis s__Bifidobacterium_adolescentis s__Faecalibacterium_prausnitzii

Table 4.2 Phylogenetic level differences in stool samples between HC and IBS subgroups. Taxonomy on display have passed the significance threshold of FDR corrected $p < 0.05$ based on Kruskal-Wallis test.

Metabolite	Microbial Species	correlation	P value	FDR
Butyrate	Roseburia intestinalis	0.552	4.7E-07	2.7E-04
Butyrate	Ruthenibacterium lactatiformans	-0.541	8.6E-07	3.1E-04
Butyrate	Roseburia faecis	0.489	1.1E-05	1.6E-03
Butyrate	Lachnospiraceae bacterium TF01-11	0.486	1.3E-05	1.7E-03
Acetate	Ruthenibacterium lactatiformans	-0.472	2.4E-05	2.5E-03

Table 4.3 Top five associations between butyrate, acetate and propionate with microbial species produced using the Lasso penalized regression model. Correlation coefficient, P values and FDR corrected P values are shown.

4.4.5 Evidence for enhanced microbial protein fermentation in IBS-C patients

Lower stool SCFA in IBS-C patients suggests reduced microbial saccharolytic activities in the gut. A question was then proposed, could the diminishing carbohydrate-fermenting activity as a consequence of prolonged colonic transit indicate an increased bacteria protein fermentation in IBS-C patients? As carbohydrates become increasingly scarce towards the distal gut, bacteria often rely on amino acids as their main source of energy.⁹¹ In terms of protein fermentation, research has shown that bacteria have a preference for catabolising branched chain amino acids (BCAA) namely valine, leucine and isoleucine for energy which produces branched chain fatty acids (BCFA), 2-methylbutyrate (2mb) and isovalerate as a result.^{94,184}

The OPLS-DA model described in section 4.4.3 (**Figure 4.4A-B**) identified a number of AA that were decreased in the stool of IBS-C-patients. Therefore, peak integrals for isoleucine, leucine, valine and their degradation products 2mb and isovalerate were extracted for comparison using Kruskal-Wallis test. As **Figure 4.7A** shows, all BCAA were significantly lower in IBS-C patients whereas leucine was higher in IBS-D patients when compared to HC. While 2mb was reduced in IBS-D patients, no difference was seen in IBS-C patients (**Figure 4.7B**). However, the ratio between 2mb and isoleucine was higher in IBS-C patients and lower in IBS-D patients compared to HC suggesting a transit time related effect on the fermentation of isoleucine leading to the production of 2mb (**Figure 4.7C**).

Next, associations between metabolites derived from protein fermentation and species abundance were investigated by reviewing the correlation data generated from the Lasso regression analysis performed by Ms Sambhawa Priya. Significant associations ($q < 0.25$) between BCAA, BCFA and microbial metabolites are illustrated in a network graph in **Figure 4.8**. Several strong positive associations were observed between *Alistipes spp.* and 2mb. *Alistipes shahii* (*A. shahii*) was also negatively associated with all three BCAA. *A. shahii* has a gene that encodes dihydrolipoamide dehydrogenase (K00382) which is involved in the degradation pathway of BCAA. The same enzyme has also been identified in *Methanobrevibacter smithii* (*M. smithii*) which was negatively correlated with valine. *M. smithii* also contains a gene that encode branched-chain amino acid aminotransferase (K00826) that converts BCAA to methyl-butanoate in the initial stage of degradation.¹⁸⁵ Literature reports that *Alistipes* has the highest number of protein fermentation pathways amongst commensal bacteria and species in this genus have been implicated in liver and heart diseases.¹⁸⁶

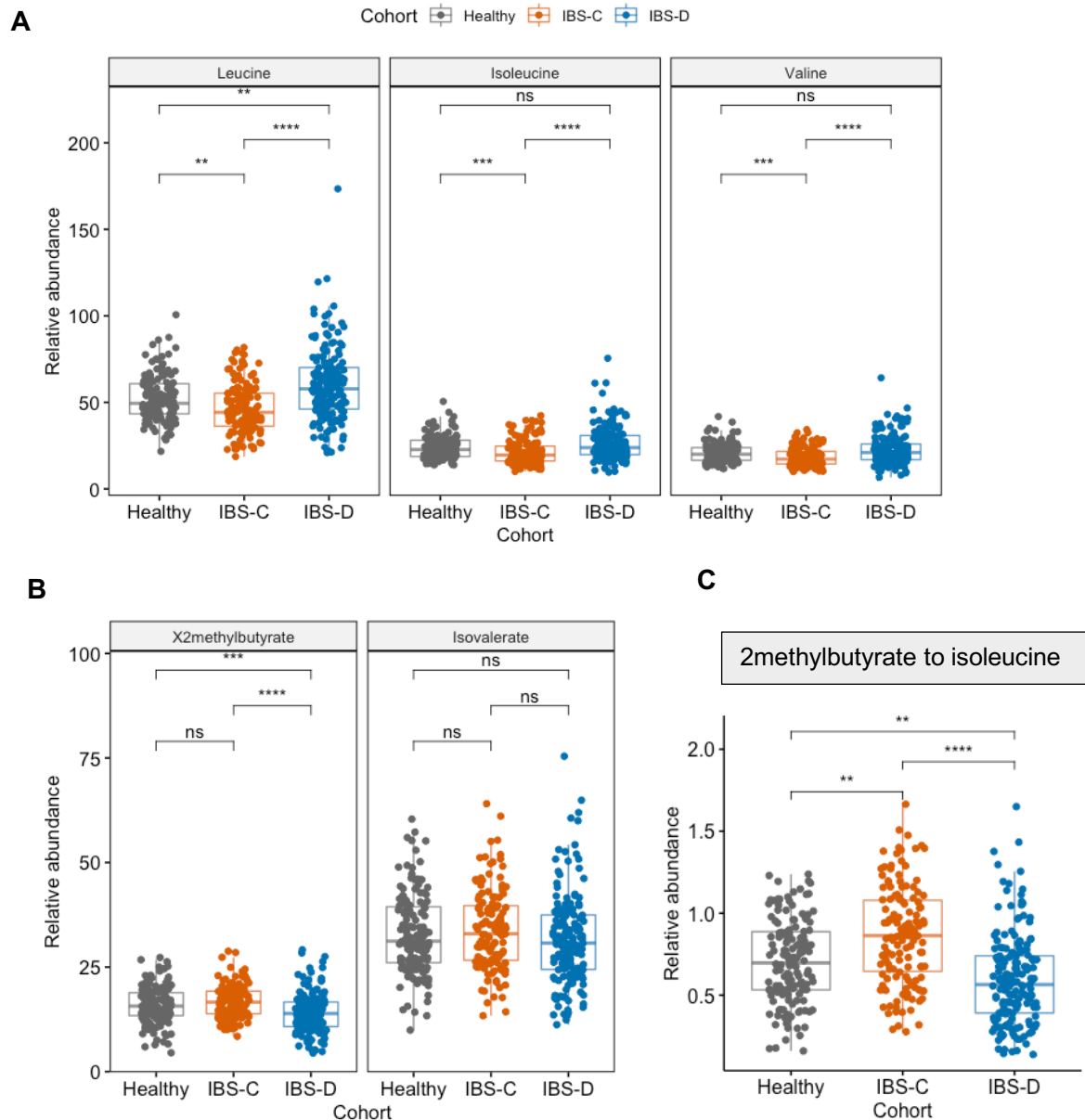


Figure 4.7. Alteration in stool BCAA and BCFA levels indicate increased protein fermentation in IBS-C patients.

(A) Boxplots of relative abundance of NMR peak integrals for leucine, isoleucine and valine in stool samples. (Kruskal-Wallis test, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$, FDR corrected)

(B) Same as (A) for 2methylbutyrate and isovalerate.

(C) Same as (A) for ratio of 2methylbutyrate and isoleucine.

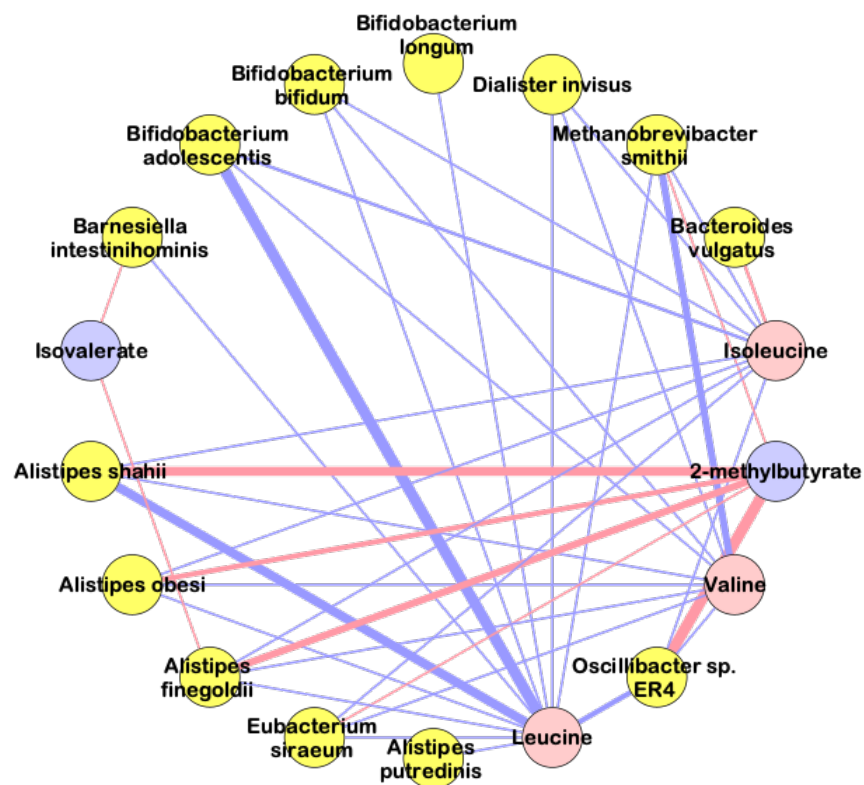


Figure 4.8. Network plot of significant positive (red edges) and negative (blue edges) associations between metabolites (blue and pink nodes) and microbial taxa (yellow node) produced from Lasso penalized regression. Thickness of the edges indicate strength of the association measured by q value.

4.4.6 Multi-omics integration identified a novel host-microbial pathway in IBS

A number of stool metabolites involved in intestinal energy metabolism were significantly decreased in IBS-C and IBS-D patients compared to HC (**Figure 4.9**). The rapidly dividing intestinal epithelial cells (IEC) are energy demanding and require a constant supply of oxidative fuel. Glutamate is an important metabolite in the tricarboxylic acid (TCA) cycle involved in energy generation. Both glutamate and aspartate are extensively metabolised by the IEC.¹⁸⁷ A recently completed clinical trial found that glutamine was able to restore intestinal permeability in patients with post-infectious-IBS-D.¹⁸⁸ The gut commensal *Bacteroides thetaiotaomicron* is capable of breaking down glutamate and interestingly this species was also elevated in IBS-D patients (**Table 4.2**).¹⁸⁹ However, no associations between microbial species with either metabolite were found.

Hypoxanthine (Hpx) and uracil are metabolites in the purine and pyrimidine salvage pathway used by IEC for nucleotide recovery and ATP generation.^{190,191} Hpx was significantly reduced in both subtypes compared to healthy patients whereas reduction of uracil was only seen in IBS-C patients (**Figure 4.9**). Lower hypoxanthine in the lumen could reflect decreased production or increased usage by the host intestinal tissue or the microbiota.

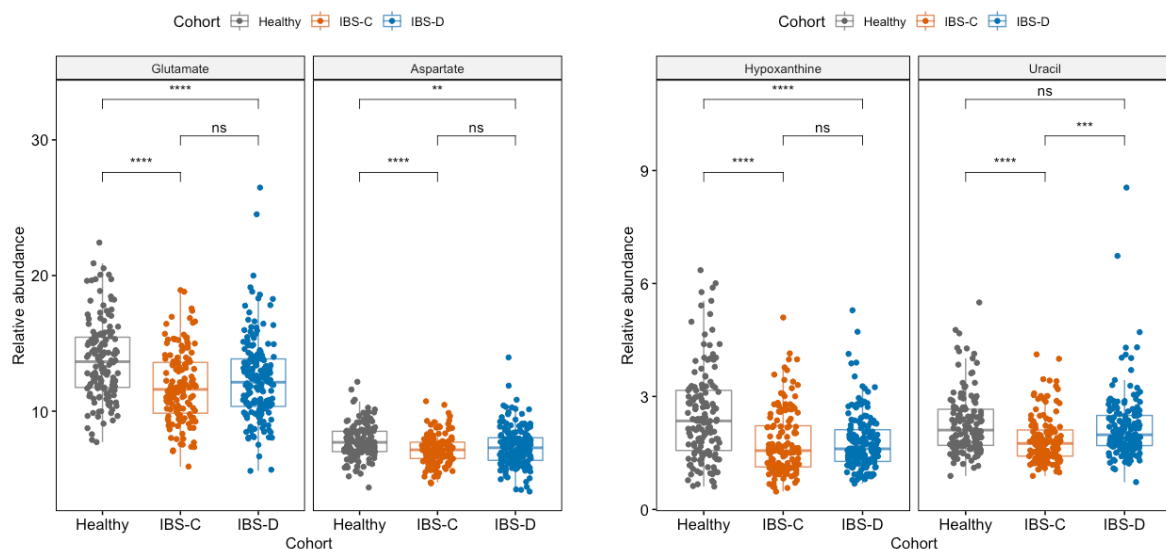


Figure 4.9. Boxplots showing the relative abundance of NMR peak integrals for glutamate, aspartate, hypoxanthine and uracil in stool samples. (Kruskal-Wallis test, ** $P < 0.01$, *** $P < 0.001$, **** $P < 0.0001$, FDR corrected)

To better understand the physiological relevance of this finding hypoxanthine was integrated with metagenomics and host transcriptomics data which led to the discovery of a host-microbial co-metabolic pathway underlying the pathophysiology of IBS. The following experiments were conducted by collaborators (**Figure 4.10**, **Figure 4.11**). A description of the main findings is provided.

A significant positive correlation was identified between hypoxanthine and a structural variable gene region in *Lachnospiraceae* sp. 3_1_46FAA responsible for DNA replication suggesting that bacteria may be using Hpx for growth. Germ-free mice mono-colonised with a similar strain of *Lachnospiraceae* were used in a feeding experiment to demonstrate that Hpx could be consumed by *Lachnospiraceae* *in vivo* (**Figure 4.10**). Host contribution to the depleted Hpx pool was determined by analysing the sigmoid biopsies of both IBS subtypes which showed elevated expression of the human xanthine dehydrogenase (XDH) gene responsible for the breakdown of Hpx in both IBS subtypes (**Figure 4.11**). There was also an increase in purine nucleoside phosphorylase (PNP) in both IBS subtypes which is involved

in the host purine salvation pathway by converting nucleotides back to Hpx (**Figure 4.11**). PNP was also found to be negatively correlated with Hpx levels. The IEC have limited capacity for *de novo* synthesis of purines and therefore relies on the purine salvage pathway for nucleotide generation.^{190,191} The process of purine salvage is energy demanding which suggests that there may be an imbalance in energy supply to the epithelial tissue to maintain its normal functioning.

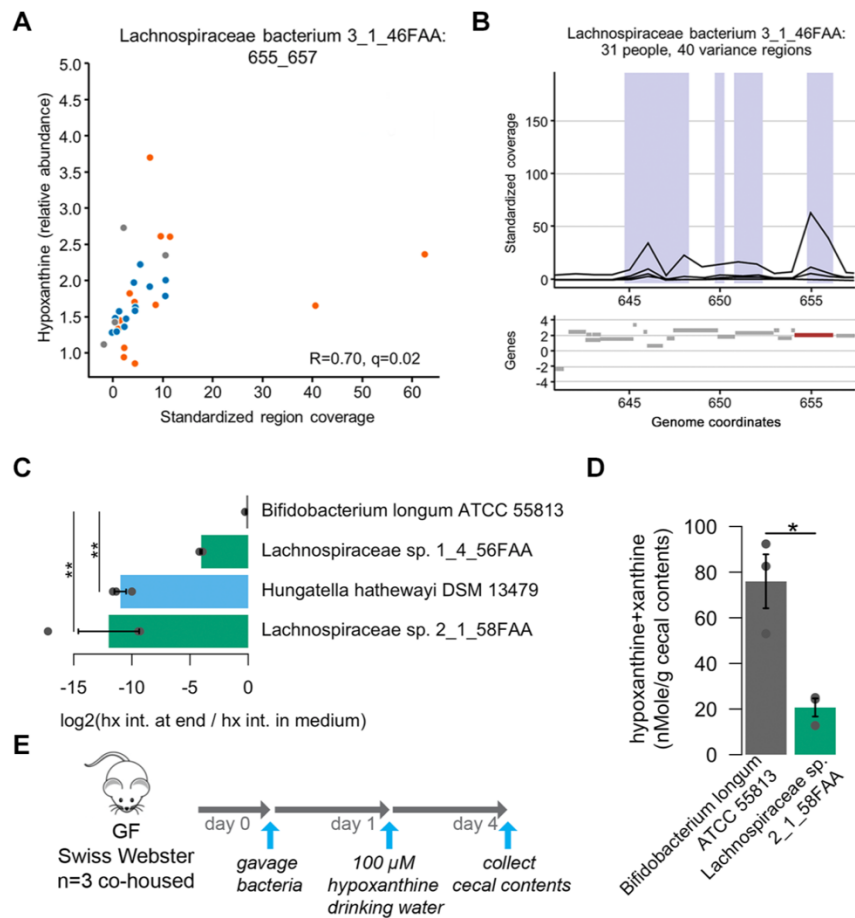


Figure 4.10. Hypoxanthine consumption by specific gut microbiome members as suggested by microbial gene region associations.

- (A) Scatter plot of metabolite intensities and standardized region coverage for SV association result for *Lachnospiraceae* sp. 3_1_46FAA genomic region positively correlated to hypoxanthine (Spearman correlation).
- (B) Genomic context of region from (A) with relevant gene highlighted in red.
- (C) *Clostridiales* strains and *B. longum* were grown in Mega medium. Hypoxanthine levels in the culture supernatant after ON growth were determined with LC-MS (ANOVA Tukey HSD on $\log_2(\text{fold change})$, $n=3$ cultures per strain).
- (D) Outline of mono-colonisation mouse experiment verifying *in vivo* hypoxanthine consumption. 3 female GF Swiss Webster mice were oral gavaged with $\sim 2 \times 10^6$ colony-forming units (CFUs) of either *B. longum* or *Lachnospiraceae* sp.2_1_58FAA and co-housed for the duration of the experiment. Hypoxanthine was supplied in drinking water to mimic exogenous production by the microbiome. On day 4 after, gavage mice were sacrificed and cecal contents were collected.
- (E) Hypoxanthine and xanthine pool size was determined in cecal contents using enzyme assays. Samples were corrected for baseline levels of H_2O_2 in the sample based on parallel reactions without Xanthine oxidase enzyme (Welch t test on averaged duplicate samples). Error bars indicate standard error of the mean (SEM). (* $P < 0.05$, ** $P < 0.01$)

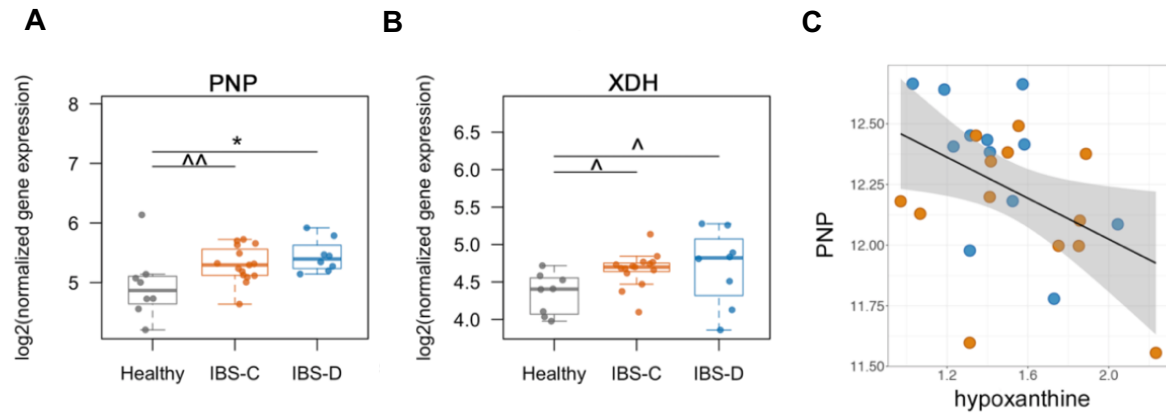


Figure 4.11. Evidence for hypoxanthine breakdown by the host colonic mucosa.

(A) Purine nucleoside phosphorylase (PNP) gene expression in colonic biopsy tissue (ANOVA Tukey HSD, *FDR<0.05, ^^FDR<0.005).

(B) Human xanthine dehydrogenase (XDH) gene expression in colonic biopsy tissue (ANOVA Tukey HSD, ^FDR<0.05).

(C) Lasso correlation plot between hypoxanthine and PNP (FDR <0.001).

4.4.7 Increased excretion of primary bile acids in IBS-D patients

BA are known for their secretory and motility stimulating effect in the colon and therefore are readily investigated in IBS patients.^{145,146} In this study, peak intensities of the nine BA most frequently implicated in IBS were extracted for univariate analysis. BA modification by the gut bacteria involves an initial deconjugation stage mediated by BSH followed by dehydroxylation into secondary BA. In humans, glycine conjugation is more common than taurine conjugation¹⁹². This is reflected in the results where considerably less tauro-conjugated BA were detected compared to glycine conjugates (**Figure 4.12A-B**). Taurochenodeoxycholic acid (TCDCA) was the only conjugated primary BA increased in IBS-D patients (**Figure 4.12A**) whereas glycine conjugated primary BA were all significantly reduced in IBS-C patients (**Figure 4.12A-B**). This pattern continued for the deconjugated primary BA in IBS-C patients where they were also significantly reduced. IBS-D patients on the contrary excreted much more primary BA. The consistent observations between the conjugated and deconjugated BA suggests that bacterial BSH activity was efficient and comparable between IBS patients and HC. No difference in secondary BA was observed (**Figure 4.12C**) except for UDCA, another secondary BA derived from bacterial epimerization of CDCA was significantly reduced in IBS-C patients. Increase in primary BA in IBS-D patients are consistently reported in the literature.^{145,146} By applying the same integration method used for associating bacteria structural variable region and metabolite like the one presented in **Figure 4.10A-B**,²⁴⁵ a number of significant associations between structural variable regions in the genome of *Blautia wexlerae* DSM 19850 and *Eubacterium rectale* DSM 17629 with CA and CDCA were identified. Some of the associated gene regions could encode a novel process for the modification of primary bile acids. In a separate physiological experiment, the Mayo team demonstrated the ability for CDCA to increase ion flux which indicates increased water secretion when CDCA was applied to the basolateral membrane of proximal colon mucosa submucosa preparations. This experiment supports the rationale that excess CDCA induced colonic water secretion which may contribute to the development of diarrhoea.

