CORA Cork Open Research Archive
Cartlann Taighde Oscailte Chorcaí

Title	Development of knowledge base and methodology for the rational microbiome modulation in irritable bowel syndrome
Authors	Das, Anubhav
Publication date	2023
Original Citation	Das, A. 2023. Development of knowledge base and methodology for the rational microbiome modulation in irritable bowel syndrome. PhD Thesis, University College Cork.
Type of publication	Doctoral thesis
Rights	© 2023, Anubhav Das. - https://creativecommons.org/licenses/by-nc-nd/4.0/
Download date	2026-10-28 13:16:04
Item downloaded from	https://hdl.handle.net/10468/15038



UCC
University College Cork, Ireland
Coláiste na hOllscoile Corcaigh

Ollscoil na hÉireann, Corcaigh

National University of Ireland, Cork



Development of knowledge base and methodology for the Rational Microbiome Modulation in Irritable Bowel Syndrome

Thesis presented by

Anubhav Das, M. Tech

for the degree of

Doctor of Philosophy

University College Cork

School of Microbiology

Head of School: Dr. David Clarke

Supervisor(s): Prof. Paul W. O'Toole, Prof. Pavel Baranov, Dr. Ian B. Jeffery

2023

Dedicated to my parents, Tutu and Ashis Das

Table of Contents

Declaration ii

List of publications.....iii

Table of abbreviations iv

Acknowledgments.....vi

Abstract ix

Chapter 1: Review of literature: Gut microbiome, host health, and scope for rational microbiome modulation. 1

Chapter 2: Differences in Faecal Microbiomes and Metabolomes of People With vs Without Irritable Bowel Syndrome and Bile Acid Malabsorption. 69

Chapter 3: The faecal mycobiome in patients with Irritable Bowel Syndrome 204

Chapter 4: The gut virome in Irritable Bowel Syndrome differs from that of controls. 242

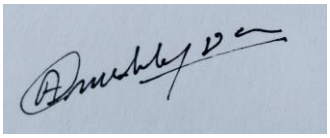
Chapter 5: Integrative microbiome analysis shows that loss of inter-kingdom interactions is a potential driver of microbiome alterations in Irritable Bowel Syndrome. 327

Chapter 6: Bacterial community-level metabolic modelling reveals differences in metabolic capabilities and interactions within the gut microbial ecosystem in Irritable Bowel Syndrome. 389

Chapter 7: Discussion.....525

Declaration

This is to certify that the work I am submitting is my own and has not been submitted for another degree, either at University College Cork or elsewhere. All external references and sources are clearly acknowledged and identified within the contents. I have read and understood the regulations of University College Cork concerning plagiarism and intellectual property.



Signed: _____

Anubhav Das (Student Id: 119222654)

March 2023

List of publications

1. Ian B. Jeffery#, **Anubhav Das#**, Eileen O'Herlihy, Simone Coughlan, Katryna Cisek, Michael Moore, Fintan Bradley, Tom Carty, Meenakshi Pradhan, Chinmay Dwibedi, Fergus Shanahan, Paul W. O'Toole*. Differences in Fecal Microbiomes and Metabolomes of People With vs Without Irritable Bowel Syndrome and Bile Acid Malabsorption. Gastroenterology 2020;158:1016-1028.e8.

DOI: 10.1053/j.gastro.2019.11.301.

2. **Anubhav Das**, Eileen O'Herlihy, Fergus Shanahan, Paul W. O'Toole, Ian B. Jeffery*. The fecal mycobiome in patients with Irritable Bowel Syndrome. Sci Rep. 2021;11:124.

DOI: 10.1038/s41598-020-79478-6.

3. Simone Coughlan, **Anubhav Das**, Eileen O'Herlihy, Fergus Shanahan, Paul W. O'Toole, and Ian B. Jeffery*. The gut virome in Irritable Bowel Syndrome differs from that of controls. Gut Microbes 2021;13:1887719.

DOI: 10.1080/19490976.2021.1887719.

Authors who share first authorship; * Corresponding author

Patents

1. O'Toole, P.W., Shanahan, F., Jeffery, I.B., O'Herlihy, E., **Das, A.** Methods of diagnosing disease. WIPO PCT (2020), 188pp. WO2020201457A1.

Table of Abbreviations

AGORA	Assembly of gut organisms through reconstruction and analysis
AUC	Area under the curve
BA	Bile Acids
BAM	Bile acid malabsorption
BAs	Bile acids
BMI	Body mass index
CAGs	Coabundant groups
CDI	Clostridium difficile infection
COBRA	Constraint based Reconstruction and Analysis
CRC	Colorectal Cancer
DTT	Dithiothreitol
FAIMS	Field asymmetric waveform ion mobility spectrometry
FBA	Flux Balance Analysis
FDR	False Discovery Rate
FFQ	Food Frequency Questionnaire
FMT	Fecal microbiota transplantation
FODMAPs	Fermentable oligosaccharides, disaccharides, monosaccharides and polyols
GABA	Gamma-Amino Butyrate
GC	Gas chromatography
GENRE	Genome-scale network reconstruction
GI	Gastrointestinal
HADS	Hospital Anxiety and Depression Scale
HMP	Human Microbiome Project
IBD	Inflammatory bowel disease
IBS	Irritable bowel syndrome
IBS-C	Constipation predominant IBS
IBS-D	Diarrhoea predominant IBS
IBS-M	Mixed IBS
IBS-QoL	IBS quality of life
IBS-SSS	IBS symptom severity scores
IBS-U	Unclassified IBS
IMGVR	Integrated Microbial Genomes/Virus
IPCO	Inference of Pathways from Co-variance analysis
ITS	Internal transcribed spacers
LBPs	Live Biotherapeutic Products
LC	Liquid chromatography
LOO-CV	Leave-one-out cross validation
MAGs	Metagenome-assembled genomes
MS	Mass spectrometry
ORFs	Open reading frames
OTU	Operational taxonomic unit

PCA	Principal component analysis
PCoA	Principal coordinate analysis
PEG	Polyethylene glycol
pFBA	Parsimonious flux balance analysis
PPIs	Proton Pump Inhibitors
PRR	Pattern recognition receptors
pVOGs	Prokaryotic viral orthologous groups
ROC	Receiver Operating characteristics
rRNA	Ribosomal RNA
RUTF	Ready-to-use therapeutic food
SCFAs	Short-chain fatty acids
SeHCAT	^{75}Selenium homocholic acid taurine
SNPs	Single Nucleotide Polymorphisms
SVMs	Support vector machines
UDCA	Ursodeoxycholic acid
VCs	Viral Clusters
VLPs	Virus-Like Particles
VOGDB	Virus orthologous groups
zOTUs	Zero-radius operational taxonomic units

Acknowledgements

My research work in the form of this thesis would not have been possible without the help of all the people who have motivated, guided, and to some extent, tolerated me during my journey. Firstly, I must thank my supervisor Prof. Paul W. O'Toole for providing me the opportunity to work with him. I will always be grateful for his mentorship, guidance, and straightforward comments, which has helped me grow as a researcher. I would like to extend my sincere gratitude towards my co-supervisors Dr. Ian B. Jeffery, and Prof. Pavel Baranov, who helped me in developing this thesis, and provided me all the necessary guidance. I will always be grateful to Ian, not only for helping me grow as a computational biologist, but also for supporting me, and keeping me motivated during stressful times. My sincere thanks to Prof. Fergus Shanahan, for providing valuable inputs, especially a clinical perspective, which has helped me refine my thesis. I would also like to thank Dr. Tarini Shankar Ghosh, who used to be my supervisor at my previous workplace and later became part of Prof. O'Toole's research group, for introducing me to this field of study, and for his technical inputs and scientific discussions, which has helped me further my technical skills. I am thankful to all my co-authors, whose work has helped me shape my own research. I extend my sincere gratefulness to Dr. Jerry Reen, and other administrative staff of the School of Microbiology, Graduate Studies Office, and University College Cork, who have helped me in all the required formalities.

I would also like to thank my employer, 4D Pharma Cork Ltd. and its management, for funding my PhD, and providing me with opportunities to grow academically as well as in industry. A special thanks to Dr. Eileen O'Herlihy, whose administrative assistance helped my relocation to Ireland, and my onboarding hassle free, and

pleasant. I am also grateful to Ms. Lucy Christie, and Ms. Jemma Bruce, for taking care of all my formal requirements.

A good workplace exits because of friendly and helpful colleagues, and it has been my pleasure to work with such a fun-loving group. I am thankful to my colleagues Luke, Fiona, Meenakshi, Marie, Chinmay, Siobhan, Darren, David, Jamie, Katryna, Simone, Jack, Kevin, Naresh, and Tony, for providing a conducive work environment along with all scientific and technical inputs. A very special thanks to my Bioinformatics group members, Dr. Emma Raftis, Stephanie, Ana, for their inputs, and other discussion encompassing a variety of topics. I am also thankful to our past research nurses of 4D Pharma Cork Ltd., Sinead, Marian, and Caroline, for their contribution in collecting samples and other data which has helped me in my analysis.

My personal friend circle has been as important in maintaining my patience and composure. I extend my heartfelt thanks to my friends Piyush, Ritika, Gaurav, Ipsita, Shashank, Priya, Girish, and Preeti for their support and not making me miss India. I also owe my gratitude to my supportive friends back in India, whose constant encouragement has helped me pass through this rigorous journey of a PhD scholar. I am thankful to my friends, and colleagues from my previous company, Swadha, Kuntal, Sunil, Deepak, Disha, Vishnu, and Ritesh, who supported me and helped me during my academic journey. I am also thankful to my past mentors, my M.Tech thesis supervisor, Prof. Mainak Das, and my supervisor during my past employment, Dr. Sharmila Mande, for instilling confidence in me to pursue research in any desired field of study.

I definitively have missed a few names, but that does not undermine their contribution in my life, and I am thankful to them too.

Lastly, coming to the most important support system of my life, my family. I am thankful to my loving wife, Bhagyashree, who has been my strength and a constant source of motivation, which has helped me overcome all the hardships I had faced during this phase of my life. I am also thankful to my extended family, my in-laws. At the end, I would like to dedicate this thesis to my mother (Mrs. Tutu Das), and my father (Mr. Ashis Das), whose struggles, encouragement, guidance, and support has brought me at this point, and I will always be indebted to them.

Abstract

The gut microbiota typically comprises a steady-state community whose composition and functions are governed in part by factors including diet, disease, medications, ethnicity, geographic location, host genotype, and other lifestyle and clinical factors. There is a large body of research that supports the role of the microbiota in disease, host gene regulation and maintenance of host physiology.

My thesis focuses on Irritable Bowel Syndrome (IBS), a chronic functional gastrointestinal disorder, associated with alterations in microbiota composition and function. IBS is marked by symptoms like pain and distension of the abdomen, abnormal bowel habits, leading to social disablement. Patients with IBS are mainly classified into four clinical sub-types based on symptoms, viz., IBS-C (Constipation predominant), IBS-D (diarrhoea predominant), IBS-M (mixed), and IBS-U (unclassified), but there is heterogeneity within IBS at various levels including symptoms as well as microbiome composition. Attributed by some investigators as a disorder of the gut-brain axis, the aetiology of IBS remains unclear.

There has been growing interest in the development of efficient microbiome based diagnostic tools and therapeutic products that rely on robust and biologically relevant biomarkers which can be identified through a deeper understanding of the disorder. Considering this, the goals of my thesis include investigation of alterations in microbiome composition, function, and structure in IBS, which could be applied to diagnose or stratify IBS patients. My research also focused on understanding inter-microbial interaction patterns, and development of statistical models which can aid in screening of potential live bio-therapeutic products (LBPs) and augment transformative therapeutics for microbiome related disorders like IBS.

To work toward the goals of my thesis, I analyzed a cohort of 80 IBS patients and 65 matched Control subjects. Using state of the art bioinformatics methods, along with statistical, and systems biology-based approaches, I explored the various microbiome data types searching for insights into the underlying biology of IBS pathophysiology.

Firstly, I took a multi-omics analysis approach to investigate alterations in microbiome community composition and structure, predicted functionality, and faecal and urine metabolome, along with dietary habits in IBS. Subsequently, I analyzed the non-bacterial components of the microbiome (mycobiome and virome) to further understand their potential roles in IBS progression. An integrative, inter-kingdom analysis of the microbiome was also conducted to explore the differences in inter-kingdom interactions in IBS as well as comparison of the ability of these components to predict IBS. Finally, I performed a systems biology-based metabolic modelling of microbiomes of the samples to evaluate changes in metabolic events, and inter-microbial metabolic interactions in IBS, followed by identification of species predicted as being capable of modulating the overall metabolic output of the microbiome community of IBS and Controls.

The findings of my analysis not only corroborated previously reported observations, but also provided novel insights which have clinical implications. I observed significant differences in faecal microbiome and metabolome of patients with IBS and Bile Acid Malabsorption (BAM) and developed predictive models to stratify IBS patients. I also observed reduced agreement within microbiome, and between omics datasets, in datasets from subjects with IBS. Inter-kingdom analyses also indicated a lack of concordance between bacteriome and non-bacterial components in IBS. Metabolic modelling analysis showed differences in predicted reaction rate profiles between IBS and Controls, along with lack of co-operation within the bacterial

community, and provided a distinct taxonomic-metabolomic signature of dysbiosis in IBS. Based on perturbation analysis of the metabolic models, I could identify species acting as potent modulators of the community metabolic output. Given the stochasticity associated with the microbiota of IBS, there is scope for these modulating species to directionally perturb the microbial community towards a more favourable structure (i.e., one promoting gut health), thus representing a therapeutic target.

Chapter 1

Review of literature: Gut microbiome, host health, and scope for rational microbiome modulation.

Contents

1.1 Introduction 3
1.2 Microbiome in health and diseases 4
1.2.1 The bacteriome 5
1.2.2 The mycobiome 7
1.2.3 The phageome/virome 9
1.3 Microbiome community composition is influenced by multiple factors 11
1.3.1 Diet influences the microbiome in a personalized manner. 11
1.3.2 Age determines the evolution of microbiome community in a host. 12
1.3.3 Host genetics 14
1.3.4 Medication 14
1.3.5 Other extrinsic factors 16
1.4 Metabolic effect of the gut microbiota 17
1.5 Modulation of the microbiome 21
1.5.1 Antibiotic intervention 21
1.5.2 Prebiotics and dietary intervention 22
1.5.3 Probiotics 22
1.5.4 Faecal microbiota transplantation (FMT) 23
1.5.5 Single strain live biotherapeutic products 24
1.6. Understanding the biology of the microbiome through bioinformatics 25
1.6.1 Looking beyond composition; Exploration of functional potential of the microbiome. 25
1.6.2 Microbial metabolomics 26
1.7 Systems Biology: A scope for rational microbiome modulation 27
1.7.1 Mathematical modelling of the microbiome 28
1.7.2 Genome scale metabolic models (GENREs): Generation protocols and available tools. 29
1.7.3 Constraint-based modelling of the microbiome. 31
1.8. Perspective: Method integration and possible applications of microbiome perturbation studies 32
1.9. Thesis scope and defined objectives 34
1.10 References 37

1.1 Introduction

Trillions of microorganisms reside within the gut. Weighing just 200g, and with a cell count equivalent to the number of eukaryotic cells within our bodies, these microorganisms, consisting of mostly bacteria with minor but significant contributions from archaea, fungi, and viruses, compose an intestinal ecosystem that is essential to our physical and mental health [1,2]. As well as the gut, the human body contains other environmental niches at different body sites like skin, mouth, stomach, lungs and vagina[3]. Whilst important in their own right, the gut microbiome is the most densely and diversely populated and is the focus of my work due to its importance for health and its implication in various disorders including enteric disorders like IBD, and IBS [4–8].

The microbiome community composition varies across the length of the human gastrointestinal tract, from stomach (least diverse microbiome) to the colon, where diversity and richness of the microbiome reaches its peak [9,10]. A plethora of microbes co-exist and interact with each other, influencing the functioning of the gut. The majority of the gut microbes are strict anaerobes and belong to the phyla *Firmicutes* and *Bacteroidetes*, with other Phyla such as *Proteobacteria* and *Actinobacteria*, making significant contributions[11].

For individuals with a shared geographical/ancestral background; within the substantial inter-individual variation observed in the microbiome community composition [12,13], at population level, there exists a microbiome community which occurs consistently within subject groups, also referred to as the core microbiome [14,15]. Outside of the core microbiome, previously published literature have proposed the existence of enterotypes [16], wherein, the microbiota composition can be divided into

three microbiota configurations based on the sub-set of microbial species that dominate the individual gut microbiome. With more in-depth analysis on different datasets by various groups, the notion of the enterotypes has been debated [17–19]. It is now accepted that the enterotypes are not discrete clusters of samples, but are part of a continuous gradient of microbiota configurations [17,19,20].

From birth, the microbiota is known to establish itself and to rapidly increase in diversity in the neonatal period [21]. As the diversity increases in early childhood, the microbiota gradually stabilizes into the microbiota structures seen in healthy adults [22]. Microbiome profiles from within a healthy population, can be described as a stable ecosystem whose composition and functions are resistant to drastic changes and continues to exist within low levels of variability over time [23]. The microbiome tends to regain a stable state after being exposed to changes in its environment, suggesting better resilience to perturbations [24,25]. In general, a healthy individual's microbiome can be characterized by increased diversity, richness, presence of beneficial microbes, and concordance among them. It is observed that the microbiome of individuals deviating from healthy often presents a marked decrease in overall diversity and an overgrowth of pathobionts accompanied by diminishing presence of symbionts [26,27].

1.2 Microbiome in health and diseases

Owing to the large proportion of bacteria in the microbiome, and extensive number of bacteria centric research, the term microbiome is often used for the bacterial community (or bacteriome). However, the microbiome is a collection of the genomes of all the microbes which co-exist in a particular site. The microbiota, dominated by bacteria, also consists of archaea, fungi (mycobiome), and viruses (virome or phageome). These other branches of the tree of life make up a small proportion (by

weight) of our gut microbiota compared to the bacteriome but add significant genetic and functional variability.

These microbiome components have been implicated in multiple ailments and show potential to be biomarkers [28–30], but the association between the state of the microbiome and host health status is often complex and bi-directional in nature [28]. The alteration in host health might trigger alterations in the microbiota, which upon microbiome state adaptation, may aid in perpetuation of the disease state through secreted metabolites. Here I discuss a few important health conditions, and their associations with the microbial community members.

1.2.1 The bacteriome

Being the major contributor to the microbiome, the bacteriome comprises of greater than 90% of the genes of non-host origin [31], and its alterations have been implicated in different gastrointestinal, and metabolic diseases [28], few of which are briefly discussed below.

Condition	Bacterial alterations	References
Colorectal Cancer (CRC)	Some common bacterial signatures have been identified including *Bacteroides fragilis*, *Fusobacterium nucleatum*, *Parvimonas micra*, *Porphyromonas asaccharolytica*, *Prevotella intermedia*, *Alistipes finegoldii*, and *Thermanaerovibrio acidaminovorans*. The oral microbiota has also been shown to be predictive of CRC.	[32,33]

Irritable bowel syndrome (IBS)	Studies have shown an important role of the gut microbiota in its progression, and a greater abundance of *Ruminococcus torques*, and *Ruminococcus gnavus* in IBS, along with a statistically non-significant but a multi-fold increase of *Akkermansia muciniphila*, a mucin degrader, in IBS.	[6,7,34,35]
Inflammatory bowel disease (IBD)	*Christensenellaceae*, *Coriobacteriaceae*, and *Faecalibacterium prausnitzii* were found to be reduced in Crohn's disease, and *Veillonella*, and *Escherichia coli* showed increased relative abundance compared to controls. *Eubacterium rectale*, and *Akkermansia* were reduced, and abundance of *Escherichia coli* was increased in ulcerative colitis. Bacterial diversity was also observed to be reduced in IBD.	[36]
Clostridioides difficile infection (CDI)	A reduced bacterial diversity has been observed in CDI, along with a decreased abundance of *Bacteroides*, *Prevotella*, and *Bifidobacterium*, and increased abundance of *Clostridium*, and *Lactobacillus*.	[37,38]

Cardiovascular diseases	Conversion of Tri-methylamine (TMA) from choline, lecithin, and carnitine mediated by gut bacteria, is converted to trimethylamine N-oxide (TMAO) by the host which may cause atherosclerosis.	[39]
Age related frailty	Gut bacterial composition correlates with health in elderly. An increase in pathobionts and a reduction in symbionts are linked with unhealthy ageing, and increased frailty.	[40,41]

Table 1: Association between gut bacteria and selected disorders.

1.2.2 The mycobiome

The mycobiome or the collection of fungi in an environment, is an understudied component of the microbiome. The fungal community makes up just 0.01% of the gut metagenome [42], but plays an important role in gut homeostasis, and pathogenesis [29]. Although research into the mycobiota is an emerging field, currently available literature suggests that the mycobiome is not only associated with health conditions, but it can also interact with the bacteriome, thereby influencing the bacterial community [43–47]. Table 2 summarizes the alterations in the gut mycobiome in selected disorders. It is interesting to note that the bacterial diversity and *Candida* abundance are consistently observed to be inversely related, which suggests a damaging influence of the fungi in bacteriome stability.

Condition	Fungal alterations	References
Colorectal Cancer (CRC)	Decrease in alpha diversity; Increase in *Malassezia*, and *Basidiomycota*-to-*Ascomycota* ratio.	[48–51]
Irritable bowel syndrome (IBS)	Overall decrease in alpha diversity; Increase in *Candida*, *Candida albicans*, *Malassezia*, *Cladosporium*, and *Saccharomyces cerevisiae*; decrease in *Mycosphaerella*, and *Aspergillus*.	[52–55]
Inflammatory bowel disease (IBD)	Crohn's disease: Increase in alpha diversity, and fungal load; Increase in *Basidiomycota*-to-*Ascomycota* ratio, *Candida* increase, and *Candida albicans*; Decrease in *Saccharomyces cerevisiae*.	[56–63]
	Ulcerative colitis: Decrease in alpha diversity; Increase in *Debaryomyces*, and *Candida albicans*; Decrease in *Alternaria alternata*, *Aspergillus flavus*, and *Aspergillus cibarius*.	[56,59,64]
Clostridioides difficile infection (CDI)	Decreased overall diversity; Increased *Ascomycota*-to-*Basidiomycota* ratio; Increased *Cladosporium*, *Aspergillus*, *Candida albicans*.	[65,66]

Graft Versus Host Disease	Increase in *Candida*, *Candida albicans*, and *Candida parapsilosis*.	[67–69]

Table 2: Gut mycobiome alterations in human disorders.

1.2.3 The phageome/virome

The phageome or virome is the collection of viruses in an environment, which like the mycobiome, is understudied and understanding its role in modulation of the microbiome is ongoing. Like the mycobiome, its signal is small compared to bacteria, but recent literature suggests a link between alterations in phage abundance and disease[70]. While analysis of bacteriome and mycobiome is straight forward, estimation of the viral community in the gut has its own challenges. Viruses can occur as part of the host genome as prophages (lysogenic cycle) and can move to the lytic cycle to produce virus particles, which move out of the host cell after killing it. The viruses within the gut microbiome may play an important role in modulating the gut microbiota by altering its density, diversity and functionality [71], thereby influencing the host physiology and progression of diseases [30,72]. The alterations in the gut virome in some selected human diseases are summarized in Table 3.

Condition	Virome alterations	References
Colorectal Cancer (CRC)		
Irritable bowel syndrome (IBS)	Reduced alpha diversity of virome, and greater heterogeneity in IBS.	[73]
Inflammatory bowel disease (IBD)	Both CD and UC: Increase in virome diversity. Increased *Caudovirales,* and decreased *Microviridae* abundance.	[74]
	Crohn's disease: Reduced virulent core, increase in temperate virions, and shift from lysogeny to lytic cycle increased in CD.	[70]
	Ulcerative colitis: Greater abundance of *Caudovirales*; Increased *Escherichia*, and *Enterobacteria* phages.	[75]
Clostridioides difficile infection (CDI)	Increased abundance of *Caudovirales*, and *Anelloviridae*. Reduced abundance of *Microviridae*	[76]

Table 3: Summary of virome alterations in selected diseases.

1.3 Microbiome community composition is influenced by multiple factors

Although a healthy microbiome remains relatively stable over the course of an adult's life, the community composition and functions are dynamic and respond to multiple factors such as host genotype, age, temperature, diet, change of location, host immune system, medications, amongst other lifestyle related, and clinical attributes [24,77].

1.3.1 Diet influences the microbiome in a personalized manner.

A balanced diet, comprising of recommended amounts of macro and micro nutrients is known to be beneficial for health, and multiple studies have indicated that a diverse diet is associated with positive health outcomes [78]. As with the human health, the microbiome health reflected by its composition is also known to be influenced by dietary habits of individuals [79–82]. Although diet influences the microbiome composition, the effect is mediated by the microbiome, which may be considered unique to each individual, and a previously published study, based on daily sampling, showed personalized diet-microbiome associations[83]. A short-term change in diet can change the microbiome of an individual, but the change is temporary and its persistence does not remain beyond a few days [81]. Despite alterations in microbiota with short-term dietary interventions, the microbiome remains stable, and inter-individual variability is not lost in the short to medium term. This has been attributed to the long term dietary habits of individuals [79,82]. A recent dietary intervention study on older adults has shown a modest but positive effect of high compliance to the Mediterranean diet on frailty and health status [84]. It is also worth noting that complete elimination of certain components like dietary fibres from regular diet may lead to increase in abundance of mucin degrading bacteria, which in turn can erode the mucus layer of the epithelium, thereby impacting the gut barrier function [85].

The gut mycobiome composition is also affected by dietary habits [86]. It has been observed that mycobiome composition is different between individuals with broadly different dietary patterns, especially between vegetarian and western diet [87]. While, an increased abundance of *Candida* is associated with a carbohydrate-rich diet [88,89], a protein rich diet is linked with reduction of *Candida* in healthy individuals [88,89]. A previous study of gut mycobiome across populations in China, linked changes in abundance of multiple gut fungal species to ethnicity-specific diets [86].

Pertaining to IBS, the dietary habits of IBS patients, including irregular eating, deviate from those of healthy subjects [90]. Moreover, there have been reports of lack of tolerance of certain food items by IBS patients, especially FODMAPs (short for fermentable oligosaccharides, disaccharides, monosaccharides and polyols) [91,92]. Fermentable saccharides have been reported to induce colon distention and production of hydrogen gas due to fermentation in healthy individuals [93,94], was also observed in IBS subjects, but colonic hypersensitivity to distension and not gas production was attributed to carbohydrate related symptoms in IBS [95]. A previous study also reported 8-fold reduction in urinary histamine (a molecular mediator of visceral hypersensitivity [96–99]) upon low FODMAP dietary intervention compared to high FODMAP diet [100]. The study also reported changes in microbiota, wherein, a low FODMAP diet increased the richness and diversity of *Actinobacteria*, and a high FODMAP diet was associated with decrease in abundance of gas consuming bacteria [100].

1.3.2 Age determines the evolution of microbiome community in a host.

Colonisation of the human body by microbes starts in the post-natal stage, but some studies report that an infant's meconium is not sterile and contains multiple microbes[101]. Such microbes act as seeding elements for the developing microbial ecosystem and further determine the course of microbiome evolution inside the host with time [77]. However considerable uncertainty still exists, with good evidence

showing that this may not be the case and arguing that the meconium is sterile at birth [101,102]. Evolution of microbiome is also determined by the mode of birth, and differences between microbiome of infants born by vaginal delivery and C-Section have been reported [103,104]. The neonatal microbiome is characterized by relative instability compared to adults, and with increased age and introduction of solid food, the microbiome starts to assume a more stable structure similar to that of adults [77].

During adult life, the microbiome remains stable over time but is changed due to external perturbations [23]. As the ageing process progresses, the relative stability of the microbiome tends to reduce[105], and in unhealthy aging associated with frailty, diversity is observed to be reduced and different from young adults [40,106]. The inter-individual variation of microbiome largely persists among elderly individuals [107]. This variability is also influenced by cohort, geographical location, and diet [40,105].

Some studies have made claims that the mycobiome is assembled in the gut of the neonate before birth, based on an analysis of the first passed meconium, formed prior to birth [108], however, the claims of *in utero* development of the mycobiome is debatable and studies have reported post-natal colonization by gut mycobiome [109,110]. During infancy, the gut mainly harbours fungi belonging to *Saccharomycetales*, and *Malasseziales*, followed by reduction in *Malasseziales* at 5 months [111]. Due to dietary changes from breastmilk to solid food, at 1-2 years of age, a rise in *Saccharomyces cerevisiae* is observed, along with species belonging to genera *Ascomycota*, *Cystofilobasidium*, and *Monographella* [112], indicating the importance of diet in the evolution of the gut mycobiome. The adult mycobiome is dominated by fungi belonging to the phyla *Ascomycota*, *Basidiomycota*, and *Zygomycota*, with dominant genera like *Candida*, *Saccharomyces*, and *Cladosporium* [42,87] . In old age, fungal

genera like *Aspergillus*, *Candida*, *Penicillium*, and *Saccharomyces* have been reported to dominate the gut [113,114].

1.3.3 Host genetics

The genotype of the host is suspected to influence the gut microbiome to some extent. Although the whole microbiome is not heritable, studies have shown evidence of lateral transfer for some members of the microbiome, in previously published twin studies [115–117]. However, another study involving greater number of individuals from multiple ancestral origins reported insignificant association between microbiome compositions and genetic ancestry as well as host genetics, suggesting a minor impact on their microbiota composition [118]. The study also emphasized on the role of shared environment in shaping the microbiome of the individuals [118]. Studies have also shown association of microbes to SNPs (Single Nucleotide Polymorphisms) of some marker genes of the host, and their associated phenotypes [116,119–121].

1.3.4 Medication

Usage of medicines have been shown to affect the gut microbiota composition and function in the human host [10,122–129]. A study previously reported that 19 out of 41 medications consumed by individuals were associated with changes in the microbiome[122]. Proton pump inhibitors, laxatives, metformin, and antibiotics showed strongest associations [122]. The interaction of drugs and gut microbes is bi-directional, i.e., while the drugs effect the composition of the microbiome, the microbiome can also influence the drug metabolism, thereby influencing their bioavailability [123].

Impact of antibiotics is well characterised and is often associated with reduced microbiome diversity during the treatment period [124]. After antibiotic use, it has been observed that a healthy microbiome recovers, but the extent of recovery depends on

the microbiome composition prior to the treatment [24,130]. Although antibiotic treatment is required to eliminate pathogens, the risk of accumulation of antibiotic resistance by the microbiome has been recognised as a potential issue [125].

Among the non-antibiotic drugs, effect of PPIs (Proton Pump Inhibitors) on microbiome are associated with decreased bacterial richness, and other alterations in microbiome like increased abundance of oral microbiota in stool [128,129]. A study suggested that the use of PPIs can increase the risk of growth of potential pathogens in the gut, and has shown that microbiome alteration by PPI is more profound than that caused by the antibiotics[128].

Metformin, a commonly used drug for treatment of type 2 diabetes, has also been linked to alterations in the gut microbiota[122,126,127]. The changes induced by metformin can be rapid, and metformin can alter the gut bacteria composition [126], which thereby lead to a better glucose tolerance in the individuals. While no clear mechanism was proposed, the author attributed the betterment of glucose tolerance to increased levels of butyrate and propionate, and metal binding proteins produced by the gut microbiota.

Although the effect of medication on the gut mycobiome is not extensively studied, there are indications that treatment with anti-fungal drugs reduce the fungal diversity in the gut, and in certain cases, increase in the abundance of pathogenic bacteria, leading to exacerbation of colitis [131,132]. Furthermore, the mycobiome is known to interact with the gut bacteria, and interdependencies have been suggested [29]. Due to this, any influence on the gut bacteria due to medications is also expected to influence the mycobiome.

1.3.5 Other extrinsic factors

Daily lifestyle of individuals is known to affect the human physiology. Such lifestyle factors have also been linked to changes in the gut microbiota. It has been shown that modern lifestyle elements, especially at an early age, influence the microbiota composition [133]. A recently published study showed that the composition of microbiota differs between populations leading a rural life and an urban life [134]. The authors also showed that the individuals living in non-industrialized environments are more similar than individuals living in industrialized setting, irrespective of their ethnicities [134]. Urbanization is also believed to impact the fungal community in the gut [29,86]. The association of gut mycobiome with urbanization related disorders like IBD, obesity, further strengthens the claims [29]. Furthermore, people with urban lifestyle show an increased prevalence of *S. cerevisiae*, and decreased prevalence of the bacteria producing short-chain fatty acids (SCFAs) in the gut [44,86,135] , but in rural settling, and minority ethnic populations, there exists multiple beneficial fungal species in the gut [86]. Two notable species reported by the authors were *Aspergillus oryzae* (a probiotic), and *Candida dubliniensis*, which was found to be positively correlated with serum HDL, and negatively with fasting glucose [86].

Body mass index (BMI) has also been linked with the gut microbiota, especially with obesity, where decreased taxonomic and functional diversity were observed [136–138]. On the contrary, few studies and meta-analysis have shown that the link between microbiota and obesity is significant, but marginal [14,139,140]. Differences in microbiota with exercise has also been reported [141]. The study showed that an increase in diversity is associated with exercise, association between protein consumption and microbial abundance, and effect of BMI among individuals doing exercise.

Studies have shown that stool consistency co-vary with the microbiome composition. Measured using Bristol stool scale, it was the top feature which co-varied with the faecal microbiota in one of the cohorts [14]. Other studies also reported strong association between stool consistency and microbiome composition [142–145]. It was observed that stool consistency and microbiota association was not dependent on water activity [146], but mainly driven by transit time [147]. Stool consistency and transit times are important for gastroenterological conditions like IBS, where it is one of the main factors considered for classification of patients into clinical subtypes. In a previous study by Tap *et. al.*, the authors observed that the microbiome signatures associated with IBS symptom severity were also associated with stool consistency and transit time [6].

1.4 Metabolic effect of the gut microbiota

The nutrition procured through food, and metabolites released by the host can be transformed by the members of gut microbiota, producing bioactive molecules with important functions in the maintenance of the host health. The alterations in gut microbiome composition leads to perturbations in the repertoire of the microbe derived metabolites, which directly or indirectly affects the host physiology. Here I have summarized and briefly discussed bacterial metabolites, focusing majorly on the ones which are frequently observed to be altered in IBS (Table 4).

Metabolite	Function	Alterations
Short chain fatty acids (SCFA)	End product of fermentation of non-digestible carbohydrates; Act as energy source for organs like colon, and liver [148].	Reduced in patients with IBD [150], and IBS (especially IBS-C)[35]

	Maintains the epithelial integrity and improves the gut barrier function [148]. Modulates gut motility via serotonergic pathways [149].	
Bile Acids (BA)	Essential for dietary fat, and fat-soluble vitamins absorption. Also functions as signalling molecules for endocrine and paracrine activities [151]. Biotransformation of primary BAs required for BA homeostasis and reabsorption. Excess bile acids in colon can cause gut irritation and diarrhoea [152,153].	Elevated levels of primary BAs observed in IBS [35]. Reduced levels of secondary BAs observed in IBS-D [154].
Hydrogen	Hydrogen is a by-product of anaerobic fermentation, and is normally produced in the gut, and excess production causes flatulence and gut discomfort [155(p2)].	Greater hydrogen production in IBS [156].
Methane	Produced mostly by hydrogenotrophic methanogens, by reducing CO_2 to CH_4 [157]. It acts as a neuromuscular transmitter, augmenting contractile activity of ileal circular muscle, and increases the amplitude of contraction, slowing down peristalsis [158,159]. Excess of methane is directly associated with slowing of gut motility [160].	Higher methane production observed in IBS patients [161].

Hydrogen sulphide	Produced by sulphate reducing bacteria. It is a signalling molecule, involved in inflammation and gastrointestinal mucosal repair [162].	Increased levels associated with colorectal cancer, ulcerative colitis [163,164], and diarrhoea [165]. Also observed in increased levels in IBS-D [166,167].
Tryptophan	An essential amino acid, which is a precursor of serotonin, and melatonin [168]. Tryptophan derived aryl-hydrocarbon receptor agonists, and indole derivatives can help in maintaining gut equilibrium [169].	Disrupted tryptophan metabolism, and lower levels of tryptophan, and its metabolites in IBS [170].
Glycine	Anti-inflammatory, immunomodulatory, and cytoprotective molecule [171]; acts as a neurotransmitter [172].	In animal models, linked with alleviation of diarrhoea and inflammatory changes in the colon [173].
Histamine	Molecular mediator involved in visceral hypersensitivity [96–98].	Faecal microbiota from IBS patients with high urinary histamine induced greater production of histamine in germ free mice [99].
Hypoxanthine	Precursor of purine sugar synthesis, to be used as energy source for epithelial cells, and	Increased microbiota mediated

	helps in development of intestinal barrier [174,175].	biotransformation in IBS, lesser availability to the host [35].
Compounds acting as Neurotransmitters	Allow communication across gut-brain axis. 5-HT/Serotonin: A metabolite of Tryptophan, known to modulate gut secretion and motility.	Higher levels in gut is associated with visceral hypersensitivity and abdominal discomfort [177].
	Dopamine: Precursor of adrenaline and nor-adrenaline. Potential to control chronic pain, intestinal inflammation, and psychological disorders [168].	Decreased in IBS [178].
	Gamma-Amino Butyrate (GABA): Inhibitory neurotransmitter; Modulates GI motility and inflammation [176].	Diminished levels of GABA, and dysregulation of GABAergic system in IBS-D [179].

Table 4: Metabolites of microbiome origin which may be relevant to IBS pathophysiology.

The fungal metabolites in the gut also influence the host physiology either directly or indirectly by interacting with the bacteriome [29]. The fungal cell wall components like beta-glucan, chitin, and mannan, can be recognized by pattern recognition receptors (PRR), which can then trigger the host immune system to illicit pro-inflammatory response [180]. Dectin-1 (a member of C-type lectin receptors, CLRs) expressed by the host can recognize beta-glucan of the fungal cell wall, which helps in homeostasis

between the mycobiome and the host immune system. Alteration in dectin-1 mediated pattern recognition due to changes in dectin-1, can promote cytokine release and immune response [180]. Indirectly, the fungi in the gut can also interact with the gut bacteria, thereby influencing the bacterial metabolite levels, and subsequently influencing the host physiology. This metabolite crosstalk is bi-directional. The gut mycobiota can produce anti-microbial peptides, alcohol, quorum-sensing molecules, and other metabolites which can influence bacterial growth, while the gut bacteria can produce fatty acids, lactic acids, and SCFAs, which modulate the fungal community [29].

1.5 Modulation of the microbiome

A common factor observed in the majority of intestinal disorders, both causative and/or secondary effect, is alteration in microbiota composition marked by overgrowth of pathogens, and pathobionts [26]. Therefore, for alleviation of symptoms, or disease management, microbiome restoration therapy via modulation of microbiome is of growing interest. In most cases, the goal is to achieve a microbiome community closer to that recognized as the "healthy Microbiota", even if that definition is context dependent. Multiple strategies exist, involving molecular, and live organismal interventions, which are discussed below.

1.5.1 Antibiotic intervention

One of the molecular interventions to perturb the microbiome for disease amelioration is the use of antibiotics. Due to its untargeted nature, and impact on the gut microbial diversity, the use of antibiotics for restoration of microbiome is discouraged [181,182]. For example, in non-constipated IBS, use of Rifaximin, a well-tolerated antibiotic,

showed significant positive impact on alleviation of IBS symptoms [183,184], but its use transiently impacted the diversity of the microbiota [185].

1.5.2 Prebiotics and dietary intervention

Prebiotics are chemical compounds found in food items, which favour the growth of beneficial microbes [186], which in turn may modulate the microbiome composition towards a microbiota configuration similar to that of a healthy microbiome. Prebiotics, including fibre rich diets and fructans, have shown a greater potential to confer healthy changes in the microbiome, however, the effect of prebiotics on the whole microbiome is marginal [28]. Dietary intervention studies, investigating the changes in microbiome, have shown effectiveness of ready-to-use therapeutic food (RUTF) in malnutrition [187,188], and the Mediterranean diet in reduction of frailty and improvement of health status in elderly people [84]. In IBS, the most popular dietary intervention is application of low FODMAP diets. Multiple trials on application of low FODMAP diets, reviewed by Staudacher and Whelan, suggest that such a strategy is effective in alleviation of IBS symptoms [189]. Also, a diet low in FODMAPs and high in tryptophan has also been suggested for IBS symptom amelioration [190].

1.5.3 Probiotics

Probiotics are live microbes, which are consumed for betterment of gut health [191], and has shown potential benefits in some cases [192]. Apparently, the efficacy of a probiotic is also dependent on existing baseline microbiota composition [193]. However, multiple studies have shown that the effect of probiotics on the microbiota are mild, and engraftment of the probiotic is doubtful [194]. In IBS, both single and multi-strain probiotics have been applied, but, meta-analysis of published reports and trials showed

efficacy in symptomatic relief to be inconclusive and requires further investigation [195,196].

1.5.4 Faecal microbiota transplantation (FMT)

Faecal microbiota transplantation is a method of introducing the microbiome of a healthy donor into a diseased individual to re-structure the altered microbiome towards a healthier status. Faecal microbiota transplantation has been applied in multiple disorders including, but not limited to, Inflammatory Bowel disease (IBD) [197,198], *Clostridioides difficile* infection (CDI) [199], and Irritable bowel syndrome (IBS) [200,201]. Although proven to be safe, and recently a microbiota consortium (RBX2660) being approved by FDA for treatment of CDI [202], the approach is associated with multiple and variable side effects, including adverse effects in some cases. Nausea, bloating, fatigue, abdominal pain, and mild diarrhoea are some of the common side effects of FMT [203,204]. A concerning factor in FMT application is the presence of the plethora of unknown organisms in the donor stool, which might harbour pathobionts, including antibiotic resistant microbes [205], and infectious viruses [206]. In another study, FMT was found to be associated with development of unrelated symptoms, including worsening of IBD symptoms [207], and onset of obesity in lean patients because of the transplanted microbiome from an overweight donor [208]. Such adverse effects further highlight its non-specificity, and risk and safety assessment should be performed [209]. Furthermore, long-term efficacy of FMT therapy should also be considered. A recent systematic analysis of FMT trials on IBS patients showed significant improvement in IBS symptom severity scores (IBS-SSS) and quality of life (IBS-QoL), but these improvements were observed to be not long lasting [210].

1.5.5 Single strain live biotherapeutic products

Functional diversity among strains of a single species indicates possibility of mixed effects on the gut microbiome. The well-known example is of *Escherichia coli*, a well-studied organism, a species which includes strains which are beneficial to the gut, conferring probiotic effects, however certain strains can be pathogenic [26]. Therefore, microbiome intervention with highly characterized strains is advantageous over multi-strain probiotics, or untargeted approaches like FMT. Furthermore, the dosage of the live biotherapeutic product can be defined and controlled and makes safety assessments more robust. A recent randomized control trial in IBS with a specific strain of *Blautia hydrogenotrophica*, MRx1234, showed beneficial effects, by alleviation of IBS symptoms like abdominal discomfort and bowel habit irregularities [211].

In summary, the microbiome modulation strategies have shown promising results, albeit some safety concerns still exist. With the improvement in knowledgebase and technology, strain selection for therapeutic purposes has become more robust and regulated, but experimental screening of the strains with therapeutic potential and the study of their effects on the microbiome and the host health can be both time consuming and expensive. With the growing development of the computational analysis methods, we can now investigate the mechanistic insights of the interactions in the microbiome ecosystem. The methods, discussed in the next section, provide us with the possibility to simulate the gut environment to study the efficacy of an LBP, which can reduce both the time and cost of LBP screening.

1.6. Understanding the biology of the microbiome through bioinformatics

1.6.1 Looking beyond composition; Exploration of functional potential of the microbiome.

Much of the microbiome research have been based on determining compositional differences in microbiome between populations and are mostly estimated using 16S rRNA gene amplicon sequences for bacteria, and 18S rRNA gene, and internal transcribed spacers (ITS-1, and ITS-2) for fungi. Additionally, microbiome composition can also be estimated using metagenomic sequences. Although compositional differences exist between healthy and disease populations, it is not sufficient to explain the underlying biology. Therefore, harnessing the functional potential information of the microbiome and its investigation is imperative to understand the mechanistic insights of the observed compositional differences.

Pipelines like Tax4Fun [212(p4)], and PICRUSt [213] make use of the microbial abundance and genomic information of the microbial species to estimate the functional potential of the microbiome. Such derivation of potential functionality from genomic information bears multiple weaknesses including the assumption that each identified species contains all the genes represented by the corresponding reference genomes, and the validity of taxonomic classification of the microbes. These issues can be largely resolved using whole genome shotgun sequences, or metagenomic analysis, wherein, metagenomic short reads or assembled contigs, are mapped to well annotated protein databases like NCBI Refseq, and UNIREF, to obtain protein/enzyme abundance in metagenomes. Analysis pipelines like HUMAnN 3.0 [214], which map the metagenomic reads to curated genomes, can extract protein and pathway abundances which are biologically meaningful.

Exploring the genetic structure of the microbiome provide insights into potential metabolic mechanisms, but such estimation of functional potential of microbiome can only provide information of presence or absence of genes/proteins in a microbe and does not indicate actual activity. Therefore, additional meta-omics approaches like metatranscriptomics (study of microbial mRNAs) and metaproteomics (analysis of synthesized proteins by the microbiome), have been applied in various studies [35,215–218]. Although, these datasets are valuable in gaining mechanistic insights, sample and data processing of such omics datasets bear challenges, and care must be taken while analysing these data. While a major challenge for metatranscriptomics data is the high degree of variation of mRNA half-life within microbial taxa [219], the major drawback of metaproteomics is its ability to yield numerous false positives due to the use of large databases, and high levels of false discovery rates during metaproteomic data analysis [220].

1.6.2 Microbial metabolomics

Microbial metabolites are the final downstream molecular outputs which interact with the host, thereby, potentially modulating its physiology. Profiling of the metabolome can be achieved by targeted, or untargeted metabolomics, both of which have their own advantages. The former is best suited for quantification of selected metabolites, while the latter provides a global picture of the metabolome. Metabolomics of host samples like serum [221–224], urine [8,79,225,226], and faeces [8,35,79], have been proven useful for biomarker identification and associations in various host conditions. The faecal metabolome is considered a better option as it contains host-gut co-metabolites, which can be useful for investigation of the microbiota derived metabolites [227].

Although metabolomics is a powerful tool for obtaining mechanistic insights of the host-microbiome interactions, the cost associated with metabolome profiling currently renders it unfeasible for large scale exploratory studies, involving thousands of subjects. Therefore, methods have been developed to extrapolate metabolic information based on integration of the metabolome and the microbiome abundance of a comparatively smaller subset of samples. The method MelonnPan [228] integrates paired metabolite and microbiome abundances to build a statistical model, generating a coefficient matrix based on Elastic-net regression, which can be used to extrapolate metabolic abundance of samples which only have information of the microbiome abundance. Another approach, IPCO (Inference of Pathways from Co-variance analysis) [229], infers microbiome functionality based on microbiome-metabolome co-variance (RLQ analysis), and has shown better performance in metabolome extrapolation. A neural network based method, mmvec [230], provides conditional probability of the presence of a metabolite given the presence of a microbe, and may be used to understand microbiome-metabolome interaction, as well as, empirically estimate the presence/absence of a metabolite based on the microbiome composition.

Although such methods can infer the metabolome which may be used to understand metabolic dynamics of the microbiome and identify biomarkers, the major limitation of such methods is the data-driven nature of the algorithms, and inability to deal with the inherent stochasticity within the microbiome and metabolome profiles, which might lead to inferences which are biologically irrelevant.

1.7 Systems Biology: A scope for rational microbiome modulation

Systems biology is a predictive approach towards understanding of events happening in various biological systems. This field of study relies on mathematical models

generated using simulations based on logical assumptions, and in some cases, biological information, which is then used to understand the interaction and growth dynamics of the members of the microbiome. With the development of technology for microbiome analysis, many research groups have proposed multiple strategies to mathematically model the microbial ecosystem in the host [231–233].

1.7.1 Mathematical modelling of the microbiome

Simplistic mathematical approaches include pairwise interaction models based on microbial abundance, Lotka-Voltera equations, or analysis of microbial networks based on correlation of abundances, or association rule mining [234–237].

Microbial network analysis, attempting to decipher inter-microbial associations using a correlation network-based approach, have shown differences in microbial co-abundance structure to be different in different populations [234,238–241]. Such correlations may be observed due to biologically meaningful interactions or may be a technical or data dependent artifact. This network-based approach has been helpful in identifying sub-networks/modules in the microbial community, which may be potential drivers/outcomes of the host physiology [84,242,243].

Although such models show differences across populations, more complex mechanistic models incorporating metabolism, and growth information are needed. An earlier study [232] proposed a potential mechanism (SMETANA), to predict biologically significant microbial crosstalk based on co-occurrence and metabolic pathways of the studied organisms. The authors first identified co-occurring communities, followed by comparison of metabolic repertoires of the participating microbes, and suggested that the major driver for co-occurrence is metabolic cross

feeding between the interacting species. The study provides deeper insights into metabolic dependencies, but the methodology produced many false positives.

Agent-based modelling to generate simulated microbial abundance have also shown promising results. One such method investigated the effect of diet and macronutrient source on microbiota diversity[231] where the bacterial population's response to diet in mice with varied macronutrient ratios and total calorie intake. The models simulated bacterial growth trends with variable macronutrient intakes, where hypothetical bacterial communities were categorized into guilds based on nutrition consumption. With generalized growth equations and rules, the possible solutions for bacterial growth were predicted. Confirmation by experimental data showed that this approach is partially validated and therefore may be used to gain biologically meaningful insights.

The methods, utilizing generalized equations, interaction rules, and optimizing over conditional variables, are limited to simulation space as they mostly ignore real biological contexts [231,236,244–246]. These also lack inclusion of microbiome properties such as the effects of microbial metabolism, and metabolite exchanges between the communities.

1.7.2 Genome scale metabolic models (GENREs): Generation protocols and available tools.

With the development of a knowledge base on bacterial genomes and their corresponding metabolism, genome scale metabolic network reconstructions (GENREs) have come to the forefront of systems biology for modelling bacterial growth, activity, and interactions [247–249]. The core architecture of such metabolic

models consists of a network of metabolic reactions based on genomic content, limits for reaction rates (flux), and an optimization objective variable.

For any genome scale model generation, the primary requirement is to have a sequenced and well annotated genome of the target organism. Using such information, draft metabolic models can be generated using ModelSEED [250], KBase [251], and Pathway Tools [252]. Although these draft models are useful, there is a void due to unknown genes and misannotated proteins, leading to gaps in the metabolic network. Furthermore, precautions should be taken to ensure that the reactions are biologically valid, and the mathematical model represents the organism as closely as possible. For this, large-scale manual intervention is required [253]. With advancement in knowledge base, metabolic reconstructions of a large set of species observed in the gut, with experimentally validated reactions, were recently published as AGORA (assembly of gut organisms through reconstruction and analysis)) [247]. Although, manually curated GENREs are preferrable, generating reconstructions of large number of species (especially the unculturable ones) is limited due to time constraints. Therefore, tools like CarveMe [254], gapseq [255], and DEMETER [256], have been developed, which not only generate biologically valid metabolic models, but also deal with the gap-filling problem. A generalized protocol has also been proposed to generate multi-strain models from a template [257].

The methods discussed above assume that the genomes of the species/strains remain intact in the environment where they are present, but there is a possibility of gain/loss of function in the species of the microbiomes, and genomic composition and structure may vary across the microbiomes. Pipelines like metaGEM can generate metabolic models based on MAGs, which can resolve these issues by capturing the gene contents of each species in different metagenomic samples [258].

1.7.3 Constraint-based modelling of the microbiome.

To understand the dynamics of multi-species metabolic interactions, it is necessary to develop the microbiome community level *in silico* metabolic models followed by constraint-based modelling of the same under defined growth conditions. The basic approach is to compartmentalize the metabolic models of each species, sharing a common environment where metabolite exchanges are allowed, given biologically valid constraints. Tools like MICOM [248], KBase [251], SteadyCom [259], optCOM [260], SMETANA [232], and cFBA [261] make use of the same concept to obtain an integrated *in silico* model representing the microbiome community. Such tools provide insights about the metabolite production, and consumption dynamics, in the simulation space, which can be used to draw inferences capable of identification of biomarkers for diagnostics, and therapeutic interventions.

Metabolic reconstruction comparison, and constraint based modelling including flux balance and flux variance analysis, have provided quantitative insights into bacterial metabolism and metabolic cross-feeding [232,233,248,262]. These approaches have also shown better performance in predicting microbial interactions, and health status, due to the inclusion of biological information [263].

Flux Balance Analysis (FBA) provides an optimized solution and flux distribution of metabolites, assuming the steady state. This helps in understanding the metabolic exchanges by the microbiome community at a given time, but the microbiome community composition, and metabolism, may change with time. To obtain insights into changes in flux over time, dynamic flux balance analysis can be a useful approach [264]. Tools like FLYCOP [265], and dOptCom [266], use this approach to predict growth rates, and changes in metabolite concentrations over time.

The methods discussed above do not consider the spatial aspect of the microbiome. To analyse microbial interaction dynamics spatiotemporally, microbial community members are placed in lattice grids, prior to metabolic simulations. Tool like COMETS [267], and BacArena [268], make use of dynamic FBA and agent based modelling, to capture the spatial and temporal changes in the microbiome.

Overall, GENREs, and their application in community level modelling using constraint based, dynamic FBA, and agent-based simulations, can provide mechanistic insights into microbial cross-feeding, and metabolite exchanges. The ability to computationally derive metabolome profiles, and potentially perform interventional studies *in silico*, render these methods more cost effective, and time efficient. The methods, although *in silico*, closely represent the microbiome and their metabolism, and can be used to investigate differences between healthy and diseased populations, thereby, augmenting microbiome and metabolome based biomarker identification.

1.8. Perspective: Method integration and possible applications of microbiome perturbation studies

Individually, each microbiome analysis strategy provides insights into possible events happening in the microbial ecosystem of hosts, but these methods alone cannot completely explain the variance across microbiomes. Much of the variance associated with the microbiota is stochastic in nature [269] and therefore, cannot be attributed to life events or disease. Therefore, there is scope to better understand the biology and the host microbiome by integrating all such methods to study microbial growth and their interaction dynamics.

Using co-occurrence patterns of microbes and investigating the commonality between functional pathways present in them, spurious data dependent correlations can be

filtered out, and biologically meaningful covarying pairs can be identified. Also, systems biology based analysis can provide a general idea about the biochemical events within and between the microbes [270]. Such biologically meaningful contexts can be fed into machine learning algorithms, which may help in betterment of the performance of the trained models.

The microbiota composition and function, along with metabolome profiles and information from systems biology based mechanistic models may be utilized for microbiota perturbation studies. Inferences from integrated analysis have the potential to determine the contribution of the microbiota to the differences in diseases and disorders. These functionally important differences may be used to provide predictors of metabolomic or immune changes. In addition, there is the potential to manipulate the microbiome to modulate the disease. With regards to management of inflammatory diseases, modulation of the microbiome is particularly worthy of consideration given the acknowledged need for novel therapeutic approaches and the recognition of the importance of microbial interactions with the human host. The use of faecal microbial transplantation (FMT) has shown promise in the treatment of IBD and other conditions [197–201]. Although the efficacy of FMT in IBD is mixed, a positive impact has been reported in various studies [271–274]. Furthermore, novel cultured gut commensals employed as Live Biotherapeutic Products (LBPs) are promising therapeutic candidates for such diseases [211].

Besides helping understand the mechanism of action of therapeutic microbes in modulating the microbiome and the host physiology, efficacy of administered drugs can also be estimated using integrated analysis approaches. It has been shown that the microbiota may have an effect on the bioavailability and biotransformation of orally administered drugs [275]. It has also been reported that efficacy of checkpoint inhibitors

in case of cancer, and potency of vaccines, are also affected by the baseline microbiota of the subjects prior to administration [276–281]. In the case of faecal microbiota transplantation and LBPs, the pre-established microbiota can augment or reduce the efficacy of treatment [282,283]. Such modulations can also be predicted by applying bioinformatics and data analysis methods on multi-omics datasets.

Basic statistical comparisons across populations coupled with machine learning and systems biology, with integrated biological contexts may provide details which can explain variances due to the microbiome. In this thesis, I propose to implement conventional as well as new analytical approaches, to enhance application of predictive biology, especially in the field of personalized therapeutics, and diagnostics.

1.9. Thesis scope and defined objectives

This thesis utilises and expands on state-of-the-art bioinformatic methods to explore the properties of, and extent of dysbiosis of the microbiome and metabolome in irritable bowel syndrome (IBS). IBS is a chronic, debilitating, gastrointestinal functional disorder, and is heterogenous and marked by symptoms like abdominal pain, distension, irregularities in bowel habits, and thereby, social disablement. The heterogeneity in IBS arises due to inconsistencies in identification of pathological factors, and in relationship between aetiology and symptoms [284] . It is estimated to affect 11.2% (95% CI: 9.8% - 12.8%) of the global population [285] , with higher occurrence in females [286], but largely varies among communities across geographies [287]. Clinically, the IBS patients are classified based on symptoms, bowel frequency, and stool consistency. Based on the symptoms, the subtypes can be classified as IBS-C (Constipation predominant), IBS-D (diarrhoea predominant), IBS-M (mixed), and

IBS-U (unclassified) [288]. The aetiology of the disease is mostly unclear, but it is often attributed as a disorder of the gut-brain axis, and its progression is associated with alterations in the microbiome [6,7,34]. With further advancement in technologies, integrated multi-omics studies have further strengthened the hypothesis of microbial involvement in IBS progression [8,35]. Additional aspects of IBS are discussed further in the forthcoming chapters.

Using an internal cohort of 80 IBS patients, and age, gender, ethnicity, and geography matched 65 Healthy controls, the thesis attempts to address the microbial associations with the disorder, along with other factors like metabolome, and diet. I also investigated the changes in other microbiome components, viz., mycobiome, and virome, in IBS and controls. I report inter-microbial associations based on co-occurrence, and potential metabolic interactions, and investigated the effect of microbial perturbations on microbiome structure. I also report biologically important microbes and their properties which can be used to cause directed perturbation in microbiome and may lead to restoration of microbiome structure closer to that of a defined healthy microbiota. Furthermore, I have developed models from the analysis, which may help in screening of novel LBP candidates, and transformative therapeutics for microbiome-related disorders like IBS.

The defined objectives of the thesis are as follows:

1. Establishment of the extent of bacterial associations in IBS and Control subjects using experimental and omics datasets.

2. Identification of important omics datasets for prediction of health status.

3. Establishment of Steady state flux models of the microbiota in a population of IBS patients and control samples.

4. Prediction of perturbations in the microbiota or metabolome based upon the addition or removal of specific taxa with association to the disorder.

5. Identification of the microbes predicted to have the ability to modulate the microbial community.

1.10 References

1. Fan Y, Pedersen O. Gut microbiota in human metabolic health and disease. Nat Rev Microbiol 2021;19:55–71.

2. Sender R, Fuchs S, Milo R. Revised Estimates for the Number of Human and Bacteria Cells in the Body. PLoS Biol 2016;14:e1002533.

3. Turnbaugh PJ, Ley RE, Hamady M, et al. The human microbiome project. Nature 2007;449:804–810.

4. Ni J, Wu GD, Albenberg L, et al. Gut microbiota and IBD: causation or correlation? Nat Rev Gastroenterol Hepatol 2017;14:573–584.

5. Lloyd-Price J, Arze C, Ananthakrishnan AN, et al. Multi-omics of the gut microbial ecosystem in inflammatory bowel diseases. Nature 2019;569:655–662.

6. Tap J, Derrien M, Törnblom H, et al. Identification of an Intestinal Microbiota Signature Associated With Severity of Irritable Bowel Syndrome. Gastroenterology 2017;152:111-123.e8.

7. Jeffery IB, O'Toole PW, Öhman L, et al. An irritable bowel syndrome subtype defined by species-specific alterations in faecal microbiota. Gut 2012;61:997–1006.

8. Jeffery IB, Das A, O'Herlihy E, et al. Differences in Fecal Microbiomes and Metabolomes of People With vs Without Irritable Bowel Syndrome and Bile Acid Malabsorption. Gastroenterology 2020;158:1016-1028.e8.

9. O'Hara AM, Shanahan F. The gut flora as a forgotten organ. EMBO Rep 2006;7:688–693.

10. Minalyan A, Gabrielyan L, Scott D, et al. The Gastric and Intestinal Microbiome: Role of Proton Pump Inhibitors. Curr Gastroenterol Rep 2017;19:42.

11. Power SE, O'Toole PW, Stanton C, et al. Intestinal microbiota, diet and health. Br J Nutr 2014;111:387–402.

12. McDonald D, Hyde E, Debelius JW, et al. American Gut: an Open Platform for Citizen Science Microbiome Research. mSystems 2018;3.

13. Pasolli E, Asnicar F, Manara S, et al. Extensive Unexplored Human Microbiome Diversity Revealed by Over 150,000 Genomes from Metagenomes Spanning Age, Geography, and Lifestyle. Cell 2019;176:649-662.e20.

14. Falony G, Joossens M, Vieira-Silva S, et al. Population-level analysis of gut microbiome variation. Science 2016;352:560–564.

15. Gupta VK, Paul S, Dutta C. Geography, Ethnicity or Subsistence-Specific Variations in Human Microbiome Composition and Diversity. Front Microbiol 2017;8:1162.

16. Arumugam M, Raes J, Pelletier E, et al. Enterotypes of the human gut microbiome. Nature 2011;473:174–180.

17. Cheng M, Ning K. Stereotypes About Enterotype: the Old and New Ideas. Genomics Proteomics Bioinformatics 2019;17:4–12.

18. Knights D, Ward TL, McKinlay CE, et al. Rethinking "enterotypes." Cell Host Microbe 2014;16:433–437.

19. Jeffery IB, Claesson MJ, O'Toole PW, et al. Categorization of the gut microbiota: enterotypes or gradients? Nat Rev Microbiol 2012;10:591–592.

20. Costea PI, Hildebrand F, Arumugam M, et al. Enterotypes in the landscape of gut microbial community composition. Nat Microbiol 2018;3:8–16.

21. Dominguez-Bello MG, Costello EK, Contreras M, et al. Delivery mode shapes the acquisition and structure of the initial microbiota across multiple body habitats in newborns. Proc Natl Acad Sci U S A 2010;107:11971–11975.

22. Stewart CJ, Ajami NJ, O'Brien JL, et al. Temporal development of the gut microbiome in early childhood from the TEDDY study. Nature 2018;562:583–588.

23. Delgado S, Suárez A, Mayo B. Identification of dominant bacteria in feces and colonic mucosa from healthy Spanish adults by culturing and by 16S rDNA sequence analysis. Dig Dis Sci 2006;51:744–751.

24. Fassarella M, Blaak EE, Penders J, et al. Gut microbiome stability and resilience: elucidating the response to perturbations in order to modulate gut health. Gut 2020.

25. Shetty SA, Hugenholtz F, Lahti L, et al. Intestinal microbiome landscaping: insight in community assemblage and implications for microbial modulation strategies. FEMS Microbiol Rev 2017;41:182–199.

26. Shanahan F, Ghosh TS, O'Toole PW. The Healthy Microbiome—What Is the Definition of a Healthy Gut Microbiome? Gastroenterology 2021;160:483–494.

27. Boulangé CL, Neves AL, Chilloux J, et al. Impact of the gut microbiota on inflammation, obesity, and metabolic disease. Genome Med 2016;8:42.

28. Schmidt TSB, Raes J, Bork P. The Human Gut Microbiome: From Association to Modulation. Cell 2018;172:1198–1215.

29. Zhang F, Aschenbrenner D, Yoo JY, et al. The gut mycobiome in health, disease, and clinical applications in association with the gut bacterial microbiome assembly. The Lancet Microbe 2022:S2666524722002038.

30. Cao Z, Sugimura N, Burgermeister E, et al. The gut virome: A new microbiome component in health and disease. EBioMedicine 2022;81:104113.

31. Qin J, Li R, Raes J, et al. A human gut microbial gene catalogue established by metagenomic sequencing. Nature 2010;464:59–65.

32. Wong SH, Yu J. Gut microbiota in colorectal cancer: mechanisms of action and clinical applications. Nat Rev Gastroenterol Hepatol 2019;16:690–704.

33. Flemer B, Warren RD, Barrett MP, et al. The oral microbiota in colorectal cancer is distinctive and predictive. Gut 2018;67:1454–1463.

34. Jeffery IB, Quigley EMM, Öhman L, et al. The microbiota link to irritable bowel syndrome: an emerging story. Gut Microbes 2012;3:572–576.

35. Mars RAT, Yang Y, Ward T, et al. Longitudinal Multi-omics Reveals Subset-Specific Mechanisms Underlying Irritable Bowel Syndrome. Cell 2020;182:1460-1473.e17.

36. Pittayanon R, Lau JT, Leontiadis GI, et al. Differences in Gut Microbiota in Patients With vs Without Inflammatory Bowel Diseases: A Systematic Review. Gastroenterology 2020;158:930-946.e1.

37. Hopkins MJ, Macfarlane GT. Changes in predominant bacterial populations in human faeces with age and with Clostridium difficile infection. J Med Microbiol 2002;51:448–454.

38. Chang JY, Antonopoulos DA, Kalra A, et al. Decreased diversity of the fecal Microbiome in recurrent Clostridium difficile-associated diarrhea. J Infect Dis 2008;197:435–438.

39. Tang WHW, Kitai T, Hazen SL. Gut Microbiota in Cardiovascular Health and Disease. Circ Res 2017;120:1183–1196.

40. Claesson MJ, Jeffery IB, Conde S, et al. Gut microbiota composition correlates with diet and health in the elderly. Nature 2012;488:178–184.

41. Ghosh TS, Shanahan F, O'Toole PW. The gut microbiome as a modulator of healthy ageing. Nat Rev Gastroenterol Hepatol 2022;19:565–584.

42. Nash AK, Auchtung TA, Wong MC, et al. The gut mycobiome of the Human Microbiome Project healthy cohort. Microbiome 2017;5:153.

43. Huseyin CE, O'Toole PW, Cotter PD, et al. Forgotten fungi-the gut mycobiome in human health and disease. FEMS Microbiol Rev 2017;41:479–511.

44. Chin VK, Yong VC, Chong PP, et al. Mycobiome in the Gut: A Multiperspective Review. Mediators Inflamm 2020;2020:9560684.

45. Gu Y, Zhou G, Qin X, et al. The Potential Role of Gut Mycobiome in Irritable Bowel Syndrome. Front Microbiol 2019;10:1894.

46. Jiang L, Stärkel P, Fan J-G, et al. The gut mycobiome: a novel player in chronic liver diseases. J Gastroenterol 2020.

47. Nagpal R, Neth BJ, Wang S, et al. Gut mycobiome and its interaction with diet, gut bacteria and alzheimer's disease markers in subjects with mild cognitive impairment: A pilot study. EBioMedicine 2020;59:102950.

48. Gao R, Kong C, Li H, et al. Dysbiosis signature of mycobiota in colon polyp and colorectal cancer. Eur J Clin Microbiol Infect Dis 2017;36:2457–2468.

49. Luan C, Xie L, Yang X, et al. Dysbiosis of fungal microbiota in the intestinal mucosa of patients with colorectal adenomas. Sci Rep 2015;5:7980.

50. Chin S-F, Megat Mohd Azlan PIH, Mazlan L, et al. Identification of Schizosaccharomyces pombe in the guts of healthy individuals and patients with colorectal cancer: preliminary evidence from a gut microbiome secretome study. Gut Pathog 2018;10:29.

51. Richard ML, Liguori G, Lamas B, et al. Mucosa-associated microbiota dysbiosis in colitis associated cancer. Gut Microbes 2018;9:131–142.

52. Das A, O'Herlihy E, Shanahan F, et al. The fecal mycobiome in patients with Irritable Bowel Syndrome. Sci Rep 2021;11:124.

53. Botschuijver S, Roeselers G, Levin E, et al. Intestinal Fungal Dysbiosis Is Associated With Visceral Hypersensitivity in Patients With Irritable Bowel Syndrome and Rats. Gastroenterology 2017;153:1026–1039.

54. Sciavilla P, Strati F, Di Paola M, et al. Gut microbiota profiles and characterization of cultivable fungal isolates in IBS patients. Appl Microbiol Biotechnol 2021;105:3277–3288.

55. Hong G, Li Y, Yang M, et al. Gut fungal dysbiosis and altered bacterial-fungal interaction in patients with diarrhea-predominant irritable bowel syndrome: An explorative study. Neurogastroenterol Motil 2020:e13891.

56. Sokol H, Leducq V, Aschard H, et al. Fungal microbiota dysbiosis in IBD. Gut 2017;66:1039–1048.

57. Liguori G, Lamas B, Richard ML, et al. Fungal Dysbiosis in Mucosa-associated Microbiota of Crohn's Disease Patients. J Crohns Colitis 2016;10:296–305.

58. Ott SJ, Kühbacher T, Musfeldt M, et al. Fungi and inflammatory bowel diseases: Alterations of composition and diversity. Scand J Gastroenterol 2008;43:831–841.

59. Chehoud C, Albenberg LG, Judge C, et al. Fungal Signature in the Gut Microbiota of Pediatric Patients With Inflammatory Bowel Disease. Inflamm Bowel Dis 2015;21:1948–1956.

60. Hoarau G, Mukherjee PK, Gower-Rousseau C, et al. Bacteriome and Mycobiome Interactions Underscore Microbial Dysbiosis in Familial Crohn's Disease. mBio 2016;7:e01250-16.

61. Nelson A, Stewart CJ, Kennedy NA, et al. The Impact of NOD2 Genetic Variants on the Gut Mycobiota in Crohn's Disease Patients in Remission and in Individuals Without Gastrointestinal Inflammation. J Crohns Colitis 2021;15:800–812.

62. Limon JJ, Tang J, Li D, et al. Malassezia Is Associated with Crohn's Disease and Exacerbates Colitis in Mouse Models. Cell Host Microbe 2019;25:377-388.e6.

63. Li Q, Wang C, Tang C, et al. Dysbiosis of gut fungal microbiota is associated with mucosal inflammation in Crohn's disease. J Clin Gastroenterol 2014;48:513–523.

64. Leonardi I, Paramsothy S, Doron I, et al. Fungal Trans-kingdom Dynamics Linked to Responsiveness to Fecal Microbiota Transplantation (FMT) Therapy in Ulcerative Colitis. Cell Host Microbe 2020;27:823-829.e3.

65. Cao Y, Wang L, Ke S, et al. Fecal Mycobiota Combined With Host Immune Factors Distinguish Clostridioides difficile Infection From Asymptomatic Carriage. Gastroenterology 2021;160:2328-2339.e6.

66. Zuo T, Wong SH, Cheung CP, et al. Gut fungal dysbiosis correlates with reduced efficacy of fecal microbiota transplantation in Clostridium difficile infection. Nat Commun 2018;9:3663.

67. Malard F, Lavelle A, Battipaglia G, et al. Impact of gut fungal and bacterial communities on the outcome of allogeneic hematopoietic cell transplantation. Mucosal Immunol 2021;14:1127–1132.

68. Marr KA, Seidel K, Slavin MA, et al. Prolonged fluconazole prophylaxis is associated with persistent protection against candidiasis-related death in allogeneic marrow transplant recipients: long-term follow-up of a randomized, placebo-controlled trial. Blood 2000;96:2055–2061.

69. Velden WJFM van der, Netea MG, Haan AFJ de, et al. Role of the mycobiome in human acute graft-versus-host disease. Biol Blood Marrow Transplant 2013;19:329–332.

70. Clooney AG, Sutton TDS, Shkoporov AN, et al. Whole-Virome Analysis Sheds Light on Viral Dark Matter in Inflammatory Bowel Disease. Cell Host Microbe 2019;26:764-778.e5.

71. Zünd M, Ruscheweyh H-J, Field CM, et al. High throughput sequencing provides exact genomic locations of inducible prophages and accurate phage-to-host ratios in gut microbial strains. Microbiome 2021;9:77.

72. Tiamani K, Luo S, Schulz S, et al. The role of virome in the gastrointestinal tract and beyond. FEMS Microbiology Reviews 2022;46:fuac027.

73. Coughlan S, Das A, O'Herlihy E, et al. The gut virome in Irritable Bowel Syndrome differs from that of controls. Gut Microbes 2021;13:1–15.

74. Norman JM, Handley SA, Baldridge MT, et al. Disease-Specific Alterations in the Enteric Virome in Inflammatory Bowel Disease. Cell 2015;160:447–460.

75. Zuo T, Lu X-J, Zhang Y, et al. Gut mucosal virome alterations in ulcerative colitis. Gut 2019;68:1169–1179.

76. Zuo T, Wong SH, Lam K, et al. Bacteriophage transfer during faecal microbiota transplantation in *Clostridium difficile* infection is associated with treatment outcome. Gut 2017:gutjnl-2017-313952.

77. Yatsunenko T, Rey FE, Manary MJ, et al. Human gut microbiome viewed across age and geography. Nature 2012;486:222–227.

78. Drescher LS, Thiele S, Mensink GBM. A new index to measure healthy food diversity better reflects a healthy diet than traditional measures. J Nutr 2007;137:647–651.

79. De Filippis F, Pellegrini N, Vannini L, et al. High-level adherence to a Mediterranean diet beneficially impacts the gut microbiota and associated metabolome. Gut 2016;65:1812–1821.

80. Leeming ER, Johnson AJ, Spector TD, et al. Effect of Diet on the Gut Microbiota: Rethinking Intervention Duration. Nutrients 2019;11.

81. David LA, Materna AC, Friedman J, et al. Host lifestyle affects human microbiota on daily timescales. Genome Biol 2014;15:R89.

82. Wu GD, Chen J, Hoffmann C, et al. Linking long-term dietary patterns with gut microbial enterotypes. Science 2011;334:105–108.

83. Johnson AJ, Vangay P, Al-Ghalith GA, et al. Daily Sampling Reveals Personalized Diet-Microbiome Associations in Humans. Cell Host Microbe 2019;25:789-802.e5.

84. Ghosh TS, Rampelli S, Jeffery IB, et al. Mediterranean diet intervention alters the gut microbiome in older people reducing frailty and improving health status: the NU-AGE 1-year dietary intervention across five European countries. Gut 2020;69:1218–1228.

85. Desai MS, Seekatz AM, Koropatkin NM, et al. A Dietary Fiber-Deprived Gut Microbiota Degrades the Colonic Mucus Barrier and Enhances Pathogen Susceptibility. Cell 2016;167:1339-1353.e21.

86. Sun Y, Zuo T, Cheung CP, et al. Population-Level Configurations of Gut Mycobiome Across 6 Ethnicities in Urban and Rural China. Gastroenterology 2021;160:272-286.e11.

87. Hallen-Adams HE, Kachman SD, Kim J, et al. Fungi inhabiting the healthy human gastrointestinal tract: a diverse and dynamic community. Fungal Ecology 2015;15:9–17.

88. Lam S, Zuo T, Ho M, et al. Review article: fungal alterations in inflammatory bowel diseases. Aliment Pharmacol Ther 2019;50:1159–1171.

89. Hoffmann C, Dollive S, Grunberg S, et al. Archaea and fungi of the human gut microbiome: correlations with diet and bacterial residents. PLoS ONE 2013;8:e66019.

90. Cozma-Petruț A, Loghin F, Miere D, et al. Diet in irritable bowel syndrome: What to recommend, not what to forbid to patients! World J Gastroenterol 2017;23:3771–3783.

91. Spiller R. Impact of Diet on Symptoms of the Irritable Bowel Syndrome. Nutrients 2021;13:575.

92. Shepherd SJ, Parker FC, Muir JG, et al. Dietary triggers of abdominal symptoms in patients with irritable bowel syndrome: randomized placebo-controlled evidence. Clin Gastroenterol Hepatol 2008;6:765–771.

93. Murray K, Wilkinson-Smith V, Hoad C, et al. Differential effects of FODMAPs (fermentable oligo-, di-, mono-saccharides and polyols) on small and large intestinal contents in healthy subjects shown by MRI. Am J Gastroenterol 2014;109:110–119.

94. Murray KA, Lam C, Rehman S, et al. Corticotropin-releasing factor increases ascending colon volume after a fructose test meal in healthy humans: a randomized controlled trial. Am J Clin Nutr 2016;103:1318–1326.

95. Major G, Pritchard S, Murray K, et al. Colon Hypersensitivity to Distension, Rather Than Excessive Gas Production, Produces Carbohydrate-Related Symptoms in Individuals With Irritable Bowel Syndrome. Gastroenterology 2017;152:124-133.e2.

96. Boeckxstaens GE, Wouters MM. Neuroimmune factors in functional gastrointestinal disorders: A focus on irritable bowel syndrome. Neurogastroenterol Motil 2017;29.

97. Olivera A, Beaven MA, Metcalfe DD. Mast cells signal their importance in health and disease. J Allergy Clin Immunol 2018;142:381–393.

98. Cenac N, Andrews CN, Holzhausen M, et al. Role for protease activity in visceral pain in irritable bowel syndrome. J Clin Invest 2007;117:636–647.

99. De Palma G, Shimbori C, Reed DE, et al. Histamine production by the gut microbiota induces visceral hyperalgesia through histamine 4 receptor signaling in mice. Sci Transl Med 2022;14:eabj1895.

100. McIntosh K, Reed DE, Schneider T, et al. FODMAPs alter symptoms and the metabolome of patients with IBS: a randomised controlled trial. Gut 2017;66:1241–1251.

101. Stinson LF, Boyce MC, Payne MS, et al. The Not-so-Sterile Womb: Evidence That the Human Fetus Is Exposed to Bacteria Prior to Birth. Front Microbiol 2019;10:1124.

102. Perez-Muñoz ME, Arrieta M-C, Ramer-Tait AE, et al. A critical assessment of the "sterile womb" and "in utero colonization" hypotheses: implications for research on the pioneer infant microbiome. Microbiome 2017;5:48.

103. Neu J, Rushing J. Cesarean versus vaginal delivery: long-term infant outcomes and the hygiene hypothesis. Clin Perinatol 2011;38:321–331.

104. Mändar R, Mikelsaar M. Transmission of mother's microflora to the newborn at birth. Biol Neonate 1996;69:30–35.

105. Tiihonen K, Ouwehand AC, Rautonen N. Human intestinal microbiota and healthy ageing. Ageing Res Rev 2010;9:107–116.

106. Biagi E, Nylund L, Candela M, et al. Through ageing, and beyond: gut microbiota and inflammatory status in seniors and centenarians. PLoS One 2010;5:e10667.

107. Claesson MJ, Cusack S, O'Sullivan O, et al. Composition, variability, and temporal stability of the intestinal microbiota of the elderly. Proc Natl Acad Sci U S A 2011;108 Suppl 1:4586–4591.

108. Willis KA, Purvis JH, Myers ED, et al. Fungi form interkingdom microbial communities in the primordial human gut that develop with gestational age. FASEB J 2019;33:12825–12837.

109. Blaser MJ, Devkota S, McCoy KD, et al. Lessons learned from the prenatal microbiome controversy. Microbiome 2021;9:8.

110. Stewart CJ, Nelson A, Scribbins D, et al. Bacterial and fungal viability in the preterm gut: NEC and sepsis. Arch Dis Child Fetal Neonatal Ed 2013;98:F298-303.

111. Fujimura KE, Sitarik AR, Havstad S, et al. Neonatal gut microbiota associates with childhood multisensitized atopy and T cell differentiation. Nat Med 2016;22:1187–1191.

112. Schei K, Avershina E, Øien T, et al. Early gut mycobiota and mother-offspring transfer. Microbiome 2017;5:107.

113. Wu L, Zeng T, Deligios M, et al. Age-Related Variation of Bacterial and Fungal Communities in Different Body Habitats across the Young, Elderly, and Centenarians in Sardinia Marco ML, ed. mSphere 2020;5:e00558-19.

114. Ahmad HF, Castro Mejia JL, Krych L, et al. *Gut Mycobiome Dysbiosis is Linked to Hypertriglyceridemia Among Home Dwelling Elderly Danes*. Microbiology; 2020. Available at: http://biorxiv.org/lookup/doi/10.1101/2020.04.16.044693 [Accessed December 1, 2022].

115. Goodrich JK, Waters JL, Poole AC, et al. Human genetics shape the gut microbiome. Cell 2014;159:789–799.

116. Goodrich JK, Davenport ER, Beaumont M, et al. Genetic Determinants of the Gut Microbiome in UK Twins. Cell Host Microbe 2016;19:731–743.

117. Hansen EE, Lozupone CA, Rey FE, et al. Pan-genome of the dominant human gut-associated archaeon, Methanobrevibacter smithii, studied in twins. Proc Natl Acad Sci U S A 2011;108 Suppl 1:4599–4606.

118. Rothschild D, Weissbrod O, Barkan E, et al. Environment dominates over host genetics in shaping human gut microbiota. Nature 2018;555:210–215.

119. Lim MY, You HJ, Yoon HS, et al. The effect of heritability and host genetics on the gut microbiota and metabolic syndrome. Gut 2017;66:1031–1038.

120. Le Roy CI, Beaumont M, Jackson MA, et al. Heritable components of the human fecal microbiome are associated with visceral fat. Gut Microbes 2018;9:61–67.

121. Kolde R, Franzosa EA, Rahnavard G, et al. Host genetic variation and its microbiome interactions within the Human Microbiome Project. Genome Med 2018;10:6.

122. Vich Vila A, Collij V, Sanna S, et al. Impact of commonly used drugs on the composition and metabolic function of the gut microbiota. Nat Commun 2020;11:362.

123. Weersma RK, Zhernakova A, Fu J. Interaction between drugs and the gut microbiome. Gut 2020;69:1510–1519.

124. Schwartz DJ, Langdon AE, Dantas G. Understanding the impact of antibiotic perturbation on the human microbiome. Genome Med 2020;12:82.

125. Francino MP. Antibiotics and the Human Gut Microbiome: Dysbioses and Accumulation of Resistances. Front Microbiol 2015;6:1543.

126. Wu H, Esteve E, Tremaroli V, et al. Metformin alters the gut microbiome of individuals with treatment-naive type 2 diabetes, contributing to the therapeutic effects of the drug. Nat Med 2017;23:850–858.

127. Mardinoglu A, Boren J, Smith U. Confounding Effects of Metformin on the Human Gut Microbiome in Type 2 Diabetes. Cell Metab 2016;23:10–12.

128. Imhann F, Bonder MJ, Vich Vila A, et al. Proton pump inhibitors affect the gut microbiome. Gut 2016;65:740–748.

129. Bruno G, Zaccari P, Rocco G, et al. Proton pump inhibitors and dysbiosis: Current knowledge and aspects to be clarified. World J Gastroenterol 2019;25:2706–2719.

130. Raymond F, Ouameur AA, Déraspe M, et al. The initial state of the human gut microbiome determines its reshaping by antibiotics. ISME J 2016;10:707–720.

131. Wheeler ML, Limon JJ, Bar AS, et al. Immunological Consequences of Intestinal Fungal Dysbiosis. Cell Host Microbe 2016;19:865–873.

132. Qiu X, Zhang F, Yang X, et al. Changes in the composition of intestinal fungi and their role in mice with dextran sulfate sodium-induced colitis. Sci Rep 2015;5:10416.

133. O'Toole PW, Claesson MJ. Gut microbiota: Changes throughout the lifespan from infancy to elderly. International Dairy Journal 2010;20:281–291.

134. Keohane DM, Ghosh TS, Jeffery IB, et al. Microbiome and health implications for ethnic minorities after enforced lifestyle changes. Nat Med 2020;26:1089–1095.

135. Li XV, Leonardi I, Putzel GG, et al. Immune regulation by fungal strain diversity in inflammatory bowel disease. Nature 2022;603:672–678.

136. Turnbaugh PJ, Hamady M, Yatsunenko T, et al. A core gut microbiome in obese and lean twins. Nature 2009;457:480–484.

137. Le Chatelier E, Nielsen T, Qin J, et al. Richness of human gut microbiome correlates with metabolic markers. Nature 2013;500:541–546.

138. Costea PI, Coelho LP, Sunagawa S, et al. Subspecies in the global human gut microbiome. Mol Syst Biol 2017;13:960.

139. Sze MA, Schloss PD. Looking for a Signal in the Noise: Revisiting Obesity and the Microbiome. mBio 2016;7:e01018-16.

140. Finucane MM, Sharpton TJ, Laurent TJ, et al. A taxonomic signature of obesity in the microbiome? Getting to the guts of the matter. PLoS One 2014;9:e84689.

141. Clarke SF, Murphy EF, O'Sullivan O, et al. Exercise and associated dietary extremes impact on gut microbial diversity. Gut 2014;63:1913–1920.

142. Vandeputte D, Falony G, Vieira-Silva S, et al. Stool consistency is strongly associated with gut microbiota richness and composition, enterotypes and bacterial growth rates. Gut 2016;65:57–62.

143. Tigchelaar EF, Bonder MJ, Jankipersadsing SA, et al. Gut microbiota composition associated with stool consistency. Gut 2016;65:540–542.

144. Vandeputte D, Kathagen G, D'hoe K, et al. Quantitative microbiome profiling links gut community variation to microbial load. Nature 2017;551:507–511.

145. Zhernakova A, Kurilshikov A, Bonder MJ, et al. Population-based metagenomics analysis reveals markers for gut microbiome composition and diversity. Science 2016;352:565–569.

146. Vandeputte D, Falony G, D'hoe K, et al. Water activity does not shape the microbiota in the human colon. Gut 2017;66:1865–1866.

147. Roager HM, Hansen LBS, Bahl MI, et al. Colonic transit time is related to bacterial metabolism and mucosal turnover in the gut. Nat Microbiol 2016;1:16093.

148. Morrison DJ, Preston T. Formation of short chain fatty acids by the gut microbiota and their impact on human metabolism. Gut Microbes 2016;7:189–200.

149. Reigstad CS, Salmonson CE, Rainey JF, et al. Gut microbes promote colonic serotonin production through an effect of short-chain fatty acids on enterochromaffin cells. FASEB J 2015;29:1395–1403.

150. Martinez KB, Leone V, Chang EB. Microbial metabolites in health and disease: Navigating the unknown in search of function. J Biol Chem 2017;292:8553–8559.

151. Keitel V, Kubitz R, Häussinger D. Endocrine and paracrine role of bile acids. World J Gastroenterol 2008;14:5620–5629.

152. Yang M, Gu Y, Li L, et al. Bile Acid-Gut Microbiota Axis in Inflammatory Bowel Disease: From Bench to Bedside. Nutrients 2021;13:3143.

153. Riemsma R, Al M, Corro Ramos I, et al. SeHCAT [tauroselcholic (selenium-75) acid] for the investigation of bile acid malabsorption and measurement of bile acid pool loss: a systematic review and cost-effectiveness analysis. Health Technol Assess 2013;17:1–236.

154. Dior M, Delagrèverie H, Duboc H, et al. Interplay between bile acid metabolism and microbiota in irritable bowel syndrome. Neurogastroenterol Motil 2016;28:1330–1340.

155. Chassard C, Gaillard-Martinie B, Bernalier-Donadille A. Interaction between H2-producing and non-H2-producing cellulolytic bacteria from the human colon. FEMS Microbiol Lett 2005;242:339–344.

156. Kumar S, Misra A, Ghoshal UC. Patients with irritable bowel syndrome exhale more hydrogen than healthy subjects in fasting state. J Neurogastroenterol Motil 2010;16:299–305.

157. Gaci N, Borrel G, Tottey W, et al. Archaea and the human gut: new beginning of an old story. World J Gastroenterol 2014;20:16062–16078.

158. Triantafyllou K, Chang C, Pimentel M. Methanogens, methane and gastrointestinal motility. J Neurogastroenterol Motil 2014;20:31–40.

159. Bin Waqar SH, Rehan A. Methane and Constipation-predominant Irritable Bowel Syndrome: Entwining Pillars of Emerging Neurogastroenterology. Cureus 2019;11:e4764.

160. Pimentel M, Lin HC, Enayati P, et al. Methane, a gas produced by enteric bacteria, slows intestinal transit and augments small intestinal contractile activity. Am J Physiol Gastrointest Liver Physiol 2006;290:G1089-1095.

161. Kunkel D, Basseri RJ, Makhani MD, et al. Methane on breath testing is associated with constipation: a systematic review and meta-analysis. Dig Dis Sci 2011;56:1612–1618.

162. Guo F-F, Yu T-C, Hong J, et al. Emerging Roles of Hydrogen Sulfide in Inflammatory and Neoplastic Colonic Diseases. Front Physiol 2016;7:156.

163. Levine J, Ellis CJ, Furne JK, et al. Fecal hydrogen sulfide production in ulcerative colitis. Am J Gastroenterol 1998;93:83–87.

164. Pitcher MC, Beatty ER, Cummings JH. The contribution of sulphate reducing bacteria and 5-aminosalicylic acid to faecal sulphide in patients with ulcerative colitis. Gut 2000;46:64–72.

165. Feuerstein JD, Moss AC, Farraye FA. Ulcerative Colitis. Mayo Clin Proc 2019;94:1357–1373.

166. Singer-Englar T, Rezaie A, Gupta K, et al. 182 - Competitive Hydrogen Gas Utilization by Methane- and Hydrogen Sulfide-Producing Microorganisms and Associated Symptoms: Results of a Novel 4-Gas Breath Test Machine. Gastroenterology 2018;154:S-47.

167. Singer-Englar T, Rezaie A, Gupta K, et al. 1089 - A Novel 4-Gas Device for Breath Testing Shows Exhaled H 2 S is Associated with Diarrhea and Abdominal Pain in a Large Scale Prospective Trial. Gastroenterology 2018;154:S-213.

168. Xiao L, Liu Q, Luo M, et al. Gut Microbiota-Derived Metabolites in Irritable Bowel Syndrome. Front Cell Infect Microbiol 2021;11:729346.

169. Zhang J, Zhu S, Ma N, et al. Metabolites of microbiota response to tryptophan and intestinal mucosal immunity: A therapeutic target to control intestinal inflammation. Med Res Rev 2021;41:1061–1088.

170. Berstad A, Raa J, Valeur J. Tryptophan: “essential” for the pathogenesis of irritable bowel syndrome? Scand J Gastroenterol 2014;49:1493–1498.

171. Zhong Z, Wheeler MD, Li X, et al. L-Glycine: a novel antiinflammatory, immunomodulatory, and cytoprotective agent. Curr Opin Clin Nutr Metab Care 2003;6:229–240.

172. López-Corcuera B, Geerlings A, Aragón C. Glycine neurotransmitter transporters: an update. Mol Membr Biol 2001;18:13–20.

173. Tsune I, Ikejima K, Hirose M, et al. Dietary glycine prevents chemical-induced experimental colitis in the rat. Gastroenterology 2003;125:775–785.