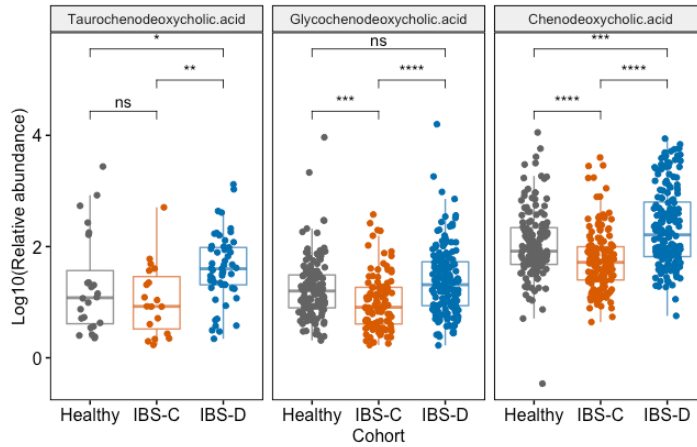
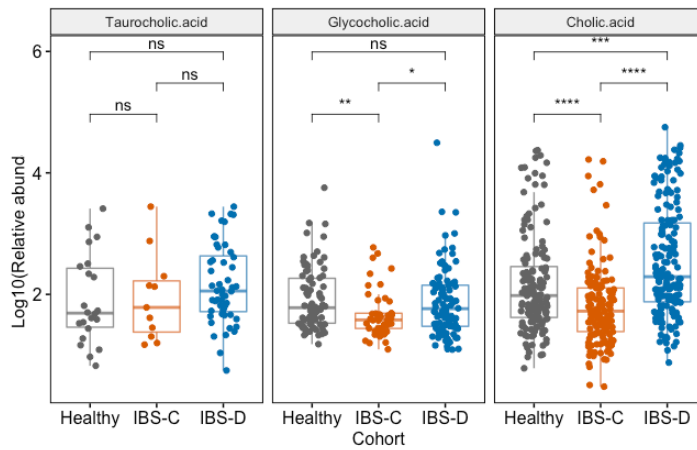
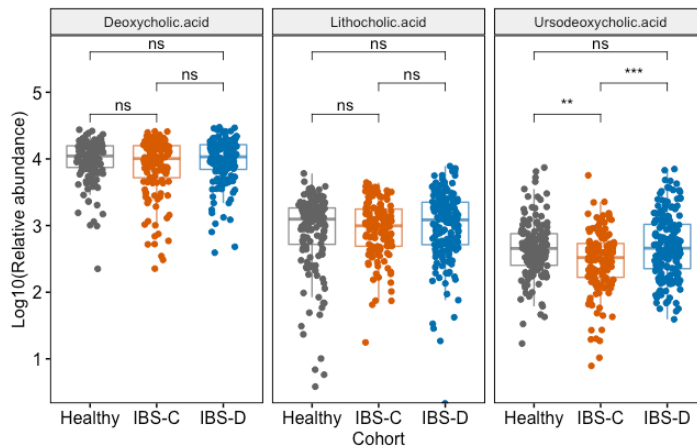
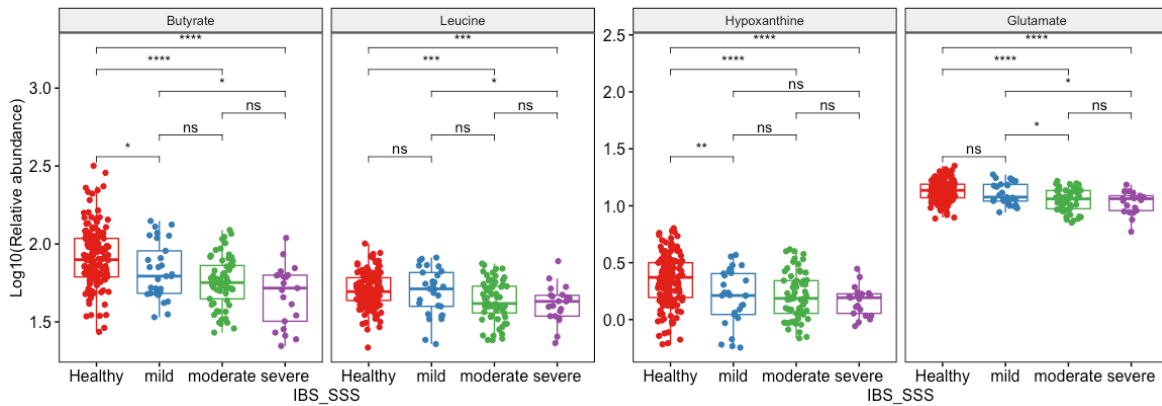
A**B****C**

Figure 4.12. Elevated stool primary bile acids (BA) in IBS-D patients. Box plots of log₁₀ relative abundance of **(A)** taurochenodeoxycholic acid (TCDCA), glycochenodeoxycholic acid (GCDCA), chenodeoxycholic acid (CDCA) and **(B)** taurocholic acid (TCA), glycocholic acid (GCA), cholic acid (CA) and **(C)** deoxycholic acid (DCA), lithocholic acid (LCA) and ursodeoxycholic acid (UDCA) in stool samples determined with LC-QToF. (Kruskal – Wallis test * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$, **** $P < 0.0001$, FDR corrected.)

4.4.8 IBS symptom severity is associated with luminal metabolite levels

Many luminal metabolites described so far demonstrate diagnostic potential in differentiating IBS subtypes from HC. Since IBS patients often experience remission and exacerbation in symptoms, the ability for metabolites to differentiate between symptom severity was investigated. An IBS Symptom Severity Score (IBS SSS) was recorded by patients on a monthly basis around the time of their stool collection. IBS SSS (0 to 500) is a cumulative metric of a patient's abdominal pain intensity, frequency, distension, dissatisfaction with bowel habits and influence of IBS on life in general. Patients were grouped according to their SSS into mild ($SSS < 100$), moderate ($100 < SSS < 300$) and severe ($SSS > 300$). Result showed that butyrate, leucine and glutamate were significantly different between mild and severe group of IBS-C patients (**Figure 4.13A**). Cholic acid was found to differentiate between mild and the other two groups in IBS-D patients whereas glutamate was only able to differentiate between mild and the moderate group in IBS-D patients (**Figure 4.13B**). Interestingly, while other metabolites returned to a healthy-like level in IBS-C patients with mild symptoms, hypoxanthine and butyrate did not. This suggests that mucosa energy deprivation may be a persistent pathological feature in IBS-C patients despite symptom improvements.

A IBS-C



B IBS-D

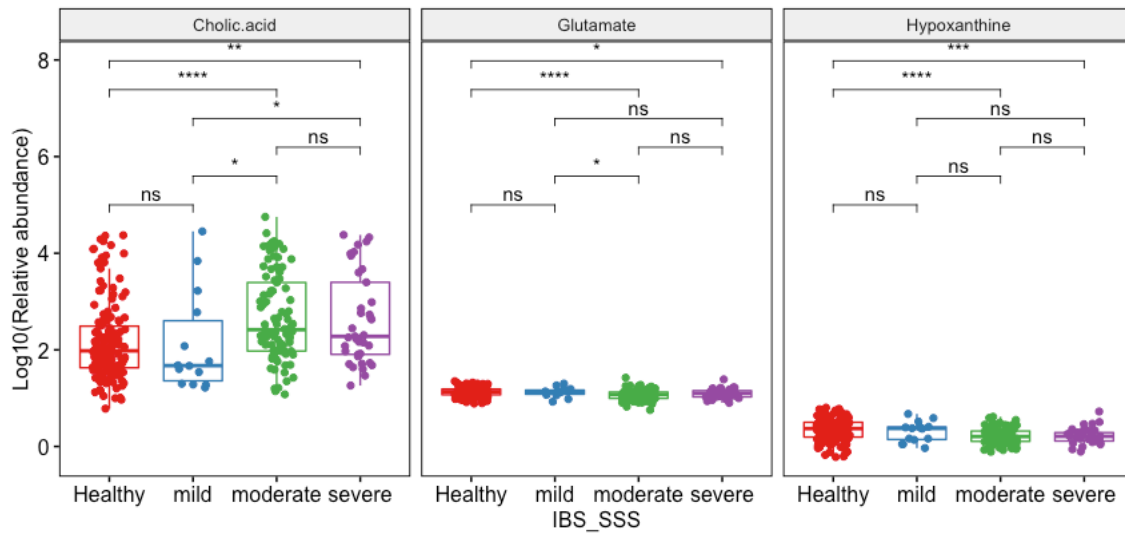


Figure 4.13. Luminal metabolites could differentiate between symptom severity in IBS-subtypes.

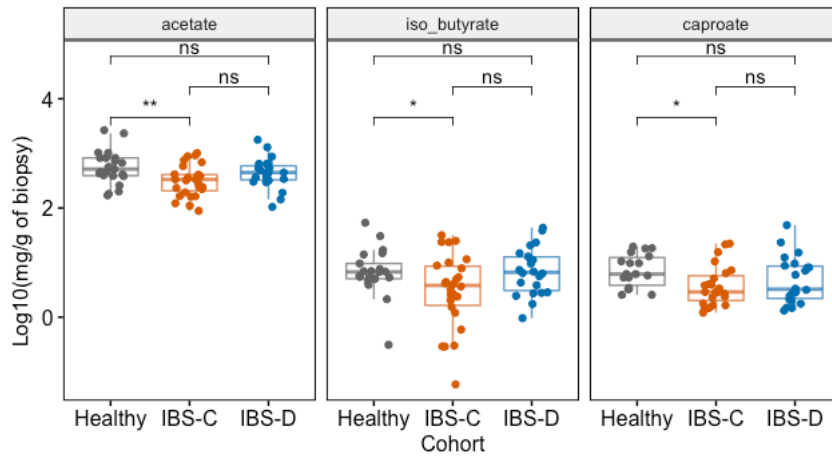
(A) Boxplots comparing log₁₀ relative abundance of NMR peak integrals for butyrate, leucine, hypoxanthine and glutamate in stool samples between healthy and symptom severity groups of IBS-C patients. (Kruskal – Wallis test * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$, **** $P < 0.0001$, FDR corrected)

(B) Same as (A) but for cholic acid, glutamate and hypoxanthine in healthy and symptom severity groups of IBS-D patients.

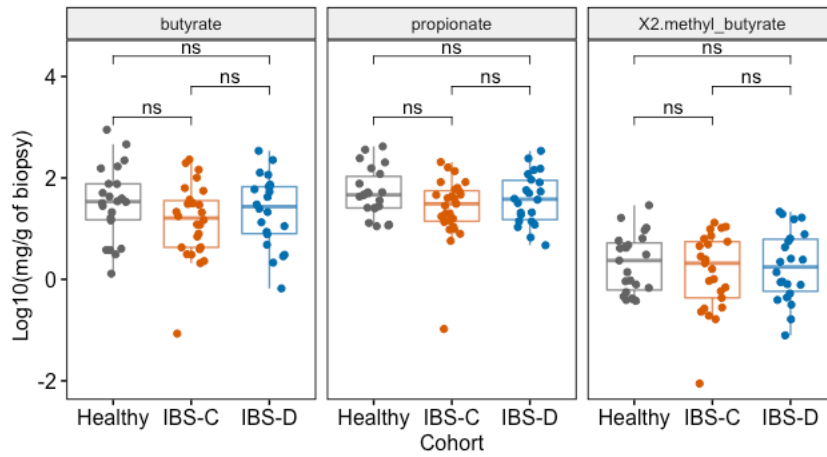
4.4.9 A rare view into the mucosa metabolome in IBS patients

Results in previous sections showed that host gene expression is altered in IBS patients and microbial metabolites in particular SCFA are known to modulate gene expression levels in the intestinal mucosa. However, compared to stool samples, little is known about the presence of microbial metabolites in the mucosa compartment. A targeted GC-TQ method was used to measure a panel of short-chain, branched-chain and medium-chain fatty acids in sigmoid biopsies. Out of the ten metabolites measured, a significant decrease in acetate, isobutyrate and caproate was observed in IBS-C patients. This result partially reflects the lower SCFA abundance in the stool sample of IBS-C patients (**Figure 4.14**).

A



B



C

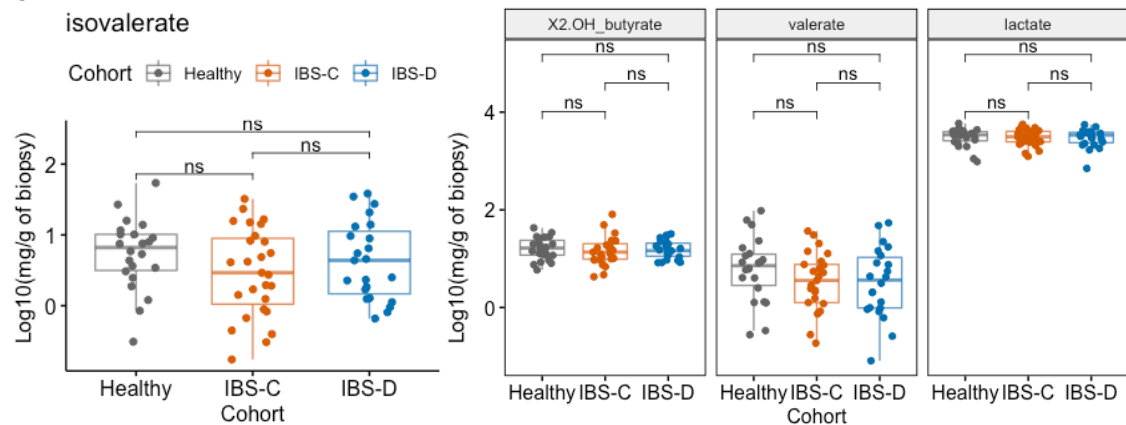


Figure 4.14. Sigmoid biopsy metabolites

- (A) Box plot showing concentration in mg per g of sigmoid biopsy sample for acetate, isobutyrate and caproate. (Kruskal-Wallis test, * $P < 0.05$, ** $P < 0.01$, FDR corrected).
 (B) Same as (A) but for butyrate, propionate and 2methylbutyrate.
 (C) Same as (A) but for isovalerate, 2-OH-butyrate, valerate and lactate.

Relative abundance of the three major SCFA were compared between the lumen and the mucosa (**Figure 4.15A,B**). Acetate dominated both the stool and the biopsy samples whereas butyrate and propionate were relatively similar in abundance. Compared to the stool samples, butyrate and propionate constituted an even smaller proportion in the biopsies. This is not surprising since butyrate is metabolised rapidly by the intestinal epithelial cells.³⁹ The relative abundance ratios for acetate : propionate : butyrate were around 7:2:1 in the stool samples and 8:1:1 in the biopsies. This result largely agreed with the literature where butyrate and propionate have similar ratios, the difference between them and acetate is more variable.^{41,193}

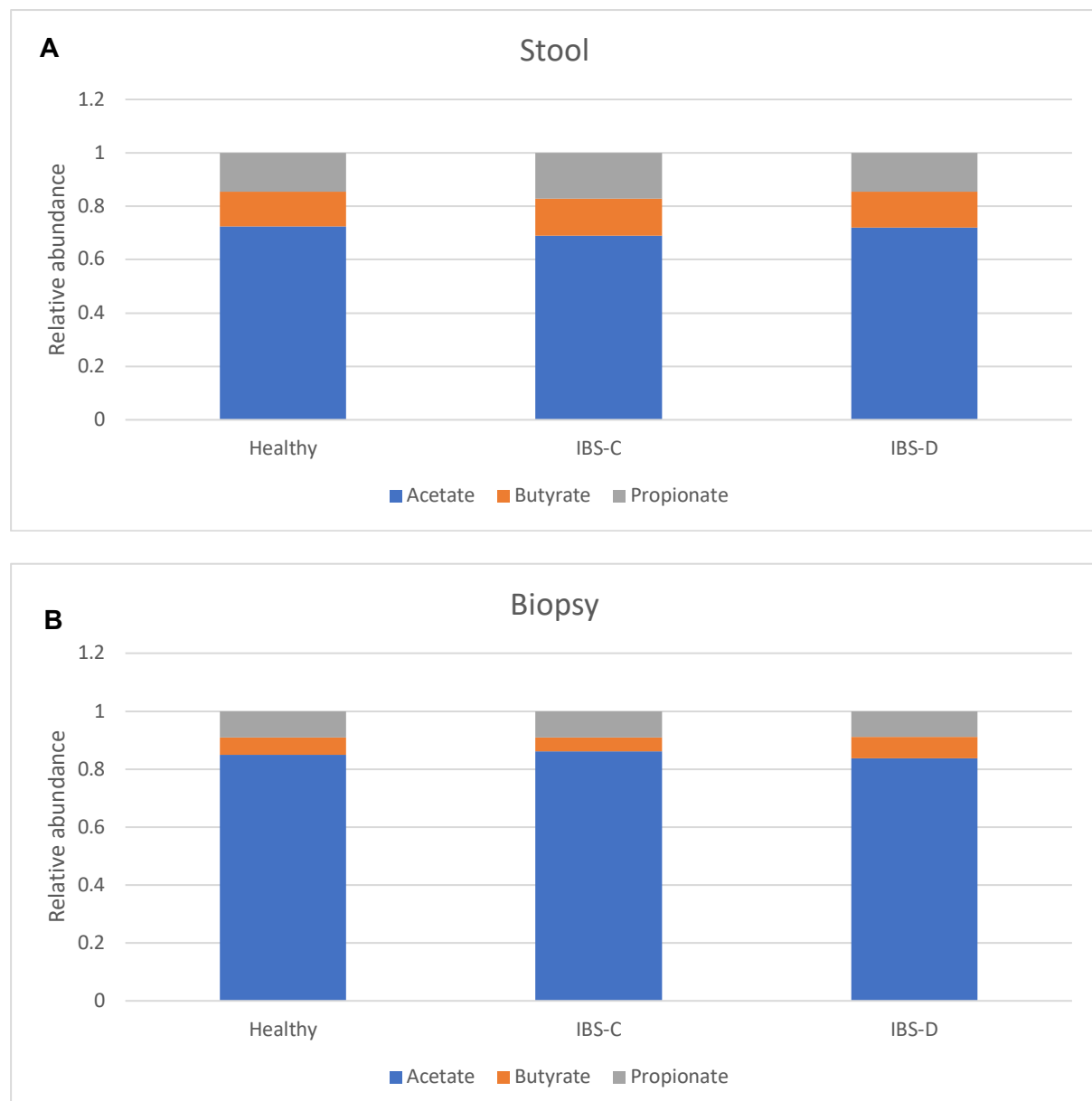


Figure 4.15. Bar charts showing the average relative abundance of acetate, butyrate and propionate in (A) the stool samples measured using 600 MHz H^1 NMR and in (B) the sigmoid biopsies measured using GC-TQ targeted quantification.

The effect of SCFA on modulating mucosa gene expressions is well documented in the literature.⁴⁵⁻⁴⁷ Using Spearman correlation, a number of significant associations between host gene expression and mucosal SCFA levels were identified in IBS-C patients (**Figure 4.16**). Isobutyrate was negatively associated with B cell regulating genes, FCER2 and FCRL1.¹⁹⁴ It was also positively correlated with OLFM4, a gene that was found to be selectively expressed in inflamed colonic epithelium.¹⁹⁵

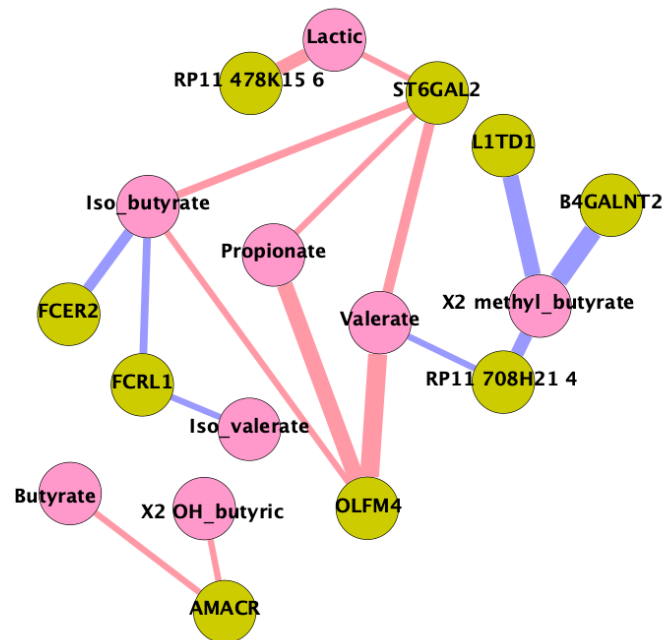


Figure 4.16. Network plot showing Spearman correlations between mucosal SCFA (pink nodes) and host genes (green nodes) that have passed the significance threshold of $r > 0.1$ and $FDR < 0.25$. Blue edges represent negative correlations and red edges represent positive correlations. Genes with $\log_2$ fold change ≥ 1 between IBS-C and HC cohorts were selected for analysis. Strength of the correlation (r) is represented by the thickness of the edges.

4.4.10 The lumen microbiome is richer than the mucosa

Although the mucosa may be harbouring a smaller variety of microbial species than the lumen as indicated by the significantly lower Chao1 richness score, these species are just as evenly distributed as the luminal community as indicated by the Shannon evenness index (**Figure 4.17**). There was a borderline significance of lower Shannon score in the lumen of IBS-D patients compared to the mucosa which suggest that the mucosa community is slightly more stable, perhaps by adhering to the host tissue their community structure is less likely to be disturbed by the luminal flow. However, a previous study observed significantly higher Shannon index in the faeces of IBS patients compared to their sigmoid colonic mucosa.²⁰ This disagreement may be caused by the colon cleansing procedure used in this study which is known to alter the mucosal microbiome.¹⁹⁶

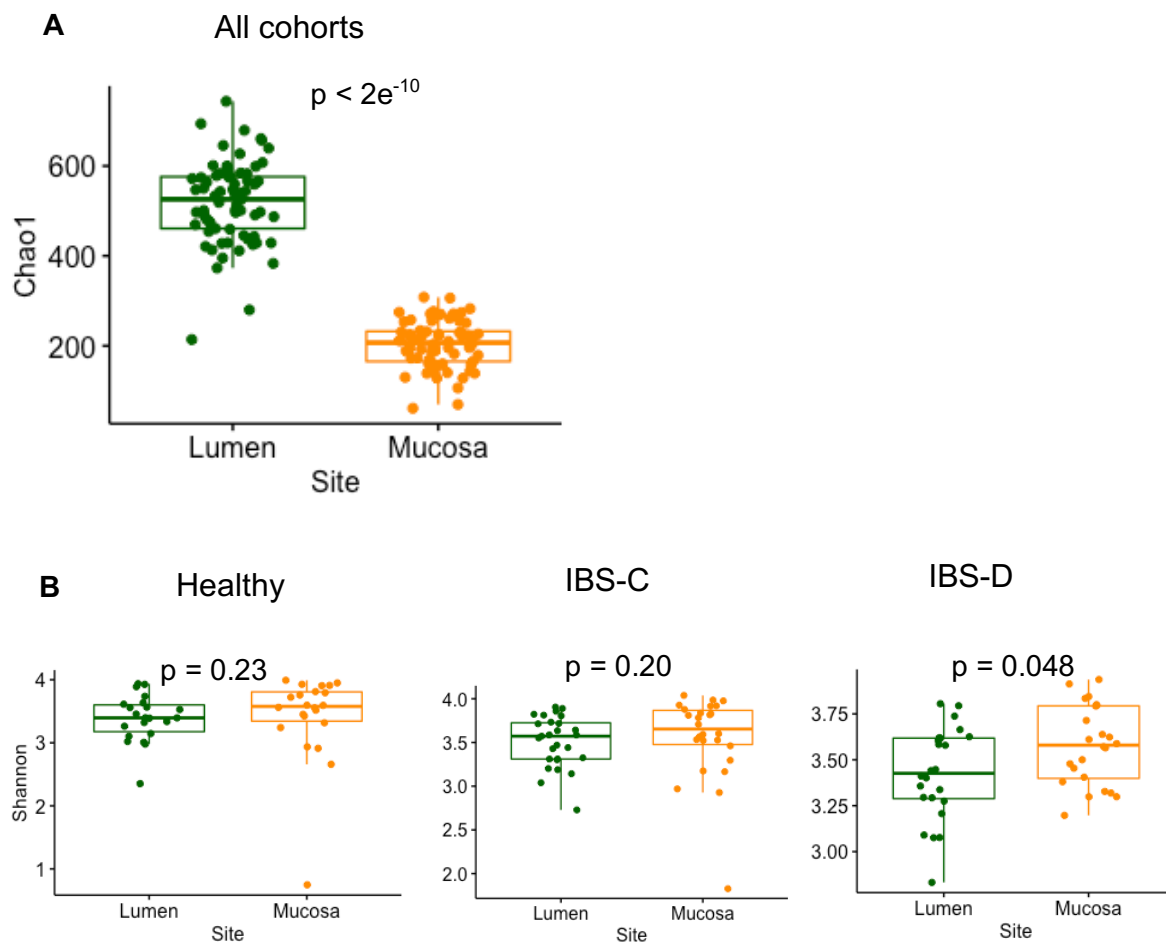


Figure 4.17. Comparing alpha-diversity in the lumen and the mucosa.

(A) Boxplot showing Chao1 richness score in the lumen and the mucosa of all cohorts. Only subjects who provided both stool and biopsies samples were included. (Mann-Whitney U test, p value displayed.)

(B) Same as (A) but comparing Shannon evenness index in each cohort.

4.5 Discussion

Outcomes from this longitudinal multi-omics study can be divided into three main areas.

4.5.1 IBS-subtype specific metabolic alterations are associated with the colonic microbiota and host physiology

If the cohorts were to be put on a metabolic abundance scale of SCFA and primary bile acids, IBS-C and IBS-D would sit at opposite end of the scale with HC in the middle. This was probably a reflection of the extreme differences in intestinal transit between the subtypes. It has been shown that SCFA concentrations in the distal colon are lower than the proximal section of the gut where saccharolytic activity is high.⁴¹ This effect is being exacerbated by longer transit which leads to further depletion of carbohydrates towards the distal colon and could result in the disappearance of butyrate-producing bacteria where carbohydrates are their primary food source. Indeed, this study found that reduced butyrate in IBS-C patients was associated with the disappearance of *Roseburia spp.* which are prominent butyrate-producers in the gut. Lower abundance of the *Roseburia* genus have also been reported in patients with IBD and colorectal cancer.¹⁹⁸⁻²⁰⁰ Interestingly, butyrate was negatively associated with *Ruthenibacterium lactatiformans*. This lactate-producing strain from the *Ruminococcaceae* family has sialidase encoding genes which suggests their potential role at cleaving the sialic acid residues from complex glycoproteins, such as mucins.^{183,201} One study showed that an increase of *Roseburia spp.* was associated with a decrease in mucin-degraders when a diet rich in β -mannans was fed to mice.¹⁹⁷

The longer colonic transit in IBS-C patients is believed to be the main driver of enhanced proteolytic activity as there was an increase in BCAA-degrading species and reduced BCAA in the stool samples. Bacteria protein fermentation often leads to the production of harmful metabolites such as ammonia, hydrogen sulphide and *p-cresol* which have been associated with colon cancer.⁶⁶ IBS-C patients also had increased luminal microbial diversity and richness which challenges the concept that higher microbial diversity is associated with a healthy gut.

The stool metabolome in IBS-D patients was characterised by higher primary bile acids. Primary BA and SCFA are known to stimulate colonic secretion and colonic muscle contractile response which was also demonstrated by the ionic flux experiments in this study.^{104,108} Once again, as many studies have shown in the past, excess primary BA contribute to an increase in colon motility and water secretion which can lead to diarrhoea.^{109,145} It is possible that BA

malabsorption which was not determined in IBS-D patients in this study could also cause excess excretion of primary BA. However, the fact there was no change in secondary BA suggests a lack of microbial biotransformation of primary BA may be partly responsible for this effect.

4.5.2 Comparing the metabolome in the lumen and the mucosa

For the first time, SCFA ratios could be directly compared between the lumen and the mucosa where butyrate and propionate were proportionally reduced in the mucosa indicating their rapid metabolism at the intestinal epithelium. While the effects of SCFA have been extensively studied, little is known for branched chain and medium chain fatty acids probably due to their lower abundance in the lumen. As this study showed, isobutyrate was associated with the expression of host genes involved in immune responses, highlighting a need to further explore such effects. The disagreement between faecal and biopsy microbiome diversity between this study and previous publications was suggested to be caused by differences in how biopsies were collected from the colon (*i.e.* the inclusion of a washout procedure prior to endoscopy or not). The sampling site is also important given the spatial variability of microbiome in the colon.²² The research community interested in the study of mucosa microbiome is now calling for a more precise sampling device which also aims to minimise complications associated with conventional endoscopy.¹⁹⁶

4.5.3 Purine salvage underlies the pathophysiology of both IBS subtypes

Using multi-omics integration, this study identified purine salvation as a novel host-microbial co-metabolic pathway in IBS. Purine nucleotides make up one of the base pairs in DNA which are needed in large quantities by the intestinal mucosa to sustain rapid cellular turnover and maintain barrier function. Cells rely on two different pathways for the production of purine nucleotides, the preferred pathway uses exogenous hypoxanthine via purine salvage and the other one is the *de novo* synthesis pathway which involves breaking down ATP.^{190,191} It was hypothesised that when hypoxanthine supply in the gut lumen becomes limited, as in the case of IBS patients, intestinal epithelial cells respond by activating the ATP-consuming *de novo* synthesis pathway to salvage hypoxanthine. As a result, there could be less ATP available for the intestinal mucosa to maintain its normal cellular function which in turn reduced its capacity for mucosa repair. This lower energy state seemed to be a persistent pathological feature in IBS patients since hypoxanthine level did not change between patients with mild symptoms and those on the severe end of the spectrum. There are many more hypoxanthine associated

bacterial gene regions not investigated in this study. They serve as valuable resources for the research community to identify microbial species involved in the degradation and biosynthesis of hypoxanthine which in turn pave the way for the development of a new generation of probiotics for IBS.

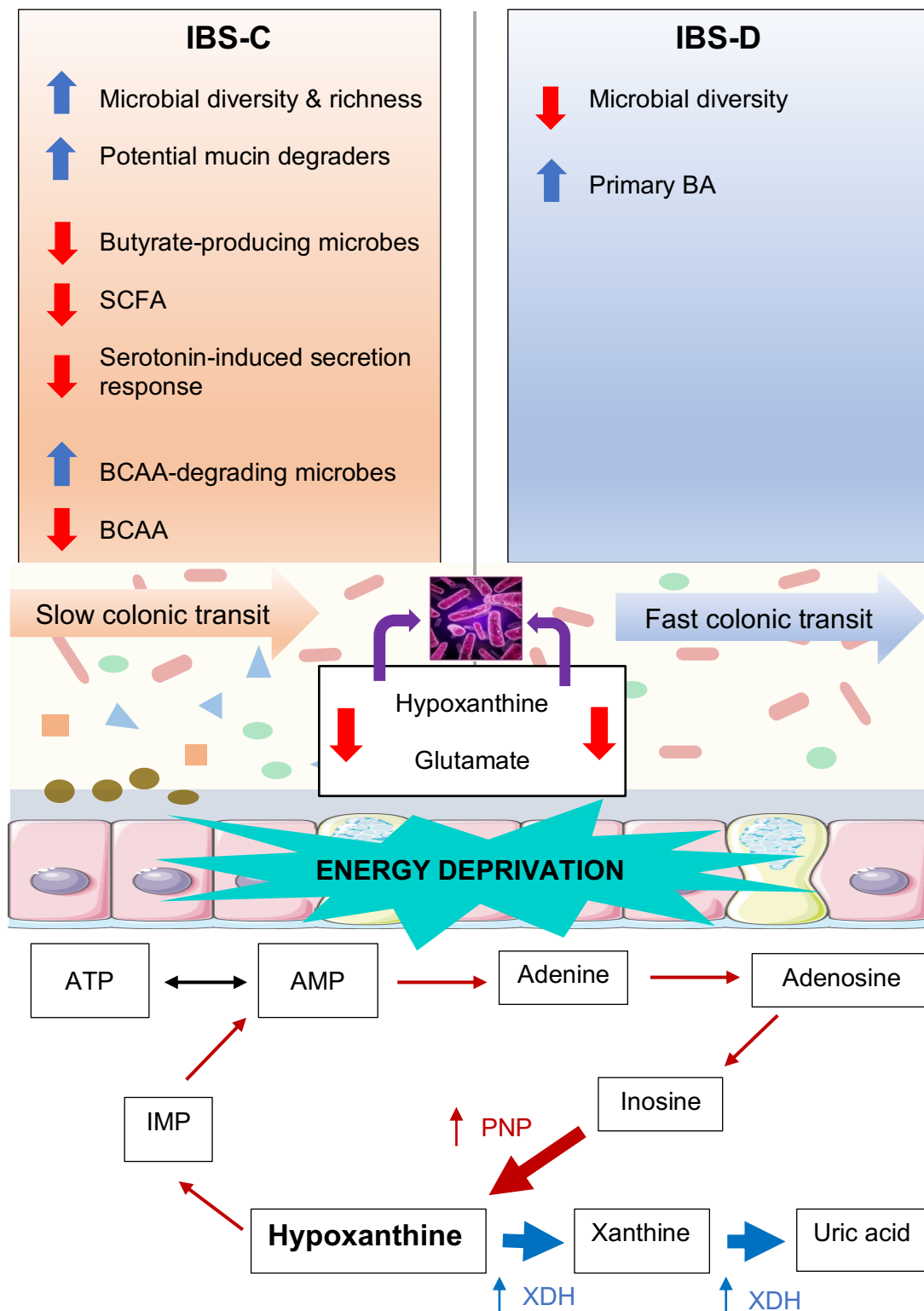


Figure 4.18 Summary of the main findings described in this chapter and the proposed host-microbial co-metabolic pathway underlying the pathophysiology of IBS. Red arrows represent the host purine salvage pathway and blue arrows represent hypoxanthine degradation by the host xanthine hydrogenase (XDH) enzyme.

4.6 Study limitation and future work

While this study identified several strong physiological relevant alterations in the metabolome and the microbiome, their causal relationship requires further mechanistic study to elucidate.

SCFA and BA disturbances in IBS patients have been reported in the past.^{205,109} However, this study was the first to report lower hypoxanthine in both IBS subtypes therefore findings should be validated in an independent cohort. The purine salvage pathway requires further elucidation since there are many more enzymes involved in the pathway not investigated. Although XDH gene expression was increased, there was no direct measure of its enzyme activity. XDH breaks down hypoxanthine to produce uric acid which does not have a detectable NMR-proton at physiological pH.²⁰² However it can be detected using LC-MS/MS with high precision and sensitivity.²⁰³

Due to instrument availability and time constrain, only ten metabolites mostly SCFA were measured in the biopsies. Literature reports on the mucosa metabolome is still limited, therefore a profiling method would be better suited at detecting a wide range of metabolites such as the GC-TOFMS method used in the study of colorectal cancer.²⁰⁴

4.7 Conclusion

Using a combination of targeted and untargeted metabolomics techniques together with multi-omics integration, this chapter identified several microbial-mediated metabolic disturbances in the transverse gut compartments specific to each IBS subtype. In addition, a novel host-microbial co-metabolic pathway involving purine salvage was identified to underlie the pathophysiology of IBS. Hypotheses generated from this study have opened up new avenues of research into therapeutics that could restore the functionality of the microbiome in IBS patients.

Chapter 5

Metabolite diffusion through the mucus compartment

5.1 Introduction

In Chapter 4, SCFA ratios in the lumen were observed to differ from those of the mucosa. In this chapter the role of mucus layer in hindering the trans-compartmental passage of SCFA was investigated for its role in limiting the supply of SCFA to the mucosa.

Mucus is a dynamic gel-like compartment separating the epithelium from the microbiota and harsh chemical substances in the gut lumen.²⁴ Aside from its protective functionality, mucus plays an important role in microbial-host symbiosis as it provides a rich source of carbohydrates in the form of mucin glycans for bacteria to feed on and to use them as anchoring sites to establish a stable community.^{28,29} Several *in vivo* and *ex vivo* experiments showed that the consistency of the mucus layer depends on the amount of water and bulky substance in the lumen.^{206,207} In one mouse study, investigators demonstrated that formation of the sterile inner mucus layer in the distal colon is facilitated by the depletion of water in the lumen. In the proximal colon, where the faecal mixture still retains a large amount of water, no inner sterile layer was found.²⁰⁶ It is probable that the loose mucus structure in the proximal colon facilitates SCFA diffusion since it has been shown that SCFA are efficiently absorbed in the proximal colon with very little remain in the distal colon.^{208,209}

In 1986, Smith *et al.* were the first group to discover that mucus decreased the rate of butyrate diffusion in an experiment using mucus scraping from the distal colon of pigs.²¹⁰ Although more advanced mucosal diffusion systems have been developed since, most have been used to study the diffusion of drug molecules.²¹¹ As such, knowledge of the diffusion of SCFA through the mucus layer remains limited. Moreover, in the Smith *et al.* work, no cell layer was incorporated in the model system and therefore the fate of the SCFA after passing through the mucus layer is unclear. Specifically, it is unknown what proportion of diffused SCFA migrate to the basolateral side. To investigate this, an *in vitro* model was built to study the diffusion of acetate, propionate and butyrate through goblet cells under normal conditions with

mucus as well as under mucus-free conditions. The diffusion rate and the consumption preference by the goblet cells for the three SCFAs was also explored.

5.2 Methods

5.2.1 Cell culture

Mucus producing cell line HT29-MTX-E12 at passage 51 (lot no: 16D017) was purchased from Public Healthy England, ECACC. Cell culture reagents were purchased from Thermo Fisher Scientific UK (Paisely, UK), unless otherwise stated. DMEM (high glucose, no glutamine), Fetal Bovine Serum (FBS), L-Glutamine (200mM), non-essential amino acids (100X concentration), Penicillin-Streptomycin (Pen-Strep, 100x concentration), Phosphate-buffered saline (PBS), T12 and 24-wells Transwell plate with 0.4µm pore polyester membrane inserts (Corning, Inc.). Cell culture medium was made with DMEM + 2mM Glutamine + 1% non-essential amino acids + 10% foetal bovine serum + 1% Pen-Strep. Sodium acetate, sodium butyrate and sodium propionate were purchased from Sigma-Aldrich. Cells were allowed to grow in an incubator operated at 37°C with 5% CO₂ until desired numbers of cells were obtained. Cells at passage number 54 were used for cytotoxicity assay and mucus diffusion experiment.

5.2.2 Confocal fluorescence microscopy

Fluorescent dye for staining extracellular mucus SBA-Alexa Fluor 488 (Cat.no. L11272) and cell membrane FM4-64 (Cat.no. T13320) were purchased from Thermo Fisher Scientific UK. For imaging a 35 mm dish was used. Cells were seeded at a density of 50,000 cells/cm² to three Transwell inserts in a 12-well Transwell plate. Cells were allowed to grow for ten days before imaging. On the day of imaging, culture media was removed from the Transwells, cells were washed with 0.5 mL of ice-cold PBS to remove FBS which can produce unspecific staining. The washed insert was placed in a 35 mm imaging dish and incubated on ice for 10 minutes with 200 µL of 100 µg/mL SBA-Alexa Fluor 488. Then, 200 µL of 2.5 µg/mL of FM4-64 diluted in ice-cold H₂O and PBS was added for another minute of incubation on ice. The membrane insert was carefully cut out with a scalpel on a glass slide then moved back into the centre of the imaging dish. Image was taken on a Leica SP5 upright confocal fluorescence microscope with excitation wavelengths set at 519 nm and 558 nm. As shown by the 2D and 3D image (**Figure 5.2**) taken on a 1 mm x 1 mm area of the insert, mucus coverage was not homogenous as some areas showed intense mucus staining and others were without any. The nucleus staining (red dye) also showed varied intensity indicating uneven spread of cells which might have caused some areas to produce more mucus than the others.

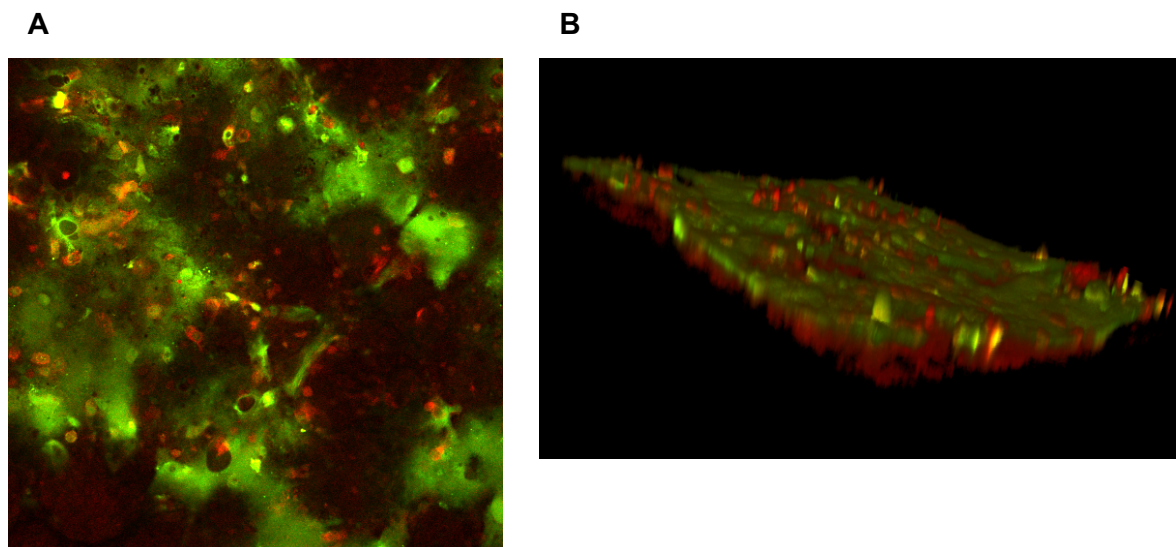


Figure 5.1. Confocal fluorescence microscopy of a 1mm x 1mm section of membrane covered with HT29-MTX-E12 cells grown for 10 days. View from the (A) 2D - X and Y axis and (B) 3D - X, Y, Z axis. Green dye stained for extracellular mucus. Red dye stained for cell membrane.

5.2.3 Lactate dehydrogenase (LDH) cytotoxicity assay

Optimum concentrations for sodium acetate, butyrate and propionate treatment solutions were determined using the LDH cytotoxicity assay. Upon cell death, LDH which is an enzyme present in the cytoplasm of many cell types is released into the cell culture medium. Thus, the level of LDH activity corresponds to the level of cell death and can be measured via a colorimetric method. The LDH test kit and Triton X-100 solution was purchased from Sigma-Aldrich. P53 cells were seeded into a 96-well plate at 30,000 cells/cm² and reached confluency after a day. A series of treatment solutions were made as shown in **Table 5.1**. Each treatment solution and a negative control (culture medium) was added to the wells in triplicate. Treatment lasted for 21 hours. On the day of the assay, three wells had a positive control (Triton X-100 solution, used for lysing the cells) added to them.

Solution	Acetate Concentration (mM)	Propionate Concentration (mM)	Butyrate Concentration (mM)	Preparation method
T1	200	67	67	0.17uL acetate, propionate and butyrate + 0.49mL media
T2	100	33	33	0.42mL of each stock + 3.75mL media
T3	75	25	25	0.75mL of T2 + 0.25mL media
T4	50	17	17	0.5mL T2 + 0.5mL media
T5	25	8	8	0.25mL T2 + 0.75mL media
T6	10	3	3	0.1 mL T2 + 0.9mL media
T7	1	0.3	0.3	0.1mL T6 + 0.9mL media
T8	0.1	0.03	0.03	0.1mL T7 + 0.9mL media
T9	0.01	0.003	0.003	0.1mL T8 + 0.9mL media

Table 5.1. Table of treatment solutions made for the LDH cytotoxicity assay.

An ELISA microplate reader was used to measure the absorbance of LDH activity at 450nm. Percentage of cell death was calculated using the following formula:

$$\text{Cell death (\%)} = \frac{(\text{sample value} - \text{negative control value})}{(\text{positive control value} - \text{negative control value})} \times 100\%$$

Result of the cytotoxicity assay is shown in **Figure 5.3**. An optimum concentration was one that provided the maximum amount of SCFA with a cell death of less than 15% to allow better detection in the basolateral compartment as lower concentration might result in all SCFA on the apical side to be consumed by the cells leaving trace amount flowing through to the bottom of the transwell. Therefore, a treatment solution consisting of 50 mM of acetate and 17 mM of each butyrate and propionate was appropriate.

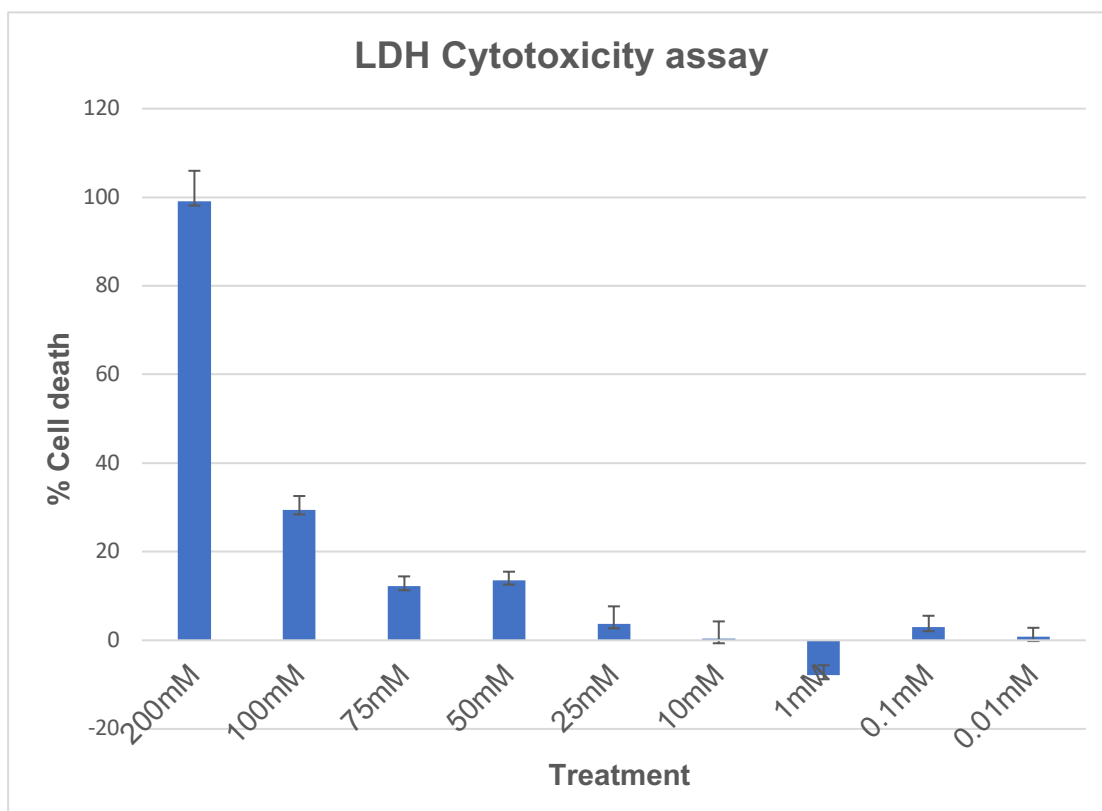


Figure 5.2. Bar charts with standard deviation error bars showing the average % cell death for each treatment solution. Labels indicate the concentration for acetate in each solution.

5.2.4 Mucus diffusion experiment

Given that the cell line did not produce a homogenous mucus layer as shown **Figure 5.1**, two additional experimental conditions were included with different volumes of reconstituted porcine mucus added on top of the endogenous mucus. The following mucus conditions were proposed:

1. Cells without endogenous mucus.
2. Cells with endogenous mucus.
3. Cells with endogenous mucus and 0.3 mm thick of reconstituted mucus.
4. Cells with endogenous mucus and 0.9 mm thick of reconstituted mucus.

The rate of SCFA diffusion was studied across the experimental groups changing the thickness of the mucus layer. For each mucus conditions, culture media at the apical and basolateral side was sampled at 15 min, 1 hr and 3 hr. Each time point had its own triplicate of wells to sample from to avoid having to replenish cells with culture media which would affect the original concentration of the SCFA. Thus, each mucus condition had 3 sampling time

points and each time point required a triplicate of wells making the total number of wells to 36. The experimental procedures consisted of the following: Cells were allowed to grow for 28 days as literature shows this resulted in a mucus coverage of 95% and above.²¹² The seeding density was 50,000 cells/cm². Cells were grown on 24-well Transwell inserts which had a smaller diameter at 6.5 mm compared to the 12-well 12 mm Transwell used for confocal imaging. Diffusion experiments were carried out on the 29th day post-seeding. For experimental condition 1, endogenous mucus was removed by incubating with 10 mM *N*-acetylcysteine (NAC) dissolved in PBS at 37 °C for 1 hour as recommended in literature;²¹¹ cells were allowed to grow normally for experimental condition 2; MUC2 mucine powder purchased from Sigma-Aldrich (M2378) were dissolved in culture media to a concentration of 2 mg/mL and this reconstituted porcine mucus was added to the wells of experimental condition 3 and 4. Each experimental condition was carried out in triplicate of wells. At baseline, 100 µL of culture media from the apical side and 600 µL from the basolateral side were collected and the Transwells were replenished with fresh media. After addition of the SCFA treatment solution at 50 mM of acetate and 17 mM of butyrate and propionate to the apical compartment, 100 µL from the apical side and 600 µL from the basolateral side were collected from wells dedicated to each time point. To measure the amount of SCFA trapped inside the mucus layer, the mucus layer in experimental conditions 2, 3 and 4 was also collected by incubating the Transwells with 10 mM of NAC at 37 °C for 1 hour.

5.2.5 GC-MS SCFA Quantitation

SCFA concentrations in the apical, basolateral and mucus compartment of the Transwells were quantified using the targeted GC-MS method as described in Chapter 2. The sample preparation procedure and instrument operating conditions were the same as the ones used for the human studies, the only differences were the calibration curves and the dilution factor where they were adopted for the SCFA concentration range used specifically for this diffusion experiment. Initially, an 11-point calibration curve ranged from 0.8 mM to 400 mM of acetate and 0.04 mM to 20 mM of butyrate/propionate was tested and we found that instrument saturation started to occur at around 2 mM where the curve is no longer linear (**Figure 5.3**).