174. Lee JS, Wang RX, Alexeev EE, et al. Hypoxanthine is a checkpoint stress metabolite in colonic epithelial energy modulation and barrier function. J Biol Chem 2018;293:6039–6051.

175. Lee JS, Wang RX, Goldberg MS, et al. Microbiota-Sourced Purines Support Wound Healing and Mucous Barrier Function. iScience 2020;23:101226.

176. Auteri M, Zizzo MG, Serio R. GABA and GABA receptors in the gastrointestinal tract: from motility to inflammation. Pharmacol Res 2015;93:11–21.

177. Cremon C, Carini G, Wang B, et al. Intestinal serotonin release, sensory neuron activation, and abdominal pain in irritable bowel syndrome. Am J Gastroenterol 2011;106:1290–1298.

178. Dunlop BW, Nemeroff CB. The role of dopamine in the pathophysiology of depression. Arch Gen Psychiatry 2007;64:327–337.

179. Aggarwal S, Ahuja V, Paul J. Dysregulation of GABAergic Signalling Contributes in the Pathogenesis of Diarrhea-predominant Irritable Bowel Syndrome. J Neurogastroenterol Motil 2018;24:422–430.

180. Iliev ID, Funari VA, Taylor KD, et al. Interactions between commensal fungi and the C-type lectin receptor Dectin-1 influence colitis. Science 2012;336:1314–1317.

181. Becattini S, Taur Y, Pamer EG. Antibiotic-Induced Changes in the Intestinal Microbiota and Disease. Trends in Molecular Medicine 2016;22:458–478.

182. Langdon A, Crook N, Dantas G. The effects of antibiotics on the microbiome throughout development and alternative approaches for therapeutic modulation. Genome Med 2016;8:39.

183. Pimentel M, Lembo A, Chey WD, et al. Rifaximin therapy for patients with irritable bowel syndrome without constipation. N Engl J Med 2011;364:22–32.

184. Ford AC, Harris LA, Lacy BE, et al. Systematic review with meta-analysis: the efficacy of prebiotics, probiotics, synbiotics and antibiotics in irritable bowel syndrome. Aliment Pharmacol Ther 2018;48:1044–1060.

185. Fodor AA, Pimentel M, Chey WD, et al. Rifaximin is associated with modest, transient decreases in multiple taxa in the gut microbiota of patients with diarrhoea-predominant irritable bowel syndrome. Gut Microbes 2019;10:22–33.

186. Gibson GR, Hutkins R, Sanders ME, et al. Expert consensus document: The International Scientific Association for Probiotics and Prebiotics (ISAPP) consensus statement on the definition and scope of prebiotics. Nat Rev Gastroenterol Hepatol 2017;14:491–502.

187. Potani I, Spiegel-Feld C, Brixi G, et al. Ready-to-Use Therapeutic Food (RUTF) Containing Low or No Dairy Compared to Standard RUTF for Children with Severe Acute Malnutrition: A Systematic Review and Meta-Analysis. Adv Nutr 2021;12:1930–1943.

188. Subramanian S, Huq S, Yatsunenko T, et al. Persistent gut microbiota immaturity in malnourished Bangladeshi children. Nature 2014;510:417–421.

189. Staudacher HM, Whelan K. The low FODMAP diet: recent advances in understanding its mechanisms and efficacy in IBS. Gut 2017;66:1517–1527.

190. Bruta K, Vanshika, Bhasin K, et al. The role of serotonin and diet in the prevalence of irritable bowel syndrome: a systematic review. transl med commun 2021;6:1.

191. Hill C, Guarner F, Reid G, et al. The International Scientific Association for Probiotics and Prebiotics consensus statement on the scope and appropriate use of the term probiotic. Nat Rev Gastroenterol Hepatol 2014;11:506–514.

192. Ford AC, Quigley EMM, Lacy BE, et al. Efficacy of prebiotics, probiotics, and synbiotics in irritable bowel syndrome and chronic idiopathic constipation: systematic review and meta-analysis. Am J Gastroenterol 2014;109:1547–1561; quiz 1546, 1562.

193. Buffie CG, Bucci V, Stein RR, et al. Precision microbiome reconstitution restores bile acid mediated resistance to Clostridium difficile. Nature 2015;517:205–208.

194. Kristensen NB, Bryrup T, Allin KH, et al. Alterations in fecal microbiota composition by probiotic supplementation in healthy adults: a systematic review of randomized controlled trials. Genome Med 2016;8:52.

195. McFarland LV, Karakan T, Karatas A. Strain-specific and outcome-specific efficacy of probiotics for the treatment of irritable bowel syndrome: A systematic review and meta-analysis. EClinicalMedicine 2021;41:101154.

196. Li B, Liang L, Deng H, et al. Efficacy and Safety of Probiotics in Irritable Bowel Syndrome: A Systematic Review and Meta-Analysis. Front Pharmacol 2020;11:332.

197. Bennet JD, Brinkman M. Treatment of ulcerative colitis by implantation of normal colonic flora. Lancet 1989;1:164.

198. Anderson JL, Edney RJ, Whelan K. Systematic review: faecal microbiota transplantation in the management of inflammatory bowel disease. Aliment Pharmacol Ther 2012;36:503–516.

199. Bakken JS, Borody T, Brandt LJ, et al. Treating Clostridium difficile infection with fecal microbiota transplantation. Clin Gastroenterol Hepatol 2011;9:1044–1049.

200. Cui J, Lin Z, Tian H, et al. Long-Term Follow-Up Results of Fecal Microbiota Transplantation for Irritable Bowel Syndrome: A Single-Center, Retrospective Study. Front Med (Lausanne) 2021;8:710452.

201. El-Salhy M, Hatlebakk JG, Gilja OH, et al. Efficacy of faecal microbiota transplantation for patients with irritable bowel syndrome in a randomised, double-blind, placebo-controlled study. Gut 2020;69:859–867.

202. Mullard A. FDA advisory committee votes for approval of first microbiome-based drug, despite data problems. Nat Rev Drug Discov 2022;21:786–787.

203. Baxter M, Colville A. Adverse events in faecal microbiota transplant: a review of the literature. J Hosp Infect 2016;92:117–127.

204. Wang S, Xu M, Wang W, et al. Systematic Review: Adverse Events of Fecal Microbiota Transplantation. PLoS One 2016;11:e0161174.

205. DeFilipp Z, Bloom PP, Torres Soto M, et al. Drug-Resistant E. coli Bacteremia Transmitted by Fecal Microbiota Transplant. N Engl J Med 2019;381:2043–2050.

206. Cammarota G, Ianiro G, Tilg H, et al. European consensus conference on faecal microbiota transplantation in clinical practice. Gut 2017;66:569–580.

207. Qazi T, Amaratunga T, Barnes EL, et al. The risk of inflammatory bowel disease flares after fecal microbiota transplantation: Systematic review and meta-analysis. Gut Microbes 2017;8:574–588.

208. Alang N, Kelly CR. Weight gain after fecal microbiota transplantation. Open Forum Infect Dis 2015;2:ofv004.

209. Merrick B, Allen L, Masirah M Zain N, et al. Regulation, risk and safety of Faecal Microbiota Transplant. Infect Prev Pract 2020;2:100069.

210. Samuthpongtorn C, Kantagowit P, Pittayanon R, et al. Fecal microbiota transplantation in irritable bowel syndrome: A meta-analysis of randomized controlled trials. Front Med (Lausanne) 2022;9:1039284.

211. Quigley EMM, Markinson L, Stevenson A, et al. Randomised clinical trial: efficacy and safety of the live biotherapeutic product MRX1234 in patients with irritable bowel syndrome. Aliment Pharmacol Ther 2022:apt.17310.

212. Aßhauer KP, Wemheuer B, Daniel R, et al. Tax4Fun: predicting functional profiles from metagenomic 16S rRNA data. Bioinformatics 2015;31:2882–2884.

213. Langille MGI, Zaneveld J, Caporaso JG, et al. Predictive functional profiling of microbial communities using 16S rRNA marker gene sequences. Nat Biotechnol 2013;31:814–821.

214. Beghini F, McIver LJ, Blanco-Míguez A, et al. Integrating taxonomic, functional, and strain-level profiling of diverse microbial communities with bioBakery 3. eLife 2021;10:e65088.

215. Zhang X, Li L, Butcher J, et al. Advancing functional and translational microbiome research using meta-omics approaches. Microbiome 2019;7:154.

216. Zhang X, Deeke SA, Ning Z, et al. Metaproteomics reveals associations between microbiome and intestinal extracellular vesicle proteins in pediatric inflammatory bowel disease. Nat Commun 2018;9:2873.

217. Schirmer M, Franzosa EA, Lloyd-Price J, et al. Dynamics of metatranscription in the inflammatory bowel disease gut microbiome. Nat Microbiol 2018;3:337–346.

218. Abu-Ali GS, Mehta RS, Lloyd-Price J, et al. Metatranscriptome of human faecal microbial communities in a cohort of adult men. Nat Microbiol 2018;3:356–366.

219. Steiner PA, De Corte D, Geijo J, et al. Highly variable mRNA half-life time within marine bacterial taxa and functional genes. Environ Microbiol 2019;21:3873–3884.

220. Issa Isaac N, Philippe D, Nicholas A, et al. Metaproteomics of the human gut microbiota: Challenges and contributions to other OMICS. Clinical Mass Spectrometry 2019;14:18–30.

221. Liu R, Hong J, Xu X, et al. Gut microbiome and serum metabolome alterations in obesity and after weight-loss intervention. Nat Med 2017;23:859–868.

222. Chen P, Zhou G, Lin J, et al. Serum Biomarkers for Inflammatory Bowel Disease. Front Med (Lausanne) 2020;7:123.

223. Tan B, Qiu Y, Zou X, et al. Metabonomics Identifies Serum Metabolite Markers of Colorectal Cancer. J Proteome Res 2013;12:3000–3009.

224. Nybacka S, Simrén M, Störsrud S, et al. Changes in serum and urinary metabolomic profile after a dietary intervention in patients with irritable bowel syndrome. PLoS One 2021;16:e0257331.

225. Sands BE. Biomarkers of Inflammation in Inflammatory Bowel Disease. Gastroenterology 2015;149:1275-1285.e2.

226. Mallafré-Muro C, Llambrich M, Cumeras R, et al. Comprehensive Volatilome and Metabolome Signatures of Colorectal Cancer in Urine: A Systematic Review and Meta-Analysis. Cancers (Basel) 2021;13:2534.

227. Chen MX, Wang S-Y, Kuo C-H, et al. Metabolome analysis for investigating host-gut microbiota interactions. Journal of the Formosan Medical Association 2019;118:S10–S22.

228. Mallick H, Franzosa EA, Mclver LJ, et al. Predictive metabolomic profiling of microbial communities using amplicon or metagenomic sequences. Nat Commun 2019;10:3136.

229. Das M, Ghosh TS, Jeffery IB. IPCO: Inference of Pathways from Co-variance analysis. BMC Bioinformatics 2020;21:62.

230. Morton JT, Aksenov AA, Nothias LF, et al. Learning representations of microbe-metabolite interactions. Nat Methods 2019;16:1306–1314.

231. Holmes AJ, Chew YV, Colakoglu F, et al. Diet-Microbiome Interactions in Health Are Controlled by Intestinal Nitrogen Source Constraints. Cell Metab 2017;25:140–151.

232. Zelezniak A, Andrejev S, Ponomarova O, et al. Metabolic dependencies drive species co-occurrence in diverse microbial communities. Proc Natl Acad Sci U S A 2015;112:6449–6454.

233. Bauer E, Thiele I. From Network Analysis to Functional Metabolic Modeling of the Human Gut Microbiota. mSystems 2018;3.

234. Faust K, Sathirapongsasuti JF, Izard J, et al. Microbial co-occurrence relationships in the human microbiome. PLoS Comput Biol 2012;8:e1002606.

235. Xiao Y, Angulo MT, Friedman J, et al. Mapping the ecological networks of microbial communities. Nat Commun 2017;8:2042.

236. Kuntal BK, Gadgil C, Mande SS. Web-gLV: A Web Based Platform for Lotka-Volterra Based Modeling and Simulation of Microbial Populations. Front Microbiol 2019;10:288.

237. Tandon D, Haque MM, Mande SS. Inferring Intra-Community Microbial Interaction Patterns from Metagenomic Datasets Using Associative Rule Mining Techniques Ahmed N, ed. PLoS ONE 2016;11:e0154493.

238. Faust K, Lima-Mendez G, Lerat J-S, et al. Cross-biome comparison of microbial association networks. Front Microbiol 2015;6:1200.

239. Yadav D, Ghosh TS, Mande SS. Global investigation of composition and interaction networks in gut microbiomes of individuals belonging to diverse geographies and age-groups. Gut Pathog 2016;8:17.

240. Fernandez M, Riveros JD, Campos M, et al. Microbial “social networks.” BMC Genomics 2015;16 Suppl 11:S6.

241. Ghosh TS, Gupta SS, Bhattacharya T, et al. Gut microbiomes of Indian children of varying nutritional status. PLoS One 2014;9:e95547.

242. Shankar V, Agans R, Holmes B, et al. Do gut microbial communities differ in pediatric IBS and health? Gut Microbes 2013;4:347–352.

243. Raman AS, Gehrig JL, Venkatesh S, et al. A sparse covarying unit that describes healthy and impaired human gut microbiota development. Science 2019;365.

244. Shashkova T, Popenko A, Tyakht A, et al. Agent Based Modeling of Human Gut Microbiome Interactions and Perturbations. PLoS One 2016;11:e0148386.

245. Zeng Q, Wu S, Sukumaran J, et al. Models of microbiome evolution incorporating host and microbial selection. Microbiome 2017;5:127.

246. Lin C, Culver J, Weston B, et al. GutLogo: Agent-based modeling framework to investigate spatial and temporal dynamics in the gut microbiome. PLoS One 2018;13:e0207072.

247. Magnúsdóttir S, Heinken A, Kutt L, et al. Generation of genome-scale metabolic reconstructions for 773 members of the human gut microbiota. Nat Biotechnol 2017;35:81–89.

248. Diener C, Gibbons SM, Resendis-Antonio O. MICOM: Metagenome-Scale Modeling To Infer Metabolic Interactions in the Gut Microbiota. mSystems 2020;5:e00606-19.

249. Machado D, Maistrenko OM, Andrejev S, et al. Polarization of microbial communities between competitive and cooperative metabolism. Nat Ecol Evol 2021;5:195–203.

250. Henry CS, DeJongh M, Best AA, et al. High-throughput generation, optimization and analysis of genome-scale metabolic models. Nat Biotechnol 2010;28:977–982.

251. Arkin AP, Cottingham RW, Henry CS, et al. KBase: The United States Department of Energy Systems Biology Knowledgebase. Nat Biotechnol 2018;36:566–569.

252. Karp PD, Midford PE, Billington R, et al. Pathway Tools version 23.0 update: software for pathway/genome informatics and systems biology. Brief Bioinform 2021;22:109–126.

253. Mendoza SN, Olivier BG, Molenaar D, et al. A systematic assessment of current genome-scale metabolic reconstruction tools. Genome Biol 2019;20:158.

254. Machado D, Andrejev S, Tramontano M, et al. Fast automated reconstruction of genome-scale metabolic models for microbial species and communities. Nucleic Acids Res 2018;46:7542–7553.

255. Zimmermann J, Kaleta C, Waschina S. gapseq: informed prediction of bacterial metabolic pathways and reconstruction of accurate metabolic models. Genome Biol 2021;22:81.

256. Heinken A, Magnúsdóttir S, Fleming RMT, et al. DEMETER: Efficient simultaneous curation of genome-scale reconstructions guided by experimental data and refined gene annotations. Bioinformatics 2021;37:3974–3975.

257. Norsigian CJ, Fang X, Seif Y, et al. A workflow for generating multi-strain genome-scale metabolic models of prokaryotes. Nat Protoc 2020;15:1–14.

258. Zorrilla F, Buric F, Patil KR, et al. metaGEM: reconstruction of genome scale metabolic models directly from metagenomes. Nucleic Acids Research 2021;49:e126–e126.

259. Chan SHJ, Simons MN, Maranas CD. SteadyCom: Predicting microbial abundances while ensuring community stability. PLoS Comput Biol 2017;13:e1005539.

260. Zomorrodi AR, Maranas CD. OptCom: a multi-level optimization framework for the metabolic modeling and analysis of microbial communities. PLoS Comput Biol 2012;8:e1002363.

261. Khandelwal RA, Olivier BG, Röling WFM, et al. Community flux balance analysis for microbial consortia at balanced growth. PLoS One 2013;8:e64567.

262. Magnúsdóttir S, Thiele I. Modeling metabolism of the human gut microbiome. Curr Opin Biotechnol 2018;51:90–96.

263. Heinken A, Ravcheev DA, Baldini F, et al. Systematic assessment of secondary bile acid metabolism in gut microbes reveals distinct metabolic capabilities in inflammatory bowel disease. Microbiome 2019;7:75.

264. Zhuang K, Izallalen M, Mouser P, et al. Genome-scale dynamic modeling of the competition between Rhodoferax and Geobacter in anoxic subsurface environments. ISME J 2011;5:305–316.

265. García-Jiménez B, García JL, Nogales J. FLYCOP: metabolic modeling-based analysis and engineering microbial communities. Bioinformatics 2018;34:i954–i963.

266. Zomorrodi AR, Islam MM, Maranas CD. d-OptCom: Dynamic multi-level and multi-objective metabolic modeling of microbial communities. ACS Synth Biol 2014;3:247–257.

267. Dukovski I, Bajić D, Chacón JM, et al. A metabolic modeling platform for the computation of microbial ecosystems in time and space (COMETS). Nat Protoc 2021;16:5030–5082.

268. Bauer E, Zimmermann J, Baldini F, et al. BacArena: Individual-based metabolic modeling of heterogeneous microbes in complex communities Maranas CD, ed. PLoS Comput Biol 2017;13:e1005544.

269. Erb-Downward JR, Falkowski NR, D'Souza JC, et al. Critical Relevance of Stochastic Effects on Low-Bacterial-Biomass 16S rRNA Gene Analysis Ravel J, ed. mBio 2020;11:e00258-20.

270. Heinken A, Basile A, Thiele I. Advances in constraint-based modelling of microbial communities. Current Opinion in Systems Biology 2021;27:100346.

271. Basso PJ, Câmara NOS, Sales-Campos H. Microbial-Based Therapies in the Treatment of Inflammatory Bowel Disease - An Overview of Human Studies. Front Pharmacol 2018;9:1571.

272. Tan P, Li X, Shen J, et al. Fecal Microbiota Transplantation for the Treatment of Inflammatory Bowel Disease: An Update. Front Pharmacol 2020;11:574533.

273. El Hage Chehade N, Ghoneim S, Shah S, et al. Efficacy of Fecal Microbiota Transplantation in the Treatment of Active Ulcerative Colitis: A Systematic Review and Meta-Analysis of Double-Blind Randomized Controlled Trials. Inflamm Bowel Dis 2023;29:808–817.

274. Xiang L, Ding X, Li Q, et al. Efficacy of faecal microbiota transplantation in Crohn's disease: a new target treatment? Microb Biotechnol 2020;13:760–769.

275. Doestzada M, Vila AV, Zhernakova A, et al. Pharmacomicrobiomics: a novel route towards personalized medicine? Protein Cell 2018;9:432–445.

276. Gong J, Chehrazi-Raffle A, Placencio-Hickok V, et al. The gut microbiome and response to immune checkpoint inhibitors: preclinical and clinical strategies. Clin Transl Med 2019;8:9.

277. Shaikh FY, Gills JJ, Sears CL. Impact of the microbiome on checkpoint inhibitor treatment in patients with non-small cell lung cancer and melanoma. EBioMedicine 2019;48:642–647.

278. Sears CL, Pardoll DM. The intestinal microbiome influences checkpoint blockade. Nat Med 2018;24:254–255.

279. Hagan T, Cortese M, Rouphael N, et al. Antibiotics-Driven Gut Microbiome Perturbation Alters Immunity to Vaccines in Humans. Cell 2019;178:1313-1328.e13.

280. Jamieson AM. Influence of the microbiome on response to vaccination. Hum Vaccin Immunother 2015;11:2329–2331.

281. Vlasova AN, Takanashi S, Miyazaki A, et al. How the gut microbiome regulates host immune responses to viral vaccines. Curr Opin Virol 2019;37:16–25.

282. Lam S, Bai X, Shkoporov AN, et al. Roles of the gut virome and mycobiome in faecal microbiota transplantation. The Lancet Gastroenterology & Hepatology 2022;7:472–484.

283. Danne C, Rolhion N, Sokol H. Recipient factors in faecal microbiota transplantation: one stool does not fit all. Nat Rev Gastroenterol Hepatol 2021;18:503–513.

284. Enck P, Aziz Q, Barbara G, et al. Irritable bowel syndrome. Nat Rev Dis Primers 2016;2:16014.

285. Mearin F, Lacy BE, Chang L, et al. Bowel Disorders. Gastroenterology 2016:S0016-5085(16)00222–5.

286. Lovell RM, Ford AC. Effect of gender on prevalence of irritable bowel syndrome in the community: systematic review and meta-analysis. Am J Gastroenterol 2012;107:991–1000.

287. Oka P, Parr H, Barberio B, et al. Global prevalence of irritable bowel syndrome according to Rome III or IV criteria: a systematic review and meta-analysis. Lancet Gastroenterol Hepatol 2020;5:908–917.

288. Drossman DA, Hasler WL. Rome IV-Functional GI Disorders: Disorders of Gut-Brain Interaction. Gastroenterology 2016;150:1257–1261.

Chapter 2

Differences in Faecal Microbiomes and Metabolomes of People With vs Without Irritable Bowel Syndrome and Bile Acid Malabsorption.

This chapter has been published as

Ian B. Jeffery[#], **Anubhav Das**[#], Eileen O'Herlihy, Simone Coughlan, Katryna Cisek, Michael Moore, Fintan Bradley, Tom Carty, Meenakshi Pradhan, Chinmay Dwibedi, Fergus Shanahan, Paul W. O'Toole*. Differences in Fecal Microbiomes and Metabolomes of People With vs Without Irritable Bowel Syndrome and Bile Acid Malabsorption. Gastroenterology 2020;158:1016-1028.e8.

Authors who share first authorship; * Corresponding author

Author contributions: IBJ, EOH, FS, and PWOT conceived the study idea and design. IBJ, EOH, FS, and PWOT drafted and reviewed the manuscript. IBJ, AD, SC, KC, MP, and CD performed the bioinformatics analysis. MAM, FB, and TC conducted the SeHCAT assay for BAM and reviewed the associated data.

AD Contribution:

Data processing: Metagenomic read processing and quality control checks. Metagenomic compositional and functional estimation. Urine and Faecal metabolomic data processing. Dietary data processing.

Analysis: Hypothesis testing, multivariate analyses. Multi-omic data integration, network generation, and analysis. Bacterial Co-occurrence network generation and analysis. Development of Machine learning models and validation with feature selection and design of the machine learning protocols for IBS and BAM.

Others: Contribution to manuscript by drafting relevant methods, and reporting inferences from data analysis.

Table of Contents

2.1 Abstract 72

2.2 Introduction 74

2.3 Methods 76

2.3.1 Participant recruitment and inclusion criteria. 76

2.3.2 Microbiome profiling and metagenomics. 78

2.3.3 Metabolomics. 78

2.3.4 Bile Acid Malabsorption and ^{75}Seleno-homotaurocholic acid (SeHCAT) assay. 78

2.3.5 Microbiome Sequencing and Bioinformatics. 79

2.4 Results 80

2.4.1 The microbiome differs between IBS and controls but not across IBS clinical subtypes. 80

2.4.2 IBS patients with Bile Acid Malabsorption have altered faecal microbiome. 83

2.4.3 Altered urine and faecal metabolomes in IBS. 85

2.4.4 Integrative analysis of diet, microbiome and metabolome reveals major alterations in IBS. 89

2.5 Discussion 93

2.6 Acknowledgements 98

2.7 References 98

2.8 Supplementary information 107

2.8.1 Supplementary methods 107

2.8.3 Supplementary Figures 121

2.8.4 Supplementary Tables 127

2.1 Abstract

Background & Aims: Irritable bowel syndrome (IBS) is a heterogeneous disorder but diagnoses and determination of subtypes are made based on symptoms. We profiled the faecal microbiomes of patients with and without IBS to identify biomarkers of this disorder.

Methods: We collected faecal and urine samples from 80 patients with IBS (Rome IV criteria; 16–70 years old) and 65 matched individuals without IBS (control individuals), along with anthropometric, medical, and dietary information. Shotgun and 16S ribosomal RNA amplicon sequencing were performed on faeces, whereas urine and faecal metabolites were analysed by gas chromatography and liquid chromatography–mass spectrometry. Co-occurrence networks were generated based on significant Spearman correlations between data. Bile acid malabsorption (BAM) was identified in patients with diarrhoea by retention of radiolabelled Selenium-75 homocholic acid taurine.

Results: Patients with IBS had significant differences in network connections between diet and faecal microbiomes compared with control individuals; these were accompanied by differences in faecal metabolomes. We did not find significant differences in faecal microbiota composition among patients with different IBS symptom subtypes. Faecal metabolome profiles could discriminate patients with IBS from control individuals. Urine metabolomes also differed significantly between patients with IBS and control individuals, but most discriminatory metabolites were related to diet or medications. Faecal metabolomes, but not microbiomes, could distinguish patients with IBS with vs those without BAM.

Conclusions: Despite the heterogeneity of IBS, patients have significant differences in urine and faecal metabolomes and faecal microbiome vs control individuals, independent of symptom-based subtypes of IBS. Faecal metabolome analysis can be used to distinguish patients with IBS with vs those without BAM. These findings might be used for developing microbe-based treatments for these disorders.

2.2 Introduction

Like most chronic noncommunicable disorders, irritable bowel syndrome (IBS) appears to be heterogeneous [1]. It ranges in severity from nuisance bowel disturbance to social disablement, accompanied by marked symptomatic heterogeneity [2]. Although frequently considered a disorder of the brain-gut axis [3,4], it is unclear if IBS begins in the gut, in the brain, or in both. The occurrence of postinfectious IBS [5] indicates that a proportion of cases are initiated in the end organ, albeit with susceptibility risk factors, some of which may be psychosocial. Observations from recent studies further links the brain-gut axis to the pathophysiology of IBS [6,7]. In one of the studies, altered levels of serotonin and 5-hydroxyindoleacetate in plasma were observed for clusters of IBS patients based on microbiome and metabolome profiles[6]. Modulation of mucosal serotonin reuptake transporter expression by gut bacteria in IBS, via mast cell prostaglandin E2 pathway, as shown recently [7], also suggests link between gut health and neurotransmitters. Advances in microbiome science, with emerging evidence for a modifying influence by the microbiota on neurodevelopment and, perhaps, on behaviour, have broadened the concept of the mind/body link to encompass the microbiota-gut-brain axis [8].

However, progress in understanding and treating IBS has been limited by the absence of reliable biomarkers; IBS is still defined by symptoms. Moreover, the current approach to stratification of patients into clinical subtypes based on predominant symptoms (diarrhoea-predominant [IBS-D] or constipation-predominant [IBS-C] IBS) has significant limitations, including failure to inform treatment of patients who alternate between subtypes, sometimes within days [9]. Similar observations were made in a current study, wherein, nearly 1/3rd of the participants showed changes in symptom severity, while in more than 50% of the participants the predominant stool

pattern changed [10]. Pharmaceutical agents designed to tackle symptoms that are polar opposites have the potential for severe unwanted adverse effects if prescribed for a patient that is misclassified [11]. Finally, a proportion of patients incorrectly diagnosed as having IBS-D actually have bile acid (BA) malabsorption (BAM) [12]. Additionally, diarrhoea due to bile acids is not limited to IBS-D. A recent study investigated mucosal biofilms, observed in a sub-set of IBS and ulcerative colitis (UC) patients, were enriched in *Ruminococcus gnavus* and *Escherichia coli*, and accumulated bile acids, which correlated with faecal bile acid excretion, thereby, linked to diarrhoea [13].

Multiple groups have independently reported gut microbiota alterations in patients with IBS [14–17]. Interaction of the microbiome with diet, antibiotics, and enteric infections, all of which may be involved in IBS, is consistent with the hypothesis that microbiome alterations could activate or perpetuate pathophysiologic mechanisms in the syndrome [18,19]. However, robust microbiome signatures or biomarkers that separate patients with IBS from control individuals and that help inform therapies are lacking, although signatures have been reported for IBS severity [17]. Furthermore, most studies have used 16S ribosomal RNA (rRNA) gene profiling and did not analyse bacterial metabolites. To work toward an evidence-based stratification of patients with IBS, we performed a metagenomic study of faecal samples along with metabolomic analyses of urine and faeces in 80 patients with IBS (Rome IV criteria) compared with 65 control individuals. Our findings show that microbiome and metabolomic signatures are evident in IBS, but these are independent of the traditional clinical symptom-based subsets of IBS (IBS-D vs IBS-C, alternating IBS, or mixed IBS [IBS-M]) and also distinct from that of BAM.

2.3 Methods

2.3.1 Participant recruitment and inclusion criteria.

Eighty patients aged 16–70 years with IBS (Rome IV criteria) and 65 control individuals of the same age range, ethnicity, and geographic region were recruited (Table 1). Patients were recruited at the specialty gastroenterology clinics of the Cork University Hospital and were confirmed as having IBS by application of the Rome IV criteria, which include Bristol Stool Scores to identify IBS subtypes (IBS-C, IBS-D, IBS-M). Subtyping included the average Bristol Stool Score reported by the participants for the previous week. Primary exclusion criteria were a significant acute or chronic coexisting illness (cardiovascular, gastrointestinal, endocrine, immunological, metabolic) or malignant disease or any concomitant end-stage organ disease, abdominal surgery other than hernia repair or appendectomy, receiving treatment in an experimental trial in the last 30 days, antibiotics treatment in the previous 6 weeks, taking systemic steroids within the last month, use of an antipsychotic in 3 months before entering the study, or having had a major psychiatric disorder within the past 2 years. Potential participants older than 55 years were excluded if they had not undergone sigmoidoscopy or colonoscopy in the previous 5 years. Study participants who had undergone colonoscopy were not recruited until at least 2 months after the procedure. Individuals were excluded from the study if there had been a change in their diet within the previous 3 months that included recent use of prebiotic or probiotic supplements; if they were following a diet low in fermentable oligosaccharides, disaccharides, monosaccharides, and polyols (FODMAPs); or if they were lactose intolerant or had celiac disease. Control individuals were recruited through spouses and families of the patients and staff of the Cork University hospital, local university, and general public. Control individuals were interviewed for their

complete medical history, and to exclude undiagnosed IBS, they were asked if they had any history of IBS symptoms, and Rome IV criteria were checked.

Table 1: Descriptive statistics of control and IBS subjects studied

	Control (n = 65)	IBS (n = 80)
Age range, years (mean)	19-65 (45)	17-66 (39)
Sex (male/female)	16/49	15/65
BMI Class, n (%)		
Normal	25 (38)	31 (39)
Obese Class I	11 (17)	14 (18)
Obese Class II	3 (5)	5 (6)
Obese Class III	1 (2)	3 (4)
Overweight	21 (32)	22 (28)
Underweight	3 (5)	3 (4)
HADS: Anxiety, n (%)		
Normal (0-10)	59 (91)	58 (72)
Abnormal (11-21)	6 (9)	22 (28)
HADS: Depression, n (%)		
Normal (0-10)	64 (98)	70 (88)
Abnormal (11-21)	1 (2)	10 (12)
Bristol Stool Score, n (%)		
Normal	54 (83)	18 (22)
Constipated	8 (12)	22 (28)
Diarrhoea	3 (5)	40 (50)
IBS subtype, n (%)		
IBS-C		30 (38)
IBS-D	N/A	21 (26)
IBS-M		29 (36)
SeHCAT assayed, n (%)	9 (14)	46 (57)
Dietary group (FFQ), n (%)		
Omnivore	63 (97)	74 (92)
Vegetarian	1 (2)	2 (2)
Pescatarian	1 (2)	1 (1)
Gluten-free	0 (0)	4 (5)
Drinks alcohol, n (%)		
Current	54 (83)	57 (71)
Previous	0 (0)	1 (1)
Never	10 (15)	22 (28)
Smoker, n (%)		
Current	10* (15)	14* (18)
Previous	13 (20)	18 (22)
Never	42 (65)	48 (60)

* 1 subject in each group smoked e-cigarettes
BMI: Body Mass Index, FFQ: Food Frequency Questionnaire, HADS: Hospital Anxiety and Depression Scale, N/A: not applicable.

Clinical subtyping [20] of the patients included IBS-C, IBS-M, and IBS-D. Faecal and urine samples were collected from all participants for microbiome and metabolomic profiling. Further details are available in the Supplementary Materials and Methods.

2.3.2 Microbiome profiling and metagenomics.

Genomic DNA was extracted and amplified from frozen faecal samples (0.25 g) by using the method described by Brown et al. [21] Sequencing was performed on an Illumina (San Diego, CA) MiSeq instrument (2 × 250–base pair paired ends). For shotgun sequencing, 1 μg (concentration > 5 ng/μL) of high-molecular-weight DNA was sequenced on an Illumina HiSeq platform (HiSeq 2500) (2 × 250–base pair paired ends).

2.3.3 Metabolomics.

Urinary high–field asymmetric waveform ion mobility spectrometry (FAIMS) analysis was performed as described [22]. Gas chromatography (GC)–/liquid chromatography (LC)–mass spectrometry (MS) was performed on urine and faecal samples by Metabolon (Potsdam, Germany) (Supplementary Material).

2.3.4 Bile Acid Malabsorption and ^{75}Seleno-homotaurocholic acid (SeHCAT) assay.

^{75}Seleno-homotaurocholic acid taurine (SeHCAT) assay was administered as a single-capsule dose containing less than 0.1 mg of tauroselcholic acid (GE Healthcare, Little Chalfont, UK) and with a radioactivity dose of 370 kBq at the reference date. A baseline whole-body absorption reading using an uncollimated gamma counter (Siemens [Munich, Germany] Ecam camera) was taken for each participant 2–3 hours after capsule administration. A follow-up scan was taken 7 days later, and the

proportion of BA retention was calculated; a value of ≤15% retention indicated mild to severe BAM, with a SeHCAT score of 15%–20% representing a borderline classification, as discussed by Watson et al [23].

2.3.5 Microbiome Sequencing and Bioinformatics.

Standard procedures were used for 16S rRNA gene amplicon sequencing [24] and shotgun analysis [25]. Data were analysed with Human Microbiome Project pipelines, as described in the Supplementary Material.

2.4 Results

2.4.1 The microbiome differs between IBS and controls but not across IBS clinical subtypes.

Descriptive statistics for the study participants are shown in Table 1. Microbiota profiling by 16S rRNA gene amplicon sequencing and principal co-ordinate analysis (PCoA) of the microbiota composition data confirmed that the microbiota of participants with IBS was distinct from that of control individuals (Figure 1A), although with overlap. An informative way to reduce the complexity of metagenomic data while increasing the biological signal is to assemble the reads into co-abundant gene groups (CAGs) [26] representing strain-level variables and commonly referred to as metagenomic species. Analysis of these strain-level variables significantly differentiated patients with IBS from control individuals, as shown in Figure 1B. Machine learning applied to the metagenomic species (CAGs) data set produced a prediction model for IBS with an area under the curve (AUC) of 0.814 (sensitivity, 0.875; specificity, 0.497) based on 136 predictive features (Supplementary Table 1). Taxonomic assignment of the CAGs was sparse, with the majority of features unclassified, but assigned features were broadly consistent with the species-level analysis. Among the differentially abundant OTUs, the OTUs classified as *Ruminococcus gnavus* and *Lachnospiraceae* were significantly more abundant in IBS, whereas *Barnesiella intestinihominis* and *Coprococcus catus* were among taxa significantly less abundant in IBS based on pairwise comparison (see Supplementary Table 2). These alterations are consistent with previous studies [15–17], where at OTU level, the taxa that were significantly differentially abundant belonged to the *Ruminococcaceae*, *Lachnospiraceae*, and *Bacteroidetes* families/phylum. Clinical

subtypes of IBS did not separate in a PCoA of microbiota beta (interindividual) diversity derived from 16S rRNA gene profiling data (Figure 1C).

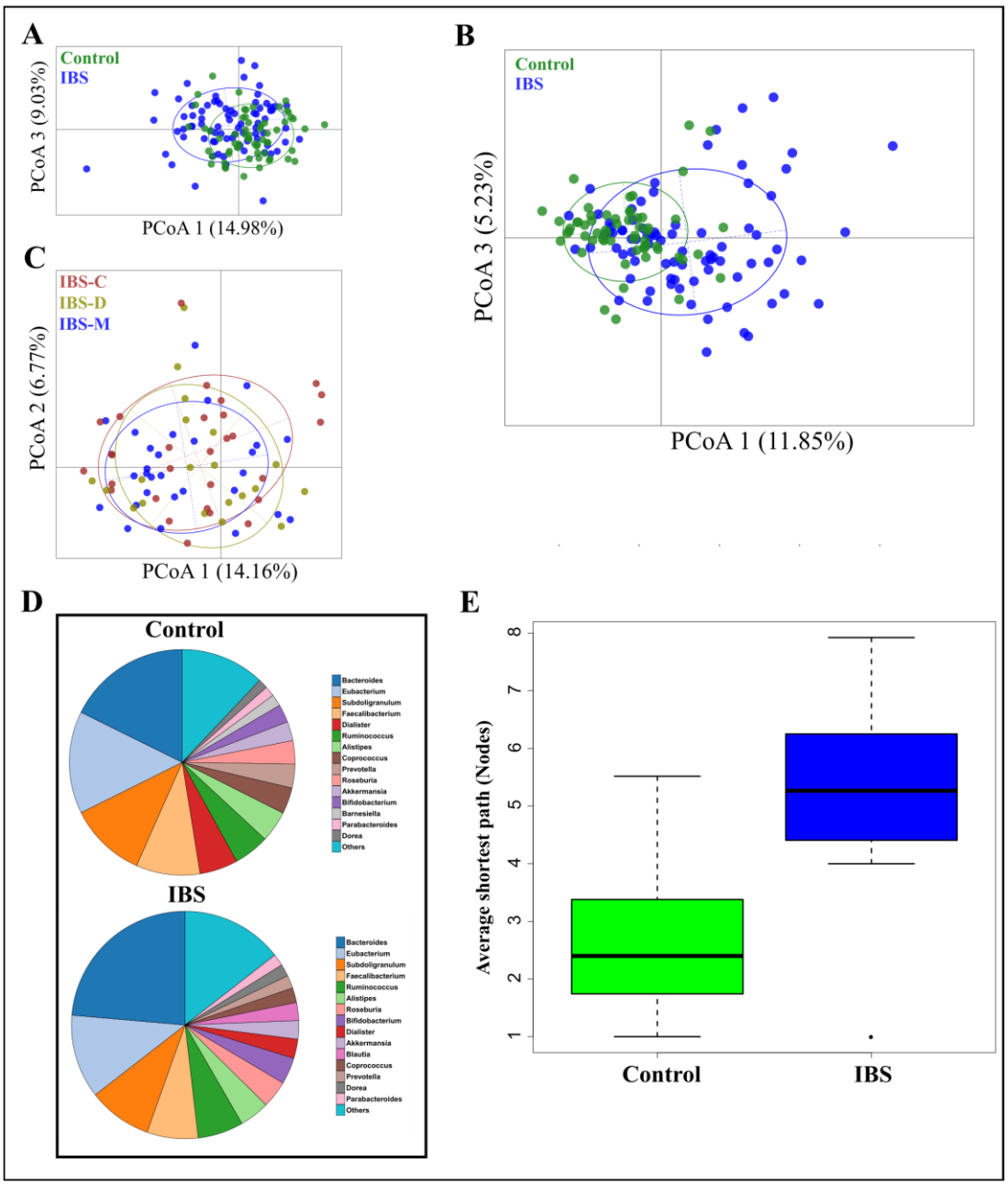


Figure 1. Microbiota compositional analysis of the control and IBS groups. (A) PCoA of microbiota beta diversity showing significant difference between the control and IBS groups. PCoA was performed with Spearman distance at the 16S rRNA gene based genus level abundance (P = 0.001; control, n = 63; IBS, n = 78). (B) PCoA of the metagenomic species analysis (CAGs) showing a significant difference between

the control and IBS groups. (P = 0.001; control, n = 59; IBS, n = 80). (C) PCoA of the microbiota composition showing no significant difference between IBS clinical subtypes. PCoA was performed with a Spearman distance at 16S rRNA gene based operational taxonomic unit level abundance (P = 0.976; IBS-C, n = 29; IBS-D, n = 20; IBS-M, n = 29). (D) Shotgun genus profile of control and IBS groups (control, n = 58; IBS, n = 78). (E) Connectivity of microbiota network of the control and IBS groups. A significantly higher level of connectivity was associated with the control network using the Wilcoxon rank-sum test (P = 2.818 × 10–15). P values for data/tests presented in panels A, B, and C were calculated with permutational multivariate analysis of variance (R function/package :adonis/vegan).

Metagenomic shotgun sequencing corroborated 16S rRNA gene profiling in separating patients with IBS from control individuals (Supplementary Figure 1). Moreover, the metagenomically derived composition at the genus and species level (as assigned by using shotgun sequence data) underscored the microbiota composition differences between patients with IBS and control individuals (Figure 1D). Pairwise comparison of the annotated metagenome data set identified 232 shotgun pathways stratified per organism that were significantly more abundant in the IBS group (Supplementary Table 3). These notably included a number of amino acid biosynthesis/degradation pathways, the altered activity of which may be relevant to IBS pathophysiology [27]. Other pathways that were less abundant in the metagenome of patients with IBS included galactose degradation, sulphate reduction, sulphate assimilation, and cysteine biosynthesis, collectively indicative of a reduced sulphur metabolism in IBS. The genes encoding 12 pathways were more abundant in patients with IBS, including those for starch V degradation. Of a total of 232 functional

pathways that were significantly more abundant in the IBS group, 113 were associated with the *Lachnospiraceae* family or *Ruminococcus* species.

The structure of the microbiome in a group of participants may be represented by microbial co-occurrence networks [28] that are based on patterns of co-abundance of taxa and that vary reproducibly in a manner constrained by factors including trophic interactions and co-operativity. These networks showed a significantly higher level of connectivity (shorter path lengths and denser co-occurrence network) in the control individuals than in patients with IBS (Figure 1E and Supplementary Figure 2). Furthermore, the faecal microbiota alpha diversity (intraindividual microbiota diversity) of patients with IBS was significantly lower than that of control individuals (Supplementary Figure 3A), with mean values for observed species of 345 ± 83 and 263 ± 73, respectively (Supplementary Figure 3B). There was no significant difference in microbiota alpha diversity values across the IBS clinical subtypes (Supplementary Figure 3C).

2.4.2 IBS patients with Bile Acid Malabsorption have altered faecal microbiome.

Because some patients with IBS-D may have BAM, which will influence transit time and possibly microbiota composition, we tested 19 of 21 patients with IBS-D, 26 of 29 patients with IBS-M, and 9 of 65 control individuals for SeHCAT retention, the criterion standard for quantifying BAM [29]. Failure to retain ≥15% of the labelled BA analogue was classified as BAM: 10%–15% retention was classified as mild BAM, 5%–10% as moderate BAM, and ≤5% as severe BAM, in accordance with current guidelines [30]. Of the 45 patients with IBS tested (Figure 2A), 18 were diagnosed with BAM, of whom 4 had severe BAM, 7 had moderate BAM, and 7 had mild BAM. A further 5 patients were classified as having borderline BAM (16%–20% retention). A

positive BAM classification was reported in 40% of the IBS population tested. Mild BAM was identified in 1 control individual, in whom IBS was later diagnosed. As expected, a positive BAM diagnosis was more common in IBS-D (74%) than in IBS-M (35%) (P value = .03) (Figure 2A). Only patients with IBS in the severe BAM category showed a significant separation in their microbiota from the microbiota of patients with normal, mild, moderate, or borderline BAM diagnoses (Figure 2B).

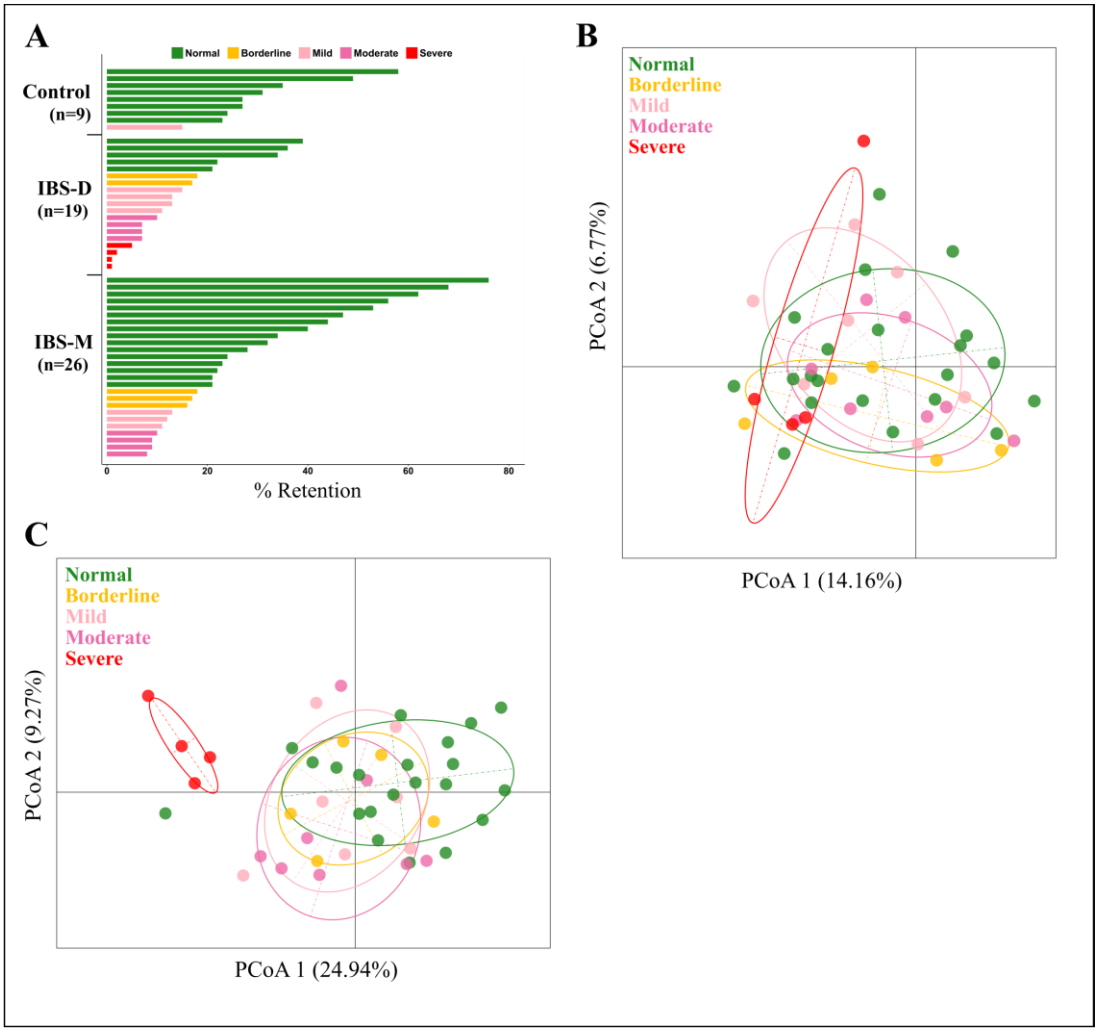


Figure 2. BAM distribution and microbiome and metabolomic profiles in SeHCAT-assayed participants. (A) SeHCAT retention rate in control individuals and patients with IBS. (B) PCoA of the microbiota composition showing no significant difference between BAM classes in patients with IBS tested by using Spearman distance at 16S rRNA gene based operational taxonomic unit level abundance (P = .289). (C) PCoA

of the faecal MS metabolomics showing a significant difference between BAM classes in tested patients with IBS , performed by using Spearman distance (P = .001). P values we calculated using permutational multivariate analysis of variance (R function/package :adonis/vegan).

To further investigate the biological impact of BAM, we performed untargeted faecal metabolite analysis by GC- and LC-MS. The faecal metabolome of patients with IBS with severe BAM was significantly different from that of the other BAM grades in those with IBS who had undergone SeHCAT (Figure 2C). We applied machine learning to the faecal metabolome data set for the patients with IBS who underwent SeHCAT. The predictive model successfully identified the participants who had BAM, with an AUC of 0.83 (sensitivity, 0.88; specificity, 0.67) across all 4 BAM grades. The model performed with 100% accuracy for severe and moderate BAM, 71.4% for mild BAM and 80% for borderline BAM (Supplementary Table 4 and Supplementary Figure 4). Optimization of the threshold increased the specificity to 0.714 without diminishing the sensitivity. Among the main predictive metabolites for BAM were glycerophospholipids and oligopeptides. Elevated levels of these compounds have been associated with altered fatty acid metabolism and disease [27,31,32]. Machine learning applied to the microbiome data sets was not predictive, indicating that these faecal metabolomic alterations are independent of the microbiota.

2.4.3 Altered urine and faecal metabolomes in IBS.

Because metabolomic analysis was informative for distinguishing participants with BAM, we extended its application to all participants, focusing initially on urine as a non-invasive test sample. We compared 2 methods: FAIMS analysis for volatile organics, and combined GC/LC-MS. The FAIMS technique did not identify discriminatory metabolites directly but separated samples/participants by

characteristic plumes of ionized metabolites. In unsupervised analysis, FAIMS readily identified urine samples from control individuals and patients with IBS (Figure 3A) but could not distinguish among IBS clinical subtypes (Supplementary Figure 5). GC/LC-MS analysis of the urine metabolome also separated patients with IBS from control individuals (Figure 3B) and with greater accuracy than FAIMS (Supplementary Figure 6A and B). Machine learning identified urine metabolomics features that are predictive of IBS (AUC, 1.000; sensitivity, 1.000, specificity, 0.97). Features that were highly predictive included dietary components such as epicatechin sulphate and medicagenic acid 3-O-b-d-glucuronide but also an acylglycine (N-undecanoylglycine) and an acylcarnitine (decanoylcarnitine) (Supplementary Table 5), discussed later in the Network Analysis section. Pairwise comparison of control and IBS urine metabolomes identified 127 differentially abundant features (Supplementary Table 6). Eighty-nine urine metabolites were significantly less abundant in patients with IBS, including a number of amino acids such as l-arginine, a precursor for the biosynthesis of nitric oxide that is associated both with mucosal defense and, perhaps, IBS pathophysiology [33]. Another 38 metabolites were present at significantly higher levels in IBS, including an acylglycine (N-undecanoylglycine) and an acylcarnitine (decanoylcarnitine). Elevated levels of metabolites from these groups are associated with altered fatty acid oxidation/metabolism and disease [34–36].

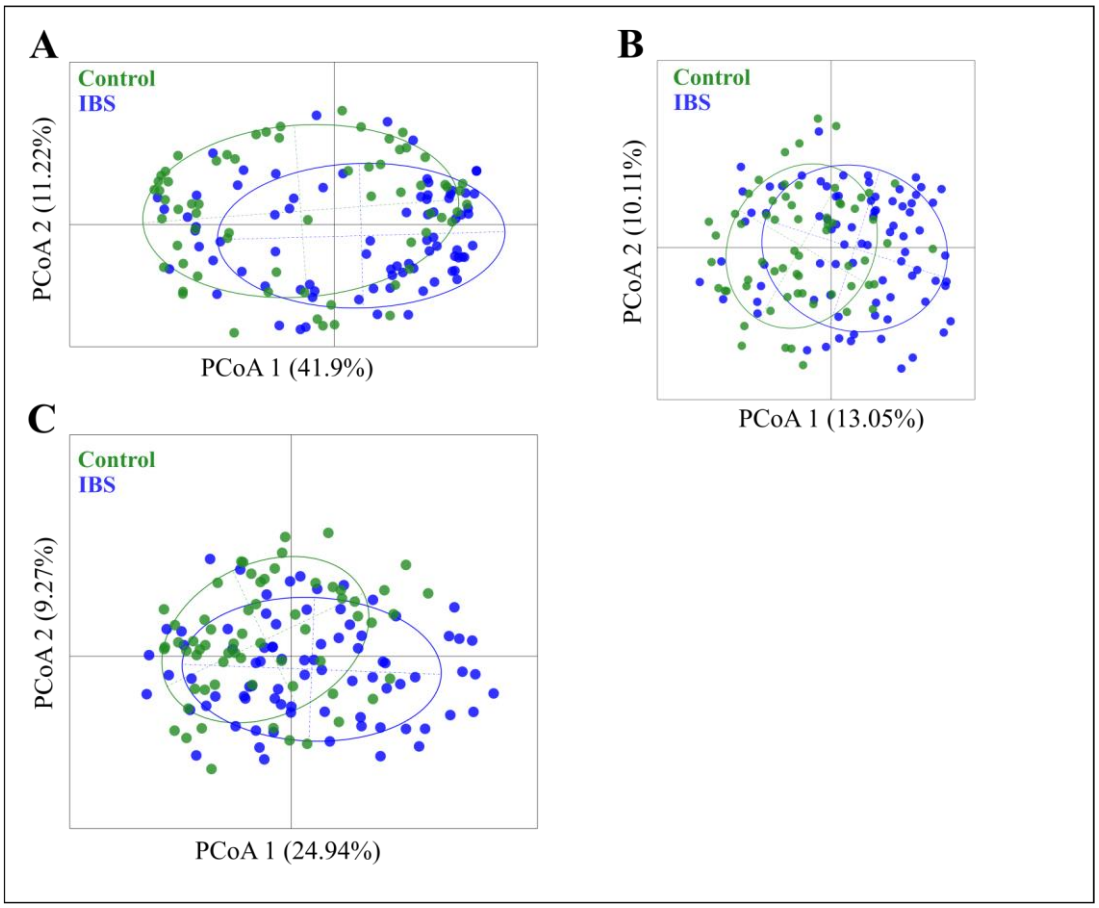


Figure 3. Comparison of control and IBS urine and faecal metabolomes. (A) PCoA of urine volatile organic compounds (FAIMS) metabolomes (P = .001; control, n = 65; IBS, n = 80). (B) PCoA of urine MS metabolomics using Spearman distance (P = .001; control, n = 63; IBS, n = 80). (C) PCoA of faecal MS metabolomics using Spearman distance (P = .001; control, n = 63; IBS, n = 80). P values we calculated using permutational MANOVA (R function/package :adonis/vegan).

Analysis of the faecal metabolome by GC/LC-MS separated patients with IBS from control individuals (Figure 3C), again without difference among the clinical IBS subtypes (Supplementary Figure 7). Machine learning identified 158 faecal metabolites predictive of IBS (AUC, 0.686; sensitivity, 0.737; specificity, 0.476) (Supplementary Table 7), which included the amino acids l-alanine, l-arginine, tyrosine, l-phenylalanine; dodecanedioic acid, an indicator of fatty acid oxidation defects [37], and 2 purine nucleosides, adenosine and inosine, previously reported as

biomarkers of IBS [38]. The adenosine ribonucleotide de novo biosynthesis functional pathway was significantly more abundant in species associated with IBS (Supplementary Table 3), which concords with adenosine being the second highest ranked predictive metabolite for IBS.

Pairwise comparison analysis of faecal metabolites identified 128 differentially abundant features, including 77 that were significantly depleted in IBS (Supplementary Table 8). Fifty-one faecal metabolites were significantly more abundant, including tyrosine (which was a predictive feature of IBS), lysine, and 3 BAs: [ST hydroxy] (25R)-3α,7α-dihydroxy-5β-cholestan-27-oyl taurine; [ST (2:0)] 5β-Chola-3,11-dien-24-oic acid, and ursodeoxycholic acid (UDCA). Because excess BAs affect water absorption and can lead to diarrhoea [39], the level of BA metabolites in the subgroups was analysed and showed a significant difference in the patients with IBS-D (Total BAs, sulphated BAs, UDCA, and conjugated BAs) (Supplementary Table 9). These differences were associated with an altered functional potential, reflected by the ursodeoxycholate biosynthesis and glycocholate metabolism pathway gene abundances correlating with the secondary BAs, UDCA, and total BA levels (Supplementary Table 10). Primary BAs and taurine/glycine conjugated BAs were not significantly different across the groups. Similar findings (in a smaller IBS/control cohort) were reported by Dior and colleagues [40] for secondary BAs, sulphated BAs, and UDCA- and taurine:glycine-conjugated BAs. These results are also corroborated by a recent study, wherein higher levels of faecal bile acids were seen in individuals with diarrhoea (functional diarrhoea + IBS-D) compared to constipated phenotype (functional constipation + IBS-C) [41]. Thus, the differences in faecal microbiome composition and predicted function in patients with IBS and control individuals are mirrored by differences in the measured metabolome in the 2 sample types.

2.4.4 Integrative analysis of diet, microbiome and metabolome reveals major alterations in IBS.

Weighted correlation network analysis [42] was used to explore the interactions between diet, microbiome, and metabolome in participants with IBS and control individuals (Figure 4). Nodes derived from each of the 6 data types are indicated and comprise multiple features listed in Supplementary Tables 11–16. Comparing the network in Figure 4 to an ordination plot based on between-class analysis (Supplementary Figure 8) shows that the large subnetwork is dominated by control-associated features and that 2 smaller subnetworks are dominated by IBS-associated features. The control individual subnetwork features a high density of connections and a multiplicity of feature types, contrasting with the sparser connections in the 2 IBS-related subnetworks. The taxonomy of the metagenomic species (CAGs) most predictive for IBS span the 4 species nodes (B1–4) in the control subnetwork and both IBS subnetworks (as shown in Figure 4 and detailed in Supplementary Tables 1 and 12). Species in the control subnetwork that are depleted in patients with IBS include *Paraprevotella* and *Bacteroides* species (B4) and *B. intestinihominis*, *Eubacterium eligens*, *Ruminococcus lactaris*, *Eubacterium biforme*, and *Desulfovibrio desulfuricans*, along with *Coprococcus* and *Eubacterium* species (B1). Species associated with IBS consisted of the *Ruminococcaceae*, *Lachnospiraceae*, *Alistipes*, and *Bacteroides* species (B3) in IBS subnetwork I and *Bifidobacterium adolescentis* and *Streptococcus* species (B2) in IBS subnetwork II. Collectively, this loss of *Bacteroidetes* species resonates with previous findings [16].

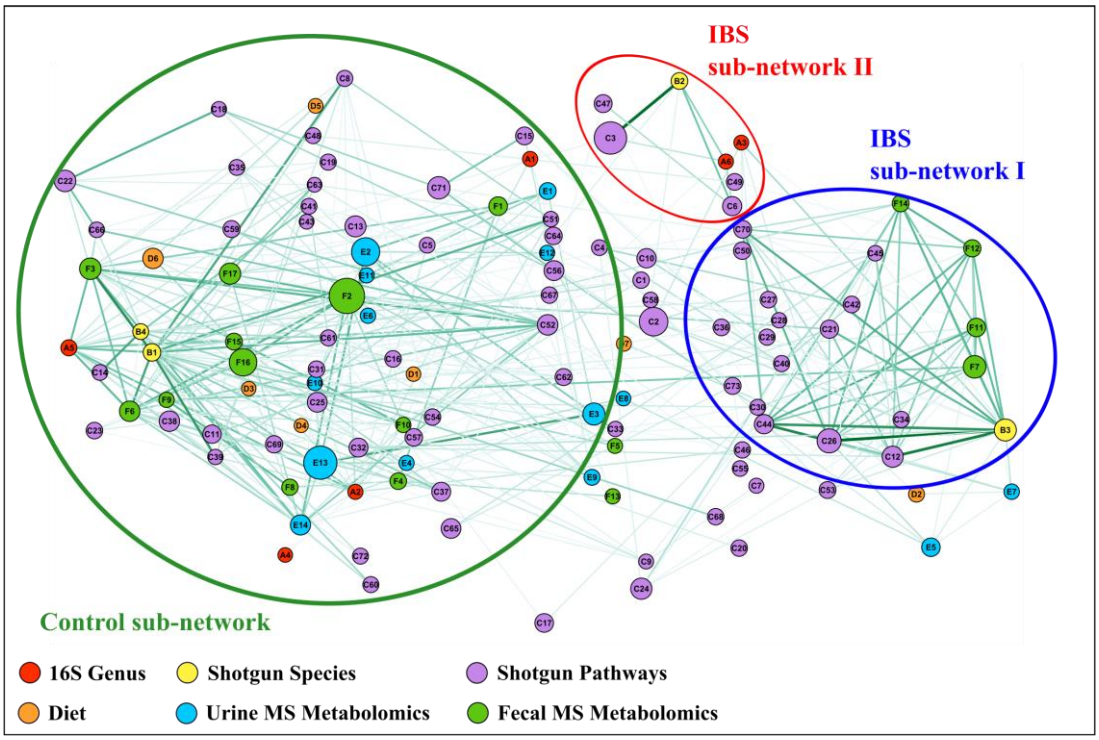


Figure 4. Visualization of structure and co-abundance across multiple-omic data sets. Nodes represent co-abundant variables from each of the 5/6 data sets, as clustered based on weighted correlation network analysis. Edges (lines) represent significant interactions between diet, microbiome, and metabolome in the IBS and control groups. The overall structure shows a large single-control-associated subnetwork and 2 IBS-associated subnetworks.

Differential association with diet is a major distinguishing feature of the control and IBS subnetworks. Underpinned by 5 dietary data nodes embedded in the subnetwork (nodes D1 and D3–6), and 1 on the periphery (node D7), the main structure in the control subnetwork represents the covariation of the normal microbiota with habitual diet. The food types in the dietary node D6, representing a high-fibre diet, are consumed significantly less frequently by patients with IBS, which is in line with findings in a previous meta-analysis of fibre intake in patients with IBS [43]. This D6 node is well connected to bacterial taxa and both faecal and urine metabolites in the control subnetwork. The only dietary node remote from the control individuals (D2)

is associated with the IBS population and is dominated by processed and high-glycaemic-index foods, but it is very weakly associated with bacterial function and is not integral to the IBS-associated subnetwork. These observations are reflected in the pairwise comparison of the dietary components consumed by the IBS and control populations (Supplementary Table 17).

The meat/high-protein dietary node D7 was associated with both the control and IBS subnetworks but to very different nodes and functional features, so we can infer that the control microbiota and the IBS microbiota use the protein in the diet in different ways. Supporting this protein utilization hypothesis, there is a relative overrepresentation of glycine metabolic pathways in the linked nodes of the IBS subnetwork (Supplementary Table 13). In addition, faecal metabolomics detected higher levels of peptides and amino acids in patients with IBS (Supplementary Table 16). Thus, altered dietary intake and utilization is one of several significant elements underpinning the major differences between the subnetworks of the control individuals and those with IBS.

Bacterial abundance data from clade-specific gene information generated more granular associations than 16S rRNA gene-based genus data and also showed associations particular to patients with IBS. Seven pathway nodes associated with the control subnetwork reflect the functionality of health-associated taxa, including *B. intestinihominis* (C8), *E. biforme* (C35), and *E. eligens* (C35). These nodes include phosphoglyceride and amino acid biosynthesis pathways (Supplementary Tables 2 and 13). Prominent in the IBS subnetworks are 4 nodes comprising microbiome functional pathways and 4 comprising faecal metabolites. The pathways reflect the functionality of the IBS-associated taxa described earlier and include a *Lachnospiraceae* bacterium (C12), *Ruminococcus torques* (C12), and *R. gnavus* (C44). *R. gnavus* is known to be

involved in tryptophan conversion to tryptamine, which is associated with slower colon transit [44]. Included in the 4 faecal metabolomics nodes are the amino acids l-alanine, l-lysine, l-methionine, l-phenylalanine, and tyrosine (F7), increased levels of which have been previously associated with IBS and the genera reported here [27]. l-alanine, l-phenylalanine, and tyrosine were predictive of IBS (Supplementary Table 6). The metabolite p-cresol sulphate, which is also predictive (F7), was significantly correlated with the main axis of the control to IBS separation (Figure 4) (p-cresol sulphate: coefficient, 0.258; P = .01). Although its excretion in the urine was reduced in the IBS group, as evidenced by the inverse trend in the urine metabolic data set (node E14 in the control subnetwork) (coefficient, –0.323; P = .003), this compound has been correlated with faecal water cytotoxicity [45]. A number of BAs—UDCA (F11), sulfolithocholylglycine, and [ST hydroxy] (25R)-3α,7α-dihydroxy-5β-cholestan-27-oyl taurine (F14)—and the bile pigment I-urobilin (F12) were also more abundant in IBS subnetwork I. Thus, a major feature discriminating patients with IBS is the presence of altered microbiota-encoded pathways linked to altered faecal metabolites.

The (smaller) IBS-associated subnetwork II is linked to 2 nodes comprising genera (A6), species (B2), and associated functionalities (C3 and C6) that are dominated by commensals of the small intestine, including *Streptococcus* and *Veillonella*. An in-depth analysis of medication, as presented in Supplementary Materials (Supplementary Table 18), identified proton pump inhibitor use (P < .05) as being associated with alterations in the microbiota and provided evidence that these medications are altering the microbiota. Linear regression was used to test for the influence of these medications on individual nodes in the network. This analysis confirmed the increased abundance of taxa, reflecting the known effect of proton pump inhibitors on the microbiome [46] (Q < 0.05 for both nodes A6 and B2). The functional nodes C3 and C36 were also significantly associated. A trend

was also identified for laxatives but failed to reach significance in the multivariate model (Supplementary Table 18) ($P < .1$), and no nodes were significantly associated with laxative use.

Two urine metabolome nodes (E5 and E7) are peripherally associated with IBS subnetwork I. They include a number of metabolites of host or bacterial origin including allantoin, and cis-4-decenedioic acid, decanoylcarnitine, and dodecanedioylcarnitine, respectively. These metabolites and compounds are indicative of urea metabolism [47] and fatty acid oxidation defects associated with disease [48], which may be relevant for IBS pathology. However, as noted, the strongest features that separate the IBS urine metabolome (Supplementary Table 15) are associated with the metabolism of dietary or medication components (nodes E5 and E7). This analysis suggests that the identification of patients with IBS by the application of FAIMS or MS to urine may be driven, in large part, by dietary adjustments or medications. This integrative modelling of diet, microbiome, and metabolome defines subnetworks in IBS.

2.5 Discussion

We have shown that the microbiome of patients with IBS is distinct from that of control individuals and is reflected in faecal metabolome profiles. However, metagenome and metabolome configurations do not separate the clinical subtypes of IBS (IBS-C, IBS-D, IBS-M). A recent study, integrating microbiome and metabolome data, showed the inability of microbiome and metabolome profiles to segregate the conventional clinical subtypes of IBS [6], which further strengthens our observation. We also show that the subset of patients with IBS-D and IBS-M with BAM have an altered faecal metabolome that can potentially be used to distinguish these individuals without requiring a SeHCAT test. Integrated analysis of the diet-microbiome-metabolome axis

shows how dietary ingredients interact with gut microbes in patients with IBS compared with control individuals and shows a different set of faecal metabolites that may drive or perpetuate the syndrome.

We identified a strain-level microbiome signature for IBS comprising 136 metagenomic species. The separation between the microbiota of patients with IBS and control individuals by unsupervised analysis exceeds that of earlier reports [15,17]. The limitations of 16S rRNA gene amplicon data sets and the relatively mild disease symptoms may account for failure to identify a microbiome signature in 1 report [17]. Moreover, microbiome alterations were significantly associated with physician-diagnosed IBS, but they were less significant in self-reported Rome criteria IBS [49].

Although urine metabolomics was highly discriminatory for IBS, the machine learning and network analysis showed that the compounds identified were predominantly associated with diet or medication. In contrast, the faecal metabolome correlated well with taxonomic and functional data for the microbiota, and machine learning identified discriminatory metabolites that corresponded to nodes in the network analysis that aligned as expected with IBS or control associations. The differentially abundant metabolites including BAs, amino acids, peptides, and p-cresol sulphate may contribute to IBS symptoms. p-Cresol, a precursor of p-cresol sulphate, has been shown to reduce the oxygen consumption of colonocytes and to be cytotoxic [45], and bacterial production of bioactive substances from dietary protein has been implicated in inflammation and tissue permeability in the gut [50]. Prominent among the differentially abundant metabolites were various sulphate, glucuronide, carnitine, glycine, and glutamine conjugates. This suggests a widespread up-regulation of phase 2 metabolism [51] in IBS. Supporting this, we found γ-glutamyl-cysteine, the precursor

of glutathione, to be enriched in the IBS faecal metabolome. Glutathione quenches free radical reactions and is a key metabolite in oxidative stress responses, which are prominent in the gut. Glutathione is also a key co-substrate for phase 2 conjugation of electrophilic substrates originating from diet, microbiome. or medication [51].

The microbiome among IBS clinical subtypes does not significantly differ, and the clinical value of assigning patients to symptom-based categories is limited. However, we were able to identify a subset of patients with IBS-D and IBS-M with BAM who were distinguishable by a metabolomic signature. Transit time (reflected by Bristol Stool Score) is a major covariate with microbiome composition [52], but the tendency for patients with IBS to alternate between the mixed, constipation, and diarrhoea subtypes [53] may mask or average out microbiota associations with transit time and may explain the modest microbiota alterations previously observed in IBS [16]. Regardless, all 3 subtypes can be distinguished from control individuals by a common faecal microbiome signature based on metagenomic species analysis. We detected BAM by SeHCAT in more than half of the tested participants with IBS-D and IBS-M combined. Differences in the microbiome were evident only in the small group with severe BAM. The unrecognized presence of appreciable numbers of participants with BAM may have contributed to low treatment success rates compared with placebo in previous trials of various IBS therapeutics [54]. It is also noteworthy that 89% (8/9) of the healthy volunteers returned normal bile retention scores from the SeHCAT assay, and because the purpose of deploying this assay was to discriminate individuals with low-retention IBS from healthy control individuals, this was an appropriate control group. Only participants with severe BAM had a significantly altered microbiome. This observation is based on analysis of a small number of samples and would need to be validated in future studies. However, we were able to identify a faecal

metabolomic signature for all participants with a diagnosis of BAM. This faecal metabolomics signature for BAM has the potential for clinical application because it requires instrumentation that is more convenient, more accessible, and less expensive than SeHCAT (which is not currently available in the United States).

The ecologic networks we constructed allowed identification of differences in fiber and protein metabolism in individuals with IBS compared to control individuals. People with IBS often manage their symptoms with popular diets, including low FODMAP, but the evidence for clinical improvement for this is limited [55]. Related to FODMAP, participants were not stratified for sugar malabsorption, which is a limitation of the present study. Overall fibre intake was significantly lower in participants with IBS, as expected, and multiple OTUs classified as *Bacteroidetes* phylum were detected in correspondingly lower abundance. Besides being able to adapt to different diets, *Bacteroides* species can also use fibre as an energy source when a high-fibre diet is consumed, but some species such as *Bacteroides thetaiotaomicron* can shift to mucin glycan metabolism as an energy source [56], linked with a reduced mucus layer in gnotobiotic mice. Some of the other fibre-associated taxa, *Lachnospiraceae* and *Ruminococcus* species, that are paradoxically more abundant in patients with IBS, are also mucin degrading [57], which could impair barrier function. Although hydrogen production and species specific contributions were not estimated in our study, we observed that a number of these bacterial taxa can potentially produce hydrogen (*Clostridium* species, *Bacteroides* species, and *Eubacterium* species [58]), and gas production is a clinical consideration for symptomology in IBS [59].

Amino acid metabolism in the large intestine plays a role in whole-body protein metabolism [60]. Differential metabolism of dietary protein by an altered gut microbiome in IBS might also contribute to the pathogenesis of IBS, even though the syndrome is not associated with a high-protein diet. However, a large number of peptides and amino acids are differentially present in the IBS faecal metabolome, and bacterial metabolism of proteins and amino acids has been implicated in other diseases such as type 2 diabetes and obesity, either directly or through further metabolism, such as into short-chain fatty acids [61]. Nitrogen limitation in the gut is also a determinant of microbiota composition and may shape host-microbiome interactions [62]. Due to majority of amino acid absorption is considered complete at the end of terminal ileum [60], limited amino acid availability is a normal feature of the distal bowel. The loss of this regulation in IBS concords with lower microbiota crosstalk, as shown by the decreased connectivity in the microbiota networks.

The findings of the current study have clinical implications. First, the microbiome and faecal metabolome and the urine metabolome offer objective biomarkers for IBS that could be translated to clinical practice. Second, a faecal metabolomic profile has been linked with BAM, which can distinguish it from non–BAM-related IBS. Third, the traditional Rome subtyping of IBS is not supported by differences in the microbiome and metabolome, and it may be time to look for an alternative basis for disease classification. Fourth, although the results in no way detract from the concept of an altered brain-gut axis in IBS, they point toward disturbances of the diet-microbiome-metabolome axis, which are consistent with the complaints of many patients and should inform the design of future therapeutic interventions in IBS. We contend that current stratification into clinical subtypes of IBS should not form the basis for therapeutic decisions, because the altered microbiota (compared with control

individuals) is similar in the subtypes, consistent with alternating between constipation and diarrheal forms in many patients. A more informative stratification would be achieved by faecal microbiota and metabolome profiling (thereby excluding patients with BAM, who would be prescribed BA sequestrants) and then treatment to rectify the altered microbiome, metabolome, and diet interaction networks. The metagenomic and metabolomic signatures that distinguish patients with IBS from control individuals identified here may be targeted by these microbiota-directed therapies.

2.6 Acknowledgements

We are grateful to the 4D Pharma Cork clinical team, Caroline O'Leary, Marian O'Meara, and Sinead Kelly, who recruited the participants to the study and collected the biological samples and clinical metadata. We would also like to thank gastroenterologist Dr Syed Akbar Zulquernain and Dr Elizabeth Kenny at Cork University Hospital, who assisted with IBS patient recruitment. We acknowledge the input of 4D Pharma Cork laboratory staff, Luke O'Brien, Fiona Grady, Marie McBrien, and Siobhan O'Neill, for sequencing and metabolomic analysis. We also thank the radiographers Fionnuala McDonnell, Lean O'Donoghue, Pauline Beecher, Roisin Wilson, and Shane Hayes at Cork University Hospital, who performed the body scans for the SeHCAT assay.

2.7 References

1. Enck P, Aziz Q, Barbara G, et al. Irritable bowel syndrome. Nat Rev Dis Primers 2016;2:16014.

2. Soares RLS. Irritable bowel syndrome: a clinical review. World J Gastroenterol 2014;20:12144–12160.

3. Van Oudenhove L, Aziz Q. The role of psychosocial factors and psychiatric disorders in functional dyspepsia. Nat Rev Gastroenterol Hepatol 2013;10:158–167.

4. Koloski NA, Jones M, Kalantar J, et al. The brain--gut pathway in functional gastrointestinal disorders is bidirectional: a 12-year prospective population-based study. Gut 2012;61:1284–1290.

5. Schwille-Kiuntke J, Mazurak N, Enck P. Systematic review with meta-analysis: post-infectious irritable bowel syndrome after travellers' diarrhoea. Aliment Pharmacol Ther 2015;41:1029–1037.

6. Mujagic Z, Kasapi M, Jonkers DM, et al. Integrated fecal microbiome-metabolome signatures reflect stress and serotonin metabolism in irritable bowel syndrome. Gut Microbes 2022;14:2063016.

7. Gao J, Xiong T, Grabauskas G, et al. Mucosal Serotonin Reuptake Transporter Expression in Irritable Bowel Syndrome Is Modulated by Gut Microbiota Via Mast Cell-Prostaglandin E2. Gastroenterology 2022;162:1962-1974.e6.

8. Quigley EMM. The Gut-Brain Axis and the Microbiome: Clues to Pathophysiology and Opportunities for Novel Management Strategies in Irritable Bowel Syndrome (IBS). J Clin Med 2018;7:E6.

9. Drossman DA, Morris CB, Schneck S, et al. International survey of patients with IBS: symptom features and their severity, health status, treatments, and risk taking to achieve clinical benefit. J Clin Gastroenterol 2009;43:541–550.

10. Wang T, Rijnaarts I, Hermes GDA, et al. Fecal Microbiota Signatures Are Not Consistently Related to Symptom Severity in Irritable Bowel Syndrome. Dig Dis Sci 2022.

11. Lacy BE, Everhart KK, Weiser KT, et al. IBS patients' willingness to take risks with medications. Am J Gastroenterol 2012;107:804–809.

12. Slattery SA, Niaz O, Aziz Q, et al. Systematic review with meta-analysis: the prevalence of bile acid malabsorption in the irritable bowel syndrome with diarrhoea. Aliment Pharmacol Ther 2015;42:3–11.

13. Baumgartner M, Lang M, Holley H, et al. Mucosal Biofilms Are an Endoscopic Feature of Irritable Bowel Syndrome and Ulcerative Colitis. Gastroenterology 2021;161:1245-1256.e20.

14. Carroll IM, Ringel-Kulka T, Keku TO, et al. Molecular analysis of the luminal- and mucosal-associated intestinal microbiota in diarrhea-predominant irritable bowel syndrome. Am J Physiol Gastrointest Liver Physiol 2011;301:G799-807.

15. Rajilić-Stojanović M, Biagi E, Heilig HGHJ, et al. Global and deep molecular analysis of microbiota signatures in fecal samples from patients with irritable bowel syndrome. Gastroenterology 2011;141:1792–1801.

16. Jeffery IB, O'Toole PW, Öhman L, et al. An irritable bowel syndrome subtype defined by species-specific alterations in faecal microbiota. Gut 2012;61:997–1006.

17. Tap J, Derrien M, Törnblom H, et al. Identification of an Intestinal Microbiota Signature Associated With Severity of Irritable Bowel Syndrome. Gastroenterology 2017;152:111-123.e8.

18. Collins SM. A role for the gut microbiota in IBS. Nat Rev Gastroenterol Hepatol 2014;11:497–505.

19. Ohman L, Simrén M. Intestinal microbiota and its role in irritable bowel syndrome (IBS). Curr Gastroenterol Rep 2013;15:323.

20. Longstreth GF, Thompson WG, Chey WD, et al. Functional bowel disorders. Gastroenterology 2006;130:1480–1491.

21. Brown JR-M, Flemer B, Joyce SA, et al. Changes in microbiota composition, bile and fatty acid metabolism, in successful faecal microbiota transplantation for Clostridioides difficile infection. BMC Gastroenterol 2018;18:131.

22. Arasaradnam RP, Westenbrink E, McFarlane MJ, et al. Differentiating coeliac disease from irritable bowel syndrome by urinary volatile organic compound analysis--a pilot study. PLoS One 2014;9:e107312.

23. Watson L, Lalji A, Bodla S, et al. Management of bile acid malabsorption using low-fat dietary interventions: a useful strategy applicable to some patients with diarrhoea-predominant irritable bowel syndrome? Clin Med (Lond) 2015;15:536–540.

24. Flemer B, Warren RD, Barrett MP, et al. The oral microbiota in colorectal cancer is distinctive and predictive. Gut 2018;67:1454–1463.

25. Borrel G, McCann A, Deane J, et al. Genomics and metagenomics of trimethylamine-utilizing Archaea in the human gut microbiome. ISME J 2017;11:2059–2074.

26. Nielsen HB, Almeida M, Juncker AS, et al. Identification and assembly of genomes and genetic elements in complex metagenomic samples without using reference genomes. Nat Biotechnol 2014;32:822–828.

27. Shankar V, Homer D, Rigsbee L, et al. The networks of human gut microbe-metabolite associations are different between health and irritable bowel syndrome. ISME J 2015;9:1899–1903.

28. Faust K, Raes J. Microbial interactions: from networks to models. Nat Rev Microbiol 2012;10:538–550.

29. Kurien M, Thurgar E, Davies A, et al. Challenging current views on bile acid diarrhoea and malabsorption. Frontline Gastroenterol 2018;9:92–97.

30. Summers JA, Peacock J, Coker B, et al. Multicentre prospective survey of SeHCAT provision and practice in the UK. BMJ Open Gastroenterol 2016;3:e000091.

31. Huang H, Zhang A, Cao H, et al. Metabolomic analyses of faeces reveals malabsorption in cirrhotic patients. Dig Liver Dis 2013;45:677–682.

32. Zhou C, Jia H-M, Liu Y-T, et al. Metabolism of glycerophospholipid, bile acid and retinol is correlated with the early outcomes of autoimmune hepatitis. Mol Biosyst 2016;12:1574–1585.

33. Wallace JL. Nitric oxide in the gastrointestinal tract: opportunities for drug development. Br J Pharmacol 2019;176:147–154.

34. Hoppel C. The role of carnitine in normal and altered fatty acid metabolism. Am J Kidney Dis 2003;41:S4-12.

35. Liu X, Liu Y, Cheng M, et al. Metabolomic Responses of Human Hepatocytes to Emodin, Aristolochic Acid, and Triptolide: Chemicals Purified from Traditional Chinese Medicines. J Biochem Mol Toxicol 2015;29:533–543.

36. Vishwanath VA. Fatty Acid Beta-Oxidation Disorders: A Brief Review. Ann Neurosci 2016;23:51–55.

37. Korman SH, Waterham HR, Gutman A, et al. Novel metabolic and molecular findings in hepatic carnitine palmitoyltransferase I deficiency. Mol Genet Metab 2005;86:337–343.

38. Ponnusamy K, Choi JN, Kim J, et al. Microbial community and metabolomic comparison of irritable bowel syndrome faeces. J Med Microbiol 2011;60:817–827.

39. Riemsma R, Al M, Corro Ramos I, et al. SeHCAT [tauroselcholic (selenium-75) acid] for the investigation of bile acid malabsorption and measurement of bile acid pool loss: a systematic review and cost-effectiveness analysis. Health Technol Assess 2013;17:1–236.

40. Dior M, Delagrèverie H, Duboc H, et al. Interplay between bile acid metabolism and microbiota in irritable bowel syndrome. Neurogastroenterol Motil 2016;28:1330–1340.

41. James SC, Fraser K, Young W, et al. Concentrations of Fecal Bile Acids in Participants with Functional Gut Disorders and Healthy Controls. Metabolites 2021;11:612.

42. Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. BMC Bioinformatics 2008;9:559.

43. Moayyedi P, Quigley EMM, Lacy BE, et al. The effect of fiber supplementation on irritable bowel syndrome: a systematic review and meta-analysis. Am J Gastroenterol 2014;109:1367–1374.

44. 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 2018;23:775-785.e5.

45. Beaumont M, Portune KJ, Steuer N, et al. Quantity and source of dietary protein influence metabolite production by gut microbiota and rectal mucosa gene expression: a randomized, parallel, double-blind trial in overweight humans. Am J Clin Nutr 2017;106:1005–1019.

46. Imhann F, Bonder MJ, Vich Vila A, et al. Proton pump inhibitors affect the gut microbiome. Gut 2016;65:740–748.

47. Schicho R, Shaykhutdinov R, Ngo J, et al. Quantitative metabolomic profiling of serum, plasma, and urine by (1)H NMR spectroscopy discriminates between patients with inflammatory bowel disease and healthy individuals. J Proteome Res 2012;11:3344–3357.

48. Nizioł J, Bonifay V, Ossoliński K, et al. Metabolomic study of human tissue and urine in clear cell renal carcinoma by LC-HRMS and PLS-DA. Anal Bioanal Chem 2018;410:3859–3869.

49. Vich Vila A, Imhann F, Collij V, et al. Gut microbiota composition and functional changes in inflammatory bowel disease and irritable bowel syndrome. Sci Transl Med 2018;10:eaap8914.

50. Diether NE, Willing BP. Microbial Fermentation of Dietary Protein: An Important Factor in Diet¯Microbe¯Host Interaction. Microorganisms 2019;7:E19.

51. Swanson HI. Drug Metabolism by the Host and Gut Microbiota: A Partnership or Rivalry? Drug Metab Dispos 2015;43:1499–1504.

52. Falony G, Joossens M, Vieira-Silva S, et al. Population-level analysis of gut microbiome variation. Science 2016;352:560–564.

53. Drossman DA, Morris CB, Hu Y, et al. A prospective assessment of bowel habit in irritable bowel syndrome in women: defining an alternator. Gastroenterology 2005;128:580–589.

54. Ford AC, Harris LA, Lacy BE, et al. Systematic review with meta-analysis: the efficacy of prebiotics, probiotics, synbiotics and antibiotics in irritable bowel syndrome. Aliment Pharmacol Ther 2018;48:1044–1060.

55. Dionne J, Ford AC, Yuan Y, et al. A Systematic Review and Meta-Analysis Evaluating the Efficacy of a Gluten-Free Diet and a Low FODMAPs Diet in

Treating Symptoms of Irritable Bowel Syndrome. Am J Gastroenterol 2018;113:1290–1300.

56. Desai MS, Seekatz AM, Koropatkin NM, et al. A Dietary Fiber-Deprived Gut Microbiota Degrades the Colonic Mucus Barrier and Enhances Pathogen Susceptibility. Cell 2016;167:1339-1353.e21.

57. Crost EH, Tailford LE, Le Gall G, et al. Utilisation of mucin glycans by the human gut symbiont Ruminococcus gnavus is strain-dependent. PLoS One 2013;8:e76341.

58. Magnúsdóttir S, Heinken A, Kutt L, et al. Generation of genome-scale metabolic reconstructions for 773 members of the human gut microbiota. Nat Biotechnol 2017;35:81–89.

59. Wilder-Smith CH, Olesen SS, Materna A, et al. Fermentable Sugar Ingestion, Gas Production, and Gastrointestinal and Central Nervous System Symptoms in Patients With Functional Disorders. Gastroenterology 2018;155:1034-1044.e6.

60. Wielen N van der, Moughan PJ, Mensink M. Amino Acid Absorption in the Large Intestine of Humans and Porcine Models. J Nutr 2017;147:1493–1498.

61. Neis EPJG, Dejong CHC, Rensen SS. The role of microbial amino acid metabolism in host metabolism. Nutrients 2015;7:2930–2946.

62. Holmes AJ, Chew YV, Colakoglu F, et al. Diet-Microbiome Interactions in Health Are Controlled by Intestinal Nitrogen Source Constraints. Cell Metab 2017;25:140–151.

2.8 Supplementary information

2.8.1 Supplementary methods

Subject recruitment and sample collection.

Exclusion criteria included the use of antibiotics in the 6 weeks before study enrollment, other chronic illnesses including gastrointestinal diseases, severe psychiatric disease, and abdominal surgery other than hernia repair or appendectomy. Standard-of-care blood analysis was carried out for all participants if recent results were not available. Celiac disease was excluded by serology in all participants. The inclusion/exclusion criteria for the control population were the same as for the IBS population with the exception of having to fulfill the Rome IV criteria for IBS. Gastrointestinal symptom history, psychological symptoms, diet, medical history, and medication data were collected on each participant (both participants with IBS and control individuals) by using the following questionnaires: Bristol Stool Score, Hospital Anxiety and Depression Scale [1], Food Frequency Questionnaire (FFQ) [2]. Patients with IBS-D and IBS-M reporting diarrhea and a subset of consenting control individuals were assessed for BAM by SeHCAT, a radiolabeled synthetic BA that is used to clinically diagnose BAM and that is not metabolized by bacteria and passes through the enterohepatic circulation in the same way as endogenous BAs.

Ethical approval was granted by the local ethics committee before the study commenced, and all participants provided written informed consent to take part. Participants collected a freshly voided faecal sample at home using a collection kit and brought the sample to the clinic that day, when a fresh urine sample was also collected. Samples were kept at 4°C until they were brought to the laboratory for storage at –80°C, which occurred within a few hours of the sample collection.

16S rRNA gene amplicon sequencing

16S rRNA gene amplicon's polymerase chain reaction (PCR) forward primer (S-D-Bact-0341-b-S-17):

5′-

TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGCCTACGGGNGGCWGCAG

16S amplicon PCR reverse primer (S-D-Bact-0785-a-A-21):

5′-

GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGGACTACHVGGGTATCTAATCC

The amplicon size was 531 base pairs (bp). The products were purified, and forward and reverse barcodes were attached by a second round of adapter PCR.

The modifications from the methods described by Brown et al [3] included bead beating tubes containing 0.5 g of 0.1-mm zirconia beads and 4 × 3.5–mm glass beads. Faecal samples were homogenized by bead beating for 3 × 60–s cycles and cooled on ice between each cycle. Genomic DNA was visualized on 0.8% agarose gel and quantified by using the SimpliNano Spectrometer (Biochrom, Holliston, MA). The PCR master mix used 2X Phusion Taq High-Fidelity Mix (Thermo Fisher Scientific, Waltham, MA) and 15 ng of DNA. The resulting PCR products were purified and quantified, and equimolar amounts of each amplicon were then pooled before being sent for sequencing to the commercial supplier (GATC Biotech AG, Konstanz, Germany) on the MiSeq (2 × 250–bp) chemistry platforms.

Shotgun metagenomic sequencing.

Genomic DNA was extracted as described earlier. The DNA quality was checked on 0.8% agarose gel and quantified using the Simplinano (Thermo Fisher Scientific). One μg (concentration > 5 ng/μL) of high-molecular-weight DNA for each sample was sent to GATC Biotech for sequencing on an Illumina HiSeq (HiSeq 2500) using 2 × 250–bp paired-end chemistry. This returned 2,714,158,144 raw reads (2,612,201,598 processed reads), of which 45.6% were mapped to an average of 222,945 gene families per sample, with a mean count value of 8,924,302 ± 2,569,353 per sample.

Urine and Faecal Metabolomics.

For FAIMS analysis, frozen (–80°C) urine samples were thawed overnight at 4°C, and 5 mL of each urine sample was aliquoted into a 20-mL glass vial and placed into an ATLAS sampler (Owlstone, Cambridge, UK) attached to the Lonestar FAIMS instrument (Owlstone). The sample was heated to 40°C and sequentially run 3 times. Each sample run had a flow rate over the sample of 500 mL/min of clean, dry air. Further make-up air was added to create a total flow rate of 2.5 L/min. The FAIMS was scanned from 0% to 99% dispersion field in 51 steps and from +6V to –6V compensation voltage in 512 steps, and both positive and negative ions were detected to produce an untargeted volatile organic compound profile for each sample.

Untargeted metabolomic analysis was performed on urine samples by LC-MS and solid phase microextraction GC, and metabolites were identified by using electrospray ionization MS. Short-chain fatty acid (SCFA) analysis was also performed by LC-MS/MS.

Faecal samples were dried and resuspended to a final concentration of 10 mg per 400 μL before analysis. GC-MS and SCFA analysis were performed with wet samples.

Untargeted metabolomics and SCFA analysis were carried out as described previously for urine MS metabolomics.

For urine metabolomics, the values of metabolites were normalized with reference to urine creatinine levels in each sample. For faecal metabolomics, levels of metabolites identified using LC-MS were normalized by weight during preparation. For metabolites identified using GC-MS, the values were normalized with wet weights per sample.

Bioinformatic Analysis.

The 16S rRNA gene amplicon reads were trimmed using flash [4]. USEARCH was used to generate the operational taxonomic unit (OTU) table [5]. UPARSE was used to cluster into OTUs at 97% similarity [6]. The UCHIME chimera removal algorithm was used with Chimeraslayer to remove chimeric sequences [7]. The Ribosomal Database Project (RDP) taxonomic classifier was used to assign taxonomy to the representative OTU sequences [5], and microbiota compositional (abundance and diversity) information was generated. Miseq 16S rRNA gene sequencing data were returned for 144 participants. Data generated for 3 samples (2 IBS and 1 control) were removed because the number of reads returned from sequencing was too low for analysis, leaving 141 samples (control, n = 63; IBS, n = 78). Metagenomic shotgun sequencing as well as urine and faecal metabolomics were also performed. Significant differences in beta-diversity were defined by Permutational MANOVA (R function/package adonis/vegan) [8].

Shotgun metagenomic sequencing

The number of raw read pairs obtained after sequencing varied from 5,247,013 to 21,280,723 (mean ± standard deviation, 9,763,159 ± 2,408,048). Six control samples were not sequenced due to data not passing quality control or not having a sample

available (control, n = 59; IBS, n = 80). Reads were processed in accordance with the standard operating procedures of Human Microbiome Project (HMP) Consortium [9] Metagenomic composition, and functional profiles were generated using HUMAnN2 pipeline [10]. For each sample, multiple profiles were obtained, including microbial composition profiles from clade-specific gene information (using MetaPhlAn2), gene family abundance, pathways stratified per organism, total pathway coverage, and abundance.

Co-abundant Genes Analysis

Clusters of CAGs representing metagenomically defined species variables were identified by using gene family abundances generated through the HUMAnN2 pipeline, based on metagenomic sequences of the samples. Coabundance analysis of the gene families was performed with a modified canopy clustering algorithm [11]. This was run with default parameters for 139 samples (IBS, n = 80; controls, n = 59) using the relative abundance of 1,706,571 gene families (UniRef90 database) stratified by species with the HUMAnN2 methodology [10].

Microbial co-occurrence networks

Microbial co-occurrence networks were generated based on Spearman correlation. Normalized OTU abundance profiles of each subject subgroup were selected, and OTUs present in at least 30% of the samples were considered. For each pair of microbes, Spearman correlation coefficient was calculated. The statistical robustness of the observed significant correlations was tested with the ReBoot method [12]. Gene family clusters were identified and filtered so that at least 90% of the cluster signal originated from more than 3 samples and contained more than 2 gene families [11]. The clusters remaining after filtering were termed coabundant groups or CAGs. The

abundance indices of the CAGs were generated by singular value decomposition as implemented in principal component analysis by using the dudi.pca command with default parameters (ade4 package in R, version 3.5.1). The first principal component was extracted as the index, and directionality was corrected by the index being compared to the median CAG gene abundance using the Spearman correlation of all values within a CAG. CAGs returning a negative correlation were corrected by inverting the principal component values for that CAG. The principal component values were then scaled by subtracting the minimum value. Because each CAG is composed of multiple gene families, taxonomy was assigned to a CAG by reporting the most common genera and species associated with the gene families in the CAGs, along with the percentage of the CAG that they composed. For CAGs where a genus or species represented greater than 60% of the gene families, a taxonomy was assigned.

The ReBoot method [12] used a Python script developed in-house, wherein a null distribution of correlations between the microbe pairs was generated by permuting and renormalizing the OTU profile 1000 times and subsequently recalculating the Spearman correlation between pairs. Similar distribution was formed using bootstrapped resampling (70% of samples) with replacement for 1000 iterations and was compared to the null distribution using the Fisher z-test. The P values of the comparisons were noted and were corrected with the Benjamini-Hochberg false discovery rate method with a family-wise error rate of 0.05. Correlations with corrected P values of <.05 were considered significant, and the top 10% of positive correlations were considered for network visualization and analysis.

Metabolomics data and dietary analysis.

Faecal and urine MS metabolomics data were returned for all participants with IBS (n = 80) and all but 2 control individuals (n = 63) because these did not pass quality control or no sample was available. The number of individuals common to all the data sets was 134, consisting of 78 patients with IBS and 56 control individuals. Dietary habits were assessed for all participants with the use of a 146-item FFQ. This FFQ is an adapted version of the European Prospective Investigation into Cancer and Nutrition (EPIC) cancer study questionnaire [13].

Urine metabolome (Field Asymmetric Waveform Ion Mobility Spectrometry)

Each urine sample analyzed with FAIMS yielded a profile with approximately 52,224 data points. A pooled profile containing these data points for each sample was generated for preprocessing, to reduce the noise, size, and complexity of the data.

The signals for each sample at each dispersion field (DF) were smoothed with the Savitzky-Golay filter [14] (window size, 9; degree, 3). The signals were trimmed based on an optimized cutoff of 0.007 for positive mode and –0.007 for negative mode outputs, to obtain the region of interest and reduce the baseline noise. Signals were aligned to the trimmed signals at each DF, by using cross-correlation, by using the mean signal as reference to make them comparable. Because the initial DFs of the FAIMS signal, and higher DFs were noninformative, signals corresponding to 17th DF through the 42nd DF of both, positive, and negative modes were considered. These preprocessing steps were performed with customized programs developed in Python, version 2.7.11, with relevant packages (Scipy v-1.1 [15], and Numpy v-1.15.2 [16]). To further reduce the complexity and to retain informative data, kurtosis normality tests

were performed on each feature vector and features with raw P value > .1, were considered, and the final profile was generated for various statistical analyses.

Urine metabolome (Mass Spectrometry)

A total of 2,887 metabolites were returned from untargeted urine metabolomics analysis, of which 594 were identified. Only the identified features with peak values normalized by creatinine levels in urine (mg/dL) were considered for further analysis.

Faecal metabolome (Mass Spectrometry)

A total of 2,933 metabolites were returned from untargeted faecal metabolomics analysis carried out by the service provider, of which 753 were identified. Metabolites identified by using LC-MS were not normalized, because the faecal samples were already normalized with dry weight (10 mg per 400 μL) during sample preparation. Metabolites identified with GC-MS were normalized with corresponding sample wet weights. Only the identified metabolites were considered for further analyses. Machine learning analysis was carried out as described previously for the urine metabolome. Summary statistics for all data sets were generated with the Wilcoxon rank-sum test with Q-value adjustment for multiple testing.

Weighted Correlation Network Analysis (WGCNA) and Machine Learning

For each data type, unsupervised feature clustering was performed with the weighted correlation network analysis and machine learning (WGCNA) package [17], version 1.60, in R, version 3.3.1 [18], with modifications as described below.

The FFQ is a validated dietary tool [19] that consists of 12 food categories with 147 specific food items and 10 frequency categories. Its quantitative interpretation leads to a numeric score for major food classes and macronutrients, providing the granularity required for microbiome interaction analysis. For clustering into indexes for the

network analysis, the WGCNA methodology [17] was applied. The minimum number of features per module was set at 5.

After removal of outlying samples, only features present in >10% of the samples were considered for the WGCNA analysis. A signed network, based on Spearman correlation, was generated and made scale free based on an optimized soft power, followed by generation of adjacency matrix of features. A network in the form of a dendrogram was generated, and preliminary feature clustering was achieved by dynamic tree cutting, with a condition of 5 or 10 minimum number of features per cluster. Eigenvalues were calculated for each of the premodules, and Spearman distances between the modules were calculated. These premodules were clustered into final modules at a Spearman distance < 0.25, and eigenvalues of each final module were calculated using all the samples. WGCNA module eigenvalues from all data types were pooled for further analysis. To investigate the relation between modules of the WGCNA, a co-occurrence network was generated by using Spearman correlation. Positive correlations between WGCNA modules with adjusted P values (Benjamini-Hochberg) < .05 were considered and visualized using Gephy, version 0.9.2 (http://gephi.github.io/). The spatial arrangement of the nodes was based on loadings of the modules in between-class analysis among the 3 sample subgroups, which was performed using the ade4 (1.7.5) package in R, version 3.3.1).

Machine learning

An in-house machine learning pipeline was applied to each data type (16S rRNA gene, shotgun, urine, and faecal MS metabolomics and BAM-faecal MS metabolomics) with a 2-step approach within a 10-fold cross-validation. Within each validation fold, least absolute shrinkage and selection operator (LASSO) feature selection was carried out,

followed by random forest modelling [20], and an optimized model was validated against the cross-validation test data, which is external to the cross-validation training subset.

First, feature selection was performed with the LASSO algorithm to improve the accuracy and interpretability of models by selecting the relevant features. An optimal lambda (penalty parameter) range was generated and assessed by the cross-validated LASSO using the caret and glmnet packages. Samples were assigned weights based on class probabilities. This step reduces skewness due to imbalances in class composition. The assigned weights for the training samples are used in subsequent applicable steps. If the number of selected features is ≥5, then the selected features are considered for downstream classifier generation; otherwise, all the features are considered for the classifier generation step. The training data were filtered to include only the features selected by the LASSO algorithm. After the LASSO step, an optimized random forest classifier (with 1500 trees) is generated with the selected features (or all features), which can be used to predict the external test fold. This step is also performed using caret, and internal cross-validations, with an objective of maximizing the receiver operating characteristic AUC by tuning the mtry parameter. For tuning, if the number of selected features is ≥5, mtry ranges from 1 to the square root of the number of selected features, or else the range is from 1 to 6. Both LASSO feature selection and random forest modeling were performed within a 10-fold cross-validation, which generated an internal 10-fold prediction model that predicts the IBS or control classification of samples. This 10-fold cross-validation procedure was repeated 10 times, and the average AUC, sensitivity, and specificity are reported. The optimized model is then used to predict the cross-validation test subset, and final classifier performance metrics are calculated from across the 10 folds of the cross-validation (AUC, sensitivity, and specificity). For BAM faecal metabolomic data

analysis, machine learning was performed in a similar manner, with the only difference being that, instead of 10-fold cross-validation, leave-one-out cross-validation was used at every cross-validation step. Data used to predict BAM consisted of borderline to severe BAM, as defined by SeHCAT retention of <15%, and normal bile acid retention samples, as defined as SeHCAT retention of >20%. The models were implemented with R software, version 3.5.1, using package glmnet [21] for LASSO feature selection, and R packages randomForest [22] and caret for the classification [23].

The ability of urine FAIMS metabolomics to differentiate between health classes was tested with support vector machines, using Python 2.7 and Scikit-Learn [24] (version 0.19.2). Each of the selected features of the FAIMS profile, using the kurtosis normality test, were centered and scaled. The samples were split into a training and test set, for 10-fold cross-validation. The classifier used was a support vector machine with a linear kernel. Class weights were balanced. Other parameters were set to default.

Supplementary References

1. Zigmond AS, Snaith RP. The hospital anxiety and depression scale. Acta Psychiatr Scand 1983;67:361–370.

2. Power SE, Jeffery IB, Ross RP, et al. Food and nutrient intake of Irish community-dwelling elderly subjects: who is at nutritional risk? J Nutr Health Aging 2014;18:561–572.

3. Brown JR-M, Flemer B, Joyce SA, et al. Changes in microbiota composition, bile and fatty acid metabolism, in successful faecal microbiota transplantation for Clostridioides difficile infection. BMC Gastroenterol 2018;18:131.

4. Magoč T, Salzberg SL. FLASH: fast length adjustment of short reads to improve genome assemblies. Bioinformatics 2011;27:2957–2963.

5. Edgar RC. Search and clustering orders of magnitude faster than BLAST. Bioinformatics 2010;26:2460–2461.

6. Edgar RC. UPARSE: highly accurate OTU sequences from microbial amplicon reads. Nat Methods 2013;10:996–998.

7. Edgar RC, Haas BJ, Clemente JC, et al. UCHIME improves sensitivity and speed of chimera detection. Bioinformatics 2011;27:2194–2200.

8. Anderson MJ. A new method for non-parametric multivariate analysis of variance: NON-PARAMETRIC MANOVA FOR ECOLOGY. Austral Ecology 2001;26:32–46.

9. Human Microbiome Project Consortium. Structure, function and diversity of the healthy human microbiome. Nature 2012;486:207–214.

10. Franzosa EA, McIver LJ, Rahnavard G, et al. Species-level functional profiling of metagenomes and metatranscriptomes. Nat Methods 2018;15:962–968.

11. Nielsen HB, Almeida M, Juncker AS, et al. Identification and assembly of genomes and genetic elements in complex metagenomic samples without using reference genomes. Nat Biotechnol 2014;32:822–828.

12. Faust K, Sathirapongsasuti JF, Izard J, et al. Microbial co-occurrence relationships in the human microbiome. PLoS Comput Biol 2012;8:e1002606.

13. McKeown NM, Day NE, Welch AA, et al. Use of biological markers to validate self-reported dietary intake in a random sample of the European Prospective Investigation into Cancer United Kingdom Norfolk cohort. Am J Clin Nutr 2001;74:188–196.

14. Savitzky Abraham, Golay MJE. Smoothing and Differentiation of Data by Simplified Least Squares Procedures. Anal Chem 1964;36:1627–1639.

15. Virtanen P, Gommers R, Oliphant TE, et al. SciPy 1.0: fundamental algorithms for scientific computing in Python. Nat Methods 2020;17:261–272.

16. Harris CR, Millman KJ, Walt SJ van der, et al. Array programming with NumPy. Nature 2020;585:357–362.

17. Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. BMC Bioinformatics 2008;9:559.

18. Core Team R. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria 2017:https://www.R-project.org/.

19. Claesson MJ, Jeffery IB, Conde S, et al. Gut microbiota composition correlates with diet and health in the elderly. Nature 2012;488:178–184.

20. Flemer B, Warren RD, Barrett MP, et al. The oral microbiota in colorectal cancer is distinctive and predictive. Gut 2018;67:1454–1463.

21. Friedman J, Hastie T, Tibshirani R. Regularization Paths for Generalized Linear Models via Coordinate Descent. J Stat Soft 2010;33. Available at: http://www.jstatsoft.org/v33/i01/ [Accessed October 10, 2022].

22. Liaw A, Wiener M. Classification and Regression by randomForest. R News 2002;2:18–22.

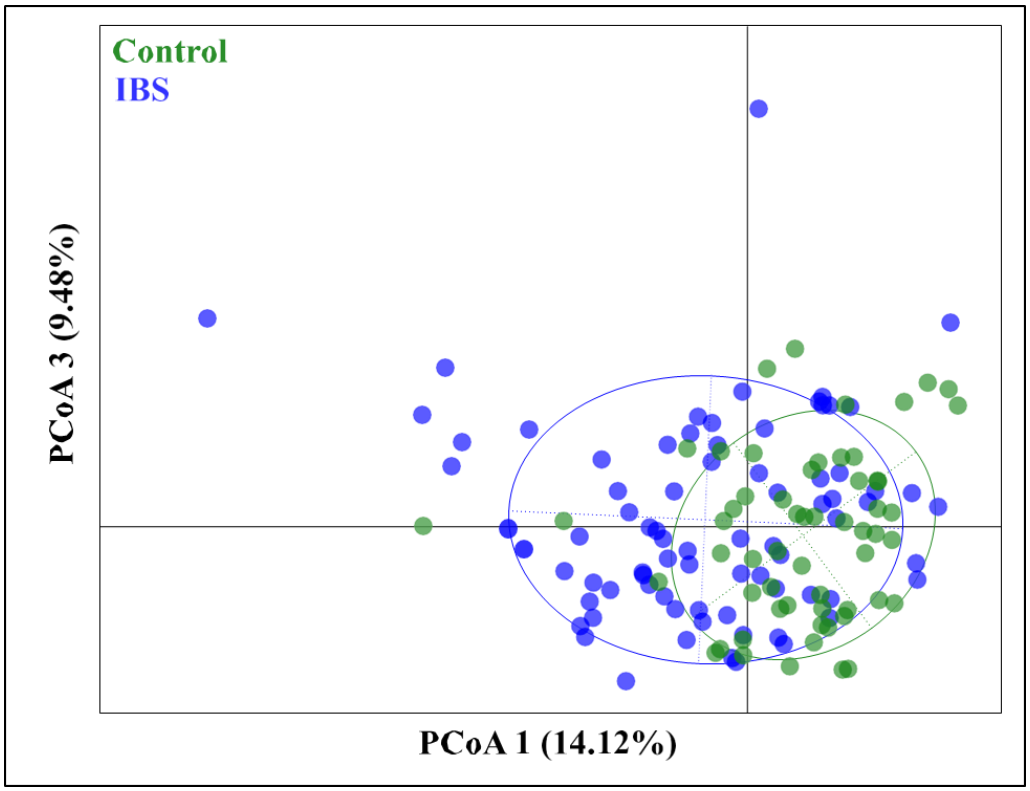
23. Kuhn M. Building Predictive Models in *R* Using the **caret** Package. J Stat Soft 2008;28. Available at: http://www.jstatsoft.org/v28/i05/ [Accessed October 10, 2022].

24. Pedregosa F, Varoquaux G, Gramfort A, et al. Scikit-learn: Machine Learning in Python. Journal of Machine Learning Research 2011;12:2825–2830.

2.8.3 Supplementary Figures

Supplementary Figure 1

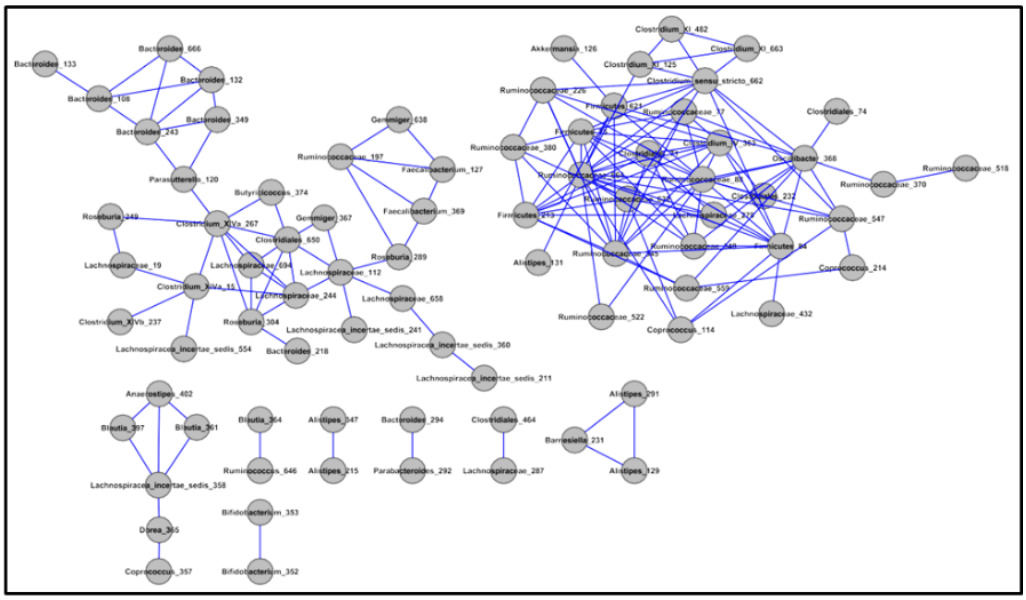
PCoA of microbiota diversity shows a significant difference between the control and IBS groups. PCoA performed using Spearman distance at the shotgun genus level (P = .001; control, n = 58; IBS, n = 78).



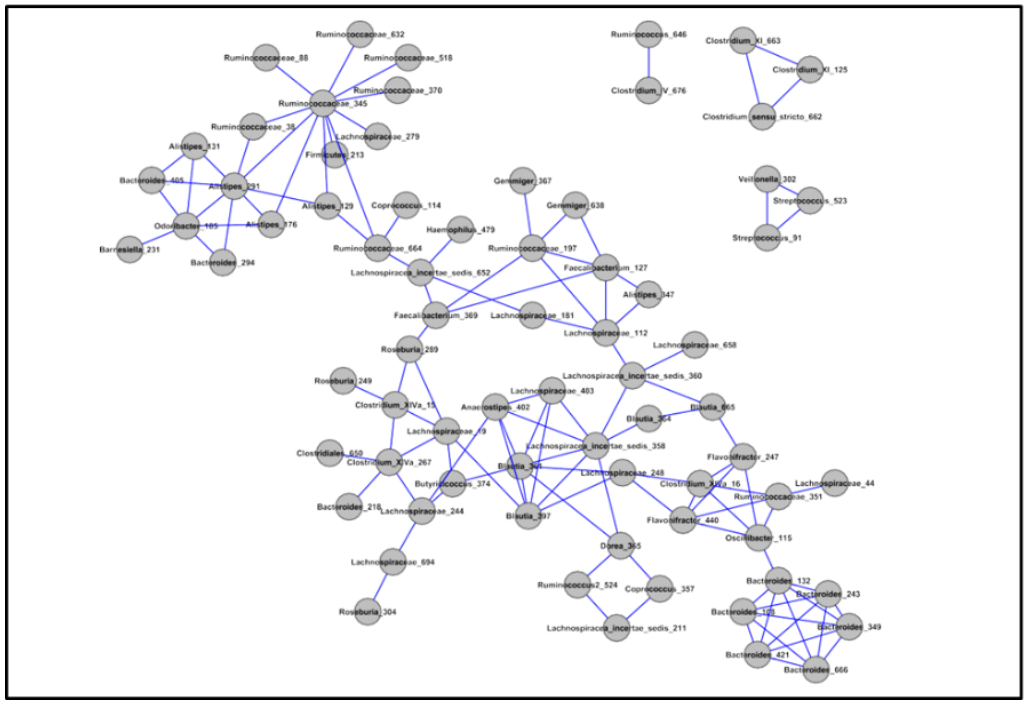
Supplementary Figure 2

Comparison of co-occurrence network of (A) control and (B) IBS groups based on 16S rRNA gene based OTU profile (control, n = 63; IBS, n = 78).

A

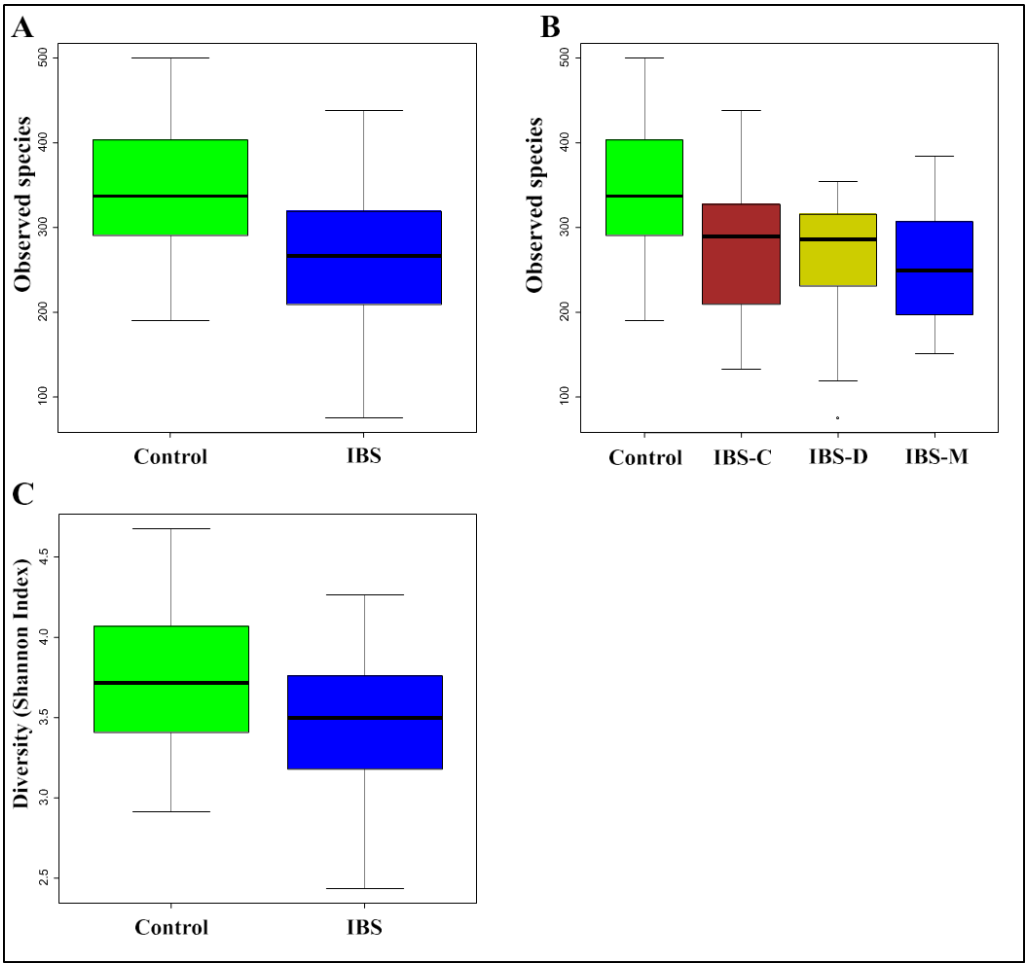


B



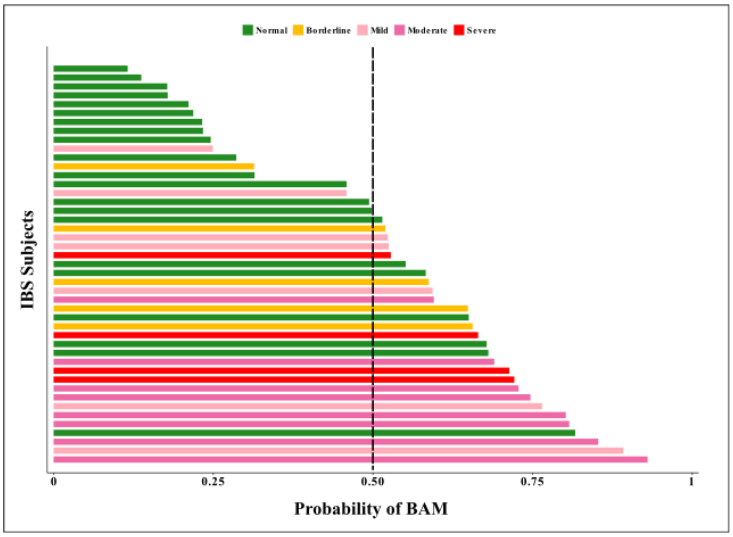
Supplementary Figure 3

Microbiota diversity of the IBS and control groups. (A) The diversity (observed richness) of the IBS group was significantly different from that of the control group based on Wilcoxon rank-sum test (P = 9.215 × 10–8; control, n = 63; IBS, n = 78). (B) The diversities (observed richness) of the IBS clinical subtypes were significantly different from that of the control group based on Kruskal-Wallis (P = 1.2810–6; control, n = 63; IBS-C, n = 29; IBS-D, n = 20; IBS-M, n = 29). (C) The diversity (Shannon index) of the control was significantly different from the IBS group using differences based on Wilcoxon (P = .00032; control, n = 63; IBS, n = 78).



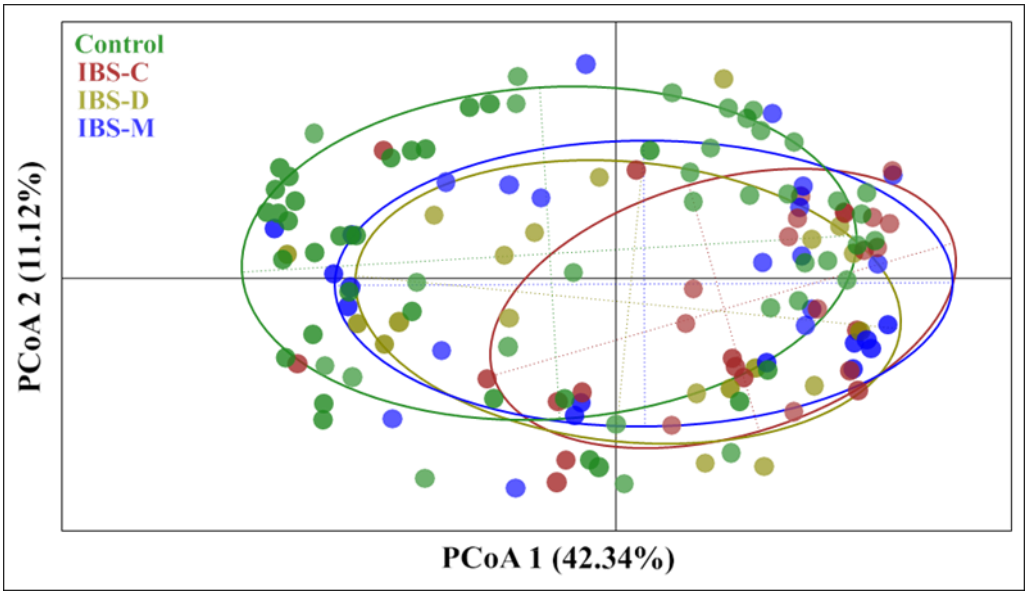
Supplementary Figure 4

Prediction probabilities for BAM, based on faecal metabolites and random forest, on SeHCAT-assayed IBS patients (IBS-BAM, n = 24; IBS non-BAM, n = 21).



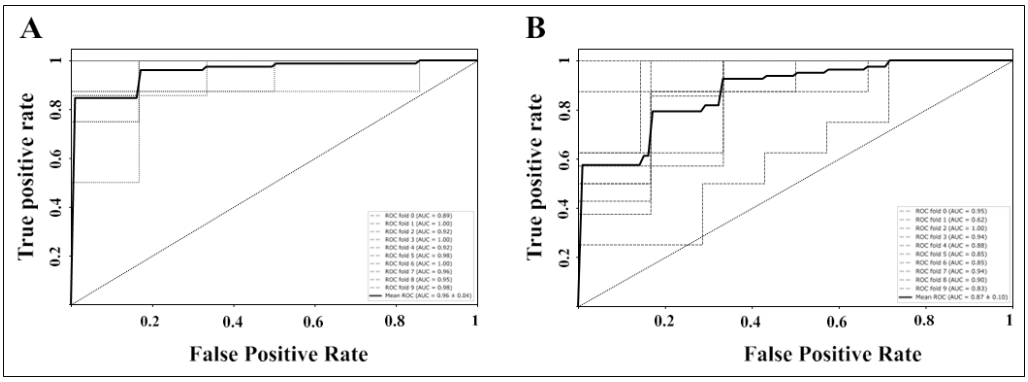
Supplementary Figure 5

PCoA of FAIMS urine metabolomics using Spearman distance shows a significant difference between control and IBS clinical subtypes (Adonis P = .001; control, n = 63; IBS-C, n = 29; IBS-D, n = 20; IBS-M, n = 29).



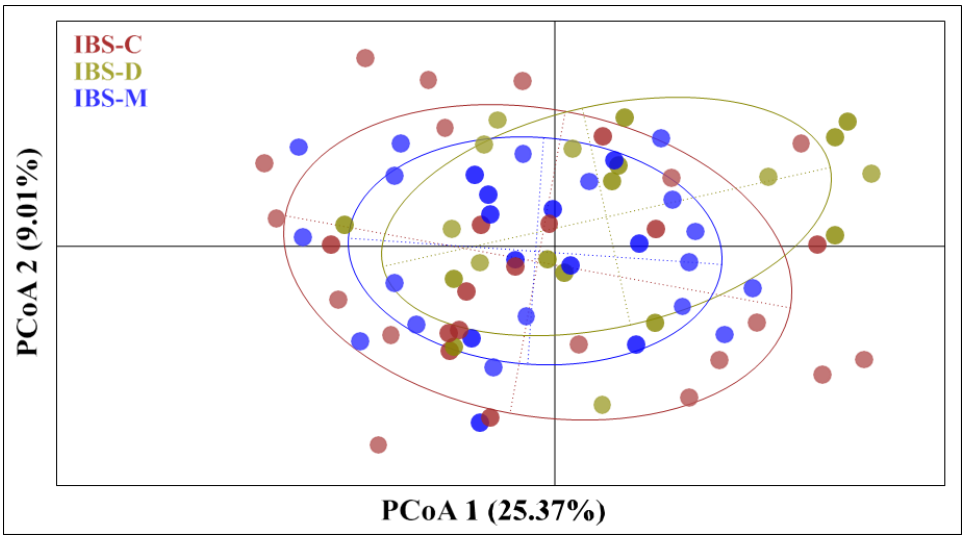
Supplementary Figure 6

Comparison of support vector machine (SVM) classifier performance of IBS and control samples when input is FAIMS or LC/GC-MS metabolomics. Urine metabolomic receiver operating characteristic (ROC) curves to distinguish IBS from control status. (A) ROC curve analysis using an SVM classifier with 10-fold cross-validation on urine LC/GC-MS metabolomics (control, n = 61; IBS, n = 78), where 85% (52/61) of the control group and 95% (74/78) of the IBS group were correctly predicted. (B) ROC curve analysis using an SVM classifier with 10-fold cross-validation on urine FAIMS metabolomics (control, n = 63; IBS, n = 78), where 70% (44/63 of the control group and 83% [65/78] of the IBS group) were correctly predicted.



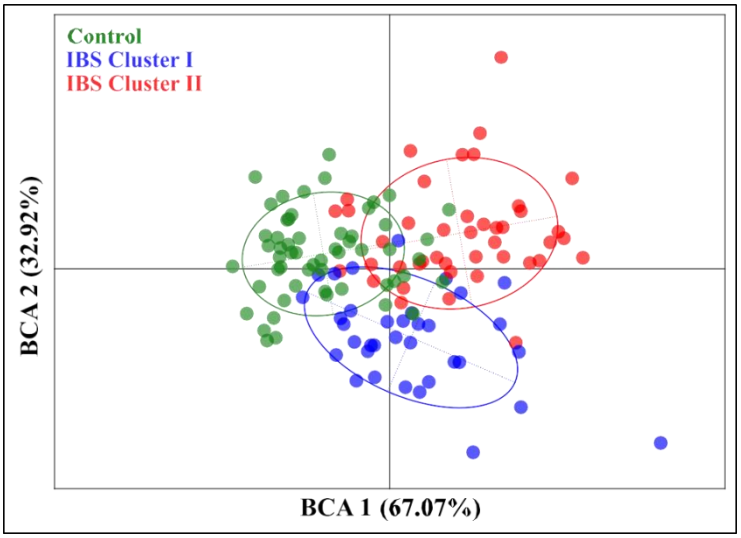
Supplementary Figure 7

PCoA of faecal metabolomics using Spearman distance shows no significant difference between the IBS clinical subtypes (P = .202; IBS-C, n = 29; IBS-D, n = 20; IBS-M, n = 29).



Supplementary Figure 8

Between-class analysis (BCA) showing separation of control individuals and patients with IBS based on microbiome and metabolome data used for WGCNA analysis (control, n = 63; IBS cluster I, n = 35; IBS cluster II, n = 43).



2.8.4 Supplementary Tables

Supplementary Table 1: Strain level (CAG) Machine learning is predictive of IBS versus Control.

	LASSO Optimisation	Random Forest Optimisation	Model Performance
AUC	0.754 (0.146)	0.897 (0.09)	0.814 (0.134)
Sensitivity	0.814 (0.162)	0.95 (0.074)	0.875 (0.102)
Specificity	0.525 (0.241)	0.57 (0.205)	0.497 (0.217)

10-fold Cross Validation

	Predicted IBS	Predicted Control
IBS	70	10
Control	30	29

Rank #	CAG ID	LASSO coefficients	Random Forest feature importance
1	unclassified_00060	0.001381	60.04
2	unclassified_13382	0.068289	57.19
3	Ambiguous_02465	0.010803	55.91
4	unclassified_10544	0.030574	43.01
5	unclassified_01797	0.020433	42.42
6	unclassified_01214	0.001162	40.69
7	unclassified_04033	0.008943	40.54
8	Ambiguous_00664	0.001472	39.75
9	unclassified_07453	0.027742	39.38
10	unclassified_09604	0.025831	38.55
11	unclassified_04421	0.018453	37.31
12	unclassified_02178	0.014262	33.23
13	unclassified_04275	0.022114	32.93
14	unclassified_00992	0.003028	32.52
15	unclassified_08180	0.03671	32.5
16	unclassified_02378	0.00303	30.66
17	unclassified_14410	0.028182	28.63
18	unclassified_14848	0.00442	28.44
19	Escherichia_coli_08281	0.007697	26.73
20	unclassified_01723	0.002795	25.28
21	unclassified_01973	0.003755	23.46
22	unclassified_07490	0.017293	23.05
23	unclassified_04642	0.010974	22.99
24	unclassified_12490	0.041094	22.65
25	unclassified_04705	0.004598	22.01
26	unclassified_01929	0.013678	21.88
27	unclassified_04761	0.025652	21.43
28	unclassified_13688	0.010278	20.66
29	Clostridium_spp_04742	0.005228	19.73
30	Streptococcus_spp_01624	0.001426	19.23
31	unclassified_12615	0.036959	18.59
32	unclassified_10766	0.05376	17.8
33	unclassified_11165	0.035285	17.52
34	unclassified_00496	0.001305	17.34
35	unclassified_07581	0.007595	15.91

36	unclassified_10074	0.012338	15.41
37	unclassified_01227	0.000621	13.73
38	unclassified_01850	0.004519	13.48
39	unclassified_01534	0.001799	12.87
40	unclassified_00657	0.001686	12.77
41	unclassified_03784	0.012933	12.67
42	Streptococcus_anginosus_14524	0.01304	12.16
43	unclassified_04216	0.003356	12.02
44	Parabacteroides_johnsonii_04505	0.007269	11.48
45	unclassified_02737	0.0006	10.34
46	Streptococcus_gordonii_00694	0.00061	10.11
48	Ambiguous_00350	0.011386	10
49	Ambiguous_01019	0.008179	10
50	unclassified_00612	0.004216	10
51	Clostridium_spp_00680	0.003678	10
52	Ambiguous_00176	0.002303	10
53	Ambiguous_00008	0.000835	10
47	Ambiguous_01504	0.000006	10
54	unclassified_07058	0.001504	9.82
55	Clostridium_spp_11230	0.002081	9.47
56	Ambiguous_01105	0.002674	9.4
57	unclassified_02000	0.003605	9.28
58	unclassified_01034	0.005573	9.27
59	unclassified_06517	0.041237	8.95
60	Clostridium_bolteae_00697	0.001039	8.78
61	Turicibacter_sanguinis_07698	0.041323	8.57
62	unclassified_04716	0.004963	8.29
63	unclassified_06120	0.023365	8.22
64	Clostridiales_bacterium_1_7_47FAA_00444	0.000169	8.15
65	unclassified_00404	0.004334	8.15
66	Ambiguous_06054	0.000061	8.14
67	Clostridium_spp_09935	0.008296	8.09
68	unclassified_03271	0.001025	8
69	Ambiguous_03591	0.007581	7.86
70	unclassified_11816	0.030684	7.6
71	Ambiguous_03760	0.004159	7.52
72	Clostridiales_bacterium_1_7_47FAA_00369	0.000864	7.46
73	unclassified_04974	0.003624	7.35
74	Streptococcus_anginosus_02750	0.000721	6.82
75	unclassified_08690	0.003226	6.72
76	unclassified_06706	0.00206	6.56
77	Paraprevotella_xylaniphila_07441	0.00209	6.41
78	unclassified_04992	0.005196	6.09
79	unclassified_08989	0.011704	6.08
80	unclassified_02911	0.002799	6
81	unclassified_02952	0.006054	5.87
82	unclassified_00342	0.000084	5.49
83	Eubacterium_sp_3_1_31_00679	0.001407	5.12
84	Lachnospiraceae_bacterium_5_1_57FAA_01560	0.000291	5
85	Escherichia_coli_01241	0.000114	4.84
86	unclassified_02624	0.002928	4.72
87	Clostridiaceae_bacterium_JC118_03657	0.005134	4.58
88	unclassified_09127	0.001119	4.5
89	unclassified_05532	0.000001	4.48

90	unclassified_09184	0.005517	4.45
91	Bacteroides_spp_03730	0.000523	4.4
92	Paraprevotella_xylaniphila_08998	0.002821	4.3
93	unclassified_03065	0.001211	4.27
94	Ambiguous_01649	0.000779	4.26
95	Streptococcus_mutans_09018	0.005574	4.26
96	Ambiguous_13545	0.00493	4.22
97	unclassified_08505	0.004519	4.12
98	Escherichia_coli_00201	0.000672	3.9
99	unclassified_03041	0.004803	3.78
100	unclassified_05056	0.007699	3.77
101	unclassified_01365	0.000379	3.38
102	Bacteroides_plebeius_08099	0.009286	3.37
103	Ambiguous_05609	0.008937	3.32
104	unclassified_05684	0.00422	3.25
105	unclassified_02242	0.002019	3.21
106	Clostridium_clostridioforme_06211	0.061218	3.16
107	Klebsiella_pneumoniae_01817	0.012099	2.92
108	Clostridium_hathewayi_06002	0.000291	2.87
109	Ambiguous_03727	0.000144	2.8
110	Bacteroides_fragilis_14807	0.011963	2.71
111	unclassified_01340	0.001622	2.66
112	unclassified_08925	0.000758	2.57
113	unclassified_08324	0.000257	2.48
114	Prevotella_disiens_10832	0.004206	2.48
115	Clostridium_leptum_11975	0.002101	2.35
116	unclassified_01283	0.004063	2.09
117	Pseudoflavonifractor_capillosus_03569	0.000849	2.06
118	unclassified_12165	0.006268	2.02
119	unclassified_07203	0.000139	1.84
120	Bacteroides_intestinalis_14747	0.001208	1.73
121	unclassified_08104	0.000055	1.6
122	unclassified_14839	0.000932	1.54
123	Enterococcus_faecalis_01189	0.00061	1.52
124	Streptococcus_infantis_14065	0.00542	1.24
125	Lachnospiraceae_bacterium_1_4_56FAA_13504	0.000698	1.09
126	Alistipes_shahii_15132	0.000646	1.04
127	Clostridium_spp_10114	0.000481	1.03
128	unclassified_13766	0.000045	0.94
129	Ambiguous_06549	0.00035	0.73
130	unclassified_14263	0.00382	0.7
131	Eubacterium_sp_3_1_31_05331	0.001123	0.55
132	Clostridium_asparagiforme_06161	0.000488	0.4
133	Streptococcus_mutans_07592	0.000826	0.33
134	unclassified_12188	0.003405	0.26
135	Clostridium_symbiosum_14754	0.002328	0.17
136	Streptococcus_sanguinis_11557	0.001	0

LASSO and Random Forest (RF) statistics of CAGs predictive of IBS versus Control
Analysis had 2 classes: Control and IBS and included 139 samples (IBS: n= 80 and Control: n=59)
Metrics reported are the mean and the standard deviation of values from Cross Validation.
Taxonomy is assigned where greater than 60% of the gene families are associated with a genus level.
LASSO coefficients are absolute values for the CAG dataset

Supplementary Table 2: Shotgun species differentially abundant between the IBS and Control groups. Significantly different species are shown. For complete table please refer to the published manuscript (doi: 10.1053/j.gastro.2019.11.301).

Species	IBS (IQR)	Control (IQR)	Wilcoxon Statistic	p-value	q-value
Ruminococcus_gnavus	0.0136 (0 - 0.187)	0 (0 - 0)	1209	< 0.001	< 0.001
Clostridium_bolteae	0.016 (0 - 0.0873)	0 (0 - 0.00248)	1189	< 0.001	< 0.001
Clostridiales_bacterium_1_7_47FAA	0 (0 - 0.0122)	0 (0 - 0)	1401	< 0.001	< 0.001
Anaerotruncus_colihominis	0 (0 - 0.0266)	0 (0 - 0)	1457	< 0.001	0.00029
Lachnospiraceae_bacterium_1_4_56FAA	0.000465 (0 - 0.0453)	0 (0 - 0)	1433	< 0.001	0.00029
Flavonifractor_plautii	0.000835 (0 - 0.0266)	0 (0 - 0)	1480.5	< 0.001	0.000871
Clostridium_clostridioforme	0 (0 - 0.0209)	0 (0 - 0)	1612	0.0001	0.000871
Clostridium_hathewayi	0.00177 (0 - 0.0316)	0 (0 - 0)	1468	0.000106	0.000871
Clostridium_symbiosum	0.00164 (0 - 0.0882)	0 (0 - 0)	1515	0.000201	0.00147
Ruminococcus_torques	0.557 (0.266 - 1.33)	0.249 (0.107 - 0.568)	1428	0.000245	0.00161
Alistipes_senegalensis	0 (0 - 0.016)	0.0155 (0 - 0.0885)	3027	0.000365	0.00218
Prevotella_copri	0 (0 - 0)	0 (0 - 0.596)	2835	0.000607	0.00309
Eggerthella_lenta	0 (0 - 0.00447)	0 (0 - 0)	1645.5	0.000612	0.00309
Lachnospiraceae_bacterium_5_1_57FAA	0 (0 - 0)	0 (0 - 0)	1885	0.00116	0.00546
Lachnospiraceae_bacterium_3_1_46FAA	0.0729 (0.0207 - 0.2)	0.0212 (0.00171 - 0.0787)	1534.5	0.00135	0.0059
Clostridium_asparagiforme	0 (0 - 0.0113)	0 (0 - 0)	1651	0.00177	0.00705
Barnesiella_intestinihominis	0.558 (0 - 1.75)	1.41 (0.587 - 2.35)	2968.5	0.00182	0.00705
Clostridium_citroniae	0.00289 (0 - 0.0237)	0 (0 - 0.00399)	1630	0.00194	0.00709
Eubacterium_eligens	0.669 (0.0405 - 1.27)	1.18 (0.395 - 2.12)	2947	0.00258	0.00874
Lachnospiraceae_bacterium_7_1_58FAA	0.0273 (0.0102 - 0.0683)	0.0121 (0.00511 - 0.0273)	1579.5	0.00266	0.00874
Coprococcus_sp_ART55_1	0 (0 - 0)	0 (0 - 4.25)	2817.5	0.00376	0.0118
Lachnospiraceae_bacterium_3_1_57FAA_CT1	0.000675 (0 - 0.0517)	0 (0 - 0.000522)	1675	0.004	0.0119
Clostridium_ramosum	0 (0 - 0)	0 (0 - 0)	1927.5	0.00532	0.0152
Coprococcus_catus	0.238 (0.0985 - 0.426)	0.338 (0.239 - 0.512)	2877	0.0068	0.0186
Eubacterium_biforme	0 (0 - 0.37)	0.222 (0 - 0.86)	2815	0.00721	0.0189
Ruminococcus_lactaris	0 (0 - 0.488)	0.41 (0 - 0.99)	2814.5	0.00986	0.0249

Bacteroides_massiliensis	0 (0 - 0)	0 (0 - 1.19)	2729	0.0108	0.0253
Lachnospiraceae_bacterium_2_1_58FAA	0.00245 (0 - 0.0446)	0 (0 - 0.0101)	1735	0.0111	0.0253
Haemophilus_parainfluenzae	0 (0 - 0.0112)	0.00638 (0 - 0.0493)	2788.5	0.0115	0.0253
Clostridium_nexile	0 (0 - 0.00897)	0 (0 - 0)	1846.5	0.0119	0.0253
Clostridium_innocuum	0 (0 - 0.00333)	0 (0 - 0)	1869.5	0.012	0.0253
Bacteroides_xylanisolvens	0.00587 (0 - 0.103)	0.0561 (0.00379 - 0.163)	2807	0.0144	0.0296
Oxalobacter_formigenes	0 (0 - 0)	0 (0 - 0)	2575	0.0167	0.0332
Alistipes_putredinis	1.29 (0 - 3.26)	3.05 (0.483 - 4.23)	2796.5	0.0177	0.0342
Paraprevotella_clara	0 (0 - 0.014)	0 (0 - 0.179)	2714	0.0192	0.036
Odoribacter_splanchnicus	0.357 (0 - 0.687)	0.573 (0.0488 - 0.883)	2772	0.0217	0.0395
Eubacterium_sp_3_1_31	0 (0 - 0)	0 (0 - 0)	1951	0.0266	0.0472

Shotgun compositional analysis performed on 139 samples (IBS: n= 78 and Control: n=58)
Median abundance % represented as inter-quartile range (IQR)

Supplementary Table 3: Shotgun pathway differentially abundant between IBS and the Control groups. Significantly different features are shown. For complete table please refer to the published manuscript (doi: 10.1053/j.gastro.2019.11.301).

Pathway_Species	Pathway names	IBS (IQR)	Control (IQR)	Wilcoxon Statistic	p-value	q-value
PWY_6700_unclassified	queuosine biosynthesis	0.00641 (0.00467 - 0.0083)	0.0102 (0.0082 - 0.0155)	3496	0	0
NONMEVIPP_PWY_unclassified	methylerythritol phosphate pathway I	0.0124 (0.00846 - 0.015)	0.017 (0.0138 - 0.0199)	3421	0	0.000142
PWY_5667_unclassified	CDP-diacylglycerol biosynthesis I	0.00867 (0.00609 - 0.0115)	0.0129 (0.00984 - 0.0159)	3395	0	0.000142
PWY_6737_unclassified	starch degradation V	0.0158 (0.00983 - 0.02)	0.0221 (0.0166 - 0.0268)	3398	0	0.000142
PWY0_1319_unclassified	CDP-diacylglycerol biosynthesis II	0.00867 (0.00609 - 0.0115)	0.0129 (0.00984 - 0.0159)	3395	0	0.000142
PWY_2942_unclassified	L-lysine biosynthesis III	0.00753 (0.00574 - 0.00975)	0.0113 (0.00881 - 0.014)	3374	0	0.000159
PWY_6387_unclassified	UDP-N-acetylmuramoyl-pentapeptide biosynthesis I (meso-diaminopimelate containing)	0.0155 (0.0115 - 0.0191)	0.022 (0.0168 - 0.0277)	3376	0	0.000159
PWY_724_unclassified	superpathway of L-lysine, L-threonine and L-methionine biosynthesis II	0.00666 (0.00519 - 0.00858)	0.00967 (0.00778 - 0.0117)	3369	0	0.000159
PWY_6386_unclassified	UDP-N-acetylmuramoyl-pentapeptide biosynthesis II (lysine-containing)	0.016 (0.0119 - 0.0197)	0.0227 (0.0174 - 0.0286)	3360	0	0.000166
PWY_6703_unclassified	preQ0 biosynthesis	0.00304 (0.00198 - 0.00418)	0.00503 (0.00351 - 0.00691)	3357	0	0.000166
PWY_5097_unclassified	L-lysine biosynthesis VI	0.00973 (0.00701 - 0.0127)	0.014 (0.0107 - 0.0175)	3340	0	0.000219
PWY0_1296_Clostridium_bolteae	purine ribonucleosides degradation	0 (0 - 0.0000507)	0 (0 - 0)	1467	0	0.00024
UNINTEGRATED_unclassified	UNINTEGRATED	8.68 (6.59 - 9.76)	10.6 (9.11 - 11.8)	3328	0	0.00024
PWY_7187_unclassified	pyrimidine deoxyribonucleotides de novo biosynthesis II	0.00302 (0.00241 - 0.00404)	0.00453 (0.0035 - 0.00555)	3323	0	0.000249
PWY_6124_unclassified	inosine-5'-phosphate biosynthesis II	0.00091 (0.000715 - 0.0013)	0.00153 (0.00111 - 0.00211)	3318	0	0.000258
PEPTIDOGLYCANSYN_PWY_unclassified	peptidoglycan biosynthesis I (meso-diaminopimelate containing)	0.0132 (0.00956 - 0.0165)	0.0179 (0.014 - 0.0252)	3305	0	0.000305
PWY_5686_unclassified	UMP biosynthesis	0.0146 (0.0108 - 0.0187)	0.0198 (0.0161 - 0.0242)	3302	0	0.000305
PWY_6151_unclassified	S-adenosyl-L-methionine cycle I	0.0121 (0.00814 - 0.0148)	0.0159 (0.0129 - 0.0189)	3299	0	0.000305
PWY_7219_unclassified	adenosine ribonucleotides de novo biosynthesis	0.0197 (0.0155 - 0.0268)	0.0308 (0.021 - 0.0373)	3300	0	0.000305
UNINTEGRATED_Ruminococcus_gnavus	UNINTEGRATED	0 (0 - 0.21)	0 (0 - 0)	1431.5	0	0.000326
ANAGLYCOLYSIS_PWY_unclassified	glycolysis III (from glucose)	0.00221 (0.00115 - 0.00326)	0.00375 (0.00288 - 0.00472)	3285.5	0	0.000365
COA_PWY_1_unclassified	coenzyme A biosynthesis I	0.0129 (0.00896 - 0.016)	0.0179 (0.0131 - 0.0211)	3273	0	0.000431
PWY_6123_unclassified	inosine-5'-phosphate biosynthesis I	0.00117 (0.000849 - 0.00167)	0.00198 (0.0014 - 0.00266)	3274	0	0.000431
PWY_5686_Lachnospiraceae_bacterium_7_1_58FAA	UMP biosynthesis	0 (0 - 0.0000914)	0 (0 - 0)	1490	0	0.000471
ASPASN_PWY_unclassified	superpathway of L-aspartate and L-asparagine biosynthesis	0.000953 (0.000508 - 0.0012)	0.00134 (0.000972 - 0.00189)	3245	0	0.000492
COA_PWY_1_Lachnospiraceae_bacterium_7_1_58FAA	coenzyme A biosynthesis I	0 (0 - 0.000138)	0 (0 - 0)	1511	0	0.000492
HISDEG_PWY_unclassified	L-histidine degradation I	0.00142 (0.000609 - 0.00296)	0.00363 (0.00176 - 0.00524)	3239	0	0.000492

PWY_6121_unclassified	5-aminoimidazole ribonucleotide biosynthesis I	0.0133 (0.00989 - 0.0166)	0.0182 (0.0137 - 0.0221)	3248	0	0.000492
PWY_6122_unclassified	5-aminoimidazole ribonucleotide biosynthesis II	0.0137 (0.0103 - 0.0177)	0.0189 (0.0139 - 0.0227)	3240	0	0.000492
PWY_6277_unclassified	superpathway of 5-aminoimidazole ribonucleotide biosynthesis	0.0137 (0.0103 - 0.0177)	0.0189 (0.0139 - 0.0227)	3240	0	0.000492
PWY_6737_Ruminococcus_gnavus	starch degradation V	0 (0 - 0.0000431)	0 (0 - 0)	1571	0	0.000492
PWY_7111_Lachnospiraceae_bacterium_7_1_58FAA	pyruvate fermentation to isobutanol (engineered)	0.0000524 (0 - 0.000133)	0 (0 - 0.000026)	1336	0	0.000492
PWY_7219_Lachnospiraceae_bacterium_7_1_58FAA	adenosine ribonucleotides de novo biosynthesis	0.0000374 (0 - 0.000151)	0 (0 - 0)	1372	0	0.000492
PWY_7219_Ruminococcus_gnavus	adenosine ribonucleotides de novo biosynthesis	0 (0 - 0.0000285)	0 (0 - 0)	1529.5	0	0.000492
PWY_7221_unclassified	guanosine ribonucleotides de novo biosynthesis	0.0135 (0.00938 - 0.0172)	0.0179 (0.0143 - 0.0237)	3256	0	0.000492
PWY0_1296_Ruminococcus_gnavus	purine ribonucleosides degradation	0 (0 - 0.0000432)	0 (0 - 0)	1600	0	0.000492
THRESYN_PWY_unclassified	superpathway of L-threonine biosynthesis	0.00446 (0.00339 - 0.00534)	0.00613 (0.00445 - 0.00756)	3257	0	0.000492
TRNA_CHARGING_PWY_unclassified	tRNA charging	0.0138 (0.0111 - 0.0192)	0.0199 (0.0143 - 0.0263)	3239	0	0.000492
UNINTEGRATED_Clostridium_bolteae	UNINTEGRATED	0 (0 - 0.359)	0 (0 - 0)	1435	0	0.000492
VALSYN_PWY_Lachnospiraceae_bacterium_7_1_58FAA	L-valine biosynthesis	0.0000524 (0 - 0.000133)	0 (0 - 0.000026)	1336	0	0.000492
PWY_841_unclassified	superpathway of purine nucleotides de novo biosynthesis I	0.0018 (0.00142 - 0.00242)	0.00305 (0.00178 - 0.00397)	3237	0	0.000499
PWY_2942_Lachnospiraceae_bacterium_7_1_58FAA	L-lysine biosynthesis III	0 (0 - 0.0000645)	0 (0 - 0)	1524	0	0.000521
PWY_5973_unclassified	cis-vaccenate biosynthesis	0.00195 (0.000762 - 0.00314)	0.00338 (0.00249 - 0.00421)	3231.5	0	0.000521
PWY_7221_Lachnospiraceae_bacterium_7_1_58FAA	guanosine ribonucleotides de novo biosynthesis	0 (0 - 0.0000568)	0 (0 - 0)	1475	0	0.000521
DENOVOPURINE2_PWY_unclassified	superpathway of purine nucleotides de novo biosynthesis II	0.00181 (0.00153 - 0.0026)	0.00318 (0.00192 - 0.00433)	3225	0	0.000576
PWY_6545_unclassified	pyrimidine deoxyribonucleotides de novo biosynthesis III	0.00126 (0.000884 - 0.00184)	0.00221 (0.00158 - 0.00312)	3222	0	0.000596
PWY0_1296_unclassified	purine ribonucleosides degradation	0.0113 (0.00861 - 0.0148)	0.0152 (0.0117 - 0.0209)	3220	0	0.000596
PWY0_166_unclassified	superpathway of pyrimidine deoxyribonucleotides de novo biosynthesis (E. coli)	0.00237 (0.00171 - 0.00372)	0.00396 (0.00313 - 0.0053)	3221	0	0.000596
PWY_7663_unclassified	gondoate biosynthesis (anaerobic)	0.00177 (0.000653 - 0.00263)	0.00284 (0.00208 - 0.00353)	3202	0	0.000826
PWY_5695_unclassified	urate biosynthesis/inosine 5'-phosphate degradation	0.00353 (0.00208 - 0.00572)	0.00623 (0.00364 - 0.00997)	3199	0	0.000857
NONMEVIPP_PWY_Ruminococcus_torques	methylerythritol phosphate pathway I	0.000113 (0 - 0.00033)	0 (0 - 0.000079)	1405.5	0	0.00113
PWY_3001_unclassified	superpathway of L-isoleucine biosynthesis I	0.00515 (0.00409 - 0.0064)	0.00742 (0.00549 - 0.0085)	3180	0	0.00118
PANTOSYN_PWY_unclassified	pantothenate and coenzyme A biosynthesis I	0.0028 (0.00173 - 0.00405)	0.00417 (0.00311 - 0.00545)	3169	0	0.00142
PWY_6386_Lachnospiraceae_bacterium_7_1_58FAA	UDP-N-acetylmuramoyl-pentapeptide biosynthesis II (lysine-containing)	0 (0 - 0.0000798)	0 (0 - 0)	1582	0	0.00145
PWY_6387_Lachnospiraceae_bacterium_7_1_58FAA	UDP-N-acetylmuramoyl-pentapeptide biosynthesis I (meso-diaminopimelate containing)	0 (0 - 0.0000712)	0 (0 - 0)	1582	0	0.00145
PWY_7219_Clostridium_bolteae	adenosine ribonucleotides de novo biosynthesis	0 (0 - 0.0000501)	0 (0 - 0)	1568	0	0.00145
PWY_6122_Ruminococcus_gnavus	5-aminoimidazole ribonucleotide biosynthesis II	0 (0 - 0.0000227)	0 (0 - 0)	1687.5	0	0.00154
PWY_6277_Ruminococcus_gnavus	superpathway of 5-aminoimidazole ribonucleotide biosynthesis	0 (0 - 0.0000227)	0 (0 - 0)	1687.5	0	0.00154
PWY_6737_Clostridium_clostridioforme	starch degradation V	0 (0 - 0.000027)	0 (0 - 0)	1685.5	0	0.00154
UNINTEGRATED_Lachnospiraceae_bacterium_1_4_56FAA	UNINTEGRATED	0 (0 - 0.0687)	0 (0 - 0)	1566.5	0	0.00154
PWY_5188_Lachnospiraceae_bacterium_7_1_58FAA	tetrapyrrole biosynthesis I (from glutamate)	0 (0 - 0.0000684)	0 (0 - 0)	1554	0	0.00155

CALVIN_PWY_unclassified	Calvin-Benson-Bassham cycle	0.00666 (0.00482 - 0.00804)	0.00836 (0.00703 - 0.0101)	3152	0	0.00166
PWY_4984_Lachnospiraceae_bacterium_7_1_58FAA	urea cycle	0 (0 - 0.0000923)	0 (0 - 0)	1510	0	0.00172
PWY_7184_unclassified	pyrimidine deoxyribonucleotides de novo biosynthesis I	0.00137 (0.000978 - 0.00222)	0.00233 (0.00168 - 0.00388)	3149	0	0.00172
PWY_7199_unclassified	pyrimidine deoxyribonucleosides salvage	0.00137 (0.000839 - 0.0022)	0.00215 (0.00147 - 0.00285)	3147	0	0.00174
UNINTEGRATED_Lachnospiraceae_bacterium_2_1_58FAA	UNINTEGRATED	0 (0 - 0.0361)	0 (0 - 0)	1667	0.000108	0.00185
PANTO_PWY_Ruminococcus_gnavus	phosphopantothenate biosynthesis I	0 (0 - 0.0000557)	0 (0 - 0)	1642.5	0.000116	0.00193
PWY_6737_Clostridium_bolteae	starch degradation V	0 (0 - 0.0000318)	0 (0 - 0)	1651	0.000115	0.00193
PWY_6151_Ruminococcus_gnavus	S-adenosyl-L-methionine cycle I	0 (0 - 0.0000345)	0 (0 - 0)	1712	0.00012	0.00198
PWY_6609_Ruminococcus_gnavus	adenine and adenosine salvage III	0 (0 - 0.0000122)	0 (0 - 0)	1718	0.000143	0.00232
PEPTIDOGLYCANSYN_PWY_Lachnospiraceae_bacterium_7_1_58FAA	peptidoglycan biosynthesis I (meso-diaminopimelate containing)	0 (0 - 0.0000774)	0 (0 - 0)	1629	0.000152	0.00243
PWY_7219_Clostridium_symbiosum	adenosine ribonucleotides de novo biosynthesis	0 (0 - 0.0000445)	0 (0 - 0)	1608.5	0.000158	0.00248
PWY_6122_Clostridium_clostridioforme	5-aminoimidazole ribonucleotide biosynthesis II	0 (0 - 0)	0 (0 - 0)	1769	0.000162	0.00249
PWY_6277_Clostridium_clostridioforme	superpathway of 5-aminoimidazole ribonucleotide biosynthesis	0 (0 - 0)	0 (0 - 0)	1769	0.000162	0.00249
UNINTEGRATED_Lachnospiraceae_bacterium_3_1_46FAA	UNINTEGRATED	0.062 (0 - 0.202)	0 (0 - 0.0573)	1481	0.000188	0.00284
PWY_7219_Lachnospiraceae_bacterium_3_1_46FAA	adenosine ribonucleotides de novo biosynthesis	0.0000183 (0 - 0.000202)	0 (0 - 0)	1515.5	0.000203	0.00304
PWY_6125_unclassified	superpathway of guanosine nucleotides de novo biosynthesis II	0.0022 (0.00171 - 0.00324)	0.00331 (0.00255 - 0.00535)	3105	0.00021	0.00309
PWY_5667_Lachnospiraceae_bacterium_7_1_58FAA	CDP-diacylglycerol biosynthesis I	0 (0 - 0.0000801)	0 (0 - 0)	1598.5	0.000226	0.00325
PWY0_1319_Lachnospiraceae_bacterium_7_1_58FAA	CDP-diacylglycerol biosynthesis II	0 (0 - 0.0000801)	0 (0 - 0)	1598.5	0.000226	0.00325
CENTFERM_PWY_unclassified	pyruvate fermentation to butanoate	0 (0 - 0.000114)	0.000128 (0 - 0.000282)	3033	0.000264	0.00361
NONMEVIPP_PWY_Lachnospiraceae_bacterium_3_1_46FAA	methylerythritol phosphate pathway I	0 (0 - 0.000137)	0 (0 - 0)	1587	0.000257	0.00361
PWY_6590_unclassified	superpathway of Clostridium acetobutylicum acidogenic fermentation	0 (0 - 0.000145)	0.000161 (0 - 0.000354)	3034	0.000259	0.00361
PWY_7220_Clostridiales_bacterium_1_7_47FAA	adenosine deoxyribonucleotides de novo biosynthesis II	0 (0 - 0)	0 (0 - 0)	1798	0.000267	0.00361
PWY_7222_Clostridiales_bacterium_1_7_47FAA	guanosine deoxyribonucleotides de novo biosynthesis II	0 (0 - 0)	0 (0 - 0)	1798	0.000267	0.00361
PWY_7219_Clostridium_clostridioforme	adenosine ribonucleotides de novo biosynthesis	0 (0 - 0.00002)	0 (0 - 0)	1723	0.000285	0.0038
UNINTEGRATED_Clostridium_clostridioforme	UNINTEGRATED	0 (0 - 0.238)	0 (0 - 0)	1702.5	0.000338	0.00441
UNINTEGRATED_Prevotella_copri	UNINTEGRATED	0 (0 - 0)	0 (0 - 0.318)	2854	0.000335	0.00441
PWY_7221_Ruminococcus_gnavus	guanosine ribonucleotides de novo biosynthesis	0 (0 - 0)	0 (0 - 0)	1771	0.000343	0.00442
PWY_6386_Ruminococcus_gnavus	UDP-N-acetylmuramoyl-pentapeptide biosynthesis II (lysine-containing)	0 (0 - 0)	0 (0 - 0)	1772	0.000352	0.00447
PWY_6387_Ruminococcus_gnavus	UDP-N-acetylmuramoyl-pentapeptide biosynthesis I (meso-diaminopimelate containing)	0 (0 - 0)	0 (0 - 0)	1773	0.000362	0.00447
PWY_6737_Lachnospiraceae_bacterium_3_1_46FAA	starch degradation V	0 (0 - 0.000114)	0 (0 - 0)	1569.5	0.000362	0.00447
PWY0_1296_Clostridium_clostridioforme	purine ribonucleosides degradation	0 (0 - 0)	0 (0 - 0)	1773	0.000362	0.00447
PWY_7219_Prevotella_copri	adenosine ribonucleotides de novo biosynthesis	0 (0 - 0)	0 (0 - 0.00058)	2850	0.000367	0.00448
PWY_5667_Ruminococcus_gnavus	CDP-diacylglycerol biosynthesis I	0 (0 - 0.0000191)	0 (0 - 0)	1744	0.000383	0.00453
PWY_7111_Clostridium_bolteae	pyruvate fermentation to isobutanol (engineered)	0 (0 - 0.0000345)	0 (0 - 0)	1660.5	0.000376	0.00453
PWY0_1319_Ruminococcus_gnavus	CDP-diacylglycerol biosynthesis II	0 (0 - 0.0000191)	0 (0 - 0)	1744	0.000383	0.00453
NONOXIPENT_PWY_Clostridium_bolteae	pentose phosphate pathway (non-oxidative branch)	0 (0 - 0.0000225)	0 (0 - 0)	1717	0.00039	0.00456
PWY_7229_unclassified	superpathway of adenosine nucleotides de novo biosynthesis I	0.00617 (0.00471 - 0.00799)	0.00829 (0.00609 - 0.0109)	3063	0.000428	0.00492
UNINTEGRATED_Lachnospiraceae_bacterium_7_1_58FAA	UNINTEGRATED	0.157 (0 - 0.239)	0 (0 - 0.155)	1502	0.000429	0.00492
PWY_6737_Lachnospiraceae_bacterium_2_1_58FAA	starch degradation V	0 (0 - 0)	0 (0 - 0)	1827	0.000439	0.00498
PWY_6122_Clostridium_bolteae	5-aminoimidazole ribonucleotide biosynthesis II	0 (0 - 0.0000142)	0 (0 - 0)	1751	0.000459	0.00511
PWY_6277_Clostridium_bolteae	superpathway of 5-aminoimidazole ribonucleotide biosynthesis	0 (0 - 0.0000142)	0 (0 - 0)	1751	0.000459	0.00511
PANTO_PWY_unclassified	phosphopantothenate biosynthesis I	0.00787 (0.00518 - 0.0101)	0.00981 (0.00729 - 0.0139)	3058	0.000465	0.00512

THISYNARA_PWY_unclassified	superpathway of thiamin diphosphate biosynthesis III (eukaryotes)	0.000524 (0.000233 - 0.000888)	0.000835 (0.000431 - 0.00134)	3053	0.000504	0.00551
PWY0_1296_Lachnospiraceae_bacterium_3_1_46FAA	purine ribonucleosides degradation	0 (0 - 0.000154)	0 (0 - 0)	1601.5	0.000569	0.00615
PWY_6121_Ruminococcus_gnavus	5-aminoimidazole ribonucleotide biosynthesis I	0 (0 - 0)	0 (0 - 0)	1802.5	0.000605	0.00642
PWY_7221_Clostridium_clostridioforme	guanosine ribonucleotides de novo biosynthesis	0 (0 - 0)	0 (0 - 0)	1802.5	0.000605	0.00642
PWY_2942_Ruminococcus_gnavus	L-lysine biosynthesis III	0 (0 - 0)	0 (0 - 0)	1772	0.000614	0.00645
PWY_6897_Escherichia_coli	thiamin salvage II	0 (0 - 0.000205)	0 (0 - 0)	1598.5	0.000634	0.00655
UNINTEGRATED_Clostridiales_bacterium_1_7_47FAA	UNINTEGRATED	0 (0 - 0)	0 (0 - 0)	1773	0.00063	0.00655
NONMEVIPP_PWY_Lachnospiraceae_bacterium_7_1_58FAA	methylerythritol phosphate pathway I	0 (0 - 0.0000615)	0 (0 - 0)	1682	0.000698	0.00689
PEPTIDOGLYCANSYN_PWY_Lachnospiraceae_bacterium_3_1_46FAA	peptidoglycan biosynthesis I (meso-diaminopimelate containing)	0 (0 - 0.000136)	0 (0 - 0)	1635	0.000685	0.00689
PWY_6121_Clostridium_bolteae	5-aminoimidazole ribonucleotide biosynthesis I	0 (0 - 0)	0 (0 - 0)	1806.5	0.000675	0.00689
PWY_6163_Lachnospiraceae_bacterium_3_1_46FAA	chorismate biosynthesis from 3-dehydroquinate	0 (0 - 0.000137)	0 (0 - 0)	1636	0.000699	0.00689
PWY_6700_Prevotella_copri	queuosine biosynthesis	0 (0 - 0)	0 (0 - 0.000255)	2815	0.000688	0.00689
PWY_7221_Prevotella_copri	guanosine ribonucleotides de novo biosynthesis	0 (0 - 0)	0 (0 - 0.000402)	2814	0.000704	0.00689
PWY_5097_Prevotella_copri	L-lysine biosynthesis VI	0 (0 - 0)	0 (0 - 0.000589)	2813	0.00072	0.00692
PWY_6121_Clostridium_clostridioforme	5-aminoimidazole ribonucleotide biosynthesis I	0 (0 - 0)	0 (0 - 0)	1856	0.000716	0.00692
ARO_PWY_unclassified	chorismate biosynthesis I	0.0121 (0.00844 - 0.0155)	0.0144 (0.0125 - 0.018)	3030	0.000732	0.00696
PWY_2942_Prevotella_copri	L-lysine biosynthesis III	0 (0 - 0)	0 (0 - 0.000514)	2812	0.000736	0.00696
ANAEROFRUCAT_PWY_unclassified	homolactic fermentation	0.000931 (0.000346 - 0.00226)	0.00178 (0.00105 - 0.00325)	3026	0.000772	0.00703
PWY_6122_Ruminococcus_torques	5-aminoimidazole ribonucleotide biosynthesis II	0.0000587 (0 - 0.000253)	0 (0 - 0.0000671)	1531.5	0.000781	0.00703
PWY_6277_Ruminococcus_torques	superpathway of 5-aminoimidazole ribonucleotide biosynthesis	0.0000587 (0 - 0.000253)	0 (0 - 0.0000671)	1531.5	0.000781	0.00703
PWY_7111_Clostridium_symbiosum	pyruvate fermentation to isobutanol (engineered)	0 (0 - 0.0000329)	0 (0 - 0)	1715	0.000787	0.00703
PWY_7111_Lachnospiraceae_bacterium_1_4_56FAA	pyruvate fermentation to isobutanol (engineered)	0 (0 - 0.0000173)	0 (0 - 0)	1754.5	0.000787	0.00703
VALSYN_PWY_Clostridium_symbiosum	L-valine biosynthesis	0 (0 - 0.0000329)	0 (0 - 0)	1715	0.000787	0.00703
VALSYN_PWY_Lachnospiraceae_bacterium_1_4_56FAA	L-valine biosynthesis	0 (0 - 0.0000173)	0 (0 - 0)	1754.5	0.000787	0.00703
PANTO_PWY_Lachnospiraceae_bacterium_2_1_58FAA	phosphopantothenate biosynthesis I	0 (0 - 0)	0 (0 - 0)	1783	0.000813	0.00721
PANTO_PWY_Lachnospiraceae_bacterium_3_1_46FAA	phosphopantothenate biosynthesis I	0 (0 - 0.000224)	0 (0 - 0)	1603	0.00086	0.00757
PWY_1042_Alistipes_senegalensis	glycolysis IV (plant cytosol)	0 (0 - 0)	0 (0 - 0.0000773)	2807.5	0.00095	0.00776
PWY_1042_unclassified	glycolysis IV (plant cytosol)	0.00674 (0.00521 - 0.0103)	0.00932 (0.00716 - 0.0114)	3015	0.000929	0.00776
PWY_5686_Ruminococcus_gnavus	UMP biosynthesis	0 (0 - 0)	0 (0 - 0)	1829	0.000925	0.00776
PWY_6386_Lachnospiraceae_bacterium_3_1_46FAA	UDP-N-acetylmuramoyl-pentapeptide biosynthesis II (lysine-containing)	0 (0 - 0.000134)	0 (0 - 0)	1641	0.000932	0.00776
PWY_6387_Lachnospiraceae_bacterium_3_1_46FAA	UDP-N-acetylmuramoyl-pentapeptide biosynthesis I (meso-diaminopimelate containing)	0 (0 - 0.000125)	0 (0 - 0)	1642	0.00095	0.00776
PWY_6608_unclassified	guanosine nucleotides degradation III	0.00112 (0.000637 - 0.00162)	0.0016 (0.00113 - 0.00218)	3015	0.000929	0.00776
PWY_6897_unclassified	thiamin salvage II	0.000728 (0.000211 - 0.00193)	0.00191 (0.000663 - 0.00348)	3015	0.000906	0.00776
PWY_7219_Lachnospiraceae_bacterium_2_1_58FAA	adenosine ribonucleotides de novo biosynthesis	0 (0 - 0)	0 (0 - 0)	1830	0.000951	0.00776
PWY0_1296_Clostridiales_bacterium_1_7_47FAA	purine ribonucleosides degradation	0 (0 - 0)	0 (0 - 0)	1830	0.000951	0.00776
PYRIDNUCSYN_PWY_Alistipes_senegalensis	NAD biosynthesis I (from aspartate)	0 (0 - 0)	0 (0 - 0.0000477)	2822	0.000929	0.00776
PEPTIDOGLYCANSYN_PWY_Ruminococcus_gnavus	peptidoglycan biosynthesis I (meso-diaminopimelate containing)	0 (0 - 0)	0 (0 - 0)	1831	0.000977	0.00788
PWY_5097_Ruminococcus_gnavus	L-lysine biosynthesis VI	0 (0 - 0)	0 (0 - 0)	1800	0.000979	0.00788
HISDEG_PWY_Clostridium_symbiosum	L-histidine degradation I	0 (0 - 0.0000378)	0 (0 - 0)	1727	0.00102	0.00819
PWY_6737_Clostridium_nexile	starch degradation V	0 (0 - 0)	0 (0 - 0)	1833	0.00103	0.00819
PWY_7111_Lachnospiraceae_bacterium_3_1_46FAA	pyruvate fermentation to isobutanol (engineered)	0 (0 - 0.000163)	0 (0 - 0)	1630	0.00106	0.00821
UNINTEGRATED_Anaerotruncus_colihominis	UNINTEGRATED	0 (0 - 0.0845)	0 (0 - 0)	1735	0.00104	0.00821
VALSYN_PWY_Lachnospiraceae_bacterium_3_1_46FAA	L-valine biosynthesis	0 (0 - 0.000163)	0 (0 - 0)	1630	0.00106	0.00821

COMPLETE_ARO_PWY_unclassified	superpathway of aromatic amino acid biosynthesis	0.0113 (0.00793 - 0.0146)	0.014 (0.0121 - 0.0171)	3006	0.00107	0.00826
PWY_6126_unclassified	superpathway of adenosine nucleotides de novo biosynthesis II	0.00376 (0.00249 - 0.00626)	0.0063 (0.00407 - 0.00855)	3005	0.00109	0.00833
PWY_7221_Clostridium_symbiosum	guanosine ribonucleotides de novo biosynthesis	0 (0 - 0)	0 (0 - 0)	1805	0.00111	0.00847
SULFATE_CYS_PWY_unclassified	superpathway of sulfate assimilation and cysteine biosynthesis	0 (0 - 0.000316)	0.000348 (0 - 0.000681)	2955	0.00112	0.00851
COA_PWY_1_Ruminococcus_gnavus	coenzyme A biosynthesis I	0 (0 - 0)	0 (0 - 0)	1885	0.00116	0.00869
PWY_7221_Lachnospiraceae_bacterium_2_1_58FAA	guanosine ribonucleotides de novo biosynthesis	0 (0 - 0)	0 (0 - 0)	1885	0.00116	0.00869
UNINTEGRATED_Flavonifractor_plautii	UNINTEGRATED	0 (0 - 0.0595)	0 (0 - 0)	1708	0.0012	0.00892
GLUCONEO_PWY_unclassified	gluconeogenesis I	0 (0 - 0)	0 (0 - 0.000509)	2878	0.00121	0.00894
PEPTIDOGLYCANSYN_PWY_Dorea_formicigenerans	peptidoglycan biosynthesis I (meso-diaminopimelate containing)	0.000118 (0.0000391 - 0.000188)	0.0000523 (0 - 0.0000868)	1538.5	0.00127	0.00919
PWY_5345_unclassified	superpathway of L-methionine biosynthesis (by sulfhydrylation)	0 (0 - 0.000269)	0.000278 (0 - 0.000587)	2907	0.00125	0.00919
PWY_6151_Prevotella_copri	S-adenosyl-L-methionine cycle I	0 (0 - 0)	0 (0 - 0.000419)	2780	0.00127	0.00919
COA_PWY_unclassified	coenzyme A biosynthesis I	0.00195 (0.00102 - 0.00281)	0.00274 (0.00193 - 0.00357)	2993	0.00131	0.00939
PWY_5676_unclassified	acetyl-CoA fermentation to butanoate II	0 (0 - 0.000384)	0.000349 (0 - 0.000738)	2955	0.00132	0.00942
PWY_6163_unclassified	chorismate biosynthesis from 3-dehydroquinate	0.0125 (0.00817 - 0.0161)	0.0148 (0.0126 - 0.0183)	2992	0.00133	0.00942
PWY_6121_Ruminococcus_torques	5-aminoimidazole ribonucleotide biosynthesis I	0.0000675 (0 - 0.000275)	0 (0 - 0.0000799)	1569.5	0.00137	0.00968
PWY_1042_Lachnospiraceae_bacterium_3_1_46FAA	glycolysis IV (plant cytosol)	0 (0 - 0.000148)	0 (0 - 0)	1692	0.00139	0.00972
PWY_1269_Alistipes_senegalensis	CMP-3-deoxy-D-manno-octulosonate biosynthesis I	0 (0 - 0)	0 (0 - 0.0000637)	2813.5	0.00143	0.00996
ARO_PWY_Lachnospiraceae_bacterium_3_1_46FAA	chorismate biosynthesis I	0 (0 - 0.000143)	0 (0 - 0)	1694	0.00144	0.00998
PWY_6386_Dorea_formicigenerans	UDP-N-acetylmuramoyl-pentapeptide biosynthesis II (lysine-containing)	0.000127 (0.0000421 - 0.0002)	0.0000649 (0 - 0.0001)	1544.5	0.00147	0.0101
PWY_6163_Clostridium_symbiosum	chorismate biosynthesis from 3-dehydroquinate	0 (0 - 0)	0 (0 - 0)	1858.5	0.00153	0.0105
COA_PWY_1_Lachnospiraceae_bacterium_3_1_46FAA	coenzyme A biosynthesis I	0 (0 - 0.000126)	0 (0 - 0)	1729	0.00163	0.011
PWY_4242_unclassified	pantothenate and coenzyme A biosynthesis III	0.00153 (0.000733 - 0.00238)	0.00235 (0.00155 - 0.003)	2979	0.00162	0.011
PWY_5690_unclassified	TCA cycle II (plants and fungi)	0.000228 (0.0000986 - 0.000334)	0.000308 (0.000183 - 0.000568)	2976	0.00166	0.011
PWY_7219_Lachnospiraceae_bacterium_1_4_56FAA	adenosine ribonucleotides de novo biosynthesis	0 (0 - 0)	0 (0 - 0)	1861.5	0.00166	0.011
PWY0_1296_Eubacterium_eligens	purine ribonucleosides degradation	0.000272 (0 - 0.000601)	0.00049 (0.000165 - 0.00121)	2976.5	0.00163	0.011
DTDPRHAMSYN_PWY_Coprococcus_catus	dTDP-L-rhamnose biosynthesis I	0 (0 - 0.0000999)	0.0000813 (0 - 0.000126)	2941	0.0017	0.0112
PWY_7219_Alistipes_senegalensis	adenosine ribonucleotides de novo biosynthesis	0 (0 - 0)	0 (0 - 0.0000724)	2841	0.00172	0.0113
PWY_6122_Flavonifractor_plautii	5-aminoimidazole ribonucleotide biosynthesis II	0 (0 - 0.0000315)	0 (0 - 0)	1745.5	0.00175	0.0114
PWY_6277_Flavonifractor_plautii	superpathway of 5-aminoimidazole ribonucleotide biosynthesis	0 (0 - 0.0000315)	0 (0 - 0)	1745.5	0.00175	0.0114
PWY_6703_Barnesiella_intestinihominis	preQ0 biosynthesis	0.0000794 (0 - 0.000487)	0.000364 (0.000113 - 0.000656)	2960	0.00177	0.0114
PWY_7357_Escherichia_coli	thiamin formation from pyrithiamine and oxythiamine (yeast)	0.0000172 (0 - 0.000274)	0 (0 - 0.00002)	1634.5	0.0018	0.0115
PWY_7197_unclassified	pyrimidine deoxyribonucleotide phosphorylation	0.00157 (0.00108 - 0.00214)	0.0021 (0.00148 - 0.00331)	2971	0.00182	0.0116
PWY_7221_Lachnospiraceae_bacterium_3_1_46FAA	guanosine ribonucleotides de novo biosynthesis	0 (0 - 0.000128)	0 (0 - 0)	1678	0.00185	0.0117
NONOXIPENT_PWY_Ruminococcus_gnavus	pentose phosphate pathway (non-oxidative branch)	0 (0 - 0)	0 (0 - 0)	1836	0.00189	0.0118
PWY_6122_Lachnospiraceae_bacterium_3_1_57FAA_CT1	5-aminoimidazole ribonucleotide biosynthesis II	0 (0 - 0.0000249)	0 (0 - 0)	1756	0.0019	0.0118
PWY_6277_Lachnospiraceae_bacterium_3_1_57FAA_CT1	superpathway of 5-aminoimidazole ribonucleotide biosynthesis	0 (0 - 0.0000249)	0 (0 - 0)	1756	0.0019	0.0118
PWY_6527_unclassified	stachyose degradation	0.00523 (0.00393 - 0.0076)	0.00717 (0.00534 - 0.0099)	2969	0.00188	0.0118

COBALSYN_PWY_unclassified	adenosylcobalamin salvage from cobinamide I	0.0009 (0.000478 - 0.00157)	0.00144 (0.000936 - 0.00204)	2967	0.00193	0.0119
PWY_7111_Clostridiales_bacterium_1_7_47FAA	pyruvate fermentation to isobutanol (engineered)	0 (0 - 0)	0 (0 - 0)	1838	0.00199	0.0121
UNINTEGRATED_Clostridium_symbiosum	UNINTEGRATED	0 (0 - 0.174)	0 (0 - 0)	1705	0.00197	0.0121
PWY_5667_Ruminococcus_torques	CDP-diacylglycerol biosynthesis I	0.000305 (0.00014 - 0.000741)	0.000163 (0.0000809 - 0.000322)	1562	0.00207	0.0125
PWY0_1319_Ruminococcus_torques	CDP-diacylglycerol biosynthesis II	0.000305 (0.00014 - 0.000741)	0.000163 (0.0000809 - 0.000322)	1562	0.00207	0.0125
PWY_6121_Lachnospiraceae_bacterium_3_1_46FAA	5-aminoimidazole ribonucleotide biosynthesis I	0 (0 - 0.000138)	0 (0 - 0)	1709.5	0.00214	0.0128
COMPLETE_ARO_PWY_Lachnospiraceae_bacterium_3_1_46FAA	superpathway of aromatic amino acid biosynthesis	0 (0 - 0.000136)	0 (0 - 0)	1721	0.00218	0.0129
PWY_7228_unclassified	superpathway of guanosine nucleotides de novo biosynthesis I	0.00256 (0.00191 - 0.00381)	0.00372 (0.00261 - 0.00549)	2959	0.00218	0.0129
PWY_7383_unclassified	anaerobic energy metabolism (invertebrates, cytosol)	0.00124 (0.000886 - 0.00209)	0.00178 (0.00137 - 0.0024)	2959	0.00218	0.0129
BRANCHED_CHAIN_AA_SYN_PWY_unclassified	superpathway of branched amino acid biosynthesis	0.0073 (0.00558 - 0.0099)	0.0096 (0.00716 - 0.0114)	2957	0.00224	0.0131
PWY_7111_Clostridium_hathewayi	pyruvate fermentation to isobutanol (engineered)	0 (0 - 0.000012)	0 (0 - 0)	1808	0.00224	0.0131
SO4ASSIM_PWY_unclassified	sulfate reduction I (assimilatory)	0 (0 - 0.000189)	0.000167 (0 - 0.000471)	2911	0.00228	0.0133
PWY_5667_Lachnospiraceae_bacterium_3_1_46FAA	CDP-diacylglycerol biosynthesis I	0 (0 - 0.000123)	0 (0 - 0)	1691	0.00234	0.0134
PWY_6121_Flavonifractor_plautii	5-aminoimidazole ribonucleotide biosynthesis I	0 (0 - 0.0000305)	0 (0 - 0)	1787	0.00235	0.0134
PWY0_1319_Lachnospiraceae_bacterium_3_1_46FAA	CDP-diacylglycerol biosynthesis II	0 (0 - 0.000123)	0 (0 - 0)	1691	0.00234	0.0134
UNINTEGRATED_Clostridium_asparagiforme	UNINTEGRATED	0 (0 - 0.0534)	0 (0 - 0)	1772.5	0.00232	0.0134
PWY_6387_Dorea_formicigenerans	UDP-N-acetylmuramoyl-pentapeptide biosynthesis I (meso-diaminopimelate containing)	0.000117 (0.0000381 - 0.000188)	0.0000518 (0 - 0.000099)	1577.5	0.00242	0.0137
HISTSYN_PWY_Bifidobacterium_longum	L-histidine biosynthesis	0 (0 - 0.000348)	0 (0 - 0)	1666.5	0.00245	0.0139
PWY_5686_Prevotella_copri	UMP biosynthesis	0 (0 - 0)	0 (0 - 0.000261)	2741	0.00251	0.014
PWY_6163_Clostridium_bolteae	chorismate biosynthesis from 3-dehydroquinate	0 (0 - 0)	0 (0 - 0)	1888	0.00251	0.014
PWY_7219_Clostridiales_bacterium_1_7_47FAA	adenosine ribonucleotides de novo biosynthesis	0 (0 - 0)	0 (0 - 0)	1888	0.00251	0.014
PWY0_1296_Lachnospiraceae_bacterium_2_1_58FAA	purine ribonucleosides degradation	0 (0 - 0)	0 (0 - 0)	1889	0.00258	0.0143
ANAGLYCOLYSIS_PWY_Alistipes_senegalensis	glycolysis III (from glucose)	0 (0 - 0)	0 (0 - 0.0000535)	2699	0.00274	0.0144
P162_PWY_unclassified	L-glutamate degradation V (via hydroxyglutarate)	0 (0 - 0.000097)	0.000101 (0 - 0.000248)	2878	0.00276	0.0144
PWY_5667_Clostridium_symbiosum	CDP-diacylglycerol biosynthesis I	0 (0 - 0)	0 (0 - 0)	1892	0.00279	0.0144
PWY_6122_Lachnospiraceae_bacterium_3_1_46FAA	5-aminoimidazole ribonucleotide biosynthesis II	0 (0 - 0.000152)	0 (0 - 0)	1685	0.00279	0.0144
PWY_6151_Coprococcus_sp_ART55_1	S-adenosyl-L-methionine cycle I	0 (0 - 0)	0 (0 - 0.00132)	2831	0.0028	0.0144
PWY_6163_Clostridium_clostridioforme	chorismate biosynthesis from 3-dehydroquinate	0 (0 - 0)	0 (0 - 0)	1892	0.00279	0.0144
PWY_6277_Lachnospiraceae_bacterium_3_1_46FAA	superpathway of 5-aminoimidazole ribonucleotide biosynthesis	0 (0 - 0.000152)	0 (0 - 0)	1685	0.00279	0.0144
PWY_6703_Ruminococcus_lactaris	preQ0 biosynthesis	0 (0 - 0.000472)	0.000313 (0 - 0.00112)	2896	0.0027	0.0144
PWY_7111_Eubacterium_eligens	pyruvate fermentation to isobutanol (engineered)	0.000291 (0.0000056 - 0.000699)	0.0006 (0.000222 - 0.00138)	2944	0.00266	0.0144
PWY_7220_unclassified	adenosine deoxyribonucleotides de novo biosynthesis II	0.00178 (0.00117 - 0.00316)	0.00306 (0.00187 - 0.00451)	2943	0.00275	0.0144
PWY_7222_unclassified	guanosine deoxyribonucleotides de novo biosynthesis II	0.00178 (0.00117 - 0.00316)	0.00306 (0.00187 - 0.00451)	2943	0.00275	0.0144
PWY0_1297_Ruminococcus_gnavus	superpathway of purine deoxyribonucleosides degradation	0 (0 - 0)	0 (0 - 0)	1890	0.00265	0.0144
PWY0_1319_Clostridium_symbiosum	CDP-diacylglycerol biosynthesis II	0 (0 - 0)	0 (0 - 0)	1892	0.00279	0.0144
UNINTEGRATED_Clostridium_hathewayi	UNINTEGRATED	0 (0 - 0.135)	0 (0 - 0)	1755	0.00272	0.0144
VALSYN_PWY_Eubacterium_eligens	L-valine biosynthesis	0.000291 (0.0000056 - 0.000699)	0.0006 (0.000222 - 0.00138)	2944	0.00266	0.0144
PWY_6163_Ruminococcus_gnavus	chorismate biosynthesis from 3-dehydroquinate	0 (0 - 0)	0 (0 - 0)	1894	0.00294	0.0151
PWY_7237_Clostridium_symbiosum	myo-, chiro- and scillo-inositol degradation	0 (0 - 0.0000168)	0 (0 - 0)	1805	0.00294	0.0151
PWY_7221_Eubacterium_eligens	guanosine ribonucleotides de novo biosynthesis	0.000313 (0 - 0.000718)	0.00064 (0.000165 - 0.00139)	2934	0.003	0.0153
PWY_7111_Ruminococcus_gnavus	pyruvate fermentation to isobutanol (engineered)	0 (0 - 0.000012)	0 (0 - 0)	1807	0.00307	0.0155
PWY_7219_Eubacterium_eligens	adenosine ribonucleotides de novo biosynthesis	0.000359 (0 - 0.000874)	0.000821 (0.000196 - 0.00177)	2933.5	0.00307	0.0155