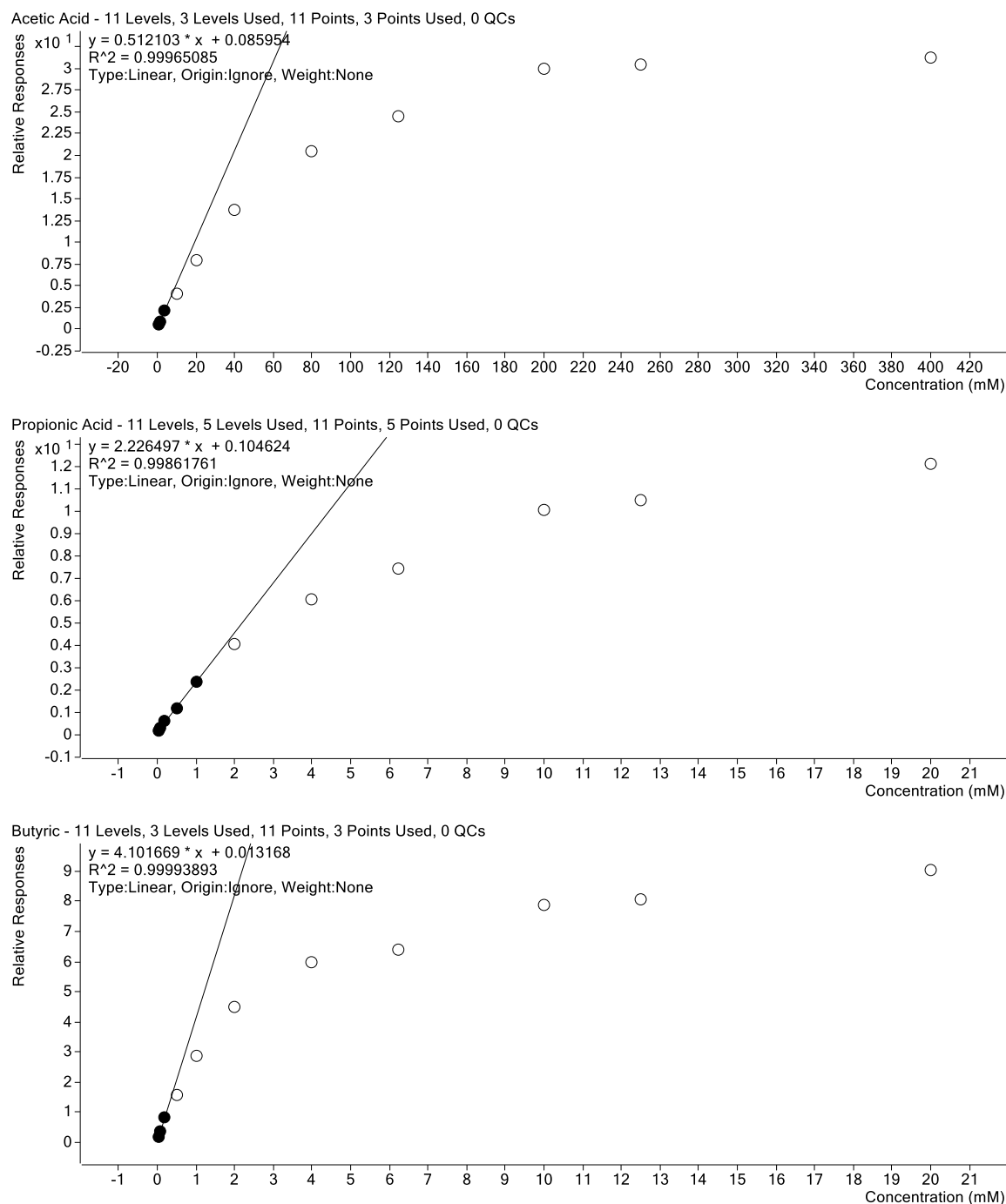


Figure 5.3. GC-MS calibration curve for (A) acetate, (B) propionate and (C) butyrate. Concentration in mM is displayed on the X-axis whereas the Y-axis is showing the relative response of SCFA and the internal standard. The curved trajectory formed by the white circles indicated that beyond 2mM, the relationship between the relative response and the SCFA concentration is no longer linear.

A new calibration curve was made with concentrations for acetate ranged from 0.8 μM to 400 μM and butyrate/propionate from 0.16 μM to 80 μM (Figure 5.5). Several dilution series were trialled on the QC samples of the three compartments to ensure the final concentrations of SCFA fall within the range of the calibration curves. Samples in the apical compartment were

found to be much more concentrated than samples in the mucus and the basolateral compartment. The optimum dilution factor for the apical samples was 1000 times and for the less SCFA concentrated mucus and basolateral samples, the dilution factor was 70 times.

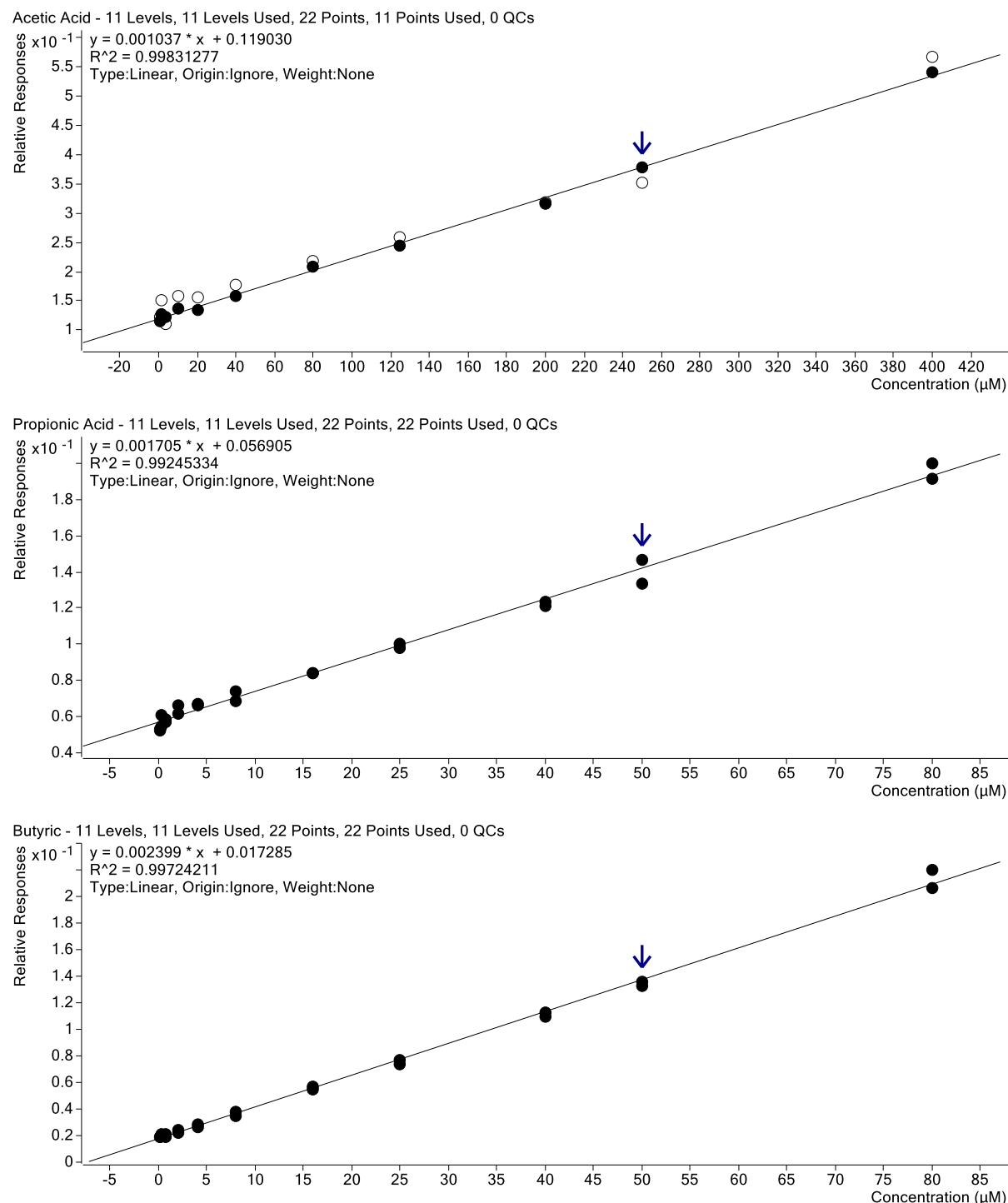
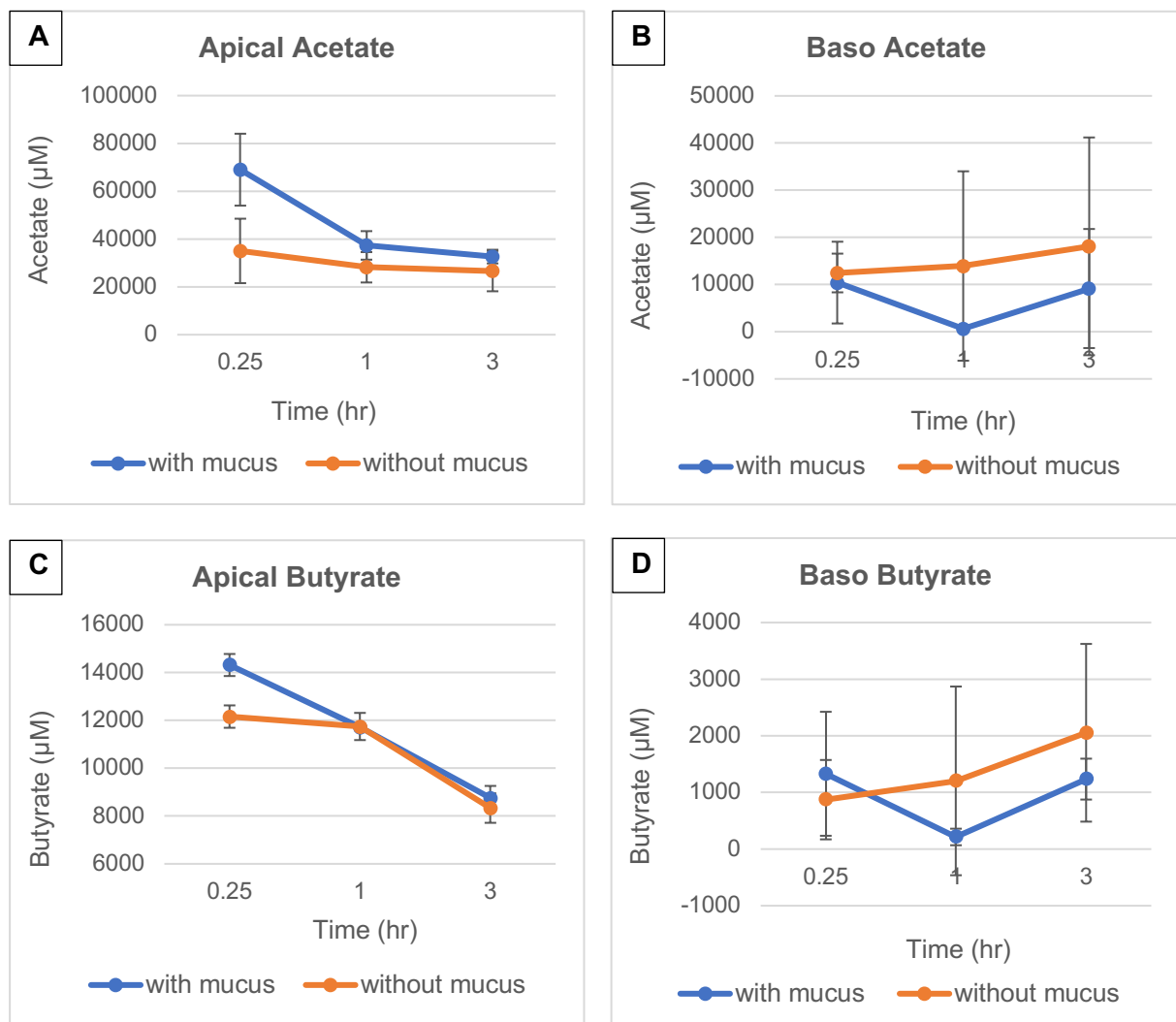


Figure 5.3. Optimised GC-MS calibration curves for (A) acetate, (B) propionate and (C) butyrate where all 11 calibration points fall within the fitted linear line. Concentration in μM is displayed on the X-axis whereas the Y-axis is showing the relative response of SCFA and the internal standard.

5.3 Results

5.3.1 Mucus reduced the flow of SCFA

By measuring SCFA concentration in the apical and the basolateral (BL) compartment, the rate of diffusion was compared between experiments with and without a mucus layer. It was evident that mucus reduced the flow of SCFA as shown in **Figure 5.4A,C,E** where across all time points more acetate, butyrate and propionate remained in the apical compartment of Transwells with mucus. Since less SCFA were detected in the apical compartment of Transwells without a mucus layer the amount of SCFA that diffused into the BL compartment was investigated. At two of the three sampling points (1 hr and 3 hr), higher concentrations of SCFA were detected in the BL side of Transwells without mucus (**Figure 5.4B,D,F**). However, results at 15 min (0.25 hr) showed higher concentration in the BL compartment of Transwells with mucus which suggests that more SCFA move into the BL compartment when a mucus layer is present.



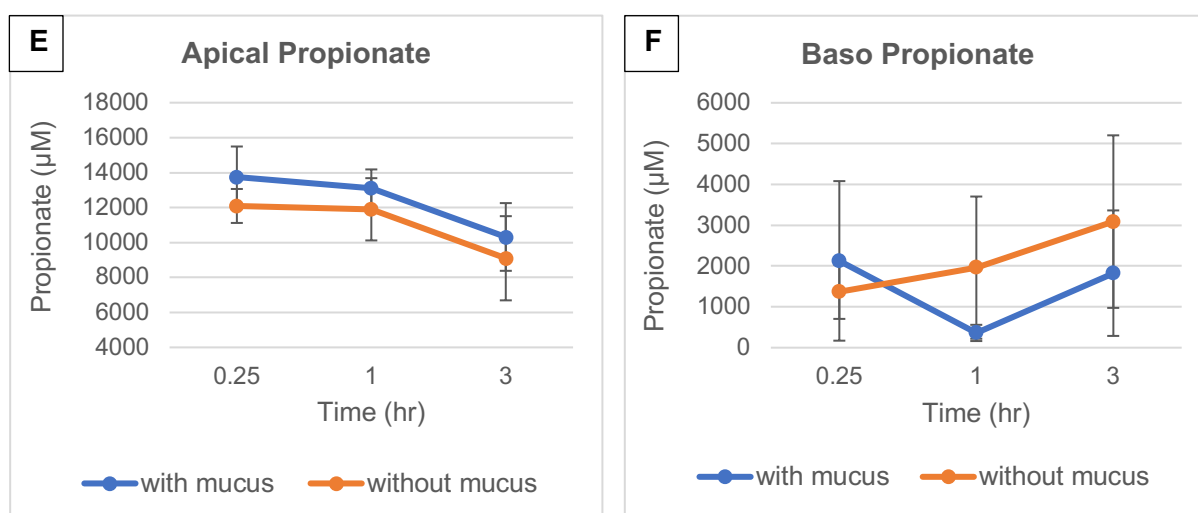


Figure 5.4. Line graphs showing the average concentration (μM) of (A-B) acetate, (C-D) butyrate and (E-F) propionate in the apical compartments and the basolateral compartments over time (hr). Blue line indicates experiments with mucus and orange line indicate experiments without mucus. Error bars indicate $\pm$ standard deviation of three replicates of measurements.

5.3.2 Mucus temporarily traps SCFA

In addition to measuring SCFA concentrations in the apical and the BL compartments, the amount of SCFA, if any, trapped inside the mucus during their passage between these two compartments was also measured. As shown below in **Figure 5.5A, B and C**, the concentration against time graph in the BL compartment formed a V-shape whereas an inverted V-shaped was formed for the mucus compartment. The opposite trend of the two compartments suggests that SCFA were able to diffuse freely across the mucus layer without much obstruction during the first 15 min and as more SCFA enters the mucus it created more competition for routes through the mucus hence more SCFA are found inside the mucus at 1 hr and less SCFA were flowing into the BL compartment. At the same time, SCFA decreased in the BL compartment suggesting that cells may be consuming SCFA from the BL compartment. This sudden drop in concentration in the BL compartment coincided with an increase in mucus concentration suggests that as SCFA were trapped inside the mucus, SCFA that have already diffused into the BL compartment started to enter the cells.

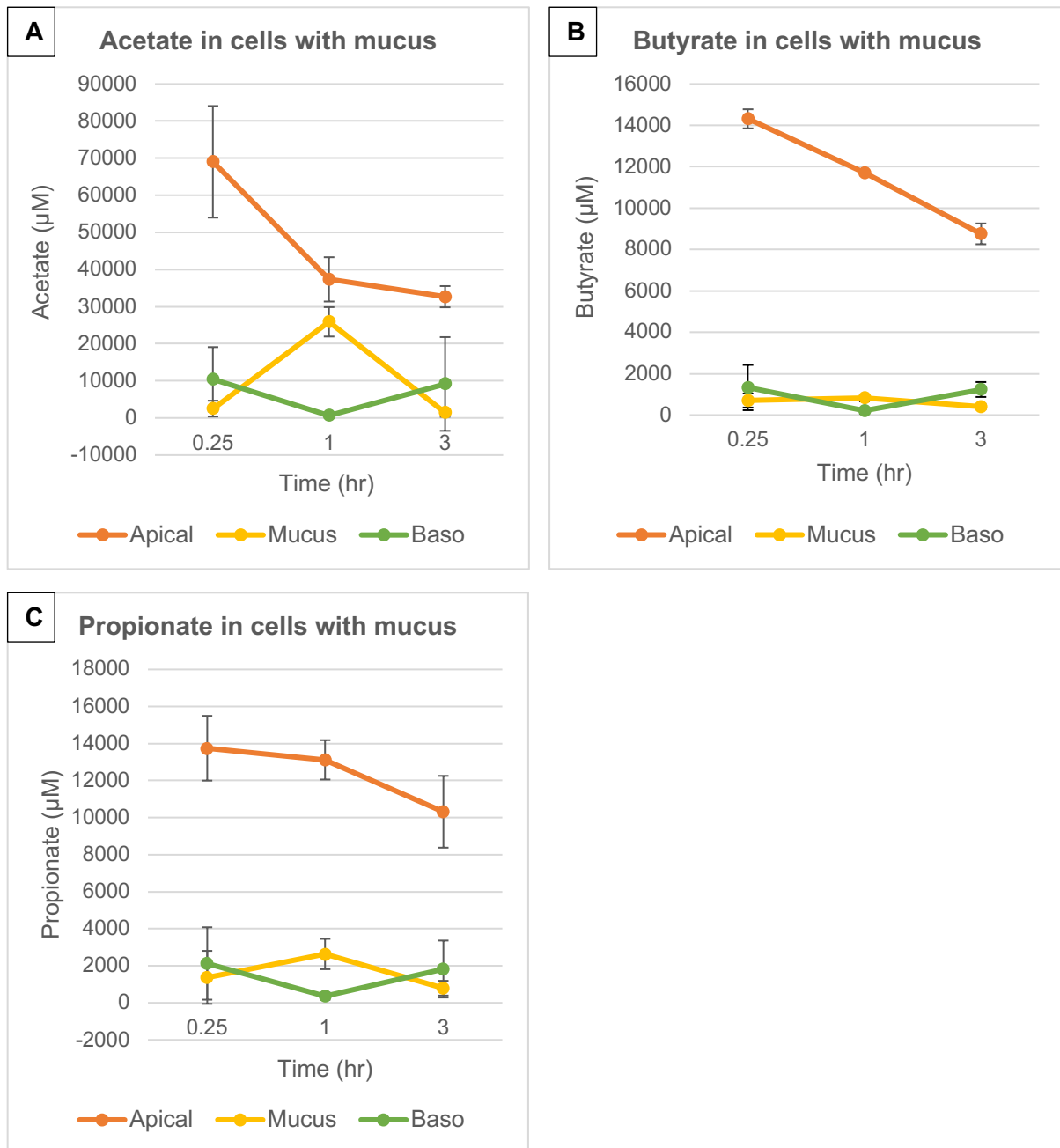


Figure 5.5. Line graphs comparing the average concentration (μM) of (A) acetate, (B) butyrate and (C) propionate in the apical (orange line), BL (green line) compartment and the mucus layer (yellow line) over time (hr). Error bars indicate $\pm$ standard deviation of three replicates of measurements.

5.3.3 Estimating SCFA uptake by the goblet cells

After successfully passing through the mucus layer, SCFA can either continue to diffuse passively into the BL compartment or be taken up by the cells via active transport. SCFA uptake was estimated by calculating the percentage remaining in the Transwells, which included concentration in the apical, BL and the mucus compartment out of the initial treatment concentration (**Figure 5.6A-C**). Butyrate and propionate were taken up rapidly by the cells during the first 15 min in Transwells without a mucus layer as the percentage remaining in the Transwells as a whole was 20% lower than in Transwells with mucus. From there onwards, percentage differences between the two mucus conditions became smaller and at 1 hr and 3 hr, the amount of butyrate remaining in the Transwells were the same under both conditions. These results show that without the obstruction of a mucus layer, SCFA were able to reach the cells and be consumed much faster. Based on our experiment, at the 3 hr mark the metabolic capacity of the cells for SCFA was the same in both experiments with mucus and without mucus. However, whether this trend continues beyond 3 hours remains to be determined.

The results for acetate in the Transwells with mucus were difficult to interpret (**Figure 5.6A**). Here, the percentage of acetate remaining in the system at 0.25 hr and 1 hr was above 100%, which was mostly contributed by high acetate concentration exceeding that of the treatment solution even after adjusting for concentration in the mucus and media blank samples. One possible explanation could be that adding SCFA to cells with mucus triggered a release of compounds with a mass spectral signal that overlaps with acetate. However, tracing the origin of this additional acetate signal was beyond the scope of our study.

From the figures below the consumption preference for the three SCFA by HT29-MTX-E12 cells was investigated. Under normal conditions where cells have an intact mucus layer, there was a fairly constant decrease in the rate of consumption of acetate and butyrate (**Figure 5.6A-B**). In contrast, propionate was metabolised much slower during the first 1 hr. In experiments where cells had their mucus layer stripped away, the amount of butyrate and propionate consumed barely changed during the first hour perhaps due to the quick influx of SCFA at the start of the treatment causing a temporary saturation of the cells' metabolic capacity. At the end of the experiment (3hr), there was approximately 90% of acetate, 60% of butyrate and 75% of propionate remaining in the Transwells showing that HT29-MTX-E12 cells have a preference for metabolising butyrate followed by propionate and acetate.