UNINTEGRATED_Alistipes_senegalensis	UNINTEGRATED	0 (0 - 0)	0 (0 - 0.0557)	2823	0.00321	0.0161
PWY_7456_Coprococcus_sp_ART55_1	mannan degradation	0 (0 - 0)	0 (0 - 0.0017)	2822	0.00327	0.0163
PWY_5667_Clostridium_bolteae	CDP-diacylglycerol biosynthesis I	0 (0 - 0)	0 (0 - 0)	1842	0.00333	0.0165
PWY_6609_Alistipes_senegalensis	adenine and adenosine salvage III	0 (0 - 0)	0 (0 - 0.0000493)	2720	0.00335	0.0165
PWY0_1319_Clostridium_bolteae	CDP-diacylglycerol biosynthesis II	0 (0 - 0)	0 (0 - 0)	1842	0.00333	0.0165
DTDPRHAMSYN_PWY_Eggerthella_lenta	dTDP-L-rhamnose biosynthesis I	0 (0 - 0)	0 (0 - 0)	1901	0.00354	0.0168
GALACTUROCAT_PWY_unclassified	D-galacturonate degradation I	0.000718 (0.000477 - 0.000989)	0.000953 (0.000665 - 0.00121)	2926	0.00351	0.0168
PANTO_PWY_Coprococcus_sp_ART55_1	phosphopantothenate biosynthesis I	0 (0 - 0)	0 (0 - 0.00103)	2818	0.00349	0.0168
PWY_5659_Coprococcus_sp_ART55_1	GDP-mannose biosynthesis	0 (0 - 0)	0 (0 - 0.00163)	2820.5	0.00357	0.0168
PWY_6151_Barnesiella_intestinihominis	S-adenosyl-L-methionine cycle I	0.000125 (0 - 0.000411)	0.000331 (0.000128 - 0.0006)	2915	0.00349	0.0168
PWY_6151_Eubacterium_eligens	S-adenosyl-L-methionine cycle I	0.000375 (0 - 0.000799)	0.000622 (0.000203 - 0.00149)	2921	0.00356	0.0168
PWY_6305_unclassified	putrescine biosynthesis IV	0.00134 (0.000913 - 0.00206)	0.00181 (0.00136 - 0.00253)	2927	0.00346	0.0168
PWY_7219_Coprococcus_sp_ART55_1	adenosine ribonucleotides de novo biosynthesis	0 (0 - 0)	0 (0 - 0.00158)	2821.5	0.00352	0.0168
PWY_7219_Flavonifractor_plautii	adenosine ribonucleotides de novo biosynthesis	0 (0 - 0.0000552)	0 (0 - 0)	1791.5	0.00342	0.0168
PWY_7221_Ruminococcus_torques	guanosine ribonucleotides de novo biosynthesis	0.0000312 (0 - 0.000251)	0 (0 - 0.0000344)	1648	0.00344	0.0168
TRPSYN_PWY_Coprococcus_sp_ART55_1	L-tryptophan biosynthesis	0 (0 - 0)	0 (0 - 0.00129)	2822.5	0.00346	0.0168
HSERMETANA_PWY_unclassified	L-methionine biosynthesis III	0.000828 (0.000622 - 0.00151)	0.00133 (0.000792 - 0.00222)	2923	0.00366	0.017
NONMEVIPP_PWY_Lachnospiraceae_bacterium_1_1_57FAA	methylerythritol phosphate pathway I	0 (0 - 0)	0 (0 - 0)	1837.5	0.00361	0.017
PWY_5484_unclassified	glycolysis II (from fructose 6-phosphate)	0.000523 (0.000135 - 0.00178)	0.00113 (0.000569 - 0.00222)	2923	0.00363	0.017
PWY_621_Coprococcus_sp_ART55_1	sucrose degradation III (sucrose invertase)	0 (0 - 0)	0 (0 - 0.00264)	2818.5	0.0037	0.0171
UNINTEGRATED_Coprococcus_sp_ART55_1	UNINTEGRATED	0 (0 - 0)	0 (0 - 0.923)	2816.5	0.00382	0.0176
PWY_2942_Coprococcus_sp_ART55_1	L-lysine biosynthesis III	0 (0 - 0)	0 (0 - 0.00118)	2814.5	0.00395	0.0182
GLYCOGENSYNTH_PWY_Coprococcus_sp_ART55_1	glycogen biosynthesis I (from ADP-D-Glucose)	0 (0 - 0)	0 (0 - 0.00155)	2813.5	0.00402	0.0183
THRESYN_PWY_Coprococcus_sp_ART55_1	superpathway of L-threonine biosynthesis	0 (0 - 0)	0 (0 - 0.00116)	2813.5	0.00402	0.0183
COA_PWY_1_Clostridiales_bacterium_1_7_47FAA	coenzyme A biosynthesis I	0 (0 - 0)	0 (0 - 0)	1917.5	0.0041	0.0184
COA_PWY_Clostridiales_bacterium_1_7_47FAA	coenzyme A biosynthesis I	0 (0 - 0)	0 (0 - 0)	1918.5	0.00421	0.0184
GALACT_GLUCUROCAT_PWY_unclassified	superpathway of hexuronide and hexuronate degradation	0.000628 (0.000439 - 0.00109)	0.000997 (0.000653 - 0.00132)	2912.5	0.00423	0.0184
PWY_3001_Coprococcus_sp_ART55_1	superpathway of L-isoleucine biosynthesis I	0 (0 - 0)	0 (0 - 0.00107)	2811.5	0.00415	0.0184
PWY_5667_Clostridium_clostridioforme	CDP-diacylglycerol biosynthesis I	0 (0 - 0)	0 (0 - 0)	1918.5	0.00421	0.0184
PWY_6123_Coprococcus_sp_ART55_1	inosine-5'-phosphate biosynthesis I	0 (0 - 0)	0 (0 - 0.0012)	2811.5	0.00415	0.0184
PWY_7111_Coprococcus_sp_ART55_1	pyruvate fermentation to isobutanol (engineered)	0 (0 - 0)	0 (0 - 0.00099)	2810.5	0.00422	0.0184
PWY_7111_Lachnospiraceae_bacterium_1_1_57FAA	pyruvate fermentation to isobutanol (engineered)	0 (0 - 0.0000348)	0 (0 - 0)	1789	0.00417	0.0184
PWY_7208_Coprococcus_sp_ART55_1	superpathway of pyrimidine nucleobases salvage	0 (0 - 0)	0 (0 - 0.00111)	2811.5	0.00415	0.0184
PWY0_1319_Clostridium_clostridioforme	CDP-diacylglycerol biosynthesis II	0 (0 - 0)	0 (0 - 0)	1918.5	0.00421	0.0184
UDPNAGSYN_PWY_Coprococcus_sp_ART55_1	UDP-N-acetyl-D-glucosamine biosynthesis I	0 (0 - 0)	0 (0 - 0.00137)	2810.5	0.00422	0.0184
VALSYN_PWY_Lachnospiraceae_bacterium_1_1_57FAA	L-valine biosynthesis	0 (0 - 0.0000348)	0 (0 - 0)	1789	0.00417	0.0184
PWY_5097_Coprococcus_sp_ART55_1	L-lysine biosynthesis VI	0 (0 - 0)	0 (0 - 0.0012)	2809.5	0.00429	0.0185
PWY_6700_Coprococcus_sp_ART55_1	queuosine biosynthesis	0 (0 - 0)	0 (0 - 0.00117)	2809.5	0.00429	0.0185
PWY_6527_Coprococcus_sp_ART55_1	stachyose degradation	0 (0 - 0)	0 (0 - 0.00172)	2808.5	0.00436	0.0186
PWY_6737_Clostridium_hathewayi	starch degradation V	0 (0 - 0)	0 (0 - 0)	1880	0.00437	0.0186
PWY0_1296_Clostridium_asparagiforme	purine ribonucleosides degradation	0 (0 - 0)	0 (0 - 0)	1919.5	0.00432	0.0186
PWY_5104_Coprococcus_sp_ART55_1	L-isoleucine biosynthesis IV	0 (0 - 0)	0 (0 - 0.0012)	2807.5	0.00443	0.0187
PWY_6122_Lachnospiraceae_bacterium_2_1_58FAA	5-aminoimidazole ribonucleotide biosynthesis II	0 (0 - 0)	0 (0 - 0)	1920.5	0.00444	0.0187
PWY_6277_Lachnospiraceae_bacterium_2_1_58FAA	superpathway of 5-aminoimidazole ribonucleotide biosynthesis	0 (0 - 0)	0 (0 - 0)	1920.5	0.00444	0.0187
BRANCHED_CHAIN_AA_SYN_PWY_Coprococcus_sp_ART55_1	superpathway of branched amino acid biosynthesis	0 (0 - 0)	0 (0 - 0.00108)	2806.5	0.00451	0.0189
PWY_5103_Coprococcus_sp_ART55_1	L-isoleucine biosynthesis III	0 (0 - 0)	0 (0 - 0.000978)	2806.5	0.00451	0.0189
ILEUSYN_PWY_Coprococcus_sp_ART55_1	L-isoleucine biosynthesis I (from threonine)	0 (0 - 0)	0 (0 - 0.00114)	2804.5	0.00466	0.0193

PYRIDNUCSYN_PWY_Coprococcus_sp_ART55_1	NAD biosynthesis I (from aspartate)	0 (0 - 0)	0 (0 - 0.00108)	2804.5	0.00466	0.0193
VALSYN_PWY_Coprococcus_sp_ART55_1	L-valine biosynthesis	0 (0 - 0)	0 (0 - 0.00114)	2804.5	0.00466	0.0193
PWY_6124_Coprococcus_sp_ART55_1	inosine-5'-phosphate biosynthesis II	0 (0 - 0)	0 (0 - 0.00113)	2803.5	0.00473	0.0195
PWY0_1297_Clostridiales_bacterium_1_7_47FAA	superpathway of purine deoxyribonucleosides degradation	0 (0 - 0)	0 (0 - 0)	1923.5	0.0048	0.0197
NONOXIPENT_PWY_Eubacterium_eligens	pentose phosphate pathway (non-oxidative branch)	0.000394 (0 - 0.000819)	0.000617 (0.000177 - 0.00142)	2899	0.00486	0.0199
GLYCOGENSYNTH_PWY_Eubacterium_biforme	glycogen biosynthesis I (from ADP-D-Glucose)	0 (0 - 0.000268)	0.000125 (0 - 0.000507)	2840	0.00497	0.0202
PWY_6737_Clostridium_symbiosum	starch degradation V	0 (0 - 0)	0 (0 - 0)	1845	0.00501	0.0202
UNINTEGRATED_Eubacterium_eligens	UNINTEGRATED	0.192 (0.0259 - 0.367)	0.324 (0.12 - 0.706)	2900	0.00496	0.0202
VALSYN_PWY_Clostridium_bolteae	L-valine biosynthesis	0 (0 - 0.0000142)	0 (0 - 0)	1823.5	0.00498	0.0202
PWY_7111_unclassified	pyruvate fermentation to isobutanol (engineered)	0.0115 (0.00851 - 0.0145)	0.0136 (0.0108 - 0.0164)	2899	0.0051	0.0205
PWY_5667_Ruminococcus_obeum	CDP-diacylglycerol biosynthesis I	0.0000317 (0 - 0.000122)	0.0000943 (0.0000231 - 0.00024)	2881	0.00517	0.0206
PWY_7219_Barnesiella_intestinihominis	adenosine ribonucleotides de novo biosynthesis	0.000173 (0 - 0.000674)	0.000512 (0.000192 - 0.000716)	2894	0.00513	0.0206
PWY0_1319_Ruminococcus_obeum	CDP-diacylglycerol biosynthesis II	0.0000317 (0 - 0.000122)	0.0000943 (0.0000231 - 0.00024)	2881	0.00517	0.0206
ILEUSYN_PWY_unclassified	L-isoleucine biosynthesis I (from threonine)	0.0117 (0.00851 - 0.0145)	0.0138 (0.0115 - 0.0164)	2897	0.00524	0.0207
VALSYN_PWY_unclassified	L-valine biosynthesis	0.0117 (0.00851 - 0.0145)	0.0138 (0.0115 - 0.0164)	2897	0.00524	0.0207
PWY_1042_Ruminococcus_obeum	glycolysis IV (plant cytosol)	0 (0 - 0)	0 (0 - 0.000149)	2796.5	0.0053	0.0209
PWY_7221_Lachnospiraceae_bacterium_1_4_56FAA	guanosine ribonucleotides de novo biosynthesis	0 (0 - 0)	0 (0 - 0)	1927.5	0.00532	0.0209
PWY0_1298_unclassified	superpathway of pyrimidine deoxyribonucleosides degradation	0.000285 (0.000138 - 0.00042)	0.000382 (0.000236 - 0.000607)	2895.5	0.00535	0.0209
PWY_4984_Flavonifractor_plautii	urea cycle	0 (0 - 0.0000361)	0 (0 - 0)	1823	0.00561	0.0219
X1CMET2_PWY_Bacteroides_massiliensis	N10-formyl-tetrahydrofolate biosynthesis	0 (0 - 0)	0 (0 - 0.00036)	2741	0.00562	0.0219
PWY_7111_Barnesiella_intestinihominis	pyruvate fermentation to isobutanol (engineered)	0.000125 (0 - 0.000436)	0.00033 (0.000123 - 0.000561)	2886	0.00584	0.0225
VALSYN_PWY_Barnesiella_intestinihominis	L-valine biosynthesis	0.000125 (0 - 0.000436)	0.00033 (0.000123 - 0.000561)	2886	0.00584	0.0225
PWY_5103_unclassified	L-isoleucine biosynthesis III	0.00655 (0.00495 - 0.00942)	0.00852 (0.00635 - 0.0104)	2888	0.00592	0.0228
PWY_6471_unclassified	peptidoglycan biosynthesis IV (Enterococcus faecium)	0 (0 - 0)	0 (0 - 0)	1903	0.00604	0.0231
PWY0_1296_Eubacterium_biforme	purine ribonucleosides degradation	0 (0 - 0.000317)	0.000114 (0 - 0.000641)	2824.5	0.00604	0.0231
SALVADEHYPOX_PWY_unclassified	adenosine nucleotides degradation II	0.00109 (0.000787 - 0.00145)	0.00141 (0.00112 - 0.00192)	2886	0.00608	0.0232
PWY_6737_Dorea_formicigenerans	starch degradation V	0.000204 (0.0000983 - 0.000285)	0.000116 (0.0000639 - 0.000195)	1640	0.00619	0.0235
PWY_6121_Lachnospiraceae_bacterium_1_1_57FAA	5-aminoimidazole ribonucleotide biosynthesis I	0 (0 - 0.0000328)	0 (0 - 0)	1829	0.00629	0.0238
PWY_7111_Flavonifractor_plautii	pyruvate fermentation to isobutanol (engineered)	0 (0 - 0.0000355)	0 (0 - 0)	1817	0.00632	0.0238
VALSYN_PWY_Flavonifractor_plautii	L-valine biosynthesis	0 (0 - 0.0000355)	0 (0 - 0)	1817	0.00632	0.0238
PWY_7357_Ruminococcus_obeum	thiamin formation from pyrithiamine and oxythiamine (yeast)	0.0000515 (0 - 0.000208)	0.000128 (0.0000665 - 0.000336)	2869.5	0.00645	0.0242
PANTO_PWY_Ruminococcus_torques	phosphopantothenate biosynthesis I	0.000375 (0.000178 - 0.0007)	0.000285 (0.000106 - 0.000385)	1644	0.00658	0.0245
PWY_1042_Barnesiella_intestinihominis	glycolysis IV (plant cytosol)	0.000203 (0 - 0.000564)	0.000417 (0.000167 - 0.000662)	2875	0.00656	0.0245
ARO_PWY_Clostridium_bolteae	chorismate biosynthesis I	0 (0 - 0)	0 (0 - 0)	1947	0.00667	0.0246
COMPLETE_ARO_PWY_Clostridium_bolteae	superpathway of aromatic amino acid biosynthesis	0 (0 - 0)	0 (0 - 0)	1947	0.00667	0.0246
PWY_6121_Eubacterium_biforme	5-aminoimidazole ribonucleotide biosynthesis I	0 (0 - 0.000322)	0.000154 (0 - 0.000682)	2820	0.0067	0.0246
PWY_7219_Anaerotruncus_colihominis	adenosine ribonucleotides de novo biosynthesis	0 (0 - 0)	0 (0 - 0)	1907	0.00663	0.0246
PWY_5686_Barnesiella_intestinihominis	UMP biosynthesis	0.000113 (0 - 0.000461)	0.000344 (0.00011 - 0.000565)	2871	0.00674	0.0247
PWY_6527_Faecalibacterium_prausnitzii	stachyose degradation	0 (0 - 0)	0 (0 - 0.000995)	2765	0.0069	0.0252

PWY_1042_Lachnospiraceae_bacterium_1_1_57FAA	glycolysis IV (plant cytosol)	0 (0 - 0)	0 (0 - 0)	1910	0.0071	0.0258
PWY_7111_Clostridium_clostridioforme	pyruvate fermentation to isobutanol (engineered)	0 (0 - 0)	0 (0 - 0)	1920	0.00724	0.0261
PWY66_422_Eubacterium_biforme	D-galactose degradation V (Leloir pathway)	0 (0 - 0.000342)	0.000163 (0 - 0.000623)	2815	0.00721	0.0261
VALSYN_PWY_Clostridium_clostridioforme	L-valine biosynthesis	0 (0 - 0)	0 (0 - 0)	1920	0.00724	0.0261
PWY_6386_Lachnospiraceae_bacterium_1_1_57FAA	UDP-N-acetylmuramoyl-pentapeptide biosynthesis II (lysine-containing)	0 (0 - 0)	0 (0 - 0)	1864	0.00739	0.0264
PWY0_1296_Coprococcus_catus	purine ribonucleosides degradation	0.0000464 (0 - 0.000124)	0.0000949 (0.0000532 - 0.000137)	2866	0.0074	0.0264
PWY66_422_unclassified	D-galactose degradation V (Leloir pathway)	0.00592 (0.00417 - 0.00827)	0.0069 (0.0057 - 0.00939)	2871	0.00742	0.0264
UNINTEGRATED_Barnesiella_intestinihominis	UNINTEGRATED	0.162 (0 - 0.485)	0.368 (0.163 - 0.537)	2868.5	0.0074	0.0264
PWY_241_unclassified	C4 photosynthetic carbon assimilation cycle, NADP-ME type	0 (0 - 0)	0 (0 - 0.000143)	2742.5	0.00759	0.0266
PWY_6317_unclassified	galactose degradation I (Leloir pathway)	0.00592 (0.00417 - 0.00827)	0.0069 (0.0057 - 0.00939)	2870	0.00752	0.0266
PWY_6387_Lachnospiraceae_bacterium_1_1_57FAA	UDP-N-acetylmuramoyl-pentapeptide biosynthesis I (meso-diaminopimelate containing)	0 (0 - 0)	0 (0 - 0)	1865	0.00754	0.0266
PWY_7111_Lachnospiraceae_bacterium_2_1_58FAA	pyruvate fermentation to isobutanol (engineered)	0 (0 - 0)	0 (0 - 0)	1952	0.00759	0.0266
VALSYN_PWY_Clostridium_hathewayi	L-valine biosynthesis	0 (0 - 0)	0 (0 - 0)	1887.5	0.00755	0.0266
PWY_4981_unclassified	L-proline biosynthesis II (from arginine)	0.00162 (0.000595 - 0.00345)	0.00283 (0.00117 - 0.0045)	2867	0.00782	0.0273
PWY_7111_Ruminococcus_lactaris	pyruvate fermentation to isobutanol (engineered)	0 (0 - 0.000202)	0.000125 (0 - 0.000338)	2826.5	0.00798	0.0278
PEPTIDOGLYCANSYN_PWY_Lachnospiraceae_bacterium_1_1_57FAA	peptidoglycan biosynthesis I (meso-diaminopimelate containing)	0 (0 - 0)	0 (0 - 0)	1891.5	0.00822	0.0284
PWY_6122_Eubacterium_biforme	5-aminoimidazole ribonucleotide biosynthesis II	0 (0 - 0.000389)	0.000167 (0 - 0.000642)	2805	0.00833	0.0284
PWY_6277_Eubacterium_biforme	superpathway of 5-aminoimidazole ribonucleotide biosynthesis	0 (0 - 0.000389)	0.000167 (0 - 0.000642)	2805	0.00833	0.0284
PWY_6608_Odoribacter_splanchnicus	guanosine nucleotides degradation III	0.0000571 (0 - 0.000193)	0.000171 (0 - 0.000303)	2845	0.00823	0.0284
PWY_6609_Lachnospiraceae_bacterium_2_1_58FAA	adenine and adenosine salvage III	0 (0 - 0)	0 (0 - 0)	1926	0.00833	0.0284
PWY_7219_Bacteroides_massiliensis	adenosine ribonucleotides de novo biosynthesis	0 (0 - 0)	0 (0 - 0.00056)	2742	0.00821	0.0284
PWY_7221_Barnesiella_intestinihominis	guanosine ribonucleotides de novo biosynthesis	0.0000759 (0 - 0.000478)	0.000342 (0.000114 - 0.000576)	2856	0.00834	0.0284
PWY_7221_Flavonifractor_plautii	guanosine ribonucleotides de novo biosynthesis	0 (0 - 0.0000288)	0 (0 - 0)	1851.5	0.00856	0.0291
NONOXIPENT_PWY_Ruminococcus_lactaris	pentose phosphate pathway (non-oxidative branch)	0 (0 - 0)	0 (0 - 0.000296)	2742	0.00878	0.0292
PWY_3841_Bacteroides_massiliensis	folate transformations II	0 (0 - 0)	0 (0 - 0.000393)	2730	0.00867	0.0292
PWY_5667_Barnesiella_intestinihominis	CDP-diacylglycerol biosynthesis I	0.00013 (0 - 0.000513)	0.000316 (0.000143 - 0.000605)	2852	0.00878	0.0292
PWY_6609_Eubacterium_biforme	adenine and adenosine salvage III	0 (0 - 0.000335)	0.000121 (0 - 0.00067)	2799.5	0.00871	0.0292
PWY_7219_Clostridium_hathewayi	adenosine ribonucleotides de novo biosynthesis	0 (0 - 0)	0 (0 - 0)	1872	0.00867	0.0292
PWY_7357_Eubacterium_biforme	thiamin formation from pyrithiamine and oxythiamine (yeast)	0 (0 - 0)	0 (0 - 0.000193)	2706	0.00867	0.0292
PWY0_1319_Barnesiella_intestinihominis	CDP-diacylglycerol biosynthesis II	0.00013 (0 - 0.000513)	0.000316 (0.000143 - 0.000605)	2852	0.00878	0.0292
THISYNARA_PWY_Ruminococcus_obeum	superpathway of thiamin diphosphate biosynthesis III (eukaryotes)	0 (0 - 0.0000872)	0.000078 (0 - 0.000176)	2833	0.00881	0.0292
PWY_7219_Lachnospiraceae_bacterium_1_1_57FAA	adenosine ribonucleotides de novo biosynthesis	0 (0 - 0.0000818)	0 (0 - 0)	1814	0.00884	0.0293
PANTO_PWY_Eggerthella_lenta	phosphopantothenate biosynthesis I	0 (0 - 0)	0 (0 - 0)	1881	0.00902	0.0297
PWY_5121_unclassified	superpathway of geranylgeranyl diphosphate biosynthesis II (via MEP)	0 (0 - 0)	0 (0 - 0.000128)	2657	0.00899	0.0297
PWY_7219_Ruminococcus_torques	adenosine ribonucleotides de novo biosynthesis	0.000353 (0.000193 - 0.000936)	0.000225 (0.000127 - 0.000465)	1669	0.00912	0.0299
PWY_7219_Paraprevotella_clara	adenosine ribonucleotides de novo biosynthesis	0 (0 - 0.0000432)	0 (0 - 0.000387)	2758	0.00918	0.03
VALSYN_PWY_Dorea_formicigenerans	L-valine biosynthesis	0.000131 (0.0000661 - 0.0002)	0.0000863 (0.0000473 - 0.000133)	1671	0.00929	0.0303
PWY_6122_Lachnospiraceae_bacterium_1_1_57FAA	5-aminoimidazole ribonucleotide biosynthesis II	0 (0 - 0.0000438)	0 (0 - 0)	1828	0.00943	0.0306

PWY_6277_Lachnospiraceae_bacterium_1_1_57FAA	superpathway of 5-aminoimidazole ribonucleotide biosynthesis	0 (0 - 0.0000438)	0 (0 - 0)	1828	0.00943	0.0306
PWY_5097_Barnesiella_intestinihominis	L-lysine biosynthesis VI	0.000202 (0 - 0.000581)	0.000436 (0.000171 - 0.000668)	2846	0.00961	0.0311
PWY_2723_Escherichia_coli	trehalose degradation V	0 (0 - 0.0000739)	0 (0 - 0)	1781.5	0.00986	0.0317
PYRIDNUCSYN_PWY_unclassified	NAD biosynthesis I (from aspartate)	0.00198 (0.00123 - 0.00265)	0.00236 (0.0017 - 0.00321)	2849	0.00986	0.0317
PWY_7219_Eubacterium_biforme	adenosine ribonucleotides de novo biosynthesis	0 (0 - 0.000456)	0.000158 (0 - 0.000928)	2792	0.01	0.0321
PWY_6151_Eubacterium_biforme	S-adenosyl-L-methionine cycle I	0 (0 - 0.000483)	0.000196 (0 - 0.0007)	2790	0.0103	0.0329
UNINTEGRATED_Eubacterium_biforme	UNINTEGRATED	0 (0 - 0.192)	0.11 (0 - 0.235)	2790	0.0103	0.0329
PWY_7357_unclassified	thiamin formation from pyrithiamine and oxythiamine (yeast)	0.00307 (0.00178 - 0.00543)	0.00454 (0.00291 - 0.00637)	2845	0.0104	0.033
PWY_5188_Coprococcus_catus	tetrapyrrole biosynthesis I (from glutamate)	0 (0 - 0.0000326)	0.000028 (0 - 0.0000766)	2794.5	0.0106	0.0336
PWY_5686_Ruminococcus_obeum	UMP biosynthesis	0.0000767 (0 - 0.000177)	0.000124 (0.0000616 - 0.000249)	2838	0.0108	0.034
PWY_7111_Clostridium_asparagiforme	pyruvate fermentation to isobutanol (engineered)	0 (0 - 0)	0 (0 - 0)	1890	0.0108	0.034
DTDPRHAMSYN_PWY_Eubacterium_biforme	dTDP-L-rhamnose biosynthesis I	0 (0 - 0.0000352)	0.0000106 (0 - 0.000117)	2759	0.011	0.0346
PWY_1042_Eggerthella_lenta	glycolysis IV (plant cytosol)	0 (0 - 0)	0 (0 - 0)	1939	0.0112	0.0351
PWY_6700_Paraprevotella_clara	queuosine biosynthesis	0 (0 - 0)	0 (0 - 0.000342)	2726.5	0.0112	0.0351
UNINTEGRATED_Clostridium_citroniae	UNINTEGRATED	0 (0 - 0.0845)	0 (0 - 0)	1795	0.0115	0.0358
ARO_PWY_Dorea_formicigenerans	chorismate biosynthesis I	0.0001 (0 - 0.000171)	0.0000586 (0 - 0.0000962)	1701	0.0115	0.0359
PWY_6122_Eubacterium_eligens	5-aminoimidazole ribonucleotide biosynthesis II	0.000314 (0 - 0.000728)	0.000572 (0.00017 - 0.0012)	2833	0.0117	0.0362
PWY_6277_Eubacterium_eligens	superpathway of 5-aminoimidazole ribonucleotide biosynthesis	0.000314 (0 - 0.000728)	0.000572 (0.00017 - 0.0012)	2833	0.0117	0.0362
PWY_6737_Roseburia_inulinivorans	starch degradation V	0.000516 (0.0000853 - 0.00175)	0.000276 (0.0000219 - 0.000721)	1690	0.0117	0.0362
PANTO_PWY_Bacteroides_xylanisolvens	phosphopantothenate biosynthesis I	0 (0 - 0.00015)	0.0000913 (0 - 0.000371)	2797.5	0.0118	0.0365
PWY_2942_Eggerthella_lenta	L-lysine biosynthesis III	0 (0 - 0)	0 (0 - 0)	1917	0.0119	0.0365
PWY_6151_Bacteroides_massiliensis	S-adenosyl-L-methionine cycle I	0 (0 - 0)	0 (0 - 0.0004)	2714.5	0.0119	0.0365
COA_PWY_1_Ruminococcus_torques	coenzyme A biosynthesis I	0.000285 (0.000109 - 0.000644)	0.000164 (0.0000489 - 0.000362)	1693	0.0121	0.0366
PWY_5686_Eubacterium_biforme	UMP biosynthesis	0 (0 - 0.000297)	0.000118 (0 - 0.000531)	2779	0.012	0.0366
PWY_6386_Ruminococcus_torques	UDP-N-acetylmuramoyl-pentapeptide biosynthesis II (lysine-containing)	0.000366 (0.000152 - 0.000861)	0.00023 (0.0000872 - 0.000405)	1691	0.012	0.0366
GLYCOLYSIS_unclassified	glycolysis I (from glucose 6-phosphate)	0.000572 (0.000142 - 0.00214)	0.00128 (0.000583 - 0.00265)	2832	0.0122	0.0368
ARO_PWY_Lachnospiraceae_bacterium_1_1_57FAA	chorismate biosynthesis I	0 (0 - 0)	0 (0 - 0)	1920	0.0127	0.0379
NONMEVIPP_PWY_Paraprevotella_clara	methylerythritol phosphate pathway I	0 (0 - 0)	0 (0 - 0.000313)	2718.5	0.0127	0.0379
PWY_2942_Flavonifractor_plautii	L-lysine biosynthesis III	0 (0 - 0)	0 (0 - 0)	1898	0.0126	0.0379
PWY_6163_Lachnospiraceae_bacterium_1_1_57FAA	chorismate biosynthesis from 3-dehydroquinate	0 (0 - 0)	0 (0 - 0)	1898	0.0126	0.0379
PWY_7219_Eggerthella_lenta	adenosine ribonucleotides de novo biosynthesis	0 (0 - 0)	0 (0 - 0)	1898	0.0126	0.0379
PWY_6121_Eubacterium_eligens	5-aminoimidazole ribonucleotide biosynthesis I	0.000339 (0 - 0.00074)	0.000526 (0.000181 - 0.00124)	2825.5	0.0128	0.0382
TCA_unclassified	TCA cycle I (prokaryotic)	0.0000488 (0 - 0.000226)	0.000173 (0 - 0.000483)	2802	0.0128	0.0382
PWY_5695_Lachnospiraceae_bacterium_3_1_57FAA_CT1	urate biosynthesis/inosine 5'-phosphate degradation	0 (0 - 0)	0 (0 - 0)	1946	0.0131	0.0385
PWY_6387_Barnesiella_intestinihominis	UDP-N-acetylmuramoyl-pentapeptide biosynthesis I (meso-diaminopimelate containing)	0.000121 (0 - 0.000453)	0.000347 (0.000104 - 0.000553)	2818	0.0131	0.0385
PWY_6936_Eubacterium_biforme	seleno-amino acid biosynthesis	0 (0 - 0.00025)	0.0000634 (0 - 0.000465)	2761	0.0131	0.0385
UNINTEGRATED_Roseburia_hominis	UNINTEGRATED	0.23 (0.155 - 0.33)	0.303 (0.227 - 0.417)	2825.5	0.013	0.0385
PWY_2942_Bacteroides_massiliensis	L-lysine biosynthesis III	0 (0 - 0)	0 (0 - 0.000345)	2707.5	0.0133	0.0391
ARGININE_SYN4_PWY_unclassified	L-ornithine de novo biosynthesis	0 (0 - 0.000111)	0.0000637 (0 - 0.000168)	2772.5	0.0135	0.0395
PWY_5100_Eubacterium_biforme	pyruvate fermentation to acetate and lactate II	0 (0 - 0.000376)	0.000138 (0 - 0.000697)	2768	0.0135	0.0395
PWY_6527_Lachnospiraceae_bacterium_3_1_57FAA_CT1	stachyose degradation	0 (0 - 0)	0 (0 - 0)	1902	0.0136	0.0396

GLUCUROCAT_PWY_unclassified	superpathway of β-D-glucuronide and D-glucuronate degradation	0.000751 (0.000425 - 0.0012)	0.00107 (0.000643 - 0.00137)	2822.5	0.0137	0.0397
PWY_6609_Coprococcus_catus	adenine and adenosine salvage III	0.0000683 (0 - 0.000172)	0.000126 (0.0000688 - 0.000197)	2816.5	0.0137	0.0397
PWY_6737_Clostridium_asparagiforme	starch degradation V	0 (0 - 0)	0 (0 - 0)	1931.5	0.0137	0.0397
UNINTEGRATED_Bacteroides_massiliensis	UNINTEGRATED	0 (0 - 0)	0 (0 - 0.428)	2712	0.014	0.0404
PEPTIDOGLYCANSYN_PWY_Barnesiella_intestinihominis	peptidoglycan biosynthesis I (meso-diaminopimelate containing)	0.000123 (0 - 0.000469)	0.000352 (0.0000938 - 0.000593)	2811	0.0143	0.0405
PWY_5667_Ruminococcus_lactaris	CDP-diacylglycerol biosynthesis I	0 (0 - 0.000104)	0.0000533 (0 - 0.00028)	2768.5	0.0143	0.0405
PWY_6386_Barnesiella_intestinihominis	UDP-N-acetylmuramoyl-pentapeptide biosynthesis II (lysine-containing)	0.000133 (0 - 0.000464)	0.000362 (0.000117 - 0.000573)	2811	0.0143	0.0405
PWY_6700_Barnesiella_intestinihominis	queuosine biosynthesis	0.000178 (0 - 0.000563)	0.000401 (0.000155 - 0.00058)	2813.5	0.0142	0.0405
PWY_7282_Bacteroides_fragilis	4-amino-2-methyl-5-phosphomethylpyrimidine biosynthesis (yeast)	0 (0 - 0.000108)	0 (0 - 0)	1852	0.0142	0.0405
PWY0_1319_Ruminococcus_lactaris	CDP-diacylglycerol biosynthesis II	0 (0 - 0.000104)	0.0000533 (0 - 0.00028)	2768.5	0.0143	0.0405
PWY66_399_unclassified	gluconeogenesis III	0.000175 (0 - 0.00044)	0.000336 (0 - 0.000942)	2803	0.0142	0.0405
PANTO_PWY_Paraprevotella_clara	phosphopantothenate biosynthesis I	0 (0 - 0.0000509)	0 (0 - 0.000372)	2728	0.0144	0.0406
PANTO_PWY_Lachnospiraceae_bacterium_1_1_57FAA	phosphopantothenate biosynthesis I	0 (0 - 0.0000863)	0 (0 - 0)	1830	0.0144	0.0407
PWY_5667_Clostridium_nexile	CDP-diacylglycerol biosynthesis I	0 (0 - 0)	0 (0 - 0)	1960	0.0147	0.0412
PWY0_1319_Clostridium_nexile	CDP-diacylglycerol biosynthesis II	0 (0 - 0)	0 (0 - 0)	1960	0.0147	0.0412
PWY_2942_Barnesiella_intestinihominis	L-lysine biosynthesis III	0.00013 (0 - 0.000585)	0.000425 (0.000138 - 0.000612)	2809	0.0149	0.0414
PWY_5855_Escherichia_coli	ubiquinol-7 biosynthesis (prokaryotic)	0 (0 - 0.000103)	0 (0 - 0)	1802.5	0.0151	0.0414
PWY_5856_Escherichia_coli	ubiquinol-9 biosynthesis (prokaryotic)	0 (0 - 0.000103)	0 (0 - 0)	1802.5	0.0151	0.0414
PWY_5857_Escherichia_coli	ubiquinol-10 biosynthesis (prokaryotic)	0 (0 - 0.000103)	0 (0 - 0)	1802.5	0.0151	0.0414
PWY_6708_Escherichia_coli	ubiquinol-8 biosynthesis (prokaryotic)	0 (0 - 0.000103)	0 (0 - 0)	1802.5	0.0151	0.0414
PWY_6737_Lachnospiraceae_bacterium_7_1_58FAA	starch degradation V	0 (0 - 0)	0 (0 - 0)	1893.5	0.0148	0.0414
PWY_7111_Clostridium_citroniae	pyruvate fermentation to isobutanol (engineered)	0 (0 - 0)	0 (0 - 0)	1961	0.015	0.0414
VALSYN_PWY_Clostridium_citroniae	L-valine biosynthesis	0 (0 - 0)	0 (0 - 0)	1961	0.015	0.0414
HISTSYN_PWY_Lachnospiraceae_bacterium_7_1_58FAA	L-histidine biosynthesis	0 (0 - 0)	0 (0 - 0)	1936.5	0.0152	0.0416
ARGSYN_PWY_Escherichia_coli	L-arginine biosynthesis I (via L-ornithine)	0 (0 - 0.000139)	0 (0 - 0)	1839	0.0155	0.0425
CALVIN_PWY_Ruminococcus_torques	Calvin-Benson-Bassham cycle	0.0000569 (0 - 0.000357)	0 (0 - 0.0000924)	1755	0.0157	0.0428
PANTO_PWY_Barnesiella_intestinihominis	phosphopantothenate biosynthesis I	0.000134 (0 - 0.000584)	0.000375 (0.0000885 - 0.000625)	2805	0.0157	0.0428
PWY_7221_Bacteroides_massiliensis	guanosine ribonucleotides de novo biosynthesis	0 (0 - 0)	0 (0 - 0.000451)	2696.5	0.0158	0.0429
UBISYN_PWY_Escherichia_coli	superpathway of ubiquinol-8 biosynthesis (prokaryotic)	0 (0 - 0.000111)	0 (0 - 0)	1813	0.0159	0.0431
PWY_6125_Eggerthella_lenta	superpathway of guanosine nucleotides de novo biosynthesis II	0 (0 - 0)	0 (0 - 0)	1964	0.0161	0.0434
PWY_6737_Lachnospiraceae_bacterium_1_1_57FAA	starch degradation V	0 (0 - 0.0000376)	0 (0 - 0)	1845.5	0.0161	0.0434
PWY0_1296_Roseburia_inulinivorans	purine ribonucleosides degradation	0.000235 (0.0000364 - 0.000771)	0.000104 (0 - 0.000394)	1720	0.0163	0.0439
PWY0_162_unclassified	superpathway of pyrimidine ribonucleotides de novo biosynthesis	0.00204 (0.00166 - 0.00355)	0.00306 (0.00191 - 0.00456)	2808	0.0164	0.044
PWY_5188_Flavonifractor_plautii	tetrapyrrole biosynthesis I (from glutamate)	0 (0 - 0)	0 (0 - 0)	1900.5	0.0168	0.0449
PWY_6609_Eggerthella_lenta	adenine and adenosine salvage III	0 (0 - 0)	0 (0 - 0)	1941.5	0.0168	0.0449
UNINTEGRATED_Lachnospiraceae_bacterium_3_1_57FAA_CT1	UNINTEGRATED	0 (0 - 0.164)	0 (0 - 0)	1840.5	0.017	0.0453
PWY_4981_Eggerthella_lenta	L-proline biosynthesis II (from arginine)	0 (0 - 0)	0 (0 - 0)	1921	0.0172	0.0456
PWY0_1586_Eggerthella_lenta	peptidoglycan maturation (meso-diaminopimelate containing)	0 (0 - 0)	0 (0 - 0)	1921	0.0172	0.0456
PWY_5667_Eubacterium_eligens	CDP-diacylglycerol biosynthesis I	0.000143 (0 - 0.00054)	0.000286 (0.0000566 - 0.000829)	2797	0.0173	0.0457
PWY0_1319_Eubacterium_eligens	CDP-diacylglycerol biosynthesis II	0.000143 (0 - 0.00054)	0.000286 (0.0000566 - 0.000829)	2797	0.0173	0.0457
PWY_5097_Paraprevotella_clara	L-lysine biosynthesis VI	0 (0 - 0)	0 (0 - 0.000309)	2708	0.0175	0.0459

PWY_6163_Flavonifractor_plautii	chorismate biosynthesis from 3-dehydroquinate	0 (0 - 0)	0 (0 - 0)	1943.5	0.0175	0.0459
PWY0_1296_Clostridium_hathewayi	purine ribonucleosides degradation	0 (0 - 0)	0 (0 - 0)	1922	0.0175	0.0459
COA_PWY_1_Roseburia_inulinivorans	coenzyme A biosynthesis I	0.00024 (0 - 0.000831)	0.0000936 (0 - 0.000258)	1732.5	0.0176	0.046
PEPTIDOGLYCANSYN_PWY_Flavonifractor_plautii	peptidoglycan biosynthesis I (meso-diaminopimelate containing)	0 (0 - 0)	0 (0 - 0)	1969	0.0179	0.0466
PWY_5097_Roseburia_hominis	L-lysine biosynthesis VI	0 (0 - 0.000114)	0.0000777 (0 - 0.000217)	2767	0.018	0.0466
PWY_6386_Flavonifractor_plautii	UDP-N-acetylmuramoyl-pentapeptide biosynthesis II (lysine-containing)	0 (0 - 0)	0 (0 - 0)	1969	0.0179	0.0466
PWY_6387_Flavonifractor_plautii	UDP-N-acetylmuramoyl-pentapeptide biosynthesis I (meso-diaminopimelate containing)	0 (0 - 0)	0 (0 - 0)	1969	0.0179	0.0466
UNINTEGRATED_Eggerthella_lenta	UNINTEGRATED	0 (0 - 0)	0 (0 - 0)	1904.5	0.0181	0.0467
PWY_6700_Bacteroides_massiliensis	queuosine biosynthesis	0 (0 - 0)	0 (0 - 0.000379)	2683	0.0182	0.0471
ENTBACSYN_PWY_Escherichia_coli	enterobactin biosynthesis	0 (0 - 0.000363)	0 (0 - 0)	1816.5	0.0185	0.0472
PWY_5667_Bacteroides_xylanisolvens	CDP-diacylglycerol biosynthesis I	0 (0 - 0.00008)	0.0000409 (0 - 0.000238)	2747	0.0185	0.0472
PWY_6936_Eubacterium_ventriosum	seleno-amino acid biosynthesis	0 (0 - 0.000117)	0 (0 - 0.0000282)	1796	0.0185	0.0472
PWY0_1319_Bacteroides_xylanisolvens	CDP-diacylglycerol biosynthesis II	0 (0 - 0.00008)	0.0000409 (0 - 0.000238)	2747	0.0185	0.0472
ILEUSYN_PWY_Dorea_formicigenerans	L-isoleucine biosynthesis I (from threonine)	0.000123 (0 - 0.0002)	0.0000638 (0 - 0.000127)	1737	0.0186	0.0474
PWY_6151_Ruminococcus_lactaris	S-adenosyl-L-methionine cycle I	0 (0 - 0.000141)	0.000105 (0 - 0.000277)	2756	0.0186	0.0474
COMPLETE_ARO_PWY_Lachnospiraceae_bacterium_1_1_57FAA	superpathway of aromatic amino acid biosynthesis	0 (0 - 0)	0 (0 - 0)	1947.5	0.019	0.0482
PWY_6703_Bacteroides_massiliensis	preQ0 biosynthesis	0 (0 - 0)	0 (0 - 0.000264)	2671	0.0193	0.0488
PWY_7111_Bacteroides_massiliensis	pyruvate fermentation to isobutanol (engineered)	0 (0 - 0)	0 (0 - 0.00032)	2679	0.0194	0.0488
VALSYN_PWY_Bacteroides_massiliensis	L-valine biosynthesis	0 (0 - 0)	0 (0 - 0.00032)	2679	0.0194	0.0488
VALSYN_PWY_Ruminococcus_lactaris	L-valine biosynthesis	0 (0 - 0.000131)	0.0000852 (0 - 0.000222)	2755	0.0193	0.0488
PWY_5505_unclassified	L-glutamate and L-glutamine biosynthesis	0 (0 - 0)	0 (0 - 0.0000582)	2614.5	0.0198	0.0493
PWY_5667_Paraprevotella_clara	CDP-diacylglycerol biosynthesis I	0 (0 - 0.0000281)	0 (0 - 0.000278)	2703	0.0197	0.0493
PWY_5695_Roseburia_inulinivorans	urate biosynthesis/inosine 5'-phosphate degradation	0.00028 (0.00005 - 0.000871)	0.000136 (0 - 0.000359)	1737.5	0.0199	0.0493
PWY_6122_Eggerthella_lenta	5-aminoimidazole ribonucleotide biosynthesis II	0 (0 - 0)	0 (0 - 0)	1916	0.0199	0.0493
PWY_6277_Eggerthella_lenta	superpathway of 5-aminoimidazole ribonucleotide biosynthesis	0 (0 - 0)	0 (0 - 0)	1916	0.0199	0.0493
PWY_7221_Lachnospiraceae_bacterium_3_1_57FAA_CT1	guanosine ribonucleotides de novo biosynthesis	0 (0 - 0)	0 (0 - 0)	1929	0.02	0.0493
PWY0_1319_Paraprevotella_clara	CDP-diacylglycerol biosynthesis II	0 (0 - 0.0000281)	0 (0 - 0.000278)	2703	0.0197	0.0493
UNINTEGRATED_Clostridium_nexile	UNINTEGRATED	0 (0 - 0.14)	0 (0 - 0)	1898	0.0198	0.0493
UNINTEGRATED_Paraprevotella_clara	UNINTEGRATED	0 (0 - 0.0495)	0 (0 - 0.292)	2705	0.02	0.0493
PANTO_PWY_Ruminococcus_lactaris	phosphopantothenate biosynthesis I	0 (0 - 0.000108)	0.000092 (0 - 0.000242)	2738	0.0202	0.0496
PWY_5097_Bacteroides_massiliensis	L-lysine biosynthesis VI	0 (0 - 0)	0 (0 - 0.000376)	2684	0.0202	0.0496

Shotgun functional analysis performed on 139 samples (IBS: n= 78 and Control: n=58)
Median abundance % represented as inter-quartile range (IQR)

Supplementary Table 4: Faecal Metabolomics Machine learning is predictive of BAM status in IBS.

	LASSO Optimisation	Random Forest Optimisation	Model Performance
AUC	0.922 (0.016)	0.983 (0.012)	0.828
Sensitivity	0.911 (0.015)	0.958 (0.036)	0.875
Specificity	0.745 (0.046)	0.862 (0.056)	0.667
	Leave One Out Cross Validation		
	Predicted BAM	Predicted Normal	
BAM	21	3	
Normal	7	14	
Rank #	Metabolite ID	LASSO coefficients	Random Forest feature importance
1	Dimethyl benzyl carbinyl butyrate	1.647	100
2	1,3-di-(5Z,8Z,11Z,14Z,17Z-eicosapentaenoyl)-2-hydroxy-glycerol (d5)	0.993	82.66
3	Glu-Glu-Gly-Tyr	-2.298	73.44
4	PG(34:0)	0.673	61.39
5	MG(22:2(13Z,16Z)/0:0/0:0)	-0.93	54.37
6	PC(22:2(13Z,16Z)/15:0)	-0.698	53.75
7	Guanosine	-0.326	42.99
8	Hypoxanthine	-0.031	35.8
9	Thiophanate-methyl	0.408	23.35
10	PC(18:1(9Z)/15:0)	-0.687	15.91
11	Arg-Ile-Ile-Tyr	0.248	14.93
12	Gallic acid	-0.097	12.33
13	PE(22:5(7Z,10Z,13Z,16Z,19Z)/24:0)	-0.241	10.68
14	Momordol	0.01	6.81
15	N-Methylindolo[3,2-b]-5alpha-cholest-2-ene	-0.222	0

LASSO and Random Forest (RF) statistics of metabolites predictive of BAM status
Analysis had 2 classes: BAM and Normal included 40 IBS samples (BAM: n= 24 and Normal: n=21)
Metrics reported are the mean and the standard deviation of values from Cross Validation.

Supplementary Table 5: Urine Metabolomics Machine learning is predictive of IBS versus Control

	LASSO Optimisation	Random Forest Optimisation	Model Performance
AUC	1 (0)	0.999 (0.001)	1 (0)
Sensitivity	0.992 (0.027)	1 (0)	1 (0)
Specificity	0.881 (0.142)	0.976 (0.064)	0.969 (0.066)

10-fold Cross Validation

	Predicted IBS	Predicted Control
IBS	80	0
Control	2	61

LASSO and Random Forest (RF) statistics of metabolites predictive of IBS versus Control

Analysis had 2 classes: Control and IBS and included 143 samples (IBS: n= 80 and Control: n=63)

Metrics reported are the mean and the standard deviation of values from Cross Validation.

Data used was log10 transformed

For all the external cross validation folds, lasso did not return more than 5 features. Therefore, all the trained models are based on random forest with all the features.

Metabolites presented are the most predictive as defined by a AUC of greater than 0.65 when tested on the full dataset (applied as a feature selection methodology).

Rank #	Metabolite ID	AUC (Prediction of/higher in Controls)	Rank #	Metabolite ID	AUC (Prediction of/higher in IBS)
1	Tricetin 3'-methyl ether 7,5'-diglucuronide	0.86	1	A 80987	1
2	Alloathyriol	0.86	2	Medicagenic acid 3-O-b-D-glucuronide	1

3	Torasemide	0.85
4	(-)-Epigallocatechin sulfate	0.8
5	Tetrahydrodipicolinate	0.78
6	Silicic acid	0.75
7	Delphinidin 3-(6"-O-4-malyl-glucosyl)-5-glucoside	0.75
8	Creatinine	0.75
9	L-Arginine	0.74
10	Leucyl-Methionine	0.74
11	Gln-Met-Pro-Ser	0.73
12	Ala-Asn-Cys-Gly	0.72
13	Isoleucyl-Proline	0.71
14	3,4-Methylenesebacic acid	0.71
15	(4-Hydroxybenzoyl)choline	0.71
16	Diazoxide	0.7
17	(1S,3R,4S)-3,4-Dihydroxycyclohexane-1-carboxylate	0.69
18	2-Hydroxypyridine	0.69
19	Ala-Lys-Phe-Cys	0.69
20	3-Methyldioxyindole	0.68
21	N-Carboxyacetyl-D-phenylalanine	0.68
22	Urea	0.67
23	Ferulic acid 4-sulfate	0.67
24	3-Indolehydracrylic acid	0.67
25	Demethyloleuropein	0.67
26	5'-Guanosyl-methylene-triphosphate	0.67
27	Linalyl formate	0.67
28	4-Methoxyphenylethanol sulfate	0.67
29	Allyl nonanoate	0.66
30	D-Galactopyranosyl-(1->3)-D-galactopyranosyl-(1->3)-L-arabinose	0.66

3	N-Undecanoylglycine	0.99
4	Ala-Leu-Trp-Gly	0.98
5	Gamma-glutamyl-Cysteine	0.92
6	Butoctamide hydrogen succinate	0.91
7	(-)-Epicatechin sulfate	0.89
8	1,4,5-Trimethyl-naphtalene	0.86
9	Trp-Ala-Pro	0.83
10	Dodecanedioylcarnitine	0.77
11	1,6,7-Trimethylnaphthalene	0.76
12	Sumiki's acid	0.76
13	Phe-Gly-Gly-Ser	0.75
14	2-hydroxy-2-(hydroxymethyl)-2H-pyran-3(6H)-one	0.73
15	5-((2-iodoacetamido)ethyl)-1-aminonapthalene sulfate	0.72
16	Thiethylperazine	0.72
17	dCTP	0.71
18	Dimethylallylpyrophosphate/Isopentenyl pyrophosphate	0.71
19	Asp-Met-Asp-Pro	0.7
20	3,5-Di-O-galloyl-1,4-galactarolactone	0.7
21	Decanoylcarnitine	0.69
22	[FA (18:0)] N-(9Z-octadecenoyl)-taurine	0.67
23	UDP-4-dehydro-6-deoxy-D-glucose	0.66
24	Delphinidin 3-O-3",6"-O-dimalonylglucoside	0.66
25	Osmundalin	0.65
26	Cysteinyl-Cysteine	0.65

31	Met-Met-Thr-Trp	0.66
32	Cys-Pro-Pro-Tyr	0.66
33	methylphosphonate	0.66
34	2-Phenylethyl octanoate	0.66
35	Hippuric acid	0.65
36	Glutarylcarnitine	0.65
37	Cys-Phe-Phe-Gln	0.65

Supplementary Table 6: Urine metabolites differentially abundant between the IBS and Control groups. Significantly different metabolites are shown. For complete table please refer to the published manuscript (doi: 10.1053/j.gastro.2019.11.301).

Metabolite	IBS (A.U.)	Control (A.U.)	Log2FC	Wilcoxon Statistic	p-value	q-value
N-Undecanoylglycine	212.2	16.5	3.686	28	< 0.001	< 0.001
Gamma-glutamyl-Cysteine	614.2	84.8	2.856	410	< 0.001	< 0.001
Alloathyriol	1.5	453.1	-8.265	4101	< 0.001	< 0.001
Trp-Ala-Pro	6.1	0.2	4.646	763.5	< 0.001	< 0.001
A 80987	730.8	0.1	12.885	0	< 0.001	< 0.001
Medicagenic acid 3-O-b-D-glucuronide	475.4	12.8	5.212	0	< 0.001	< 0.001
Ala-Leu-Trp-Gly	420.3	120.5	1.802	83	< 0.001	< 0.001
Butoctamide hydrogen succinate	319	3.1	6.677	423	< 0.001	< 0.001
(-)-Epicatechin sulfate	274	209.8	0.385	506	< 0.001	< 0.001
1,4,5-Trimethyl-naphtalene	15.2	0	8.739	658.5	< 0.001	< 0.001
Tricetin 3'-methyl ether 7,5'-diglucuronide	0.6	22.5	-5.156	4094	< 0.001	< 0.001
Torasemide	0.5	38.2	-6.289	4023	< 0.001	< 0.001
(-)-Epigallocatechin sulfate	129.1	165.2	-0.356	3826	< 0.001	< 0.001
Dodecanedioylcarnitine	61.9	9.7	2.679	1054	< 0.001	< 0.001
1,6,7-Trimethylnaphthalene	17.2	0.1	7.234	1082.5	< 0.001	< 0.001
Tetrahydrodipicolinate	1.8	71.6	-5.324	3671	< 0.001	< 0.001

Sumiki's acid	84667.1	58728.8	0.528	1181	< 0.001	< 0.001
Silicic acid	4	734.3	-7.527	3556	< 0.001	< 0.001
Delphinidin 3-(6"-O-4-malyl-glucosyl)-5-glucoside	0.2	16.2	-6.341	3548	< 0.001	< 0.001
L-Arginine	0	13.7	-8.547	3540	< 0.001	< 0.001
Leucyl-Methionine	9.5	60.7	-2.682	3526	< 0.001	< 0.001
Phe-Gly-Gly-Ser	420	359.6	0.224	1250	< 0.001	< 0.001
Gln-Met-Pro-Ser	179.8	272.8	-0.601	3507	< 0.001	< 0.001
Creatinine	729604.9	752607.5	-0.045	3500	< 0.001	< 0.001
Ala-Asn-Cys-Gly	177.5	229.7	-0.372	3431	< 0.001	< 0.001
2-hydroxy-2-(hydroxymethyl)-2H-pyran-3(6H)-one	508.8	256	0.991	1329	< 0.001	< 0.001
Thiethylperazine	38	9.7	1.974	1365	< 0.001	< 0.001
5-((2-iodoacetamido)ethyl)-1-aminonapthalene sulfate	627.5	257.7	1.284	1366.5	< 0.001	< 0.001
dCTP	379	323.1	0.231	1390	< 0.001	< 0.001
Isoleucyl-Proline	10391.2	12988.5	-0.322	3362	< 0.001	< 0.001
3,4-Methylenesebacic acid	452826.1	482052	-0.09	3344	< 0.001	< 0.001
Dimethylallylpyrophosphate/Isopentenyl pyrophosphate	15680	9743.9	0.686	1425	< 0.001	< 0.001
(4-Hydroxybenzoyl)choline	68.6	112.2	-0.711	3329	< 0.001	< 0.001
Diazoxide	145.7	212	-0.541	3318	< 0.001	< 0.001
3,5-Di-O-galloyl-1,4-galactarolactone	638.5	539.3	0.243	1458	< 0.001	< 0.001
2-Hydroxypyridine	37.8	164.2	-2.121	3300	< 0.001	< 0.001
Decanoylcarnitine	152.9	46.7	1.71	1463	< 0.001	< 0.001
Asp-Met-Asp-Pro	894.5	744.1	0.266	1473	< 0.001	< 0.001
3-Methyldioxyindole	203	326	-0.683	3250	0.00022	0.00161
(1S,3R,4S)-3,4-Dihydroxycyclohexane-1-carboxylate	749.3	1010.2	-0.431	3244	0.000243	0.00173
Ala-Lys-Phe-Cys	47.3	107.7	-1.186	3238	0.000269	0.00187
3-Indolehydracrylic acid	972.6	1898.6	-0.965	3216	0.000385	0.00261
[FA (18:0)] N-(9Z-octadecenoyl)-taurine	197	178.2	0.145	1545	0.000404	0.00267
Ferulic acid 4-sulfate	1569	3452.2	-1.138	3174	0.000746	0.00482
Urea	188415	198969.2	-0.079	3172	0.000769	0.00487
N-Carboxyacetyl-D-phenylalanine	307.4	438.4	-0.512	3166	0.000843	0.00522

4-Methoxyphenylethanol sulfate	476.3	889.1	-0.9	3155	0.000996	0.00604
UDP-4-dehydro-6-deoxy-D-glucose	192.4	171.7	0.164	1606	0.00104	0.00606
Linalyl formate	20.8	30.6	-0.555	3153	0.00103	0.00606
Demethyloleuropein	9.1	21.5	-1.233	3148	0.00111	0.0063
5'-Guanosyl-methylene-triphosphate	337.4	428.7	-0.346	3140	0.00125	0.00683
Allyl nonanoate	18.4	24	-0.385	3140	0.00125	0.00683
2-Phenylethyl octanoate	67.9	184.7	-1.444	3132	0.0014	0.00754
beta-Cellobiose	163.4	117.1	0.48	1628	0.00145	0.00762
D-Galactopyranosyl-(1->3)-D-galactopyranosyl-(1->3)-L-arabinose	271.6	756.8	-1.479	3125	0.00156	0.00805
Cys-Phe-Phe-Gln	41.1	62	-0.593	3114	0.00182	0.00927
Hippuric acid	89463.1	125800	-0.492	3108	0.00199	0.00993
Cys-Pro-Pro-Tyr	51.1	73.6	-0.527	3098	0.00229	0.0112
Met-Met-Thr-Trp	112	151.5	-0.436	3085	0.00275	0.0132
methylphosphonate	476.1	515.7	-0.115	3084	0.00279	0.0132
3'-Sialyllactosamine	84.8	129.1	-0.606	3082	0.00287	0.0134
2,4,6-Octatriynoic acid	1438.5	1703.3	-0.244	3079	0.00299	0.0137
Delphinidin 3-O-3",6"-O-dimalonylglucoside	229.7	164.6	0.481	1681	0.00307	0.0139
L-Valine	8240.3	7936.7	0.054	1685	0.00325	0.0142
Met-Met-Cys	192.1	163.5	0.233	1685	0.00325	0.0142
Cysteinyl-Cysteine	14357	11017.4	0.382	1687	0.00334	0.0144
(all-E)-1,8,10-Heptadecatriene-4,6-diyne-3,12-diol	378	788.6	-1.061	3068	0.00348	0.0145
L-Lysine	135.9	76.8	0.823	1689	0.00343	0.0145
Pivaloylcarnitine	1262.9	1788	-0.502	3059	0.00393	0.0159
Lenticin	113	217.8	-0.946	3059	0.00393	0.0159
Phenol glucuronide	405.7	287.8	0.495	1701	0.00403	0.0159
Tyrosyl-Cysteine	957.9	802.2	0.256	1705	0.00426	0.0159
Osmundalin	533.8	317.1	0.751	1703	0.00414	0.0159
Tetrahydroaldosterone-3-glucuronide	781.6	975.4	-0.32	3054	0.0042	0.0159
N-Methylpyridinium	3882.3	13043	-1.748	3055	0.00414	0.0159
L-prolyl-L-proline	3080.2	5296.2	-0.782	3056	0.00409	0.0159

Glutarylcarnitine	698.7	864.8	-0.308	3042	0.00492	0.018
[FA (15:4)] 6,8,10,12-pentadecatetraenal	2303.5	3781	-0.715	3042	0.00492	0.018
Methyl bisnorbiotinyl ketone	2259.7	1986.7	0.186	1720	0.00519	0.0187
Acetoin	1239.3	785.6	0.658	1726	0.00561	0.02
LysoPC(18:2(9Z,12Z))	0.8	48.3	-5.859	3029	0.00584	0.0205
Hexyl 2-furoate	17	24.7	-0.537	3021	0.00647	0.0225
N-carbamoyl-L-glutamate	331.9	423.8	-0.353	3018	0.00673	0.0231
L-Homoserine	4000.7	5333.1	-0.415	3012	0.00726	0.0246
L-Asparagine	300	384.8	-0.359	3011	0.00736	0.0246
Tiglylcarnitine	314.5	762.4	-1.278	3008	0.00764	0.025
Thymine	110	76.8	0.519	1751	0.00774	0.025
3-hydroxypyridine	271.4	556.5	-1.036	3007	0.00774	0.025
Menadiol disuccinate	793.6	2024.6	-1.351	3005	0.00793	0.0254
9-Decenoylcarnitine	1951.1	2609.8	-0.42	2996	0.00888	0.0275
Pyrocatechol sulfate	27377.5	40427.9	-0.562	2996	0.00888	0.0275
sedoheptulose anhydride	4159	10851.9	-1.384	2995	0.00899	0.0275
(+)-gamma-Hydroxy-L-homoarginine	272.2	398.3	-0.549	2997	0.00877	0.0275
Thioridazine	884.1	1048.3	-0.246	2984	0.0103	0.0312
Cys-Glu-Glu-Glu	37.3	56.8	-0.609	2977	0.0112	0.0329
Marmesin rutinoside	17.8	36.1	-1.025	2977	0.0112	0.0329
L-Serine	991.6	1146.2	-0.209	2978	0.0111	0.0329
L-Urobilinogen	8.5	139.7	-4.035	2976	0.0113	0.033
Isobutyrylglycine	2274.1	2694.4	-0.245	2974	0.0116	0.0334
S-Adenosylhomocysteine	135.5	454.5	-1.746	2968	0.0125	0.0356
2,3-dioctanoylglyceramide	887.1	1277.1	-0.526	2966	0.0128	0.0357
3-Methoxy-4-hydroxyphenylglycol glucuronide	0.5	10.7	-4.335	2966.5	0.0127	0.0357
sulfoethylcysteine	5602.3	8425.3	-0.589	2965	0.013	0.0358
Hydroxyphenylacetylglycine	460.1	568.5	-0.305	2962	0.0134	0.0367
Pyrroline hydroxycarboxylic acid	13972.6	16170.5	-0.211	2961	0.0136	0.0368

1-(alpha-Methyl-4-(2-methylpropyl)benzeneacetate)-beta-D-Glucopyranuronic acid	131.6	259.6	-0.98	2956	0.0144	0.0383
2-Methylbutylacetate	1958.5	2726.3	-0.477	2956	0.0144	0.0383
N1-Methyl-4-pyridone-3-carboxamide	6162.4	9041.6	-0.553	2955	0.0146	0.0384
Cortolone-3-glucuronide	520.3	620.8	-0.255	2953	0.0149	0.039
Asn-Cys-Gly	255.4	231	0.145	1813	0.0164	0.0413
N6,N6,N6-Trimethyl-L-lysine	2282.7	2591.3	-0.183	2946	0.0162	0.0413
Benzylamine	66.3	218.7	-1.722	2947	0.016	0.0413
5-Hydroxy-L-tryptophan	177.9	218.1	-0.294	2945	0.0164	0.0413
Armillaric acid	25	44.5	-0.833	2941	0.0172	0.0429
Leucine/Isoleucine	979.3	1135.6	-0.214	2939	0.0176	0.0435
2-Butylbenzothiazole	441.4	381.3	0.211	1821	0.018	0.0441
D-Sedoheptulose 7-phosphate	297.5	497.5	-0.742	2936	0.0182	0.0442
[Fv Dimethoxy,methyl(9:1)] (2S)-5,7-Dimethoxy-3',4'-methylenedioxyflavanone	651.1	1201.2	-0.883	2935	0.0184	0.0444
Oxoadipic acid	487.5	617.2	-0.341	2934	0.0186	0.0445
Thr-Cys-Cys	2325.9	2798.8	-0.267	2933	0.0188	0.0446
Creatine	4511	15140.7	-1.747	2930	0.0195	0.0458
Hydroxybutyrylcarnitine	156.7	259.7	-0.729	2929	0.0197	0.0459
5'-Dehydroadenosine	168.5	106.9	0.656	1833	0.0206	0.0462
Phe-Thr-Val	47.6	82.5	-0.793	2925	0.0206	0.0462
dUDP	149.3	319.2	-1.096	2925	0.0206	0.0462
L-Glutamine	616.2	706.6	-0.197	2926	0.0204	0.0462
Kaempferol 3-(2",3"-diacetyl-4"-p-coumaroylrhamnoside)	32.8	113.1	-1.788	2927	0.0201	0.0462

Metabolomic analysis performed on 139 samples (IBS: n= 78 and Control: n=61)

Median concentration represented as arbitary unit (A.U.)

Log2FC, log2 fold change between the groups

Supplementary Table 7: Faecal Metabolomics Machine learning is predictive of IBS versus Control

	LASSO Optimisation	**Random Forest Optimisation**	**Model Performance**
AUC	0.683 (0.139)	0.909 (0.084)	0.686 (0.132)
Sensitivity	0.624 (0.177)	0.903 (0.108)	0.737 (0.181)
Specificity	0.608 (0.202)	0.706 (0.188)	0.476 (0.122)

10-fold Cross Validation

	Predicted IBS	Predicted Control
IBS	59	21
Control	33	30

Rank #	**Metabolite ID**	**LASSO coefficients**	**Random Forest feature importance**
1	L-Phenylalanine	-0.788	88.34
2	Adenosine	0.345	78.31
3	MG(20:3(8Z,11Z,14Z)/0:0/0:0)	0.33	64.62
4	L-Alanine	-0.752	56.24
5	3,6-Dimethoxy-19-norpregna-1,3,5,7,9-pentaen-20-one	0.292	53.14
6	Glu-Ile-Ile-Phe	-0.569	49.57
7	Glu-Ala-Gln-Ser	-0.948	48.99
8	2,4,8-Eicosatrienoic acid isobutylamide	0.179	43.67
9	Piperidine	-0.161	38.43
10	Staphyloxanthin	0.251	37.03
11	beta-Carotinal	0.368	35.35
12	Hexoses	0.107	35.21
13	Ile-Arg-Ile	0.663	35.06
14	11-Deoxocucurbitacin I	-0.141	34.94

15	1-(Malonylamino)cyclopropanecarboxylic acid	0.353	31.96
16	PG(37:2)	0.908	31.75
17	[PR] gamma-Carotene/ beta,psi-Carotene	0.122	31.31
18	20-hydroxy-E4-neuroprostane	0.126	29.99
19	Ethylphenyl acetate	0.185	29.86
20	Dodecanedioic acid	0.089	28.24
21	Ile-Lys-Cys-Gly	-0.12	27.87
22	Tuberoside	0.873	27.39
23	D-galactal	0.223	26.84
24	3,6-Dihydro-4-(4-methyl-3-pentenyl)-1,2-dithiin	0.146	21.83
25	demethylmenaquinone-6	0.079	20.51
26	L-Arginine	0.071	20.33
27	PC(o-16:1(9Z)/14:1(9Z))	-0.09	19.9
28	Mesobilirubinogen	-0.155	19.84
29	Traumatic acid	0.172	19.82
30	alpha-Tocopherol succinate	0.123	18.74
31	3-Methylcrotonylglycine	0.182	18.39
32	(S)-(E)-8-(3,6-Dimethyl-2-heptenyl)-4',5,7-trihydroxyflavanone	0.072	18.03
33	xi-7-Hydroxyhexadecanedioic acid	0.031	17.96
34	beta-Pinene	0.025	16.94
35	Leu-Ser-Ser-Tyr	-0.041	16.69
36	Orotic acid	-0.143	16.59
37	Heptane-1-thiol	-0.047	15.82
38	Glu-Asp-Asp	0.038	15.43
39	LysoPE(18:2(9Z,12Z)/0:0)	0.02	15.28
40	LysoPE(22:0/0:0)	0.282	15.14

41	Creatine	0.209	15.03
42	Inosine	0.027	13.46
43	SM(d32:2)	-0.077	13.19
44	Arg-Leu-Val-Cys	0.043	12.52
45	PS(O-18:0/15:0)	-0.229	12.45
46	Pyridoxamine	-0.105	11.89
47	N-Heptanoylglycine	0.045	11.53
48	Hematoporphyrin IX	-0.161	11.4
49	3beta,5beta-Ketotriol	-0.096	10.59
50	2-Phenylpropionate	0.026	10
51	trans-2-Heptenal	0.014	9.63
52	LysoPC(0:0/18:0)	0.028	9.08
53	Linoleoyl ethanolamide	-0.025	8.93
54	LysoPE(24:0/0:0)	0.044	8.8
55	2-Methyl-3-hydroxyvaleric acid	-0.119	8.58
56	Quasiprotopanaxatriol	0.162	8.56
57	N-oleoyl isoleucine	0.059	8.49
58	(-)-(E)-1-(4-Hydroxyphenyl)-7-phenyl-6-hepten-3-ol	0.028	8.44
59	[FA hydroxy(4:0)] N-(3S-hydroxy-butanoyl)-homoserine lactone	0.024	8.43
60	Riboflavin cyclic-4',5'-phosphate	0.092	8
61	Arg-Lys-Trp-Val	-0.626	7.86
62	PC(20:1(11Z)/P-16:0)	0.033	7.8
63	3,5-Dihydroxybenzoic acid	0.083	7.67
64	Tyrosine	-0.012	7.43
65	2,3-Epoxymenaquinone	0.005	7.02
66	His-Met-Val-Val	-0.018	6.86
67	PI(41:2)	-0.021	6.84
68	Phenol	-0.018	6.74

69	3,3'-Dithiobis[2-methylfuran]	-0.053	6.73
70	Ala-Leu-Trp-Pro	0	6.7
71	1,2,3-Tris(1-ethoxyethoxy)propane	-0.051	6.48
72	Vanilpyruvic acid	-0.052	6.43
73	2-Hydroxy-3-carboxy-6-oxo-7-methylocta-2,4-dienoate	0.035	6.2
74	Secoeremopetasitolide B	0.023	5.77
75	2-O-Benzoyl-D-glucose	0.033	5.65
76	Ile-Leu-Phe-Trp	0.094	5.49
77	(R)-lipoic acid	0.036	5.18
78	PA(20:4(5Z,8Z,11Z,14Z)e/2:0)	0.013	5.15
79	PE(P-16:0e/0:0)	0.003	5.15
80	Benzyl isobutyrate	0.001	5.04
81	Hexyl 2-furoate	-0.099	5.04
82	Trp-Ala-Ser	0.012	4.95
83	LysoPC(15:0)	-0.093	4.72
84	4-Hydroxycrotonic acid	-0.007	4.72
85	3-Feruloyl-1,5-quinolactone	0.05	4.6
86	Furfuryl octanoate	0.178	4.44
87	PC(22:2(13Z,16Z)/15:0)	-0.006	4.26
88	(-)-1-Methylpropyl 1-propenyl disulfide	-0.021	4.07
89	PC (36:6)	0.073	4.05
90	Leucyl-Glycine	-0.096	3.96
91	CE(16:2)	0.041	3.81
92	Triterpenoid	0	3.79
93	Violaxanthin	0.002	3.79
94	[FA hydroxy(17:0)] heptadecanoic acid	-0.059	3.6
95	2-Hydroxyundecanoate	0.077	3.6
96	Chorismate	-0.003	3.52

97	delta-Dodecalactone	0.161	3.34
98	3-O-Protocatechuoylceanothic acid	0.058	3.31
99	PG(16:1(9Z)/16:1(9Z))	-0.004	3.17
100	p-Cresol sulfate	-0.003	3.15
101	Quercetin 3'-sulfate	0.02	3.03
102	PS(26:0))	-0.02	2.94
103	Ala-Leu-Phe-Trp	0.016	2.93
104	L-Glutamic acid 5-phosphate	-0.003	2.87
105	N,2,3-Trimethyl-2-(1-methylethyl)butanamide	-0.058	2.86
106	Isoamyl isovalerate	-0.06	2.85
107	n-Dodecane	-0.029	2.81
108	PC(14:1(9Z)/14:1(9Z))	-0.089	2.8
109	Lucyoside Q	0.007	2.76
110	Endomorphin-1	-0.017	2.51
111	3-Hydroxy-10'-apo-b,y-carotenal	0.013	2.5
112	Pyrroline hydroxycarboxylic acid	0.014	2.39
113	S-Propyl 1-propanesulfinothioate	0.019	2.38
114	N-Methylindolo[3,2-b]-5alpha-cholest-2-ene	-0.007	2.31
115	Tocopheronic acid	0.05	2.26
116	1-(2,4,6-Trimethoxyphenyl)-1,3-butanedione	0.018	2.24
117	Homogentisic acid	0.011	2.22
118	LysoPE(18:1(9Z)/0:0)	0.008	2.19
119	N-stearoyl valine	0.009	2.17
120	trans-Carvone oxide	0.07	2.14
121	1,1'-Thiobis-1-propanethiol	0.002	2.14
122	2-(Ethylsulfonylmethyl)phenyl methylcarbamate	0.076	2.05
123	menaquinone-4	0.004	2.04

124	Benzeneacetamide-4-O-sulphate	0.01	2
125	N5-acetyl-N5-hydroxy-L-ornithine	0.001	1.98
126	Succinic acid	0	1.97
127	Asn-Lys-Val-Pro	0.083	1.92
128	LysoPC(14:1(9Z))	0.003	1.88
129	Phenol glucuronide	-0.015	1.71
130	2-methyl-Butanoic acid, 2-methylbutyl ester	0.01	1.67
131	3-O-Caffeoyl-1-O-methylquinic acid	0.004	1.66
132	[FA hydroxy(24:0)] 3-hydroxy-tetracosanoic acid	0.01	1.63
133	N-(2-hydroxyhexadecanoyl)-sphinganine-1-phospho-(1'-myo-inositol)	0.146	1.56
134	gamma-Dodecalactone	0.117	1.54
135	PA(22:1(11Z)/0:0)	-0.074	1.49
136	Butyl butyrate	0.025	1.44
137	TG(20:5(5Z,8Z,11Z,14Z,17Z)/18:1(9Z)/22:5(7Z,10Z,13Z,16Z,19Z))[iso6]	-0.035	1.38
138	Clausarinol	0.03	1.36
139	4-Methyl-2-pentanone	0.006	1.31
140	Trigonelline	0.02	1.18
141	Arg-Val-Pro-Tyr	0.008	1.17
142	2,3-Methylenesuccinic acid	0.016	1.04
143	Serinyl-Threonine	0.005	1.04
144	Lycoperoside D	-0.009	1.03
145	Geraniol	0.012	1
146	1-18:2-lysophosphatidylglycerol	0.098	0.89
147	omega-6-Hexadecalactone, Ambrettolide	0.031	0.83
148	gamma-Glutamyl-S-methylcysteinyl-beta-alanine	0.008	0.79

149	FA oxo(22:0)	0.005	0.53
150	D-Ribose	-0.021	0.53
151	LysoPC(17:0)	0.036	0.47
152	PA(O-36:4)	0.02	0.38
153	C19 Sphingosine-1-phosphate	-0.018	0.34
154	4-Hydroxy-5-(dihydroxyphenyl)-valeric acid-O-methyl-O-sulphate	0.016	0.29
155	PE(14:1(9Z)/14:0)	0.015	0.28
156	Citronellyl tiglate	0.052	0.27
157	Ethyl methylphenylglycidate (isomer 1)	-0.038	0.24
158	N-Acetyl-leu-leu-tyr	0.003	0
158	PS(O-34:3)	-0.002	0

LASSO and Random Forest (RF) statistics of metabolites predictive of IBS versus Control

Analysis had 2 classes: Control and IBS and included 143 samples (IBS: n= 80 and Control: n=63)

Metrics reported are the mean and the standard deviation of values from Cross Validation.

Data used was log10 transformed

Supplementary Table 8: Faecal metabolites differentially abundant between the IBS and Control groups. Significantly different metabolites are shown. For complete table please refer to the published manuscript (doi: 10.1053/j.gastro.2019.11.301).

Metabolite	IBS (A.U.)	Control (A.U.)	Log2FC	Wilcoxon Statistic	p-value	q-value
2-Phenylpropionate	1323182.1	3247921.9	-1.296	3374	0	0.00505
3-Buten-1-amine	280286.1	167168.2	0.746	1388	0	0.00505
Adenosine	125862.8	222491.1	-0.822	3340	0	0.00505
I-Urobilin	2129046.3	508459.2	2.066	1444	0	0.00505
2,3-Epoxymenaquinone	245989.6	547357.5	-1.154	3313	0	0.00505
[FA (22:5)] 4,7,10,13,16-Docosapentaynoic acid	516717.6	1051721	-1.025	3309	0	0.00505
3,6-Dimethoxy-19-norpregna-1,3,5,7,9-pentaen-20-one	961706.2	2013326.2	-1.066	3298	0	0.00505
Cucurbitacin S	1617422.3	812194.6	0.994	1462	0.0001	0.00505
N-Heptanoylglycine	581244.7	1189914.8	-1.034	3296	0.0001	0.00505
11-Deoxocucurbitacin I	1509367.4	1026985.6	0.556	1478	0.000132	0.00599
Staphyloxanthin	125908.3	208397.6	-0.727	3264	0.000174	0.00716
Piperidine	536820.9	366827.8	0.549	1501	0.000196	0.00722
Leu-Ser-Ser-Tyr	194085.5	88714.9	1.129	1509	0.000224	0.00722
L-Urobilin	31844915.4	58134193.3	-0.868	3249	0.000224	0.00722
L-Phenylalanine	2052003.6	1343878.1	0.611	1513	0.000239	0.00722
Ala-Leu-Trp-Pro	323939	638393.1	-0.979	3238	0.000269	0.0074
3-Feruloyl-1,5-quinolactone	524541.4	876281.8	-0.74	3236	0.000278	0.0074
PG(P-16:0/14:0)	426308.9	798780.6	-0.906	3223	0.000343	0.00832
3-deoxy-D-galactose	226693.6	145983.2	0.635	1536	0.000349	0.00832
MG(20:3(8Z,11Z,14Z)/0:0/0:0)	89430.2	214373.1	-1.261	3215	0.000391	0.00857
Mesobilirubinogen	696662.3	251218.8	1.472	1544	0.000397	0.00857
L-Alanine	1429957.1	1081997.7	0.402	1548	0.000424	0.00872
Tyrosine	533603.6	368180.1	0.535	1564	0.000546	0.0106
PG(O-30:1)	140723.9	291063.6	-1.048	3192	0.000564	0.0106
beta-Pinene	171.8	276.9	-0.689	3187	0.00061	0.011
2,4,8-Eicosatrienoic acid isobutylamide	53648.9	167764.9	-1.645	3170.5	0.000787	0.0135

Glutarylglycine	1561150.4	2367236.8	-0.601	3169	0.000805	0.0135
[PR] gamma-Carotene/ beta,psi-Carotene	39594.5	55014.4	-0.475	3155	0.000996	0.0161
Neuromedin B (1-3)	1195664.8	414438.1	1.529	1610	0.00111	0.0173
Heptane-1-thiol	435910.8	336879.8	0.372	1613	0.00116	0.0174
Violaxanthin	688839.8	991237.9	-0.525	3143	0.00119	0.0174
Isolimonene	6.8	19	-1.492	3138	0.00128	0.0182
Ile-Lys-Cys-Gly	422439.2	241750.4	0.805	1625	0.00138	0.0187
His-Met-Val-Val	377162.4	223544.7	0.755	1626	0.0014	0.0187
Allyl caprylate	9.6	7.7	0.326	1632	0.00153	0.0196
Hydroxyprolyl-Tryptophan	323127	123183.6	1.391	1633	0.00156	0.0196
Dodecanedioic acid	671845.4	956268.6	-0.509	3122	0.00162	0.0199
2-O-Benzoyl-D-glucose	220717.8	469968.1	-1.09	3119	0.0017	0.0199
2-Ethylsuberic acid	384419	749840	-0.964	3118	0.00172	0.0199
D-Urobilin	1792754.5	301418.7	2.572	1641.5	0.00176	0.0199
20-hydroxy-E4-neuroprostane	125388	208519	-0.734	3113	0.00185	0.02
PG(O-31:1)	525453.8	924227.6	-0.815	3113	0.00185	0.02
Anigorufone	754382	1783246.9	-1.241	3110	0.00193	0.0203
Nonyl acetate	13.1	8.2	0.677	1658	0.00223	0.0229
L-Arginine	32851.3	72856.2	-1.149	3095	0.00239	0.0239
PG(P-32:1)	164475.2	226435.7	-0.461	3094	0.00242	0.0239
Glu-Ala-Gln-Ser	375851.9	273805.4	0.457	1668	0.00256	0.0247
PG(31:0)	160964.6	277244.2	-0.784	3087	0.00267	0.0252
Cucurbitacin I	793831.5	470668	0.754	1683	0.00316	0.0275
Arg-Lys-Phe-Val	479994	2477823.7	-2.368	3075	0.00316	0.0275
Genipinic acid	269618.2	535154	-0.989	3072.5	0.00327	0.0275
Hexoses	63587.8	102387.4	-0.687	3072.5	0.00327	0.0275
Lys-Phe-Phe-Phe	144955.5	76014.5	0.931	1686	0.00329	0.0275
PI(41:2)	523352.3	289816	0.853	1686	0.00329	0.0275
D-galactal	236791	433511.6	-0.872	3071	0.00334	0.0275
Traumatic acid	235655	352893.3	-0.583	3066	0.00357	0.0287

Adenine	312165.1	445818.4	-0.514	3065	0.00362	0.0287
PC(22:2(13Z,16Z)/15:0)	249100.9	131882.9	0.917	1695	0.00372	0.0287
2-Phenylethyl beta-D-glucopyranoside	330200	576025.7	-0.803	3061	0.00382	0.0287
PG(37:2)	208672.4	309558.5	-0.569	3060	0.00387	0.0287
Glycerol tributanoate	1818865.6	790191.7	1.203	1699	0.00393	0.0287
Arg-Leu-Pro-Arg	1113239.6	805486.6	0.467	1699	0.00393	0.0287
2-O-p-Coumaroyl-D-glucose	177559.4	309984.6	-0.804	3057	0.00403	0.029
3,4-Dihydroxyphenyllactic acid methyl ester	172842.1	321573.9	-0.896	3055	0.00414	0.0293
PG(P-28:0)	70315.5	138650.1	-0.98	3054	0.0042	0.0293
PG(34:0)	80115.9	135649.9	-0.76	3050	0.00443	0.0298
L-Lysine	391680.5	290959.3	0.429	1710	0.00455	0.0298
Ribitol	139100.6	308432.9	-1.149	3048	0.00455	0.0298
LysoPE(18:2(9Z,12Z)/0:0)	41861	70972.6	-0.762	3048	0.00455	0.0298
PA(20:4(5Z,8Z,11Z,14Z)e/2:0)	117279.2	179176	-0.611	3046	0.00467	0.0298
5-Dehydroshikimate	270282	486194.1	-0.847	3046	0.00467	0.0298
Threoninyl-Isoleucine	302458.5	194748	0.635	1715	0.00486	0.0301
L-Methionine	296185.5	228939.8	0.372	1717	0.00499	0.0301
PS(26:0))	3551762.1	1704565.8	1.059	1717	0.00499	0.0301
alpha-Pinene	92.1	215.6	-1.227	3041	0.00499	0.0301
Fenchene	12.1	26.4	-1.124	3039	0.00512	0.0305
Glu-Ile-Ile-Phe	171216.3	125559	0.447	1721	0.00526	0.0305
Gln-Phe-Phe-Phe	367594.4	170906.7	1.105	1721	0.00526	0.0305
Ursodeoxycholic acid	12666176	6449124.1	0.974	1726	0.00561	0.0318
PC(34:2)	112528.4	208697.3	-0.891	3032	0.00561	0.0318
3,17-Androstanediol glucuronide	469180.9	755540.9	-0.687	3031	0.00569	0.0318
Pyridoxamine	56652.2	41022	0.466	1730.5	0.00595	0.0324
[ST hydroxy] (25R)-3alpha,7alpha-dihydroxy-5beta-cholestan-27-oyl taurine	319975.1	229268.9	0.481	1732	0.00607	0.0324
PA(42:2)	1782124.8	686161.8	1.377	1732	0.00607	0.0324
[FA (16:0)] 2-bromo-hexadecanal	515055.9	256899.2	1.004	1733	0.00615	0.0324
3,6-Dihydro-4-(4-methyl-3-pentenyl)-1,2-dithiin	479922.9	701686.5	-0.548	3025	0.00615	0.0324

3-Methylcrotonylglycine	161596.9	287502.4	-0.831	3024	0.00623	0.0324
xi-7-Hydroxyhexadecanedioic acid	48647.5	70410.5	-0.533	3020	0.00656	0.0337
Camphene	7.7	17.7	-1.192	3017	0.00681	0.0345
2-Hydroxy-3-carboxy-6-oxo-7-methylocta-2,4-dienoate	375469	560318.7	-0.578	3014	0.00708	0.0345
7C-aglycone	1658154.1	2581551	-0.639	3014	0.00708	0.0345
1-(3-Aminopropyl)-4-aminobutanal	1007823.6	194401.6	2.374	1744	0.00708	0.0345
Benzyl isobutyrate	79.6	152.6	-0.938	3014	0.00708	0.0345
(S)-(E)-8-(3,6-Dimethyl-2-heptenyl)-4',5,7-trihydroxyflavanone	213117.3	551116.6	-1.371	3010	0.00745	0.0346
1,3-di-(5Z,8Z,11Z,14Z,17Z-eicosapentaenoyl)-2-hydroxy-glycerol (d5)	100264.5	212921.2	-1.087	3010	0.00745	0.0346
SM(d18:0/18:0)	80949.8	49417.7	0.712	1748	0.00745	0.0346
L-Homoserine	292067.7	226802.4	0.365	1749	0.00754	0.0346
17beta-(Acetylthio)estra-1,3,5(10)-trien-3-ol acetate	630945.8	1067932	-0.759	3009	0.00754	0.0346
[ST (2:0)] 5beta-Chola-3,11-dien-24-oic Acid	695679.3	320859.8	1.116	1750	0.00764	0.0346
PG(33:2)	75974.9	50361.9	0.593	1750	0.00764	0.0346
PE(22:4(7Z,10Z,13Z,16Z)/P-16:0)	81995.8	100782	-0.298	3006	0.00783	0.0351
Protoporphyrinogen IX	255656.7	187102.1	0.45	1756	0.00824	0.0366
alpha-Tocopherol succinate	47245.3	108160.8	-1.195	3001	0.00834	0.0367
Methyl (9Z)-6'-oxo-6,5'-diapo-6-carotenoate	899831	218922.6	2.039	1760	0.00866	0.037
PG(16:1(9Z)/16:1(9Z))	103173.2	74458	0.471	1760	0.00866	0.037
PC(o-22:1(13Z)/20:4(8Z,11Z,14Z,17Z))	52447	106165.5	-1.017	2998	0.00866	0.037
PG(31:2)	136942.9	86564.2	0.662	1761	0.00877	0.0371
alpha-phellandrene	61.7	199.1	-1.69	2992	0.00933	0.0391
[PS (12:0/13:0)] 1-dodecanoyl-2-tridecanoyl-sn-glycero-3-phosphoserine (ammonium salt)	7612945.4	10637361.3	-0.483	2991	0.00945	0.0393
Glu-Asp-Asp	340635.1	461045.6	-0.437	2989	0.00968	0.0399
PG(33:1)	215737.5	257078.3	-0.253	2984	0.0103	0.0416
PA(O-20:0/22:6(4Z,7Z,10Z,13Z,16Z,19Z))	235161.1	348850.6	-0.569	2984	0.0103	0.0416
[FA oxo(19:0)] 18-oxo-nonadecanoic acid	191012.9	264717.7	-0.471	2979	0.0109	0.0438
PG(16:1(9Z)/18:0)	714625.3	987020.3	-0.466	2978	0.0111	0.0438
Leu-Val	450170.2	278770.9	0.691	1781	0.0112	0.0438
demethylmenaquinone-6	576044.2	696566.9	-0.274	2977	0.0112	0.0438

PC(o-16:1(9Z)/14:1(9Z))	400429.2	269886.5	0.569	1782	0.0113	0.0439
PG(P-32:0)	306573.5	512926.2	-0.743	2974	0.0116	0.0444
(24E)-3beta,15alpha,22S-Triacetoxylanosta-7,9(11),24-trien-26-oic acid	390549.4	641541.7	-0.716	2973	0.0118	0.0444
PA(33:5)	2066319.8	1269332.8	0.703	1785	0.0118	0.0444
LysoPC(0:0/18:0)	210818	418649.1	-0.99	2970	0.0122	0.0457
Ile-Arg-Ile	56540.6	70311.6	-0.314	2968	0.0125	0.0464
Lauryl acetate	3.3	4.8	-0.525	2967	0.0126	0.0466
Glu-Glu-Gly-Tyr	292531	208064.8	0.492	1793	0.013	0.0473
3-(Methylthio)-1-propanol	215.9	162	0.414	1794	0.0131	0.0475
(-)-(E)-1-(4-Hydroxyphenyl)-7-phenyl-6-hepten-3-ol	1827847.9	2763203.9	-0.596	2962	0.0134	0.0479
Dimethyl benzyl carbinyl butyrate	5.4	12.7	-1.232	2962	0.0134	0.0479
Methyl 2,3-dihydro-3,5-dihydroxy-2-oxo-3-indoleacetic acid	1058511.5	1884600.2	-0.832	2960	0.0137	0.0486

Metabolomic analysis performed on 139 samples (IBS: n = 78 and Control: n = 61)

Median concentration represented as arbitary unit (A.U.)

Log2FC, log2 fold change between the groups

Supplementary Table 9: Wilcox Rank Sum Statistical analysis of bile acids (BAs) between the subgroups.

Subgroup	Total BAs	Primary BAs	Secondary BAs	Sulfated BAs	UDCA	Conjugated BAs
Control	7.11 (0.285)	5.038 (4.446)	94.962 (4.446)	8.336 (6.14)	47.186 (22.212)	13.374 (9.323)
IBS-C	7.22 (0.322)	5.216 (4.271)	94.784 (4.271)	9.028 (9.923)	55.094 (21.022)	14.244 (12.293)
IBS-D	7.37 (0.31) *	3.593 (4.117)	96.407 (4.117)	4.126 (3.507) *	67.022 (15.419) **	7.719 (6.03) *
IBS-M	7.18 (0.345)	4.127 (3.121)	95.873 (3.121)	9.603 (10.878)	51.007 (22.764)	13.73 (11.995)

Statistical analysis was performed on 139 samples (IBS: n = 78 and Control: n = 61)

Significance after p value adjustment (Benjamini-Hochberg), was observed only in Control vs IBS-D.

* p-adj < 0.05, ** p-adj < 0.01.

Total bile acids are represented as mean of log10 values.

Others bile acid categories are presented as a percentage of the total bile acids (without Log 10 transformation).

Supplementary Table 10: Spearman correlation analysis between bile acids (BAs) and secondary BA synthesis pathways.

Pathway	Total BAs	Primary BAs	Secondary BAs	Sulfated BAs	UDCA	Conjugated BAs
ursodeoxycholate biosynthesis (PWY_7588)	0.258*	0.007	0.26*	-0.113	0.298**	-0.133
glycocholate metabolism (PWY_6518)	0.362**	-0.045	0.37**	-0.125	0.42**	-0.156

Statistical analysis was performed on 135 samples (IBS: n = 78 and Control: n = 57)

Supplementary Table 11: 16S rRNA gene based Genus nodes and features in the WGCNA

Node	Genus	Node	Genus	Node	Genus	Node	Genus
A1	Ruminococcus	A2	Bilophila	A4	Alloprevotella	A5	Enterorhabdus
A1	Eubacterium	A2	Erysipelotrichaceae_incertae_sedis	A4	Roseburia	A5	Phascolarctobacterium
A1	Clostridium sensu stricto	A2	Pseudoflavonifractor	A4	Gemmiger	A5	Desulfovibrio
A1	Gordonibacter	A2	Odoribacter	A4	Faecalibacterium	A6	Haemophilus
A1	Klebsiella	A2	Anaerotruncus	A5	Cloacibacillus	A6	Veillonella
A1	Bifidobacterium	A3	Holdemania	A5	Methanobrevibacter	A6	Lactococcus
A1	Clostridium XI	A3	Lachnospiracea_incertae_sedis	A5	Paraprevotella	A6	Streptophyta
A1	Turicibacter	A3	Flavonifractor	A5	Slackia	A6	Streptococcus
A1	Hydrogenoanaerobacterium	A3	Eggerthella	A5	Victivallis	A6	Lactobacillus
A1	Collinsella	A3	Ruminococcus2	A5	Akkermansia	A6	Parasutterella
A1	Achromobacter	A3	Blautia	A5	Howardella	A6	Granulicatella
A1	Escherichia/Shigella	A3	Anaerostipes	A5	Sporobacter	A6	Rothia
A1	Acidaminococcus	A3	Bacteroides	A5	Mogibacterium	A6	Anaerosporobacter
A2	Porphyromonas	A3	Clostridium XVIII	A5	Coprococcus	A6	Dialister
A2	Oscillibacter	A3	Dorea	A5	Clostridium IV	A6	Actinomyces
A2	Pseudomonas	A4	Clostridium XlVa	A5	Olsenella		
A2	Alistipes	A4	Clostridium XlVb	A5	Butyricimonas		
A2	Barnesiella	A4	Catenibacterium	A5	Anaerofilum		
A2	Parabacteroides	A4	Butyricicoccus	A5	Prevotella		
A2	Acetanaerobacterium	A4	Sutterella	A5	Allisonella		

Supplementary Table 12: Shotgun species nodes and features in the WGCNA.

Node	Species	Node	Species	Node	Species	Node	Species
B1	Methanobrevibacter_smithii	B2	Bifidobacterium_pseudocatenulatum	B3	Eggerthella_lenta	B3	Ruminococcus_torques
B1	Bacteroides_coprocola	B2	Bacteroides_xylanisolvens	B3	Anaerotruncus_colihominis	B3	Bacteroides_fragilis
B1	Bacteroides_eggerthii	B2	Parabacteroides_johnsonii	B3	Eubacterium_sp_3_1_31	B3	Dorea_longicatena
B1	Bacteroides_finegoldii	B2	Streptococcus_anginosus	B3	Blautia_hydrogenotrophica	B3	Roseburia_inulinivorans
B1	Bacteroides_plebeius	B2	Streptococcus_australis	B3	Flavonifractor_plautii	B3	Bifidobacterium_longum
B1	Bacteroidales_bacterium_ph8	B2	Streptococcus_gordonii	B3	Holdemania_filiformis	B3	Ruminococcus_sp_5_1_39BFAA
B1	Barnesiella_intestinihominis	B2	Streptococcus_infantis	B3	Lachnospiraceae_bacterium_1_4_56FAA	B3	Eubacterium_hallii
B1	Coprobacter_fastidiosus	B2	Streptococcus_mitis_oralis_pneumoniae	B3	Bifidobacterium_animalis	B3	Roseburia_intestinalis
B1	Odoribacter_splanchnicus	B2	Streptococcus_parasanguinis	B3	Lachnospiraceae_bacterium_2_1_58FAA	B3	Bacteroides_ovatus
B1	Parabacteroides_goldsteinii	B2	Streptococcus_salivarius	B3	Lachnospiraceae_bacterium_7_1_58FAA	B3	Bacteroides_dorei
B1	Alistipes_finegoldii	B2	Streptococcus_sanguinis	B3	Clostridium_hathewayi	B3	Ruminococcus_bromii
B1	Alistipes_indistinctus	B2	Streptococcus_vestibularis	B3	Clostridium_symbiosum	B3	Dialister_invisus
B1	Alistipes_onderdonkii	B2	Clostridium_bartlettii	B3	Ruminococcaceae_bacterium_D16	B3	Bacteroides_uniformis
B1	Alistipes_putredinis	B2	Turicibacter_sanguinis	B3	Clostridium_nexile	B3	Eubacterium_rectale
B1	Alistipes_senegalensis	B2	Veillonella_atypica	B3	Adlercreutzia_equolifaciens	B3	Faecalibacterium_prausnitzii
B1	Alistipes_shahii	B2	Veillonella_dispar	B3	Lachnospiraceae_bacterium_3_1_57FAA_CT1	B4	Bacteroides_caccae
B1	Clostridium_sp_L2_50	B2	Veillonella_parvula	B3	Clostridium_clostridioforme	B4	Bacteroides_cellulosilyticus
B1	Eubacterium_eligens	B2	Burkholderiales_bacterium_1_1_47	B3	Anaerostipes_hadrus	B4	Bacteroides_clarus
B1	Eubacterium_siraeum	B2	Parasutterella_excrementihominis	B3	Clostridium_leptum	B4	Bacteroides_faecis
B1	Coprococcus_catus	B2	Haemophilus_parainfluenzae	B3	Lachnospiraceae_bacterium_5_1_63FAA	B4	Bacteroides_intestinalis
B1	Coprococcus_eutactus	B3	Erysipelotrichaceae_bacterium_21_3	B3	Lachnospiraceae_bacterium_1_1_57FAA	B4	Bacteroides_massiliensis
B1	Coprococcus_sp_ART55_1	B3	Pseudoflavonifractor_capillosus	B3	Lachnospiraceae_bacterium_3_1_46FAA	B4	Bacteroides_salyersiae
B1	Lachnospiraceae_bacterium_8_1_57FAA	B3	Clostridium_ramosum	B3	Clostridium_bolteae	B4	Bacteroides_stercoris
B1	Roseburia_hominis	B3	Clostridium_innocuum	B3	Streptococcus_thermophilus	B4	Bacteroides_vulgatus
B1	Ruminococcus_callidus	B3	Erysipelotrichaceae_bacterium_2_2_44A	B3	Alistipes_sp_AP11	B4	Parabacteroides_distasonis

B1	Ruminococcus_lactaris	B3	Lachnospiraceae_bacterium_5_1_57FAA	B3	Eubacterium_ramulus	B4	Parabacteroides_merdae
B1	Eubacterium_biforme	B3	Erysipelotrichaceae_bacterium_6_1_45	B3	Eubacterium_ventriosum	B4	Paraprevotella_clara
B1	Oxalobacter_formigenes	B3	Lactococcus_lactis	B3	Ruminococcus_gnavus	B4	Paraprevotella_xylaniphila
B1	Desulfovibrio_desulfuricans	B3	Clostridium_citroniae	B3	Dorea_formicigenerans	B4	Prevotella_copri
B1	Akkermansia_muciniphila	B3	Clostridiaceae_bacterium_JC118	B3	Collinsella_aerofaciens	B4	Phascolarctobacterium_succinatutens
B2	Rothia_mucilaginosa	B3	Gordonibacter_pamelaeae	B3	Bacteroides_thetaiotaomicron	B4	Sutterella_wadsworthensis
B2	Bifidobacterium_adolescentis	B3	Clostridium_asparagiforme	B3	Coprococcus_comes	B4	Bilophila_wadsworthia
B2	Bifidobacterium_bifidum	B3	Bacteroides_nordii	B3	Escherichia_coli	B4	Desulfovibrio_piger
B2	Bifidobacterium_catenulatum	B3	Clostridiales_bacterium_1_7_47FAA	B3	Ruminococcus_obeum		

Supplementary Table 13: Shotgun pathway features in the WGCNA nodes. Nodes discussed in the main text are shown. For complete table please refer to the published manuscript (doi: 10.1053/j.gastro.2019.11.301).

Node	Pathway	Species	Node	Pathway	Species
C3	superpathway of guanosine nucleotides de novo biosynthesis I	COMPLETE__Haemophilus_parainfluenzae	C6	L-arginine biosynthesis IV (archaebacteria)	Veillonella_parvula
C3	pyruvate fermentation to isobutanol (engineered)	COMPLETE__Streptococcus_parasanguinis	C6	L-valine biosynthesis	Veillonella_parvula
C3	S-adenosyl-L-methionine cycle I	COMPLETE__Streptococcus_salivarius	C6	L-valine biosynthesis	Veillonella_parvula
C3	(5Z)-dodec-5-enoate biosynthesis	Haemophilus_parainfluenzae	C6	peptidoglycan maturation (meso-diaminopimelate containing)	Veillonella_parvula
C3	5-aminoimidazole ribonucleotide biosynthesis I	Haemophilus_parainfluenzae	C6	pyrimidine deoxyribonucleotide phosphorylation	Veillonella_parvula
C3	5-aminoimidazole ribonucleotide biosynthesis II	Haemophilus_parainfluenzae	C6	pyrimidine deoxyribonucleotide phosphorylation	Veillonella_parvula
C3	5-aminoimidazole ribonucleotide biosynthesis II	Haemophilus_parainfluenzae	C6	pyruvate fermentation to isobutanol (engineered)	Veillonella_parvula
C3	adenine and adenosine salvage III	Haemophilus_parainfluenzae	C6	superpathway of guanosine nucleotides de novo biosynthesis I	Veillonella_parvula
C3	adenosine deoxyribonucleotides de novo biosynthesis II	Haemophilus_parainfluenzae	C6	superpathway of pyrimidine nucleobases salvage	Veillonella_parvula
C3	adenosine ribonucleotides de novo biosynthesis	Haemophilus_parainfluenzae	C8	adenosine ribonucleotides de novo biosynthesis	Barnesiella_intestinihominis

C3	ADP-L-glycero-β-D-manno-heptose biosynthesis	Haemophilus_parainfluenzae	C8	CDP-diacylglycerol biosynthesis I	Barnesiella_intestinihominis
C3	CDP-diacylglycerol biosynthesis II	Haemophilus_parainfluenzae	C8	CDP-diacylglycerol biosynthesis II	Barnesiella_intestinihominis
C3	chorismate biosynthesis from 3-dehydroquinate	Haemophilus_parainfluenzae	C8	coenzyme A biosynthesis I	Barnesiella_intestinihominis
C3	chorismate biosynthesis from 3-dehydroquinate	Haemophilus_parainfluenzae	C8	glycolysis IV (plant cytosol)	Barnesiella_intestinihominis
C3	chorismate biosynthesis from 3-dehydroquinate	Haemophilus_parainfluenzae	C8	guanosine ribonucleotides de novo biosynthesis	Barnesiella_intestinihominis
C3	chorismate biosynthesis I	Haemophilus_parainfluenzae	C8	L-lysine biosynthesis III	Barnesiella_intestinihominis
C3	D-galactose degradation V (Leloir pathway)	Haemophilus_parainfluenzae	C8	L-lysine biosynthesis VI	Barnesiella_intestinihominis
C3	glycolysis VI (metazoan)	Haemophilus_parainfluenzae	C8	L-valine biosynthesis	Barnesiella_intestinihominis
C3	guanosine deoxyribonucleotides de novo biosynthesis II	Haemophilus_parainfluenzae	C8	peptidoglycan biosynthesis I (meso-diaminopimelate containing)	Barnesiella_intestinihominis
C3	guanosine ribonucleotides de novo biosynthesis	Haemophilus_parainfluenzae	C8	phosphopantothenate biosynthesis I	Barnesiella_intestinihominis
C3	guanosine ribonucleotides de novo biosynthesis	Haemophilus_parainfluenzae	C8	preQ0 biosynthesis	Barnesiella_intestinihominis
C3	inosine-5'-phosphate biosynthesis I	Haemophilus_parainfluenzae	C8	pyruvate fermentation to isobutanol (engineered)	Barnesiella_intestinihominis
C3	inosine-5'-phosphate biosynthesis I	Haemophilus_parainfluenzae	C8	queuosine biosynthesis	Barnesiella_intestinihominis
C3	inosine-5'-phosphate biosynthesis I	Haemophilus_parainfluenzae	C8	S-adenosyl-L-methionine cycle I	Barnesiella_intestinihominis
C3	inosine-5'-phosphate biosynthesis II	Haemophilus_parainfluenzae	C8	UDP-N-acetylmuramoyl-pentapeptide biosynthesis I (meso-diaminopimelate containing)	Barnesiella_intestinihominis
C3	inosine-5'-phosphate biosynthesis II	Haemophilus_parainfluenzae	C8	UDP-N-acetylmuramoyl-pentapeptide biosynthesis II (lysine-containing)	Barnesiella_intestinihominis
C3	inosine-5'-phosphate biosynthesis III	Haemophilus_parainfluenzae	C8	UMP biosynthesis	Barnesiella_intestinihominis
C3	L-tryptophan biosynthesis	Haemophilus_parainfluenzae	C8	urate biosynthesis/inosine 5'-phosphate degradation	Barnesiella_intestinihominis
C3	pantothenate and coenzyme A biosynthesis III	Haemophilus_parainfluenzae	C12	L-valine biosynthesis	Clostridium_hathewayi
C3	preQ0 biosynthesis	Haemophilus_parainfluenzae	C12	L-valine biosynthesis	Clostridium_symbiosum
C3	purine ribonucleosides degradation	Haemophilus_parainfluenzae	C12	L-valine biosynthesis	Clostridium_symbiosum
C3	pyrimidine deoxyribonucleosides salvage	Haemophilus_parainfluenzae	C12	purine ribonucleosides degradation	COMPLETE__Lachnospiraceae_bacterium_1_1_57FAA
C3	pyrimidine deoxyribonucleotide phosphorylation	Haemophilus_parainfluenzae	C12	superpathway of 5-aminoimidazole ribonucleotide biosynthesis	COMPLETE__Lachnospiraceae_bacterium_3_1_46FAA

C3	pyrimidine deoxyribonucleotides de novo biosynthesis I	Haemophilus_parainfluenzae	C12	adenosine ribonucleotides de novo biosynthesis	Lachnospiraceae_bacterium_1_1_57FAA
C3	pyrimidine deoxyribonucleotides de novo biosynthesis II	Haemophilus_parainfluenzae	C12	adenosine ribonucleotides de novo biosynthesis	Lachnospiraceae_bacterium_1_1_57FAA
C3	pyruvate fermentation to isobutanol (engineered)	Haemophilus_parainfluenzae	C12	CDP-diacylglycerol biosynthesis II	Lachnospiraceae_bacterium_1_1_57FAA
C3	pyruvate fermentation to isobutanol (engineered)	Haemophilus_parainfluenzae	C12	CDP-diacylglycerol biosynthesis II	Lachnospiraceae_bacterium_1_1_57FAA
C3	pyruvate fermentation to isobutanol (engineered)	Haemophilus_parainfluenzae	C12	chorismate biosynthesis I	Lachnospiraceae_bacterium_1_1_57FAA
C3	pyruvate fermentation to isobutanol (engineered)	Haemophilus_parainfluenzae	C12	guanosine ribonucleotides de novo biosynthesis	Lachnospiraceae_bacterium_1_1_57FAA
C3	queuosine biosynthesis	Haemophilus_parainfluenzae	C12	guanosine ribonucleotides de novo biosynthesis	Lachnospiraceae_bacterium_1_1_57FAA
C3	S-adenosyl-L-methionine cycle I	Haemophilus_parainfluenzae	C12	guanosine ribonucleotides de novo biosynthesis	Lachnospiraceae_bacterium_1_1_57FAA
C3	S-adenosyl-L-methionine cycle I	Haemophilus_parainfluenzae	C12	purine ribonucleosides degradation	Lachnospiraceae_bacterium_1_1_57FAA
C3	seleno-amino acid biosynthesis	Haemophilus_parainfluenzae	C12	pyruvate fermentation to isobutanol (engineered)	Lachnospiraceae_bacterium_1_1_57FAA
C3	starch degradation V	Haemophilus_parainfluenzae	C12	pyruvate fermentation to isobutanol (engineered)	Lachnospiraceae_bacterium_1_1_57FAA
C3	starch degradation V	Haemophilus_parainfluenzae	C12	pyruvate fermentation to isobutanol (engineered)	Lachnospiraceae_bacterium_1_1_57FAA
C3	superpathway of 5-aminoimidazole ribonucleotide biosynthesis	Haemophilus_parainfluenzae	C12	pyruvate fermentation to isobutanol (engineered)	Lachnospiraceae_bacterium_1_1_57FAA
C3	superpathway of 5-aminoimidazole ribonucleotide biosynthesis	Haemophilus_parainfluenzae	C12	queuosine biosynthesis	Lachnospiraceae_bacterium_1_1_57FAA
C3	superpathway of 5-aminoimidazole ribonucleotide biosynthesis	Haemophilus_parainfluenzae	C12	starch degradation V	Lachnospiraceae_bacterium_1_1_57FAA
C3	superpathway of adenosine nucleotides de novo biosynthesis I	Haemophilus_parainfluenzae	C12	starch degradation V	Lachnospiraceae_bacterium_1_1_57FAA
C3	superpathway of guanosine nucleotides de novo biosynthesis I	Haemophilus_parainfluenzae	C12	starch degradation V	Lachnospiraceae_bacterium_1_1_57FAA
C3	superpathway of guanosine nucleotides de novo biosynthesis II	Haemophilus_parainfluenzae	C12	UDP-N-acetylmuramoyl-pentapeptide biosynthesis I (meso-diaminopimelate containing)	Lachnospiraceae_bacterium_1_1_57FAA
C3	superpathway of L-alanine biosynthesis	Haemophilus_parainfluenzae	C12	UDP-N-acetylmuramoyl-pentapeptide biosynthesis I (meso-diaminopimelate containing)	Lachnospiraceae_bacterium_1_1_57FAA
C3	superpathway of purine deoxyribonucleosides degradation	Haemophilus_parainfluenzae	C12	UDP-N-acetylmuramoyl-pentapeptide biosynthesis II (lysine-containing)	Lachnospiraceae_bacterium_1_1_57FAA

C3	superpathway of purine deoxyribonucleosides degradation	Haemophilus_parainfluenzae	C12	UDP-N-acetylmuramoyl-pentapeptide biosynthesis II (lysine-containing)	Lachnospiraceae_bacterium_1_1_57FAA
C3	superpathway of pyrimidine nucleobases salvage	Haemophilus_parainfluenzae	C12	5-aminoimidazole ribonucleotide biosynthesis I	Lachnospiraceae_bacterium_3_1_46FAA
C3	superpathway of pyrimidine nucleobases salvage	Haemophilus_parainfluenzae	C12	5-aminoimidazole ribonucleotide biosynthesis I	Lachnospiraceae_bacterium_3_1_46FAA
C3	tetrapyrrole biosynthesis I (from glutamate)	Haemophilus_parainfluenzae	C12	5-aminoimidazole ribonucleotide biosynthesis I	Lachnospiraceae_bacterium_3_1_46FAA
C3	UDP-N-acetyl-D-glucosamine biosynthesis I	Haemophilus_parainfluenzae	C12	5-aminoimidazole ribonucleotide biosynthesis II	Lachnospiraceae_bacterium_3_1_46FAA
C3	UDP-N-acetylmuramoyl-pentapeptide biosynthesis II (lysine-containing)	Haemophilus_parainfluenzae	C12	5-aminoimidazole ribonucleotide biosynthesis II	Lachnospiraceae_bacterium_3_1_46FAA
C3	UDP-N-acetylmuramoyl-pentapeptide biosynthesis II (lysine-containing)	Haemophilus_parainfluenzae	C12	CDP-diacylglycerol biosynthesis I	Lachnospiraceae_bacterium_3_1_46FAA
C3	superpathway of branched amino acid biosynthesis	Streptococcus_australis	C12	CDP-diacylglycerol biosynthesis I	Lachnospiraceae_bacterium_3_1_46FAA
C3	5-aminoimidazole ribonucleotide biosynthesis I	Streptococcus_parasanguinis	C12	chorismate biosynthesis from 3-dehydroquinate	Lachnospiraceae_bacterium_3_1_46FAA
C3	5-aminoimidazole ribonucleotide biosynthesis I	Streptococcus_parasanguinis	C12	chorismate biosynthesis from 3-dehydroquinate	Lachnospiraceae_bacterium_3_1_46FAA
C3	5-aminoimidazole ribonucleotide biosynthesis I	Streptococcus_parasanguinis	C12	chorismate biosynthesis I	Lachnospiraceae_bacterium_3_1_46FAA
C3	5-aminoimidazole ribonucleotide biosynthesis II	Streptococcus_parasanguinis	C12	glycolysis IV (plant cytosol)	Lachnospiraceae_bacterium_3_1_46FAA
C3	5-aminoimidazole ribonucleotide biosynthesis II	Streptococcus_parasanguinis	C12	glycolysis IV (plant cytosol)	Lachnospiraceae_bacterium_3_1_46FAA
C3	adenosine ribonucleotides de novo biosynthesis	Streptococcus_parasanguinis	C12	L-lysine biosynthesis III	Lachnospiraceae_bacterium_3_1_46FAA
C3	adenosine ribonucleotides de novo biosynthesis	Streptococcus_parasanguinis	C12	methylerythritol phosphate pathway I	Lachnospiraceae_bacterium_3_1_46FAA
C3	adenosine ribonucleotides de novo biosynthesis	Streptococcus_parasanguinis	C12	methylerythritol phosphate pathway I	Lachnospiraceae_bacterium_3_1_46FAA
C3	adenosine ribonucleotides de novo biosynthesis	Streptococcus_parasanguinis	C12	peptidoglycan biosynthesis I (meso-diaminopimelate containing)	Lachnospiraceae_bacterium_3_1_46FAA
C3	CDP-diacylglycerol biosynthesis I	Streptococcus_parasanguinis	C12	peptidoglycan biosynthesis I (meso-diaminopimelate containing)	Lachnospiraceae_bacterium_3_1_46FAA
C3	CMP-3-deoxy-D-manno-octulosonate biosynthesis I	Streptococcus_parasanguinis	C12	phosphopantothenate biosynthesis I	Lachnospiraceae_bacterium_3_1_46FAA
C3	folate transformations II	Streptococcus_parasanguinis	C12	phosphopantothenate biosynthesis I	Lachnospiraceae_bacterium_3_1_46FAA
C3	folate transformations II	Streptococcus_parasanguinis	C12	superpathway of 5-aminoimidazole ribonucleotide biosynthesis	Lachnospiraceae_bacterium_3_1_46FAA

C3	galactose degradation I (Leloir pathway)	Streptococcus_parasanguinis	C12	superpathway of 5-aminoimidazole ribonucleotide biosynthesis	Lachnospiraceae_bacterium_3_1_46FAA
C3	glycolysis III (from glucose)	Streptococcus_parasanguinis	C12	5-aminoimidazole ribonucleotide biosynthesis II	Ruminococcus_torques
C3	inosine-5'-phosphate biosynthesis II	Streptococcus_parasanguinis	C12	Calvin-Benson-Bassham cycle	Ruminococcus_torques
C3	L-histidine degradation III	Streptococcus_parasanguinis	C12	coenzyme A biosynthesis I	Ruminococcus_torques
C3	lipid IVA biosynthesis	Streptococcus_parasanguinis	C12	coenzyme A biosynthesis I	Ruminococcus_torques
C3	L-isoleucine biosynthesis III	Streptococcus_parasanguinis	C12	methylerythritol phosphate pathway I	Ruminococcus_torques
C3	L-proline biosynthesis II (from arginine)	Streptococcus_parasanguinis	C12	superpathway of aromatic amino acid biosynthesis	Ruminococcus_torques
C3	L-valine biosynthesis	Streptococcus_parasanguinis	C12	superpathway of aromatic amino acid biosynthesis	Ruminococcus_torques
C3	L-valine biosynthesis	Streptococcus_parasanguinis	C35	pyruvate fermentation to isobutanol (engineered)	COMPLETE__Eubacterium_eligens
C3	peptidoglycan biosynthesis I (meso-diaminopimelate containing)	Streptococcus_parasanguinis	C35	5-aminoimidazole ribonucleotide biosynthesis I	Eubacterium_eligens
C3	purine ribonucleosides degradation	Streptococcus_parasanguinis	C35	5-aminoimidazole ribonucleotide biosynthesis II	Eubacterium_eligens
C3	pyrimidine deoxyribonucleotide phosphorylation	Streptococcus_parasanguinis	C35	adenosine ribonucleotides de novo biosynthesis	Eubacterium_eligens
C3	sucrose degradation III (sucrose invertase)	Streptococcus_parasanguinis	C35	CDP-diacylglycerol biosynthesis I	Eubacterium_eligens
C3	superpathway of 5-aminoimidazole ribonucleotide biosynthesis	Streptococcus_parasanguinis	C35	CDP-diacylglycerol biosynthesis II	Eubacterium_eligens
C3	superpathway of adenosine nucleotides de novo biosynthesis II	Streptococcus_parasanguinis	C35	chorismate biosynthesis from 3-dehydroquinate	Eubacterium_eligens
C3	superpathway of L-alanine biosynthesis	Streptococcus_parasanguinis	C35	chorismate biosynthesis I	Eubacterium_eligens
C3	superpathway of L-isoleucine biosynthesis I	Streptococcus_parasanguinis	C35	guanosine ribonucleotides de novo biosynthesis	Eubacterium_eligens
C3	superpathway of purine nucleotides de novo biosynthesis I	Streptococcus_parasanguinis	C35	L-valine biosynthesis	Eubacterium_eligens
C3	UDP-N-acetylmuramoyl-pentapeptide biosynthesis I (meso-diaminopimelate containing)	Streptococcus_parasanguinis	C35	pentose phosphate pathway (non-oxidative branch)	Eubacterium_eligens
C3	UDP-N-acetylmuramoyl-pentapeptide biosynthesis I (meso-diaminopimelate containing)	Streptococcus_parasanguinis	C35	purine ribonucleosides degradation	Eubacterium_eligens
C3	UMP biosynthesis	Streptococcus_parasanguinis	C35	S-adenosyl-L-methionine cycle I	Eubacterium_eligens
C3	urate biosynthesis/inosine 5'-phosphate degradation	Streptococcus_parasanguinis	C35	superpathway of 5-aminoimidazole ribonucleotide biosynthesis	Eubacterium_eligens

C3	adenine and adenosine salvage III	Streptococcus_salivarius	C35	superpathway of aromatic amino acid biosynthesis	Eubacterium_eligens
C3	coenzyme A biosynthesis I	Streptococcus_salivarius	C35	UDP-N-acetylmuramoyl-pentapeptide biosynthesis I (meso-diaminopimelate containing)	Eubacterium_eligens
C3	coenzyme A biosynthesis I	Streptococcus_salivarius	C35	UDP-N-acetylmuramoyl-pentapeptide biosynthesis II (lysine-containing)	Eubacterium_eligens
C3	coenzyme A biosynthesis I	Streptococcus_salivarius	C36	superpathway of adenosine nucleotides de novo biosynthesis I	COMPLETE__Streptococcus_thermophilus
C3	glucose and glucose-1-phosphate degradation	Streptococcus_salivarius	C36	5-aminoimidazole ribonucleotide biosynthesis I	Streptococcus_thermophilus
C3	glycolysis IV (plant cytosol)	Streptococcus_salivarius	C36	5-aminoimidazole ribonucleotide biosynthesis II	Streptococcus_thermophilus
C3	heme biosynthesis II (anaerobic)	Streptococcus_salivarius	C36	adenine and adenosine salvage III	Streptococcus_thermophilus
C3	lactose and galactose degradation I	Streptococcus_salivarius	C36	adenosine ribonucleotides de novo biosynthesis	Streptococcus_thermophilus
C3	lactose and galactose degradation I	Streptococcus_salivarius	C36	chorismate biosynthesis from 3-dehydroquinate	Streptococcus_thermophilus
C3	L-histidine biosynthesis	Streptococcus_salivarius	C36	chorismate biosynthesis I	Streptococcus_thermophilus
C3	L-histidine degradation I	Streptococcus_salivarius	C36	D-galactose degradation V (Leloir pathway)	Streptococcus_thermophilus
C3	L-isoleucine biosynthesis I (from threonine)	Streptococcus_salivarius	C36	folate transformations II	Streptococcus_thermophilus
C3	L-lysine biosynthesis I	Streptococcus_salivarius	C36	galactose degradation I (Leloir pathway)	Streptococcus_thermophilus
C3	L-lysine biosynthesis III	Streptococcus_salivarius	C36	guanosine ribonucleotides de novo biosynthesis	Streptococcus_thermophilus
C3	N10-formyl-tetrahydrofolate biosynthesis	Streptococcus_salivarius	C36	inosine-5'-phosphate biosynthesis I	Streptococcus_thermophilus
C3	peptidoglycan biosynthesis I (meso-diaminopimelate containing)	Streptococcus_salivarius	C36	inosine-5'-phosphate biosynthesis II	Streptococcus_thermophilus
C3	pyrimidine deoxyribonucleosides salvage	Streptococcus_salivarius	C36	L-valine biosynthesis	Streptococcus_thermophilus
C3	superpathway of adenosine nucleotides de novo biosynthesis I	Streptococcus_salivarius	C36	purine ribonucleosides degradation	Streptococcus_thermophilus
C3	superpathway of aromatic amino acid biosynthesis	Streptococcus_salivarius	C36	pyrimidine deoxyribonucleotide phosphorylation	Streptococcus_thermophilus
C3	superpathway of aromatic amino acid biosynthesis	Streptococcus_salivarius	C36	pyruvate fermentation to isobutanol (engineered)	Streptococcus_thermophilus
C3	superpathway of aromatic amino acid biosynthesis	Streptococcus_salivarius	C36	superpathway of 5-aminoimidazole ribonucleotide biosynthesis	Streptococcus_thermophilus
C3	superpathway of fatty acid biosynthesis initiation (E. coli)	Streptococcus_salivarius	C36	superpathway of aromatic amino acid biosynthesis	Streptococcus_thermophilus

C3	superpathway of N-acetylglucosamine, N-acetylmannosamine and N-acetylneuraminate degradation	Streptococcus_salivarius	C36	superpathway of guanosine nucleotides de novo biosynthesis I	Streptococcus_thermophilus
C3	superpathway of purine nucleotides de novo biosynthesis II	Streptococcus_salivarius	C36	superpathway of pyrimidine nucleobases salvage	Streptococcus_thermophilus
C3	L-valine biosynthesis	Streptococcus_vestibularis	C36	UDP-N-acetyl-D-glucosamine biosynthesis I	Streptococcus_thermophilus
C3	superpathway of L-aspartate and L-asparagine biosynthesis	Streptococcus_vestibularis	C36	UDP-N-acetylmuramoyl-pentapeptide biosynthesis II (lysine-containing)	Streptococcus_thermophilus
C3	chorismate biosynthesis I	Veillonella_atypica	C44	L-valine biosynthesis	Clostridium_citroniae
C3	gondoate biosynthesis (anaerobic)	Veillonella_atypica	C44	L-valine biosynthesis	Clostridium_citroniae
C3	inosine-5'-phosphate biosynthesis III	Veillonella_atypica	C44	pentose phosphate pathway (non-oxidative branch)	Clostridium_citroniae
C3	superpathway of L-lysine, L-threonine and L-methionine biosynthesis II	Veillonella_atypica	C44	CDP-diacylglycerol biosynthesis II	Lachnospiraceae_bacterium_1_4_56FAA
C3	chorismate biosynthesis I	Veillonella_dispar	C44	phosphopantothenate biosynthesis I	Lachnospiraceae_bacterium_1_4_56FAA
C3	superpathway of L-serine and glycine biosynthesis I	Veillonella_dispar	C44	superpathway of purine deoxyribonucleosides degradation	Lachnospiraceae_bacterium_1_4_56FAA
C3	superpathway of L-threonine biosynthesis	Veillonella_dispar	C44	adenosine ribonucleotides de novo biosynthesis	Lachnospiraceae_bacterium_2_1_58FAA
C3	superpathway of pyrimidine deoxyribonucleotides de novo biosynthesis (E. coli)	Veillonella_dispar	C44	adenosine ribonucleotides de novo biosynthesis	Lachnospiraceae_bacterium_2_1_58FAA
C6	5-aminoimidazole ribonucleotide biosynthesis I	Burkholderiales_bacterium_1_1_47	C44	guanosine ribonucleotides de novo biosynthesis	Lachnospiraceae_bacterium_2_1_58FAA
C6	5-aminoimidazole ribonucleotide biosynthesis II	Burkholderiales_bacterium_1_1_47	C44	phosphopantothenate biosynthesis I	Lachnospiraceae_bacterium_2_1_58FAA
C6	coenzyme A biosynthesis I	Burkholderiales_bacterium_1_1_47	C44	purine ribonucleosides degradation	Lachnospiraceae_bacterium_2_1_58FAA
C6	coenzyme A biosynthesis I	Burkholderiales_bacterium_1_1_47	C44	purine ribonucleosides degradation	Lachnospiraceae_bacterium_2_1_58FAA
C6	guanosine deoxyribonucleotides de novo biosynthesis II	Burkholderiales_bacterium_1_1_47	C44	pyruvate fermentation to isobutanol (engineered)	Lachnospiraceae_bacterium_2_1_58FAA
C6	guanosine ribonucleotides de novo biosynthesis	Burkholderiales_bacterium_1_1_47	C44	pyruvate fermentation to isobutanol (engineered)	Lachnospiraceae_bacterium_2_1_58FAA
C6	L-arginine biosynthesis I (via L-ornithine)	Burkholderiales_bacterium_1_1_47	C44	pyruvate fermentation to isobutanol (engineered)	Lachnospiraceae_bacterium_2_1_58FAA
C6	lipid IVA biosynthesis	Burkholderiales_bacterium_1_1_47	C44	5-aminoimidazole ribonucleotide biosynthesis I	Ruminococcus_gnavus
C6	L-lysine biosynthesis III	Burkholderiales_bacterium_1_1_47	C44	5-aminoimidazole ribonucleotide biosynthesis II	Ruminococcus_gnavus

C6	L-ornithine biosynthesis	Burkholderiales_bacterium_1_1_47	C44	5-aminoimidazole ribonucleotide biosynthesis II	Ruminococcus_gnavus
C6	L-ornithine biosynthesis	Burkholderiales_bacterium_1_1_47	C44	adenine and adenosine salvage III	Ruminococcus_gnavus
C6	methylerythritol phosphate pathway I	Burkholderiales_bacterium_1_1_47	C44	adenine and adenosine salvage III	Ruminococcus_gnavus
C6	pyruvate fermentation to isobutanol (engineered)	Burkholderiales_bacterium_1_1_47	C44	adenosine ribonucleotides de novo biosynthesis	Ruminococcus_gnavus
C6	queuosine biosynthesis	Burkholderiales_bacterium_1_1_47	C44	CDP-diacylglycerol biosynthesis I	Ruminococcus_gnavus
C6	queuosine biosynthesis	Burkholderiales_bacterium_1_1_47	C44	chorismate biosynthesis from 3-dehydroquinate	Ruminococcus_gnavus
C6	superpathway of 5-aminoimidazole ribonucleotide biosynthesis	Burkholderiales_bacterium_1_1_47	C44	coenzyme A biosynthesis I	Ruminococcus_gnavus
C6	superpathway of 5-aminoimidazole ribonucleotide biosynthesis	Burkholderiales_bacterium_1_1_47	C44	guanosine ribonucleotides de novo biosynthesis	Ruminococcus_gnavus
C6	superpathway of guanosine nucleotides de novo biosynthesis I	Burkholderiales_bacterium_1_1_47	C44	guanosine ribonucleotides de novo biosynthesis	Ruminococcus_gnavus
C6	superpathway of guanosine nucleotides de novo biosynthesis II	Burkholderiales_bacterium_1_1_47	C44	L-lysine biosynthesis III	Ruminococcus_gnavus
C6	superpathway of polyamine biosynthesis I	Burkholderiales_bacterium_1_1_47	C44	L-lysine biosynthesis VI	Ruminococcus_gnavus
C6	tetrapyrrole biosynthesis I (from glutamate)	Burkholderiales_bacterium_1_1_47	C44	peptidoglycan biosynthesis I (meso-diaminopimelate containing)	Ruminococcus_gnavus
C6	tetrapyrrole biosynthesis I (from glutamate)	Burkholderiales_bacterium_1_1_47	C44	purine ribonucleosides degradation	Ruminococcus_gnavus
C6	tRNA processing	Burkholderiales_bacterium_1_1_47	C44	pyruvate fermentation to isobutanol (engineered)	Ruminococcus_gnavus
C6	5-aminoimidazole ribonucleotide biosynthesis II	Veillonella_parvula	C44	S-adenosyl-L-methionine cycle I	Ruminococcus_gnavus
C6	adenosine deoxyribonucleotides de novo biosynthesis II	Veillonella_parvula	C44	starch degradation V	Ruminococcus_gnavus
C6	adenosine deoxyribonucleotides de novo biosynthesis II	Veillonella_parvula	C44	starch degradation V	Ruminococcus_gnavus
C6	adenosine ribonucleotides de novo biosynthesis	Veillonella_parvula	C44	superpathway of 5-aminoimidazole ribonucleotide biosynthesis	Ruminococcus_gnavus
C6	adenosine ribonucleotides de novo biosynthesis	Veillonella_parvula	C44	superpathway of 5-aminoimidazole ribonucleotide biosynthesis	Ruminococcus_gnavus
C6	guanosine deoxyribonucleotides de novo biosynthesis II	Veillonella_parvula	C44	UDP-N-acetylmuramoyl-pentapeptide biosynthesis I (meso-diaminopimelate containing)	Ruminococcus_gnavus
C6	guanosine ribonucleotides de novo biosynthesis	Veillonella_parvula	C44	UDP-N-acetylmuramoyl-pentapeptide biosynthesis II (lysine-containing)	Ruminococcus_gnavus
C6	L-arginine biosynthesis II (acetyl cycle)	Veillonella_parvula	C44	UMP biosynthesis	Ruminococcus_gnavus

Supplementary Table 14: Dietary nodes and features in the WGCNA.

Node	Dietary component	Node	Dietary component
D1	Tomato based sauces e.g. pasta sauces	D6	Green salad, lettuce
D1	White Rice	D6	Jam, marmalade, honey, syrup (teaspoon)
D1	Curry-type sauces	D6	Non-ready to eat - Porridge, Readybrek
D1	Baked beans	D6	Mineral supplements (details)
D1	White/yellow or green pastas (e.g. spagetti, noodles)	D6	Strawberries, raspberries, kiwi fruit
D1	Cocoa, Hot Chocolate (cup)	D6	Cucumber, celery
D1	Lasagne (meat based)	D6	Sweet peppers
D1	Port, Sherry, Vermouth, liqueurs (glass)	D6	Wine (glass)
D2	Sugar added to tea coffee, cereals (teaspoons)	D6	Peas
D2	Boiled, instant or jacket potatoes	D6	Grapes
D2	White bread and rolls (including ciabatta & pannini)	D6	Blueberries
D2	Chocolates, singles or squares	D6	Peanuts or other nuts
D2	Chocolate coated sweet biscuits e.g. digestive (one)	D6	Wheat-free; Rye brad; spelt bread (specify)
D2	Mashed potatoes	D6	Mushrooms
D2	Plain biscuit e.g. marietta, digestives, rich tea (one)	D6	Full-fat yoghurt or Greek style yoghurt (125g carton)
D2	High fibre (All Bran/flakes, Weetabix, Shredded wheat)	D6	Spinach
D2	Chips / Roast potatoes	D6	Salad cream, mayonnaise (tablespoon)
D2	Sweets, toffees, mints	D6	Vegetable soups: homemade/fresh (1 bowl)
D2	Crisps or other packet snacks	D6	Dried fruit e.g. raisins
D2	Fruit squash (small glass)	D6	Beetroot
D2	Cornflakes, Rice Krispies	D6	Sweetcorn
D2	Breaded chicken, chicken nuggets, chicken burger	D6	Green beans, broad beans, runner beans
D2	Ice cream, choc ices, frozen desserts	D6	Courgettes
D2	Fizzy Soft drinks e.g. (Coca Cola - Glass)	D6	Cream (tablespoon)
D2	Takeaway (specify)	D6	Peanut butter (teaspoon)
D2	Sugar-coated cereals (e.g. Frosties, Crunchy nut cornflakes)	D6	Muesli (e.g. Country Store, Aplen)
D3	Lite Butter e.g. Dawn Lite, Connacht Gold	D6	Brown Rice

D3	Cream & Veg. Oil spread e.g. Kerrymaid, Dairy Gold	D6	Coleslaw
D3	Olive oil spread e.g. Golden Olive	D6	Dried lentils, beans, peas
D3	Sunflower margarine e.g. Flora	D6	Other salad dressing (tablespoon)
D3	Cholesterol Lowering e.g. Flora Pro-Active, DairyGold Heart	D6	Spirits e.g. Gin, Whiskey (Single measure)
D3	Low-fat margarine (e.g. lowlow)	D6	Melon
D3	Low alcohol / alcohol free beer / larger (half pint)	D6	French dressing (tablespoon)
D4	Oily fish, fresh or canned (e.g. mackerel, kippers, tuna, salmon, sardines, herring)	D6	Potato Salad
D4	White fish, fresh or frozen (e.g. cod, haddock, plaice, sole)	D6	Frozen fruit (specify)
D4	Fish fried (batter/breadcrumbs), baked, grilled, fingers/cakes	D6	Peaches, plums
D4	Shellfish (e.g. crab, prawns, mussels)	D6	Pickles, chutney (tablespoon)
D4	Liver, heart, kidney, paté	D6	Wholemeal pasta
D5	Carrots	D6	Leeks
D5	Eggs as boiled, fried, scrambled, poached (one)	D6	Apricots
D5	Apples	D6	Grapefruit
D5	Oranges, satsumas, mandarins	D6	Cottage cheese
D5	Broccoli, spring greens, kale	D6	Tofu, soya meat, TVP, vegeburger
D5	Pears	D6	Lasagne (vegetarian)
D5	Parsnips, turnips	D7	Chicken or other poultry e.g. turkey: roast
D5	Cauliflower	D7	Sugar substitute e.g. canderel (teaspoon)
D5	Cabbage	D7	Low calorie or diet soft fizzy drink (glass)
D5	Brussel sprouts	D7	Sauces e.g. white, cheese, gravy (tablespoon)
D6	Whole milk (cup) - cow, goat, soya, rice milk - specify	D7	Bacon / Ham
D6	Coffee ground (cup)	D7	Processed meat (Corned beef, Luncheon meats, sausages)
D6	Tea (cup)	D7	Beef (roast / steak)
D6	Onions	D7	Beef burger (1 burger)
D6	Bananas	D7	Beef: stew
D6	Tomatoes	D7	Pork (roast / chops / escalopes)
D6	Garlic	D7	Lamb (roast / chops / stew)

Supplementary Table 15: Urine MS metabolomic nodes and features in the WGCNA.

Node	Urine metabolite	Node	Urine metabolite
E1	2-(2-Phenylacetoxy)propionylglycine	E7	UDP-4-dehydro-6-deoxy-D-glucose
E1	Luteolinidin	E7	A 80987
E1	N-Succinyl-2-amino-6-ketopimelate	E7	Medicagenic acid 3-O-b-D-glucuronide
E1	2-Methoxycarbonylphenyl beta-D-glucopyranoside	E7	1,4,5-Trimethyl-naphtalene
E1	Citpressine II	E7	1,6,7-Trimethylnaphthalene
E1	[Fv Dimethoxy,methyl(9:1)] (2S)-5,7-Dimethoxy-3',4'-methylenedioxyflavanone	E7	Methyl bisnorbiotinyl ketone
E1	Dolichyl diphosphate	E8	4'-Methoxychalcone
E1	Torasemide	E8	Methylgallic acid-O-sulphate
E1	Glu-Asp-Asp	E8	Lys-Ala-Gly
E1	ADP	E8	Dicaffeoylputrescine
E1	Trp-Cys-Tyr	E8	Ala-Trp-Asn-Asn
E1	5-Carboxypyranopelargonidin 3-O-beta-glucopyranoside	E8	Gallic acid
E1	Garciduol B	E8	Tetrahydrodipicolinate
E1	Cys-Glu-Glu-Glu	E8	3-[4-Hydroxy-3-(3-methyl-2-butenyl)phenyl]-2-propenal
E1	Melitric acid B	E8	4,9-dimethyldodeca-2,4,6,8,10-pentaene-1,12-dial
E1	Demethyloleuropein	E8	2-Hydroxydecanedioic acid
E1	dTDP-3-amino-2,3,6-trideoxy-C-methyl-D-erythro-hexopyranos-4-ulose	E9	N-Carboxyacetyl-D-phenylalanine
E1	Marmesin rutinoside	E9	3-Methylthiopropionic acid
E1	Cys-Trp-Cys-Tyr	E9	L-Tyrosine methyl ester 4-sulfate
E1	Negletein 6-[rhamnosyl-(1->2)-fucoside]	E9	Phlorisobutyrophenone 2-glucoside
E1	Malvidin 3-laminaribioside	E9	Swertiamarin
E1	5,2',4'-Trihydroxy-3,7,5'-trimethoxyflavone 2'-galactosyl-(1->4)-glucoside	E9	D-Galactosyl-3-N-acetyl-beta-D-galactosamine
E1	Tricetin 3'-methyl ether 7,5'-diglucuronide	E9	2,3,4,5,6-Penta-O-acetyl-D-glucose
E1	Kaempferol 3-[2''',3''',5'''-triacetyl-alpha-L-arabinofuranosyl-(1->6)-glucoside	E9	Petunidin 3-glucoside
E1	Delphinidin 3-(6''-O-4-malyl-glucosyl)-5-glucoside	E9	Asp-Trp-Cys-Gly
E2	4-[2,2'-Bithiophen-5-yl]-3-butyn-1-ol	E9	Ala-Trp-Cys-Cys
E2	3-Oxo-carbofuran	E9	Asn-Thr-Asp-His

E2	Fumaric acid
E2	Pivaloylcarnitine
E2	Lenticin
E2	N-Phenylacetylaspartic acid
E2	L-Phosphoarginine
E2	Isoeugenol benzyl ether
E2	N-(3-oxooctanoyl)-L-homoserine
E2	L-Hexanoylcarnitine
E2	Caffeic acid 3-sulfate
E2	Glucose 6-phosphate
E2	Dihydrocaffeic acid 3-sulfate
E2	Dehydro-4-methoxycyclobrassinin
E2	5'-Dehydroadenosine
E2	Naringenin
E2	Phenylethylamine
E2	Ferulic acid 4-sulfate
E2	Alloathyriol
E2	(1xi,3S)-1,2,3,4-Tetrahydro-1-methyl-beta-carboline-1,3-dicarboxylic acid
E2	6-Phosphogluconic acid
E2	N-[4'-hydroxy-(E)-cinnamoyl]-L-aspartic acid
E2	p-Anisidine
E2	N4-Acetylcytidine
E2	5'-(3',4'-Dihydroxyphenyl)-gamma-valerolactone sulfate
E2	D-Sedoheptulose 7-phosphate
E2	Pyruvaldehyde
E2	2-Phenylethanol glucuronide
E2	2,6 Dimethylheptanoyl carnitine
E2	5-(3',4',5'-Trihydroxyphenyl)-gamma-valerolactone-4'-O-sulphate
E2	2-(3,4-Dihydroxybenzoyloxy)-4,6-dihydroxybenzoate
E2	4'-Hydroxy-5,5'-diisopropyl-2,2'-dimethyl-3,4-biphenylquinone
E9	Thr-Thr-Thr-Trp
E9	Gln-Leu-Met-Met
E9	His-Trp-Pro-Pro
E9	Afzelechin-(4alpha->8)-afzelechin
E9	Hippuric acid
E9	3-(3-Hydroxyphenyl)-3-hydroxypropanoic acid
E9	Prenyl benzoate
E10	hesperetin 3'-O-sulfate
E10	Propanal
E10	2-Phenylethyl octanoate
E10	Beta-Citryl-L-glutamic acid
E10	3beta,6beta-Dihydroxynortropane
E10	Silicic acid
E10	Malonic acid
E10	Tiglylcarnitine
E10	4,6-Decadiyn-1-ol isovalerate
E10	Inositol cyclic phosphate
E10	[FA (15:4)] 6,8,10,12-pentadecatetraenal
E10	Citronellyl cinnamate
E10	2-Hydroxyundecanoate
E10	4-(Trimethylammonio)but-2-enoate
E11	2-trans,4-cis-Decadienoylcarnitine
E11	Trp-Ala-Ser
E11	Isopentenyladenine-9-N-glucoside
E11	Leu-Met-Asp
E11	1-(alpha-Methyl-4-(2-methylpropyl)benzeneacetate)-beta-D-Glucopyranuronic acid
E11	Sesaminol glucoside
E11	Dalpanin
E11	Italidipyrone
E11	Cys-Phe-Phe-Gln

E2	Pyrroline hydroxycarboxylic acid
E2	2,3-Dihydro-6-hydroxy-2-(4-hydroxyphenyl)-5,7-dimethoxy-4H-1- benzopyran-4-one
E2	Melanin
E2	3'-UMP
E2	2-O-p-Coumaroyl-D-glucose
E2	Miraxanthin-III
E2	Deoxyribose
E2	4-hydroxymidazolam
E2	Sucrose
E2	beta-Cellobiose
E2	Theogallin
E2	2,3-dioctanoylglyceramide
E2	Trigonelline
E2	Phe-Thr-Val
E2	[Fv Methyl,tetramethoxy(4:0)] 3',4'-Methylenedioxy-2,4,6,beta-tetramethoxychalcone
E2	Menadiol disuccinate
E2	5-(3',5'-Dihydroxyphenyl)-gamma-valerolactone 3-O-glucuronide
E2	S-Adenosylhomocysteine
E2	Hydroxylansoprazole
E2	dUDP
E2	Thiethylperazine
E2	dTDP
E2	Armillaric acid
E2	hydroxymethyl-dCDP
E2	Cys-Thr-Thr-Cys
E2	Asn-Cys-Pro-Pro
E2	2-Hydroxyglutarate
E2	L-Arabinono-1,5-lactone
E2	Tolmetin glucuronide
E2	D-olivose
E11	Harderoporphyrin
E11	Arg-Phe-Trp-Cys
E11	Phe-Trp-Pro-Tyr
E11	Arg-Trp-Asp-His
E12	Chlorphenesin carbamate
E12	D-Lactaldehyde
E12	[Fv] Ovalichromene A
E12	Garcimangosone D
E12	Cys-Trp-Cys-His
E12	1-Methyluric acid
E12	7-Methyluric acid
E12	Benzophenone
E12	Sorbitol
E12	(+/-)-3-[(2-methyl-3-furyl)thio]-2-butanone
E12	Gluconic acid
E12	4-[(2-Furanylmethyl)thio]-2-pentanone
E12	4-(2-Pyrazinylethenyl)phenol
E13	3,4-Methylenesebacic acid
E13	Creatinine
E13	Isoleucyl-Proline
E13	Ile-Pro
E13	Diazoxide
E13	Valyl-Isoleucine
E13	Isobutyryl-L-carnitine
E13	L-proline amide
E13	N-(1-Deoxy-1-fructosyl)glycine
E13	phosphate-ester-of-dihydrofuran
E13	Ethanolamine
E13	alpha-(p-Methoxyphenyl)-4-pyridineacrylic acid
E13	Ascorbic acid-2-sulfate

E2	Artobiloxanthone
E2	Asparagusic acid
E2	(S,S)-Tartaric acid
E2	L-Arabinose
E2	Pentose
E2	Gln-Met-Pro-Ser
E2	CDP-glycerol
E2	Cys-Pro-Pro-Tyr
E2	6"-O-Malonyldaidzin
E2	N-Methylpyridinium
E2	D-Fucono-1,4-lactone
E2	(-)-1-Methylpropyl 1-propenyl disulfide
E2	Acetylcysteine
E2	2-Hydroxypyridine
E2	3-hydroxypyridine
E2	Rhamnose
E2	2-Deoxy-D-glucose
E2	5-Ethyl-1,2,3,4,5,6-hexahydro-7H-cyclopenta[b]pyridin-7-one
E2	3'-Sialyllactosamine
E2	Pyridoxal
E2	2-Furoylglycine
E2	N-Butyryl-L-homoserine lactone
E2	Hexanoylglycine
E2	Shikimic acid
E2	3-Isopropylmalate
E2	Allitridin
E2	L-sorbosone 1,4-lactone
E2	D-Mannosamine
E2	Hexose
E2	6-Thiourate

E13	4-Methoxybenzyl phenylacetate
E13	Succinic acid
E13	Ile-Lys
E13	Leucyl-Methionine
E13	Thr-Gly-Ser
E13	3-Methoxy-4-hydroxyphenylethyleneglycol sulfate
E13	L-Homoserine
E13	Glutarylcarnitine
E13	D-Threitol
E13	(R)-Roemerine
E13	N-(1-Deoxy-1-fructosyl)threonine
E13	L-Octanoylcarnitine
E13	(+-)-Dalbergioidin
E13	Testosterone
E13	Thymine
E13	D-1-Piperideine-2-carboxylic acid
E13	Pimelylcarnitine
E13	N-Acetylneuraminic acid
E13	9-Decenoylcarnitine
E13	Glutaconic acid
E13	nva-FMDP
E13	Alizapride
E13	N-Acetylputrescinium
E13	Creatine
E13	Leucine/Isoleucine
E13	2,4,6-Octatriynoic acid
E13	6-Keto-decanoylcarnitine
E13	L-Asparagine
E13	5-(3',4',5'-Trihydroxyphenyl)-gamma-valerolactone-O-methyl-O-sulphate
E13	5,5'-dehydrodivanillate

E2	DL-O-Phosphoserine
E2	Pyrocatechol sulfate
E2	Dimethylglycine
E2	sedoheptulose anhydride
E2	5'-Hydroxycotinine
E2	L-Dopa
E2	Diethanolamine
E2	Glycyl-Lysine
E2	Benzylamine
E2	(Z)-[3-(Methylsulfinyl)-1-propenyl] 2-propenyl disulfide
E2	1,3,7-Trimethyluric acid
E2	L-prolyl-L-proline
E2	5-methyl-DL-tryptophan
E2	S-(2,5-Dimethyl-3-furanyl) 2-furancarbothioate
E2	6-Acetyl-D-glucose
E2	Acetaldehyde butyl phenethyl acetal
E3	gamma-Fagarine
E3	Urea
E3	3-Methylglutarylcarnitine
E3	4-Hydroxy-5-methyl-3(2H)-thiophenone
E3	4-Hydroxyproline
E3	S-Adenosylmethioninamine
E3	Succinyladenosine
E3	[Fv methoxy(6:0)] 2,3,4,2',4',6'-Hexamethoxychalcone
E3	Trp-Val-Val
E3	Cys-Lys-Phe-Gly
E3	3-Hydroxyphenylacetic acid
E3	Cholesterol sulfate
E3	Ala-Lys-Phe-Cys
E3	Met-Tyr-Tyr
E13	Glycine
E13	L-Aspartic acid
E13	4-Hydroxy-5-(dihydroxyphenyl)-valeric acid-O-methyl-O-sulphate
E13	L-Malic acid
E13	Glycolic acid
E13	Leu-Leu-Thr
E13	Hypoxanthine
E13	2-Methylnicotinamide
E13	Trp-Gly-Pro
E13	3-Methoxy-4-hydroxyphenylglycol glucuronide
E13	Ala-Asn-Cys-Gly
E13	Urocanic acid
E13	Thioridazine
E13	Ethyl methyl trisulfide
E13	Riboflavin
E13	5,7-Dihydroxy-3',4'-dimethoxy-5'-prenylflavanone
E13	1-Sinapoyl-D-glucose
E13	cyclic N-Acetylserotonin glucuronide
E13	Asp-Cys-Cys-Gly
E13	Glu-Cys-Phe
E13	Cyclopropanecarboxylate
E13	Isobutyrylglycine
E13	3-Dehydroxycarnitine
E13	Oxoglutaric acid
E13	L-Glutamine
E13	L-Lysine
E13	L-Glutamic acid
E13	L-Methionine
E13	3,5,7,9,11,13,15,17-octaoxo-eicosanoyl-[acp]
E13	D-Xylulose

E3	L-Lactic acid
E3	(7'R,8'R)-4,7'-Epoxy-3'-methoxy-4',5,9,9'-lignanetetrol 9'-glucoside
E3	His-Leu-Trp-Gly
E3	LysoPC(18:2(9Z,12Z))
E3	Divanillyltetrahydrofuran ferulate
E3	Trp-Trp-His
E3	Pimelic acid
E3	Acetylhomoserine
E3	Arg-Ile-Ile-Tyr
E3	Mesobilirubinogen
E3	2-Dimethylamino-5,6-dimethylpyrimidin-4-ol
E3	Valyl-Glycine
E3	D-Glucuronic acid
E3	Cysteine-S-sulfate
E3	Isopropyl hexanoate
E3	Isopropyl isovalerate
E3	Levulinic acid
E3	2-Methylbutylacetate
E3	2-Ethyl-1-hexanol
E3	cis-4-Decenal
E3	Isoamylbutyrate
E3	3,7-Dimethyl-1,3,6-octatriene
E3	Ethyl isobutyrate
E3	Citronellol
E3	2-Nonanone
E3	2-Octanone
E3	d-Piperitone
E3	5-Methyl-2,3-hexanedione
E3	Heptyl alcohol
E3	Allyl hexanoate
E13	N1-Methyl-4-pyridone-3-carboxamide
E13	L-Arabitol
E13	Dopamine
E13	Glycerol 1,2-cyclic phosphate
E13	3,5-Dihydroxybenzoic acid
E13	Glu-Phe-Trp
E13	(-)-Epigallocatechin 3'-glucuronide
E13	L-Alanine
E13	Asn-Phe-Phe-Gly
E13	Oxalic acid
E13	Cys-Leu-Pro-Tyr
E13	L-Histidine
E13	succinate semialdehyde-thiamine diphosphate
E13	delta-Guanidinovaleric acid
E13	[FA amino(8:0)] 3-amino-octanoic acid
E13	Oxoadipic acid
E13	Tetrahydroaldosterone-3-glucuronide
E13	(1S,3R,4S)-3,4-Dihydroxycyclohexane-1-carboxylate
E13	Cortolone-3-glucuronide
E13	8-Hydroxyluteolin 8-glucoside 3'-sulfate
E13	1-N6-ethanoadenine
E13	L-Carnitine
E13	Isoscutellarein 4'-methyl ether 8-(2''-sulfatoglucuronide)
E13	3-Methyldioxyindole
E13	2-Dehydro-D-xylonate
E13	S-2-Propenyl 1-propanesulfinothioate
E13	[FA hydroxy(4:0/5:0)] 2,3,4,4-tetrahydroxy-pentanoic acid
E13	S-Propyl 1-propanesulfinothioate
E13	3-(4-Methyl-3-pentenyl)thiophene
E13	Homocysteinesulfinic acid

E3	cis-3-Hexenyl butyrate
E3	Linalyl formate
E3	3,5-Dimethyl-1,2-cyclopentadione
E3	Heptane
E3	n-Nonane
E3	n-Decane
E3	n-Dodecane
E3	n-Hexadecane
E3	Allyl nonanoate
E3	Methyl citronellate
E3	2-methyl-Butanoic acid, 2-methylbutyl ester
E3	Butanoic acid, 1-methylethyl ester
E3	Isobutyl acetoacetate
E3	2-methyl-Propanoic acid, 3-methylbutyl ester
E3	Hexyl 2-furoate
E3	3-methyl-4-Heptanone
E3	2-Heptanone
E3	1,1-dimethoxy-Heptane
E3	Butanoic acid, 1-methylhexyl ester-
E4	Homocarnosine
E4	2-Methoxy-6-methylpyrazine
E4	Taurine
E4	dIMP
E4	Tyr-His-His
E4	Quercetin 7-glucuronide 3-rhamnoside
E4	2-Pyrroloylglycine
E4	Phosphodimethylethanolamine
E4	3-Methylhistidine
E4	1-Methylhistidine
E4	Menadione
E13	6'-Sialyllactose
E13	Uric acid
E13	cis-Aconitic acid
E13	L-Arginine
E13	N-Acetyl-L-aspartic acid
E13	Caffeic acid
E13	5-Hydroxyisourate
E13	2-Ketobutyric acid
E13	Pelargonidin 3-rutinoside-7-(6-(p-hydroxybenzoyl)glucoside)
E13	Allyl cinnamate
E13	Azelaic acid
E13	N-Alpha-acetyllysine
E13	N6-Acetyl-L-lysine
E13	Dimethyltryptamine
E13	N6,N6,N6-Trimethyl-L-lysine
E13	Glutarylglycine
E13	2 -(Butylamido)-4-hydroxybutanoic acid
E13	Homocitrulline
E13	N-carbamoyl-L-glutamate
E13	5,6-Dihydroxyindole-2-carboxylic acid
E13	4-Methyldibenzothiophene
E13	L-Serine
E13	Glyceric acid
E13	Spermine dialdehyde
E13	Capryloylglycine
E13	Asymmetric dimethylarginine
E13	(+)-gamma-Hydroxy-L-homoarginine
E13	4-(Methylnitrosamino)-1-(3-pyridyl)-1-butanone
E13	Hydroxyphenylacetylglycine
E13	Galactaric acid

E4	carboxyspermidine
E5	Dimethylallylpyrophosphate/Isopentenyl pyrophosphate
E5	Gamma-glutamyl-Cysteine
E5	Phenol glucuronide
E5	Tyrosyl-Cysteine
E5	Estrone-3,4-quinone
E5	Osmundalin
E5	Asn-Cys-Gly
E5	N-vanillate-L-glutamate
E5	p-nitrophenyl-α-D-glucopyranoside
E5	uridine 3'5'-cyclic monophosphate
E5	dUMP
E5	Phe-Gly-Gly-Ser
E5	Met-Met-Cys
E5	[PI (2:0)] 1,2-diacyl-sn-glycero-3-phospho-(1'-myo-inositol)
E5	Sumiki's acid
E5	Uridine 5'-diphosphate
E5	2-Keto-glutaramic acid
E5	5-((2-iodoacetamido)ethyl)-1-aminonapthalene sulfate
E5	Tizoxanide glucuronide
E5	Glisoxepide
E5	Acetoin
E5	Asp-Met-Asp-Pro
E5	11-beta-Hydroxyandrosterone-3-glucuronide
E5	His-Leu-Lys-Ser
E5	3,5-Di-O-galloyl-1,4-galactarolactone
E5	Kaempferol 3-sulfate-7-alpha-arabinopyranoside
E5	Pro-Trp-Gly-Tyr
E5	Allantoin
E5	Potassium 2-(1'-ethoxy) ethoxypropanoate
E13	D-Glucaric acid
E13	Valyl-Proline
E13	N-a-Acetyl-L-arginine
E13	Propionylcarnitine
E13	Uracil
E13	3-Hydroxysebacic acid
E13	O-Succinyl-L-homoserine
E13	Pantothenic acid
E13	5-Hydroxy-L-tryptophan
E13	N-Acetyl-D-galactosamine
E13	(R)-2,3-Dihydro-3,5-dihydroxy-2-oxo-3-indoleacetic acid
E13	(4-Hydroxybenzoyl)choline
E14	sulfoethylcysteine
E14	D-Ribulose 5-phosphate
E14	N-acetyltryptophan
E14	Asparaginyl-Aspartate
E14	Phenylacetylglutamine
E14	Asparenyol
E14	2-O-Benzoyl-D-glucose
E14	Indoleacetyl glutamine
E14	Quinoline
E14	Dihyroxy-1H-indole glucuronide I
E14	2-Deoxy-L-ribono-1,4-lactone
E14	(Â±)-2-Hydroxy-2-phenylacetonitrile
E14	2,8-Dihydroxyquinoline-beta-D-glucuronide
E14	glucosamine-1,6-diphosphate
E14	6-Hydroxy-5-methoxyindole glucuronide
E14	(+)-Chebulic acid
E14	Grevilline D
E14	Quinestrol

E5	Delphinidin 3-O-3",6"-O-dimalonylglucoside
E5	Psoralen
E5	5-(Methylthio)-2-[(methylthio)methyl]-2-pentenal
E5	2-Butylbenzothiazole
E5	Camalexin
E5	cis-4-Decenedioic acid
E5	4,6-Dihydroxy-2-quinolinecarboxylic acid
E5	Cysteinyl-Cysteine
E6	Deoxycytidine
E6	4-Methoxyphenylethanol sulfate
E6	L-Valine
E6	1-(3-aminopropyl)-pyrrolinium
E6	Thr-Cys-Cys
E6	Cys-Phe-Gly-Gly
E6	Asp-Asp-Gly-His
E6	Asp-Pro-Ser-His
E6	Glu-Asp-Asp-Asp
E6	Asn-Met-Asp-Asp
E6	Asp-Met-Asp-Asp
E6	Imidazolelactic acid
E6	5'-Guanosyl-methylene-triphosphate
E6	methylphosphonate
E7	N-Undecanoylglycine
E7	Butoctamide hydrogen succinate
E7	Decanoylcarnitine
E7	(-)-Epicatechin sulfate
E7	Trp-Ala-Pro
E7	Dodecanedioylcarnitine
E7	[FA (18:0)] N-(9Z-octadecenoyl)-taurine
E7	Ala-Leu-Trp-Gly
E14	Urate D-ribonucleotide
E14	Axillarin 7-sulfate
E14	6-Hydroxyluteolin 6-xyloside
E14	Met-Trp-Cys
E14	Viniferifuran
E14	Xanthine
E14	4-Deoxy-beta-D-gluc-4-enuronosyl-(1,3)-N-acetyl-D-galactosamine 4-sulfate
E14	D-Galactopyranosyl-(1->3)-D-galactopyranosyl-(1->3)-L-arabinose
E14	Cyanidin-3-galactoside
E14	Phe-Tyr-Tyr
E14	8-methylsulfinyloctyl glucosinolate
E14	Protoporphyrin IX
E14	Cys-Met-Trp-Gln
E14	Met-Met-Thr-Trp
E14	Gln-Leu-Met-Trp
E14	Asn-Met-Tyr-Tyr
E14	D-Urobilinogen
E14	L-Urobilin
E14	L-Urobilinogen
E14	Arg-Phe-Trp-Pro
E14	Kaempferol 3-(2",3"-diacetyl-4"-p-coumaroylrhamnoside)
E14	Luteolin 7-(6"-malonylneohesperidoside)
E14	Malvidin 3-glucoside-5-(6-acetylglucoside)
E14	8-Hydroxy-7-methylguanine
E14	p-Cresol sulfate
E14	O-methoxycatechol-O-sulphate
E14	3-Indolehydracrylic acid
E14	dolichol
E14	1-Deoxy-D-xylulose 5-phosphate
E14	4-hydroxybenzoic acid-4-O-sulphate

E7	dCTP	E14	N-Acetyl-L-tyrosine

Supplementary Table 16: Faecal MS metabolomic nodes and features in the WGCNA.

Node	Faecal metabolite	Node	Faecal metabolite
F1	Adenine	F7	L-Glutamine
F1	Quercetin 3'-sulfate	F7	L-Lysine
F1	LysoPC(14:1(9Z))	F7	Glu-Ala-Gln-Ser
F1	Lucyoside Q	F7	L-Methionine
F1	12-O-Palmitoyl-16-hydroxyphorbol 13-acetate	F7	Acetoin
F1	PG(29:1)	F7	L-Alanine
F1	PS(P-32:0)	F7	Asn-Ser-His-His
F1	PS(O-20:0/12:0)	F7	Glu-Glu-Gly-Tyr
F1	PE(18:4(6Z,9Z,12Z,15Z)/P-18:0)	F7	1-18:1-lysophosphatidylglycerol
F1	PC(32:2)	F7	[FA amino(8:0)] 3-amino-octanoic acid
F1	SM(d18:0/18:0)	F7	Alanyl-Alanine
F1	PS(18:0/14:0)	F7	Arg-Lys-Val-Tyr
F1	N-(15Z-tetracosenoyl)-4E,6E-hexadecasphingadienine-1-phosphoethanolamine	F7	Tuberoside
F1	PS(O-34:3)	F7	L-Phenylalanine
F1	PG(P-18:0/17:2(9Z,12Z))	F7	S-Propyl 1-propanesulfinothioate
F1	PS(O-34:2)	F7	3-deoxy-D-galactose
F1	PG(16:1(9Z)/18:0)	F7	Pyridoxamine
F1	1-18:0-2-16:3-monogalactosyldiacylglycerol	F7	Hexanoylglycine
F1	1-18:1-2-16:1-monogalactosyldiacylglycerol	F7	Citrulline
F1	PG(O-36:4)	F7	Tyrosine
F1	PG(O-36:3)	F7	p-Cresol sulfate
F1	PG(35:2)	F7	N6-Acetyl-L-lysine
F1	PS(34:1)	F7	N5-acetyl-N5-hydroxy-L-ornithine
F1	PG(13:0/22:1(11Z))	F7	Dimethylglycine

F1	PA(O-20:0/22:6(4Z,7Z,10Z,13Z,16Z,19Z))	F7	Gluconic acid
F1	PC(18:4(6Z,9Z,12Z,15Z)/P-18:1(11Z))	F7	N-Acetylhistidine
F1	PA(P-42:4)	F7	L-Serine
F1	PC(o-16:1(9Z)/20:4(8Z,11Z,14Z,17Z))	F7	1-Isothiocyanato-7-(methylthio)heptane
F1	PS(O-18:0/18:4(6Z,9Z,12Z,15Z))	F7	O-methoxycatechol-O-sulphate
F1	PC(18:1(11Z)/P-18:1(11Z))	F7	L-Tryptophan
F1	PC(20:1(11Z)/P-16:0)	F7	Valyl-Serine
F1	PG(36:3)	F7	Serinyl-Threonine
F1	PA(41:0)	F7	Valyl-Valine
F1	PG(36:1)	F7	Threoninyl-Valine
F1	PC (36:6)	F7	1-Decanol
F1	3-O-Sulfogalactosylceramide (d18:1/16:0)	F7	Furfuryl octanoate
F1	PI(P-18:0/22:6(4Z,7Z,10Z,13Z,16Z,19Z))	F7	trans-Carvone oxide
F1	3-Hydroxyphenyl-valeric acid	F8	Xylitol 5-phosphate
F2	Deoxycytidine	F8	Uridine
F2	Traumatic acid	F8	Succinic acid
F2	Myristic acid	F8	Adenosine
F2	[FA amino(13:0)] 2R-aminotridecanoic acid	F8	Inosine
F2	(1xi,3xi)-1,2,3,4-Tetrahydro-1-methyl-beta-carboline-3-carboxylic acid	F8	Niacinamide
F2	Dodecanedioic acid	F8	Guanosine
F2	Benzeneacetamide-4-O-sulphate	F8	Pyrroline hydroxycarboxylic acid
F2	L-Proline	F8	Leucine/Isoleucine
F2	L-Valine	F8	Asn-Cys-Cys
F2	Dehydro-4-methoxycyclobrassinin	F8	Ala-Asn-Asn-Gly
F2	Phenylacetylglutamine	F8	Cys-Thr-Ser-Ser
F2	16-Hydroxy hexadecanoic acid	F8	L-Glutamic acid
F2	3-Hydroxytetradecanedioic acid	F8	Guanine
F2	2-pyridylethylamine	F8	3-sulfopropanoate
F2	Linoleamide	F8	Cys-Lys-Cys-His

F2	Thr-Tyr	F8	N-Acetyl-L-aspartic acid
F2	Oleic acid	F8	carboxyspermidine
F2	2-Phenylethyl beta-D-glucopyranoside	F8	2-Hydroxyundecanoate
F2	[FA hydroxy(17:0)] heptadecanoic acid	F8	2-acetamidoglucal
F2	cis-Hydroxy Perhexiline	F8	trans-2-Heptenal
F2	17-Hydroxylinolenic acid	F9	2-(Ethylsulfonylmethyl)phenyl methylcarbamate
F2	Avenoleic acid	F9	Nicotinic acid
F2	8-Deoxy-11-hydroxy-13-chlorogrosheimin	F9	Tyrosyl-Cysteine
F2	7C-aglycone	F9	Hypoxanthine
F2	Gravelliferone	F9	Riboflavin
F2	Ricinoleic acid	F9	D-Ribose
F2	2-(3-carboxy-3-(trimethylammonio)propyl)-L-histidine	F9	Xanthine
F2	xi-10-Hydroxyoctadecanoic acid	F9	Met-Trp-Val-Val
F2	uridine 3'5'-cyclic monophosphate	F9	Psoralen
F2	trans-O-Methylgrandmarin	F9	cis-2-Methylaconitate
F2	Thr-Cys-Ser	F9	Uracil
F2	Eicosenoic acid	F9	Pantothenic acid
F2	Ala-Gln-Pro	F9	N-Acetyl-D-galactosamine
F2	Lys-Thr-Ala	F9	Octanoic acid
F2	[FA (21:2)] 12Z,15Z-heneicosadienoic acid	F10	[FA hydroxy(14:0)] 11S-hydroxy-tetradecanoic acid
F2	1-Acetoxy-2-hydroxy-16-heptadecyn-4-one	F10	[FA hydroxy(16:0)] (+)-15S-hydroxy-hexadecanoic acid
F2	Thr-Cys-Cys	F10	Stearic acid
F2	2-Methyl-3-hydroxyvaleric acid	F10	[FA hydroxy(20:0)] 3-hydroxy-eicosanoic acid
F2	gamma-Glutamyl-S-methylcysteinyl-beta-alanine	F10	FA (22:2)
F2	4-Hydroxy-5-(dihydroxyphenyl)-valeric acid-O-methyl-O-sulphate	F10	FA oxo(22:0)
F2	Erucic acid	F10	[FA hydroxy(22:0)] 13-hydroxy-docosanoic acid
F2	Ala-Val-Val-Ala	F10	Tetracosanoic acid
F2	Tetracosapentaenoic acid (24:5n-6)	F10	FA hydroxy(24:0)
F2	[FA oxo,hydroxy,dimethyl(2:0)] 9-oxo-15R-hydroxy-16,16-dimethyl-10Z,13E-prostadienoic acid	F10	N-stearoyl valine

F2	trans-2-Tetradecenoylcarnitine	F10	[FA hydroxy(24:0)] 3-hydroxy-tetracosanoic acid
F2	Mammea B/BA	F10	N-oleoyl histidine
F2	[ST Hydrox] 3alpha-Hydroxy-5beta-chol-7-en-24-oic Acid	F10	13'-Carboxy-gamma-tocopherol
F2	10'-Apo-beta-carotenal	F10	delta-Decalactone
F2	Glu-Asp-Asp	F11	Creatinine
F2	[PI (2:0)] 1,2-diacyl-sn-glycero-3-phospho-(1'-myo-inositol)	F11	Palmitoleic acid
F2	Pyranomammea B	F11	Arachidonic acid
F2	2,3-Methylenesuccinic acid	F11	11,14,17-Eicosatrienoic acid
F2	N-oleoyl isoleucine	F11	Eicosadienoic acid
F2	7b-Hydroxy-3-oxo-5b-cholanoic acid	F11	Linoleoyl ethanolamide
F2	Uridine 5'-diphosphate	F11	[FA trihydroxy(18:1)] 5,8,12-trihydroxy-9-octadecenoic acid
F2	11'-Carboxy-gamma-chromanol	F11	Adrenic acid
F2	Caprylic acid	F11	Sorbitan laurate
F2	N-Acetylcadaverine	F11	11,12,15-TriHETRE
F2	Methandriol dipropionate	F11	[ST (2:0)] 5beta-Chola-3,11-dien-24-oic Acid
F2	Ala-Lys-Thr-Thr	F11	Trigonelline
F2	Asn-Cys-Cys-Ser	F11	2-Hydroxylauroylcarnitine
F2	3beta,5beta-Ketotriol	F11	12b-Hydroxy-5b-cholanoic acid
F2	[ST (4:0)] (5Z,7E)-(1S,3R,24R)-22-oxa-9,10-seco-5,7,10(19)-cholestatriene-1,3,24,25-tetrol	F11	Kinocoumarin
F2	Varanic acid	F11	Ursodeoxycholic acid
F2	Stigmastentriol	F11	MG(22:2(13Z,16Z)/0:0/0:0)
F2	N-Acetyl-leu-leu-tyr	F11	1-(3-Aminopropyl)-4-aminobutanal
F2	4-Thiocyanatophenol	F11	[ST dimethyl(4:0/3:0)] (5Z,7E,17Z)-(1S,3R)-26,27-dimethyl-9,10-seco-5,7,10(19),17(20)-cholestatetraen-22-yne-1,3,25-triol
F2	1-16:0-lysophosphatidylethanolamine	F11	Lys-Val-Val-Pro
F2	Asp-Pro-Ser-His	F11	Ala-Lys-Gln-Pro
F2	Arg-Val-Pro-Ser	F11	Asn-Leu-Leu-Pro
F2	Pubesenolide	F11	Arachidyl carnitine
F2	(3beta,5alpha,9alpha,22E,24R)-3,5,9-Trihydroxy-23-methylergosta-7,22-dien-6-one	F11	Ile-Leu-Val-Asn
F2	13'-Carboxy-alpha-tocopherol	F11	Hematoporphyrin IX

F2	Ribitol	F11	PG(31:2)
F2	6-Deoxohomodolichosterone	F11	PG(16:1(9Z)/16:1(9Z))
F2	Cholesterol sulfate	F11	PG(33:2)
F2	Leu-Pro-Ser-Tyr	F11	Silicic acid
F2	(22R,23R)-28-homocastasterone	F11	1-18:0-2-18:2-monogalactosyldiacylglycerol
F2	3,5-Dihydroxybenzoic acid	F11	PA(42:2)
F2	Lys-Met-Phe-Gly	F11	1-18:0-2-18:1-monogalactosyldiacylglycerol
F2	LysoPC(15:0)	F11	PC(22:2(13Z,16Z)/15:0)
F2	PA(22:1(11Z)/0:0)	F11	Nonyl acetate
F2	Cys-Leu-Leu-Phe	F11	L-Menthol
F2	Orotic acid	F11	Phenol
F2	L-Lactic acid	F11	n-Dodecane
F2	His-Lys-Lys-Pro	F12	Neuromedin B (1-3)
F2	20-Hydroxy-3,7,11,15,23-pentaoxolanost-8-en-26-oic acid	F12	Glycerol tributanoate
F2	Isovalerylglycine	F12	(-)-Epigallocatechin
F2	Nb-Hexacosanoyltryptamine	F12	Glutathione, reduced
F2	3-cis-Hydroxy-b,e-Caroten-3'-one	F12	[FA (16:0)] 2-bromo-hexadecanal
F2	His-Leu-Trp-Pro	F12	Methyl (9Z)-6'-oxo-6,5'-diapo-6-carotenoate
F2	Isorhamnetin 3,7,4'-tri-O-sulfate	F12	3-Dehydroxycarnitine
F2	demethylmenaquinone-6	F12	Ile-Lys-Cys-Gly
F2	GDP-4-dehydro-3,6-dideoxymannose	F12	His-Lys-Val-Ser
F2	Phenylpyruvic acid	F12	Ile-Trp-Gln-Gln
F2	DG(15:0/0:0/18:1n9)	F12	His-Met-Phe-Phe
F2	DG(15:0/0:0/18:0)	F12	Bilirubin
F2	Cryptoxanthin 5,6:5',8'-diepoxide	F12	Gln-Phe-Phe-Phe
F2	GDP-4-dehydro-6-deoxy-L-mannose	F12	Lys-Phe-Phe-Phe
F2	Arg-Lys-Trp-Val	F12	D-Urobilin
F2	CE(14:1(9Z))	F12	I-Urobilin
F2	Glycerol triundecanoate	F12	D-Urobilinogen
F2	Endomorphin-1	F12	Mesobilirubinogen

F2	CE(16:2)	F12	Arg-Trp-Pro-Tyr
F2	DG(15:0/0:0/22:4n6)	F12	3-O-Protocatechuoylceanothic acid
F2	Pyridoxal	F12	Lyciumoside III
F2	TG(12:0/12:0/12:0)	F12	PA(33:5)
F2	Phosphoric acid	F12	PI(25:0)
F2	Uric acid	F12	delta-Dodecalactone
F2	Homogentisic acid	F12	4-Methyl-2-pentanone
F2	DG(20:2n6/0:0/20:3n6)	F12	Isoamyl isovalerate
F2	DG(39:4)	F12	Allyl caprylate
F2	DG(20:1n9/0:0/18:2n6)	F12	2-methyl-Butanoic acid, hexyl ester
F2	SM(d32:2)	F13	Vitamin D3
F2	CerP(d18:1/20:0)	F13	[ST methyl(2:0)] 24-methylcholesta-5,24-dien-3beta-ol
F2	SM(d18:1/14:0)	F13	5-Dehydroavenasterol
F2	N-(2-hydroxy-eicosanoyl)-tetradecasphing-4-enine-1-phosphoethanolamine	F13	Avenasterol
F2	Dephospho-CoA	F13	Cer(d18:1/14:0)
F2	PC(o-16:1(9Z)/14:1(9Z))	F13	Cer(d33:1)
F2	PC(14:1(9Z)/16:1(9Z))	F13	N-[(4E,8E)-1,3-dihydroxyoctadeca-4,8-dien-2-yl]hexadecanamide
F2	N,2,3-Trimethyl-2-(1-methylethyl)butanamide	F13	sphing-8(E)-enine-16:0, ceramide
F2	PS(O-18:0/15:0)	F13	N-(eicosanoyl)-4E,6E-pentadecasphingadienine
F2	3-nonaprenyl-4,5-dihydroxybenzoate	F13	Ceramide (d18:1/9Z-18:1)
F2	PS(36:4)	F13	[PC (10:2/12:2)] 1-decyl-2-dodecyl-sn-glycero-3-phosphocholine
F2	N-(hexadecanoyl)-4R-hydroxysphinganine-1-phospho-(1'-myo-inositol)	F13	Cer(d18:0/18:1(11Z))
F2	N-(2-hydroxyhexadecanoyl)-sphinganine-1-phospho-(1'-myo-inositol)	F13	N-(2-hydroxyoctadecanoyl)-4,8-sphingadienine
F2	1-20:3-2-18:3-monogalactosyldiacylglycerol	F14	Tocopheronic acid
F2	PS(38:4)	F14	L-trans-4-Methyl-2-pyrrolidinecarboxylic acid
F2	PS(39:6)	F14	5'-Hydroxy-3',4',7-trimethoxyflavan
F2	4-Hydroxycrotonic acid	F14	Hydroxyprolyl-Tryptophan
F2	(5R)-6-Hydroxy-5-isopropenyl-2-methylhexanoate	F14	[ST Trioxo(2:0)] 3,7,12-Trioxochola-1,5-dien-24-oic Acid
F2	Valeric acid	F14	Leu-Ser-Ser-Tyr

F2	PC(24:1(15Z)/P-18:1(9Z))	F14	Ala-Lys-Lys-His
F2	PC(42:3)	F14	His-Met-Val-Val
F2	Allyl cinnamate	F14	Cucurbitacin S
F2	chlorophyll a'	F14	11-Deoxocucurbitacin I
F2	PC(o-22:1(13Z)/22:2(13Z,16Z))	F14	1-18:2-lysophosphatidylglycerol
F2	2-oxindole-3-acetate	F14	Sulfolithocholylglycine
F2	2-Hydroxy-7-methoxy-2H-1,4-benzoxazin-3(4H)-one	F14	Cucurbitacin I
F2	2-(3-Phenylpropyl)pyridine	F14	Arg-Ile-Asn-Ile
F2	PI(41:2)	F14	Glu-Ile-Ile-Phe
F2	TG(20:4(5Z,8Z,11Z,14Z)/18:2(9Z,12Z)/22:6(4Z,7Z,10Z,13Z,16Z,19Z))[iso6]	F14	Arg-Phe-Gly-Arg
F2	3-Methyl-3H-imidazo[4,5-f]quinoxalin-2-amine	F14	Lys-Lys-Val-Tyr
F2	TG(20:5(5Z,8Z,11Z,14Z,17Z)/22:6(4Z,7Z,10Z,13Z,16Z,19Z)/o-18:0)	F14	Arg-Leu-Pro-Arg
F2	cis-4-Decenedioic acid	F14	[ST hydrox] (25R)-3alpha,7alpha-dihydroxy-5beta-cholestan-27-oyl taurine
F2	Dodecanoic acid	F14	Ala-Lys-Phe-Trp
F2	TG(20:5(5Z,8Z,11Z,14Z,17Z)/18:1(9Z)/22:5(7Z,10Z,13Z,16Z,19Z))[iso6]	F14	Triterpenoid
F2	TG(24:1(15Z)/24:0/20:3n6)	F14	Mesoporphyrin IX
F2	Ethyl menthane carboxamide	F14	Protoporphyrinogen IX
F2	Cysteinyl-Cysteine	F14	Lys-Gly
F2	Undecanal	F14	p-Anisaldehyde
F2	Methyl-2-methylbutyrate	F15	2,4,8-Eicosatrienoic acid isobutylamide
F2	3-(Methylthio)-1-propanol	F15	Gallic acid
F2	cis-Jasmone	F15	Vanillin
F2	Methyl isobutyrate	F15	o-Cresol
F2	gamma-Dodecalactone	F15	Dihydrocarvylacetate
F2	Citronellal	F15	beta-Pinene
F2	Menthyl isovalerate	F15	alpha-Pinene
F2	Citronellyl tiglate	F15	Dimethyl benzyl carbinyl butyrate
F2	Hexyl 2-furoate	F15	alpha-phellandrene
F2	Acelan A	F15	alpha-terpinene

F3	alpha,alpha-Dimethylphenethyl alcohol
F3	Heptadecane
F3	Lauryl acetate
F3	Methyl caprate
F3	Octanoic acid, methyl ester
F3	Methyl caproate
F3	[PR] gamma-Carotene/ beta,psi-Carotene
F3	Ganglioside GA2 (d18:1/24:0)
F3	CerP(d18:1/18:0)
F3	PG(20:0/22:1(11Z))
F3	protochlorophyll a
F3	PE(22:4(7Z,10Z,13Z,16Z)/P-16:0)
F3	PG(O-34:3)
F3	PS(P-16:0/15:0)
F3	PE(18:3(6Z,9Z,12Z)/P-18:0)
F3	PG(34:0)
F3	PC(o-22:1(13Z)/20:4(8Z,11Z,14Z,17Z))
F3	alpha-Tocopherol succinate
F3	PG(P-28:0)
F3	Cer(d31:0)
F3	Staphyloxanthin
F3	Glutamyl-Arginine
F3	MG(20:3(8Z,11Z,14Z)/0:0/0:0)
F3	TG(14:0/14:1(9Z)/18:4(6Z,9Z,12Z,15Z))
F3	PG(P-18:0/15:1(9Z))
F3	PG(P-32:1)
F3	[FA oxo(19:0)] 18-oxo-nonadecanoic acid
F3	PG(31:0)
F3	PG(O-30:1)
F3	PG(33:1)
F15	p-Vinylguaiacol
F15	Germacrene D
F15	Fenchene
F15	Benzyl acetoacetate
F15	Acoradien
F15	Nezukon
F15	Neryl propionate
F15	Junipen
F15	Isolimonene
F15	Neoclovene
F15	Camphene
F15	Humulene
F15	.alpha.-Himachalene
F15	.beta.-Selinene
F15	Santolina alcohol
F16	3,3'-Dithiobis[2-methylfuran]
F16	L-Glutamic acid 5-phosphate
F16	Dopamine 3-O-sulfate
F16	FA oxo(14:2) cyclopentanebutanoic acid
F16	[FA hydroxy(14:1)] 12-hydroxy-13-tetradecenoic acid
F16	Pentadecanoic acid
F16	[FA hydroxy(14:0)] 3,11-dihydroxy-tetradecanoic acid
F16	Phenol glucuronide
F16	Heptadecanoic acid
F16	[FA hydroxy(16:0)] 4,12-dihydroxy-hexadecanoic acid
F16	[FA (18:3)] 12,13S-epoxy-9Z,11,15Z-octadecatrienoic acid
F16	Arachidic acid
F16	Cortisone
F16	C19 Sphingosine-1-phosphate
F16	Secoeremopetasitolide B