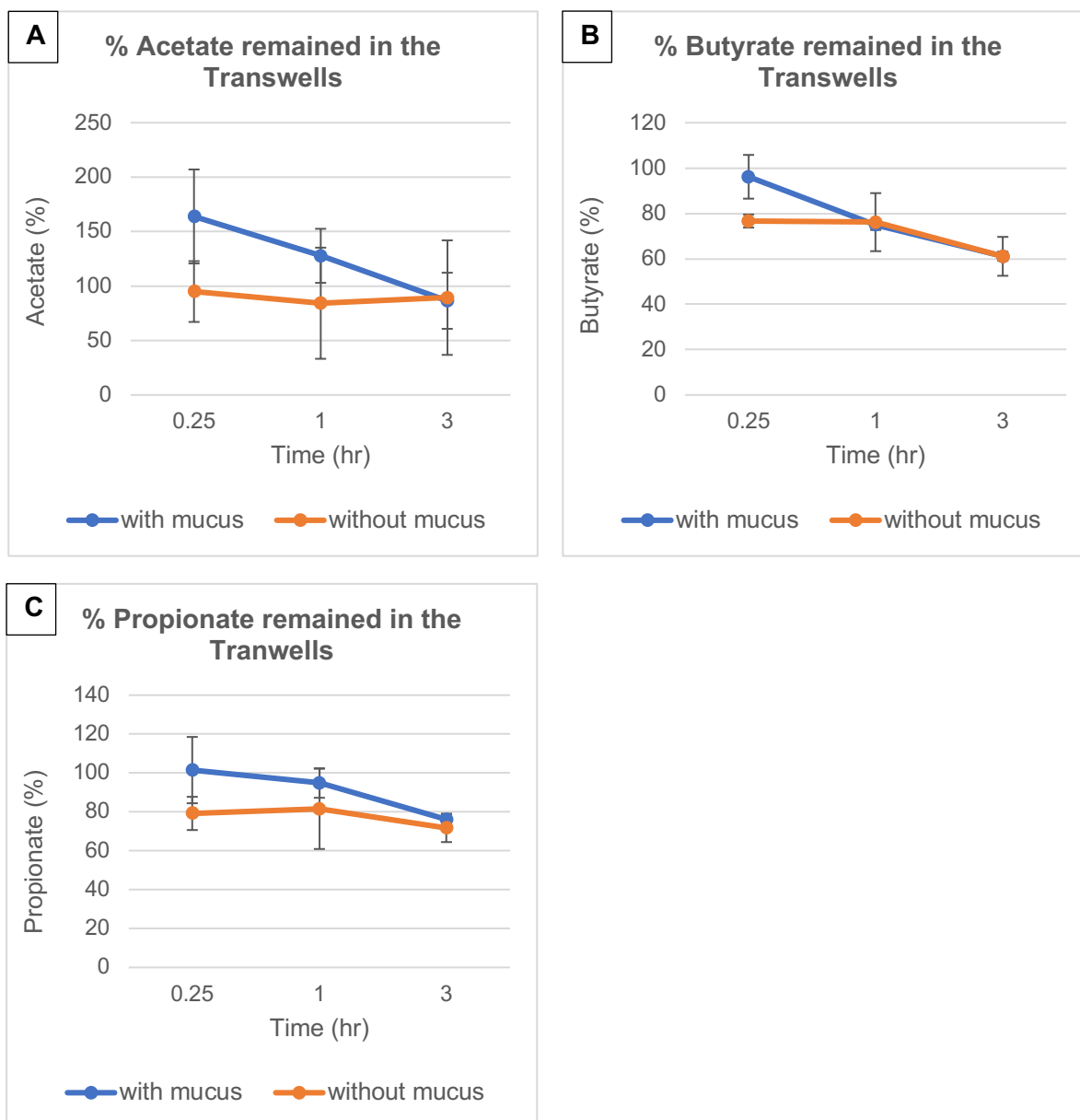
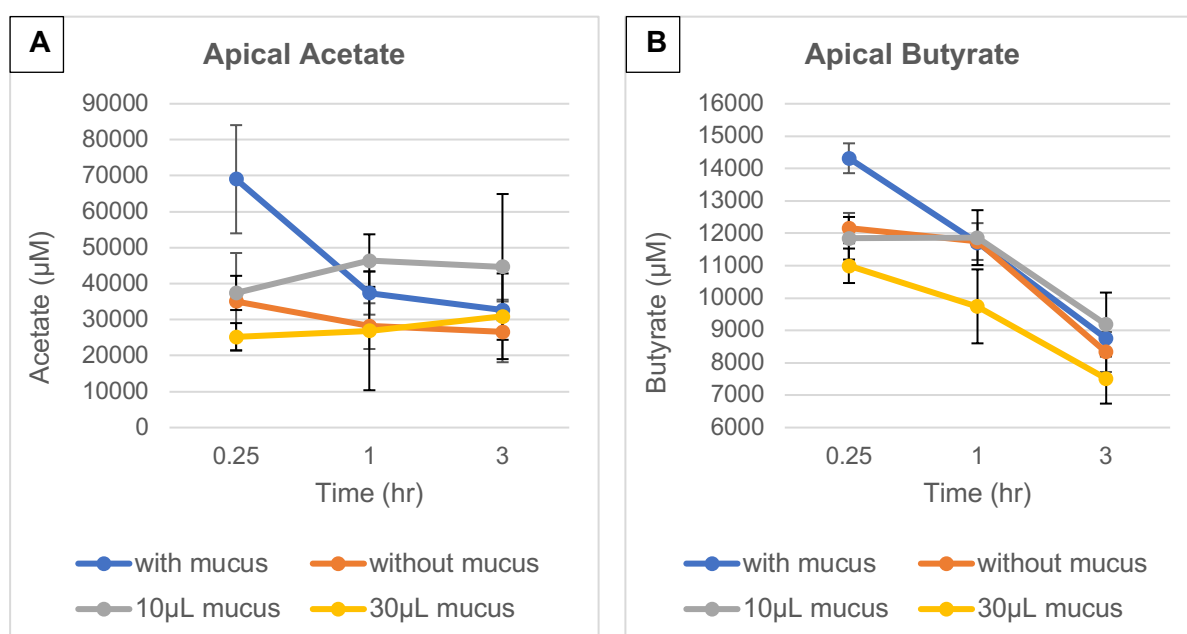


Figure 5.6. Line graphs comparing the average percentage (%) of (A) acetate, (B) butyrate and (C) propionate in the remained in Transwells post treatment. Percentage was calculated as the sum of concentrations in the apical, BL and mucus compartment experiments over the concentration of treatment solution multiplied by a hundred. Blue lines indicate experiments with mucus and orange lines indicate experiments without mucus. Error bars indicate +/- standard deviation of three replicates of measurements.

5.3.4 Diffusion through reconstituted porcine mucus

Experiments using reconstituted porcine mucus produced unexpected results. It was hypothesised that SCFA would diffuse faster from the apical compartment with a thinner mucus. However, results showed that (**Figure 5.7B-C**) butyrate and propionate diminished from the apical compartment more rapidly in a linear fashion in Transwells with 30 μ L of reconstituted mucus than in Transwells without a mucus layer. This may be due to the unexpected dilution effect made by the reconstituted mucus solution meaning that instead of reinforcing the original mucus layer it was perhaps not compatible and ended up disrupting the endogenous mucus structure thus causing more leakage. The reconstituted Sigma mucin was more viscous and more penetrable compared the native collected intestinal mucus where drug molecules diffused through more easily as studies have found.^{213,214}



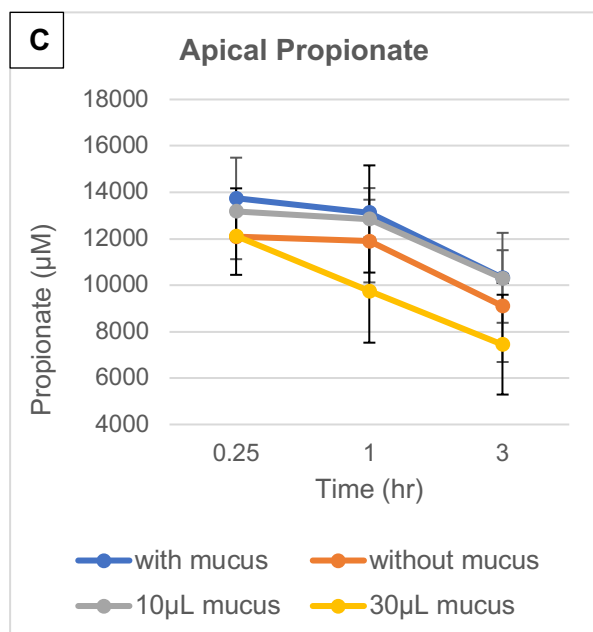


Figure 5.7. Line graphs showing the average concentration (µM) of (A) acetate, (B) butyrate and (C) propionate in the apical compartments over time (hr). Blue line indicates experiments with mucus, orange line indicates experiments without mucus, grey line indicates experiments with 10µL of reconstituted porcine mucus and yellow line indicates experiments with 30µL of porcine mucus. Error bars indicate +/- standard deviation of three replicates of measurements.

5.4 Discussion

In this proof-of-concept study a simple *in vitro* model was used to study the diffusion of SCFA across goblet cells under normal conditions with mucus and mucus-free conditions. Under normal conditions, SCFA were observed to diffuse rapidly after their initial addition, then after a short while, some were trapped inside the mucus thus temporarily halt their passage towards the mucosa. It was suggested that during this period (from 15min to 1hr), less SCFA were entering the cells from the apical compartment but rather from the BL compartment. However, currently all of the transporters identified on the BL membrane are for SCFA efflux out of the cells, there has not been a mechanism that describes the entry of SCFA into the cells basolaterally.⁵⁰ Currently, two main routes of entry have been described for SCFA uptake into the cells. One uses a Na^+/H^+ transporter or a HCO_3^- exchanger where SCFA exist in the ionised form and another one which allows them to diffuse directly through the cell membrane in their protonated form.⁵⁵ Therefore, it is possible that SCFA maybe entering the cells through the baso-membrane in its protonated form. Further evidence supporting this theory can be found from studies that measured the intracellular pH (pHi) of intestinal epithelial cells (IEC). One study showed that when SCFA was added to the apical and the basolateral borders of IEC, there was a sudden drop in intracellular pH (pHi) that gradually returned to baseline indicating that membranes on both sides were permeable to SCFA.²¹⁵ In a separate study, authors found that when SCFA was added to the apical side of HT29 monolayer, a parental cell line to the one used in this study, pHi hardly changed but when it was added to the BL side a marked decrease in pHi was observed suggesting that BL membrane was more permeable to SCFA than the apical membrane.²¹⁶ Based on preliminary result from this *in vitro* study and the literature a hypothesis for SCFA diffusion dynamics through the mucus layer is developed as illustrated in **Figure 5.8**.

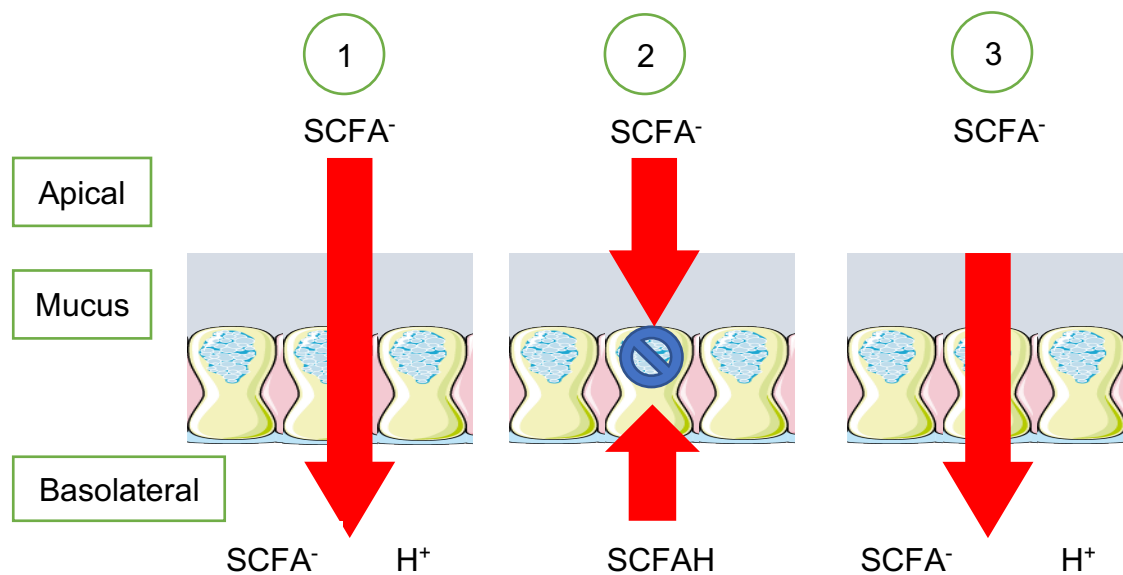


Figure 5.8. SCFA diffusion dynamics through the mucus layer.

- 1) At the start of the diffusion experiment, SCFA were flowing through the mucus with little obstruction.
- 2) As more SCFA entered the mucus, the mucus gel became saturated with SCFA molecules such that it restricted their movement inside the mucus gel resulting in a scenario where SCFA were being trapped. During this time SCFA entered the cell in their protonated form as SCFAH from the basolateral side.
- 3) However, it seemed like this trapping effect was only temporary, as SCFA started to flow into the BL compartment again.

Findings from this proof-of-concept study can be extrapolated to gain new perspective on the role of the mucus layer. Inside a normal healthy gut, the epithelium is lined with a continuous layer of mucus protecting it from pathogens and harsh chemicals in the lumen. It is well established that mucus thickness varies along the longitudinal compartments of the gut where it is relatively thin in the SI and adopt a double-layered structure in the colon.²⁴ Based on the results of this study, the absorption of SCFA into the epithelial cells is complicated by the presence of a mucus layer. Although it is unclear whether the uptake of SCFA into the IEC from the blood stream (the BL side) occurs under physiological conditions since SCFA are rapidly metabolised by the IEC leaving only a small amount enters the circulation.⁴¹

Understanding how mucus influences SCFA diffusion is important in studying conditions where the mucus layer has been disrupted such as in ulcerative colitis.²⁷ It is not known if a quick influx of SCFA into areas where the mucus structure has been compromised can have a detrimental effect on the host. By slowing down their descent to the epithelium, mucus may be exerting a buffering effect to a sudden influx of SCFA which may explain why there was

minimum change to pH_i when SCFA was added to the luminal membrane.²¹⁷ On the mucosal surface facing the lumen, there is a pH microclimate maintained by the secretion of HCO₃⁻.²¹⁸ Mucus is an integral part of this pH microclimate as one study showed that after washing the colonic mucosal tissue for 2 min in 20 mM of NAC, the same mucolytic agent used in this Chapter, there was a significant reduction in the thickness of the pH microclimate.²¹⁹ The effect of HCO₃⁻ excretion at the mucosal surface in response to SCFA might be doing more than just maintaining pH as study has shown that a fast increase in pH and Ca²⁺ removal in the presence of HCO₃⁻ promote the expansion of the tightly stratified inner mucus layer.²²⁰

5.5 Limitations and future work

The *in vitro* mucus model used in this work is relatively simple and inexpensive to set-up. However, it is acknowledged that there are several limitations that need to be addressed in future experiments.

To achieve maximum mucus coverage, only mucus producing cells were used but the human gut epithelium consists of a variety of cells types with enterocytes being the most abundant and goblet cells only constituting approximately 20%.²²¹ Therefore, a co-culture model with a mixture of enterocytes and goblet cells would be more appropriate at simulating the absorptive mechanism and SCFA usage of the colonic epithelium.²²² Although the cytotoxicity of the SCFA treatment solutions were studied, their impact on barrier integrity remained unknown, especially with the use of a mucolytic agent for removing the mucus layer. It is therefore advisable that for future experiments, transepithelial electrical resistance (TEER), a measure of barrier integrity should be recorded to ensure any change in TEER is taken into consideration when interpreting the results as a compromised barrier could lead to more SCFA flowing between the compartments.

To confirm the mechanism of SCFA entry into the cells, SCFA membrane transporters (MCT1) and intracellular pH should be measured to see if a decrease in pH_i and increase in MCT1 expression correlates with a decrease in SCFA in the BL compartment. At the end of the 3hr experiment, there were still more than 50% of SCFA left in the apical compartment. It would be interesting to extend the experiment to 24 hr with regular time intervals whenever possible to establish a temporal profile of SCFA diffusion through the mucus.

Using reconstituted porcine mucus was found to not be the best approach for increasing mucus thickness due to potential compatibility issue. It would be ideal to instead manipulate

the thickness of the endogenous mucus. One study found that by adding *N*-[(3,5-Difluorophenyl)acetyl]-L-alanyl-2-phenylglycine-1,1-dimethylethyl ester (DAPT, a Notch γ -secretase inhibitor that promotes goblet cell differentiation) in a semi-wet culture interface with mechanical shaking it was possible to increase mucus thickness.²²³

The colonic mucus layer is comprised of a loose outer layer and a tight inner layer. To the best of our knowledge, no Transwell system using cell line have been able to recreate this double-layered structure. However, recent developments in Gut-on-a-Chip technology have not only successfully created a double layered mucus but a system closely mimicking the physiology and chemical environment of the human gut.²²⁴ Another advantage of this system is the use of organoids from healthy volunteers which is better at recapitulating normal cellular behaviour than cancer cell lines. This is especially important where SCFA is used as they stimulate mucin release of healthy goblet cells but can have apoptotic effect on cancer cells therefore it is not always possible to infer the effect of SCFA on cancer line to healthy epithelial tissue.⁴⁵⁻⁴⁷

Chapter 6

General discussion

6.1 Summary of key results

Based on results from the two multi-omics studies described in Chapters 3 and 4, this thesis has identified unique metabolic and microbial signatures in the longitudinal and transverse compartments of the human GI tract implicated in irritable bowel syndrome and modified by the Western diet.

6.1.1 New learnings from exploring the SI microbiota

The SI microbiome has long been underexplored mainly due to the difficulties in accessing this part of the GI tract. The majority of studies exploring this region have relied on ileal effluent from patients undergoing ileostomy which may not be representative of a healthy individual.^{6,9,10,11} One study found that the process of ileostomy introduced oxygen to the distal section of the SI and as a result strict anaerobes were being replaced by facultative anaerobes.²²⁵ The study presented in Chapter 3 was the first that examined the microbiome and the metabolome of the upper SI of healthy volunteers undergoing a short-term Western diet intervention. This allowed the composition and metabolic activities of the SI microbiota in human to be studied and their relevance to host health.

In the duodenum, large amounts of amino acids and simple carbohydrates are present which is believed to be the primary source of nutrients for the SI microbiota. Compared to the colonic region, the microbial community of the SI was considerably less diverse than the colon as it was disproportionally dominated by *Streptococcus spp.* The extensive repertoire of *Streptococcus spp.* genes involved in the metabolism of amino acids, sugars and lipids grants them a biochemical advantage to thrive in the SI.¹² However, there seems to be variation in metabolic capacity between different *Streptococcus* species indicating a possible scenario of inter-species competition for nutrients in the SI. Competition for coexistence has been identified in *Streptococcus mutants* and *Streptococcus sanguinis* in the dental biofilm and therefore there is reason to suggest competition between *Streptococcus* species in the SI.²²⁶

Streptococcus and *Veillonella* were the most abundant genera in the duodenal aspirates which was in agreement with the ileostoma effluent study.²²⁷ Given that the SI is a primary location

for immune adaptation, these resident bacteria could play an important role in modulating the host's immune response. In fact, *Streptococcus* and *Veillonella* isolates from human SI microbiota have shown to stimulate cytokine production in human dendritic cells.²²⁸

Streptococcus.spp. is an incredibly diverse group of bacteria where some strains are capable of displaying virulent behaviour.^{168,169} Therefore, it is possible that, WD induced such behaviour in *Streptococcus.spp.* at the strain level and may even be responsible for the development of GI symptoms which was seen in all participants during the WD trial.²²⁹ *S. parasanguinis* was associated with an increased amount of choline in the duodenum. This coupled with the existence of choline-binding-proteins on the cell wall of their close relatives, raises interesting questions about the role of this species in modulating choline availability in the SI.¹⁶³ Choline is an important nutrient required for the synthesis of neurotransmitters and membrane components phosphatidylcholine and is therefore essential for maintaining brain function and cellular structure. Competition for choline in the SI can also have an impact on choline availability for the colonic microbiota. Microbial metabolism of choline in the colon produces trimethylamine (TMA) which is further oxidized in the liver to produce trimethylamine N-oxide (TMAO), a microbial-host co-metabolite implicated in atherosclerosis.²³⁰ Therefore, choline availability in the colon could modify the community structure and the flow of these potential harmful signals to the host. Collectively, these findings suggest that certain strains of *S. parasanguinis* may be involved in the pathogenesis of SIBO and can affect choline availability for the host and the colonic microbiome. Understanding how microbiota modulate nutrient flow to the host in the SI is important as variation in their activities could manipulate the nutritional environment of the gut and the molecules available to the host. Indeed, increased consumption by the microbiota could lead to nutrient deficiencies in the host.^{165,230}

6.1.2 The microbiome in the large intestines may be more diverse but it is not always a good sign

As we travel through the SI and entering the colon, intestinal transit slows down, the lumen becomes more anaerobic and a diverse community of microbes starts to establish. The colonic microbiota was the focus of the IBS study (Chapter 4). Patients were monitored for a period of 6 months where multiple omics data were collected. Stool microbiome was studied for their composition as well as their metabolic function using metagenomics sequencing and metabolomic profiling. Furthermore, advanced integration methods developed by one of the study collaborators were instrumental in identifying novel associations between bacterial gene regions and stool metabolites which eventually led to the discovery of a novel host-microbial co-metabolic pathway in IBS.

Distinction between the duodenum and the colonic microbiota could be found at all phylogenetic levels. The microbiota of the large intestines was higher in diversity and more evenly distributed at the genus level compared to the SI microbiome. The facultative anaerobes in the SI were being replaced by obligate anaerobes, which are complex carbohydrate-fermenting specialists in the genus of *Bacteroides*, *Blautia*, *Faecalibacterium*, *Collinsella* and *Ruminococcus*. SCFA were abundant in the colon as products of complex polysaccharides fermentation. As the Western diet study (Chapter 3) showed, these microbial products were positively correlated with fibre consumption. However, no difference in microbial abundance nor community structure were detected following the dietary switch, highlighting the resilience of the commensal microbiota and its adaptability to short-term changes in fibre and carbohydrate consumption. This could mean that a more persistent dietary intervention is needed to shift the microbiome of long term WD-consuming individuals to a healthier profile to prevent permanent loss of fibre-fermenting species which are associated with the development of obesity and type 2 diabetes.²³¹

Interestingly, irrespective of fibre intakes, patients with constipation-dominant IBS had significantly less SCFA in their stool which coincided with a reduction in butyrate-producing species. Lower butyrate in IBS-C patients could potentially impair their mucosal function as butyrate is an important fuel for the intestinal epithelial cells.^{42,43} Collectively contribute to the lower energy state. Furthermore, increased protein fermentation was evident in IBS-C patients based on the reduction of BCAA and increased abundance of BCAA-degrading species. Increased microbial diversity was also observed in IBS-C patients and possibly as a consequence of prolonged colonic transit based on previous *in vitro* and *in vivo* evidence.^{16,17} It was hypothesised that increased microbial diversity drove an overgrowth of proteolytic bacteria which may threaten the survival of saccharolytic commensals and thus creating an imbalance in the gut microbiome community. The causal relationship between microbiome and intestinal transit requires further investigation. Given the apparent metabolic and microbiome disturbances in IBS, it may sound logical to use prebiotics/probiotics treatments to promote favourable changes in the metabolic activity of the microbiome, however, results from clinical studies have shown inconsistencies in patient response, which seemed to be subtype specific and could potentially be associated with other unknown phenotypes.²³² Therefore, stratifying patients into responders and non-responders and characterising the clinical and molecular phenotypes in each of these groups is important as there is no one-size-fits-all solution when it comes to modulating the microbiome. In the IBS study (Chapter 4), several stool metabolites could differentiate patients according to their symptom severities which suggest the possibility of developing it into a diagnostic test to guide probiotic selection.

6.1.3 We must not forget about the mucus compartment

Metabolites in the gut lumen are absorbed through mucosa into the systemic circulation. Chapter 5 showed that the passage of bacterial metabolites was interrupted by the mucus layer. By studying the temporal change of SCFA concentrations in the apical, mucus and the basolateral compartments, a diffusion dynamic of SCFA through the mucus was hypothesised such that SCFA could potentially enter the cells through the basolateral membrane when their flow are restricted in the apical compartment. While the model lacked the physiological representation of the colon, it nevertheless shed new light on the role of the mucus layer and reiterated its importance in the crosstalk between the microbiota and the host.

The mucus compartment not only offer protection for the gut epithelium it is also a feeding ground for many gut commensals such as *Akkermansia*.³⁰ Sometimes mucus integrity could be compromised as a result of overfeeding by the microbiota in response to a shortage of fibre.²³³ As shown in Chapter 4, microbial species with mucin-degrading potential were elevated in IBS-C patients which may be driven by a lack of carbohydrates as a result of prolonged colonic transit. However, not all species are damaging to the mucus layer, some of them appear to have restorative properties as microbial-derived hypoxanthine was recently found to induce goblet cell mucin production and promoted mucosa-healing in a colitis mouse model.²³⁴ Therefore, results of the bacterial genes and hypoxanthine associations in Chapter 4 could be used to select hypoxanthine-producing species for designing new probiotic treatment for IBS.

6.2 Future work

While study limitations and suggestions for follow-up experiments were discussed in individual chapters, findings from this thesis as a whole have also opened up new avenues of research.

The interconnectedness of the SI and the colonic microbiota meant that more effort should be devoted to understanding the metabolic function of the SI microbiota in modulating nutrient availability as well as its role in IBS. The virulent potential of *Streptococcus* strains should be further explored given the overlap of symptom characteristics between SIBO and IBS. Since the link between the SI microbiome and IBS is not well established, initial studies should aim to generate hypotheses by integrating datasets from multiple omics platforms to build a network of system-level host-microbial interaction in the SI and the colon in the context of disease pathophysiology. Longitudinal sampling is recommended as it limits the intra-individual heterogeneity seen in cross-sectional studies. While metagenomics study allows for

the discovery of novel metabolic functions, in vitro and animal models are still needed to confirm the physiological significance of newly identified metabolic pathways.

The intestinal motility of constipated and diarrhoea patients may be modelled on a microfluidic gut-on-a-chip device. By adjusting the intensity of peristalsis and luminal flow, researchers could simulate the colonic transit of IBS-C and IBS-D patients to examine its effect on the mucus layer and the microbiome. The fact that each device can be made using patient's own cells highlights its potential for developing personalised diagnostics and therapeutics to treat IBS. Furthermore, the translational potential of gut-on-a-chip devices could mean that one day we will no longer have to rely on animal models to mimic the human physiology and as a research community we should all strive towards that goal.