F3	PG(37:2)	F16	Solanidine
F3	Trp-Val-Val	F16	Leu-Leu-Gly-Pro
F3	N-[2-(1H-Indol-3-yl)ethyl]docosanamide	F16	3-Epidemissidine
F3	1,3-di-(5Z,8Z,11Z,14Z,17Z-eicosapentaenoyl)-2-hydroxy-glycerol (d5)	F16	Ile-Arg-Ile
F3	PG(O-33:1)	F16	3'-Geranyl-2',4,4',6'-tetrahydroxychalcone
F3	CerP(d18:1/16:0)	F16	(R)-Kanzonol Y
F3	PG(O-16:0/17:2(9Z,12Z))	F16	4alpha-Formyl-5alpha-cholesta-8-en-3beta-ol
F3	PG(P-32:0)	F16	Tomatidine
F3	PC(34:2)	F16	Leu-Pro-Pro-Pro
F3	Cer(d18:0/14:0)	F16	DHAP(18:0e)
F3	[FA oxo,hydroxy(2:0)] 9-oxo-11R,15S-dihydroxy-1a,1b-dihomo-13E-prostaenoic acid	F16	LPA(18:0e/0:0)
F3	Nb-Palmitoyltryptamine	F16	Leupeptin
F3	Ala-Trp-Pro-Tyr	F16	(3beta,5alpha,6alpha,7alpha,22E,24R)-5,6-Epoxyergosta-8,14,22-triene-3,7-diol
F3	Cer(d33:0)	F16	3,5-Dihydroxyergosta-7,22-dien-6-one
F3	5,6-Dihydroxyprostaglandin F1a	F16	His-Ile-Gly-Ile
F3	3,17-Androstanediol glucuronide	F16	Momordol
F3	2-Ethylsuberic acid	F16	Ascorbyl stearate
F3	PG(P-16:0/14:0)	F16	menaquinone-4
F3	MG(22:1(13Z)/0:0/0:0)	F16	27-Norcholestanehexol
F3	Violaxanthin	F16	Asn-Lys-Val-Pro
F3	PG(O-31:1)	F16	Quasiprotopanaxatriol
F3	3-Feruloyl-1,5-quinolactone	F16	Ala-Lys-Lys-Asn
F3	MG(24:1(15Z)/0:0/0:0)	F16	Lys-Lys-Val-Ser
F3	(3b,4b,11b,14b)-11-Ethoxy-3,4-epoxy-14-hydroxy-12-cyathen-15-al 14-xyloside	F16	His-Leu-Val-Val
F3	Azelaic acid	F16	LysoSM(d18:0)
F3	Glutarylglycine	F16	Lys-Lys-Pro-Pro
F3	Caproic acid	F16	N-Methylindolo[3,2-b]-5alpha-cholest-2-ene
F3	Nb-Stearoyltryptamine	F16	Homodolichosterone
F3	Pimelic acid	F16	Brassinolide

F4	O-Phosphoethanolamine	F16	[PR] 17beta,21beta-epoxy-16alpha-ethoxyhopan-3beta-ol
F4	PE(P-16:0e/0:0)	F16	Rubraflavone D
F4	LysoPE(15:0/0:0)	F16	Ananasic acid
F4	Sodium glycocholate	F16	[ST hydroxy(2:0)] (5Z,7E)-(1R,2R,3R)-2-(4-hydroxybutoxy)-9,10-seco-5,7,10(19)-cholestatriene-1,3-diol
F4	Minabeolide-2	F16	Leu-Val-Val-Tyr
F4	LysoPE(18:2(9Z,12Z)/0:0)	F16	26-Hydroxy-24-epi-brassinolide
F4	LysoPE(18:1(9Z)/0:0)	F16	[PR] Neurosporaxanthin
F4	LysoPC(17:0)	F16	Cyclopassifloic acid D
F4	LysoPC(18:1(9Z))	F16	Ganoderiol G
F4	LysoPC(0:0/18:0)	F16	Theasapogenol A
F4	Ala-Leu-Phe-Trp	F16	Leu-Leu-Thr-Tyr
F4	DG(20:2n6/0:0/18:4n3)	F16	His-Lys-Lys-Val
F4	DG(18:3n6/0:0/20:3n6)	F16	Nb-Lignoceroyltryptamine
F4	DG(20:2n6/0:0/20:4n3)	F16	Arg-Leu-Lys-Val
F4	ubiquinone-7	F16	[ST hydroxy(6:2/3:0)] (5Z,7E)-(1S,2R,3R)-2-(6-hydroxyhexyl)-9,10-seco-5,7,10(19)-cholestatriene-1,3,25-triol
F4	PE(14:0/16:0)	F16	13-beta-D-Glucosyloxydocosanoate
F4	PC(16:1(9Z)/14:0)	F16	2-Hexaprenyl-6-hydroxyphenol
F4	cyclopropane phosphatidylethanolamine (dihexadec-9,10-cyclo-anoyl, n-C16:0 cyclo)	F16	PC(O-19:0)
F4	PC(31:1)	F16	Arg-Ile-Gln-Ile
F4	PC(33:4)	F16	Torulene
F4	PS(34:6)	F16	LysoPE(22:0/0:0)
F4	PS(35:6)	F16	Arg-Lys-Val-His
F4	PS(36:6)	F16	Arg-Leu-Val-Arg
F5	Pro-Thr-Val-Val	F16	DG(31:3)
F5	Arg-Leu-Val-Cys	F16	DG(31:2)
F5	LysoPC(18:2(9Z,12Z))	F16	Arg-Lys-Lys-Gln
F5	Arg-Val-Pro-Tyr	F16	LysoPE(24:0/0:0)
F5	Arg-Ile-Ile-Tyr	F16	1,1'-(1,4-Dihydro-4-nonyl-3,5-pyridinediyl)bis[1-dodecanone]
F5	Ile-Leu-Phe-Trp	F16	7,8-Dihydroparasiloxanthin

F5	Arg-Pro-Arg-Arg	F16	PA(P-28:0)
F5	Nummularine A	F16	DG(15:0/0:0/18:3n6)
F5	PS(26:0))	F16	Vitamin K2
F5	Lycoperoside D	F16	5-Oxoavermectin "2b" aglycone
F5	PS(36:5)	F16	[PR] bacteriohopane-32,33, 34-triol-35-carbamate
F5	Solamargine	F16	[SP (16:1)] N-(3E-hexadecenoyl)-deoxysphing-4-enine-1-sulfonate
F6	(-)-(E)-1-(4-Hydroxyphenyl)-7-phenyl-6-hepten-3-ol	F16	PE(P-28:0)
F6	xi-7-Hydroxyhexadecanedioic acid	F16	PE(14:1(9Z)/14:0)
F6	[FA (22:5)] 4,7,10,13,16-Docosapentaynoic acid	F16	PE(29:2)
F6	2,3-Epoxymenaquinone	F16	PE(14:1(9Z)/15:0)
F6	2-Hydroxy-4-oxopentanoic acid	F16	PC(14:1(9Z)/14:1(9Z))
F6	4-Deoxycohumulone	F16	1-18:1-2-18:2-phosphatidate
F6	9,10,13-Trihydroxystearic acid	F16	PE(18:3(6Z,9Z,12Z)/15:0)
F6	3,6-Dimethoxy-19-norpregna-1,3,5,7,9-pentaen-20-one	F16	DG(22:2n6/0:0/22:4n6)
F6	Thiophanate-methyl	F16	N-Acetyl-L-glutamate 5-semialdehyde
F6	Trp-Ala-Ser	F16	PC(18:1(9Z)/15:0)
F6	17beta-(Acetylthio)estra-1,3,5(10)-trien-3-ol acetate	F16	S-Carboxymethyl-L-cysteine
F6	L-Histidinol	F16	PC(20:5(5Z,8Z,11Z,14Z,17Z)/P-18:1(11Z))
F6	(S)-(E)-8-(3,6-Dimethyl-2-heptenyl)-4',5,7-trihydroxyflavanone	F16	PC(40:6)
F6	Ala-Lys-Phe-Gly	F16	TG(18:3(9Z,12Z,15Z)/18:3(9Z,12Z,15Z)/18:4(6Z,9Z,12Z,15Z))
F6	Riboflavin cyclic-4',5'-phosphate	F16	PE(22:5(7Z,10Z,13Z,16Z,19Z)/24:0)
F6	2-Phenylpropionate	F16	4-Hydroxybutyric acid
F6	Leu-Trp-Gly-Pro	F16	(Z)-[3-(Methylsulfinyl)-1-propenyl] 2-propenyl disulfide
F6	Capillene	F16	L-prolyl-L-proline
F6	Cys-Cys-Pro-Tyr	F16	[FA (14:2)] 5Z,8Z-tetradecadienoic acid
F6	Ala-Leu-Trp-Pro	F16	Hexenal
F6	PA(20:4(5Z,8Z,11Z,14Z)e/2:0)	F16	Geranylaceton(isomer 2)
F6	Asp-Phe-Phe-Val	F16	Butyl butyrate
F6	Arg-Lys-Phe-Val	F16	Octanal

F6	Ile-Lys-Lys-Tyr	F16	n-Decane
F6	L-Carnitine	F17	Chorismate
F6	L-Urobilin	F17	2-Hydroxy-3-carboxy-6-oxo-7-methylocta-2,4-dienoate
F6	L-Urobilinogen	F17	N'-Formylkynurenine
F6	(24E)-3beta,15alpha,22S-Triacetoxylanosta-7,9(11),24-trien-26-oic acid	F17	Methyl 2,3-dihydro-3,5-dihydroxy-2-oxo-3-indoleacetic acid
F6	Arg-Phe-Phe-Arg	F17	2-(Ethylsulfinylmethyl)phenyl methylcarbamate
F6	Lys-Phe-Phe-Trp	F17	Genipinic acid
F6	2,3,16-triacetate-15-oxo-23,24-dihydroxycucurbitacin	F17	1-O-Galloylglycerol
F6	[PS (12:0/13:0)] 1-dodecanoyl-2-tridecanoyl-sn-glycero-3-phosphoserine (ammonium salt)	F17	Biotin
F6	Momorcharaside B	F17	1-(2,4,6-Trimethoxyphenyl)-1,3-butanedione
F6	Lys-Trp-Tyr-Tyr	F17	Histidinyl-Cysteine
F6	(3a,5b)-24-oxo-24-[(2-sulfoethyl)amino]cholan-3-yl-b-D-Glucopyranosiduronic acid	F17	L-Homocystine
F6	PA(O-36:4)	F17	Anigorufone
F6	1,2-Di-(9Z,12Z-octadecadienoyl)-sn-glycero-3-phosphate	F17	4-Phenyl-1H,3H-naphtho[1,8-cd]pyran-1,3-dione
F6	(S)-N-(4,5-Dihydro-1-methyl-4-oxo-1H-imidazol-2-yl)alanine	F17	2-O-Benzoyl-D-glucose
F6	2-(1-Naphthyl)acetamide	F17	Argininosuccinic acid
F6	N-Heptanoylglycine	F17	2-Hydroxy-4-(4-methoxyphenyl)-1H-phenalen-1-one
F6	3-O-(Rhaa1-2Glcb)-26-O-(Glcb)-(25R)-furosta-5,20(22)-dien-3beta,26-diol	F17	Pro-Pro-Pro
F6	1-18:3-2-16:3-digalactosyldiacylglycerol	F17	Hydroxytyrosol 1-O-glucoside
F6	PI(41:7)	F17	2-O-p-Coumaroyl-D-glucose
F6	Nonanoic acid	F17	2-Deoxy-L-ribono-1,4-lactone
F6	omega-6-Hexadecalactone, Ambrettolide	F17	Glucocaffeic acid
F6	5-Methyl-2,3-hexanedione	F17	Glycolic acid
F6	Ethylphenyl acetate	F17	3-O-Caffeoyl-1-O-methylquinic acid
F6	Ethyl methylphenylglycidate (isomer 1)	F17	Leu-Leu-Met
F6	Benzyl isobutyrate	F17	20-hydroxy-E4-neuroprostane
F6	n-Hexadecane	F17	Clausarinol
F6	Geranyl isovalerate	F17	D-galactal

F7	Ile-Pro	F17	Dopamine
F7	Leu-Val	F17	3-Methylcrotonylglycine
F7	Asparaginyl-Valine	F17	2-Naphthalenethiol
F7	Threoninyl-Isoleucine	F17	(-)-1-Methylpropyl 1-propenyl disulfide
F7	hexahomomethionine	F17	Biliverdin
F7	3-Oxopentanoic acid	F17	5-Dehydroshikimate
F7	Isoleucyl-Isoleucine	F17	Isovalerylsarcosine
F7	Val-Ala-Gly	F17	Shikimic acid
F7	Valyl-Lysine	F17	Suberic acid
F7	Val-Ala-Ala	F17	L-Arginine
F7	L-Homoserine	F17	Ascorbic acid
F7	Hexadecasphinganine	F17	3-Isopropylmalate
F7	Taurine	F17	Hexoses
F7	3-Buten-1-amine	F17	1,1'-Thiobis-1-propanethiol
F7	cinchoninone	F17	1-(Malonylamino)cyclopropanecarboxylic acid
F7	Sphingosine	F17	[FA hydroxy(4:0)] N-(3S-hydroxy-butanoyl)-homoserine lactone
F7	1,2,3-Tris(1-ethoxyethoxy)propane	F17	Furfuryl propyl disulfide
F7	2-Methyl-3-ketovaleric acid	F17	Leucyl-Glycine
F7	(2E,5E,12Z,15Z)-1-Hydroxy-2,5,12,15-heneicosatetraen-4-one	F17	5-(Methylthio)-2-[(methylthio)methyl]-2-pentenal
F7	Creatine	F17	2-(Methylthiomethyl)-3-phenyl-2-propenal
F7	Dopamine glucuronide	F17	Malonic acid
F7	Ornithine	F17	3,4,6-trihydroxy-cis-cinnamate
F7	Glycine	F17	4-Methyldibenzothiophene
F7	Heptane-1-thiol	F17	3,6-Dihydro-4-(4-methyl-3-pentenyl)-1,2-dithiin
F7	Lys-Phe-Gly	F17	cis-(homo)2aconitate
F7	Urocanic acid	F17	(R)-lipoic acid
F7	Leu-Pro-Tyr	F17	N-Acetyl-L-phenylalanine
F7	3-Hydroxy-10'-apo-b,y-carotenal	F17	Vanilpyruvic acid
F7	Piperidine	F17	3,4-Dihydroxyphenyllactic acid methyl ester

F7	beta-Carotinal	F17	Triacetin

Supplementary Table 17: Dietary components differentially abundant between the IBS and Control groups. Significantly different components are shown. For complete table please refer to the published manuscript (doi: 10.1053/j.gastro.2019.11.301).

Dietary component	IBS (IQR)	Control (IQR)	Wilcoxon Statistic	p-value	q-value
Leeks	0.07 (0 - 0.233)	0.233 (0.07 - 0.617)	3366	0.000095	0.00281
Onions	2 (0.07 - 2)	2 (2 - 5)	3390.5	0	0.00281
Cucumber, celery	0.233 (0 - 1.75)	2 (0.233 - 2)	3390	0	0.00281
Other salad dressing (tablespoon)	0 (0 - 0.07)	0.07 (0 - 1)	3297	0.000192	0.0031
Blueberries	0.07 (0 - 1)	1 (0.07 - 2)	3332	0.000191	0.0031
Tomatoes	1 (0 - 2)	2 (2 - 5)	3330.5	0.00021	0.0031
Cornflakes, Rice Krispies	0.07 (0 - 2)	0 (0 - 0.07)	1697	0.000606	0.00768
Wine (glass)	0.07 (0 - 2)	1 (0.07 - 2)	3242.5	0.000921	0.0102
Garlic	0.233 (0 - 2)	2 (1 - 5)	3229.5	0.00113	0.0111
Spinach	0.07 (0 - 1)	1 (0.07 - 2)	3215	0.00135	0.0119
Pickles, chutney (tablespoon)	0 (0 - 0.07)	0.233 (0 - 1)	3141	0.00209	0.016
Coffee ground (cup)	0 (0 - 5)	5 (0 - 14)	3168.5	0.00217	0.016
Dried lentils, beans, peas	0 (0 - 0.233)	0.233 (0.035 - 1)	3158	0.00244	0.0166
Low calorie or diet soft fizzy drink (glass)	0.07 (0 - 2)	0 (0 - 0.07)	1818.5	0.00437	0.0277
Cream (tablespoon)	0.07 (0 - 0.233)	0.07 (0.07 - 1)	3089.5	0.00652	0.034
French dressing (tablespoon)	0 (0 - 0.07)	0.07 (0 - 1)	3057.5	0.00646	0.034
Beetroot	0.07 (0 - 0.233)	0.233 (0.07 - 1.5)	3096.5	0.00637	0.034
Brussel sprouts	0.07 (0 - 0.07)	0.07 (0 - 0.233)	3061	0.0081	0.0378
Courgettes	0.07 (0 - 0.233)	0.233 (0.07 - 1)	3080.5	0.00772	0.0378

Median consumption per week

Dietary analysis performed on 141 samples (IBS: n = 78 and Control: n = 63)

Median consumption per week represented as inter-quartile range (IQR)

Supplementary Table 18: Medication usage and its effect on beta-diversity.

	Association Analysis IBS versus Healthy Chi-squared statistic					Univariate Analysis IBS versus Healthy Permutational MANOVA (R library vegan)				
	Chi^2 Statistic	DF	IBS N value	Healthy N value	P-value	Df	Sum Of Sqs	R-squared	F	P-value
PPI	6.32	1	20	5	0.012	1	0.087	0.015	2.16	0.004
Laxative	9.66	1	13	0	0.0019	1	0.069	0.012	1.72	0.015
5 HT agonist	5.95	1	9	0	0.015	1	0.068	0.012	1.68	0.019
SSRI	1.69	1	8	2	0.19	1	0.068	0.012	1.7	0.024
Probiotic	8.17	1	14	1	0.0043	1	0.058	0.01	1.43	0.048
Angiotensin Antagonist	0.00	1	3	2	1	1	0.056	0.01	1.39	0.058
Beta-blocker	1.72	1	4	0	0.19	1	0.051	0.009	1.27	0.14
Calcium-blocker	0.97	1	3	0	0.32	1	0.058	0.01	1.44	0.057
5 HT inhibitor	5.07	1	8	0	0.024	1	0.057	0.01	1.41	0.069
COX-2 inhibitor	0.75	1	2	2	1	1	0.048	0.009	1.19	0.15
Antidopaminergic	0.00	1	3	0	0.32	1	0.051	0.009	1.26	0.13
Diabetes Medication	0.00	1	2	1	1	1	0.044	0.008	1.08	0.28
H2 blocker	4.20	1	7	0	0.04	1	0.05	0.009	1.23	0.14
Antispasm	4.20	1	7	0	0.04	1	0.054	0.01	1.33	0.076
NSAID	0.80	1	8	3	0.37	1	0.052	0.009	1.28	0.11
Thyroid hormone	0.11	1	4	5	0.74	1	0.051	0.009	1.25	0.13
Antidepressant	6.86	1	10	0	0.01	1	0.05	0.009	1.24	0.13
Progestogen	1.61	1	6	1	0.2	1	0.044	0.008	1.08	0.29
Antihistamine	0.09	1	3	1	0.77	1	0.048	0.008	1.18	0.19
AntiSeizure	0.00	1	2	1	1	1	0.045	0.008	1.12	0.26
Oestro-progestogen	0.59	1	9	4	0.44	1	0.046	0.008	1.13	0.19
Progestogen subdermal implant	0.75	1	0	2	0.39	1	0.046	0.008	1.14	0.2
Benzodiazepine	0.32	1	2	0	0.57	1	0.045	0.008	1.1	0.26
beta3 Antagonist	0.32	1	2	0	0.57	1	0.038	0.007	0.93	0.53

Ace Inhibitor	0.53	1	1	3	0.47	Ace Inhibitor	1	0.041	0.007	1.02	0.37
Opiod	1.69	1	8	2	0.19	Opiod	1	0.04	0.007	0.98	0.43
NonOpiod	0.41	1	7	3	0.52	NonOpiod	1	0.037	0.007	0.92	0.55
Reflux Supplement	0.32	1	2	0	0.57	Reflux Supplement	1	0.038	0.007	0.93	0.54
Antidiarrhea	0.32	1	2	0	0.57	Antidiarrhea	1	0.035	0.006	0.87	0.66
Corticosteroid	1.61	1	6	1	0.2	Corticosteroid	1	0.033	0.006	0.81	0.8
Selective b2 agonist	0.02	1	4	2	0.88	Selective b2 agonist	1	0.033	0.006	0.81	0.79
Anticholinergic	3.35	1	6	0	0.07	Anticholinergic	1	0.033	0.006	0.81	0.85
Leukotriene Antagonist	0.32	1	2	0	0.57	Leukotriene Antagonist	1	0.033	0.006	0.82	0.79
Statins	1.68	1	4	8	0.19	Statins	1	0.033	0.006	0.82	0.81
Heart Rate Lowering	0.97	1	3	0	0.32	Heart Rate Lowering	1	0.034	0.006	0.84	0.74
Rheumatology	0.06	1	2	3	0.81	Rheumatology	1	0.032	0.006	0.78	0.86
H1 Antagonist	0.00	1	2	1	1	H1 Antagonist	1	0.025	0.004	0.61	1
PolyPharmacy (3+ medications)	7.98	1	22	5	0.0047	PolyPharmacy (3+ medications)	1	0.08	0.014	1.98	0.008

Multivariate Analysis
IBS versus Healthy
Permutational MANOVA (R library vegan)

	Df	Sum Of Sqs	R-squared	F	P-value	
PPI	1	0.087	0.015	2.22	0.002	**
Laxative	1	0.057	0.01	1.45	0.053	.
5 HT agonist	1	0.043	0.008	1.09	0.3	
SSRI	1	0.046	0.008	1.16	0.21	
Probiotic	1	0.047	0.008	1.21	0.16	
Angiotensin Antagonist	1	0.064	0.011	1.62	0.026	*
Beta-blocker	1	0.038	0.007	0.95	0.51	
Calcium-blocker	1	0.048	0.008	1.21	0.17	
5 HT inhibitor	1	0.054	0.009	1.37	0.057	.

COX-2 inhibitor	1	0.047	0.008	1.2	0.17	
Antidopaminergic	1	0.047	0.008	1.19	0.18	
Diabetes Medication	1	0.05	0.009	1.27	0.11	
H2 blocker	1	0.037	0.007	0.95	0.49	
Antispasm	1	0.049	0.009	1.25	0.11	
NSAID	1	0.042	0.007	1.07	0.3	
Thyroid hormone	1	0.055	0.01	1.39	0.07	.
Antidepressant	1	0.032	0.006	0.83	0.77	
Progestogen	1	0.039	0.007	0.99	0.41	
Antihistamine	1	0.05	0.009	1.27	0.12	
AntiSeizure	1	0.038	0.007	0.97	0.46	
Oestro-progestogen	1	0.053	0.009	1.34	0.08	.
Progestogen subdermal implant	1	0.042	0.007	1.08	0.28	
Benzodiazepine	1	0.034	0.006	0.88	0.68	
beta3 Antagonist	1	0.033	0.006	0.85	0.7	
Ace Inhibitor	1	0.038	0.007	0.95	0.5	
Opiod	1	0.047	0.008	1.19	0.18	
NonOpiod	1	0.042	0.007	1.07	0.28	
Reflux Supplement	1	0.029	0.005	0.73	0.92	
Antidiarrhea	1	0.041	0.007	1.04	0.35	
Corticosteroid	1	0.022	0.004	0.55	1	
Selective b2 agonist	1	0.041	0.007	1.05	0.33	
Anticholinergic	1	0.039	0.007	0.99	0.43	
Leukotriene Antagonist	1	0.037	0.007	0.95	0.5	
Statins	1	0.037	0.006	0.94	0.55	
Heart Rate Lowering	1	0.033	0.006	0.85	0.72	
Rheumatology	1	0.033	0.006	0.84	0.72	
H1 Antagonist	1	0.035	0.006	0.89	0.62	
PolyPharmacy (3+ medications)	1	0.033	0.006	0.83	0.76	
IBS Versus Healthy	**1**	**0.072**	**0.013**	**1.83**	**0.01**	**

Residual	101	3.971	0.699

Analysis of medication was performed on 141 samples (IBS: n = 78 and Control: n = 63)
Medication usage was represented as a Binary matrix
Spearman dissimilarity was used in the Permutational MANOVA (adonis function; Vegan R Library)

Chapter 3

The faecal mycobiome in patients with Irritable Bowel Syndrome.

This chapter has been published as

Anubhav Das, Eileen O'Herlihy, Fergus Shanahan, Paul W. O'Toole, Ian B. Jeffery*. The fecal mycobiome in patients with Irritable Bowel Syndrome. Sci Rep. 2021;11:124. DOI: https://doi.org/10.1038/s41598-020-79478-6

*Corresponding author

Author contributions: IBJ., EOH., FS and PWOT. conceived the study idea and design. AD, IBJ, EOH, FS and PWOT drafted and reviewed the manuscript. AD and IBJ, performed the bioinformatics analysis.

AD Contribution:

Data processing: ITS-1 read processing and quality control checks. Clustering of ITS-1 reads.

Analysis: Hypothesis testing, multivariate and all statistical analyses. Co-inertia analysis. Development of Machine learning models and subsequent analysis.

Others: Contribution to manuscript by drafting Introduction, Methods, Results, and Discussion sections.

Table of Contents

3.1 Abstract 207

3.2 Introduction 207

3.3 Methods 209

3.3.1 Study Participants 209

3.3.2 Internal transcribed spacer (ITS) sequencing for mycobiome analysis 211

3.3.3 Fungal zOTU clustering and data processing 211

3.3.4 Statistical analysis 213

3.3.5 Machine learning analysis 213

3.3.6 Co-inertia analysis 214

3.4 Results 214

3.4.1 The faecal mycobiome in IBS and controls 214

3.4.2 Predictive value of the mycobiome for IBS 218

3.4.3 Co-variance of the mycobiome with faecal metabolome and bacterial metagenome 220

3.4.4 Gut mycobiome is not associated with Bile Acid Malabsorption 222

3.5 Discussion 223

3.6 Acknowledgements 225

3.7 References 225

3.8 Supplementary information 230

3.8.1 Supplementary Figures 230

3.8.2 Supplementary Tables 235

3.1 Abstract

Alterations of the gut microbiota have been reported in various gastrointestinal disorders, but knowledge of the mycobiome is limited. We investigated the gut mycobiome of 80 patients with Irritable Bowel Syndrome (IBS) in comparison with 64 control subjects. The fungal-specific internal transcribed spacer 1 (ITS-1) amplicon was sequenced, and mycobiome zero-radius operational taxonomic units (zOTUs) were defined representing known and unknown species and strains. The fungal community was sparse and individual-specific in all (both IBS and control) subjects. Although beta-diversity differed significantly between IBS and controls, no difference was found among clinical subtypes of IBS or in comparison with the mycobiome of subjects with bile acid malabsorption (BAM), a condition which may overlap with IBS with diarrhoea.

The mycobiome alterations co-varied significantly with the bacteriome and metabolome but were not linked with dietary habits. As a putative biomarker of IBS, the predictive power of the faecal mycobiome in machine learning models was significantly better than random but insufficient for clinical diagnosis. The mycobiome presents limited therapeutic and diagnostic potential for IBS, despite co-variation with bacterial components which do offer such potential.

3.2 Introduction

The gut microbiota has been implicated in various gastrointestinal and extra-intestinal diseases, including Irritable Bowel Syndrome (IBS) [1–6]. Most studies have focused on the prokaryotes in the gut (Archaea, and Bacteria), but little is known about the role of eukaryotes (mainly fungi) in health and disease.

Because of its clinical heterogeneity and the unclear aetiology of IBS, robust biomarkers and therapeutic targets for IBS are difficult to identify[7]. Multiple possible causes or triggers including bacterial dysbiosis, altered activity of the gut-brain axis, and environmental factors like food, may contribute to the pathophysiology of IBS[5,8,9]. One such factor which is understudied, and which has recently attracted attention, is the gut mycobiome (the fungal community of the gut). In recent studies, an altered gut mycobiota in IBS has been reported[7,10–15], but this has been based on small study populations and disparate analytical methods.

Furthermore, targeting the mycobiome for treatment of IBS specific symptoms bears certain caveats. For example, *Candida* associated infectious diarrhoea due to overgrowth of *Candida* also presents IBS-like symptoms[16], and some studies showed absence of any conclusive link between *Candida* overgrowth and IBS[17]. Also, *Candida* abundance is directly associated with short term high carbohydrate consumption[7]. Such mycosis may lead to a false diagnosis of IBS, and its incorrect treatment[18]. A recent study showed that isolates of *Candida albicans* are genotypically, and phenotypically different between IBS and Controls [19]. Such genomic and phenotypic variations in patients were also reported in a separate study, especially for the gene responsible for fungal toxin candidalysin encoding [20]. Therefore, these differences are also required to be considered before anti-fungal treatments in IBS. Symptoms like visceral hypersensitivity are observed only in a subset of subjects with IBS and this has been associated with the gut mycobiome[10]. Therefore, treatment with antifungal agents to ameliorate such symptoms may benefit only a subset of IBS sufferers and may not universally cure IBS. Thus, it is imperative to thoroughly investigate and establish the gut mycobiome as an potential target for IBS therapy, since there is very little evidence of role of gut mycobiome in IBS[16].

We analysed the faecal mycobiome composition by sequencing the ITS1 region amplified from the faecal DNA of patients with IBS and from control subjects. We also assessed the association of the mycobiome with our previously published bacteriome and metabolome datasets[6]. The results suggest that despite a significant difference in the mycobiome between IBS and controls, it is a weak predictive biomarker for IBS.

3.3 Methods

3.3.1 Study Participants

Eighty patients with irritable bowel syndrome (Rome IV Criteria), along with sixty-five control participants were recruited to this study. Both populations (IBS, and controls) belonged to same age range, ethnicity, and geographic region. The IBS population was further classified into clinical subtypes, viz., IBS-C (constipation predominant), IBS-D (diarrhoea predominant), and IBS-M (mixed). A subset of the IBS population (n=45) who underwent a selenium-75 homocholic acid taurine (SeHCAT) assay for bile acid malabsorption detection were also analysed. A summary of the study population characteristics is given in Table 1. The study was approved by the Cork Research Ethics Committee, all research was performed in accordance with relevant guidelines/regulations, and all participants gave written informed consent to take part. Further details of participant recruitment, the metagenomic, faecal metabolomic, and dietary data used in this analysis is available in Jeffery et al 2020[6].

Table 1: Descriptive statistics of control and IBS subjects studied

	Control (n = 65)	IBS (n = 80)
Age range, years (mean)	19-65 (45)	17-66 (39)
Sex (male/female)	16/49	15/65
BMI Class, n (%)		
Normal	25 (38)	31 (39)
Obese Class I	11 (17)	14 (18)
Obese Class II	3 (5)	5 (6)
Obese Class III	1 (2)	3 (4)
Overweight	21 (32)	22 (28)
Underweight	3 (5)	3 (4)
HADS: Anxiety, n (%)		
Normal (0-10)	59 (91)	58 (72)
Abnormal (11-21)	6 (9)	22 (28)
HADS: Depression, n (%)		
Normal (0-10)	64 (98)	70 (88)
Abnormal (11-21)	1 (2)	10 (12)
Bristol Stool Score, n (%)		
Normal	54 (83)	18 (22)
Constipated	8 (12)	22 (28)
Diarrhoea	3 (5)	40 (50)
IBS subtype, n (%)		
IBS-C		30 (38)
IBS-D	N/A	21 (26)
IBS-M		29 (36)
SeHCAT assayed, n (%)	9 (14)	46 (57)
Dietary group (FFQ), n (%)		
Omnivore	63 (97)	74 (92)
Vegetarian	1 (2)	2 (2)
Pescatarian	1 (2)	1 (1)
Gluten-free	0 (0)	4 (5)
Drinks alcohol, n (%)		
Current	54 (83)	57 (71)
Previous	0 (0)	1 (1)
Never	10 (15)	22 (28)
Smoker, n (%)		
Current	10* (15)	14* (18)
Previous	13 (20)	18 (22)
Never	42 (65)	48 (60)

* 1 subject in each group smoked e-cigarettes
BMI: Body Mass Index, FFQ: Food Frequency Questionnaire, HADS: Hospital Anxiety and Depression Scale, N/A: not applicable.

3.3.2 Internal transcribed spacer (ITS) sequencing for mycobiome analysis

DNA was extracted from frozen faecal samples as described previously by Jeffery et al[6] and prepared for ITS-1 sequencing for mycobiome analysis following the method by Huseyin et al.[21]. Briefly, DNA was extracted from 0.25g of frozen faecal samples using the Qiagen DNeasy Blood and Tissue Kit according to manufacturer's instructions. PCR was performed on 100ng of DNA faecal extracts using Kapa Robust HotStart Readymix, ITS1 (Forward: 5' - *TCGTCGGCAGCGTCAGATGTGTATAAGAGACAG***CTTGGTCATTTAGAGGA AGTAA** - 3'), and ITS2 (Reverse: 5' – *GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAG***GCTGCGTTCTTCATCG ATGC** – 3') primers containing an Illumina adaptor sequence for indexing. Two technical replicates were pooled and then purified with AMPure SPRIselect. Forward and reverse barcodes were attached by a second round of adapter PCR using the 16S Sequencing Library Preparation Nextera protocol developed by Illumina (San Diego, California, USA). The resulting tagged samples were purified with AMPure SPRIselect, quantified using Qubit High Sensitivity dsDNA kit (Life Technologies) and equimolar amounts of each sample were then pooled before being sent for sequencing GATC Biotech AG, Konstanz, Germany on the MiSeq (2×250 bp) chemistry platforms.

3.3.3 Fungal zOTU clustering and data processing

zOTUs (zero-radius Operational Taxonomic Units) were identified and profiles were generated using Usearch V11 pipeline codes. Firstly, raw reads were merged using default parameters, followed by primer sequence trimming and sequence orientation using in house python code and primer sequence information. Merged sequences,

which were at least 100 base pairs long after primer trimming were retained and pooled into a single file for zOTU identification using UNOISE algorithm [22].

For zOTU identification, sequences were filtered based on base quality, overall sequence quality, and length, followed by dereplication to obtain unique sequences. These unique sequences were passed through the UNOISE algorithm, and zOTUs were obtained.

The pooled sequences prior to quality filtering were re-mapped to these zOTUs, and the abundance profile was generated.

An additional zOTU filtering step was performed owing to presence of zOTUs of bacterial origin. For this, bacterial genomes from Chocophlan database [23] were pooled and sequence alignment of zOTUs were performed against it. zOTUs which mapped to bacterial genomes with high confidence and identity were rejected and a profile with 1631 zOTUs and their abundances in samples were obtained. The profile was normalized and filtered based on occurrence. zOTUs which were present in >10% of samples were selected. The final profile with 138 zOTUs and 144 samples (IBS: n= 80, controls: n= 64) was transformed using logit function, or log10 transformation for subsequent analyses.

Logit transformation function:

$f(x) = \ln\left(\frac{x}{1-x}\right)$ where, x is the normalized abundance of zOTU

Log10 transformation:

$f(x) = log10(x + 1E - 6) + 6$

where, x is the scalar normalized abundance of zOTU and 1E -6 is used as a minimal non-zero and the addition of six ensure that all values are positive. Taxonomic

classification of the representative sequences of zOTUs was done using UNITE database [24] of fungal sequences, and qiime2's [25] feature-classifier, with default parameters.

3.3.4 Statistical analysis

All statistical analyses were performed in R (v 3.5.1). Beta-diversity was calculated using the Bray-Curtis dissimilarity metric and visualised with PCoA and co-inertia analysis. Significant differences in beta-diversity are defined by Permutational MANOVA as implemented by the Adonis functionality in the vegan R package. Spearman correlation was used to identify features that were significantly correlated with the PCoA secondary eigenvector. The p-values were adjusted using Benjamini-Hochberg FDR method, and adjusted p values < 0.05 were considered significant. Wilcoxon rank sum test was used to identify significant and differentially abundant features. Significance was defined as nominal unadjusted P-value of 0.05.

3.3.5 Machine learning analysis

Machine learning classification was performed to investigate the predictive power of the mycobiome to distinguish between IBS and Controls, and to identify health-associated features. For this, support vector machines (SVMs) with linear kernel were employed on the logit transformed abundance, using Caret package in R. Models were trained using class weights. 10-fold cross validations for 100 iterations were performed, either using true labels, or shuffled class labels. For each 10-fold CV using shuffled class labels, we obtained a distribution of coefficients for each feature, which was considered as null distribution. A single sample z test of this distribution was compared to the true coefficient values. Significantly different features after p value

adjustments using Benjamini-Hochberg False Discovery Rate (FDR) method were identified and plotted (Fig. 3B).

3.3.6 Co-inertia analysis

Association between the faecal mycobiome and other datasets (bacteriome, metabolome and diet) from the same study population previously reported by Jeffery et al [8] was performed using co-inertia analyses.

For association analysis of each dataset, samples common to the mycobiome and the compared dataset was used. For shotgun species/genus data from 138 study participants (IBS: n = 80, control: n= 58) was transformed using Hellinger transformation. The dietary data based on the Food Frequency Questionnaire (FFQ) from 144 individuals (IBS: n = 80, control: n = 64) was used in the analysis where the frequencies were converted to per week consumption of food items. For the faecal metabolome, profiles from 142 participants (IBS: n = 80, Control: n = 62) were included and these were log10 transformed. The mycobiome profiles in all the comparisons was log10 transformed.

3.4 Results

3.4.1 The faecal mycobiome in IBS and controls

Next generation sequencing of the fungal ITS-1 region amplicon was performed on 144 (IBS; n=80 and control: n=64) faecal samples. The sequence data returned yielded 10,498,738 (median: 68044, IQR: 57684 - 87378) read pairs of length 251 base pairs (bp). After data pre-processing, 70.2% or 7,368,688 (median: 46099, IQR: 39274 - 59185) read pairs were retained and had a median length of 254 bp (IQR: 188 - 442). Of these reads, 95.2% or 7013893 (median: 44631, IQR: 37051 - 56672) mapped to 1631 fungal zOTUs (per sample median: 60, IQR: 46 - 75). Across the IBS and control

samples, the genera *Saccharomyces*, *Candida*, *Debaryomyces*, and *Penicillium* were the most abundant fungal genera. The variation in taxonomic prevalence was found to be sample specific (Supplementary Figure 1).

Beta-diversity of IBS and controls was investigated using Principal Coordinate Analysis (PCoA) (Fig1A). Permutational MANOVA showed significant differences between the mycobiome of IBS and controls (p-value = 0.004), with IBS and the control samples separating along the secondary principal axis. To identify features associated with this split, the abundances of the fungi were correlated with the second axis. Features identified as significant are presented in Figure 1B. Among the zOTUs significantly and positively correlated with the second principal axis, seven out of sixteen zOTUs belonged to the genera *Candida*, and *Saccharomyces*. Of these, two zOTUs (one classified as *Candida* spp. and another as *Candida albicans*), were significantly more abundant in IBS (p-value < 0.05). Three zOTUs classified as *Hanseniaspora ovarum*, commonly found in fresh fruits[26], were significantly less abundant in IBS (p-value < 0.05). A total of thirteen out of twenty-three fungal zOTUs that were significantly negatively associated with the IBS-control separation belonged to (at the genus level) *Penicillium*, *Agaricus*, and *Hanseniaspora*.

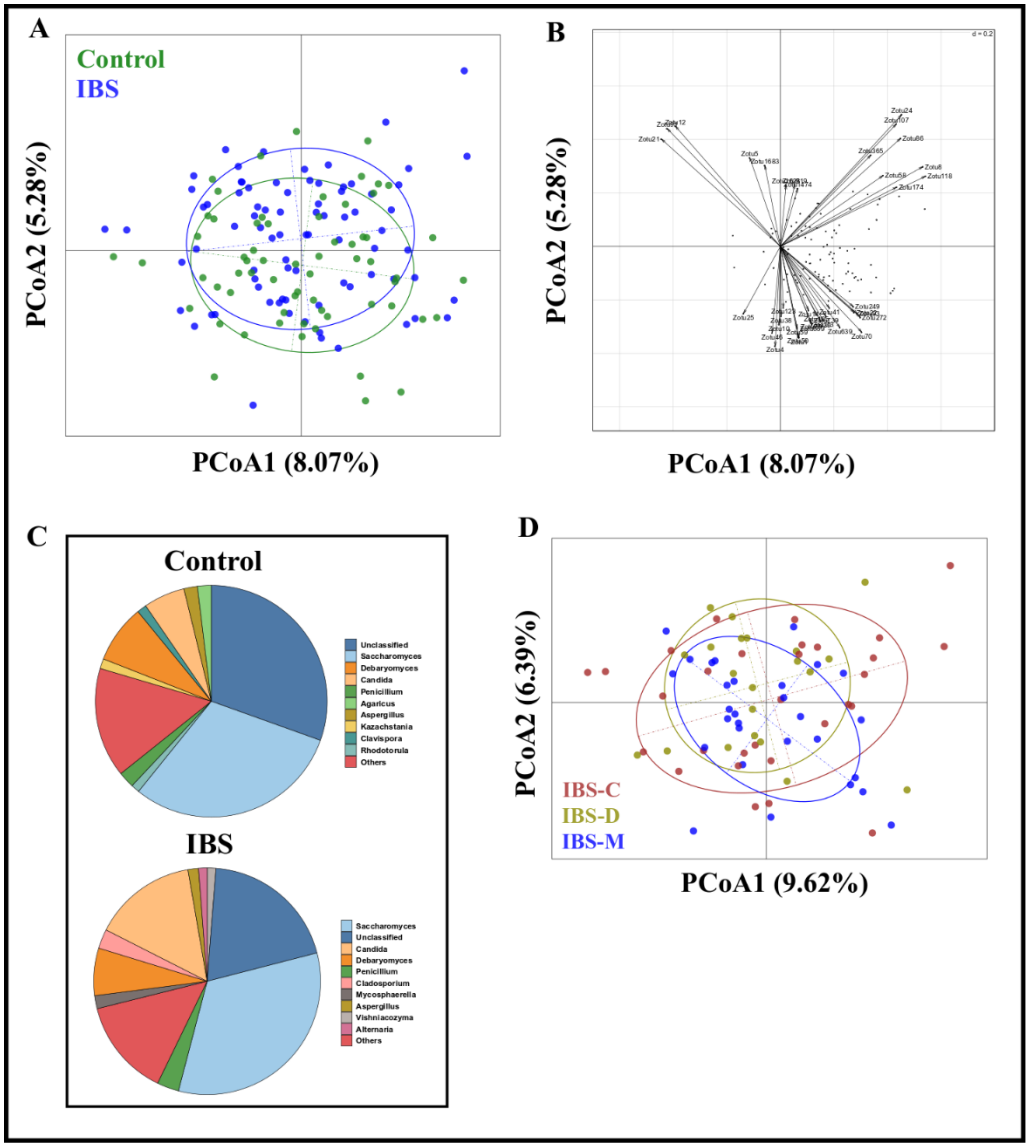


Figure 1: Mycobiome compositional differences between IBS and Control subjects. A) Principal coordinate analysis of sample data (n=144) using Bray-Curtis distance using log10 transformed abundance of zOTUs. A significant separation was observed between IBS (n =80), and control (n=64) samples, along the second principal axis. Significance was determined using permutational multivariate analysis of variance based on Bray-Curtis distance (Adonis; p-value = 0.004) B) Features and their correlation with second principal axis, showing direction of association. Features showing significant spearman correlation (adjusted $p < 0.05$) between their abundance and second axis are shown with labelled arrows. C) Pie charts showing top 10

abundant fungal genera in Control (n=64) and IBS populations (n=80). D) Principal coordinate analysis based on Bray Curtis distance and log10 transformed zOTU abundance profile of clinical sub-types of IBS population. No significant differences were observed between the 3 clinical subtypes (IBS-C: n=30, IBS-D: n=21 and IBS-M: n=29), (Adonis; p-value = 0.694).

Genus level differences in abundances of the mycobiome were detected between the IBS and control subjects (Figure 1C). Statistical analysis of IBS versus controls identified that *Malassezia*, *Candida*, and *Cladosporium* at the genus level were significantly more abundant in IBS (p-value < 0.05). No significant difference was observed between the IBS clinical subtypes (p-value = 0.694; Figure 1D). These differences must be placed in the context of the high level of variance in mycobiome feature abundance seen within the IBS and control groups.

Diversity analyses showed no significant differences in richness or Shannon diversity between IBS and controls (Figure 2 A and B), or among the IBS clinical subtypes (Fig. 2 C and D). Independent of clinical stratification, it is interesting to note that the largest trend as defined by the first principal component was based on differences in alpha-diversity (Supplementary Figure 2 A and B). Therefore, mycobiome diversity rather than taxonomy explains a major part of the variance in the dataset. While both Shannon diversity and richness showed significant positive correlation with the first axis, and the coefficient was much larger for the latter (Spearman rho: 0.596, 0.794 respectively).

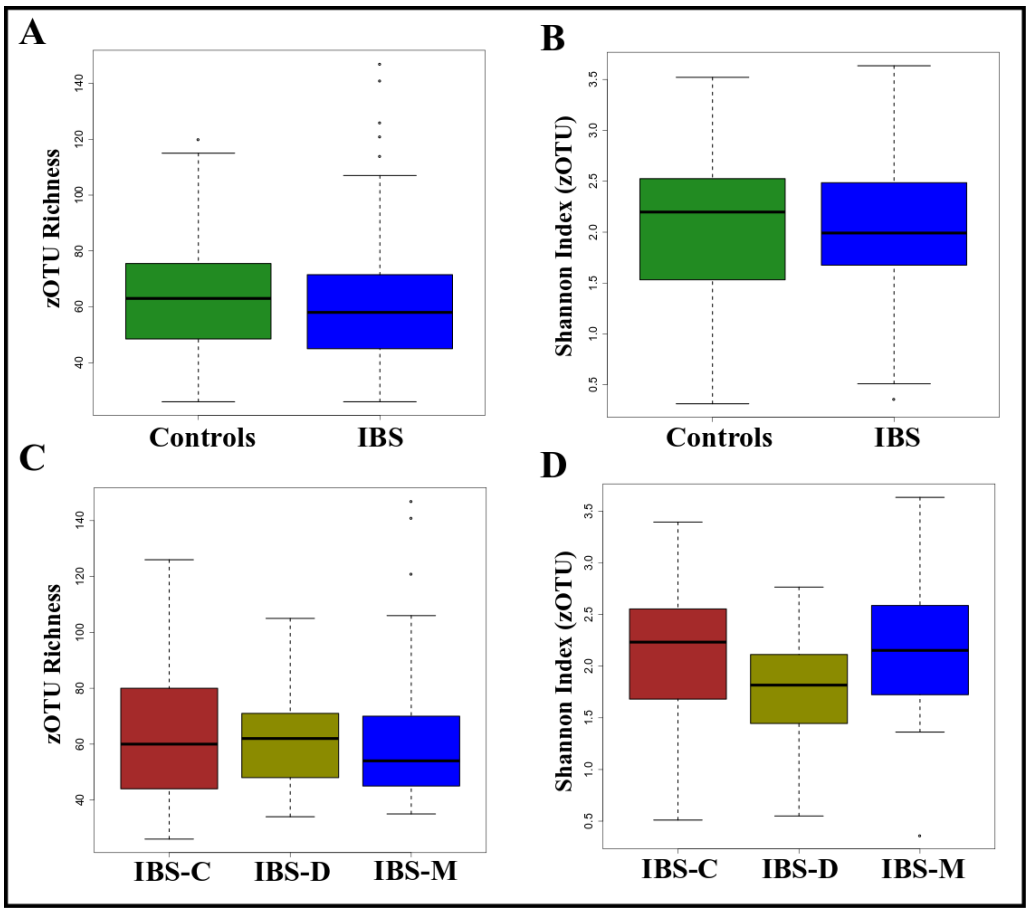


Figure 2: Mycobiome diversity differences in the analysed population. Two diversity measures were calculated, viz., Richness, and Shannon diversity index. None of the differences were significant. A) and B) zOTU richness, and Shannon diversity differences between IBS (n=80), and control (n=64) populations. Wilcoxon rank sum test returned nonsignificant p-values of 0.301 and 0.84, respectively. C) and D) zOTU richness, and Shannon diversity differences between clinical subtypes of IBS (IBS-C: n = 30, IBS-D: n = 21, IBS-M: n = 29). Kruskal-Walis test returned non-significant p-values of 0.75 and 0.17, respectively.

3.4.2 Predictive value of the mycobiome for IBS

We investigated the power of the mycobiome to predict IBS using a supervised machine learning model, Support Vector Machine (SVM) with a 10-fold cross validation. The returned values for sensitivity, specificity, and area under the curve

(AUCs) varied depending on the random seed used and therefore we repeated the analysis for 100 iterations where the class labels were maintained and compared to 100 iterations where the class labels were randomised. Based on Receiver Operating characteristics (ROC) curve analysis, the obtained AUC using true class labels was significantly different from shuffled labels (Figure 3A), suggesting that the prediction results obtained were not by chance. However, although significantly better than random, the classifier performance was not robust enough to distinguish between IBS and controls to diagnostic level accuracies. While the AUC (0.619 +/- 0.025), and specificity (0.355 +/- 0.053) were significantly higher than random, the sensitivity (0.776 +/- 0.032) was significantly lower than random. SVM coefficients, univariate statistics, and other information including taxonomic classification are summarized in Supplementary Table 1.

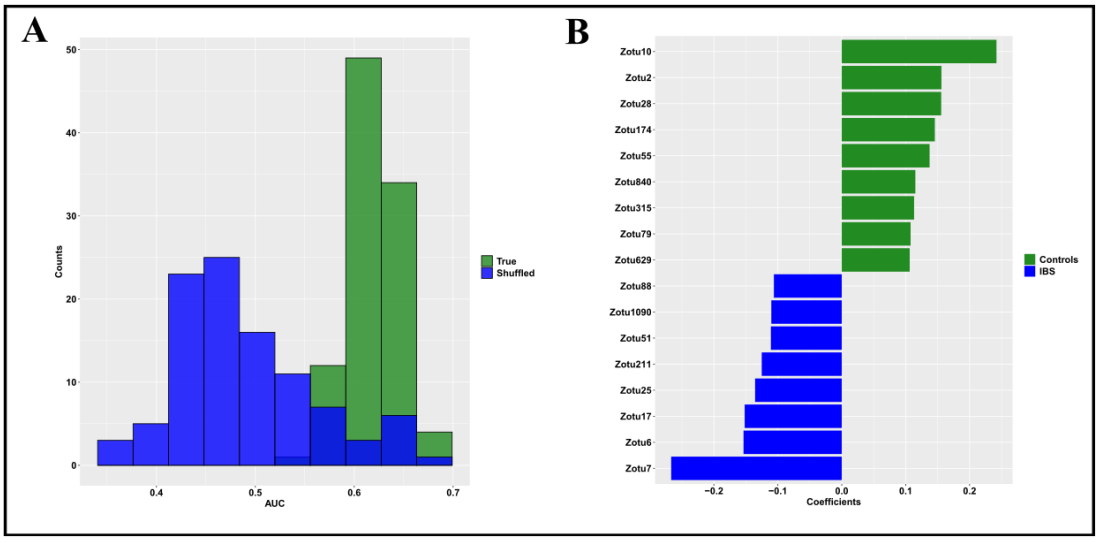


Figure 3: Predictive power of mycobiome to distinguish between IBS and controls. A) Distribution of mean Area under the curve (AUC) of each 10-fold cross validation, across 100 iterations, using true, and shuffled classes. The difference was significant between True (0.619 +/- 0.025) and Shuffled (0.489 +/- 0.071) classes (p-value < 0.001). B) Feature coefficients from trained models, significantly different from

random are shown. zOTUs with negative coefficients are associated with IBS, and those with positive coefficients are associated with controls (green).

Feature coefficients from corresponding models (True and shuffled classes) were obtained and compared using the z-test using package BSDA in R. Feature coefficients that were significantly different from random (adjusted p-value < 0.05, Benjamini-Hochberg FDR) are shown in Figure 3B. The two most predictive mycobiome zOTUs identified were zOTU7 (Cohen'sD = 0.374; classified as *Cladosporium cladosporioides*) which was more abundant in patients with IBS, and zOTU10 (Cohen's D = -0.711; an unidentified fungus) which was less abundant in IBS. These two zOTUs were confirmed to be significantly differentially abundant based on Wilcoxon rank sum test (Supplementary Figure 3A and B).

3.4.3 Co-variance of the mycobiome with faecal metabolome and bacterial metagenome

Co-inertia analysis based on PCA was performed to examine concordance between the faecal mycobiome and bacteriome, metabolome, and habitual dietary intake across the cohort and within IBS and control populations (Figure 4, Supplementary Table 2). Significance was tested using monte-carlo permutation methodology. The mycobiome showed significant co-variance with the bacteriome, and metabolome, but not with diet. Co-variance between the mycobiome and metabolomic and metagenomic data was higher in control subjects than in IBS patients (Supplementary Table 2).

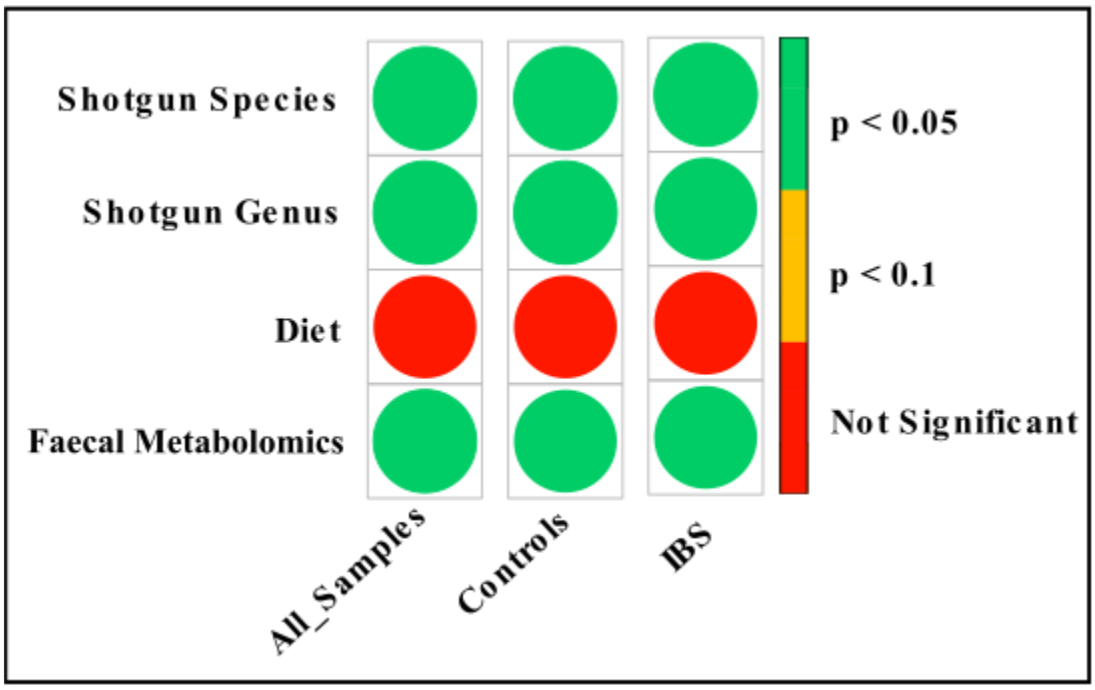


Figure 4: Co-inertia analysis between mycobiome composition with bacterial composition (All subjects: n=138, controls: n=58, IBS: n=80), food consumption (All subjects; n=144, controls: n=64, IBS: n=80), and faecal metabolome (All subjects; n=142, controls: n=62, IBS: n=80). Mycobiome showed significant association with bacterial composition and faecal metabolome but failed to show significance with food consumption. Degree of association as RV coefficient is presented in Supplementary Table 2.

To identify the variables that co-vary between the bacterial metagenome and the mycobiome, the co-inertia loadings of the two datasets were visualized (Supplementary Figure 4). The ordination of mycobiome and bacterial metagenomic datasets based on co-inertia are shown in Supplementary Figure 4A and B, respectively. Features in the direction of positive y-axis are associated with controls, while features in the opposite direction are associated with IBS (Supplementary Figure 4 C and D). Mycobiome zOTUs (Supplementary Figure 4 C) lying in the same

direction as shotgun species (Supplementary Figure 4 D) co-vary. Similarly, mycobiome and the faecal metabolome samples were visualized, based on co-inertia (Supplementary Figure 5A, and B). Mycobiome zOTUs (Supplementary Figure 5 C), and faecal metabolites (Supplementary Figure 5 D) in the direction of negative x-axis are associated with controls, while features lying in the opposite direction are associated with IBS. Further details of the feature IDs and names of the metabolites are given in Supplementary Table 3 and the taxonomy of the zOTUs are listed in Supplementary Table 1. These putative associations may indicate the major drivers of stratification within the datasets and the concordance between them, but more functional investigation is required to confidently infer the biological importance of these associations.

3.4.4 Gut mycobiome is not associated with Bile Acid Malabsorption

Bile acid malabsorption (BAM) is a condition which causes diarrhoea and which is difficult to distinguish from patients with diarrhoea-predominant IBS[27]. In our previous analysis, we observed that only the bacteriome of subjects with severe BAM (SeHCAT retention < 5%), showed a significant separation from that of patients in other BAM severity categories[6]. To investigate the differences in mycobiome between BAM classes, a subset of IBS patients (n=45) who underwent a SeHCAT assay to detect BAM (reported previously by Jeffery et al[6]), were analysed. No significant difference in the overall mycobiome was observed between BAM positive ($\leq$ 20% SeHCAT retention) and normal bile acid retention groups (Fig. 5).

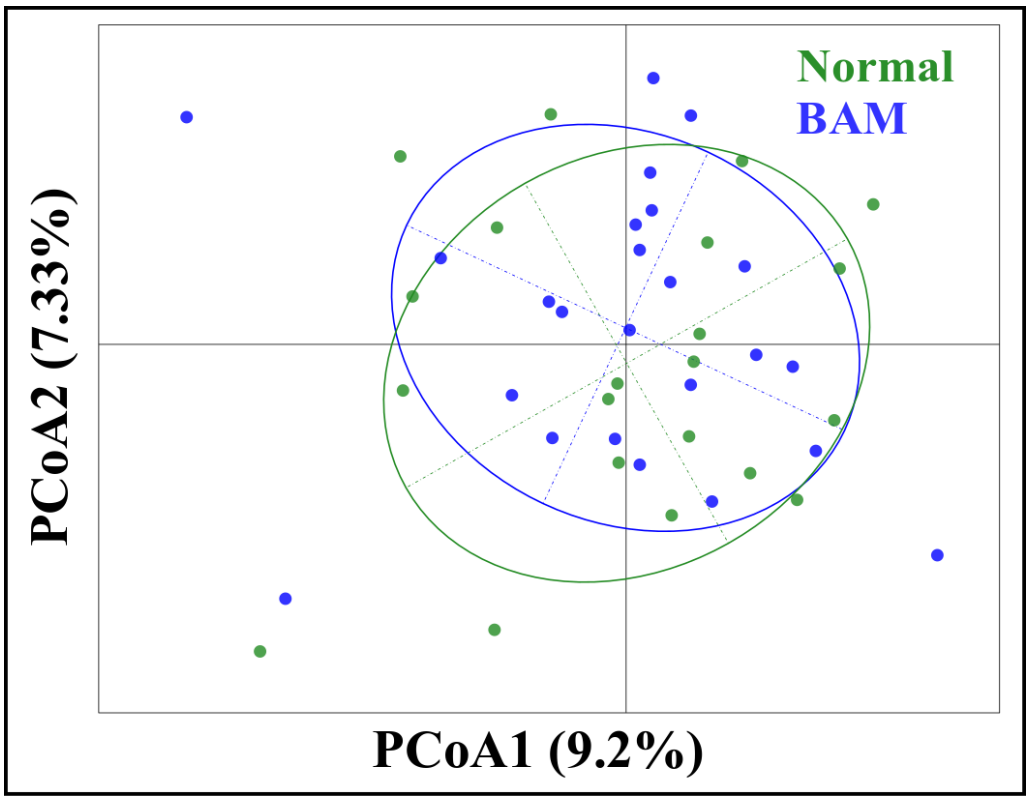


Figure 5: Principal coordinate analysis of IBS patients who had undergone SeHCAT assay to detect bile acid malabsorption (BAM). Samples were categorized as normal (n=21) (> 20% SeHCAT retention) or having borderline to severe BAM (<20% SeHCAT retention) (n = 24). There was no significant difference between the BAM classes (Adonis p-value = 0.985).

3.5 Discussion

This study shows that the faecal mycobiome varies considerably from person to person both in alpha-diversity and in its taxonomic composition. Nevertheless, the mycobiome differentiates patients with IBS from a control population. There were no differences in the mycobiome among the clinical subtypes of IBS or in the patients with BAM. We observed significant co-variation between mycobiome, and bacterial metagenome and faecal metabolome.

Intolerance to food and triggering or exacerbation of IBS symptoms is known to occur in some patients [28]. Such symptoms, including abdominal distress, are attributed to the intolerance of IBS patients to diet that is high in FODMAP content[28]. Therefore, a low FODMAP diet is sometimes recommended to patients with IBS. In our previously published study, we observed that clusters of food items high in FODMAP content were associated with the Control group[6]. Identification of depleted levels in IBS subjects of a yeast (*Hanseniospora ovarum*) which is commonly found in fresh fruits, appears to be in accordance to the food correlation data. However, there was no association with the overall dietary data, therefore suggesting no link between food intake and the fungal component of the microbiota. This contrasts with some published reports which linked habitual diet to mycobiome composition[29,30].

Multi-omics integrative studies pertaining to IBS are limited, but some reports show applicability of integrating bacteriome and metabolome for a better understanding of the disorder as potential biomarkers or a means to stratify patients[6,31]. Our analysis suggests that the mycobiome is a poor diagnostic biomarker for IBS, but it could be a potential contributor to a diagnostic score if combined with the bacteriome and metabolome.

Due to the wide variance in the mycobiome taxonomic composition by abundance and by alpha diversity, we experienced a limited ability to identify specific taxa associated with the observed beta-diversity separation between IBS and controls. The current study population is from a single geographic region and further studies in larger populations and in other geographical regions are therefore required. In this initial investigation of the mycobiome, we sequenced the ITS-1 region, but future studies should consider sequencing both the ITS-1 and ITS-2 regions.

While the mycobiome may have pathogenetic importance in IBS, it does not seem to have a sufficiently distinct signature for diagnostic utility. However, it exhibits substantial heterogeneity and may identify clinically important disease subsets or help stratify patients if used in combination with the bacteriome and metabolome.

3.6 Acknowledgements

We are grateful to the 4D pharma Cork clinical team, Caroline O'Leary, Marian O'Meara and Sinead Kelly who recruited the subjects to the study and collected the biological samples and clinical metadata. We would also like to thank gastroenterologist Dr Syed Akbar Zulquernain and Dr Elizabeth Kenny in Cork University Hospital who assisted with IBS patient recruitment. We acknowledge the input of 4D pharma Cork lab staff, Luke O'Brien, Fiona Grady, Marie McBrien and Siobhan O'Neill for sample processing for analysis. We also thank Michael Moore, Fintan Bradley, Tom Carty, Fionnuala McDonnell, Lean O'Donoghue, Pauline Beecher, Roisin Wilson and Shane Hayes at Cork University Hospital who performed the SeHCAT assay.

3.7 References

1. Ni J, Wu GD, Albenberg L, et al. Gut microbiota and IBD: causation or correlation? Nat Rev Gastroenterol Hepatol 2017;14:573–584.

2. Lloyd-Price J, Arze C, Ananthakrishnan AN, et al. Multi-omics of the gut microbial ecosystem in inflammatory bowel diseases. Nature 2019;569:655–662.

3. Hartstra AV, Bouter KEC, Bäckhed F, et al. Insights into the role of the microbiome in obesity and type 2 diabetes. Diabetes Care 2015;38:159–165.

4. Tap J, Derrien M, Törnblom H, et al. Identification of an Intestinal Microbiota Signature Associated With Severity of Irritable Bowel Syndrome. Gastroenterology 2017;152:111-123.e8.

5. Jeffery IB, O'Toole PW, Öhman L, et al. An irritable bowel syndrome subtype defined by species-specific alterations in faecal microbiota. Gut 2012;61:997–1006.

6. Jeffery IB, Das A, O'Herlihy E, et al. Differences in Fecal Microbiomes and Metabolomes of People With vs Without Irritable Bowel Syndrome and Bile Acid Malabsorption. Gastroenterology 2020;158:1016-1028.e8.

7. Gu Y, Zhou G, Qin X, et al. The Potential Role of Gut Mycobiome in Irritable Bowel Syndrome. Front Microbiol 2019;10:1894.

8. Koloski NA, Jones M, Kalantar J, et al. The brain--gut pathway in functional gastrointestinal disorders is bidirectional: a 12-year prospective population-based study. Gut 2012;61:1284–1290.

9. Gibson PR, Varney J, Malakar S, et al. Food components and irritable bowel syndrome. Gastroenterology 2015;148:1158-1174.e4.

10. Botschuijver S, Roeselers G, Levin E, et al. Intestinal Fungal Dysbiosis Is Associated With Visceral Hypersensitivity in Patients With Irritable Bowel Syndrome and Rats. Gastroenterology 2017;153:1026–1039.

11. Hong G, Li Y, Yang M, et al. Gut fungal dysbiosis and altered bacterial-fungal interaction in patients with diarrhea-predominant irritable bowel syndrome: An explorative study. Neurogastroenterol Motil 2020:e13891.

12. Li Y, Hong G, Yang M, et al. Fecal bacteria can predict the efficacy of rifaximin in patients with diarrhea-predominant irritable bowel syndrome. Pharmacol Res 2020;159:104936.

13. Santelmann H, Howard JM. Yeast metabolic products, yeast antigens and yeasts as possible triggers for irritable bowel syndrome. Eur J Gastroenterol Hepatol 2005;17:21–26.

14. Levine J, Dykoski RK, Janoff EN. Candida-associated diarrhea: a syndrome in search of credibility. Clin Infect Dis 1995;21:881–886.

15. Richard ML, Sokol H. The gut mycobiota: insights into analysis, environmental interactions and role in gastrointestinal diseases. Nat Rev Gastroenterol Hepatol 2019;16:331–345.

16. Wang ZK, Yang YS, Stefka AT, et al. Review article: fungal microbiota and digestive diseases. Aliment Pharmacol Ther 2014;39:751–766.

17. Middleton SJ, Coley A, Hunter JO. The role of faecal Candida albicans in the pathogenesis of food-intolerant irritable bowel syndrome. Postgrad Med J 1992;68:453–454.

18. Häuser W, Marschall U, Layer P, et al. The Prevalence, Comorbidity, Management and Costs of Irritable Bowel Syndrome. Dtsch Arztebl Int 2019;116:463–470.

19. Sciavilla P, Strati F, Di Paola M, et al. Gut microbiota profiles and characterization of cultivable fungal isolates in IBS patients. Appl Microbiol Biotechnol 2021;105:3277–3288.

20. Thiel IAM van, Stavrou AA, Jong A de, et al. Genetic and phenotypic diversity of fecal Candida albicans strains in irritable bowel syndrome. Sci Rep 2022;12:5391.

21. Huseyin CE, Rubio RC, O'Sullivan O, et al. The Fungal Frontier: A Comparative Analysis of Methods Used in the Study of the Human Gut Mycobiome. Front Microbiol 2017;8:1432.

22. Edgar RC, Flyvbjerg H. Error filtering, pair assembly and error correction for next-generation sequencing reads. Bioinformatics 2015;31:3476–3482.

23. Franzosa EA, McIver LJ, Rahnavard G, et al. Species-level functional profiling of metagenomes and metatranscriptomes. Nat Methods 2018;15:962–968.

24. Nilsson RH, Larsson K-H, Taylor AFS, et al. The UNITE database for molecular identification of fungi: handling dark taxa and parallel taxonomic classifications. Nucleic Acids Research 2019;47:D259–D264.

25. Bolyen E, Rideout JR, Dillon MR, et al. Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. Nat Biotechnol 2019;37:852–857.

26. Jankowski M, Jagielski T, Misiak G, et al. Hand dermatitis with Hanseniaspora uvarum as a plausible causative agent. Postepy Dermatol Alergol 2018;35:641–643.

27. Wedlake L, A'Hern R, Russell D, et al. Systematic review: the prevalence of idiopathic bile acid malabsorption as diagnosed by SeHCAT scanning in

patients with diarrhoea-predominant irritable bowel syndrome. Aliment Pharmacol Ther 2009;30:707–717.

28. Dolan R, Chey WD, Eswaran S. The role of diet in the management of irritable bowel syndrome: a focus on FODMAPs. Expert Rev Gastroenterol Hepatol 2018;12:607–615.

29. Hoffmann C, Dollive S, Grunberg S, et al. Archaea and fungi of the human gut microbiome: correlations with diet and bacterial residents. PLoS ONE 2013;8:e66019.

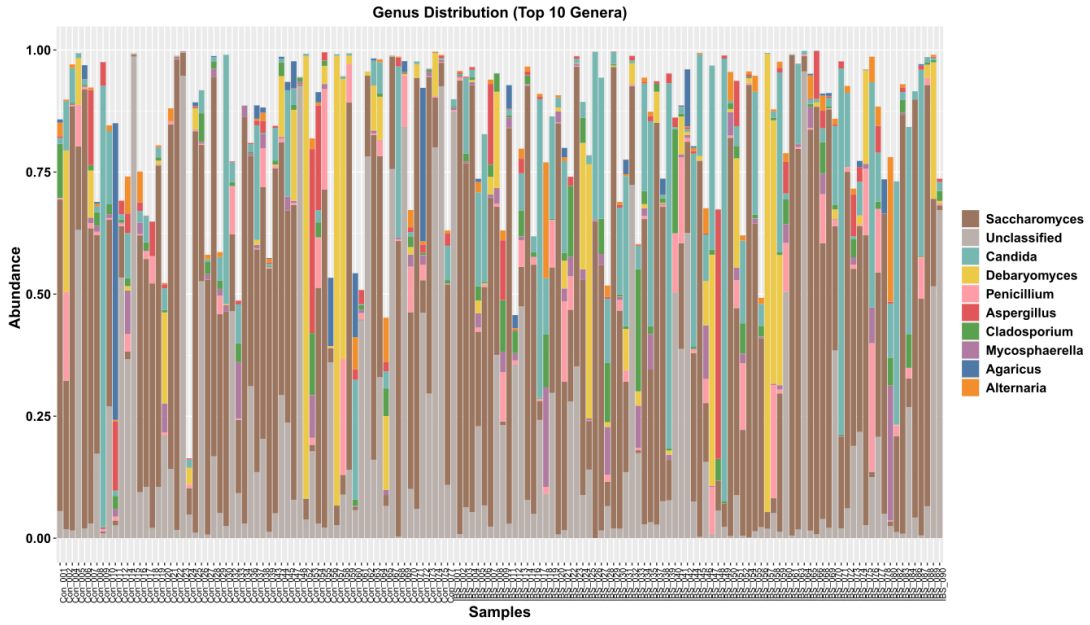
30. Auchtung TA, Fofanova TY, Stewart CJ, et al. Investigating Colonization of the Healthy Adult Gastrointestinal Tract by Fungi. mSphere 2018;3.

31. Hollister EB, Oezguen N, Chumpitazi BP, et al. Leveraging Human Microbiome Features to Diagnose and Stratify Children with Irritable Bowel Syndrome. J Mol Diagn 2019;21:449–461.

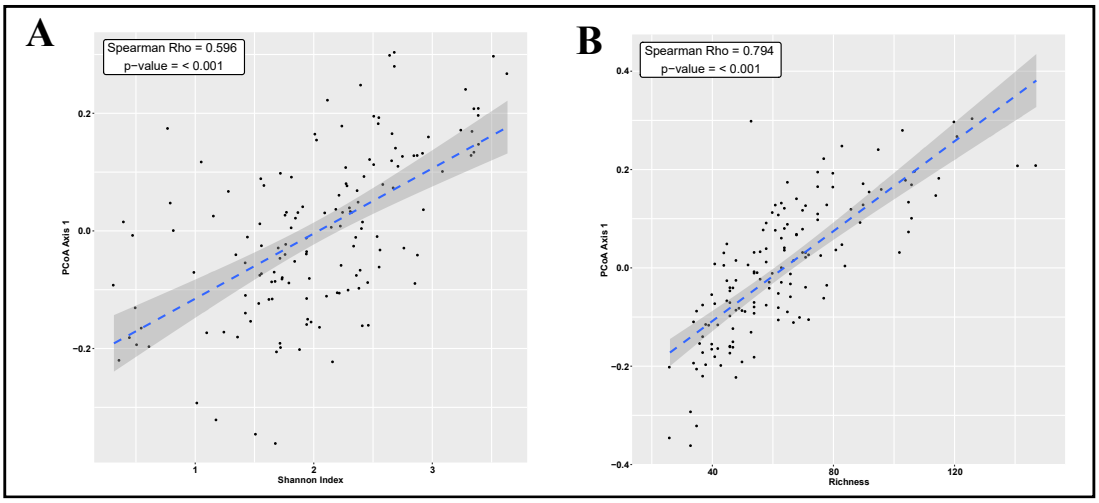
3.8 Supplementary information

3.8.1 Supplementary Figures

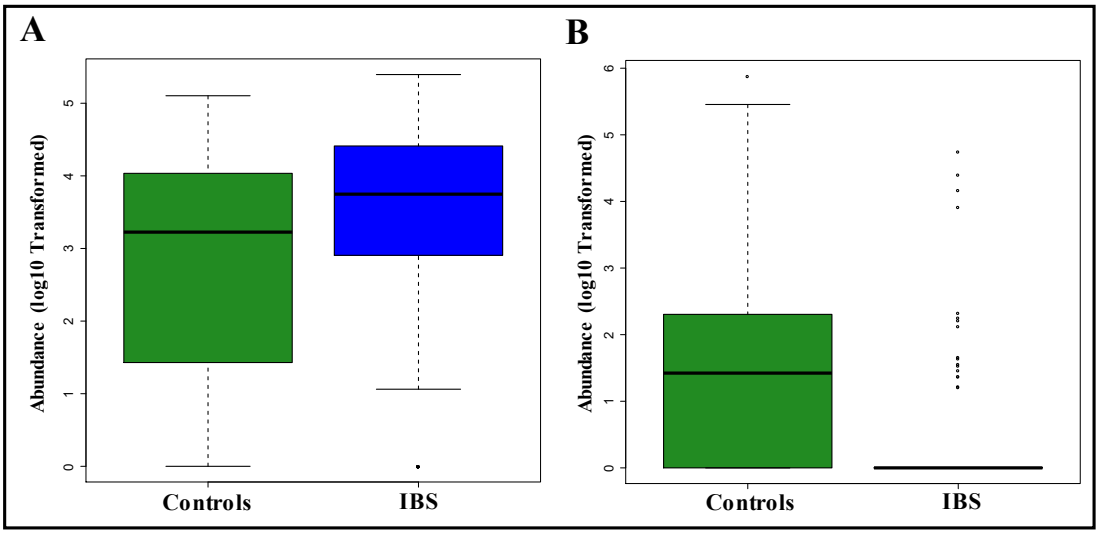
Supplementary Figure 1: Distribution of top ten abundant genera across samples. Genera are arranged in decreasing order of cumulative abundance.



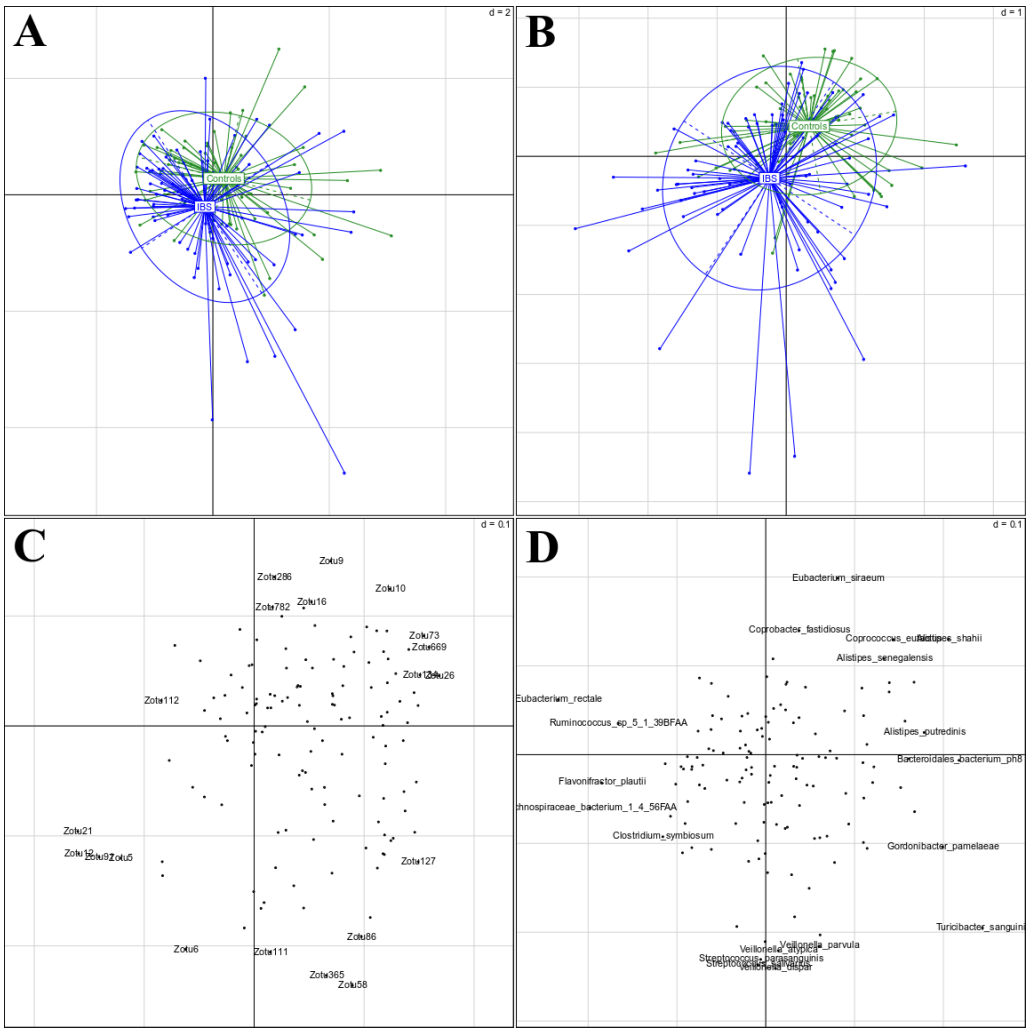
Supplementary Figure 2: Association of zOTU diversity to 1st principal axis, within the analyzed population (n=144). A) Linear association between Shannon diversity of zOTUs and the first principal axis. Spearman correlation was high and significant (Rho = 0.596, p-value < 0.001). B) Linear association between Richness of zOTUs and the first principal axis. Spearman correlation was high and significant (Rho = 0.794, p-value < 0.001).



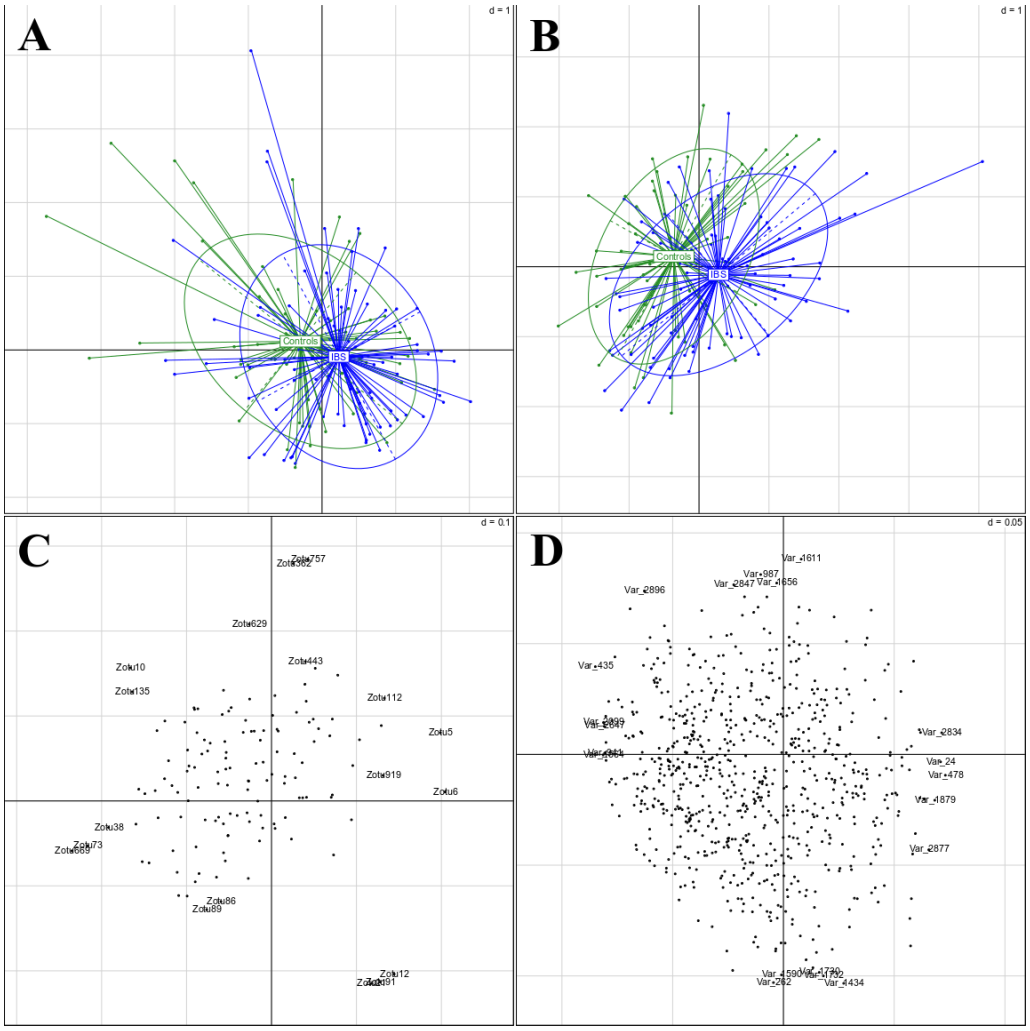
Supplementary Figure 3: Boxplots showing differences between IBS and controls with the 2 significant zOTUs based on SVM models. A) Boxplot of zOTU7, which was significantly higher in IBS (Wilcoxon rank sum test, p-value < 0.05). This zOTU was classified as *Cladosporium cladosporioides*. B) Boxplot of zOTU10 was significantly lower in IBS (Wilcoxon rank sum test, p-value < 0.05). This zOTU was an unclassified fungus.



Supplementary Figure 4: Co-inertia analysis plots showing co-variation between mycobiome (A) and species level metagenomic dataset (B). C: Highlights the loadings of fungal zOTUs and D highlights the loadings of bacterial species.



Supplementary Figure 5: Co-inertia plots showing covariance between mycobiome (A) and faecal metabolome (B). C: Highlights the loadings of fungal zOTUs. D: highlights the loadings of faecal metabolites. (Metabolites corresponding to variable IDs are given in Supplementary Table 3).



3.8.2 Supplementary Tables

Supplementary Table 1: Statistical summary of fungal zOTUs. The table summarizes results of various analyses performed to identify differences, or association between IBS and controls, including: Wilcoxon tests, SVM model coefficients, feature correlation with principal axes, taxonomic classification using QIIME2. This is a truncated table. For full table, please refer to the published manuscript (doi: 10.1038/s41598-020-79478-6).