6.3 Conclusions

This thesis has explored the role of the microbiota in the longitudinal compartments of the gut, from the duodenum to the colon, then travelling across the lumen through the mucus layer into the mucosa. Unique interactions between the host and the microbiota occur in each of these compartments and are associated with dietary changes and the development of functional GI disorders. The IBS study demonstrates the power of integrating multiple layers of omics datasets to explain physiological relevant associations and discover previously unknown disease pathways. The success of the project is a reflection of the tremendous collaborative effort made by a multidisciplinary team. The focus now should be establishing causality to understand the mechanisms behind these associations. In terms of the SI microbiome, we are still limited in our understanding of its composition and metabolic functions, particularly on its virulent potential and how it impacts nutrient availability for the host and the colonic microbiota. Once we understand how communities from different parts of the GI tract work together to influence host health, we will be better informed at finding ways to look after our gut microbiome and even modulate them to make them work for us.

References

1. Gordon JI. Honor thy gut symbionts redux. *Science*. 336(6086):1251-3. 1224686. Epub (2012).
2. Donaldson GP, Lee SM, Mazmanian SK. Gut biogeography of the bacterial microbiota. *Nat Rev Microbiol*. 14(1), (2016).
3. Pereira FC, Berry D. Microbial nutrient niches in the gut. *Environ Microbiol*. 19(4):1366–1378 (2017).
4. Hillman ET, Lu H, Yao T, Nakatsu CH. Microbial Ecology along the Gastrointestinal Tract. *Microbes Environ*, 32(4):300–313 (2017).
5. Kiela PR, Ghishan FK. Physiology of Intestinal Absorption and Secretion. *Best Pract Res Clin Gastroenterol*, 30(2):145-159 (2016).
6. Booijink CC, Zoetendal EG, Kleerebezem M, de Vos WM. Microbial communities in the human small intestine: coupling diversity to metagenomics. *Future Microbiol*, 2(3):285-95 (2007).
7. Geuking MB, Köller Y, Rupp S, McCoy KD. The interplay between the gut microbiota and the immune system. *Gut Microbes*, 5(3):411–418 (2014).
8. Gill SR, Pop M, Deboy RT, et al. Metagenomic analysis of the human distal gut microbiome. *Science*, 312(5778):1355–1359 (2016).
9. Booijink CCGM, El-Aidy S, Rajilić-Stojanović M, Heilig HGHJ, Troost FJ, Smidt H *et al*. High temporal and inter-individual variation detected in the human ileal microbiota. *Environ Microbiol* **12**: 3213–3227 (2010).
10. Fallingborg J. Intraluminal pH of the human gastrointestinal tract. *Dan Med Bull*, 46(3):183-96 (1999).
11. Zoetendal EG, Raes J, van den Bogert B, Arumugam M, Booijink CC, Troost FJ, Bork P, Wels M, de Vos WM, Kleerebezem M. The human small intestinal microbiota is driven by rapid uptake and conversion of simple carbohydrates. *ISME J*. 6(7):1415-26 Epub (2012).
12. Van den Bogert B, Boekhorst J, Herrmann R, Smid EJ, Zoetendal EG, Kleerebezem M. Comparative genomics analysis of *Streptococcus* isolates from the human small intestine reveals their adaptation to a highly dynamic ecosystem. *PLoS One*. 8(12):e83418 (2013).
13. El Kaoutari A, Armougom F, Gordon JI, Raoult D, Henrissat B. The abundance and variety of carbohydrate-active enzymes in the human gut microbiota. *Nat Rev Microbiol*. 11(7):497-504 (2013).
14. Lapébie P, Lombard V, Drula E, Terrapon N, Henrissat B. Bacteroidetes use thousands of enzyme combinations to break down glycans. *Nat Commun*. 10(1):2043 (2019).
15. Vandeputte D, Falony G, Vieira-Silva S, Tito RY, Joossens M, Raes J. Stool consistency is strongly associated with gut microbiota richness and composition, enterotypes and bacterial growth rates. *Gut*. 65(1):57-62 (2015).
16. Roager, H., Hansen, L., Bahl, M. *et al*. Colonic transit time is related to bacterial metabolism and mucosal turnover in the gut. *Nat Microbiol* **1**, 16093 (2016).
17. Tottey W, Faria-Gervasio D, Gaci N, et al. Colonic Transit Time Is a Driven Force of the Gut Microbiota Composition and Metabolism: In Vitro Evidence. *J Neurogastroenterol Motil*. 23(1):124-134 (2017).
18. Eckburg PB, Bik EM, Bernstein CN, et al. Diversity of the human intestinal microbial flora. *Science*. 308(5728):1635–1638 (2005).
19. Lavelle A, Lennon G, O'Sullivan O, Docherty N, Balfe A, Maguire A, Mulcahy HE, Doherty G, O'Donoghue D, Hyland J, Ross RP, Coffey JC, Sheahan K, Cotter PD, Shanahan F,

- Winter DC, O'Connell PR. Spatial variation of the colonic microbiota in patients with ulcerative colitis and control volunteers. *Gut*. (10):1553-61 (2015).
20. Rangel I, Sundin J, Fuentes S, Repsilber D, de Vos WM, Brummer RJ. The relationship between faecal-associated and mucosal-associated microbiota in irritable bowel syndrome patients and healthy subjects. *Aliment Pharmacol Ther*. 42(10):1211-21 (2015).
 21. Zoetendal EG, von Wright A, Vilpponen-Salmela T, Ben-Amor K, Akkermans AD, de Vos WM. Mucosa-associated bacteria in the human gastrointestinal tract are uniformly distributed along the colon and differ from the community recovered from feces. *Appl Environ Microbiol*. 68(7):3401–3407 (2002).
 22. Jones RB, Zhu X, Moan E, Murff HJ, Ness RM, Seidner DL, Sun S, Yu C, Dai Q, Fodor AA, Azcarate-Peril MA, Shrubsole MJ. Inter-niche and inter-individual variation in gut microbial community assessment using stool, rectal swab, and mucosal samples. *Sci Rep*. ;8(1):4139 (2018).
 23. Ahmed S, Macfarlane GT, Fite A, McBain AJ, Gilbert P, Macfarlane S. Mucosa-associated bacterial diversity in relation to human terminal ileum and colonic biopsy samples. *Appl Environ Microbiol*. 73(22):7435–7442 (2007).
 24. Johansson ME, Sjövall H, Hansson GC. The gastrointestinal mucus system in health and disease. *Nat Rev Gastroenterol Hepatol*. 10(6):352-361 (2013).
 25. Johansson ME, Phillipson M, Petersson J, Velcich A, Holm L, Hansson GC. The inner of the two Muc2 mucin-dependent mucus layers in colon is devoid of bacteria. *Proc Natl Acad Sci U S A*. 105(39):15064–15069 (2008).
 26. Johansson ME, Larsson JM, Hansson GC. The two mucus layers of colon are organized by the MUC2 mucin, whereas the outer layer is a legislator of host-microbial interactions. *Proc Natl Acad Sci U S A*. 108 (Suppl 1):4659–4665 (2011).
 27. Johansson ME, Gustafsson JK, Holmén-Larsson J, et al. Bacteria penetrate the normally impenetrable inner colon mucus layer in both murine colitis models and patients with ulcerative colitis. *Gut*. 63(2):281-291 (2014).
 28. Juge N. Microbial adhesins to gastrointestinal mucus. *Trends Microbiol*. 20(1):30-9 (2012).
 29. Tailford LE, Crost EH, Kavanaugh D, Juge N. Mucin glycan foraging in the human gut microbiome. *Front Genet*. 2015;6:81 (2015).
 30. van Passel MW, Kant R, Zoetendal EG, et al. The genome of *Akkermansia muciniphila*, a dedicated intestinal mucin degrader, and its use in exploring intestinal metagenomes. *PLoS One*. 6(3):e16876 (2011).
 31. Sonnenburg JL, Xu J, Leip DD, Chen CH, Westover BP, Weatherford J, Buhler JD, Gordon JI. Glycan foraging in vivo by an intestine-adapted bacterial symbiont. *Science*. 307(5717):1955-9 (2005).
 32. Espey MG Role of oxygen gradients in shaping redox relationships between the human intestine and its microbiota. *Free Radic Biol Med*. 55:130-40 (2013).
 33. Albenberg L, Esipova TV, Judge CP, Bittinger K, Chen J, Laughlin A, Grunberg S, Baldassano RN, Lewis JD, Li H, Thom SR, Bushman FD, Vinogradov SA, Wu GD. Correlation between intraluminal oxygen gradient and radial partitioning of intestinal microbiota. *Gastroenterology*. 147(5):1055-63.e8 (2014).
 34. Pédrón T, Mulet C, Dauga C, Frangeul L, Chervaux C, Grompone G, Sansonetti PJ. A crypt-specific core microbiota resides in the mouse colon. *mBio*. 22;3(3):e00116-12 (2012).
 35. Baughn AD, Malmay MH. The strict anaerobe *Bacteroides fragilis* grows in and benefits from nanomolar concentrations of oxygen. *Nature*. 427(6973):441-4 (2004).

36. Tremaroli, V., Bäckhed, F. Functional interactions between the gut microbiota and host metabolism. *Nature* 489, 242–249 (2012).
37. Cummings JH, Macfarlane GT. The control and consequences of bacterial fermentation in the human colon. *J Appl Bacteriol.* 70(6):443-59 (1991).
38. Macfarlane S, Macfarlane GT. Regulation of short-chain fatty acid production. *Proc Nutr Soc.* 62(1):67-72 (2003).
39. Macfarlane GT, Macfarlane S. Bacteria, colonic fermentation, and gastrointestinal health. *J AOAC Int.* 95(1):50-60 (2012).
40. Louis P, Flint HJ. Formation of propionate and butyrate by the human colonic microbiota. *Environ Microbiol.* 19(1):29-41 (2016).
41. Cummings JH, Pomare EW, Branch WJ, Naylor CP, Macfarlane GT. Short chain fatty acids in human large intestine, portal, hepatic and venous blood. *Gut.* (10):1221-7 (1987).
42. Roediger WE. The colonic epithelium in ulcerative colitis: an energy-deficiency disease? *Lancet.* 2(8197):712-5 (1980).
43. Roediger WE. Short chain fatty acids as metabolic regulators of ion absorption in the colon. *Acta Vet Scand Suppl.* 86:116-25 (1989).
44. Chang PV, Hao L, Offermanns S, Medzhitov R. The microbial metabolite butyrate regulates intestinal macrophage function via histone deacetylase inhibition. *Proc Natl Acad Sci USA.* 111:2247–52 (2014).
45. Shao Y, Gao Z, Marks PA, Jiang X. Apoptotic and autophagic cell death induced by histone deacetylase inhibitors. *Proc Natl Acad Sci U S A.* 101(52):18030-5 (2004).
46. Zheng L, Kelly CJ, Battista KD, et al. Microbial-Derived Butyrate Promotes Epithelial Barrier Function through IL-10 Receptor-Dependent Repression of Claudin-2. *J Immunol.* 199(8):2976-2984 (2017).
47. Wang RX, Lee JS, Campbell EL, Colgan SP. Microbiota-derived butyrate dynamically regulates intestinal homeostasis through regulation of actin-associated protein synaptopodin. *Proc Natl Acad Sci U S A.* 117(21):11648-11657 (2020).
48. Parada Venegas D, De la Fuente MK, Landskron G, González MJ, Quera R, Dijkstra G, Harmsen HJM, Faber KN, Hermoso MA. Short Chain Fatty Acids (SCFAs)-Mediated Gut Epithelial and Immune Regulation and Its Relevance for Inflammatory Bowel Diseases. *Front Immunol.* 10:1486 (2019).
49. Byrne CS, Chambers ES, Morrison DJ, Frost G. The role of short chain fatty acids in appetite regulation and energy homeostasis. *Int J Obes (Lond).* 39(9):1331-1338 (2015).
50. Sivaprakasam S, Bhutia YD, Yang S, Ganapathy V. Short-Chain Fatty Acid Transporters: Role in Colonic Homeostasis. *Compr Physiol.* 8(1):299-314 (2017).
51. Al-Mosauwi H, Ryan E, McGrane A, et al. Differential protein abundance of a basolateral MCT1 transporter in the human gastrointestinal tract. *Cell Biol Int.* 40(12):1303-1312 (2016).
52. Cummings JH. Colonic absorption: the importance of short chain fatty acids in man. *Scand J Gastroenterol Suppl.* 93:89-99 (1984).
53. Reynolds DA, Rajendran VM, Binder HJ. Bicarbonate-stimulated [¹⁴C]butyrate uptake in basolateral membrane vesicles of rat distal colon. *Gastroenterology.* 105(3):725-732 (1993).
54. Chu S, Montrose MH. Non-ionic diffusion and carrier-mediated transport drive extracellular pH regulation of mouse colonic crypts. *J Physiol.* 494 (Pt 3)(Pt 3):783-793 (1996).
55. Ruppin H, Bar-Meir S, Soergel KH, Wood CM, Schmitt MG Jr. Absorption of short-chain fatty acids by the colon. *Gastroenterology.* 78(6):1500-7 (1980).
56. Monte MJ, Marin JJ, Antelo A, Vazquez-Tato J. Bile acids: chemistry, physiology, and pathophysiology. *World J Gastroenterol.* 15(7):804-816 (2009).

57. Ridlon JM, Kang DJ, Hylemon PB. Bile salt biotransformations by human intestinal bacteria. *J Lipid Res.* 47(2):241-59 (2006).
58. Sinal CJ, Tohkin M, Miyata M, Ward JM, Lambert G, Gonzalez FJ. Targeted disruption of the nuclear receptor FXR/BAR impairs bile acid and lipid homeostasis. *Cell.* 102(6):731-44 (2000).
59. Sayin SI, Wahlström A, Felin J, Jäntti S, Marschall HU, Bamberg K, Angelin B, Hyötyläinen T, Orešič M, Bäckhed F. Gut microbiota regulates bile acid metabolism by reducing the levels of tauro-beta-muricholic acid, a naturally occurring FXR antagonist. *Cell Metab.* 17(2):225-35 (2013).
60. Li, F., Jiang, C., Krausz, K. *et al.* Microbiome remodelling leads to inhibition of intestinal farnesoid X receptor signalling and decreased obesity. *Nat Commun* 4, 2384 (2013).
61. Begley M, Gahan CG, Hill C. The interaction between bacteria and bile. *FEMS Microbiol Rev.* 29(4):625-51 (2005).
62. Zhao DL, Hirst BH. Comparison of bile salt perturbation of duodenal and jejunal isolated brush-border membranes. *Digestion.* 47(4):200-7 (1990).
63. Tanaka H, Doesburg K, Iwasaki T, Mierau I. Screening of lactic acid bacteria for bile salt hydrolase activity. *J Dairy Sci.* 82(12):2530-5 (1999).
64. Wilson KH. Efficiency of various bile salt preparations for stimulation of *Clostridium difficile* spore germination. *J Clin Microbiol.* 18(4):1017-9 (1983).
65. Sorg JA, Sonenshein AL. Inhibiting the initiation of *Clostridium difficile* spore germination using analogs of chenodeoxycholic acid, a bile acid. *J Bacteriol.* 192(19):4983-90 (2010).
66. Yao CK, Muir JG, Gibson PR. Review article: insights into colonic protein fermentation, its modulation and potential health implications. *Aliment Pharmacol Ther.* 43(2):181-96 (2016).
67. Zhu Y, Lin X, Zhao F, Shi X, Li H, Li Y, Zhu W, Xu X, Li C, Zhou G. Meat, dairy and plant proteins alter bacterial composition of rat gut bacteria. *Sci Rep.* 5:15220 (2015).
68. Govers MJ, Gannon NJ, Dunshea FR, Gibson PR, Muir JG. Wheat bran affects the site of fermentation of resistant starch and luminal indexes related to colon cancer risk: a study in pigs. *Gut.* 45(6):840-7 (1999).
69. Portune KJ, Beaumont M, Davila A-M, Tomé D, Blachier F, Sanz Y. Gut microbiota role in dietary protein metabolism and health-related outcomes: the two sides of the coin. *Trends Food Sci Technol.* 57:213–32 (2016).
70. Stoll B, Henry J, Reeds PJ, Yu H, Jahoor F, Burrin DG. Catabolism dominates the first-pass intestinal metabolism of dietary essential amino acids in milk protein-fed piglets. *J Nutr.* 128(3):606-14 (1998).
71. Bergen WG, Wu G. Intestinal nitrogen recycling and utilization in health and disease. *J Nutr.* 139(5):821-5 (2009).
72. Davila AM, Blachier F, Gotteland M, Andriamihaja M, Benetti PH, Sanz Y, Tomé D. Intestinal luminal nitrogen metabolism: role of the gut microbiota and consequences for the host. *Pharmacol Res.* 68(1):95-107 (2013).
73. Metges CC. Contribution of microbial amino acids to amino acid homeostasis of the host. *J Nutr.* 130(7):1857S-64S (2000).
74. Dai, Z., Li, X., Xi, P. *et al.* Metabolism of select amino acids in bacteria from the pig small intestine. *Amino Acids* 42, 1597–1608 (2012).
75. Dai, Z., Li, X., Xi, P. *et al.* Regulatory role for L-arginine in the utilization of amino acids by pig small-intestinal bacteria. *Amino Acids* 43, 233–244 (2012).
76. Lin R, Liu W, Piao M, Zhu H. A review of the relationship between the gut microbiota and amino acid metabolism. *Amino Acids.* 49(12):2083-2090 (2017).
77. Morales P, Fujio S, Navarrete P, Ugalde JA, Magne F, Carrasco-Pozo C, Tralma K, Quezada M, Hurtado C, Covarrubias N, Brignardello J, Henriquez D, Gotteland M.

- Impact of Dietary Lipids on Colonic Function and Microbiota: An Experimental Approach Involving Orlistat-Induced Fat Malabsorption in Human Volunteers. *Clin Transl Gastroenterol.* 7(4):e161 (2016).
78. Martinez-Guryñ K, Hubert N, Frazier K, et al. Small Intestine Microbiota Regulate Host Digestive and Absorptive Adaptive Responses to Dietary Lipids. *Cell Host Microbe.* 23(4):458-469 (2018).
 79. Fresia Fernandez, Matthew D. Collins, Vitamin K composition of anaerobic gut bacteria, *FEMS Microbiology Letters.* 41(2):175–180 (1987).
 80. Welkos SL, Toskes PP, Baer H. Importance of anaerobic bacteria in the cobalamin malabsorption of the experimental rat blind loop syndrome. *Gastroenterology.* 80(2):313-20 (1981).
 81. Oliphant K, Allen-Vercoe E. Macronutrient metabolism by the human gut microbiome: major fermentation by-products and their impact on host health. *Microbiome.* 7(1):91 (2019).
 82. Valdes Ana M, Walter Jens, Segal Eran, Spector Tim D. Role of the gut microbiota in nutrition and health *BMJ.* 361:k2179 (2018).
 83. Zmora, N., Suez, J. & Elinav, E. You are what you eat: diet, health and the gut microbiota. *Nat Rev Gastroenterol Hepatol* 16, 35–56 (2019).
 84. Sonnenburg, E., Smits, S., Tikhonov, M. *et al.* Diet-induced extinctions in the gut microbiota compound over generations. *Nature* 529, 212–215 (2016).
 85. Kopp W. How Western Diet And Lifestyle Drive The Pandemic Of Obesity And Civilization Diseases. *Diabetes Metab Syndr Obes.* 12:2221-2236 (2019).
 86. Vangay P, Johnson AJ, Ward TL, et al. US Immigration Westernizes the Human Gut Microbiome. *Cell.* 175(4):962-972.e10 (2018).
 87. Turnbaugh PJ, Bäckhed F, Fulton L, Gordon JL. Diet-induced obesity is linked to marked but reversible alterations in the mouse distal gut microbiome. *Cell Host Microbe.* 3(4):213-223 (2008).
 88. Chassaing B, Koren O, Goodrich JK, Poole AC, Srinivasan S, Ley RE, Gewirtz AT. Dietary emulsifiers impact the mouse gut microbiota promoting colitis and metabolic syndrome. *Nature.* 2015 Mar 5;519(7541):92-6. doi: 10.1038/nature14232. Epub 2015 Feb 25. Erratum in: *Nature.* 536(7615):238 (2016).
 89. Suez J, Korem T, Zeevi D, Zilberman-Schapira G, Thaiss CA, Maza O, Israeli D, Zmora N, Gilad S, Weinberger A, Kuperman Y, Harmelin A, Kolodkin-Gal I, Shapiro H, Halpern Z, Segal E, Elinav E. Artificial sweeteners induce glucose intolerance by altering the gut microbiota. *Nature.* 514(7521):181-6 (2014).
 90. Schroeder BO, Birchenough GMH, Ståhlman M, Arike L, Johansson MEV, Hansson GC, Bäckhed F. Bifidobacteria or Fiber Protects against Diet-Induced Microbiota-Mediated Colonic Mucus Deterioration. *Cell Host Microbe.* 10;23(1):27-40.e7 (2018).
 91. Diether NE, Willing BP. Microbial Fermentation of Dietary Protein: An Important Factor in Diet-Microbe-Host Interaction. *Microorganisms.* 7(1):19 (2019).
 92. David LA, Maurice CF, Carmody RN, et al. Diet rapidly and reproducibly alters the human gut microbiome. *Nature.* 505(7484):559-563 (2014).
 93. guirre M, Eck A, Koenen ME, Savelkoul PH, Budding AE, Venema K. Diet drives quick changes in the metabolic activity and composition of human gut microbiota in a validated in vitro gut model. *Res Microbiol.* 167(2):114-25 (2016).
 94. Smith EA, Macfarlane GT. Dissimilatory amino Acid metabolism in human colonic bacteria. *Anaerobe.* 3(5):327-37 (1997).
 95. Jaskiewicz J, Zhao Y, Hawes JW, Shimomura Y, Crabb DW, Harris RA. Catabolism of isobutyrate by colonocytes. *Arch Biochem Biophys.* 327(2):265-70 (1996).

96. Smith E., Macfarlane G. Enumeration of amino acid fermenting bacteria in the human large intestine: Effects of pH and starch on peptide metabolism and dissimilation of amino acids. *FEMS Microbiol. Ecol.* 25:355–368 (1998).
97. Cremin JD Jr, Fitch MD, Fleming SE. Glucose alleviates ammonia-induced inhibition of short-chain fatty acid metabolism in rat colonic epithelial cells. *Am J Physiol Gastrointest Liver Physiol.* 285(1):G105-14 (2003).
98. Salonen A., Lahti L., Salojärvi J., Holtrop G., Korpela K., Duncan S.H., Date P., Farquharson F., Johnstone A.M., Lobley G.E., et al. Impact of diet and individual variation on intestinal microbiota composition and fermentation products in obese men. *ISME J.* 8:2218–2230 (2014).
99. O'Keefe SJ, Li JV, Lahti L, Ou J, Carbonero F, Mohammed K, Posma JM, Kinross J, Wahl E, Ruder E, Vippera K, Naidoo V, Mtshali L, Tims S, Puylaert PG, DeLany J, Krasinskas A, Benefiel AC, Kaseb HO, Newton K, Nicholson JK, de Vos WM, Gaskins HR, Zoetendal EG. Fat, fibre and cancer risk in African Americans and rural Africans. *Nat Commun.* 6:6342 (2015).
100. Canavan C, West J, Card T. The epidemiology of irritable bowel syndrome. *Clin Epidemiol.* 6: 71–80, (2014).
101. Schmulson MJ, Drossman DA. What Is New in Rome IV. *J Neurogastroenterol Motil.* 23(2):151–163 (2017).
102. Bhattarai Y, Muniz Pedrogo DA, Kashyap PC. Irritable bowel syndrome: a gut microbiota-related disorder?. *Am J Physiol Gastrointest Liver Physiol.* 312(1):G52–G62 (2017).
103. Rao, M., Gershon, M. The bowel and beyond: the enteric nervous system in neurological disorders. *Nat Rev Gastroenterol Hepatol* 13, 517–528 (2016).
104. De Palma, G., Lynch, M.D., Lu, J., Dang, V.T., Deng, Y., Jury, J., Umeh, G., Miranda, P.M., Pigrau Pastor, M., Sidani, S., et al. Transplantation of fecal microbiota from patients with irritable bowel syndrome alters gut function and behavior in recipient mice. *Science Translational Medicine.* 9(379):eaaf6397 (2017).
105. Soret R, Chevalier J, De Coppet P, Poupeau G, Derkinderen P, Segain JP, Neunlist M. Short-chain fatty acids regulate the enteric neurons and control gastrointestinal motility in rats. *Gastroenterology.* 138(5):1772-82 (2010).
106. Gershon MD. 5-Hydroxytryptamine (serotonin) in the gastrointestinal tract. *Curr Opin Endocrinol Diabetes Obes.* (1):14-21 (2013).
107. Bhattarai Y, Williams BB, Battaglioli EJ, et al. Gut Microbiota-Produced Tryptamine Activates an Epithelial G-Protein-Coupled Receptor to Increase Colonic Secretion. *Cell Host Microbe.* 23(6):775-785.e5 (2018).
108. Mekhjian, H.S., Phillips, S.F. & Hofmann, A.F. Colonic absorption of unconjugated bile acids. *Digest Dis Sci* 24, 545–550 (1979).
109. Duboc H, Rainteau D, Rajca S, Humbert L, Farabos D, Maubert M, Grondin V, Jouet P, Bouhassira D, Seksik P, Sokol H, Coffin B, Sabaté JM. Increase in fecal primary bile acids and dysbiosis in patients with diarrhea-predominant irritable bowel syndrome. *Neurogastroenterol Motil.* 24(6):513-20, e246-7 (2012).
110. Plöger S, Stumpff F, Penner GB, Schulzke JD, Gäbel G, Martens H, Shen Z, Günzel D, Aschenbach JR. Microbial butyrate and its role for barrier function in the gastrointestinal tract. *Ann N Y Acad Sci.* 1258:52-9 (2012).
111. Pozuelo M, Panda S, Santiago A, Mendez S, Accarino A, Santos J, Guarner F, Azpiroz F, Manichanh C. Reduction of butyrate- and methane-producing microorganisms in patients with irritable bowel syndrome. *Sci Rep* 5: 12693, (2015).
112. Lyra A, Rinttilä T, Nikkilä J, Krogius-Kurikka L, Kajander K, Malinen E, Mättö J, Mäkelä L, Palva A. Diarrhoea-predominant irritable bowel syndrome distinguishable by 16S rRNA gene phylotype quantification. *World J Gastroenterol.* 15(47):5936-45 (2009).

113. Ghoshal UC, Srivastava D. Irritable bowel syndrome and small intestinal bacterial overgrowth: meaningful association or unnecessary hype. *World J Gastroenterol*. 20(10):2482-91 (2014).
114. Yang M, Zhang L, Hong G, Li Y, Li G, Qian W, Xiong H, Bai T, Song J, Hou X. Duodenal and rectal mucosal microbiota related to small intestinal bacterial overgrowth in diarrhea-predominant irritable bowel syndrome. *J Gastroenterol Hepatol*. 35(5):795-805 (2020).
115. Zhang, X., Li, L., Butcher, J. *et al*. Advancing functional and translational microbiome research using meta-omics approaches. *Microbiome* 7, 154 (2019).
116. J. K. Nicholson, J. C. Lindon & E. Holmes. 'Metabonomics': understanding the metabolic responses of living systems to pathophysiological stimuli via multivariate statistical analysis of biological NMR spectroscopic data, *Xenobiotica*, 1999. 29:11, 1181-1189 (1999).
117. Nicholson JK, Holmes E, Kinross J, Burcelin R, Gibson G, Jia W, Pettersson S. Host-gut microbiota metabolic interactions. *Science*. 336(6086):1262-7 (2012).
118. Nicholson JK, Buckingham MJ, Sadler PJ. High resolution ¹H n.m.r. studies of vertebrate blood and plasma. *Biochem J*. 211(3):605-615 (1983).
119. Anthony C. Dona, Beatriz Jiménez, Hartmut Schäfer, Eberhard Humpfer, Manfred Spraul, Matthew R. Lewis, Jake T. M. Pearce, Elaine Holmes, John C. Lindon, and Jeremy K. Nicholson. Precision High-Throughput Proton NMR Spectroscopy of Human Urine, Serum, and Plasma for Large-Scale Metabolic Phenotyping. *Analytical Chemistry*. 86 (19): 9887-9894 (2014).
120. J. Keeler. *Understanding NMR Spectroscopy*. 2002.
121. Dieterle F, Ross A, Schlotterbeck G, Senn H. Probabilistic quotient normalization as robust method to account for dilution of complex biological mixtures. Application in ¹H NMR metabonomics. *Anal Chem*. 78(13):4281-90 (2006).
122. Pearson, K. On lines and planes of closest fit to systems of points in space. *The London, Edinburgh and Dublin Philosophical Magazine and Journal of Science*. 2, 559-572 (1901).
123. Johnstone IM, Titterton DM. Statistical challenges of high-dimensional data. *Philos Trans A Math Phys Eng Sci*. 367(1906):4237-53 (2009).
124. Trygg, J. & Wold, S. Orthogonal projections to latent structures (OPLS). *J. Chemom*. 16, 119-128 (2002).
125. Joram M. Posma, Isabel Garcia-Perez, Timothy M. D. Ebbels, John C. Lindon, Jeremiah Stamler, Paul Elliott, Elaine Holmes, and Jeremy K. Nicholson. Optimized Phenotypic Biomarker Discovery and Confounder Elimination via Covariate-Adjusted Projection to Latent Structures from Metabolic Spectroscopy Data. *Journal of Proteome Research*. 17 (4): 1586-1595 (2018).
126. Hochberg, Y.; Benjamini, Y. More powerful procedures for multiple significance testing. *Statistics in Medicine*. 9 (7): 811–818 (1990).
127. Olivier Cloarec, Marc-Emmanuel Dumas, Andrew Craig, Richard H. Barton, Johan Trygg, Jane Hudson, Christine Blancher, Dominique Gauguier, John C. Lindon, Elaine Holmes and Jeremy Nicholson. Statistical Total Correlation Spectroscopy: An Exploratory Approach for Latent Biomarker Identification from Metabolic ¹H NMR Data Sets. *Analytical Chemistry*. 77 (5): 1282-1289 (2005).
128. Pitt JJ. Principles and applications of liquid chromatography-mass spectrometry in clinical biochemistry. *Clin Biochem Rev*. 30(1):19-34 (2009).
129. Chernushevich IV, Loboda AV, Thomson BA. An introduction to quadrupole-time-of-flight mass spectrometry. *J Mass Spectrom*. 36(8):849-65 (2001).
130. Magali H. Sarafian, Matthew R. Lewis, Alexandros Pechlivanis, Simon Ralphs, Mark J. W. McPhail, Vishal C. Patel, Marc-Emmanuel Dumas, Elaine Holmes, and Jeremy K. Nicholson. Bile Acid Profiling and Quantification in Biofluids Using Ultra-Performance

- Liquid Chromatography Tandem Mass Spectrometry. *Analytical Chemistry*. 87 (19): 9662-9670 (2015).
131. Morrison, J. D., Personal reminiscences of forty years of mass spectrometry in Australia. *Organic Mass Spectrometry*. 26 (4): 183 (1991).
 132. Hites, Ronald A. Development of Gas Chromatographic Mass Spectrometry. *Analytical Chemistry*. 88 (14): 6955–6961 (2016).
 133. García-Villalba R, Giménez-Bastida JA, García-Conesa MT, Tomás-Barberán FA, Carlos Espín J, Larrosa M. Alternative method for gas chromatography-mass spectrometry analysis of short-chain fatty acids in faecal samples. *J Sep Sci*. 35(15):1906-13 (2012).
 134. Kuczynski, J., Lauber, C., Walters, W. *et al.* Experimental and analytical tools for studying the human microbiome. *Nat Rev Genet*. 13, 47–58 (2012).
 135. Ho, N.T., Li, F., Wang, S. *et al.* metagenomeR: an R package for analysis of microbiome relative abundance data using zero-inflated beta GAMLSS and meta-analysis across studies using random effects models. *BMC Bioinformatics* 20, 188 (2019).
 136. McMurdie PJ, Holmes S. phyloseq: An R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS ONE*. 8(4), e61217 (2013).
 137. Jari Oksanen, F. Guillaume Blanchet, Michael Friendly, Roeland Kindt, Pierre Legendre, Dan McGlinn, Peter R. Minchin, R. B. O'Hara, Gavin L. Simpson, Peter Solymos, M. Henry H. Stevens, Eduard Szoecs, Helene Wagner. 'vegan': community ecology package. (2019).
 138. Andrew J. Steinberger. asteinberger9/seq_scripts: Release v1 (Version v1). Zenodo (2018).
 139. Shannon, C. E. A mathematical theory of communication. *Bell Syst. Tech. J.* 27, 379–423 (1948).
 140. Chao A. Non-parametric estimation of the number of classes in a population. *Scand J Stat*. 11:265–270 (1984).
 141. Bray JR, Curtis JT. An ordination of the upland forest communities of southern Wisconsin. *Ecol Monogr* 27(4):325–349 (1957).
 142. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Research*. 13(11):2498-504 (2003).
 143. Camilleri M, Murphy R, Chadwick VS. Dose-related effects of chenodeoxycholic acid in the rabbit colon. *Dig Dis Sci*. 25(6):433-8 (1980).
 144. Rao AS, Wong BS, Camilleri M, Odunsi-Shiyanbade ST, McKinzie S, Ryks M, Burton D, Carlson P, Lamsam J, Singh R, Zinsmeister AR. Chenodeoxycholate in females with irritable bowel syndrome-constipation: a pharmacodynamic and pharmacogenetic analysis. *Gastroenterology*. 139(5):1549-58, 1558.e1 (2010).
 145. Peleman C, Camilleri M, Busciglio I, Burton D, Donato L, Zinsmeister AR. Colonic Transit and Bile Acid Synthesis or Excretion in Patients With Irritable Bowel Syndrome-Diarrhea Without Bile Acid Malabsorption. *Clin Gastroenterol Hepatol*. 15(5):720-727.e1 (2017).
 146. Slattery SA, Niaz O, Aziz Q, Ford AC, Farmer AD. Systematic review with meta-analysis: the prevalence of bile acid malabsorption in the irritable bowel syndrome with diarrhoea. *Aliment Pharmacol Ther*. 42(1):3-11 (2015).
 147. Loren Cordain, S Boyd Eaton, Anthony Sebastian, Neil Mann, Staffan Lindeberg, Bruce A Watkins, James H O'Keefe, Janette Brand-Miller, Origins and evolution of the Western diet: health implications for the 21st century, *The American Journal of Clinical Nutrition*. 81, (2)341–354 (2005).

148. Zinöcker MK, Lindseth IA. The Western Diet-Microbiome-Host Interaction and Its Role in Metabolic Disease. *Nutrients*. 10(3):365 (2018).
149. Angelakis E, Armougom F, Carrière F, Bachar D, Laugier R, Lagier JC, Robert C, Michelle C, Henrissat B, Raoult D. A Metagenomic Investigation of the Duodenal Microbiota Reveals Links with Obesity. *PLoS One*. 10(9):e0137784 (2015).
150. De Filippo C, Cavalieri D, Di Paola M, et al. Impact of diet in shaping gut microbiota revealed by a comparative study in children from Europe and rural Africa. *Proc Natl Acad Sci U S A*. 107(33):14691-14696 (2010).
151. Martinez KB, Leone V, Chang EB. Western diets, gut dysbiosis, and metabolic diseases: Are they linked?. *Gut Microbes*. 8(2):130-142 (2017).
152. Meng, Y., Li, X., Zhang, J. *et al.* Effects of Different Diets on Microbiota in The Small Intestine Mucus and Weight Regulation in Rats. *Sci Rep* 9, 8500 (2019).
153. Turnbaugh PJ, Ridaura VK, Faith JJ, Rey FE, Knight R, Gordon JI. The effect of diet on the human gut microbiome: a metagenomic analysis in humanized gnotobiotic mice. *Sci Transl Med*. 1(6):6ra14 (2009).
154. Jang C, Hui S, Lu W, et al. The Small Intestine Converts Dietary Fructose into Glucose and Organic Acids. *Cell Metab*. 27(2):351-361.e3 (2018).
155. Grundy M.M.-L., Edwards C.H., Mackie A.R., Gidley M.J., Butterworth P.J., Ellis P.R. Re-evaluation of the mechanisms of dietary fibre and implications for macronutrient bioaccessibility, digestion and postprandial metabolism. *Br. J. Nutr.* 116:816–833 (2016).
156. Stecher B. The Roles of Inflammation, Nutrient Availability and the Commensal Microbiota in Enteric Pathogen Infection. *Microbiol Spectr*. 3(3) (2015).
157. Penttinen R, Kinnula H, Lipponen A, Bamford JK, Sundberg LR. High Nutrient Concentration Can Induce Virulence Factor Expression and Cause Higher Virulence in an Environmentally Transmitted Pathogen. *Microb Ecol*. 72(4):955-964 (2016).
158. Jang C, Hui S, Lu W, Cowan AJ, Morscher RJ, Lee G, Liu W, Tesz GJ, Birnbaum MJ, Rabinowitz JD. The Small Intestine Converts Dietary Fructose into Glucose and Organic Acids. *Cell Metab*. 27(2):351-361.e3 (2018).
159. Turnbaugh PJ, Bäckhed F, Fulton L, Gordon JI. Diet-induced obesity is linked to marked but reversible alterations in the mouse distal gut microbiome. *Cell Host Microbe*. 3(4):213-223 (2008).
160. Adike A, DiBaise JK. Small Intestinal Bacterial Overgrowth: Nutritional Implications, Diagnosis, and Management. *Gastroenterol Clin North Am*. 47(1):193-208 (2018).
161. Cade JE, Burley VJ, Warm DL, Thompson RL, Margetts BM. Food-frequency questionnaires: a review of their design, validation and utilisation. *Nutr Res Rev*. 17(1):5-22 (2004).
162. Ze, X., Duncan, S., Louis, P. *et al.* *Ruminococcus bromii* is a keystone species for the degradation of resistant starch in the human colon. *ISME J* 6, 1535–1543 (2012).
163. Hakenbeck R, Madhour A, Denapaite D, Brückner R. Versatility of choline metabolism and choline-binding proteins in *Streptococcus pneumoniae* and commensal streptococci. *FEMS Microbiol Rev*. 33(3):572-86 (2009).
164. Gosink KK, Mann ER, Guglielmo C, Tuomanen EI, Masure HR. Role of novel choline binding proteins in virulence of *Streptococcus pneumoniae*. *Infect Immun*. 68(10):5690-5 (2000).
165. Romano KA, Martinez-Del Campo A, Kasahara K, Chittim CL, Vivas EI, Amador-Noguez D, Balskus EP, Rey FE. Metabolic, Epigenetic, and Transgenerational Effects of Gut Bacterial Choline Consumption. *Cell Host Microbe*. 22(3):279-290.e7 (2017).
166. Zeisel SH, da Costa KA. Choline: an essential nutrient for public health. *Nutr Rev*. 67(11):615-23 (2009).

167. Walter J, Armet AM, Finlay BB, Shanahan F. Establishing or Exaggerating Causality for the Gut Microbiome: Lessons from Human Microbiota-Associated Rodents. *Cell*. 180(2):221-232 (2020).
168. Denaparte D, Rieger M, Köndgen S, Brückner R, Ochigava I, Kappeler P, Mätz-Rensing K, Leendertz F, Hakenbeck R. Highly Variable *Streptococcus oralis* Strains Are Common among Viridans Streptococci Isolated from Primates. *mSphere*. 1(2):e00041-15 (2016).
169. Do T, Jolley KA, Maiden MCJ, et al. Population structure of *Streptococcus oralis*. *Microbiology (Reading)*. 155(Pt 8):2593-2602 (2009).
170. Yang K, Han X. Lipidomics: Techniques, Applications, and Outcomes Related to Biomedical Sciences. *Trends Biochem Sci*. 41(11):954-969 (2016).
171. Lindon J., Nicholson J., Holmes E., Everett J. Metabonomics: metabolic processes studied by NMR spectroscopy of biofluids. *Concepts Magn Reson*. 12(5):289–320 (2000).
172. Johnson AJ, Vangay P, Al-Ghalith GA, Hillmann BM, Ward TL, Shields-Cutler RR, Kim AD, Shmagel AK, Syed AN; Personalized Microbiome Class Students, Walter J, Menon R, Koecher K, Knights D. Daily Sampling Reveals Personalized Diet-Microbiome Associations in Humans. *Cell Host Microbe*. 25(6):789-802.e5 (2019).
173. Garcia-Perez, I., Posma, J.M., Chambers, E.S. et al. Dietary metatype modelling predicts individual responses to dietary interventions. *Nat Food* 1, 355–364 (2020).
174. Oka P, Parr H, Barberio B, Black CJ, Savarino EV, Ford AC. Global prevalence of irritable bowel syndrome according to Rome III or IV criteria: a systematic review and meta-analysis. *Lancet Gastroenterol Hepatol*. 5(10):908-917 (2020).
175. Soubieres A, Wilson P, Poullis A, et al Burden of irritable bowel syndrome in an increasingly cost-aware National Health Service. *Frontline Gastroenterology*. 6:246-251 (2015).
176. Lee YJ, Park KS. Irritable bowel syndrome: emerging paradigm in pathophysiology. *World J Gastroenterol*. 20(10):2456-2469 (2014).
177. Edogawa S, Edwinston AL, Peters SA, Chikkamenahalli LL, Sundt W, Graves S, Gurunathan SV, Breen-Lyles M, Johnson S, Dyer R, Graham R, Chen J, Kashyap P, Farrugia G, Grover M. Serine proteases as luminal mediators of intestinal barrier dysfunction and symptom severity in IBS. *Gut*. 69(1):62-73 (2020).
178. Kashyap PC, Marcobal A, Ursell LK, Larauche M, Duboc H, Earle KA, Sonnenburg ED, Ferreyra JA, Higginbottom SK, Million M, Tache Y, Pasricha PJ, Knight R, Farrugia G, Sonnenburg JL. Complex interactions among diet, gastrointestinal transit, and gut microbiota in humanized mice. *Gastroenterology*. 144(5):967-77 (2013).
179. Roager HM, Hansen LB, Bahl MI, Frandsen HL, Carvalho V, Gøbel RJ, Dalgaard MD, Plichta DR, Sparholt MH, Vestergaard H, Hansen T, Sicheritz-Pontén T, Nielsen HB, Pedersen O, Lauritzen L, Kristensen M, Gupta R, Licht TR. Colonic transit time is related to bacterial metabolism and mucosal turnover in the gut. *Nat Microbiol*. 1(9):16093 (2016).
180. Duan R, Zhu S, Wang B, Duan L. Alterations of Gut Microbiota in Patients With Irritable Bowel Syndrome Based on 16S rRNA-Targeted Sequencing: A Systematic Review. *Clin Transl Gastroenterol*. 10(2):e00012 (2019).
181. Reigstad CS, Salmonson CE, Rainey JF 3rd, Szurszewski JH, Linden DR, Sonnenburg JL, Farrugia G, Kashyap PC. Gut microbes promote colonic serotonin production through an effect of short-chain fatty acids on enterochromaffin cells. *FASEB J*. (4):1395-403 (2015).
182. Lewis SJ, Heaton KW. Increasing butyrate concentration in the distal colon by accelerating intestinal transit. *Gut*. 41(2):245-51 (1997).
183. Shkoporov AN, Chaplin AV, Shcherbakova VA, et al. *Ruthenibacterium lactatiformans* gen. nov., sp. nov., an anaerobic, lactate-producing member of the family

- Ruminococcaceae isolated from human faeces. *Int J Syst Evol Microbiol*. 66(8):3041-3049 (2016).
184. Zhaolai Dai, Zhenlong Wu, Suqin Hang, Weiyun Zhu, Guoyao Wu, Amino acid metabolism in intestinal bacteria and its potential implications for mammalian reproduction, *Molecular Human Reproduction*. 21, (5) 389–409 2015.
 185. Kanehisa, M. and Goto, S.; KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Res*. 28, 27-30 (2000).
 186. Parker BJ, Wearsch PA, Veloo ACM, Rodriguez-Palacios A. The Genus *Alistipes*: Gut Bacteria With Emerging Implications to Inflammation, Cancer, and Mental Health. *Front Immunol*. 11:906 (2020).
 187. Douglas G Burrin, Barbara Stoll, Metabolic fate and function of dietary glutamate in the gut, *The American Journal of Clinical Nutrition*. 90, Issue (3) 850S–856S (2009).
 188. Zhou Q, Verne ML, Fields JZ, *et al*. Randomised placebo-controlled trial of dietary glutamine supplements for postinfectious irritable bowel syndrome. *Gut*. **68**:996-1002 (2019).
 189. Baggio L, Morrison M. The NAD(P)H-utilizing glutamate dehydrogenase of *Bacteroides thetaiotaomicron* belongs to enzyme family I, and its activity is affected by trans-acting gene(s) positioned downstream of *gdhA*. *J Bacteriol*. 178(24):7212-7220 (1996).
 190. Murray AW. The biological significance of purine salvage. *Annu Rev Biochem*. 40:811-826 (1971).
 191. Mackinnon AM, Deller DJ. Purine nucleotide biosynthesis in gastrointestinal mucosa. *Biochim Biophys Acta*. 319(1):1-4. doi:10.1016/0005-2787(73)90034-8 (1973).
 192. Chiang JYL. Bile acid metabolism and signaling in liver disease and therapy. *Liver Res*. 1(1):3-9 (2017).
 193. Chun C, Zheng L, Colgan SP. Tissue metabolism and host-microbial interactions in the intestinal mucosa. *Free Radic Biol Med*. 105:86-92 (2017).
 194. Jelacic K, Cimbri R, Nawaz F, Huang da W, Zheng X, Yang J, Lempicki RA, Pascuccio M, Van Ryk D, Schwing C, Hiatt J, Okwara N, Wei D, Roby G, David A, Hwang IY, Kehrl JH, Arthos J, Cicala C, Fauci AS. The HIV-1 envelope protein gp120 impairs B cell proliferation by inducing TGF- β 1 production and FcRL4 expression. *Nat Immunol*. 14(12):1256-65 (2013).
 195. Liu W, Li H, Hong SH, Piszczek GP, Chen W, Rodgers GP. Olfactomedin 4 deletion induces colon adenocarcinoma in *Apc*^{Min/+} mice. *Oncogene*. 2016 Oct 6;35(40):5237-5247 (2016).
 196. Tang Q, Jin G, Wang G, Liu T, Liu X, Wang B, Cao H. Current Sampling Methods for Gut Microbiota: A Call for More Precise Devices. *Front Cell Infect Microbiol*. 9;10:151 (2020).
 197. La Rosa SL, Leth ML, Michalak L, *et al*. The human gut Firmicute *Roseburia intestinalis* is a primary degrader of dietary β -mannans. *Nat Commun*. 10(1):905 (2020).
 198. Zhu C, Song K, Shen Z, Quan Y, Tan B, Luo W, Wu S, Tang K, Yang Z, Wang X. *Roseburia intestinalis* inhibits interleukin-17 excretion and promotes regulatory T cells differentiation in colitis. *Mol Med Rep*. 17(6):7567-7574 (2018).
 199. Hibberd AA, Lyra A, Ouwehand AC, *et al*. Intestinal microbiota is altered in patients with colon cancer and modified by probiotic intervention. *BMJ Open Gastroenterol*. 4(1):e000145 (2017).
 200. Nugent JL, McCoy AN, Addamo CJ, Jia W, Sandler RS, Keku TO. Altered tissue metabolites correlate with microbial dysbiosis in colorectal adenomas. *J Proteome Res*. 13(4):1921-1929 (2014).
 201. Juge N, Tailford L, Owen CD. Sialidases from gut bacteria: a mini-review. *Biochem Soc Trans*. 2016;44(1):166-175 (2016).