	Wilcoxon test		SVM		PCoA Axis 1		PCoA Axis 2			
zOTU_ID	statistic	pvalue (adjusted)	Coefficient	pvalue (adjusted)	Correlation Coefficient	pvalue (adjusted)	Correlation Coefficient	pvalue (adjusted)	Qiime2 Taxon	Confidence
zOTU1	2735.5	0.866	0.063	1.45452E-18	0.07	0.568	0.341	0	o__Saccharomycetales	0.933
zOTU10	3480.5	0.002	0.242	2.0879E-188	-0.005	0.963	0.292	0.003	k__Fungi	1
zOTU100	2683	0.843	0.088	5.10136E-37	0.194	0.076	-0.035	0.795	s__Dioszegia_hungarica	0.965
zOTU104	2377	0.742	-0.071	2.51506E-15	0.029	0.827	0.028	0.83	k__Fungi	1
zOTU106	2515	0.911	-0.041	1.51784E-15	0.421	0	0.166	0.143	s__Sclerotinia_sclerotiorum	0.979
zOTU107	2308	0.742	0.056	6.47938E-14	0.433	0	-0.458	0	s__Alternaria_metachromatica	1
zOTU1089	2733.5	0.818	0.028	6.39984E-13	0.36	0	-0.1	0.392	s__unidentified	1
zOTU1090	2478.5	0.876	-0.11	9.61804E-47	0.27	0.007	0.099	0.393	s__Aureobasidium_pullulans	0.831
zOTU11	2371	0.842	0	0.600884916	0.179	0.107	0.146	0.211	p__Ascomycota	0.729
zOTU110	2595	0.911	0.047	1.72436E-13	0.208	0.054	0.163	0.149	g__Wallemia	1
zOTU111	2402.5	0.742	-0.063	6.18941E-14	0.138	0.236	-0.07	0.569	s__Sawadaea_bicornis	0.995
zOTU112	2217	0.51	-0.049	1.43295E-12	0.171	0.128	-0.057	0.645	s__Penicillium_lanosum	0.802
zOTU114	2525	0.919	-0.044	1.68E-07	0.031	0.816	0.01	0.942	s__Kazachstania_unispora	1
zOTU1154	2692	0.866	0.039	0.014891114	-0.032	0.807	-0.067	0.587	k__Fungi	1
zOTU116	2567	0.986	0.026	0.059332198	0.291	0.003	0.126	0.287	s__Penicillium_lanosum	0.98
zOTU117	2428	0.842	-0.065	5.89154E-22	0.043	0.747	-0.04	0.763	s__Wickerhamomyces_onychis	1
zOTU118	2651.5	0.876	-0.028	0.000913124	0.542	0	-0.262	0.01	s__Alternaria_metachromatica	1
zOTU119	2622	0.876	0.013	0.877414692	0.116	0.332	-0.026	0.843	s__Mycoacia_uda	1

zOTU12	2565	0.986	0.034	2.65142E-11	-0.392	0	-0.45	0	s__Saccharomyces_cerevisiae	1
zOTU121	2472	0.876	-0.035	0.00032718	0.282	0.005	0.113	0.338	s__unidentified	0.819
zOTU123	2723.5	0.818	0.098	2.94019E-24	0.011	0.933	0.229	0.029	s__Torulaspora_delbrueckii	1
zOTU127	2603	0.919	0.044	1.58962E-15	0.303	0.002	0.123	0.299	s__Aureobasidium_pullulans	0.922
zOTU13	2783	0.685	0.05	1.02179E-06	0.022	0.874	0.207	0.056	p__Basidiomycota	0.779
zOTU134	2386	0.742	-0.039	5.47956E-07	0.234	0.025	0.14	0.23	o__Eurotiales	0.997
zOTU135	2685	0.843	0.024	9.19439E-07	0.135	0.249	0.256	0.012	s__Cladosporium_sphaerospermum	0.812
zOTU1391	2666.5	0.866	0.043	4.13549E-12	0.372	0	-0.177	0.11	s__unidentified	1
zOTU14	2528	0.952	-0.052	1.45567E-07	0.105	0.37	0.239	0.022	s__Candida_parapsilosis	1
zOTU1415	2473	0.866	-0.027	0.032683018	0.06	0.626	-0.09	0.446	k__Fungi	1
zOTU1474	2299	0.433	-0.023	0.002154952	0.065	0.596	-0.216	0.042	g__Candida	0.99
zOTU15	2398.5	0.843	0.017	0.152367039	-0.056	0.645	0.142	0.222	f__Dipodascaceae	0.999
zOTU1540	2708	0.842	0.032	3.64216E-07	0.327	0.001	-0.195	0.074	s__unidentified	1
zOTU1555	2611.5	0.911	0.034	5.27404E-11	0.445	0	-0.188	0.089	s__unidentified	1
zOTU157	2869	0.371	0.075	3.11982E-30	0.144	0.215	0.113	0.338	o__Saccharomycetales	0.727
zOTU16	2759	0.773	0.083	3.41765E-24	0.065	0.596	0.135	0.247	g__Penicillium	0.995
zOTU1628	2440	0.842	-0.034	1.33182E-07	0.021	0.876	-0.229	0.029	k__Fungi	1
zOTU163	2427	0.876	-0.021	0.013903956	-0.061	0.623	0.019	0.889	k__Fungi	1
zOTU1648	2529	0.929	-0.056	3.11149E-16	0.365	0	-0.157	0.17	s__unidentified	1
zOTU1683	2257	0.433	-0.096	7.64554E-50	-0.058	0.638	-0.303	0.002	s__Candida_albicans	0.705
zOTU1699	2652	0.866	0.07	7.40645E-40	0.263	0.01	-0.141	0.228	s__unidentified	1
zOTU17	2123	0.355	-0.152	1.51341E-64	0.198	0.069	0.127	0.281	s__Cystofilobasidium_macerans	1
zOTU174	2849	0.607	0.146	1.94959E-76	0.435	0	-0.222	0.036	s__Neoascochyta_paspali	0.993
zOTU1775	2502.5	0.911	0.024	0.014661681	-0.028	0.827	-0.048	0.706	k__Fungi	1
zOTU18	2649	0.876	-0.017	0.009003905	0.426	0	0.157	0.17	g__Aspergillus	1
zOTU1873	2686.5	0.866	0.021	0.071699261	0.095	0.414	-0.076	0.529	k__Fungi	1
zOTU19	2544	0.961	-0.026	0.000557855	0.097	0.407	0.057	0.643	o__Saccharomycetales	0.862
zOTU190	2500	0.876	-0.043	0.000166738	0.184	0.097	0.03	0.819	s__Aspergillus_penicillioides	0.999
zOTU2	2711	0.866	0.156	7.25778E-74	0.067	0.585	-0.042	0.749	s__unidentified	0.9
zOTU20	2633	0.907	-0.005	0.875228402	0.26	0.011	0.167	0.138	s__Rhodotorula_mucilaginosa	1

zOTU209	2616	0.876	0.011	0.408244759	0.201	0.064	0.026	0.843	s__Wickerhamomyces_anomalus	1
zOTU21	2428	0.876	-0.005	0.221040586	-0.444	0	-0.401	0	s__Saccharomyces_cerevisiae	1
zOTU211	2285	0.374	-0.125	3.98313E-45	0.135	0.249	-0.144	0.215	g__Fusarium	0.968
zOTU216	2866	0.355	0.054	8.17755E-14	0.014	0.916	0.124	0.294	o__Saccharomycetales	0.748
zOTU22	2667.5	0.876	0.042	5.49055E-08	0.276	0.006	0.247	0.016	s__Penicillium_spinulosum	0.719
zOTU225	2443	0.866	-0.064	1.77332E-17	0.272	0.007	-0.206	0.056	s__Neoascochyta_paspali	0.986
zOTU233	2853	0.374	0.068	2.4729E-18	-0.177	0.111	-0.041	0.758	k__Fungi	1
zOTU235	2511.5	0.911	-0.023	0.021159259	0.255	0.012	0.062	0.616	s__Cladosporium_cladosporioides	0.899
zOTU239	2516	0.911	0.002	0.023731415	0.111	0.345	-0.072	0.558	s__Cladosporium_cladosporioides	0.981
zOTU24	2315	0.818	-0.039	1.15923E-13	0.453	0	-0.493	0	s__Alternaria_metachromatica	1
zOTU245	2663	0.866	0.028	0.004418971	0.145	0.212	0.105	0.373	s__Kluyveromyces_marxianus	0.987
zOTU249	2856.5	0.388	0.05	4.92388E-22	0.273	0.007	0.226	0.032	s__Hanseniaspora_uvarum	0.955
zOTU25	2628	0.909	-0.136	2.60385E-58	-0.138	0.236	0.251	0.014	s__Candida_glaebosa	1
zOTU26	2846	0.644	0.047	1.32675E-14	0.412	0	0.174	0.12	s__Sclerotinia_sclerotiorum	0.948
zOTU268	2758	0.685	0.044	9.07128E-08	0.087	0.467	0.083	0.485	k__Fungi	1
zOTU27	2529	0.952	0.005	0.072977937	0.11	0.348	-0.113	0.338	s__Candida_tropicalis	1
zOTU272	2854	0.374	0.057	1.57197E-31	0.299	0.003	0.265	0.009	s__Hanseniaspora_uvarum	0.867
zOTU28	2770.5	0.685	0.156	6.3588E-115	0.027	0.838	0.104	0.377	s__Tausonia_pullulans	1
zOTU286	2732	0.685	0.098	1.77096E-55	-0.024	0.859	0.08	0.503	s__Tylospora_fibrillosa	1
zOTU288	2489.5	0.876	-0.01	0.19020041	0.339	0	0.069	0.576	g__Aspergillus	0.892
zOTU3	2445.5	0.876	-0.055	7.31112E-11	-0.172	0.127	0.079	0.512	s__Debaryomyces_udenii	1
zOTU30	2663	0.866	0.066	1.65286E-21	0.202	0.062	-0.061	0.623	s__Kazachstania_barnettii	1
zOTU31	2159	0.433	-0.037	8.90946E-05	0.331	0.001	-0.066	0.591	s__Vishniacozyma_victoriae	0.999
zOTU311	2603	0.911	-0.001	0.234791975	0.131	0.262	0.095	0.417	k__Fungi	1
zOTU315	2680	0.843	0.113	2.5026E-53	0.18	0.105	0.203	0.06	s__unidentified	0.996
zOTU317	2303	0.631	-0.02	0.000825282	0.145	0.212	-0.16	0.158	s__Malassezia_restricta	1
zOTU329	2725	0.742	-0.003	0.183737251	0.254	0.013	0.093	0.427	s__unidentified	0.991
zOTU33	2668	0.866	0.045	3.36433E-11	0.187	0.09	0.167	0.138	s__Cystofilobasidium_infirmominiatum	1
zOTU333	2580	0.952	0.026	0.078532555	0.08	0.501	-0.059	0.635	k__Fungi	1
zOTU34	2426	0.843	0	0.958497023	0.162	0.154	0.001	0.992	s__Clavispora_lusitaniae	0.706

zOTU36	2632.5	0.876	0.056	4.78332E-14	0.076	0.529	0.111	0.346	s__Wallemia_muriae	0.969
zOTU360	2654	0.866	0.057	3.39168E-13	0.013	0.927	0.003	0.981	s__Cyberlindnera_rhodanensis	1
zOTU362	2888	0.461	0.064	2.10393E-26	0.164	0.149	0.114	0.335	k__Fungi	1
zOTU365	2358	0.731	-0.053	9.13511E-16	0.338	0	-0.341	0	g__Microdochium	0.841
zOTU37	2746.5	0.842	0.062	9.76315E-12	0.339	0	0.171	0.128	g__Aspergillus	1
zOTU38	2909	0.584	0.036	1.49817E-09	0.003	0.978	0.261	0.01	s__Agaricus_bisporus	0.993
zOTU39	2392	0.842	-0.049	2.75141E-08	0.149	0.199	0.279	0.005	s__Aspergillus_ruber	0.722
zOTU4	2579	0.961	-0.062	2.36734E-20	-0.019	0.889	0.372	0	f__Dipodascaceae	1
zOTU40	2375	0.777	-0.024	0.048823112	0.15	0.197	0.049	0.706	s__Holtermanniella_takashimae	0.933
zOTU41	2392	0.843	-0.05	8.1069E-11	0.184	0.098	0.231	0.028	s__Penicillium_lanosum	0.98
zOTU411	2414	0.843	-0.047	5.64678E-07	0.251	0.014	0.016	0.907	s__Penicillium_lanosum	0.989
zOTU42	2791	0.685	0.08	1.61396E-26	0.222	0.036	-0.08	0.503	s__Vishniacozyma_victoriae	0.809
zOTU43	2522	0.911	-0.002	0.46806171	0.058	0.635	0.125	0.29	p__Ascomycota	0.736
zOTU44	2766	0.618	0.03	1.64966E-05	0.111	0.346	0.063	0.609	k__Fungi	1
zOTU443	2479	0.876	-0.079	5.48209E-20	0.055	0.66	0.107	0.366	k__Fungi	1
zOTU46	2659	0.876	0.014	0.001230663	-0.031	0.816	0.323	0.001	o__Saccharomycetales	0.915
zOTU476	2563	0.986	0.036	1.49307E-06	0.161	0.154	0.124	0.292	k__Fungi	1
zOTU5	1907	0.355	-0.067	8.32025E-26	-0.114	0.335	-0.332	0.001	s__Candida_albicans	0.903
zOTU50	2545	0.961	-0.061	2.31067E-15	0.066	0.591	0.338	0	s__unidentified	0.809
zOTU51	2209.5	0.374	-0.111	8.53212E-53	0.11	0.346	0.03	0.818	s__Penicillium_lanosum	0.784
zOTU549	2576	0.955	-0.048	1.9653E-10	0.001	0.994	-0.024	0.857	k__Fungi	1
zOTU55	2770.5	0.631	0.137	9.97443E-90	0.202	0.062	0.06	0.628	s__Aspergillus_japonicus	0.884
zOTU56	2621	0.876	0.022	4.15757E-07	0.247	0.016	0.101	0.39	s__unidentified	0.995
zOTU58	2644	0.876	0.051	2.74275E-17	0.384	0	-0.265	0.009	s__Sporobolomyces_roseus	1
zOTU59	2601	0.911	-0.046	0.000209442	0.063	0.611	0.307	0.002	s__Dipodascus_geotrichum	0.828
zOTU6	2038	0.355	-0.154	8.92715E-84	0.042	0.749	-0.076	0.529	s__Candida_dubliniensis	1
zOTU629	2895	0.388	0.106	3.30943E-44	0.11	0.346	0.12	0.313	k__Fungi	1
zOTU639	2803	0.433	0.034	7.74335E-11	0.222	0.035	0.303	0.002	s__Hanseniaspora_uvarum	0.905
zOTU65	2516.5	0.911	-0.058	1.69169E-19	0.277	0.006	0.117	0.332	s__unidentified	0.956
zOTU669	2640	0.876	0.023	0.002089108	0.118	0.325	0.293	0.003	s__Agaricus_bisporus	0.981

zOTU7	1962	0.355	-0.267	2.2619E-255	0.54	0	-0.14	0.23	s__Cladosporium_cladosporioides	0.995
zOTU70	2735	0.842	0.04	2.33268E-14	0.304	0.002	0.321	0.001	s__Penicillium_bialowiezense	0.952
zOTU73	2837	0.685	0.046	1.0651E-15	0.16	0.159	0.277	0.006	s__Agaricus_bisporus	0.974
zOTU739	2860	0.355	0.004	0.165197811	0.171	0.128	0.262	0.01	s__Hanseniaspora_uvarum	0.909
zOTU757	2734	0.842	0.019	0.100563755	0.197	0.072	0.045	0.73	k__Fungi	1
zOTU76	2439.5	0.866	-0.029	0.001591301	0.21	0.051	-0.05	0.692	s__Bullera_unica	1
zOTU77	2571	0.961	0.005	0.34714402	-0.015	0.91	-0.057	0.643	k__Fungi	1
zOTU78	2532	0.944	-0.003	0.558803963	0.189	0.086	0.021	0.881	s__Vishniacozyma_victoriae	0.98
zOTU782	2720	0.842	0.051	2.5956E-14	0.113	0.338	0.062	0.613	k__Fungi	1
zOTU79	2839	0.461	0.108	7.78844E-65	0.115	0.335	0.286	0.004	s__unidentified	0.859
zOTU8	2343	0.842	0.007	0.215911119	0.533	0	-0.298	0.003	s__Mycosphaerella_tassiana	0.995
zOTU80	2285	0.461	-0.065	1.87334E-13	0.151	0.193	-0.123	0.297	s__Itersonilia_perplexans	1
zOTU83	2621	0.876	0	0.673365289	0.359	0	0.116	0.333	s__Sclerotinia_sclerotiorum	0.933
zOTU84	2459	0.866	-0.044	4.14267E-10	-0.036	0.787	-0.034	0.796	s__Wickerhamomyces_onychis	1
zOTU840	2905.5	0.461	0.115	1.65477E-53	0.075	0.534	-0.123	0.297	k__Fungi	1
zOTU850	2505	0.876	-0.034	2.12581E-05	0.01	0.942	0.012	0.932	k__Fungi	1
zOTU854	2726	0.742	0.029	0.000285843	0.028	0.831	0.17	0.129	k__Fungi	1
zOTU86	2290	0.685	-0.076	6.78361E-26	0.45	0	-0.404	0	g__Microdochium	0.825
zOTU874	2232	0.355	-0.042	3.61236E-08	0.056	0.645	-0.181	0.105	s__Malassezia_restricta	0.738
zOTU88	2378	0.773	-0.106	8.67392E-34	0.14	0.23	-0.161	0.154	s__Itersonilia_perplexans	1
zOTU89	2619	0.876	0.078	2.52997E-23	0.176	0.114	-0.006	0.958	s__Penicillium_citrinum	1
zOTU9	2517	0.911	-0.042	6.31215E-07	0.087	0.462	0.18	0.105	s__Diutina_catenulata	0.999
zOTU91	2432	0.876	0.054	1.66038E-14	-0.422	0	-0.44	0	s__Saccharomyces_cerevisiae	1
zOTU919	1923	0.355	-0.095	9.71415E-32	0.054	0.665	-0.231	0.028	g__Candida	0.996
zOTU92	2896	0.374	0.047	3.14697E-11	0.184	0.098	0.11	0.346	s__Meyerozyma_caribbica	0.92
zOTU927	2396	0.784	-0.007	0.393277154	0.025	0.854	-0.1	0.392	s__Malassezia_restricta	0.84
zOTU93	2642	0.876	0.05	3.04875E-08	0.287	0.004	0.25	0.015	s__Penicillium_bialowiezense	0.937
zOTU94	2426	0.843	-0.013	0.156431382	0.224	0.033	0.005	0.963	s__Penicillium_bialowiezense	1
zOTU95	2540	0.954	0.009	0.845717736	0.278	0.006	0.008	0.949	s__Alternaria_brassicae	0.866
zOTU98	2633	0.876	0.046	3.46458E-10	0.094	0.42	-0.033	0.804	f__Nectriaceae	0.961

Supplementary Table 2: Summary of co-inertia analysis listing the RV coefficients and p-values between mycobiome, and shotgun, diet and faecal metabolomic datasets.

Coinertia	All	IBS	Controls
Shot_Sp (All subjects: N=138; IBS: n = 80; Controls: n = 58)	0.372 (0.001)	0.466 (0.022)	0.58 (0.015)
Shot_Genus (All subjects N=138; IBS: n= 80; Controls: n = 58)	0.304 (0.001)	0.39 (0.04)	0.495 (0.035)
Diet (FFQ) (All subjects: N=144; IBS: n = 80; Controls: n = 64)	0.273 (0.135)	0.35 (0.236)	0.49 (0.288)
Faecal Metabolomics (All subjects: N=142; IBS: n = 80; Controls: n = 62)	0.3 (0.001)	0.385 (0.022)	0.48 (0.001)

For each comparison, common and matched samples used. Values are represented as RV (p-value)
Normalized Mycobiome zOTU profile filtered at 10% occurrence, and log10 transformed.
Shotgun data: Hellinger transformation
Diet (FFQ): Values converted to frequency per week.
Faecal Metabolomics: Log10 transformed

Supplementary Table 3: Faecal metabolite Variable IDs and metabolite name

Faecal_Metabolite_ID	MetaboliteName
Var_24	Ile-Pro
Var_262	Heptadecanoic acid
Var_435	7C-aglycone
Var_478	Arachidonic acid
Var_941	[FA oxo,hydroxy(2:0)] 9-oxo-11R,15S-dihydroxy-1a,1b-dihomo-13E-prostaenoic acid
Var_987	3-Hydroxy-10'-apo-b,y-carotenal
Var_1434	Lys-Lys-Pro-Pro
Var_1564	Cer(d31:0)
Var_1590	Theasapogenol A
Var_1611	1-18:1-lysophosphatidylglycerol
Var_1656	LysoPC(18:1(9Z))
Var_1730	Arg-Leu-Pro-Arg
Var_1732	[ST hydrox] (25R)-3alpha,7alpha-dihydroxy-5beta-cholestan-27-oyl taurine
Var_1879	I-Urobilin
Var_2099	PG(O-30:1)
Var_2647	TG(20:5(5Z,8Z,11Z,14Z,17Z)/18:1(9Z)/22:5(7Z,10Z,13Z,16Z,19Z))[iso6]
Var_2834	Phenol
Var_2847	trans-2-Heptenal
Var_2877	Allyl caprylate
Var_2896	Acelan A

Chapter 4

The gut virome in Irritable Bowel Syndrome differs from that of controls.

This chapter has been published as

Simone Coughlan, **Anubhav Das**, Eileen O'Herlihy, Fergus Shanahan, Paul W. O'Toole, and Ian B. Jeffery*. The gut virome in Irritable Bowel Syndrome differs from that of controls. Gut Microbes 2021;13:1887719.

*Corresponding author

Author contributions: IBJ, EOH, FS and PWOT conceived the study idea and design. SC, IBJ, EOH, FS and PWOT drafted and reviewed the manuscript. SC, IBJ and AD performed the bioinformatics analysis.

AD Contribution:

Analysis: Multivariate statistical analyses of virome abundance. Predictive model development and machine learning analysis.

Others: Contribution to manuscript by drafting of applicable sections of Methods, and Results.

Table of Contents

4.1 Abstract 245
4.2 Introduction 246
4.3 Methods 247
4.3.1 Study participant recruitment 247
4.3.2 Extraction, Library Preparation and Shotgun Sequencing of Faecal VLP 249
4.3.3 Sequence Data Pre-Processing 250
4.3.3.1 Read processing. 250
4.3.3.2 Assembly and dereplication 251
4.3.3.3 Mapping and Abundance calculation 251
4.3.4 Selection of putative viral contigs 252
4.3.5 Identification of Viral Clusters (VCs) 253
4.3.6 Bacterial host prediction based on CRISPR protospacers 253
4.3.7 Statistical Analysis 254
4.3.8 Taxonomic assignment of contigs 255
4.3.9 Taxonomic assignment of viral clusters 256
4.3.10 Temperate phage prediction and differential abundance 256
4.3.11 Terminase annotation 257
4.4 Results 257
4.4.1 Taxonomy of known and unknown phage sequences in IBS 257
4.4.2 The gut virome differs between IBS and controls 260
4.4.3 The gut virome does not co-vary with the bacteriome or BAM status 266
4.4.4 No evidence of conversion from lytic to lysogenic life cycle in IBS 267
4.5 Discussion 267
4.6 Acknowledgements 271
4.7 References 271
4.8 Supplementary information 280
4.8.1 Supplementary figures 280
4.8.2 Supplementary tables 288

4.1 Abstract

Irritable Bowel Syndrome (IBS), the most common gastrointestinal disorder, is diagnosed solely on symptoms. Potentially diagnostic alterations in the bacterial component of the gut microbiome (the bacteriome) are associated with IBS, but despite the known role of the virome (particularly bacteriophages), in shaping the gut bacteriome, few studies have investigated the virome in IBS. We performed metagenomic sequencing of faecal Virus-Like Particles (VLPs) from 55 patients with IBS and 51 control individuals. We detected significantly lower alpha diversity of viral clusters comprising both known and novel viruses (viral 'dark matter') in IBS and a significant difference in beta diversity compared to controls, but not between IBS symptom subtypes. The three most abundant bacteriophage clusters belonged to the *Siphoviridae*, *Myoviridae*, and *Podoviridae* families (Order *Caudovirales*). A core virome (defined as a cluster present in at least 50% of samples) of 5 and 12 viral clusters was identified in IBS and control subjects, respectively. We also identified a subset of viral clusters that showed differential abundance between IBS and controls. The virome did not co-vary significantly with the bacteriome, with IBS clinical subtype, or with Bile Acid Malabsorption status. However, differences in the virome could be related back to the bacteriome as analysis of CRISPR spacers indicated that the virome alterations were at least partially related to the alterations in the bacteriome. We found no evidence for a shift from lytic to lysogenic replication of core viral clusters, a phenomenon reported for the gut virome of patients with Inflammatory Bowel Disease. Collectively, our data show alterations in the virome of patients with IBS, regardless of clinical subtype, which may facilitate development of new microbiome-based therapeutics.

4.2 Introduction

Irritable Bowel Syndrome (IBS) is the most common functional gastrointestinal disorder, affecting about 11% of the global population [1]**.** Diagnosis of IBS is based on symptoms, including recurrent abdominal pain and changes in bowel habits using the Rome IV diagnostic criteria. Three clinical subtypes are recognized: IBS with constipation (IBS-C), IBS with diarrhoea (IBS-D), and IBS with mixed bowel habits (IBS-M) [2]**.** The cause of IBS is unknown and perhaps multi-factorial, with input from genetics [3,4], psychological stress [5], anxiety and depression, disruption of gut-brain interactions [6], diet [7], low grade inflammation [8], and previous evidence of gastroenteritis in a subset of cases [9]. Multiple studies have shown that the bacterial component of the microbiome (the bacteriome) is altered in IBS [10–14] and that IBS-D and IBS-M can be distinguished from bile acid malabsorption (BAM), a differential diagnosis of IBS-D [15] based upon a metabolomic signature [12].

The gut virome, which is mainly composed of bacteriophages (phages) of the dsDNA *Caudovirales* order, plays a major role in shaping the composition and interactions of the bacteriome through predation and co-evolution and by facilitating horizontal gene transfer and nutrient turnover in the bacteriome [16]. Extensive inter-individual variation as well as high temporal stability of individual gut viromes have been reported [17–20]. Despite a key role in the gut microbial ecosystem, the virome is relatively understudied in gastrointestinal diseases in contrast to the bacteriome, although recent reports suggest disturbances of the gut virome in inflammatory bowel disease (IBD) [17,21–24] and in one small study of IBS [25]. The aims of the present study were, therefore, to investigate if the gut virome is

altered in patients with IBS and to determine if any alterations distinguished the symptom-based clinical subtypes of IBS.

4.3 Methods

4.3.1 Study participant recruitment

Patients with IBS (Rome IV criteria) and control subjects aged 16–70 years and of the same ethnicity and geographic region were recruited to the study. Clinical subtyping of 55 patients with IBS [26] included: 21 with constipation (IBS-C), 17 mixed type (IBS-M), and 17 with diarrhoea (IBS-D). Exclusion criteria included the use of antibiotics within 6 weeks prior to study enrolment, other chronic illnesses including gastrointestinal diseases, severe psychiatric disease, abdominal surgery other than hernia repair or appendectomy. The inclusion/exclusion criteria for the control population were the same as for the IBS population with the exception of having to fulfil the Rome IV criteria for IBS. Gastrointestinal (GI) symptom history, psychological symptoms, diet, medical history, and medication use were collected on all participants (IBS and controls) and using the Bristol Stool Score, Hospital Anxiety and Depression Scale (HADS) [27], and Food Frequency Questionnaire (FFQ) [28]. IBS-D and IBS-M patients reporting diarrhoea as well as a subset of consenting control subjects were assessed for bile acid malabsorption by SeHCAT, a radiolabelled synthetic bile acid which is used to clinical diagnosis of BAM. Ethical approval for the study was granted by the Cork Research Ethics Committee before commencing the study and all participants provided written informed consent. Fresh faecal samples were collected from all participants, and these were stored at −80°C within a few hours of the collection until processed for

metagenomic sequencing of dsDNA virus-like particles (VLPs). Characteristics of the study population are presented in Table 1.

	Control (n = 51)	IBS (n = 55)
Age range, years (mean)	20-64 (45)	18-66 (40)
Sex (male/female)	12/39	14/41
BMI Class, n (%)		
Normal	21 (41)	24 (44)
Obese Class I	10 (20)	7 (13)
Obese Class II	3 (6)	4 (7)
Obese Class III	1 (2)	3 (5)
Overweight	15 (29)	16 (29)
Underweight	1 (2)	1 (2)
HADS: Anxiety, n (%)		
Normal (0-10)	46 (90)	43 (78)
Abnormal (11-21)	5 (10)	12 (22)
HADS: Depression, n (%)		
Normal (0-10)	51 (100)	48 (87)
Abnormal (11-21)	0 (0)	7 (13)
Bristol Stool Score, n (%)		
Normal	43 (84)	15 (27)
Constipated	7 (14)	15 (27)
Diarrhoea	1 (2)	25 (46)
IBS subtype, n (%)		
IBS-C		21 (38)
IBS-D	N/A	17 (31)
IBS-M		17 (31)
SeHCAT assayed, n (%)	9 (18)	31 (56)
Dietary group (FFQ), n (%)		
Omnivore	49 (96)	52 (94)
Vegetarian	1 (2)	2 (4)
Pescatarian	1 (2)	0 (0)
Gluten-free	0 (0)	1 (2)
Drinks alcohol, n (%)		
Current	46 (90)	39 (71)
Previous	0 (0)	0 (0)
Never	5 (10)	16 (29)
Smoker, n (%)		
Current	7* (14)	7* (13)
Previous	8 (16)	14 (25)

Never	36 (70)	34 (62)

* 1 subject in each group smoked e-cigarettes; N/A not applicable

Table 1: Clinical features of control and IBS subjects

4.3.2 Extraction, Library Preparation and Shotgun Sequencing of Faecal VLP

We used the 1E protocol from Milani et al [29]. adapted from Kleiner et al [30]. to extract faecal VLP dsDNA. When compared with polyethylene glycol (PEG) and filtration methods, this protocol had the least non-viral DNA contamination and had the lowest species-specific bias. 0.5 g of stool was suspended in a sodium chloride magnesium sulphate and gelatin (SMG) buffer by vortexing and kept 1 hr on ice. The suspension was then centrifuged at 4°C for 5 min at 2500 xg. The supernatant was removed to a fresh tube and centrifuged again at 4°C for 15 min at 5000 xg. The supernatant was again transferred to a fresh tube and dithiothreitol (DTT) was added to a final concentration of 6.5 mM and incubated for 1 h at 37°C. The sample was then filtered through a 0.45 μm syringe filter and treated with 10 U DNase (Roche) for 1 hr at RT. The DNase was inactivated by heating samples for 10 min at 70°C. DNA was extracted using the Norgen Phage DNA isolation kit (Accuscience) manufacturer's instructions. This kit isolates bacteriophage DNA from the sample using spin column chromatography without the need for phenol, chloroform, or caesium chloride. The Norgen spin column binds nucleic acids depending on ionic concentrations, thus preferentially purifying DNA while removing most RNA and proteins in the flowthrough. Briefly, lysis buffer and proteinase K were added to the sample and incubated 15 min at 55°C followed by 15 min at 65°C. Isopropanol was added before transferring the sample to the spin column, centrifuging at 6000 xg for 1 min. The spin column was washed with three

rounds of wash buffer to remove any remaining impurities and the sample eluted from the column with 75 µL elution buffer. DNA concentration was quantified using Qubit High Sensitivity dsDNA kit (Life Technologies). Libraries were prepared for sequencing using the Nextera XT library prep kit (Illumina) following manufacturer's instructions. Briefly, 1 ng of DNA was enzymatically cut, and short sequences "tagged" onto resulting fragments. Subsequently, PCR adapter and index sequences were added using the short tag sequences. Samples were run on a High Sensitivity DNA Bioanalyzer 2100 chip (Agilent) to determine the average fragment size and pool in equimolar concentrations. Tagged libraries were sent for sequencing to the commercial supplier (GATC Biotech AG, Konstanz, Germany) on a flow cell of a HiSeq 2500 sequencer, producing an average of 7,295,373 ± 1,643,253 250 bp paired-end sequences per sample.

4.3.3 Sequence Data Pre-Processing

4.3.3.1 Read processing.

Read processing and subsequent bioinformatics processing were broadly based on a previously published report [17], which was further modified. Reads were checked using FastQC v0.11.8 and multiQC v1.7 [31]. Adapters were removed using Skewer v0.2.2 with adapter string 'CTGTCTCTTATACACATCT', minimum read length of 60 and maximum read length of 170. Low complexity reads were then removed using BBMap v38.22 BBDuk with parameters entropy = 0.5, entropywindow = 50 and entropyk = 5 to filter out reads with an average entropy less than 0.5. Reads that mapped to the human genome (ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA/000/001/405/GCA_000001405.15_GRCh38/seqs_for_alignment_pipelines.ucsc_ids/GCA_000001405.15_GRCh38_no_alt_analysis_set.fna.gz) were removed using BBSplit (BBMap v38.22)

(https://sourceforge.net/projects/bbmap/) with default parameters resulting in an average of 6,285,899 + 1,480,880 paired-end reads with median read length of 157. The level of bacterial contamination was quantified by calculating the percentage of quality filtered reads mapped to the silva-bac-16s-id90 supplied by SortMeRNA (derived from SILVA SSU Ref NR v.119 database) using SortMeRNA version 2.1b [32].

4.3.3.2 Assembly and dereplication

Quality filtered reads were assembled into contigs for each sample using metaSPAdes v3.13.0 [33] with default k-mer sizes of 21,33,55,77,99. Contigs less than 1 kb were removed and contigs from all samples were pooled and dereplicated using BBMap v38.22 dedupe.sh with minimum percent identity of 95% keeping the longest contigs where two or more contigs had above the minimum identity threshold. This resulted in 327,267 contigs from the 106 samples.

4.3.3.3 Mapping and Abundance calculation

Reads that passed QC filtering were mapped to the dereplicated contigs using Bowtie2 [34] v2.3.4.3 in end-to-end mapping mode. Sambamba v0.6.6 [35] view, sort, and index functions were used to convert the Sequence Alignment/Map (SAM) file to Binary Sequence Alignment/Map (BAM), sort the BAM file, and create the BAM index file in that order. Sambamba v0.6.6 depth with threshold set to 1 (-T parameter) was used to quantify the percentage of bases in each contig that had at least 1X coverage. The abundance of each contig was calculated as the number of reads mapped to the contig (read count column in sambamba depth results). Only counts for contigs passing the step to select viral contigs were retained.

4.3.4 Selection of putative viral contigs

Open reading frames (ORFs) and their corresponding protein sequences were predicted for the dereplicated contigs using Prodigal v2.6.3 [36] in metagenomic mode. HMMER hmmsearch v3.2.1 was used to search the predicted protein sequences against HMM databases of prokaryotic viral orthologous groups (pVOGs), virus orthologous groups (VOGDB) and Integrated Microbial Genomes/Virus (IMGVR) sequences with hits retained if they had a minimum full sequence E-value of 1e-5. A database of crAssphage protein sequences was compiled by downloading 22,240 protein sequences from Guerin et al. [37] and 2,684 protein sequences from Yutin et al [38]. producing 24,924 protein sequences. Predicted protein products from the dereplicated contigs were aligned to this database using Diamond v0.9.24 keeping only the top result for each protein that had a hit with a minimum E-value of 1e-10 and minimum identity of 80%.

To minimize any bacterial contamination, the dereplicated contigs were filtered to retain only those contigs that fulfilled one or more of the following criteria: a) be circular as predicted using 'find_circular.py' from VRCA [39] with read length of 150 bp and minimum contig length of 1 kb b) categories 1 to 6 from VirSorter run in virome decontamination mode with the diamond aligner and default RefSeqdb (parameters: – virome, – diamond, – db 1 c) greater than 3 ORFs with HMM hits on a contig and greater than 2 HMM hits per 10 kb of that contig to the pVOG, VOG or IMG/VR database (hits had minimum E-value of 1e-5) d) contig predicted as viral or 'unclassified' using Kaiju v1.6.3 with the kaiju nr_euk (42 GB) database downloaded on the 5th February 2019 from http://kaiju.binf.ku.dk/server and default parameters or e) one or more protein sequences on a contig with a hit to a crAssphage protein sequence as described earlier. These steps resulted in 51,703

contigs. Contigs that had less than 75% of their bases at 1X coverage in a sample had their counts set to zero in that sample, resulting in 51,178 contigs retained for downstream analysis.

4.3.5 Identification of Viral Clusters (VCs)

752,654 proteins from 51,703 contigs were clustered using vConTACT2 v0.9.11 [40] with the PC and VC inflation index set to 1.5 and all other parameters at default. 373,129 proteins from 9,208 genomes in Viral RefSeq release 97 were downloaded and added to the pool of protein sequences input to vConTACT2 in order to facilitate taxonomic placement of the viral clusters.

Where vConTACT2 refined discordant clusters into sub-clusters, the sub-clusters were considered as independent clusters. All clusters were filtered to keep those with Benjamini-Hochberg adjusted p-values less than 0.05 and that had at least one contig from our dataset. The abundance of each cluster was calculated by summing the count associated with each contig in the cluster for each sample. This produced 2,962 clusters containing at least one contig from our data (2,233/2,962 clusters had ≥2 members, 729 clusters with only one member) but only 33 of these clusters contained a Viral RefSeq genome. Five clusters were removed as they were composed entirely of contigs that did not pass the breadth of coverage filter, resulting in 2,957 contigs encompassing a total of 11,630 contigs. The maximum number of members in any cluster was 62 (mean of four and median of two contigs in clusters).

4.3.6 Bacterial host prediction based on CRISPR protospacers

Bacterial metagenome contigs which were co-assembled from metagenomes from 69 fecal samples from our IBS patient cohort [12] using Megahit v1.1.2 [41] with minimum contig length set to 200, were used to search for CRISPR spacers.

CRISPR spacers were predicted on contigs ≥1kb (n = 871,531) using PILER-CR, retaining only results with E-value <1e-05, resulting in 10,502 spacer sequences from 644 contigs. The spacer sequences were aligned to the viral contigs (n = 51,703) using blastn from Blast++ [42] version 2.9.0+ with mode 'blastn-short' and E-value <1e-05. 2,993 viral contigs had hits to 2,076 spacer sequences from 644 bacterial contigs. The 644 bacterial contigs were taxonomically classified using CAT v5.0.3 [43] with default parameters (containing diamond version 0.9.21 and using a CAT nr database downloaded from tbb.bio.uu.nl/bastiaan/CAT_prepare/CAT_prepare_20190719.tar.gz). 459 contigs were classified (71.27%). At the superkingdom level, 451 were classified as Bacteria, two as Archaea, one as Viral (k141_4221899, *Inoviridae* family), and five remained unclassified.

4.3.7 Statistical Analysis

Feature selection was performed using DESeq2 applied to the count data for the viral clusters which were prefiltered to keep clusters present in at least 10% of samples. Correction for multiple testing was applied using the Benjamini-Hochberg (BH) methodology. Significance across IBS/control stratification and IBS and BAM subgroups were performed using the Permutational MANOVA methodology from the R vegan library. Co-inertia analyses were carried out using co-inertia function from the R library ade4 to compare Metagenomic species, as defined by Metaphlan2 [44] with Archaea and Eukaryota removed (n = 326 species), for 103 samples, with the viral contig counts for the same samples. Alpha diversity (Shannon diversity) and richness (observed counts i.e., the number of phages identified in each subject) were calculated using raw counts for contigs and VCs using the 'estimate_richness' function in the phyloseq v1.28.0 R package. Raw

counts were scaled by taking the log10 of the count after adding a pseudocount of one in both datasets. PCA was performed on each matrix using the 'dudi.pca' command, coinertia was carried out using the 'coinertia' command and the significance of the RV value was assessed by calculating a simulated p-value using the 'randtest' function with 999 permutations.

4.3.8 Taxonomic assignment of contigs

The viral section of the UniprotKB TrEMBL database was downloaded on the 24th April 2019 from ftp://ftp.uniprot.org/pub/databases/uniprot/current_release/knowledgebase/taxonomic_divisions/uniprot_trembl_viruses.dat.gz. Annotation information was downloaded from https://www.uniprot.org/uniprot/?query=taxonomy%3A%22Viruses+%5B10239%5D%22+reviewed%3Ano using search term 'taxonomy:"Viruses [10239]" AND reviewed: no' and adding 'Organism ID' as a column to the table, which resulted in 3,878,757 entries. The ete3 toolkit was used in a python script to extract the official taxonomic ranks associated with each organism ID (super kingdom, kingdom, phylum, class, order, family, genus, species) and output a file with each accession number and taxonomic information. A separate script was used to extract all protein sequences from the uniprot_trembl_viruses.dat file and place them into a fasta file with accession numbers as sequence headers. Diamond v0.9.24 blastp [45] was used to align the protein sequences associated with viral contigs to the protein sequences in the fasta file with parameters – max-target-seqs 1 – evalue 1e-05 and – sensitive. The diamond results were parsed to assign taxonomy to the query proteins using the file of accession numbers and taxonomic information.

Taxonomy was assigned to contigs using a voting based lowest common ancestor approach. For contigs with at least one ORF per 10 kb of length, if all protein sequences with hits to the UniprotKB TrEMBL database on that contig agreed at the species level then the contig was assigned that taxonomy. If there was a disagreement at species level but the genus level was the same, then the most prevalent species was assigned. If there was a tie in the number of disagreements at species level e.g. two protein sequences were assigned to one species and the other two to another, then only the genus level classification was kept. The same procedure was followed to assign contigs back to order level in the case of disagreements at downstream levels. Note that species-level assignments may not be reliable as some families have very few species genomes sequenced, potentially leading to spurious species assignments.

4.3.9 Taxonomic assignment of viral clusters

As only 33/2,957 clusters contained a Viral RefSeq genome, taxonomy was assigned to viral clusters at family level using the family level taxonomy of the contigs that formed the clusters.

4.3.10 Temperate phage prediction and differential abundance

Viral contigs were assigned as belonging to putative temperate phages if they had a gene with homology to one of 28 integrase or site-specific recombinase pVOGs (E-value <1E-05, pVOGs from the list used by Clooney et al. [17]). Viral clusters were considered as potentially temperate if they contained at least one putative temperate contig. Differential abundance of temperate clusters was calculated using the Wilcoxon test on temperate cluster counts for IBS and Controls. Clusters with Benjamini Hochberg adjusted p-value <0.05 were considered differentially abundant.

4.3.11 Terminase annotation

Terminases lack bacterial homologs and so we annotated contigs that contained genes with homology to a terminase gene from a list of 72 pVOGs produced by filtering for the word 'terminase' in the 'Protein Annotations' column of http://dmk-brain.ecn.uiowa.edu/pVOGs/VOGbigtable.html.

4.4 Results

4.4.1 Taxonomy of known and unknown phage sequences in IBS

The clinical features of the study participants which included 55 patients with IBS (based on Rome IV criteria) and 51 non-IBS controls are outlined in Table 1. We isolated and performed shotgun sequencing of Virus Like Particle (VLP) dsDNA from purified fecal samples. Reads were assembled into non-redundant contigs, filtered to retain those with viral signal (Supplementary Figures 1 & 2) as described in Methods and clustered into 2,962 viral clusters (VCs), analogous to genus/sub-family taxonomic level, based on gene sharing (Bin Jang et al., 2019) with 2,233 VCs composed of more than 2 contigs (75.4%). Of note, each methodology used to identify putative viral contigs yielded many contigs unique to that approach (Supplementary Figure 1). VCs composed of contigs that had less than 75% breadth of mapped read coverage (see Methods) were removed, retaining 2,957 VCs. Only 33 clusters contained a reference genome from the Viral RefSeq database (version 97), highlighting the amount of 'viral dark matter' present in the gut virome. Thus, putative family level taxonomy was assigned based on contig level taxonomy using the UniprotKB TrEMBL database as described in Methods and putative bacterial hosts were predicted at species level by comparing CRISPR spacers on bacterial metagenomic contigs from 69 patients in our IBS cohort [12] which had bacterial metagenomic data available, to protospacers on viral contigs in VCs, although

many VCs could not be assigned bacterial host(s) (Figure 1; Supplementary Tables 1 & 2). Using the above methodology, 36% of clusters (Supplementary Table 3) and 46.9% of contigs (Supplementary Table 4) remained unclassified at family level across all samples. At family level, 46.5% of cluster and 43.7% of contig abundance were accounted for by clusters and contigs that could not be classified (Supplementary Tables 3 & 4, Supplementary Figure 3). Unclassified VCs accounted for 42.2% and 45.2% of the family level abundance in controls and IBS respectively (Supplementary Table 5).

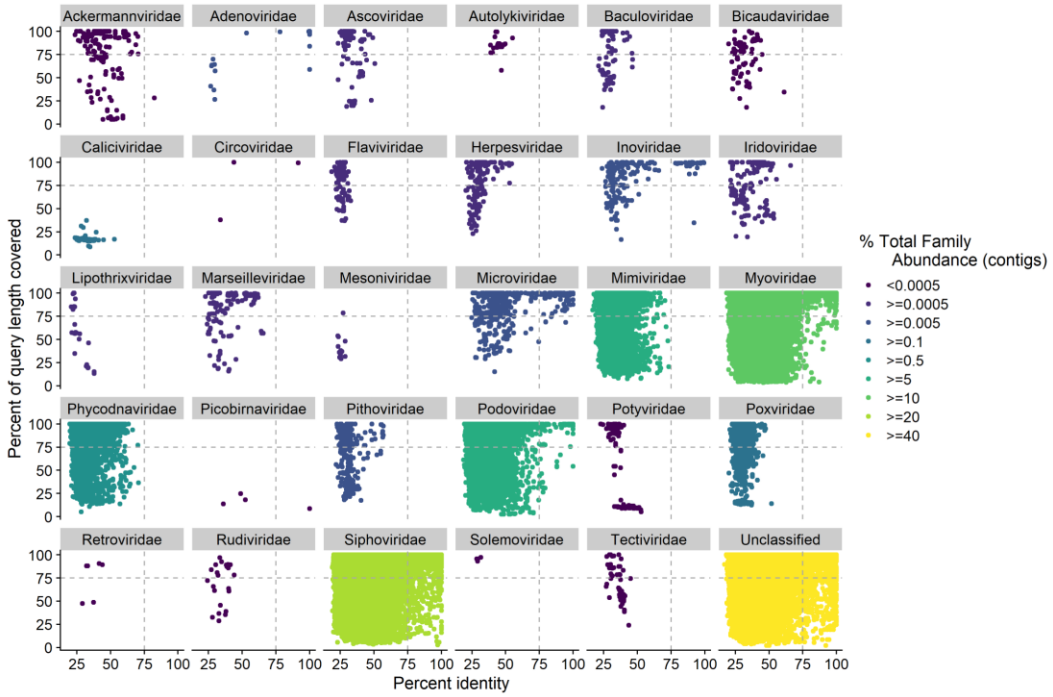


Figure 1: Confidence in taxonomic classifications, as measured by percentage query coverage and percent identity for protein sequences, predicted from the pooled contigs from IBS (n = 55) and control samples (n = 51), (n = 51,703 contigs). Each protein sequence was taxonomically classified using its top Diamond hit to the viral component of the UniProtKB TrEMBL database. Where the percentage of the query length covered was >100 due to gaps in the alignment, the coverage value

was set to a value of 100. Family abundance was calculated by summing the abundance of all contigs assigned to that family and percentage family abundance calculated as the abundance of a family divided by the abundance of all families. Dashed lines indicate 75% identity and query coverage.

Of classified clusters, the most abundant families were those of the tailed bacteriophage viruses *Siphoviridae* (27.7%), *Myoviridae* (13%), and *Podoviridae* (8.9%) respectively, which are all in the *Caudovirales* order, in agreement with previous studies of the gut virome in healthy and disease cohorts (Supplementary Table 3) [17–21,24,46,47]. The 4th most abundant family (6%) was the *Mimiviridae* family, whose members infect amoebae and protists, and whose member *Mimivirus* has been previously noted as a potentially dubious taxonomic classification in a study of the gut virome in ulcerative colitis [24,48]. This led us to examine the confidence of the taxonomic classifications in more detail. Examination of the percentage identity and percentage coverage for each protein sequence aligned to the hit protein used for taxonomic assignment (percentage query cover) revealed a wide distribution of both scores for each family indicating that many of the classifications (Figure 1 and Supplementary Figure 4) were based on low scores. While families in the *Caudovirales* order had high-scoring hits present, scores for most other families, including *Mimiviridae* were low indicating that the family-level classification was not reliable. Except for *Mimiviridae*, all other families with low scoring hits had very low abundances (<0.2% of total abundance at VC level and <0.6% at contig level; Supplementary Tables 3 & 4, Figure 1). *Ascoviridae*, *Iridoviridae*, *Marseilleviridae*, *Pithoviridae*, and Poxviridae, all families of nucleocytoplasmic large dsDNA viruses (NCLDVs; proposed order *Megavirales* [49]) of eukaryotes had mainly low scoring hits (Figure 1) and were present in very

low abundances (Figure 1; Supplementary Table 1). Furthermore, *Pithoviridae* viruses have a diameter of 500 nm, which is larger than the 450 nm filter used to remove non-viral material in this study (Methods), and so would not be expected in our sequence data [50]. Some *Mimiviridae* genomes are less than 450 nm and so could legitimately be present in our data. However, taxonomic assignments to families without the presence of high scoring hits or with low abundance are likely not as reliable as those for *Caudovirales*.

4.4.2 The gut virome differs between IBS and controls

Alpha diversity was significantly lower in the IBS faecal virome compared with controls as measured with VCs (Figure 2a; Wilcoxon $p < .05$). Among the IBS clinical subtypes, the viromes of IBS-D and IBS-M were significantly less diverse than controls, based upon both observed and Shannon diversity metrics, and the IBS-D virome was significantly less diverse than IBS-C for observed VC counts (Supplementary Figure 5a; Wilcoxon $p < .05$).

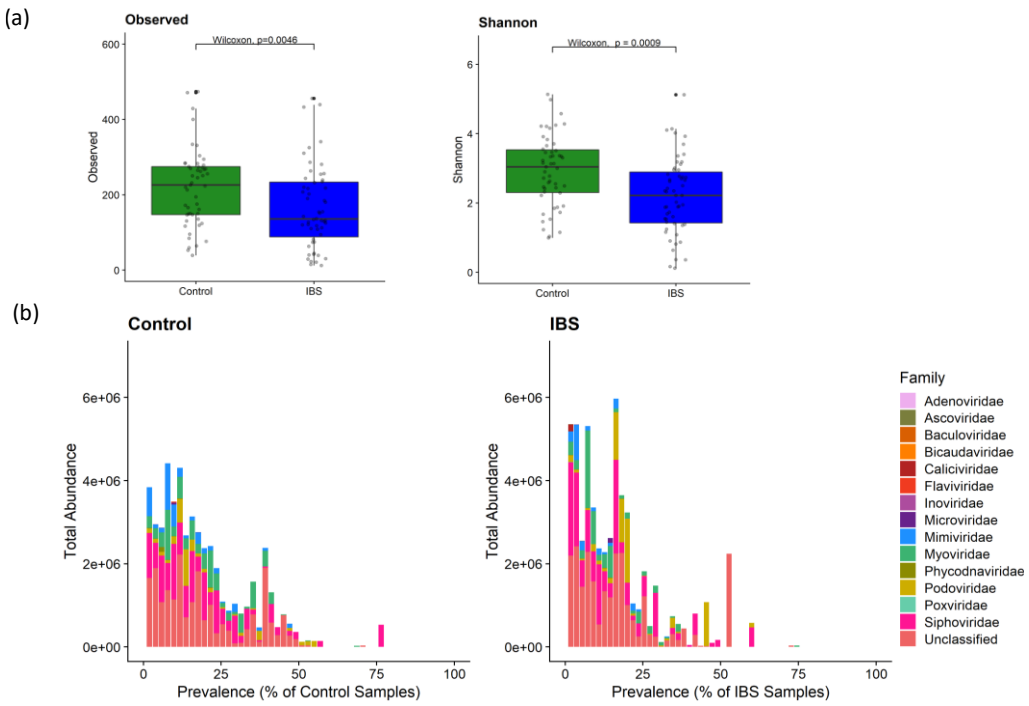


Figure 2: Difference in viral diversity between IBS and Controls. A) Boxplots of richness (observed) and alpha diversity (Shannon) estimates for IBS and control samples (Control: n = 51; IBS: n = 55 (IBS-C: n = 21; IBS-D: n = 17; IBS-M: n = 17) computed from viral cluster counts, showing significant differences (Wilcoxon rank-sum test $p < .05$) between IBS and controls. B). Barplots showing the total abundance of each viral family in control and IBS populations by prevalence in each population)

Significantly lower alpha diversity was also observed at contig level (Supplementary Figure 6) in IBS and the IBS subtypes compared to controls, showing that the grouping of contigs into VCs had not masked any differences occurring at species/strain level. Visualization of the total abundance of VCs compared with the prevalence of those VCs in IBS and controls (Figure 2b) showed that VCs with lower prevalence contribute more to the total abundance in the IBS faecal phageome while there is a larger spread of the abundance values at each prevalence level in controls.

To determine if a core virome was present in IBS or control subjects, where the core virome was operationally defined as VCs present in at least 50% of individuals in a group, the prevalence of VCs in each group and across all samples was examined (Supplementary Tables 6 & 7). We identified 5 and 12 core VCs in IBS and controls, respectively. Of these, four of the five VCs that were core VCs in IBS were also core VCs in the controls and six VCs were identified as core when considering all 106 subjects (Supplementary Table 6). The highest prevalence was attributed to two VCs, both present in 76 samples. One was taxonomically assigned as *Myoviridae* and was present in 41 IBS and 35 control samples, while the other was unclassified and found in 40 IBS and 36 control subjects. All core VCs either

had taxonomic assignments to families in the *Caudovirales* order or were unclassified but together accounted for only 6% of the total VC abundance. We could not identify a putative bacterial host for most core VCs from our analysis of CRISPR spacers with the exception of VC_1982_0, a core VC in controls only, which had some contigs with hits to *Lachnospiraceae* and VC_513_0, a core VC in all samples, which had some contigs with hits to *Bacteroides* (Supplementary Table 6).

Additional examination of all pairwise distances, as measured using pairwise Bray-Curtis dissimilarity of VC abundance between samples, for samples in the IBS group, and separately, for all samples in the control group, showed that the IBS group had a greater degree of heterogeneity as defined by larger distances between samples and this difference was statistically significant between IBS and control groups (Supplementary Figure 7; Wilcoxon rank-sum test $p < .05$). Taken together, the lower amount of core VCs and larger distances between IBS samples, demonstrate that the IBS gut virome is more individual-specific than that of controls.

Beta diversity, as measured using a PCoA of Bray-Curtis dissimilarity of VC abundance, was also significantly different between IBS and controls (PMANOVA p-value = 0.001), with the only significant split occurring at the second eigenvector (Students t-test, p-value = 0.0004) (Figure 3a). Testing beta diversity based on VCs instead of contigs accounted for more of the variance (10.7% of the total variance for VCs, see Figure 3, versus 6.4% for contigs, see Supplementary Figure 8, was accounted by the first two axes) and improved the separation between IBS and controls (Supplementary Figure 8a), demonstrating that VCs carry a stronger biological signal; however, stratification of IBS based on clinical subtypes was not

significant at VC or contig level (PMANOVA p-value = 0.1) (Supplementary Figure 5b and Supplementary Figure 8b).

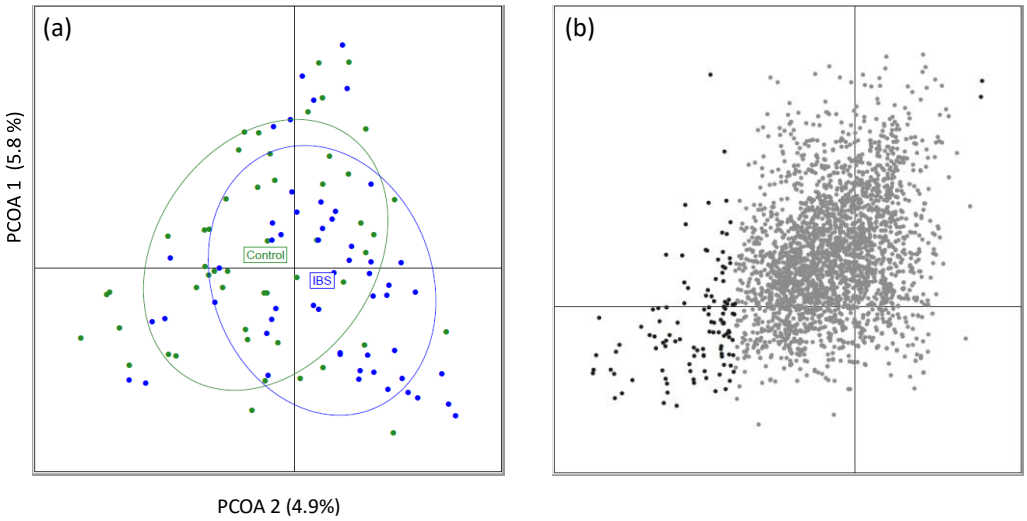


Figure 3: Difference in viral clusters between IBS and Controls. A) PCoA generated using the Bray-Curtis dissimilarity of viral cluster abundances (log-of abundance with pseudo count) between IBS and control samples showing a significant split between IBS and Control subjects (R2 = 0.0145, PMANOVA p-value = 0.001). The significant split was associated with the second eigenvector only (t.test p-value = 0.0004). PCoA1 is on the vertical axis, while PCoA2 is on the horizontal axis. B) Visualization of viral clusters using Canonical Correlation analysis with the Spearman methodology. Clusters significantly associated with the second eigenvector are highlighted as dark grey (n = 116). Of the 116 significant clusters, 114 are negatively associated with the second eigenvector showing a reduction in abundance and/or prevalence of these viral clusters across the PCoA2 axis which is associated with the IBS vs Control separation.

Feature selection analysis using the DeSeq2 methodology on viral cluster counts, identified 10 VCs that were significantly differentially abundant between IBS and non-IBS controls, seven of which showed reduced abundance in the IBS subjects (Supplementary Table 8). Of those with decreased abundance, three were

unclassified at family level, one was classified as *Mimiviridae* and three were classified in the *Caudovirales* order (two as *Siphoviridae* and one as *Podoviridae*). Of those with increased abundance, one was classified as *Mimiviridae*, one as *Podoviridae*, and one as *Siphoviridae*. Most differentially abundant VCs were present in very few samples in either IBS or Controls (median of four samples in IBS and nine samples in controls), except for VC_1982_0. This VC, which was classified as *Podoviridae*, displayed increased abundance in IBS (log2FC = 3.8) and was a core VC in control samples (present in 53% of control samples and 46% of IBS samples). The DESeq2 methodology reported large log2 fold changes between groups (mean and standard deviation of absolute log2FC of 20.6 ± 9.8) but the VCs identified were often at a low prevalence in each group. A second feature-level analysis was performed that correlated VC abundance against the secondary axis of the VCs (Supplementary Table 9) that were significantly associated with the IBS-to-control split after correction for multiple testing (Canonical Correlation analysis with Spearman methodology, Benjamini-Hochberg adjusted p-value <0.05). Of these, 114 showed decreased abundance in IBS. The findings from the two feature selection methodologies applied reflect the reduced abundance and loss of viral features in IBS as would be predicted by the overall decrease in gut virome alpha diversity in IBS compared with controls.

We used CRISPR protospacers to predict the bacterial hosts for the phages represented by the contigs in the ten VCs that were differentially abundant between IBS and control groups, and in the 116 VCs associated with the IBS to control split as defined above. CRISPR spacers prediction resulted in four phage contigs in one VC (VC_1832_0) with hits to spacers on two bacterial contigs and two phage contigs in another VC (VC_1982_0) with hits to two different bacterial contigs.

VC_1832_0, which was unclassified at family level, had a significantly lower abundance in IBS, whereas VC_1982_0, classified as *Podoviridae* and a core VC in controls, had significantly higher abundance in IBS (Figure 4). Of the four contigs in VC_1832_0 contigs, three unclassified phage contigs had matches to one unclassified bacterium and one unclassified phage contig had a predicted *Bacteroides* host. In VC_1982_0, one contig had a predicted *Lachnospiraceae* host and the other had a hit to an unclassified bacterium. Both *Lachnospiraceae* and *Bacteroides* have been reported as being present at altered abundance in IBS [51,52]. The remaining eight differentially abundant VCs did not yield any hits to spacers on bacterial contigs.

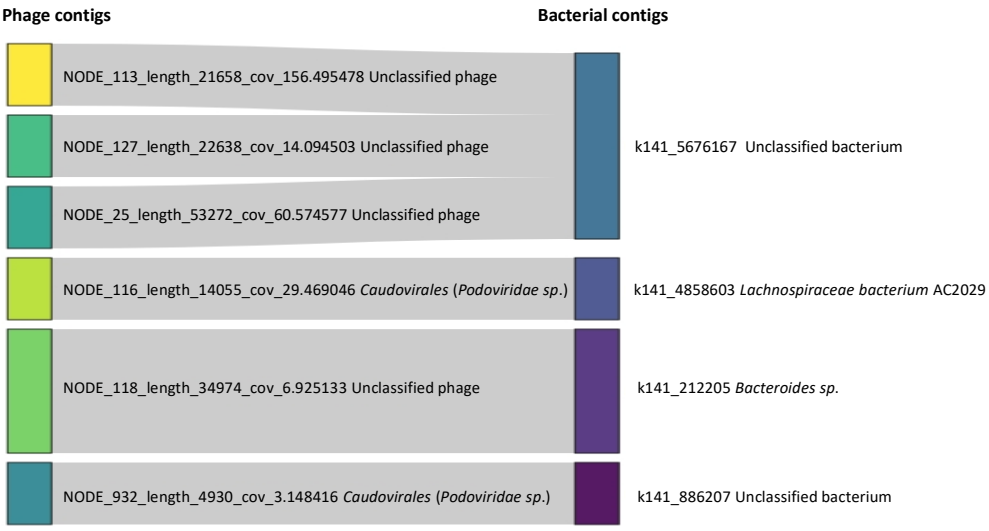


Figure 4: Sankey plot of putative phage contigs and their putative bacterial hosts based on CRISPR spacer sequences in VC_1832_0 and VC_1982_0, which were two of ten differentially abundant VCs between IBS and controls. To produce this plot, phage contigs in the 10 VCs were restricted to those that had a match based on CRISPR spacer sequences to a bacterial contig, resulting in 6 contigs, 4 from viral cluster, VC_1832_0 and 2 from VC_1982_0 (phage contigs beginning with NODE_116 and NODE_932). VC_1832_0 had significantly decreased abundance,

and VC_1982_0 had significantly increased abundance in IBS compared with control samples, as reported by DeSeq2. For phage contigs, the putative species is indicated in brackets beside the phage family name.

Analysis of bacterial hosts for the 116 VCs associated with the IBS to control split yielded 12 VCs containing contigs with hits to *Prevotella* (2 VCs), *Faecalibacterium* (4 VCs), *Blautia* and *Ruminococcus* (both associated with 1 VC), *Firmicutes* (3 VCs), *Clostridia* (I VC), and *Anaerobutyricum* (1 VC) species in a subset of VCs with decreased abundance in IBS; however, we could not predict bacterial hosts for most contigs in these VCs (Supplementary Table 9). The two VCs with hits to *Prevotella* were unclassified at family level but phage families *Myoviridae* and *Siphoviridae* were associated with the other bacterial hosts.

In summary, the reduced core virome, low count of gut virome species, and differential abundance observed in IBS signals a distinctive and less diverse virome in IBS compared with non-IBS controls.

Despite these differences, we found that the gut virome could not be used as predictive of IBS based on an XGBOOST machine learning pipeline, implemented as described previously for IBD [17] with five-fold cross-validations. Predictive models showed no predictive power when combined with gut virome cluster data with AUCs close to 0.5.

4.4.3 The gut virome does not co-vary with the bacteriome or BAM status

Bile acid malabsorption (BAM) causes diarrhoea which is clinically difficult to distinguish from IBS-D [15]. We recently showed that the faecal metabolome, but not the microbiome, could distinguish subjects with BAM from those with IBS [12]. We first used co-inertia analysis to test for co-variance between the gut virome VCs and

bacteriome for 326 bacterial species in 103 samples for which we had bacteriome information and found only limited global similarity between the gut bacteriome and virome (RV = 0.45; data not shown). Analysis using contigs instead of VCs also yielded a low RV value of 0.43 indicating that the co-variance between the VCs and the gut bacteriome of these samples was limited.

To determine if BAM status interacted with the gut virome, the BAM status of 28 patients with IBS (4 borderlines, 5 mild, 4 moderate, 12 normal, and 3 severe BAM cases) was compared with VC abundance in the faecal virome data. No significant relationship was identified (PMANOVA on Spearman distance, p-value = 0.16), although the low number of samples used was a limitation.

4.4.4 No evidence of conversion from lytic to lysogenic life cycle in IBS

A gut virome study of patients with IBD found evidence that the virulent phage core in healthy individuals is replaced with temperate phage in IBD patients [17]. To determine if this is the case in IBS, we examined 943 VCs (31.9% of VCs) that we identified as putative temperate VCs (see Methods) for differential abundance in IBS and controls. We did not find any VCs which were differentially abundant (Wilcoxon signed-rank test, Benjamini-Hochberg adjusted p-value <0.05) indicating that there is no switch in core phage lifecycles in IBS.

4.5 Discussion

Our data show that there is a limited core virome in patients with IBS and controls. Those with IBS had reduced gut virome diversity and more individual-specific viromes compared with non-IBS controls, but there were no significant differences in diversity of the gut virome among the IBS clinical subtypes. In a recent study conducted on 4198 individuals of a Japanese cohort, the possible causes of inter-

individual variation of gut virome were suggested [53]. Although disease and medication showed the strongest effects, other factors which significantly influenced the virome were age, gender, lifestyle, and diet.

The use of VCs helped attenuate the strong inter-individual variation observed when using strain level contigs as documented for IBD [17]. It detected more sharing than seen at contig level and improved stratification between IBS and controls, especially in beta diversity analysis. Of classified VCs, families in the *Caudovirales* order (*Myoviridae*, *Podoviridae*, *Siphoviridae*) predominated in both IBS and controls, with a minimal core gut virome, as reported in other gut virome studies for diseased and healthy populations [17–22,24,46,47]. The IBS gut virome is distinct from that of controls, but there is only one core VC (unclassified taxonomy) identified in IBS that was not also found as a core VC in controls.

To our knowledge, there is a single previous report on the gut virome in IBS [25] which investigated the faecal virome of 25 patients with IBS and 17 controls using the subset of reads that could be classified to known viral families by alignment to the Viral RefSeq database. The larger study presented here examines both known and unknown viral sequences. We corroborated the reduction in phage alpha diversity in IBS and lack of robust differentiation among IBS subtypes reported by Ansari et al [25]. However, we did not identify significant differences in families of eukaryotic viruses of the order *Megavirales* between patients with IBS and controls, as reported by Ansari et al [25]. with the exception of two VCs assigned to the *Mimiviridae* family, one of which was more abundant in IBS and one with less abundance, but we noted that the protein sequence hits used to classify this family had low support. All other members of the *Megavirales* order in our data had very low abundances and alignment scores and did not display significantly different

abundances between IBS and controls. Due to the challenges of taxonomic assignment, we cannot rule out the presence of eukaryotic viruses in the gut virome as these viruses have been identified in the gut virome of both healthy subjects and those with gastroenteritis previously and have been cultured from human stools [54,55]. However, they likely originate from amoebas in drinking water and a definitive role for these viruses in pathogenesis is lacking. Difficulties with taxonomic assignment of eukaryotic viruses in the gut virome have been previously reported [48]; both viral genome assembly and taxonomic assignment of viruses are challenging due to the lack of a universal marker gene, high numbers of hypervariable and repeat regions [56,57], rapid mutation rates,[57] horizontal transfer between viruses and bacteria and incomplete databases [56,58–61]. Validated methods such as vConTACT2 [40] can cluster contigs and assign taxonomy; in our data; however, many clusters remained unclassified with this approach highlighting the large amount of viral dark matter in the gut virome.

The core gut virome, dominated by lytic phage, seems to be replaced by temperate phage in IBD [17]. To determine if a similar phenomenon is involved in IBS, we examined the differential abundance of all putative temperate VCs between IBS and controls, but we did not find any evidence of a difference and conclude that replacement of the predominant phage lifecycles is not a consistent observation in IBS.

Since residual bacterial sequences may be present in VLP datasets [30], we filtered our contigs to retain only those with viral signal, which we performed by combining the results of a comprehensive set of methodologies used in other gut virome studies [17,20,46,62]. We observed low concordance across the approaches highlighting the utility of combining multiple methods. A potential limitation of our study is that

novel viruses containing features that are substantially different from those catalogued in viral databases could be missed, and that bacterial contigs with loci homologous to viral features could have been erroneously included. We also attempted to identify putative bacterial hosts for phage contigs in all VCs by identifying sequence matches on phage contigs to CRISPR spacers on bacterial contigs. For the 10 VCs that were differentially abundant between IBS and controls, we could only identify hosts for contigs in two of the ten VCs, and similarly, we could only identify bacterial hosts for 12 VCs of the 116 VCs associated with the IBS to Control split. This could potentially be improved by using contigs from more samples as we only had bacterial contig information for 66% of the virome samples.

Although we did not find global co-variance of the bacteriome and virome data, this may be due to the inherent sparseness of the data which also presented challenges for feature selection and beta diversity analysis. Additionally, we did not examine the role of ssDNA or RNA viruses in this study, although RNA viruses in the gut would be expected to be mostly of dietary (i.e., plant) based origin [23,63].

Use of a whole-virome-based approach in this study allowed us to identify changes in both the known and unknown components of the gut virome in IBS compared with controls. Further work is required to improve methods for taxonomic classification of viral sequences as well as for resolving phage genomes. With better reference databases, more formal classifications, and application of long read metagenomics to the analysis of the human gut virome using approaches such as those employed by Warwick-Dugdale and colleagues [60] would improve completeness of recovered phage genomes, enable more accurate strain-level analysis, taxonomic classification, and aid exploration of virome population

structure and host interactions to identify potential biomarkers and/or therapeutic targets for IBS.

4.6 Acknowledgements

We are grateful to the 4D pharma Cork clinical team, Caroline O'Leary, Marian O'Meara and Sinead Kelly who recruited the subjects to the study and collected the biological samples and clinical metadata. We would also like to thank gastroenterologist Dr Syed Akbar Zulquernain and Dr Elizabeth Kenny in Cork University Hospital who assisted with IBS patient recruitment. We acknowledge the input of 4D pharma Cork lab staff, Luke O'Brien, Fiona Grady, Marie McBrien and Siobhan O'Neill for sample processing for analysis. We also thank Michael Moore, Fintan Bradley, Tom Carty, Fionnuala McDonnell, Lean O'Donoghue, Pauline Beecher, Roisin Wilson, and Shane Hayes at Cork University Hospital who performed the SeHCAT assay.

4.7 References

1. Lovell RM, Ford AC. Global Prevalence of and Risk Factors for Irritable Bowel Syndrome: A Meta-analysis. Clinical Gastroenterology and Hepatology 2012;10:712-721.e4.

2. Lacy BE, Mearin F, Chang L, et al. Bowel Disorders. Gastroenterology 2016;150:1393-1407.e5.

3. Beyder A, Mazzone A, Strege PR, et al. Loss-of-Function of the Voltage-Gated Sodium Channel NaV1.5 (Channelopathies) in Patients With Irritable Bowel Syndrome. Gastroenterology 2014;146:1659–1668.

4. Ek WE, Reznichenko A, Ripke S, et al. Exploring the genetics of irritable bowel syndrome: a GWA study in the general population and replication in multinational case-control cohorts. Gut 2015;64:1774–1782.

5. Qin H-Y. Impact of psychological stress on irritable bowel syndrome. WJG 2014;20:14126.

6. Chey WD, Kurlander J, Eswaran S. Irritable Bowel Syndrome: A Clinical Review. JAMA 2015;313:949.

7. El-Salhy M, Gundersen D. Diet in irritable bowel syndrome. Nutr J 2015;14:36.

8. Collins SM. A role for the gut microbiota in IBS. Nat Rev Gastroenterol Hepatol 2014;11:497–505.

9. Thabane M, Marshall JK. Post-infectious irritable bowel syndrome. WJG 2009;15:3591.

10. Carroll IM, Ringel-Kulka T, Siddle JP, et al. Alterations in composition and diversity of the intestinal microbiota in patients with diarrhea-predominant irritable bowel syndrome: Alterations in composition and diversity of the intestinal microbiota in D-IBS. Neurogastroenterology & Motility 2012;24:521-e248.

11. Casén C, Vebø HC, Sekelja M, et al. Deviations in human gut microbiota: a novel diagnostic test for determining dysbiosis in patients with IBS or IBD. Aliment Pharmacol Ther 2015;42:71–83.

12. Jeffery IB, Das A, O'Herlihy E, et al. Differences in Fecal Microbiomes and Metabolomes of People With vs Without Irritable Bowel Syndrome and Bile Acid Malabsorption. Gastroenterology 2020;158:1016-1028.e8.

13. Jeffery IB, O'Toole PW, Öhman L, et al. An irritable bowel syndrome subtype defined by species-specific alterations in faecal microbiota. Gut 2012;61:997–1006.

14. Rajilić-Stojanović M, Biagi E, Heilig HGHJ, et al. Global and deep molecular analysis of microbiota signatures in fecal samples from patients with irritable bowel syndrome. Gastroenterology 2011;141:1792–1801.

15. Slattery SA, Niaz O, Aziz Q, et al. Systematic review with meta-analysis: the prevalence of bile acid malabsorption in the irritable bowel syndrome with diarrhoea. Aliment Pharmacol Ther 2015;42:3–11.

16. Shkoporov AN, Hill C. Bacteriophages of the Human Gut: The "Known Unknown" of the Microbiome. Cell Host & Microbe 2019;25:195–209.

17. Clooney AG, Sutton TDS, Shkoporov AN, et al. Whole-Virome Analysis Sheds Light on Viral Dark Matter in Inflammatory Bowel Disease. Cell Host Microbe 2019;26:764-778.e5.

18. Minot S, Sinha R, Chen J, et al. The human gut virome: Inter-individual variation and dynamic response to diet. Genome Res 2011;21:1616–1625.

19. Reyes A, Haynes M, Hanson N, et al. Viruses in the faecal microbiota of monozygotic twins and their mothers. Nature 2010;466:334–338.

20. Shkoporov AN, Clooney AG, Sutton TDS, et al. The Human Gut Virome Is Highly Diverse, Stable, and Individual Specific. Cell Host & Microbe 2019;26:527-541.e5.

21. Manrique P, Bolduc B, Walk ST, et al. Healthy human gut phageome. Proc Natl Acad Sci USA 2016;113:10400–10405.

22. Norman JM, Handley SA, Baldridge MT, et al. Disease-Specific Alterations in the Enteric Virome in Inflammatory Bowel Disease. Cell 2015;160:447–460.

23. Shkoporov AN, Ryan FJ, Draper LA, et al. Reproducible protocols for metagenomic analysis of human faecal phageomes. Microbiome 2018;6:68.

24. Zuo T, Lu X-J, Zhang Y, et al. Gut mucosal virome alterations in ulcerative colitis. Gut 2019;68:1169–1179.

25. Ansari MH, Ebrahimi M, Fattahi MR, et al. Viral metagenomic analysis of fecal samples reveals an enteric virome signature in irritable bowel syndrome. BMC Microbiol 2020;20:123.

26. Longstreth GF, Thompson WG, Chey WD, et al. Functional bowel disorders. Gastroenterology 2006;130:1480–1491.

27. Zigmond AS, Snaith RP. The hospital anxiety and depression scale. Acta Psychiatr Scand 1983;67:361–370.

28. Power SE, Jeffery IB, Ross RP, et al. Food and nutrient intake of Irish community-dwelling elderly subjects: who is at nutritional risk? J Nutr Health Aging 2014;18:561–572.

29. Milani C, Casey E, Lugli GA, et al. Tracing mother-infant transmission of bacteriophages by means of a novel analytical tool for shotgun metagenomic datasets: METAnnotatorX. Microbiome 2018;6:145.

30. Kleiner M, Hooper LV, Duerkop BA. Evaluation of methods to purify virus-like particles for metagenomic sequencing of intestinal viromes. BMC Genomics 2015;16:7.

31. Ewels P, Magnusson M, Lundin S, et al. MultiQC: summarize analysis results for multiple tools and samples in a single report. Bioinformatics 2016;32:3047–3048.

32. Kopylova E, Noé L, Touzet H. SortMeRNA: fast and accurate filtering of ribosomal RNAs in metatranscriptomic data. Bioinformatics 2012;28:3211–3217.

33. Nurk S, Meleshko D, Korobeynikov A, et al. metaSPAdes: a new versatile metagenomic assembler. Genome Res 2017;27:824–834.

34. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. Nat Methods 2012;9:357–359.

35. Tarasov A, Vilella AJ, Cuppen E, et al. Sambamba: fast processing of NGS alignment formats. Bioinformatics 2015;31:2032–2034.

36. Hyatt D, Chen G-L, LoCascio PF, et al. Prodigal: prokaryotic gene recognition and translation initiation site identification. BMC Bioinformatics 2010;11:119.

37. Guerin E, Shkoporov A, Stockdale SR, et al. Biology and Taxonomy of crAss-like Bacteriophages, the Most Abundant Virus in the Human Gut. Cell Host & Microbe 2018;24:653-664.e6.

38. Yutin N, Makarova KS, Gussow AB, et al. Discovery of an expansive bacteriophage family that includes the most abundant viruses from the human gut. Nat Microbiol 2018;3:38–46.

39. Crits-Christoph A, Gelsinger DR, Ma B, et al. Functional interactions of archaea, bacteria and viruses in a hypersaline endolithic community: Halite metagenome. Environ Microbiol 2016;18:2064–2077.

40. Bin Jang H, Bolduc B, Zablocki O, et al. Taxonomic assignment of uncultivated prokaryotic virus genomes is enabled by gene-sharing networks. Nat Biotechnol 2019;37:632–639.

41. Li D, Liu C-M, Luo R, et al. MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. Bioinformatics 2015;31:1674–1676.

42. Camacho C, Coulouris G, Avagyan V, et al. BLAST+: architecture and applications. BMC Bioinformatics 2009;10:421.

43. Meijenfeldt FAB von, Arkhipova K, Cambuy DD, et al. Robust taxonomic classification of uncharted microbial sequences and bins with CAT and BAT. Genome Biol 2019;20:217.

44. Segata N, Waldron L, Ballarini A, et al. Metagenomic microbial community profiling using unique clade-specific marker genes. Nat Methods 2012;9:811–814.

45. Buchfink B, Xie C, Huson DH. Fast and sensitive protein alignment using DIAMOND. Nat Methods 2015;12:59–60.

46. Gregory AC, Zablocki O, Howell A, et al. *The human gut virome database*. Bioinformatics; 2019. Available at: http://biorxiv.org/lookup/doi/10.1101/655910 [Accessed October 11, 2022].

47. McCann A, Ryan FJ, Stockdale SR, et al. Viromes of one year old infants reveal the impact of birth mode on microbiome diversity. PeerJ 2018;6:e4694.

48. Sutton TDS, Clooney AG, Hill C. Giant oversights in the human gut virome. Gut 2020;69:1357–1358.

49. Colson P, De Lamballerie X, Yutin N, et al. "Megavirales", a proposed new order for eukaryotic nucleocytoplasmic large DNA viruses. Arch Virol 2013;158:2517–2521.

50. Legendre M, Bartoli J, Shmakova L, et al. Thirty-thousand-year-old distant relative of giant icosahedral DNA viruses with a pandoravirus morphology. Proc Natl Acad Sci USA 2014;111:4274–4279.

51. Jeffery IB, Quigley EMM, Öhman L, et al. The microbiota link to irritable bowel syndrome: an emerging story. Gut Microbes 2012;3:572–576.

52. Pittayanon R, Lau JT, Yuan Y, et al. Gut Microbiota in Patients With Irritable Bowel Syndrome—A Systematic Review. Gastroenterology 2019;157:97–108.

53. Nishijima S, Nagata N, Kiguchi Y, et al. Extensive gut virome variation and its associations with host and environmental factors in a population-level cohort. Nat Commun 2022;13:5252.

54. Colson P, Aherfi S, La Scola B. Evidence of giant viruses of amoebae in the human gut. Human Microbiome Journal 2017;5–6:14–19.

55. Colson P, Fancello L, Gimenez G, et al. Evidence of the megavirome in humans. Journal of Clinical Virology 2013;57:191–200.

56. Minot S, Grunberg S, Wu GD, et al. Hypervariable loci in the human gut virome. Proc Natl Acad Sci USA 2012;109:3962–3966.

57. Minot S, Bryson A, Chehoud C, et al. Rapid evolution of the human gut virome. Proc Natl Acad Sci USA 2013;110:12450–12455.

58. Roux S, Emerson JB, Eloe-Fadrosh EA, et al. Benchmarking viromics: an *in silico* evaluation of metagenome-enabled estimates of viral community composition and diversity. PeerJ 2017;5:e3817.

59. Sutton TDS, Clooney AG, Ryan FJ, et al. Choice of assembly software has a critical impact on virome characterisation. Microbiome 2019;7:12.

60. Warwick-Dugdale J, Solonenko N, Moore K, et al. Long-read viral metagenomics captures abundant and microdiverse viral populations and their niche-defining genomic islands. PeerJ 2019;7:e6800.

61. Zuo T, Ng SC. Authors response: giant oversights in the human gut virome. Gut 2020;69:1358.2-1358.

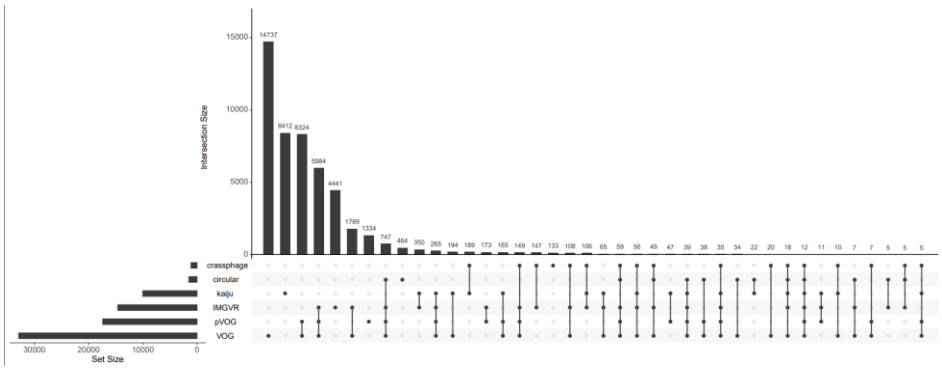
62. Draper LA, Ryan FJ, Smith MK, et al. Long-term colonisation with donor bacteriophages following successful faecal microbial transplantation. Microbiome 2018;6:220.

63. Zhang T, Breitbart M, Lee WH, et al. RNA Viral Community in Human Feces: Prevalence of Plant Pathogenic Viruses Dangl J, ed. PLoS Biol 2005;4:e3.

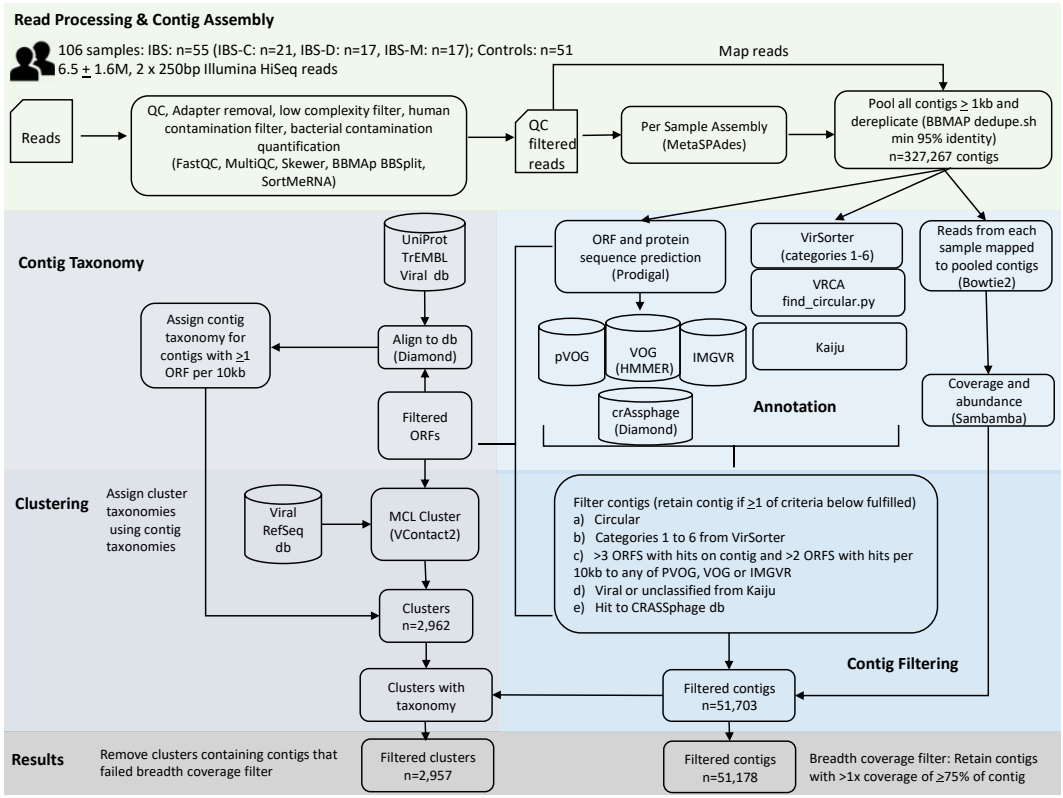
4.8 Supplementary information

4.8.1 Supplementary figures

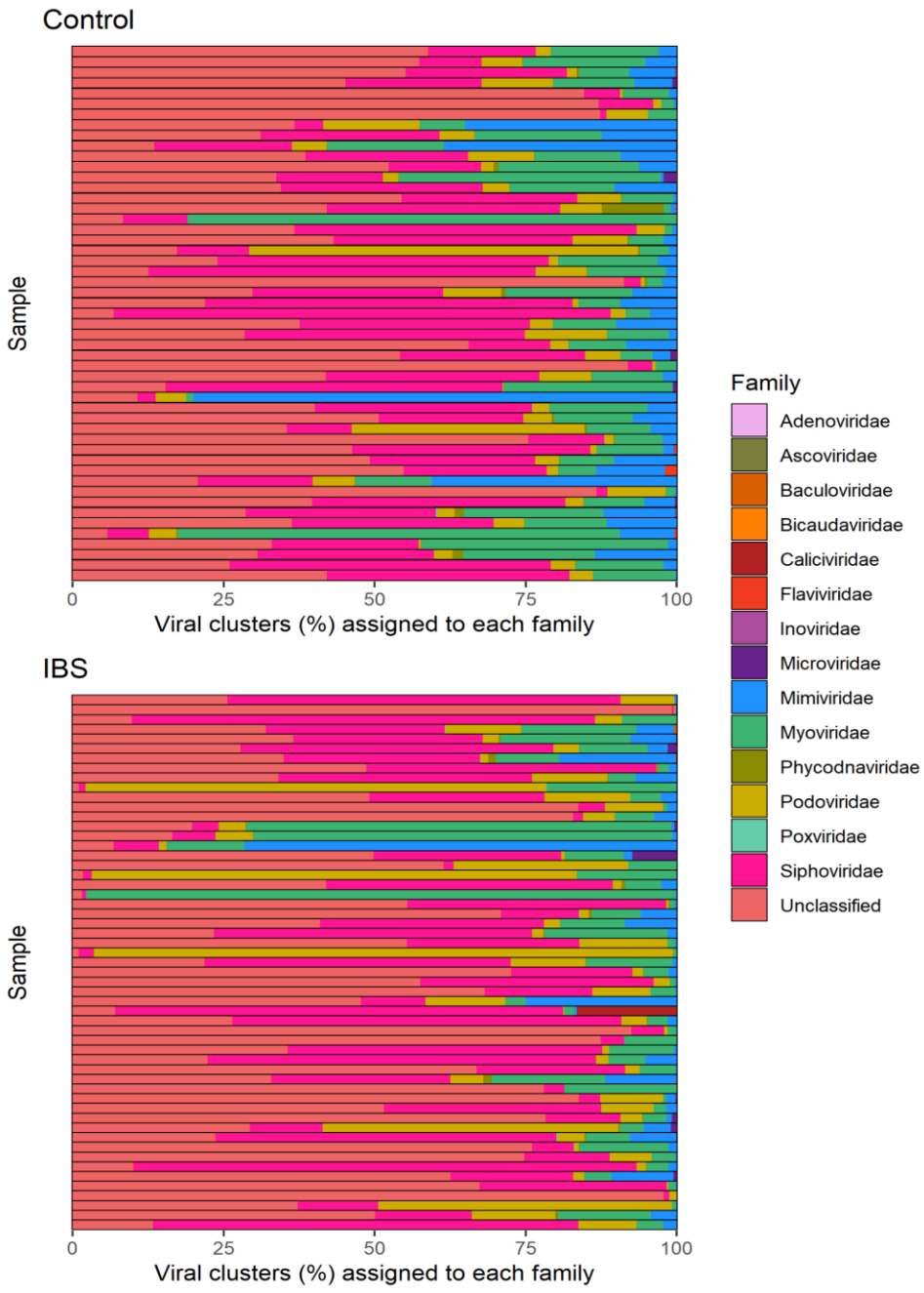
Supplementary Figure 1: Number of contigs identified as potentially viral (or viral and 'unclassified' using Kaiju) by various approaches. 51,703 contigs were identified from the union of all approaches, including VirSorter, which is not shown here. The total size of each set is represented in the left barplot (set size), the numbers of putative viral contigs identified by more than one method, or only one method in cases where there is only one filled circle associated with the method, are shown at the bottom of the plot, and the number of the same contigs identified by multiple methods (the intersection size) is given by the top barplot. The set size is the union of the numbers of contigs identified by each method (the number identified uniquely by that method in the case of one filled circle in the bottom section) added to the number of contigs that the method contributed to each intersection. For example, for VOGs, summing all values for bars that have filled circles yields a set size of 32,916 contigs, 14,737 of which were uniquely identified by the VOG approach (first bar from the left) and the rest (18,178) were identified by both the VOG approach and at least one other approach.



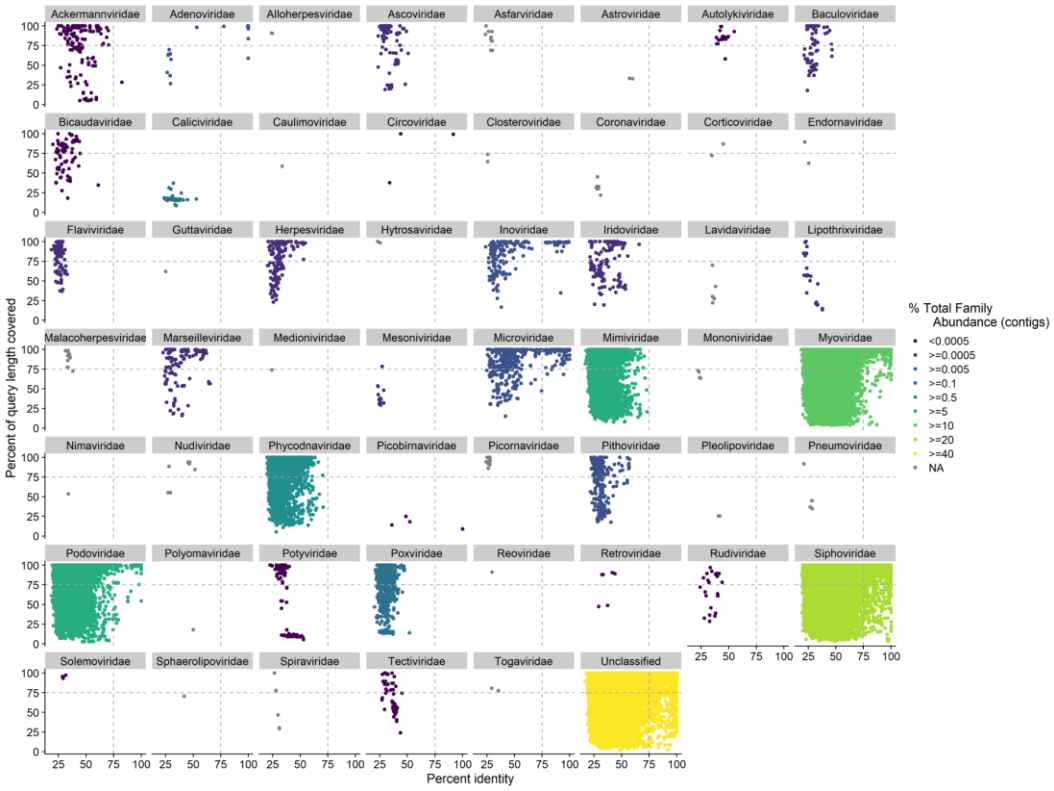
Supplementary Figure 2: Flowchart showing viral read processing, assembly, annotation and clustering steps



Supplementary Figure 3: Taxonomic composition of viral clusters in each sample in Control (n=51) and IBS (n=55) groups

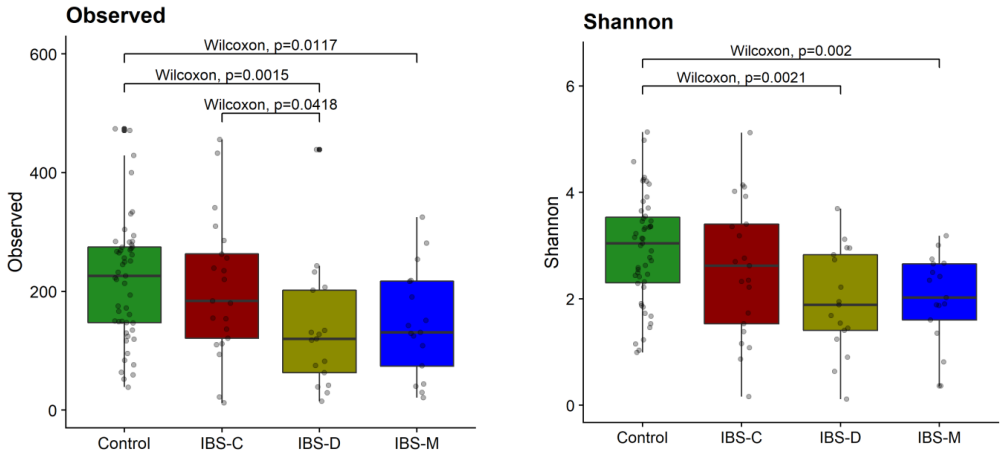


Supplementary Figure 4: Confidence in taxonomic classifications, measured in the same manner as for Figure 1 but including protein sequence hits to families that did not have enough votes on any contig to classify a contig as that family. This resulted in no abundance value at contig level for these families which is indicated as NA on the legend and by the grey dots.

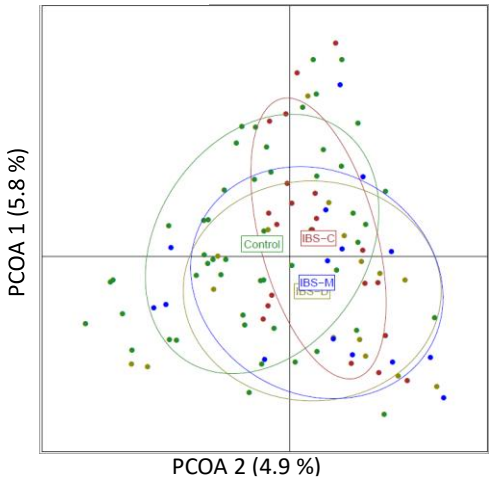


Supplementary Figure 5: Difference in the diversity of viral clusters between IBS and Controls. A) Boxplots of richness (observed) and alpha diversity (Shannon) estimates for IBS subtypes (IBS-C: n=21; IBS-D: n=17; IBS-M: n=17) and control samples (n = 51) computed from viral cluster counts. , A significant difference (Wilcoxon rank sum test $p < 0.05$) was observed between some IBS subtypes and control samples. B) PCoA as generated using the Bray-Curtis dissimilarity between IBS subtypes. Stratification of IBS based on subtypes beta-diversity does not reach significance ($R^2 = 0.0415$, PMANOVA p-value = 0.1)

(a)

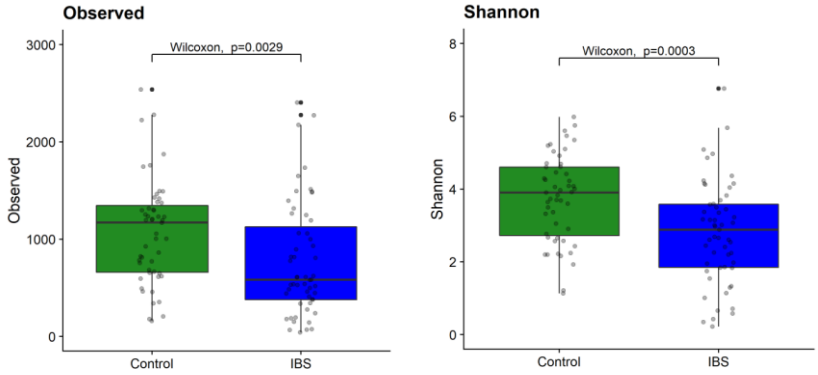


(b)

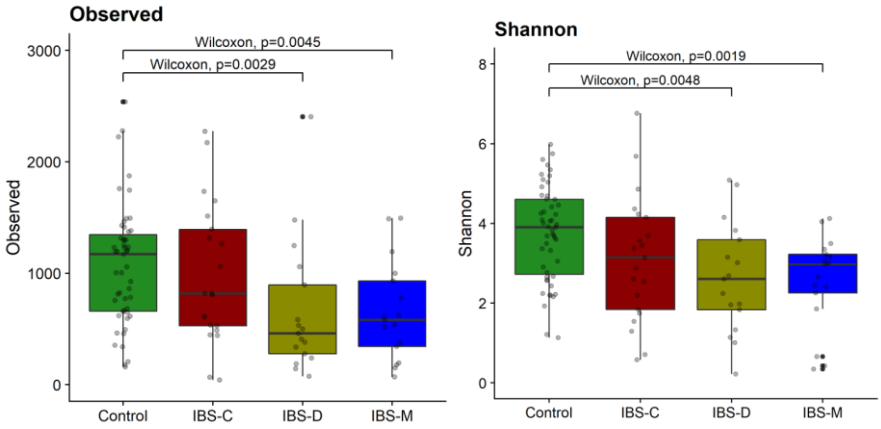


Supplementary Figure 6: Difference in the diversity of viral contigs between IBS and Control. Boxplots of richness (observed) and alpha diversity (Shannon) estimates for b) IBS and control samples computed from viral contig counts and (b) showing IBS subtypes and control samples. Significant differences (Wilcoxon rank sum test $p < 0.05$) are indicated in the plots.