202. Irwin C, van Reenen M, Mason S, Mienie LJ, Wevers RA, Westerhuis JA, Reinecke CJ. The ¹H-NMR-based metabolite profile of acute alcohol consumption: A metabolomics intervention study. *PLoS One*. 10;13(5):e0196850 (2018).
203. Kim KM, Henderson GN, Ouyang X, et al. A sensitive and specific liquid chromatography-tandem mass spectrometry method for the determination of intracellular and extracellular uric acid. *J Chromatogr B Analyt Technol Biomed Life Sci*. 877(22):2032-2038. (2009).
204. Nugent JL, McCoy AN, Addamo CJ, Jia W, Sandler RS, Keku TO. Altered tissue metabolites correlate with microbial dysbiosis in colorectal adenomas. *J Proteome Res*. 13(4):1921-1929 (2014).
205. Sun Q, Jia Q, Song L, Duan L. Alterations in fecal short-chain fatty acids in patients with irritable bowel syndrome: A systematic review and meta-analysis. *Medicine (Baltimore)*. 98(7):e14513 (2019).
206. Kamphuis, J.B.J., Mercier-Bonin, M., Eutamène, H. et al. Mucus organisation is shaped by colonic content; a new view. *Sci Rep* 7, 8527 (2017).
207. Tropini C, Moss EL, Merrill BD, et al. Transient Osmotic Perturbation Causes Long-Term Alteration to the Gut Microbiota. *Cell*. 173(7):1742-1754.e17 (2018).
208. W. von Engelhardt, J. Bartels, S. Kirschberger, H.D. Meyer zu Düttingdorf & R. Busche. Role of short-chain fatty acids in the hind gut, *Veterinary Quarterly*, 20:sup3, 52-59 (1998).
209. Topping DL, Clifton PM. Short-chain fatty acids and human colonic function: roles of resistant starch and nonstarch polysaccharides. *Physiol Rev*. 81(3):1031-64 (2001).
210. Smith GW, Wiggins PM, Lee SP, Tasman-Jones C. Diffusion of butyrate through pig colonic mucus in vitro. *Clin Sci (Lond)*. 70(3):271-276 (1986).
211. Lock JY, Carlson TL, Carrier RL. Mucus models to evaluate the diffusion of drugs and particles. *Adv Drug Deliv Rev*. 124:34-49 (2018).
212. Navabi N, McGuckin MA, Lindén SK. Gastrointestinal cell lines form polarized epithelia with an adherent mucus layer when cultured in semi-wet interfaces with mechanical stimulation. *PLoS One*. 8(7):e68761 (2013).
213. Crater JS, Carrier RL. Barrier properties of gastrointestinal mucus to nanoparticle transport. *Macromol Biosci*. 10(12):1473-83 (2010).
214. Larhed AW, Artursson P, Gräsjö J, Björk E. Diffusion of drugs in native and purified gastrointestinal mucus. *J Pharm Sci*. 86(6):660-5 (1997).
215. Sellin JH, De Soignie R. Short-chain fatty acids have polarized effects on sodium transport and intracellular pH in rabbit proximal colon. *Gastroenterology*. 114(4):737-747 (1998).
216. Busche R, Bartels J, Genz AK, von Engelhardt W. Effect of SCFA on intracellular pH and intracellular pH regulation of guinea-pig caecal and colonic enterocytes and of HT29-19a monolayers. *Comp Biochem Physiol A Physiol*. 118(2):395-398 (1997).
217. Busche R, Dittmann J, Meyer zu Düttingdorf HD, Glockenthör U, von Engelhardt W, Sallmann HP. Permeability properties of apical and basolateral membranes of the guinea pig caecal and colonic epithelia for short-chain fatty acids. *Biochim Biophys Acta*. 1565(1):55-63 (2002).
218. McNeil NI, Ling KL, Wager J. Mucosal surface pH of the large intestine of the rat and of normal and inflamed large intestine in man. *Gut*. 28(6):707-713 (1987).
219. Genz AK, v Engelhardt W, Busche R. Maintenance and regulation of the pH microclimate at the luminal surface of the distal colon of guinea-pig. *J Physiol*. 517 (Pt 2) (Pt 2):507-519 (1999).
220. Ambort D, Johansson ME, Gustafsson JK, et al. Calcium and pH-dependent packing and release of the gel-forming MUC2 mucin. *Proc Natl Acad Sci U S A*. 109(15):5645-5650 (2012).

221. Karam SM. Lineage commitment and maturation of epithelial cells in the gut. *Front Biosci.* 4:D286-D298 (1999).
222. Béduneau A, Tempesta C, Fimbel S, et al. A tunable Caco-2/HT29-MTX co-culture model mimicking variable permeabilities of the human intestine obtained by an original seeding procedure. *Eur J Pharm Biopharm.* 87(2):290-298 (2014).
223. Navabi N, McGuckin MA, Lindén SK. Gastrointestinal cell lines form polarized epithelia with an adherent mucus layer when cultured in semi-wet interfaces with mechanical stimulation. *PLoS One.* 8(7):e68761 (2013).
224. Sontheimer-Phelps A, Chou DB, Tovaglieri A, et al. Human Colon-on-a-Chip Enables Continuous In Vitro Analysis of Colon Mucus Layer Accumulation and Physiology. *Cell Mol Gastroenterol Hepatol.* 9(3):507-526 (2020).
225. Hartman AL, Lough DM, Barupal DK, Fiehn O, Fishbein T, Zasloff M, Eisen JA. Human gut microbiome adopts an alternative state following small bowel transplantation. *Proc Natl Acad Sci U S A.* 106(40):17187-92 (2009).
226. Kreth J, Merritt J, Shi W, Qi F. Competition and coexistence between *Streptococcus mutans* and *Streptococcus sanguinis* in the dental biofilm. *J Bacteriol.* 187(21):7193-203 (2005).
227. van den Bogert B, Erkus O, Boekhorst J, de Goffau M, Smid EJ, Zoetendal EG, Kleerebezem M. Diversity of human small intestinal *Streptococcus* and *Veillonella* populations. *FEMS Microbiol Ecol.* 85(2):376-88 (2013).
228. van den Bogert B, Meijerink M, Zoetendal EG, Wells JM, Kleerebezem M. Immunomodulatory properties of *Streptococcus* and *Veillonella* isolates from the human small intestine microbiota. *PLoS One.* 9(12):e114277 (2014).
229. Simrén M, Barbara G, Flint HJ, Spiegel BM, Spiller RC, Vanner S, Verdu EF, Whorwell PJ, Zoetendal EG; Rome Foundation Committee. Intestinal microbiota in functional bowel disorders: a Rome foundation report. *Gut.* 62(1):159-76 (2013).
230. Romano KA, Vivas EI, Amador-Noguez D, Rey FE. Intestinal microbiota composition modulates choline bioavailability from diet and accumulation of the proatherogenic metabolite trimethylamine-N-oxide. *mBio.* 6(2):e02481 (2015).
231. Araújo JR, Tomas J, Brenner C, Sansonetti PJ. Impact of high-fat diet on the intestinal microbiota and small intestinal physiology before and after the onset of obesity. *Biochimie.* Oct;141:97-106 (2017).
232. Dale HF, Rasmussen SH, Asiller ÖÖ, Lied GA. Probiotics in Irritable Bowel Syndrome: An Up-to-Date Systematic Review. *Nutrients.* 11(9):2048 (2019).
233. Desai MS, Seekatz AM, Koropatkin NM, et al. A Dietary Fiber-Deprived Gut Microbiota Degrades the Colonic Mucus Barrier and Enhances Pathogen Susceptibility. *Cell.* 167(5):1339-1353.e21 (2016).
234. Lee JS, Wang RX, Goldberg MS, Clifford GP, Kao DJ, Colgan SP. Microbiota-Sourced Purines Support Wound Healing and Mucous Barrier Function. *iScience.* 26;23(6):101226 (2020).
235. Berrazaga I, Micard V, Gueugneau M, Walrand S. The Role of the Anabolic Properties of Plant- versus Animal-Based Protein Sources in Supporting Muscle Mass Maintenance: A Critical Review. *Nutrients.* 11(8):1825 (2019).
236. Magnúsdóttir S, Ravcheev D, de Crécy-Lagard V, Thiele I. Systematic genome assessment of B-vitamin biosynthesis suggests co-operation among gut microbes. *Front Genet.* 20;6:148 (2015).
237. Angela McCann, Ian B Jeffery, Bouchra Ouliass, Guylaine Ferland, Xueyen Fu, Sarah L Booth, Tam T T Tran, Paul W O'Toole, Eibhlís M O'Connor, Exploratory analysis of covariation of microbiota-derived vitamin K and cognition in older adults, *The American Journal of Clinical Nutrition.* 110, (6)1404–1415 (2019).

238. Halmos EP, Power VA, Shepherd SJ, Gibson PR, Muir JG. A diet low in FODMAPs reduces symptoms of irritable bowel syndrome. *Gastroenterology*. 146(1):67-75.e5 (2014).
239. Ong DK, Mitchell SB, Barrett JS, et al. Manipulation of dietary short chain carbohydrates alters the pattern of gas production and genesis of symptoms in irritable bowel syndrome. *J Gastroenterol Hepatol*. 25(8):1366-1373 (2010).
240. Marshall JK, Thabane M, Garg AX, Clark WF, Moayyedi P, Collins SM; Walkerton Health Study Investigators. Eight year prognosis of postinfectious irritable bowel syndrome following waterborne bacterial dysentery. *Gut*. 59(5):605-11 (2010).
241. Wensaas KA, Langeland N, Hanevik K, Mørch K, Eide GE, Rortveit G. Irritable bowel syndrome and chronic fatigue 3 years after acute giardiasis: historic cohort study. *Gut*. 61(2):214-9 (2012).
242. Hoskins LC, Agustines M, McKee WB, Boulding ET, Kriaris M, Niedermeyer G. Mucin degradation in human colon ecosystems. Isolation and properties of fecal strains that degrade ABH blood group antigens and oligosaccharides from mucin glycoproteins. *J Clin Invest*. 75(3):944-53 (1985).
243. Pohl, L.; Eckle, M. Sodium 3-(trimethylsilyl)tetrauteriopropionate, a new water-soluble standard for ¹H NMR. *Angewandte Chemie International Edition in English*. 8 (5): 381 (1969).
244. Nguyen, N., Warnow, T., Pop, M. et al. A perspective on 16S rRNA operational taxonomic unit clustering using sequence similarity. *npj Biofilms Microbiomes* 2, 16004 (2016).
245. Zeevi, D., Korem, T., Godneva, A. et al. Structural variation in the gut microbiome associates with host health. *Nature* 568, 43–48 (2019).
246. Veselkov KA. et al. Recursive segment-wise peak alignment of biological 1H NMR spectra for improved metabolic biomarker recovery. *Anal Chem*. 81(1):56-66 (2009).
247. Dona, A. C. et al. A guide to the identification of metabolites in NMR-based metabonomics/metabolomics experiments. *Comput. Struct. Biotechnol. J*. 14:135-153 (2016).
248. Gabriel Al-Ghalith, & danknights. (2020, April 30). knights-lab/BURST: BURST v1.00 (Version v1.0). Zenodo. <http://doi.org/10.5281/zenodo.3779009>
249. Gohl DM, Vangay P, Garbe J, MacLean A, Hauge A, Becker A, Gould TJ, Clayton JB, Johnson TJ, Hunter R, Knights D, Beckman KB. Systematic improvement of amplicon marker gene methods for increased accuracy in microbiome studies. *Nat Biotechnol*. 34(9):942-9 (2016).
250. Al-Ghalith, G.A., Hillmann, B., Ang, K., Shields-Cutler, R., and Knights, D. SHI7 Is a Self-Learning Pipeline for Multipurpose Short-Read DNA Quality Control. *mSystems* 3. (2018)
- 251.

Appendix

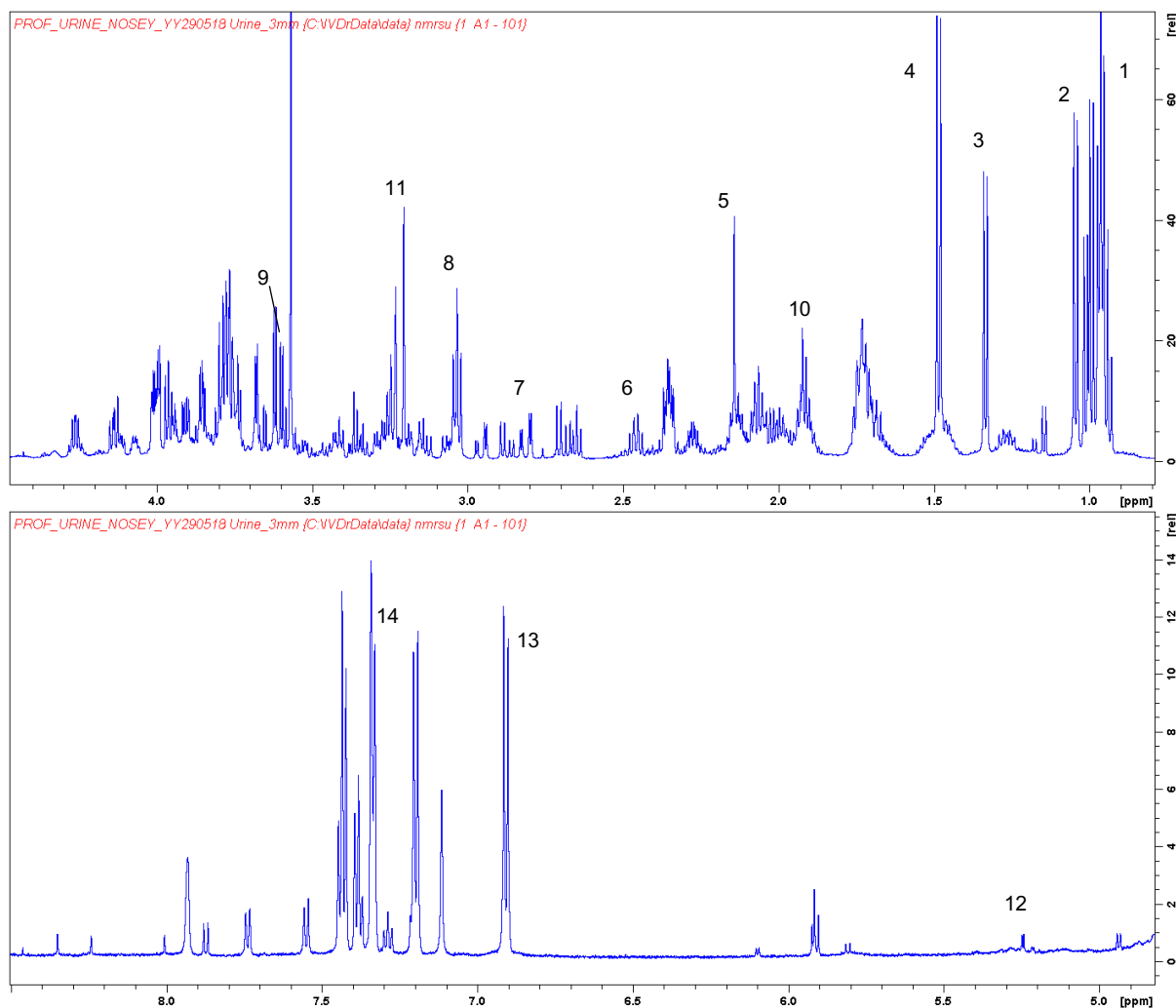


Figure 1. 600MHz ¹H NMR spectrum of a QC duodenum aspirate sample showing identified metabolite peaks for 1) Leucine, 2) Valine, 3) Lactate, 4) Alanine, 5) Methionine, 6) Glutamine, 7) Aspartate, 8) Lysine, 9) Threonine, 10) Acetate, 11) Choline, 12) Glucose, 13) Tyrosine, 14) Phenylalanine.

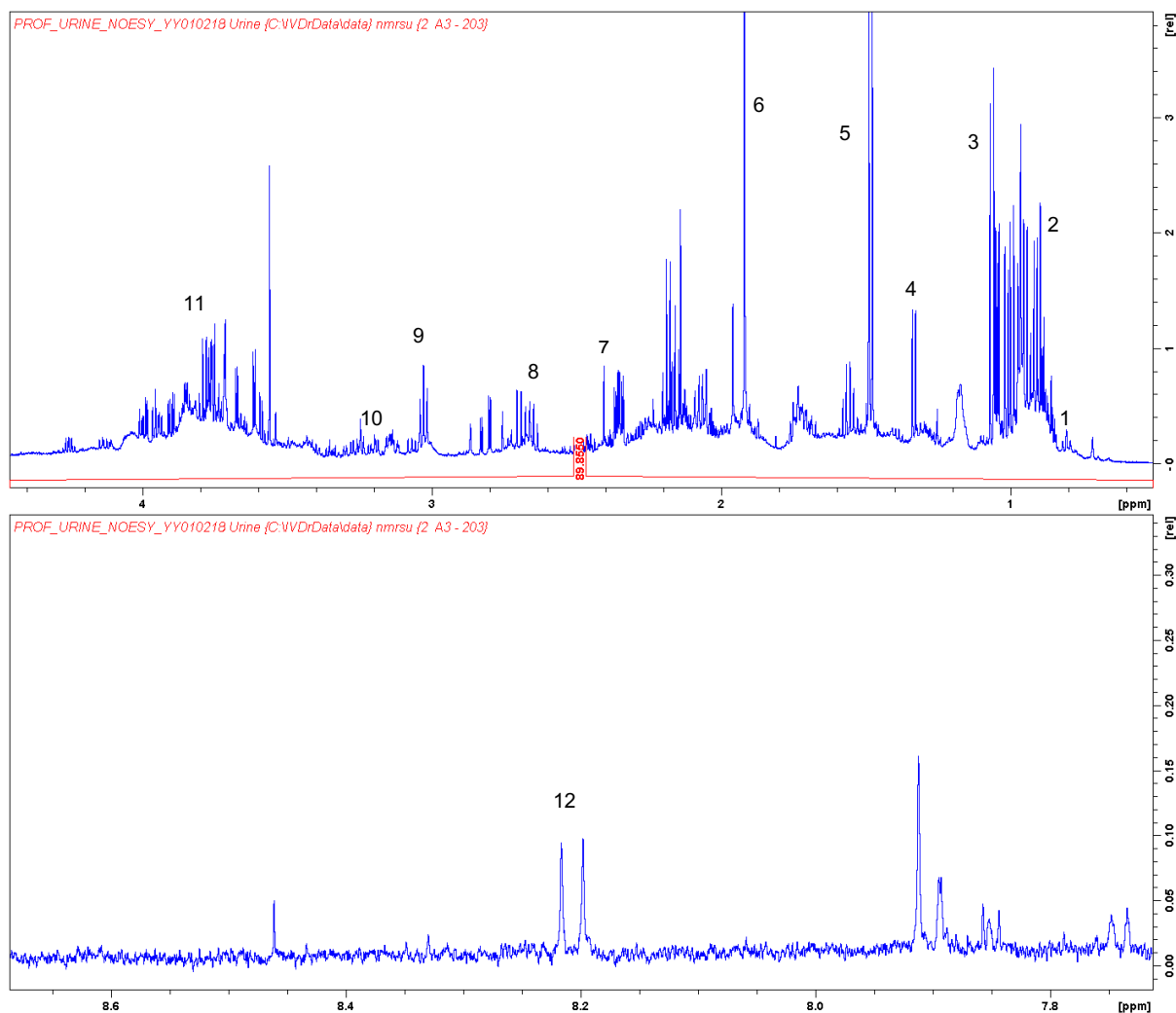


Figure 2. 600MHz ¹H NMR spectrum of a QC stool sample showing identified metabolite peaks for 1) 2-methylbutyrate, 2) Butyrate, 3) Propionate, 4) Lactate, 5) Alanine, 6) Acetate, 7) Succinate, 8) Aspartate, 9) Lysine, 10) Choline, 11) Glucose

Compound	Chemical shift ppm (multiplicities)
Alanine	1.46 (d), 3.76 (q)
Valine	0.976 (d), 1.05 (d), 2.261 (m), 3.601 (d)
Asparagine	2.84 (m), 2.94 (m), 4.00 (dd)
Methionine	2.143 (s), 2.648 (t), 3.872 (m)
Lysine	1.452 (m), 1.512 (m), 1.733 (m), 1.913 (m), 3.030 (t)
Threonine	1.316 (d), 3.575 (d), 4.244 (m)
Lactate	1.332 (d), 4.115 (q)
Aspartate	2.807 (d), 2.829 (d), 3.912 (dd)
Tyrosine	6.906 (d), 7.199 (d)
Glutamine	2.145 (m), 2.459 (m), 3.787 (t)
Leucine	0.961 (d), 0.972 (d), 1.718 (m), 3.740 (m)
Phenylalanine	7.333 (d), 7.381 (t), 7.432 (t)
Acetate	1.92 (s)
Glucose	3.416 (m), 3.539 (m), 3.715 (m), 3.771 (m), 3.842 (m), 5.238 (d)
Choline	3.205 (s), 3.526 (m), 4.070 (m)

Table 1. Metabolites identified in the duodenum aspirates for Chapter 3. Chemical shifts (ppm) and multiplicities are displayed.

Compound	Chemical shift ppm (multiplicities)
Valine	0.99 (d); 1.05 (d); 2.30 (m)
Isoleucine	0.94 (t); 1.02 (d); 1.27 (m)
Alanine	1.48 (d); 3.78 (q)
Lactate	1.33 (d); 4.11 (q)
Lysine	1.40 (m); 1.70 (m); 3.03 (t); 3.72 (t)
Succinate	2.41 (s)
Glucose	3.40 (t); 3.53 (dd); 3.71 (t); 3.73 (dd); 3.83 (m); 5.24 (d)
Acetate	1.92 (s)
Hypoxanthine	8.17 (s); 8.2 (s)
Butyrate	0.90 (t); 1.56 (t); 2.16 (t)
Aspartate	2.69 (dd); 2.82 (dd); 3.91 (dd)
Propionate	1.06 (t); 2.19 (m)
Choline	3.207 (d)
2methylbutyrate	0.86 (t); 1.05 (d); 1.38 (m); 1.5 (m); 2.21 (m)

Table 2. Metabolites identified in the stool sample for Chapter 3. Chemical shifts (ppm) and multiplicities are displayed.

Compound	Chemical shift ppm (multiplicities)
2-Methylbutyrate	0.86 (t); 1.05 (d); 1.38 (m); 1.5 (m); 2.21 (m)
Lactate	1.33 (d); 4.11 (q)
Alanine	1.48 (d); 3.78 (q)
Tyrosine	6.91 (d); 7.20 (d)
Isoleucine	0.94 (t); 1.02 (d); 1.27 (m)
Leucine	0.96 (d); 0.97 (d); 1.71 (m)
Valine	0.99 (d); 1.05 (d); 2.30 (m)
Lysine	1.40 (m); 1.70 (m); 3.03 (t); 3.72 (t)
Succinate	2.41 (s)
Glycine	3.57 (s)
b_arabinose	4.52 (d)
b_xylose	4.58 (d)
Acetate	1.92 (s)
Propionate	1.06 (t); 2.19 (m)
Butyrate	0.90 (t); 1.56 (t); 2.16 (t)
Glucose	4.645 (dd)
Isovalerate	0.92 (dd); 2.053(dd)
Uracil	5.81 (d); 7.54 (d)
Hypoxanthine	8.17 (s); 8.2 (s)
Glutamate	2.14 (m); 2.46 (m); 3.78 (m)

Table 3. Metabolites identified in the stool sample for Chapter 4. Chemical shifts (ppm) and multiplicities are displayed.

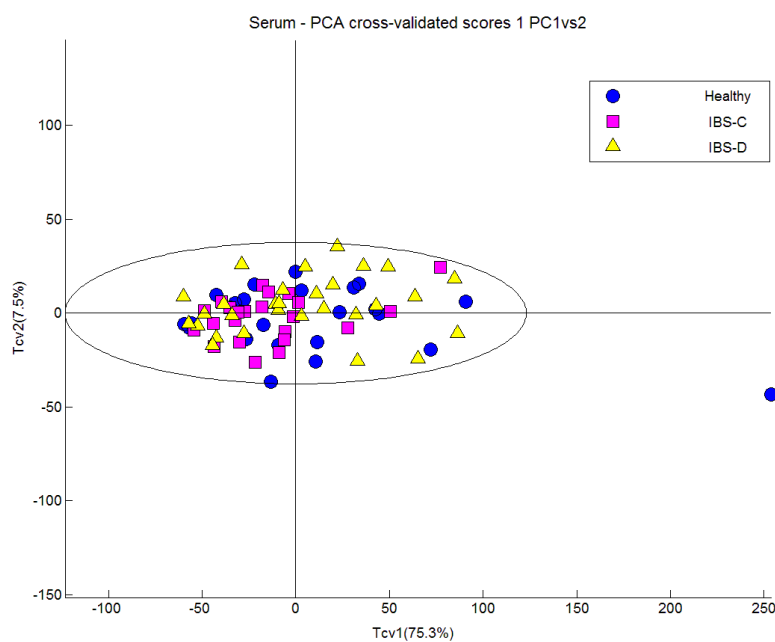


Figure 3. PCA scores plot of spectral variables from the 600 MHz ^1H NMR experiment of serum samples showing no clustering between IBS-subtypes and Healthy Controls.

Publications

1. Saffouri GB, Shields-Cutler RR, Chen J, **Yang Y**, Lekatz HR, Hale VL, Cho JM, Battaglioli EJ, Bhattarai Y, Thompson KJ, Kalari KK, Behera G, Berry JC, Peters SA, Patel R, Schuetz AN, Faith JJ, Camilleri M, Sonnenburg JL, Farrugia G, Swann JR, Grover M, Knights D, Kashyap PC. Small intestinal microbial dysbiosis underlies symptoms associated with functional gastrointestinal disorders. *Nat Commun.* 10(1):2012 (2019) doi: 10.1038/s41467-019-09964-7.
2. Mars RAT, **Yang Y**, Ward T, Houtti M, Priya S, Lekatz HR, Tang X, Sun Z, Kalari KR, Korem T, Bhattarai Y, Zheng T, Bar N, Frost G, Johnson AJ, van Treuren W, Han S, Ordog T, Grover M, Sonnenburg J, D'Amato M, Camilleri M, Elinav E, Segal E, Blekhman R, Farrugia G, Swann JR, Knights D, Kashyap PC. Longitudinal Multi-omics Reveals Subset-Specific Mechanisms Underlying Irritable Bowel Syndrome. *Cell.* 182(6):1460-1473.e17 (2020). doi: 10.1016/j.cell.2020.08.007.