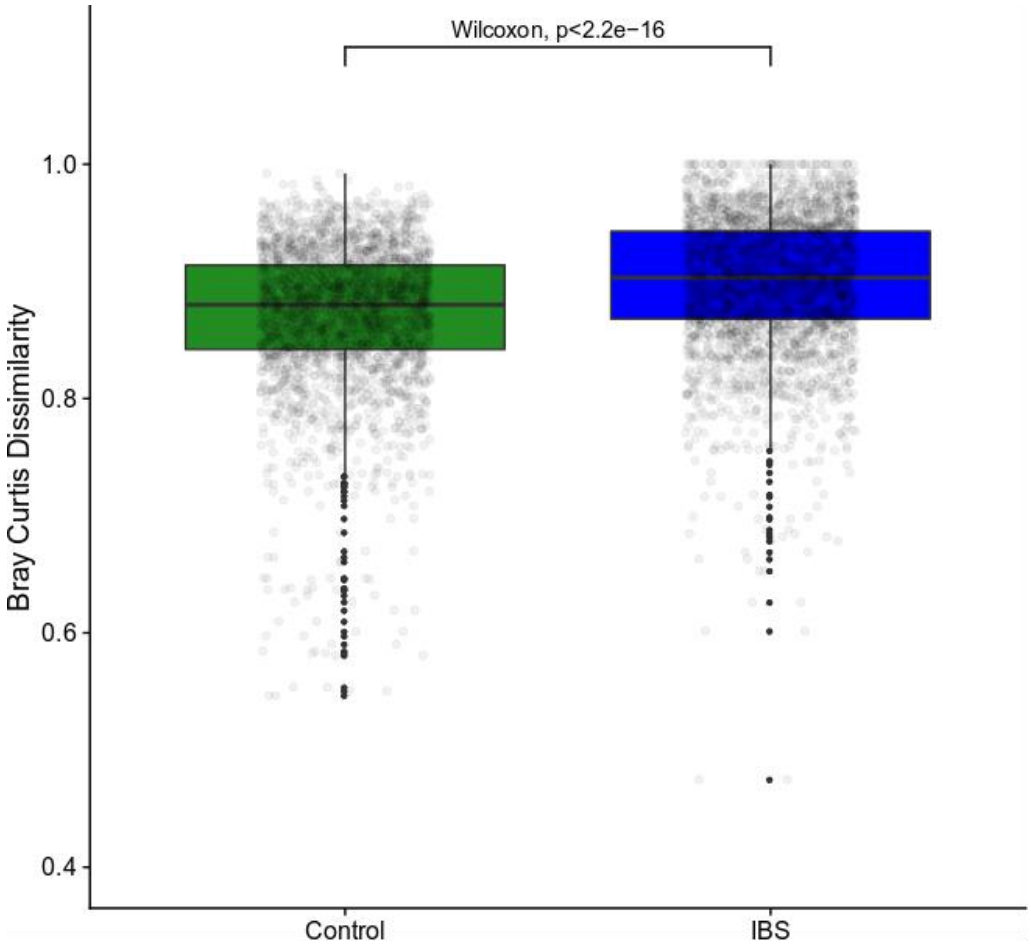
(a)



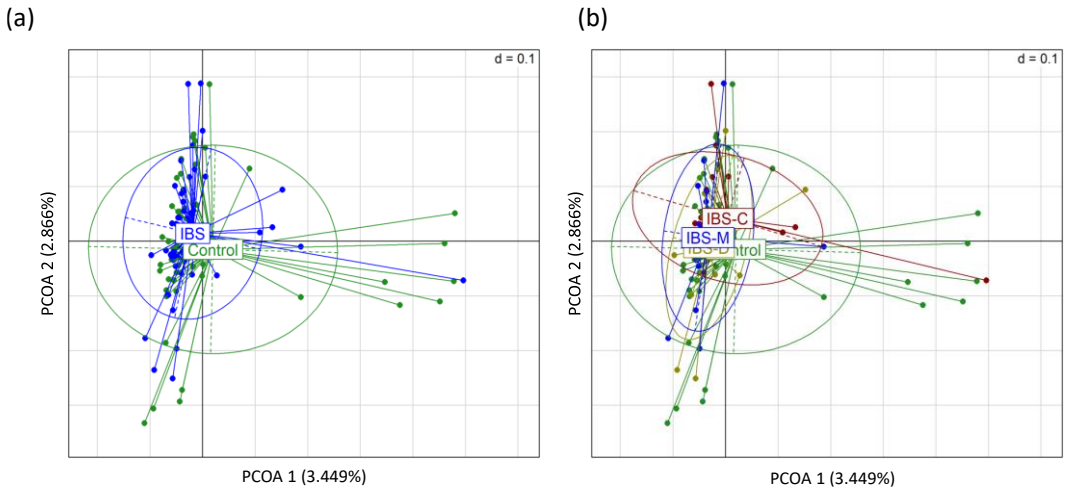
(b)



Supplementary Figure 7: Boxplots showing between sample dissimilarity for IBS and Control samples. These used Bray-Curtis dissimilarity of logged viral cluster abundances between all IBS samples and between all control samples, showing that samples in the IBS group are on average more different from each other than those in the Control group (Wilcoxon rank sum test $p < 0.05$).



Supplementary Figure 8: Significant difference in viral contig counts between IBS and Controls but not the IBS subtypes. PCoA as generated using the Bray-Curtis dissimilarity calculated from viral contig counts (log10 of count +1) showing a) significant split between IBS and Control subjects (R^2 = 0.0112, PMANOVA p-value = 0.021) and b) non-significant split between IBS subtypes (R^2 = 0.0392, PMANOVA p-value = 0.09).



4.8.2 Supplementary tables

Supplementary Table 1: Summary of viral contig characteristics including i) the cluster they are a member of, ii) the taxonomic assignment of the contig, iii) details on the taxonomic assignment and iv) details of the predicted bacterial host species for each viral contig. Viral contigs with hits to CRISPR spacers on more than one bacterial contig were assigned as targeting the bacterial species with the most votes and the additional classifications are noted in the Host Species 2, 3 and 4 columns. This is a truncated table with information of viral clusters which were considered as core (≥ 50 % prevalence) in IBS or Controls. For complete table, please refer to the published article (doi: 10.1080/19490976.2021.1887719).

		Contigs			Predicted Characteristics		
Cluster ID	**Family**	**ID**	**Length**	**Mean Count**	**Circular**	**Temperate**	**Terminase gene**
VC_1058_0	Siphoviridae	NODE_10_length_30774_cov_42.133268	30774	90.4623	FALSE	TRUE	TRUE
VC_1058_0	Siphoviridae	NODE_102_length_14891_cov_2.858843	14891	0.6321	FALSE	FALSE	TRUE
VC_1058_0	Siphoviridae	NODE_106_length_22395_cov_4.915366	22395	36.7736	FALSE	TRUE	TRUE
VC_1058_0	Siphoviridae	NODE_112_length_26892_cov_19.672340	26892	0.9434	FALSE	TRUE	TRUE
VC_1058_0	Siphoviridae	NODE_123_length_14812_cov_19.721403	14812	4.3679	FALSE	FALSE	TRUE
VC_1058_0	Siphoviridae	NODE_123_length_17864_cov_8.304250	17864	21.3868	FALSE	FALSE	TRUE
VC_1058_0	Siphoviridae	NODE_124_length_16633_cov_16.350853	16633	15.8585	FALSE	FALSE	TRUE
VC_1058_0	Siphoviridae	NODE_127_length_14468_cov_5.781056	14468	118.9057	FALSE	FALSE	TRUE
VC_1058_0	Siphoviridae	NODE_136_length_10280_cov_4.740890	10280	18.566	FALSE	FALSE	TRUE
VC_1058_0	Siphoviridae	NODE_145_length_17049_cov_4.218407	17049	20.3302	FALSE	FALSE	TRUE
VC_1058_0	Siphoviridae	NODE_152_length_12137_cov_3.597525	12137	15.0189	FALSE	FALSE	TRUE
VC_1058_0	Siphoviridae	NODE_154_length_12967_cov_8.423220	12967	18.9906	FALSE	FALSE	TRUE
VC_1058_0	Unclassified	NODE_157_length_25355_cov_7.660952	25355	28.3019	FALSE	TRUE	TRUE
VC_1058_0	Siphoviridae	NODE_220_length_7224_cov_3.470596	7224	7.4057	FALSE	FALSE	TRUE

VC_1058_0	Siphoviridae	NODE_261_length_18726_cov_7.058464	18726	29.5377	FALSE	FALSE	TRUE
VC_1058_0	Siphoviridae	NODE_290_length_8339_cov_4.246845	8339	161.7264	FALSE	FALSE	TRUE
VC_1058_0	Siphoviridae	NODE_293_length_14099_cov_5.229571	14099	3.3774	FALSE	FALSE	TRUE
VC_1058_0	Siphoviridae	NODE_343_length_8892_cov_5.029683	8892	6.1981	FALSE	FALSE	TRUE
VC_1058_0	Siphoviridae	NODE_385_length_7568_cov_4.417057	7568	13.1887	FALSE	FALSE	TRUE
VC_1058_0	Siphoviridae	NODE_386_length_12792_cov_7.005830	12792	31.5189	FALSE	FALSE	TRUE
VC_1058_0	Siphoviridae	NODE_47_length_23548_cov_4.322402	23548	40.4245	FALSE	TRUE	TRUE
VC_1058_0	Siphoviridae	NODE_57_length_26250_cov_5.753279	26250	52.283	FALSE	TRUE	TRUE
VC_1058_0	Siphoviridae	NODE_59_length_29262_cov_22.786785	29262	3.4434	FALSE	TRUE	TRUE
VC_1058_0	Siphoviridae	NODE_61_length_12847_cov_3.312598	12847	16.2075	FALSE	FALSE	TRUE
VC_1058_0	Siphoviridae	NODE_62_length_31961_cov_14.933463	31961	33.5849	FALSE	TRUE	TRUE
VC_1058_0	Siphoviridae	NODE_65_length_35564_cov_46.973410	35564	31.3396	TRUE	TRUE	TRUE
VC_1058_0	Unclassified	NODE_67_length_29210_cov_16.158359	29210	33.3491	FALSE	TRUE	TRUE
VC_1058_0	Siphoviridae	NODE_71_length_31794_cov_19.470705	31794	20.5472	TRUE	TRUE	TRUE
VC_1058_0	Siphoviridae	NODE_72_length_32816_cov_124.168720	32816	126.8302	FALSE	TRUE	TRUE
VC_1058_0	Siphoviridae	NODE_74_length_31112_cov_25.880824	31112	15.2453	FALSE	TRUE	TRUE
VC_1058_0	Siphoviridae	NODE_78_length_18545_cov_5.630868	18545	16.4717	FALSE	FALSE	TRUE
VC_1058_0	Siphoviridae	NODE_80_length_11101_cov_10.047809	11101	40.2547	FALSE	FALSE	TRUE
VC_1058_0	Siphoviridae	NODE_82_length_22396_cov_49.830336	22396	81.0472	FALSE	TRUE	TRUE
VC_1058_0	Siphoviridae	NODE_83_length_12362_cov_3.047378	12362	18.6038	FALSE	FALSE	TRUE
VC_1058_0	Siphoviridae	NODE_83_length_34713_cov_24.631421	34713	24.0566	FALSE	TRUE	TRUE
VC_1058_0	Siphoviridae	NODE_98_length_25179_cov_8.029426	25179	34.1038	FALSE	TRUE	TRUE
VC_1058_0	Siphoviridae	NODE_98_length_30604_cov_12.328307	30604	5.9528	FALSE	TRUE	TRUE
VC_1138_0	Siphoviridae	NODE_102_length_33656_cov_139.440236	33656	57.6321	FALSE	TRUE	TRUE
VC_1138_0	Unclassified	NODE_120_length_17603_cov_4.419104	17603	28.8585	FALSE	FALSE	TRUE
VC_1138_0	Siphoviridae	NODE_1210_length_2887_cov_1.828910	2887	0.8491	FALSE	FALSE	FALSE
VC_1138_0	Siphoviridae	NODE_134_length_16193_cov_3.652666	16193	31.566	FALSE	FALSE	TRUE
VC_1138_0	Unclassified	NODE_146_length_26887_cov_4.257578	26887	54.8491	FALSE	TRUE	TRUE
VC_1138_0	Siphoviridae	NODE_150_length_16780_cov_3.893172	16780	11.6509	FALSE	FALSE	FALSE

VC_1138_0	Unclassified	NODE_17_length_31350_cov_12.187930	31350	48.6509	FALSE	TRUE	TRUE
VC_1138_0	Siphoviridae	NODE_199_length_14043_cov_4.501076	14043	12.9245	FALSE	FALSE	TRUE
VC_1138_0	Siphoviridae	NODE_21_length_33409_cov_26.068778	33409	36.2547	FALSE	TRUE	TRUE
VC_1138_0	Siphoviridae	NODE_224_length_11510_cov_5.486899	11510	17.066	FALSE	FALSE	TRUE
VC_1138_0	Siphoviridae	NODE_241_length_10939_cov_3.472786	10939	1.0094	FALSE	FALSE	TRUE
VC_1138_0	Siphoviridae	NODE_2421_length_1597_cov_1.942590	1597	19.3679	FALSE	FALSE	FALSE
VC_1138_0	Siphoviridae	NODE_26_length_32200_cov_10.664746	32200	63.3774	FALSE	TRUE	TRUE
VC_1138_0	Siphoviridae	NODE_260_length_6185_cov_2.978475	6185	4.9528	FALSE	FALSE	TRUE
VC_1138_0	Siphoviridae	NODE_2944_length_1991_cov_2.338795	1991	1.3585	FALSE	FALSE	FALSE
VC_1138_0	Siphoviridae	NODE_296_length_12390_cov_28.197787	12390	22.6509	FALSE	FALSE	TRUE
VC_1138_0	Siphoviridae	NODE_326_length_4714_cov_2.520910	4714	3.066	FALSE	FALSE	FALSE
VC_1138_0	Siphoviridae	NODE_329_length_8793_cov_3.244767	8793	12.6132	FALSE	FALSE	FALSE
VC_1138_0	Siphoviridae	NODE_347_length_7839_cov_3.459819	7839	19.1604	FALSE	FALSE	TRUE
VC_1138_0	Siphoviridae	NODE_35_length_38950_cov_92.022136	38950	2.7453	FALSE	TRUE	TRUE
VC_1138_0	Siphoviridae	NODE_36_length_37427_cov_74.809634	37427	23.4057	FALSE	TRUE	TRUE
VC_1138_0	Siphoviridae	NODE_3995_length_1345_cov_1.531300	1345	0.6415	FALSE	FALSE	FALSE
VC_1138_0	Unclassified	NODE_4576_length_1781_cov_1.755648	1781	1.9057	FALSE	FALSE	FALSE
VC_1138_0	Siphoviridae	NODE_576_length_6431_cov_3.879185	6431	89.0094	FALSE	FALSE	TRUE
VC_1138_0	Siphoviridae	NODE_60_length_35027_cov_18.400681	35027	115.1132	FALSE	TRUE	TRUE
VC_1138_0	Siphoviridae	NODE_60_length_35402_cov_13.739314	35402	1255.4528	FALSE	TRUE	TRUE
VC_1138_0	Siphoviridae	NODE_61_length_17214_cov_5.379024	17214	31.1698	FALSE	FALSE	TRUE
VC_1138_0	Unclassified	NODE_679_length_2658_cov_1.641657	2658	4.0189	FALSE	FALSE	FALSE
VC_1138_0	Siphoviridae	NODE_694_length_5787_cov_2.553973	5787	238.8113	FALSE	FALSE	FALSE
VC_1138_0	Siphoviridae	NODE_7_length_36455_cov_9.445181	36455	3.5943	FALSE	TRUE	TRUE
VC_1138_0	Siphoviridae	NODE_71_length_35196_cov_24.407015	35196	6.8113	FALSE	TRUE	TRUE
VC_1138_0	Siphoviridae	NODE_81_length_2098_cov_1.619310	2098	0.4528	FALSE	FALSE	FALSE
VC_1138_0	Siphoviridae	NODE_88_length_20900_cov_6.574107	20900	2.0094	FALSE	FALSE	TRUE
VC_1138_0	Siphoviridae	NODE_95_length_24444_cov_7.178928	24444	148.9811	FALSE	TRUE	TRUE
VC_1138_0	Siphoviridae	NODE_954_length_1998_cov_1.118483	1998	2.2547	FALSE	FALSE	FALSE

VC_1138_0	Siphoviridae	NODE_96_length_11829_cov_4.593947	11829	14.9811	FALSE	FALSE	TRUE
VC_1173_0	Podoviridae	NODE_119_length_16697_cov_35.921256	16697	37.8962	FALSE	FALSE	FALSE
VC_1173_0	Podoviridae	NODE_121_length_16956_cov_37.376224	16956	1367.6604	FALSE	FALSE	FALSE
VC_1173_0	Podoviridae	NODE_126_length_17126_cov_79.415399	17126	556.7453	FALSE	FALSE	FALSE
VC_1173_0	Podoviridae	NODE_127_length_16700_cov_45.095717	16700	42.0566	FALSE	FALSE	FALSE
VC_1173_0	Podoviridae	NODE_131_length_16433_cov_333.998898	16433	84.4057	FALSE	FALSE	FALSE
VC_1173_0	Podoviridae	NODE_140_length_16986_cov_13.056848	16986	22.8679	FALSE	FALSE	FALSE
VC_1173_0	Podoviridae	NODE_150_length_17288_cov_463.599511	17288	2759.9906	FALSE	FALSE	FALSE
VC_1173_0	Podoviridae	NODE_155_length_16387_cov_64.971083	16387	29.1509	FALSE	FALSE	FALSE
VC_1173_0	Podoviridae	NODE_166_length_16630_cov_45.112455	16630	14.6038	FALSE	FALSE	FALSE
VC_1173_0	Podoviridae	NODE_174_length_16763_cov_41.622300	16763	68.6509	FALSE	FALSE	FALSE
VC_1173_0	Podoviridae	NODE_175_length_16525_cov_18.906064	16525	31.8208	FALSE	FALSE	FALSE
VC_1173_0	Podoviridae	NODE_180_length_16360_cov_533.657832	16360	27.6887	FALSE	FALSE	FALSE
VC_1173_0	Podoviridae	NODE_233_length_16545_cov_76.232215	16545	15.0566	FALSE	FALSE	FALSE
VC_1173_0	Podoviridae	NODE_239_length_16522_cov_126.310175	16522	21.3491	FALSE	FALSE	FALSE
VC_1173_0	Podoviridae	NODE_280_length_12625_cov_117.721938	12625	3.1698	FALSE	FALSE	FALSE
VC_1173_0	Podoviridae	NODE_344_length_13149_cov_22.889425	13149	21.6321	FALSE	FALSE	FALSE
VC_1173_0	Podoviridae	NODE_36_length_16941_cov_15.332146	16941	9.3396	FALSE	FALSE	FALSE
VC_1173_0	Podoviridae	NODE_42_length_16262_cov_53.200272	16262	38.8113	FALSE	FALSE	FALSE
VC_1173_0	Podoviridae	NODE_42_length_17449_cov_9.425937	17449	515.1887	FALSE	FALSE	FALSE
VC_1173_0	Podoviridae	NODE_51_length_17429_cov_27.310271	17429	2.1038	FALSE	FALSE	FALSE
VC_1173_0	Podoviridae	NODE_85_length_16083_cov_121.041667	16083	24.6604	FALSE	FALSE	FALSE
VC_1173_0	Podoviridae	NODE_93_length_12960_cov_170.181945	12960	7.4151	FALSE	FALSE	FALSE
VC_1173_0	Podoviridae	NODE_98_length_15928_cov_11.762272	15928	22.934	FALSE	FALSE	FALSE
VC_1173_0	Podoviridae	NODE_98_length_17172_cov_73.694781	17172	3105.6698	FALSE	FALSE	FALSE
VC_1207_12	Myoviridae	NODE_1_length_358852_cov_18.698227	358852	368.6038	FALSE	TRUE	FALSE
VC_1278_0	Unclassified	NODE_103_length_18613_cov_2.984390	18613	17.9811	FALSE	FALSE	FALSE
VC_1278_0	Unclassified	NODE_109_length_23838_cov_5.940730	23838	32.066	FALSE	FALSE	FALSE
VC_1278_0	Mimiviridae	NODE_112_length_26351_cov_4.095421	26351	169.4717	FALSE	FALSE	FALSE

VC_1278_0	Unclassified	NODE_1132_length_4371_cov_1.755618	4371	1.4528	FALSE	FALSE	FALSE
VC_1278_0	Unclassified	NODE_114_length_28917_cov_3.250399	28917	30.0189	FALSE	FALSE	FALSE
VC_1278_0	Unclassified	NODE_1352_length_2863_cov_2.691027	2863	2.0943	FALSE	FALSE	FALSE
VC_1278_0	Unclassified	NODE_137_length_21463_cov_3.473507	21463	466.5943	FALSE	FALSE	FALSE
VC_1278_0	Unclassified	NODE_1379_length_2296_cov_1.689577	2296	1.0472	FALSE	FALSE	FALSE
VC_1278_0	Unclassified	NODE_1413_length_3464_cov_4.751560	3464	3.9057	FALSE	FALSE	FALSE
VC_1278_0	Unclassified	NODE_143_length_5520_cov_2.350120	5520	12.6038	FALSE	FALSE	FALSE
VC_1278_0	Siphoviridae	NODE_151_length_19905_cov_3.073665	19905	18.566	FALSE	TRUE	FALSE
VC_1278_0	Unclassified	NODE_153_length_14391_cov_2.942065	14391	22.6038	FALSE	FALSE	FALSE
VC_1278_0	Unclassified	NODE_1558_length_1941_cov_1.771987	1941	1.2075	FALSE	FALSE	FALSE
VC_1278_0	Siphoviridae	NODE_16_length_44056_cov_6.921810	44056	74.3491	FALSE	FALSE	FALSE
VC_1278_0	Unclassified	NODE_1758_length_3024_cov_5.252650	3024	3.0472	FALSE	FALSE	FALSE
VC_1278_0	Unclassified	NODE_180_length_12454_cov_3.576932	12454	24.283	FALSE	FALSE	FALSE
VC_1278_0	Unclassified	NODE_192_length_10392_cov_4.501409	10392	5.283	FALSE	FALSE	FALSE
VC_1278_0	Unclassified	NODE_192_length_10808_cov_3.729573	10808	17.066	FALSE	FALSE	FALSE
VC_1278_0	Unclassified	NODE_222_length_7682_cov_3.835421	7682	285.0566	FALSE	FALSE	FALSE
VC_1278_0	Unclassified	NODE_25_length_50603_cov_4.326390	50603	8.8208	FALSE	FALSE	FALSE
VC_1278_0	Unclassified	NODE_2554_length_1483_cov_1.928468	1483	0.4717	FALSE	FALSE	FALSE
VC_1278_0	Unclassified	NODE_2577_length_3718_cov_4.211661	3718	3.8868	FALSE	FALSE	FALSE
VC_1278_0	Unclassified	NODE_2603_length_1320_cov_1.492219	1320	0.5472	FALSE	FALSE	FALSE
VC_1278_0	Siphoviridae	NODE_301_length_11701_cov_4.443803	11701	15.5849	FALSE	FALSE	FALSE
VC_1278_0	Unclassified	NODE_3188_length_2192_cov_2.683708	2192	3.0377	FALSE	FALSE	FALSE
VC_1278_0	Unclassified	NODE_320_length_8459_cov_2.653828	8459	7.6509	FALSE	FALSE	FALSE
VC_1278_0	Unclassified	NODE_322_length_8121_cov_7.787709	8121	20.5189	FALSE	FALSE	FALSE
VC_1278_0	Siphoviridae	NODE_3569_length_1626_cov_2.240995	1626	0.7925	FALSE	FALSE	FALSE
VC_1278_0	Unclassified	NODE_3606_length_1906_cov_2.438296	1906	2.6038	FALSE	FALSE	FALSE
VC_1278_0	Unclassified	NODE_369_length_6393_cov_2.373371	6393	4.0943	FALSE	FALSE	FALSE
VC_1278_0	Unclassified	NODE_386_length_4439_cov_2.323041	4439	1.8868	FALSE	FALSE	FALSE
VC_1278_0	Unclassified	NODE_4000_length_2068_cov_2.141188	2068	3.5472	FALSE	FALSE	FALSE

VC_1278_0	Unclassified	NODE_414_length_5322_cov_4.290063	5322	4.3962	FALSE	FALSE	FALSE
VC_1278_0	Siphoviridae	NODE_418_length_20081_cov_5.857221	20081	51.0943	FALSE	FALSE	FALSE
VC_1278_0	Unclassified	NODE_438_length_5720_cov_3.227362	5720	4.934	FALSE	FALSE	FALSE
VC_1278_0	Unclassified	NODE_448_length_2600_cov_2.849660	2600	1.5094	FALSE	FALSE	FALSE
VC_1278_0	Unclassified	NODE_4637_length_1240_cov_1.119194	1240	0.9906	FALSE	FALSE	FALSE
VC_1278_0	Unclassified	NODE_468_length_5689_cov_7.329696	5689	5.6981	FALSE	FALSE	FALSE
VC_1278_0	Unclassified	NODE_5247_length_1468_cov_2.130752	1468	0.4717	FALSE	FALSE	FALSE
VC_1278_0	Unclassified	NODE_5469_length_1688_cov_1.631215	1688	1.4151	FALSE	FALSE	FALSE
VC_1278_0	Unclassified	NODE_557_length_5878_cov_3.072850	5878	7.0189	FALSE	FALSE	FALSE
VC_1278_0	Unclassified	NODE_559_length_4459_cov_2.079358	4459	1.6132	FALSE	FALSE	FALSE
VC_1278_0	Unclassified	NODE_591_length_6246_cov_5.297055	6246	7.9245	FALSE	FALSE	FALSE
VC_1278_0	Unclassified	NODE_658_length_3805_cov_2.750405	3805	2.9623	FALSE	FALSE	FALSE
VC_1278_0	Unclassified	NODE_744_length_3259_cov_2.325949	3259	2.0189	FALSE	FALSE	FALSE
VC_1278_0	Unclassified	NODE_900_length_2315_cov_1.717509	2315	0.5849	FALSE	FALSE	FALSE
VC_1278_0	Unclassified	NODE_946_length_3341_cov_2.561690	3341	5.5	FALSE	FALSE	FALSE
VC_1510_0	Unclassified	NODE_105_length_7717_cov_3.924652	7717	6.1698	FALSE	FALSE	FALSE
VC_1510_0	Unclassified	NODE_282_length_9588_cov_4.678997	9588	16.7358	FALSE	FALSE	FALSE
VC_1510_0	Siphoviridae	NODE_352_length_10186_cov_3.845940	10186	11.2075	FALSE	FALSE	FALSE
VC_1510_0	Siphoviridae	NODE_380_length_5794_cov_7.987357	5794	29.2264	FALSE	FALSE	FALSE
VC_1510_0	Siphoviridae	NODE_383_length_8755_cov_4.847736	8755	14.3396	FALSE	FALSE	FALSE
VC_1510_0	Unclassified	NODE_44_length_23580_cov_4.048550	23580	51.8019	FALSE	FALSE	FALSE
VC_1510_0	Unclassified	NODE_44_length_39788_cov_22.698330	39788	182.0094	FALSE	FALSE	FALSE
VC_1510_0	Myoviridae	NODE_498_length_5384_cov_4.086850	5384	10.1604	FALSE	FALSE	FALSE
VC_1510_0	Unclassified	NODE_647_length_4909_cov_7.818711	4909	56.4245	FALSE	FALSE	FALSE
VC_1510_0	Siphoviridae	NODE_701_length_6561_cov_17.938564	6561	8.3868	FALSE	FALSE	FALSE
VC_1770_0	Unclassified	NODE_1_length_190371_cov_2212.049713	190371	17626.4811	TRUE	FALSE	FALSE
VC_1770_0	Unclassified	NODE_1_length_92684_cov_2079.497176	92684	264.7453	FALSE	FALSE	FALSE
VC_1770_0	Unclassified	NODE_1_length_95810_cov_861.269185	95810	1828.1509	FALSE	FALSE	FALSE
VC_1770_0	Unclassified	NODE_1_length_96288_cov_493.542973	96288	2110.6981	FALSE	FALSE	FALSE

VC_1770_0	Unclassified	NODE_1_length_97296_cov_4691.412780	97296	584.4434	FALSE	FALSE	FALSE
VC_1770_0	Unclassified	NODE_1_length_97523_cov_4364.725776	97523	312.6887	TRUE	FALSE	FALSE
VC_1770_0	Unclassified	NODE_11_length_61556_cov_618.841108	61556	53.5472	FALSE	FALSE	FALSE
VC_1770_0	Unclassified	NODE_11_length_63327_cov_203.711663	63327	47.3208	FALSE	FALSE	FALSE
VC_1770_0	Unclassified	NODE_13_length_88317_cov_187.808973	88317	5873.8962	FALSE	FALSE	FALSE
VC_1770_0	Unclassified	NODE_13_length_92132_cov_275.905382	92132	348.6887	FALSE	FALSE	FALSE
VC_1770_0	Unclassified	NODE_13_length_92180_cov_373.707996	92180	10.5	FALSE	FALSE	FALSE
VC_1770_0	Unclassified	NODE_13_length_98578_cov_163.443181	98578	120.6509	TRUE	FALSE	FALSE
VC_1770_0	Unclassified	NODE_13_length_99516_cov_1003.578452	99516	1059.3302	TRUE	FALSE	FALSE
VC_1770_0	Unclassified	NODE_14_length_98293_cov_1587.167159	98293	308.6981	TRUE	FALSE	FALSE
VC_1770_0	Unclassified	NODE_26_length_68557_cov_418.984063	68557	126.9528	FALSE	FALSE	FALSE
VC_1770_0	Unclassified	NODE_3_length_91826_cov_258.246449	91826	159.9717	FALSE	FALSE	FALSE
VC_1770_0	Unclassified	NODE_5_length_68852_cov_12.056230	68852	0.4434	FALSE	FALSE	FALSE
VC_1770_0	Unclassified	NODE_5_length_91979_cov_458.773574	91979	3568.7736	FALSE	FALSE	FALSE
VC_1770_0	Unclassified	NODE_5_length_95751_cov_849.410227	95751	310.4906	FALSE	FALSE	FALSE
VC_1770_0	Unclassified	NODE_6_length_92729_cov_297.605257	92729	442.1604	FALSE	FALSE	FALSE
VC_1770_0	Unclassified	NODE_6_length_96249_cov_997.055049	96249	8.2547	FALSE	FALSE	FALSE
VC_1770_0	Unclassified	NODE_6_length_97854_cov_205.797146	97854	60.3396	TRUE	FALSE	FALSE
VC_1770_0	Unclassified	NODE_62_length_69875_cov_657.317244	69875	160.1792	FALSE	FALSE	FALSE
VC_1770_0	Unclassified	NODE_8_length_95614_cov_645.546176	95614	221.1415	FALSE	FALSE	FALSE
VC_1770_0	Unclassified	NODE_9_length_91948_cov_19.457087	91948	4.0189	FALSE	FALSE	FALSE
VC_1770_0	Unclassified	NODE_9_length_95649_cov_1666.922009	95649	11.3679	FALSE	FALSE	FALSE
VC_1770_0	Unclassified	NODE_9_length_97954_cov_185.133207	97954	837.2358	TRUE	FALSE	FALSE
VC_1853_0	Unclassified	NODE_1015_length_4247_cov_134.042189	4247	6.7547	TRUE	FALSE	FALSE
VC_1853_0	Unclassified	NODE_1107_length_4235_cov_2288.146277	4235	12.7642	TRUE	FALSE	FALSE
VC_1853_0	Unclassified	NODE_257_length_4247_cov_59.799180	4247	4.7358	TRUE	FALSE	FALSE
VC_1853_0	Unclassified	NODE_371_length_4247_cov_58.131148	4247	396.9906	TRUE	FALSE	FALSE
VC_1853_0	Myoviridae	NODE_417_length_4416_cov_121.240908	4416	2.3774	TRUE	FALSE	FALSE
VC_1853_0	Unclassified	NODE_553_length_4247_cov_182.791225	4247	2.0755	TRUE	FALSE	FALSE

VC_1853_0	Myoviridae	NODE_560_length_6159_cov_7.054950	6159	8.0566	TRUE	FALSE	FALSE
VC_1853_0	Unclassified	NODE_580_length_4247_cov_18.107763	4247	5.7736	TRUE	FALSE	FALSE
VC_1853_0	Unclassified	NODE_994_length_4247_cov_25.991562	4247	12.2736	TRUE	FALSE	FALSE
VC_1982_0	Podoviridae	NODE_101_length_17267_cov_9.449383	17267	28.9623	FALSE	FALSE	FALSE
VC_1982_0	Podoviridae	NODE_110_length_16683_cov_36.959540	16683	17.4245	FALSE	FALSE	FALSE
VC_1982_0	Podoviridae	NODE_111_length_17070_cov_12.060103	17070	49.566	FALSE	FALSE	FALSE
VC_1982_0	Podoviridae	NODE_115_length_16482_cov_1091.059147	16482	22.9434	FALSE	FALSE	FALSE
VC_1982_0	Podoviridae	NODE_116_length_14055_cov_29.469046	14055	0.6509	FALSE	FALSE	FALSE
VC_1982_0	Podoviridae	NODE_126_length_12749_cov_58.717391	12749	12.6698	FALSE	FALSE	FALSE
VC_1982_0	Podoviridae	NODE_129_length_15325_cov_21.054315	15325	20.8302	FALSE	FALSE	FALSE
VC_1982_0	Podoviridae	NODE_134_length_15925_cov_19.673512	15925	35.1321	FALSE	FALSE	FALSE
VC_1982_0	Podoviridae	NODE_141_length_15066_cov_1627.561569	15066	23.6792	FALSE	FALSE	FALSE
VC_1982_0	Podoviridae	NODE_142_length_15531_cov_114.713906	15531	31.5755	FALSE	FALSE	FALSE
VC_1982_0	Podoviridae	NODE_145_length_13862_cov_30.841702	13862	9.8962	FALSE	FALSE	FALSE
VC_1982_0	Podoviridae	NODE_146_length_13545_cov_2.971069	13545	5.8585	FALSE	FALSE	FALSE
VC_1982_0	Podoviridae	NODE_148_length_9763_cov_3.416184	9763	9.4811	FALSE	FALSE	FALSE
VC_1982_0	Podoviridae	NODE_152_length_16250_cov_64.170887	16250	9.0189	FALSE	FALSE	FALSE
VC_1982_0	Podoviridae	NODE_155_length_15911_cov_31.574690	15911	51.8302	FALSE	FALSE	FALSE
VC_1982_0	Podoviridae	NODE_158_length_13602_cov_38.682737	13602	14.7925	FALSE	FALSE	FALSE
VC_1982_0	Podoviridae	NODE_180_length_13044_cov_65.728930	13044	13.5566	FALSE	FALSE	FALSE
VC_1982_0	Podoviridae	NODE_186_length_15845_cov_56.708942	15845	4.0472	FALSE	FALSE	FALSE
VC_1982_0	Podoviridae	NODE_187_length_15210_cov_9.413275	15210	22.1604	FALSE	FALSE	FALSE
VC_1982_0	Podoviridae	NODE_195_length_16388_cov_12.250230	16388	16.8019	FALSE	FALSE	FALSE
VC_1982_0	Podoviridae	NODE_198_length_17326_cov_13.068845	17326	9.283	FALSE	FALSE	FALSE
VC_1982_0	Podoviridae	NODE_203_length_14431_cov_305.301702	14431	375.3679	FALSE	FALSE	FALSE
VC_1982_0	Podoviridae	NODE_207_length_14288_cov_4.377053	14288	14.8962	FALSE	FALSE	FALSE
VC_1982_0	Podoviridae	NODE_21_length_9873_cov_3.444956	9873	6.1509	FALSE	FALSE	FALSE
VC_1982_0	Podoviridae	NODE_229_length_11176_cov_3.679245	11176	10.3208	FALSE	FALSE	FALSE
VC_1982_0	Podoviridae	NODE_247_length_10449_cov_5.828406	10449	10.6698	FALSE	FALSE	FALSE

VC_1982_0	Podoviridae	NODE_25_length_16140_cov_53.723751	16140	0.5849	FALSE	FALSE	FALSE
VC_1982_0	Podoviridae	NODE_251_length_15935_cov_21.391106	15935	35.4906	FALSE	FALSE	FALSE
VC_1982_0	Podoviridae	NODE_296_length_15355_cov_4.544114	15355	27.5377	FALSE	FALSE	FALSE
VC_1982_0	Podoviridae	NODE_307_length_15075_cov_20.242059	15075	25.5094	FALSE	FALSE	FALSE
VC_1982_0	Podoviridae	NODE_336_length_7468_cov_5.556249	7468	5.2264	FALSE	FALSE	FALSE
VC_1982_0	Podoviridae	NODE_339_length_13444_cov_5.713653	13444	53.7358	FALSE	FALSE	FALSE
VC_1982_0	Podoviridae	NODE_35_length_17132_cov_8.831034	17132	15.6226	FALSE	FALSE	FALSE
VC_1982_0	Podoviridae	NODE_37_length_20630_cov_41.534947	20630	79.8019	FALSE	FALSE	FALSE
VC_1982_0	Podoviridae	NODE_377_length_10663_cov_5.799413	10663	6.5377	FALSE	FALSE	FALSE
VC_1982_0	Podoviridae	NODE_381_length_12980_cov_12.488782	12980	25.217	FALSE	FALSE	FALSE
VC_1982_0	Podoviridae	NODE_397_length_7329_cov_6.502213	7329	3.3396	FALSE	FALSE	FALSE
VC_1982_0	Podoviridae	NODE_50_length_15721_cov_20.443221	15721	598.3491	FALSE	FALSE	FALSE
VC_1982_0	Podoviridae	NODE_56_length_14195_cov_47.892193	14195	24753.1981	FALSE	FALSE	FALSE
VC_1982_0	Podoviridae	NODE_75_length_16426_cov_8.762479	16426	15.2642	FALSE	FALSE	FALSE
VC_1982_0	Podoviridae	NODE_77_length_16439_cov_163.906059	16439	2.8019	FALSE	FALSE	FALSE
VC_1982_0	Podoviridae	NODE_77_length_28991_cov_13.667174	28991	42	FALSE	FALSE	FALSE
VC_1982_0	Podoviridae	NODE_932_length_4930_cov_3.148416	4930	7.3679	FALSE	FALSE	FALSE
VC_1983_0	Podoviridae	NODE_104_length_16616_cov_7.609130	16616	6.8868	FALSE	FALSE	FALSE
VC_1983_0	Podoviridae	NODE_108_length_16515_cov_53.662890	16515	53.9245	FALSE	FALSE	FALSE
VC_1983_0	Podoviridae	NODE_11_length_16558_cov_9.013427	16558	24.3019	FALSE	FALSE	FALSE
VC_1983_0	Podoviridae	NODE_115_length_16198_cov_21.611591	16198	16.0755	FALSE	FALSE	FALSE
VC_1983_0	Podoviridae	NODE_116_length_15731_cov_9.036272	15731	165.5283	FALSE	FALSE	FALSE
VC_1983_0	Podoviridae	NODE_12_length_16063_cov_6.529003	16063	206.8774	FALSE	FALSE	FALSE
VC_1983_0	Podoviridae	NODE_121_length_14635_cov_78.467804	14635	169.6226	FALSE	FALSE	FALSE
VC_1983_0	Podoviridae	NODE_126_length_14945_cov_249.697360	14945	26.0943	FALSE	FALSE	FALSE
VC_1983_0	Podoviridae	NODE_129_length_16754_cov_173.156409	16754	155.1698	FALSE	FALSE	FALSE
VC_1983_0	Podoviridae	NODE_132_length_15624_cov_12.744541	15624	35.3396	FALSE	FALSE	FALSE
VC_1983_0	Podoviridae	NODE_134_length_16307_cov_101.669299	16307	23.8019	FALSE	FALSE	FALSE
VC_1983_0	Podoviridae	NODE_135_length_16154_cov_8.239240	16154	33.6792	FALSE	FALSE	FALSE

VC_1983_0	Podoviridae	NODE_136_length_16149_cov_119.791776	16149	20.3491	FALSE	FALSE	FALSE
VC_1983_0	Podoviridae	NODE_139_length_14281_cov_5.757368	14281	11.0377	FALSE	FALSE	FALSE
VC_1983_0	Podoviridae	NODE_139_length_15583_cov_19.489989	15583	19.0377	FALSE	FALSE	FALSE
VC_1983_0	Podoviridae	NODE_144_length_16517_cov_10.469850	16517	29.434	FALSE	FALSE	FALSE
VC_1983_0	Podoviridae	NODE_146_length_14125_cov_22.815101	14125	17.8774	FALSE	FALSE	FALSE
VC_1983_0	Podoviridae	NODE_148_length_14594_cov_44.387259	14594	31.5755	FALSE	FALSE	FALSE
VC_1983_0	Podoviridae	NODE_160_length_7880_cov_3.331191	7880	32.0849	FALSE	FALSE	FALSE
VC_1983_0	Podoviridae	NODE_174_length_16668_cov_37.506850	16668	1.6415	FALSE	FALSE	FALSE
VC_1983_0	Podoviridae	NODE_175_length_16337_cov_34.668001	16337	40.6604	FALSE	FALSE	FALSE
VC_1983_0	Podoviridae	NODE_177_length_16301_cov_15.133934	16301	10.066	FALSE	FALSE	FALSE
VC_1983_0	Podoviridae	NODE_18_length_15895_cov_46.746342	15895	197.5189	FALSE	FALSE	FALSE
VC_1983_0	Podoviridae	NODE_187_length_13426_cov_3.802206	13426	17.1038	FALSE	FALSE	FALSE
VC_1983_0	Podoviridae	NODE_201_length_10924_cov_3.314734	10924	5.4434	FALSE	FALSE	FALSE
VC_1983_0	Podoviridae	NODE_211_length_14779_cov_116.424114	14779	27.2264	FALSE	FALSE	FALSE
VC_1983_0	Podoviridae	NODE_217_length_14746_cov_31.933024	14746	53.7075	FALSE	FALSE	FALSE
VC_1983_0	Podoviridae	NODE_242_length_15924_cov_29.335861	15924	26.1132	FALSE	FALSE	FALSE
VC_1983_0	Podoviridae	NODE_259_length_16784_cov_4.597782	16784	164.6321	FALSE	FALSE	FALSE
VC_1983_0	Podoviridae	NODE_26_length_14703_cov_5.915845	14703	30.8302	FALSE	FALSE	FALSE
VC_1983_0	Podoviridae	NODE_273_length_9000_cov_19.385462	9000	21.7453	FALSE	FALSE	FALSE
VC_1983_0	Podoviridae	NODE_276_length_14802_cov_9.908386	14802	15.6038	FALSE	FALSE	FALSE
VC_1983_0	Podoviridae	NODE_277_length_14551_cov_3.745710	14551	28.1415	FALSE	FALSE	FALSE
VC_1983_0	Podoviridae	NODE_289_length_8203_cov_3.788129	8203	6.6509	FALSE	FALSE	FALSE
VC_1983_0	Podoviridae	NODE_373_length_8607_cov_9.136695	8607	16.9717	FALSE	FALSE	FALSE
VC_1983_0	Podoviridae	NODE_51_length_17262_cov_10.375051	17262	41.6509	FALSE	FALSE	FALSE
VC_1983_0	Podoviridae	NODE_54_length_16122_cov_8.486925	16122	20.7547	FALSE	FALSE	FALSE
VC_1983_0	Podoviridae	NODE_55_length_14083_cov_7.048512	14083	9.2453	FALSE	FALSE	FALSE
VC_1983_0	Podoviridae	NODE_56_length_8149_cov_2.885590	8149	7.2925	FALSE	FALSE	FALSE
VC_1983_0	Podoviridae	NODE_57_length_16037_cov_6.636153	16037	7.6887	FALSE	FALSE	FALSE
VC_1983_0	Podoviridae	NODE_64_length_15808_cov_92.538163	15808	9.6321	FALSE	FALSE	FALSE

VC_1983_0	Podoviridae	NODE_65_length_15797_cov_6.845904	15797	6.7736	FALSE	FALSE	FALSE
VC_1983_0	Podoviridae	NODE_69_length_11065_cov_48.881087	11065	1.1698	FALSE	FALSE	FALSE
VC_1983_0	Podoviridae	NODE_72_length_15937_cov_68.043777	15937	21.4151	FALSE	FALSE	FALSE
VC_1983_0	Podoviridae	NODE_80_length_16242_cov_65.198848	16242	62.8019	FALSE	FALSE	FALSE
VC_1983_0	Podoviridae	NODE_90_length_15831_cov_31.859950	15831	1179.4906	FALSE	FALSE	FALSE
VC_1983_0	Podoviridae	NODE_96_length_16709_cov_25.814088	16709	3.9434	FALSE	FALSE	FALSE
VC_2555_0	Myoviridae	NODE_1282_length_6185_cov_4.219849	6185	9.1792	FALSE	FALSE	FALSE
VC_2555_0	Myoviridae	NODE_1404_length_3067_cov_2.366240	3067	1.3679	FALSE	FALSE	FALSE
VC_2555_0	Myoviridae	NODE_1565_length_2493_cov_1.651211	2493	2.6887	FALSE	FALSE	FALSE
VC_2555_0	Myoviridae	NODE_1831_length_2955_cov_4.549720	2955	10.4528	FALSE	FALSE	FALSE
VC_2555_0	Myoviridae	NODE_213_length_8813_cov_2.991049	8813	23.7075	FALSE	FALSE	FALSE
VC_2555_0	Myoviridae	NODE_2249_length_4104_cov_2.612235	4104	3.8019	FALSE	FALSE	FALSE
VC_2555_0	Myoviridae	NODE_2388_length_1968_cov_1.818620	1968	3.6038	FALSE	FALSE	FALSE
VC_2555_0	Myoviridae	NODE_2736_length_1888_cov_1.532700	1888	1.6226	FALSE	FALSE	FALSE
VC_2555_0	Myoviridae	NODE_275_length_9093_cov_5.639537	9093	0.8962	FALSE	FALSE	FALSE
VC_2555_0	Myoviridae	NODE_278_length_6397_cov_2.749921	6397	6.6792	FALSE	FALSE	FALSE
VC_2555_0	Myoviridae	NODE_280_length_15834_cov_3.454083	15834	25.8396	FALSE	FALSE	FALSE
VC_2555_0	Myoviridae	NODE_320_length_8976_cov_4.202771	8976	10.1698	FALSE	FALSE	FALSE
VC_2555_0	Myoviridae	NODE_3300_length_2200_cov_1.515469	2200	1.2075	FALSE	FALSE	FALSE
VC_2555_0	Myoviridae	NODE_3381_length_1834_cov_1.079539	1834	11.2358	FALSE	FALSE	FALSE
VC_2555_0	Myoviridae	NODE_3745_length_2642_cov_2.554463	2642	2.3774	FALSE	FALSE	FALSE
VC_2555_0	Myoviridae	NODE_401_length_6283_cov_5.808862	6283	6.1792	FALSE	FALSE	FALSE
VC_2555_0	Myoviridae	NODE_411_length_12053_cov_9.663627	12053	36.283	FALSE	FALSE	FALSE
VC_2555_0	Myoviridae	NODE_4179_length_1987_cov_1.401483	1987	2.5849	FALSE	FALSE	FALSE
VC_2555_0	Myoviridae	NODE_423_length_7984_cov_6.028789	7984	11.4906	FALSE	FALSE	FALSE
VC_2555_0	Myoviridae	NODE_426_length_9616_cov_3.607860	9616	22.3774	FALSE	FALSE	FALSE
VC_2555_0	Myoviridae	NODE_430_length_5973_cov_2.495914	5973	1.7925	FALSE	FALSE	FALSE
VC_2555_0	Myoviridae	NODE_498_length_4313_cov_1.973897	4313	1.2642	FALSE	FALSE	FALSE
VC_2555_0	Myoviridae	NODE_501_length_7098_cov_5.034433	7098	4.1226	FALSE	FALSE	FALSE

VC_2555_0	Myoviridae	NODE_509_length_2353_cov_1.887755	2353	1.9811	FALSE	FALSE	FALSE
VC_2555_0	Myoviridae	NODE_5191_length_1746_cov_1.580449	1746	3.5943	FALSE	FALSE	FALSE
VC_2555_0	Myoviridae	NODE_521_length_1374_cov_1.254902	1374	0.5189	FALSE	FALSE	FALSE
VC_2555_0	Myoviridae	NODE_5524_length_2017_cov_2.969239	2017	1.6698	FALSE	FALSE	FALSE
VC_2555_0	Myoviridae	NODE_60_length_62413_cov_13.418461	62413	17.3774	FALSE	FALSE	FALSE
VC_2555_0	Myoviridae	NODE_641_length_4195_cov_3.126709	4195	1.6321	FALSE	FALSE	FALSE
VC_2555_0	Unclassified	NODE_65_length_60434_cov_5.101798	60434	18.2736	FALSE	FALSE	FALSE
VC_2555_0	Myoviridae	NODE_654_length_6903_cov_2.740594	6903	10.566	FALSE	FALSE	FALSE
VC_2555_0	Myoviridae	NODE_683_length_5133_cov_2.958284	5133	5.5566	FALSE	FALSE	FALSE
VC_2555_0	Myoviridae	NODE_6851_length_1382_cov_1.272019	1382	0.3679	FALSE	FALSE	FALSE
VC_2555_0	Myoviridae	NODE_698_length_7905_cov_2.346016	7905	34.1415	FALSE	FALSE	FALSE
VC_2555_0	Myoviridae	NODE_78_length_9041_cov_3.217848	9041	2.6415	FALSE	FALSE	FALSE
VC_2555_0	Myoviridae	NODE_8305_length_1235_cov_1.088908	1235	0.3585	FALSE	FALSE	FALSE
VC_2555_0	Myoviridae	NODE_8470_length_1453_cov_2.219350	1453	0.6604	FALSE	FALSE	FALSE
VC_2555_0	Myoviridae	NODE_872_length_3483_cov_1.793144	3483	6.2264	FALSE	FALSE	FALSE
VC_2555_0	Myoviridae	NODE_884_length_3019_cov_2.725000	3019	2.434	FALSE	FALSE	FALSE
VC_2555_0	Myoviridae	NODE_982_length_3500_cov_1.892385	3500	2.6226	FALSE	FALSE	FALSE
VC_2838_0	Unclassified	NODE_1046_length_2318_cov_9.461920	2318	0.8113	TRUE	FALSE	FALSE
VC_2838_0	Unclassified	NODE_1049_length_2318_cov_2.724651	2318	1.9623	TRUE	FALSE	FALSE
VC_2838_0	Unclassified	NODE_1123_length_2318_cov_37.459667	2318	0.8491	TRUE	FALSE	FALSE
VC_2838_0	Unclassified	NODE_1532_length_2318_cov_11.294727	2318	1.1415	TRUE	FALSE	FALSE
VC_2838_0	Unclassified	NODE_1811_length_2318_cov_105.657954	2318	0.934	TRUE	FALSE	FALSE
VC_2838_0	Unclassified	NODE_1864_length_2318_cov_49.892744	2318	3.4811	TRUE	FALSE	FALSE
VC_2838_0	Unclassified	NODE_2005_length_2316_cov_4.188543	2316	2.5094	TRUE	FALSE	FALSE
VC_2838_0	Unclassified	NODE_2283_length_2318_cov_11.241100	2318	0.8491	TRUE	FALSE	FALSE
VC_2838_0	Unclassified	NODE_2351_length_2318_cov_17.107256	2318	0.8208	TRUE	FALSE	FALSE
VC_2838_0	Unclassified	NODE_2364_length_2318_cov_6.308247	2318	1.6226	TRUE	FALSE	FALSE
VC_2838_0	Unclassified	NODE_254_length_2318_cov_15.523659	2318	0.9245	TRUE	FALSE	FALSE
VC_2838_0	Unclassified	NODE_2614_length_2352_cov_5.810919	2352	0.566	TRUE	FALSE	FALSE

VC_2838_0	Unclassified	NODE_2919_length_2318_cov_8.707075	2318	2.434	TRUE	FALSE	FALSE
VC_2838_0	Unclassified	NODE_3057_length_2318_cov_6.732312	2318	0.8962	TRUE	FALSE	FALSE
VC_2838_0	Unclassified	NODE_461_length_2318_cov_10.333484	2318	2.0189	TRUE	FALSE	FALSE
VC_2838_0	Unclassified	NODE_472_length_2318_cov_9.393871	2318	1.7642	TRUE	FALSE	FALSE
VC_2838_0	Unclassified	NODE_5255_length_2318_cov_11.963948	2318	1.9623	TRUE	FALSE	FALSE
VC_2838_0	Unclassified	NODE_689_length_2318_cov_5.790446	2318	2.1604	TRUE	FALSE	FALSE
VC_2838_0	Unclassified	NODE_790_length_2318_cov_5.048671	2318	2.8679	TRUE	FALSE	FALSE
VC_2838_0	Unclassified	NODE_897_length_2318_cov_34.888238	2318	2.6132	TRUE	FALSE	FALSE
VC_2838_0	Unclassified	NODE_942_length_2318_cov_16.877873	2318	15.0189	TRUE	FALSE	FALSE
VC_2838_0	Unclassified	NODE_945_length_2342_cov_1.948729	2342	2.3585	TRUE	FALSE	FALSE
VC_513_0	Siphoviridae	NODE_11_length_41781_cov_66.677055	41781	44.8774	FALSE	FALSE	TRUE
VC_513_0	Siphoviridae	NODE_11_length_43634_cov_1310.481176	43634	41.3113	TRUE	FALSE	TRUE
VC_513_0	Siphoviridae	NODE_129_length_31595_cov_30.984950	31595	2.8019	FALSE	TRUE	TRUE
VC_513_0	Siphoviridae	NODE_13_length_45461_cov_113.301596	45461	2488.4811	TRUE	FALSE	TRUE
VC_513_0	Siphoviridae	NODE_14_length_41599_cov_50.740169	41599	1490.0189	TRUE	FALSE	TRUE
VC_513_0	Siphoviridae	NODE_14_length_44224_cov_14.152113	44224	118.3019	TRUE	FALSE	TRUE
VC_513_0	Siphoviridae	NODE_14_length_45162_cov_33.716330	45162	335.8396	TRUE	FALSE	TRUE
VC_513_0	Siphoviridae	NODE_14_length_46846_cov_46.983293	46846	165.0943	TRUE	FALSE	TRUE
VC_513_0	Siphoviridae	NODE_14_length_77162_cov_34.793143	77162	18.0283	FALSE	FALSE	TRUE
VC_513_0	Siphoviridae	NODE_142_length_42020_cov_27.095012	42020	139.6415	FALSE	FALSE	TRUE
VC_513_0	Siphoviridae	NODE_15_length_41226_cov_11.993557	41226	28.4528	FALSE	FALSE	TRUE
VC_513_0	Siphoviridae	NODE_15_length_42309_cov_829.513030	42309	74.7075	TRUE	FALSE	TRUE
VC_513_0	Siphoviridae	NODE_16_length_30920_cov_17.620843	30920	802.934	FALSE	TRUE	TRUE
VC_513_0	Siphoviridae	NODE_16_length_43104_cov_46.567422	43104	4.1792	FALSE	TRUE	TRUE
VC_513_0	Siphoviridae	NODE_166_length_32468_cov_155.797924	32468	36.9057	FALSE	TRUE	TRUE
VC_513_0	Siphoviridae	NODE_17_length_41602_cov_17.023010	41602	108.7547	TRUE	FALSE	TRUE
VC_513_0	Siphoviridae	NODE_18_length_45627_cov_41.754678	45627	206.9057	FALSE	FALSE	TRUE
VC_513_0	Siphoviridae	NODE_19_length_43018_cov_966.038724	43018	62.2736	TRUE	FALSE	TRUE
VC_513_0	Siphoviridae	NODE_19_length_44748_cov_55.262447	44748	12.2264	TRUE	FALSE	TRUE

VC_513_0	Siphoviridae	NODE_19_length_45403_cov_28.275296	45403	78.7925	FALSE	FALSE	TRUE
VC_513_0	Siphoviridae	NODE_20_length_46001_cov_6.384776	46001	17.1604	FALSE	FALSE	TRUE
VC_513_0	Siphoviridae	NODE_25_length_45240_cov_6.070712	45240	147.8585	TRUE	FALSE	TRUE
VC_513_0	Siphoviridae	NODE_28_length_43795_cov_41.642118	43795	69.4811	TRUE	FALSE	TRUE
VC_513_0	Siphoviridae	NODE_28_length_44266_cov_652.648063	44266	109.566	TRUE	FALSE	TRUE
VC_513_0	Siphoviridae	NODE_29_length_45653_cov_35.470914	45653	732.4151	FALSE	FALSE	TRUE
VC_513_0	Siphoviridae	NODE_30_length_29112_cov_106.016510	29112	32.1698	FALSE	FALSE	TRUE
VC_513_0	Siphoviridae	NODE_31_length_41853_cov_1367.382502	41853	379.717	FALSE	TRUE	TRUE
VC_513_0	Siphoviridae	NODE_31_length_44870_cov_8.324362	44870	170.5849	TRUE	FALSE	TRUE
VC_513_0	Siphoviridae	NODE_31_length_44958_cov_12.445529	44958	96.7075	FALSE	FALSE	TRUE
VC_513_0	Siphoviridae	NODE_32_length_43085_cov_57.665566	43085	141.0755	TRUE	FALSE	TRUE
VC_513_0	Siphoviridae	NODE_35_length_45491_cov_3140.043223	45491	69.0566	TRUE	FALSE	TRUE
VC_513_0	Siphoviridae	NODE_36_length_31303_cov_44.377452	31303	114.3491	FALSE	TRUE	TRUE
VC_513_0	Siphoviridae	NODE_36_length_45325_cov_10.035378	45325	78.9623	TRUE	FALSE	TRUE
VC_513_0	Siphoviridae	NODE_38_length_42878_cov_370.355128	42878	2.3491	FALSE	FALSE	TRUE
VC_513_0	Siphoviridae	NODE_40_length_45678_cov_111.558196	45678	71.4717	TRUE	FALSE	TRUE
VC_513_0	Siphoviridae	NODE_42_length_31971_cov_1193.005553	31971	2747.5	FALSE	TRUE	TRUE
VC_513_0	Siphoviridae	NODE_42_length_39553_cov_9.198535	39553	89.9906	FALSE	FALSE	TRUE
VC_513_0	Siphoviridae	NODE_42_length_45525_cov_30.875622	45525	30.4623	FALSE	FALSE	TRUE
VC_513_0	Siphoviridae	NODE_43_length_45388_cov_41.648391	45388	49.3491	TRUE	FALSE	TRUE
VC_513_0	Siphoviridae	NODE_45_length_44221_cov_658.484135	44221	164	TRUE	FALSE	TRUE
VC_513_0	Siphoviridae	NODE_46_length_42041_cov_604.521744	42041	4.0943	TRUE	FALSE	TRUE
VC_513_0	Siphoviridae	NODE_48_length_45760_cov_162.512626	45760	2.783	TRUE	FALSE	TRUE
VC_513_0	Siphoviridae	NODE_5_length_45790_cov_23.004071	45790	98.6981	TRUE	FALSE	TRUE
VC_513_0	Siphoviridae	NODE_50_length_42801_cov_70.870849	42801	10.5283	TRUE	FALSE	TRUE
VC_513_0	Siphoviridae	NODE_51_length_44347_cov_8.650515	44347	109.1792	FALSE	FALSE	TRUE
VC_513_0	Siphoviridae	NODE_52_length_41619_cov_1358.751036	41619	92.9717	FALSE	TRUE	TRUE
VC_513_0	Siphoviridae	NODE_52_length_42203_cov_610.028358	42203	93.0943	FALSE	FALSE	TRUE
VC_513_0	Siphoviridae	NODE_52_length_44226_cov_8.597956	44226	261.1038	FALSE	FALSE	TRUE

VC_513_0	Siphoviridae	NODE_54_length_33976_cov_120.872273	33976	7.3868	FALSE	FALSE	TRUE
VC_513_0	Siphoviridae	NODE_54_length_44776_cov_28.020570	44776	7.0472	TRUE	FALSE	TRUE
VC_513_0	Siphoviridae	NODE_57_length_29577_cov_366.322715	29577	1.3396	FALSE	FALSE	TRUE
VC_513_0	Siphoviridae	NODE_59_length_42525_cov_26.554495	42525	55.8868	TRUE	FALSE	TRUE
VC_513_0	Siphoviridae	NODE_63_length_44595_cov_336.971548	44595	94.8113	TRUE	FALSE	TRUE
VC_513_0	Siphoviridae	NODE_63_length_44617_cov_6.239993	44617	91.0943	TRUE	FALSE	TRUE
VC_513_0	Siphoviridae	NODE_65_length_42980_cov_31.852779	42980	167.4623	TRUE	FALSE	TRUE
VC_513_0	Siphoviridae	NODE_73_length_46742_cov_56.114358	46742	273.9057	FALSE	FALSE	TRUE
VC_513_0	Siphoviridae	NODE_77_length_42669_cov_10.584872	42669	223.6226	TRUE	FALSE	TRUE
VC_513_0	Siphoviridae	NODE_8_length_44463_cov_10.185376	44463	24.8774	TRUE	FALSE	TRUE
VC_513_0	Siphoviridae	NODE_86_length_28225_cov_104.857463	28225	45.3774	FALSE	FALSE	TRUE

Supplementary Table 2: Taxonomic classification of viral clusters. Clusters which contained contigs with varying family level classifications were assigned the family with the most votes and the additional classifications are noted in the Family 2, Family 3 and Family 4 columns. The bacterial host species associated with viral contigs in these viral clusters are summarised in the columns beginning with the text 'Host Species'. As different viral contigs in a cluster can have different predicted bacterial hosts, the percentage of contigs with hits for a particular host are presented in the Host Species Percent Votes columns. This is a truncated table with information of viral clusters with prevalence of greater than 10% in IBS or Controls. For complete table, please refer to the published article (doi: 10.1080/19490976.2021.1887719).

Cluster ID	Family	Votes(%)	Cluster ID	Family	Votes(%)	Cluster ID	Family	Votes(%)
VC_1058_0	Siphoviridae	94.59	VC_1786_0	Siphoviridae	100	VC_3018_0	Siphoviridae	93.75
VC_1078_0	Siphoviridae	94.12	VC_1803_0	Myoviridae	50	VC_3030_0	Unclassified	48.78
VC_1086_0	Siphoviridae	100	VC_1808_0	Siphoviridae	50	VC_3035_0	Siphoviridae	47.37
VC_1087_0	Siphoviridae	100	VC_1811_0	Siphoviridae	42.86	VC_3037_0	Podoviridae	60
VC_1091_0	Siphoviridae	100	VC_1815_0	Unclassified	94.12	VC_3076_0	Podoviridae	100
VC_1116_0	Podoviridae	66.67	VC_1853_0	Unclassified	77.78	VC_3099_0	Unclassified	63.64
VC_1138_0	Siphoviridae	86.11	VC_1854_0	Unclassified	53.33	VC_3141_0	Unclassified	90.91
VC_1159_0	Myoviridae	50	VC_1857_1	Siphoviridae	100	VC_3195_0	Siphoviridae	50
VC_1173_0	Podoviridae	100	VC_1858_0	Myoviridae	60.87	VC_3206_0	Siphoviridae	81.25
VC_1177_0	Unclassified	100	VC_1885_0	Unclassified	55	VC_3211_0	Siphoviridae	87.5
VC_1182_0	Myoviridae	50	VC_1890_1	Unclassified	70.59	VC_3316_0	Unclassified	50
VC_1184_0	Siphoviridae	63.64	VC_1893_0	Unclassified	100	VC_3320_0	Unclassified	54.55
VC_1189_0	Unclassified	78.57	VC_1896_0	Podoviridae	100	VC_3341_0	Unclassified	75
VC_1196_3	Siphoviridae	93.33	VC_1910_0	Mimiviridae	42.86	VC_3359_0	Myoviridae	69.23
VC_1205_0	Unclassified	88.89	VC_1911_0	Myoviridae	33.33	VC_3361_0	Unclassified	72.73
VC_1207_12	Myoviridae	100	VC_1927_0	Siphoviridae	53.85	VC_3377_0	Myoviridae	100
VC_1207_17	Unclassified	66.67	VC_1929_0	Unclassified	50	VC_3462_0	Siphoviridae	50

VC_1208_0	Unclassified	100	VC_1952_0	Unclassified	43.75	VC_3464_0	Podoviridae	100
VC_1243_0	Siphoviridae	100	VC_1955_0	Podoviridae	100	VC_3476_0	Myoviridae	81.82
VC_1252_0	Siphoviridae	62.5	VC_1968_1	Unclassified	100	VC_3483_0	Unclassified	50
VC_1254_0	Siphoviridae	77.78	VC_1978_0	Mimiviridae	68.42	VC_3558_0	Unclassified	92.31
VC_1259_0	Siphoviridae	42.86	VC_1982_0	Podoviridae	100	VC_3564_0	Unclassified	100
VC_1265_0	Siphoviridae	76.47	VC_1983_0	Podoviridae	100	VC_3583_0	Mimiviridae	66.67
VC_1273_0	Unclassified	100	VC_1985_0	Podoviridae	100	VC_3599_0	Siphoviridae	88.89
VC_1274_0	Myoviridae	50	VC_2010_0	Unclassified	100	VC_3601_0	Siphoviridae	100
VC_1275_0	Unclassified	87.5	VC_2033_0	Myoviridae	45.45	VC_3613_0	Podoviridae	61.11
VC_1278_0	Unclassified	85.11	VC_2044_0	Unclassified	55.56	VC_3641_0	Myoviridae	100
VC_1301_0	Unclassified	100	VC_2046_0	Myoviridae	77.27	VC_3688_0	Podoviridae	85.71
VC_1336_0	Myoviridae	52.63	VC_2056_0	Siphoviridae	50	VC_369_0	Myoviridae	81.48
VC_1342_0	Myoviridae	61.54	VC_2072_0	Unclassified	33.33	VC_370_0	Siphoviridae	53.33
VC_1348_0	Unclassified	95.24	VC_2141_0	Podoviridae	100	VC_3729_0	Myoviridae	84.62
VC_1363_0	Unclassified	37.93	VC_2218_0	Unclassified	52.94	VC_373_0	Siphoviridae	100
VC_1367_0	Unclassified	35.29	VC_2247_0	Siphoviridae	53.85	VC_3731_0	Siphoviridae	50
VC_1393_2	Siphoviridae	100	VC_2266_0	Unclassified	100	VC_376_0	Siphoviridae	95.45
VC_1408_0	Mimiviridae	47.62	VC_2266_7	Unclassified	100	VC_3769_0	Unclassified	90.91
VC_1418_0	Mimiviridae	41.67	VC_2272_0	Siphoviridae	66.67	VC_386_0	Siphoviridae	68.75
VC_1419_0	Mimiviridae	46.15	VC_2321_0	Podoviridae	100	VC_387_0	Siphoviridae	100
VC_1429_0	Mimiviridae	58.82	VC_2330_1	Unclassified	100	VC_3895_0	Siphoviridae	100
VC_1430_0	Podoviridae	100	VC_2352_0	Unclassified	75	VC_3999_0	Siphoviridae	50
VC_1468_0	Unclassified	100	VC_2366_0	Siphoviridae	56.41	VC_4001_0	Unclassified	90
VC_1470_0	Unclassified	92	VC_2410_0	Unclassified	100	VC_403_0	Myoviridae	57.14
VC_1475_0	Unclassified	53.85	VC_2420_0	Unclassified	44.44	VC_4099_0	Unclassified	100
VC_1477_0	Siphoviridae	77.78	VC_2453_0	Myoviridae	100	VC_4116_0	Unclassified	53.85
VC_1482_0	Unclassified	66.67	VC_2460_0	Phycodnaviridae	50	VC_4130_0	Unclassified	100
VC_1485_0	Siphoviridae	38.46	VC_2462_0	Siphoviridae	85.71	VC_4245_0	Myoviridae	37.5
VC_1486_0	Siphoviridae	80	VC_2519_0	Unclassified	100	VC_439_0	Podoviridae	100

VC_1495_0	Siphoviridae	70.59	VC_2542_0	Myoviridae	100	VC_445_0	Siphoviridae	75
VC_1503_0	Unclassified	60	VC_2555_0	Myoviridae	95.12	VC_4499_0	Siphoviridae	100
VC_1504_5	Siphoviridae	90.91	VC_2559_0	Siphoviridae	89.29	VC_4541_0	Unclassified	72.73
VC_1510_0	Unclassified	50	VC_2572_0	Siphoviridae	44.44	VC_455_0	Myoviridae	83.33
VC_1538_2	Mimiviridae	71.43	VC_2589_0	Myoviridae	46.15	VC_4576_0	Myoviridae	69.57
VC_1539_0	Siphoviridae	33.33	VC_2600_0	Podoviridae	100	VC_4621_0	Podoviridae	100
VC_1543_0	Mimiviridae	42.11	VC_2644_0	Myoviridae	81.25	VC_4643_0	Siphoviridae	100
VC_1610_0	Siphoviridae	100	VC_2667_0	Podoviridae	100	VC_4698_0	Unclassified	100
VC_1616_0	Myoviridae	71.43	VC_2678_0	Myoviridae	53.85	VC_4709_0	Unclassified	80
VC_1632_0	Myoviridae	48.15	VC_2679_0	Unclassified	66.67	VC_500_0	Myoviridae	100
VC_1640_0	Myoviridae	100	VC_2698_0	Mimiviridae	50	VC_513_0	Siphoviridae	100
VC_1641_0	Siphoviridae	55.56	VC_2751_0	Mimiviridae	76.92	VC_5183_0	Siphoviridae	50
VC_1643_0	Mimiviridae	40	VC_2753_0	Myoviridae	90	VC_5443_0	Siphoviridae	50
VC_1645_0	Myoviridae	90.91	VC_2754_0	Myoviridae	100	VC_556_0	Myoviridae	100
VC_1646_0	Myoviridae	100	VC_2761_0	Unclassified	66.67	VC_6018_0	Unclassified	42.86
VC_1656_0	Mimiviridae	40	VC_2822_0	Myoviridae	80.95	VC_6041_0	Myoviridae	83.33
VC_1661_0	Myoviridae	76.47	VC_2825_0	Myoviridae	100	VC_638_0	Unclassified	100
VC_1662_0	Unclassified	55.56	VC_2838_0	Unclassified	100	VC_639_0	Siphoviridae	50
VC_1664_0	Siphoviridae	92.86	VC_285_0	Siphoviridae	90.91	VC_640_0	Siphoviridae	90
VC_1690_0	Myoviridae	63.64	VC_2872_0	Unclassified	100	VC_650_1	Unclassified	100
VC_1745_0	Unclassified	42.86	VC_2872_1	Unclassified	100	VC_6994_0	Unclassified	100
VC_1747_0	Siphoviridae	89.47	VC_2875_0	Unclassified	100	VC_734_1	Unclassified	100
VC_1757_0	Myoviridae	81.82	VC_290_0	Siphoviridae	100	VC_839_0	Siphoviridae	73.33
VC_1761_0	Unclassified	88.89	VC_2916_0	Myoviridae	43.48	VC_839_2	Siphoviridae	41.67
VC_1767_0	Unclassified	94.12	VC_294_0	Siphoviridae	100	VC_881_0	Siphoviridae	100
VC_1770_0	Unclassified	100	VC_2967_0	Siphoviridae	100	VC_907_0	Myoviridae	100
VC_1771_0	Unclassified	100	VC_297_0	Siphoviridae	100	VC_949_0	Myoviridae	66.67
VC_1780_0	Siphoviridae	100	VC_2971_0	Unclassified	85.71			

Supplementary Table 3: Total number of VCs in each family and the abundance of that family at VC level

Family	VCs (N)	Family (% VCs)	Family (Total abundance)	Family (% abundance)
Unclassified	1064	35.9824	46362327.9825	43.7380
Siphoviridae	937	31.6875	29370008.2018	27.7076
Myoviridae	511	17.2810	13800754.6131	13.0196
Podoviridae	167	5.6476	9432402.9225	8.8985
Mimiviridae	207	7.0003	6405537.0124	6.0430
Phycodnaviridae	32	1.0822	238246.3377	0.2248
Microviridae	16	0.5411	170891.0448	0.1612
Caliciviridae	2	0.0676	166611.9063	0.1572
Flaviviridae	1	0.0338	27714.6725	0.0261
Poxviridae	9	0.3044	9917.1676	0.0094
Inoviridae	7	0.2367	9835.9887	0.0093
Baculoviridae	1	0.0338	4852.7203	0.0046
Adenoviridae	1	0.0338	375.9918	0.0004
Ascoviridae	1	0.0338	338.9989	0.0003
Bicaudaviridae	1	0.0338	184.4393	0.0002

Supplementary Table 4: Total number of contigs in each family and the abundance of that family at contig level

Family	Contigs (N)	Family (% Contigs)	Family (Total abundance)	Family (% abundance)
Unclassified	23985	46.8658	49326725.3656	46.5346
Siphoviridae	14323	27.9866	31100154.3072	29.3398
Myoviridae	7247	14.1604	12031047.4785	11.3500
Mimiviridae	2139	4.1795	6323116.1830	5.9652
Podoviridae	2323	4.5391	6167497.4878	5.8184
Phycodnaviridae	635	1.2408	663871.5773	0.6263
Caliciviridae	6	0.0117	157184.9546	0.1483
Poxviridae	109	0.2130	106540.7239	0.1005
Pithoviridae	17	0.0332	43692.6131	0.0412
Microviridae	196	0.3830	38293.7981	0.0361
Inoviridae	60	0.1172	10799.4674	0.0102
Adenoviridae	5	0.0098	7356.8042	0.0069
Baculoviridae	20	0.0391	3874.4952	0.0037
Iridoviridae	7	0.0137	3282.3851	0.0031
Ascoviridae	18	0.0352	3245.2957	0.0031
Flaviviridae	7	0.0137	3234.1504	0.0031
Herpesviridae	13	0.0254	2550.9700	0.0024
Marseilleviridae	23	0.0449	1925.0239	0.0018
Lipothrixviridae	8	0.0156	1405.7568	0.0013
Mesoniviridae	3	0.0059	1387.0627	0.0013
Tectiviridae	5	0.0098	525.8533	0.0005
Ackermannviridae	11	0.0215	460.6748	0.0004
Circoviridae	2	0.0039	409.1619	0.0004
Potyviridae	1	0.0020	363.0157	0.0003
Solemoviridae	3	0.0059	351.9911	0.0003
Autolykiviridae	2	0.0039	249.4669	0.0002

Bicaudaviridae	4	0.0078	204.6099	0.0002
Rudiviridae	3	0.0059	161.4646	0.0002
Picobirnaviridae	2	0.0039	53.2231	0.0001
Retroviridae	1	0.0020	34.6382	0.0000

Supplementary Table 5: Abundance of each viral family in Control and IBS samples based on aggregated VC abundance and the fraction of total abundance accounted for by each group

	Controls		**IBS**	
Family	**Total Abundance**	**% Abundance**	**Total Abundance**	**% Abundance**
Adenoviridae	296.4374	0.0006	79.5544	0.0001
Ascoviridae	0.0000	0.0000	338.9989	0.0006
Baculoviridae	176.4301	0.0003	4676.2902	0.0085
Bicaudaviridae	182.1907	0.0004	2.2486	0.0000
Caliciviridae	0.0000	0.0000	166611.9063	0.3029
Flaviviridae	27714.6725	0.0543	0.0000	0.0000
Inoviridae	7079.5494	0.0139	2756.4392	0.0050
Microviridae	51121.7205	0.1002	119769.3243	0.2178
Mimiviridae	4116041.3380	8.0707	2289495.6750	4.1627
Myoviridae	7673840.6980	15.0467	6126913.9150	11.1398
Phycodnaviridae	181937.9366	0.3567	56308.4011	0.1024
Podoviridae	3331763.7490	6.5329	6100639.1730	11.0921
Poxviridae	5157.7731	0.0101	4759.3944	0.0087
Siphoviridae	14089231.8000	27.6259	15280776.4000	27.7832
Unclassified	21515455.7100	42.1872	24846872.2800	45.1761

Supplementary Table 6: Prevalence of a VC identified as core VCs in any group, along with taxanomic classifications of these VCs and their predicted bacterial hosts. A core VC is defined as a VC with $\geq$ 50 % prevalence in the group.

		IBS (n=55)				Controls (n=51)				All samples (n=106)			
Cluster ID	**Family**	**Prevalence**	**Abundance**	**Prevalence(%)**	**Core VC**	**Prevalence**	**Abundance**	**Prevalence(%)**	**Core VC**	**Prevalence**	**Abundance**	**Prevalence(%)**	**Core VC**
VC_1058_0	Siphoviridae	27	169552.25	49.09	no	29	74709.23	56.86	yes	56	244261.48	52.83	yes
VC_1138_0	Siphoviridae	26	101522.8	47.27	no	29	65794	56.86	yes	55	167316.8	51.89	yes
VC_1173_0	Podoviridae	19	245365.21	34.55	no	28	145052.62	54.9	yes	47	390417.83	44.34	no
VC_1207_12	Myoviridae	41	33391.96	74.55	yes	35	27439.55	68.63	yes	76	60831.51	71.7	yes
VC_1278_0	Unclassified	24	28537.72	43.64	no	27	38781.36	52.94	yes	51	67319.08	48.11	no
VC_1510_0	Unclassified	25	20728.16	45.45	no	26	39167.81	50.98	yes	51	59895.97	48.11	no
VC_1770_0	Unclassified	29	2238660.85	52.73	yes	23	676295.51	45.1	no	52	2914956.35	49.06	no
VC_1853_0	Unclassified	40	35517.58	72.73	yes	36	31971.72	70.59	yes	76	67489.3	71.7	yes
VC_1982_0	Podoviridae	25	1030075.41	45.45	no	27	96358.92	52.94	yes	52	1126434.33	49.06	no
VC_1983_0	Podoviridae	33	109444.34	60	yes	26	81713.62	50.98	yes	59	191157.95	55.66	yes
VC_2555_0	Myoviridae	19	13697.18	34.55	no	27	20128.94	52.94	yes	46	33826.12	43.4	no
VC_2838_0	Unclassified	17	5866.27	30.91	no	26	3161.77	50.98	yes	43	9028.04	40.57	no
VC_513_0	Siphoviridae	33	471616.73	60	yes	39	536480.16	76.47	yes	72	1008096.9	67.92	yes

Supplementary Table 7: Prevalence of each VC in IBS, Controls and all samples with core VCs for each group, defined as VCs with ≥ 50 % prevalence in the group, indicated in last columns. This is a truncated table with information of viral clusters with prevalence of greater than 10% in IBS or Controls. For complete table, please refer to the published article (doi: 10.1080/19490976.2021.1887719).

		IBS (n=55)				Controls (n=51)				All samples (n=106)			
Cluster ID	**Family**	**Prevalence**	**Abundance**	**Prevalence (%)**	**Core VC**	**Prevalence**	**Abundance**	**Prevalence (%)**	**Core VC**	**Prevalence**	**Abundance**	**Prevalence (%)**	**Core VC**
VC_1058_0	Siphoviridae	27	169552.2	49.1	no	29	74709.2	56.9	yes	56	244261.5	52.8	yes
VC_1078_0	Siphoviridae	10	1612.1	18.2	no	12	3553.2	23.5	no	22	5165.3	20.8	no
VC_1086_0	Siphoviridae	10	19237.2	18.2	no	18	9629.1	35.3	no	28	28866.2	26.4	no
VC_1087_0	Siphoviridae	10	8190.3	18.2	no	16	3541.6	31.4	no	26	11731.8	24.5	no
VC_1091_0	Siphoviridae	2	15916	3.6	no	11	8698	21.6	no	13	24614.1	12.3	no
VC_1116_0	Podoviridae	10	1024540.4	18.2	no	11	45935.9	21.6	no	21	1070476.3	19.8	no
VC_1138_0	Siphoviridae	26	101522.8	47.3	no	29	65794	56.9	yes	55	167316.8	51.9	yes
VC_1159_0	Myoviridae	8	24502.8	14.5	no	12	29066	23.5	no	20	53568.8	18.9	no
VC_1173_0	Podoviridae	19	245365.2	34.5	no	28	145052.6	54.9	yes	47	390417.8	44.3	no
VC_1177_0	Unclassified	12	175747.1	21.8	no	19	66089.9	37.3	no	31	241837	29.2	no
VC_1182_0	Myoviridae	8	1811.8	14.5	no	13	13226.6	25.5	no	21	15038.4	19.8	no
VC_1184_0	Siphoviridae	8	4869.1	14.5	no	15	47678.5	29.4	no	23	52547.6	21.7	no
VC_1189_0	Unclassified	11	488509.2	20	no	14	31542.9	27.5	no	25	520052.1	23.6	no
VC_1196_3	Siphoviridae	13	83460.4	23.6	no	18	104399.3	35.3	no	31	187859.7	29.2	no
VC_1205_0	Unclassified	2	1674	3.6	no	11	26075	21.6	no	13	27749	12.3	no

VC_1207_12	Myoviridae	41	33392	74.5	yes	35	27439.6	68.6	yes	76	60831.5	71.7	yes
VC_1207_17	Unclassified	8	817	14.5	no	13	8092.9	25.5	no	21	8909.9	19.8	no
VC_1208_0	Unclassified	11	89054.1	20	no	14	23254.2	27.5	no	25	112308.3	23.6	no
VC_1243_0	Siphoviridae	8	797.8	14.5	no	12	2228.4	23.5	no	20	3026.2	18.9	no
VC_1252_0	Siphoviridae	9	1351	16.4	no	11	1166.8	21.6	no	20	2517.8	18.9	no
VC_1254_0	Siphoviridae	7	12126.8	12.7	no	14	4596.3	27.5	no	21	16723	19.8	no
VC_1259_0	Siphoviridae	11	10043.9	20	no	22	11889.5	43.1	no	33	21933.4	31.1	no
VC_1265_0	Siphoviridae	16	50070.6	29.1	no	12	3797.1	23.5	no	28	53867.7	26.4	no
VC_1273_0	Unclassified	15	231588.3	27.3	no	20	245994.3	39.2	no	35	477582.6	33	no
VC_1274_0	Myoviridae	18	31848.7	32.7	no	12	71854.3	23.5	no	30	103703	28.3	no
VC_1275_0	Unclassified	5	1228.4	9.1	no	12	4710	23.5	no	17	5938.4	16	no
VC_1278_0	Unclassified	24	28537.7	43.6	no	27	38781.4	52.9	yes	51	67319.1	48.1	no
VC_1301_0	Unclassified	16	20866.9	29.1	no	19	58144.1	37.3	no	35	79011.1	33	no
VC_1336_0	Myoviridae	9	48191.4	16.4	no	20	350434.9	39.2	no	29	398626.3	27.4	no
VC_1342_0	Myoviridae	15	125593.1	27.3	no	18	561095.1	35.3	no	33	686688.2	31.1	no
VC_1348_0	Unclassified	10	66667.4	18.2	no	21	458060.1	41.2	no	31	524727.5	29.2	no
VC_1363_0	Unclassified	19	25634	34.5	no	25	188258.7	49	no	44	213892.7	41.5	no
VC_1367_0	Unclassified	13	20614.3	23.6	no	15	70139.1	29.4	no	28	90753.4	26.4	no
VC_1393_2	Siphoviridae	7	14054.1	12.7	no	13	33939.2	25.5	no	20	47993.4	18.9	no
VC_1408_0	Mimiviridae	12	33970.7	21.8	no	20	69852.3	39.2	no	32	103823	30.2	no
VC_1418_0	Mimiviridae	5	12388.1	9.1	no	11	64429.3	21.6	no	16	76817.4	15.1	no

VC_1419_0	Mimiviridae	4	15532.8	7.3	no	14	51096.1	27.5	no	18	66629	17	no
VC_1429_0	Mimiviridae	18	15236.4	32.7	no	24	102764.9	47.1	no	42	118001.3	39.6	no
VC_1430_0	Podoviridae	12	15439	21.8	no	16	75319.7	31.4	no	28	90758.7	26.4	no
VC_1468_0	Unclassified	7	3678.7	12.7	no	11	3953.2	21.6	no	18	7631.9	17	no
VC_1470_0	Unclassified	13	62046.3	23.6	no	22	81116.9	43.1	no	35	143163.2	33	no
VC_1475_0	Unclassified	15	42784.2	27.3	no	17	53978.7	33.3	no	32	96762.9	30.2	no
VC_1477_0	Siphoviridae	7	163516.5	12.7	no	14	97845.5	27.5	no	21	261362	19.8	no
VC_1482_0	Unclassified	13	81176.7	23.6	no	7	12746.3	13.7	no	20	93923	18.9	no
VC_1485_0	Siphoviridae	8	178702.1	14.5	no	16	113090.9	31.4	no	24	291793	22.6	no
VC_1486_0	Siphoviridae	11	11217.2	20	no	12	7708.7	23.5	no	23	18925.9	21.7	no
VC_1495_0	Siphoviridae	14	45654	25.5	no	16	49818.1	31.4	no	30	95472.1	28.3	no
VC_1503_0	Unclassified	11	89588.4	20	no	13	65225	25.5	no	24	154813.4	22.6	no
VC_1504_5	Siphoviridae	14	14122.4	25.5	no	12	21982.6	23.5	no	26	36105.1	24.5	no
VC_1510_0	Unclassified	25	20728.2	45.5	no	26	39167.8	51	yes	51	59896	48.1	no
VC_1538_2	Mimiviridae	12	94285.9	21.8	no	9	101776.4	17.6	no	21	196062.3	19.8	no
VC_1539_0	Siphoviridae	13	196475.1	23.6	no	11	48687.5	21.6	no	24	245162.6	22.6	no
VC_1543_0	Mimiviridae	13	46110.2	23.6	no	19	83307.6	37.3	no	32	129417.8	30.2	no
VC_1610_0	Siphoviridae	13	14220.5	23.6	no	12	3420	23.5	no	25	17640.5	23.6	no
VC_1616_0	Myoviridae	3	5104.3	5.5	no	13	28195.6	25.5	no	16	33300	15.1	no
VC_1632_0	Myoviridae	16	2696.3	29.1	no	23	18123.8	45.1	no	39	20820.1	36.8	no
VC_1640_0	Myoviridae	11	56701.3	20	no	14	93945.3	27.5	no	25	150646.6	23.6	no

VC_1641_0	Siphoviridae	9	63317.7	16.4	no	12	205628.1	23.5	no	21	268945.7	19.8	no
VC_1643_0	Mimiviridae	7	11708.3	12.7	no	11	27005.7	21.6	no	18	38714	17	no
VC_1645_0	Myoviridae	16	47122.6	29.1	no	8	23708.2	15.7	no	24	70830.7	22.6	no
VC_1646_0	Myoviridae	11	14321.9	20	no	12	22972.5	23.5	no	23	37294.4	21.7	no
VC_1656_0	Mimiviridae	13	12938.6	23.6	no	11	17026.3	21.6	no	24	29965	22.6	no
VC_1661_0	Myoviridae	20	188721.5	36.4	no	20	25560.6	39.2	no	40	214282.2	37.7	no
VC_1662_0	Unclassified	9	1457.1	16.4	no	11	2819	21.6	no	20	4276.1	18.9	no
VC_1664_0	Siphoviridae	14	14900.5	25.5	no	7	5256.3	13.7	no	21	20156.9	19.8	no
VC_1690_0	Myoviridae	8	22391.5	14.5	no	12	178387.7	23.5	no	20	200779.2	18.9	no
VC_1745_0	Unclassified	16	18422.4	29.1	no	17	22447.1	33.3	no	33	40869.6	31.1	no
VC_1747_0	Siphoviridae	14	12250.8	25.5	no	13	12218.4	25.5	no	27	24469.2	25.5	no
VC_1757_0	Myoviridae	8	28992.3	14.5	no	11	66936.1	21.6	no	19	95928.5	17.9	no
VC_1761_0	Unclassified	13	31990.6	23.6	no	17	46545.8	33.3	no	30	78536.4	28.3	no
VC_1767_0	Unclassified	11	147682	20	no	14	125726.6	27.5	no	25	273408.6	23.6	no
VC_1770_0	Unclassified	29	2238660.8	52.7	yes	23	676295.5	45.1	no	52	2914956.4	49.1	no
VC_1771_0	Unclassified	10	1226099.9	18.2	no	12	33829.8	23.5	no	22	1259929.7	20.8	no
VC_1780_0	Siphoviridae	11	3207.4	20	no	12	12770.7	23.5	no	23	15978.1	21.7	no
VC_1786_0	Siphoviridae	22	49645.1	40	no	8	59647.6	15.7	no	30	109292.7	28.3	no
VC_1803_0	Myoviridae	6	1184	10.9	no	13	4383.4	25.5	no	19	5567.5	17.9	no
VC_1808_0	Siphoviridae	9	56223.5	16.4	no	13	49263.2	25.5	no	22	105486.6	20.8	no
VC_1811_0	Siphoviridae	12	3645.3	21.8	no	5	395.2	9.8	no	17	4040.5	16	no

VC_1815_0	Unclassified	14	70858.5	25.5	no	14	39761.4	27.5	no	28	110619.9	26.4	no
VC_1853_0	Unclassified	40	35517.6	72.7	yes	36	31971.7	70.6	yes	76	67489.3	71.7	yes
VC_1854_0	Unclassified	10	328869.9	18.2	no	20	435789.3	39.2	no	30	764659.2	28.3	no
VC_1857_1	Siphoviridae	14	41583.2	25.5	no	17	28273.6	33.3	no	31	69856.9	29.2	no
VC_1858_0	Myoviridae	17	16136.4	30.9	no	16	28226.2	31.4	no	33	44362.6	31.1	no
VC_1885_0	Unclassified	12	39396.9	21.8	no	11	38820.4	21.6	no	23	78217.3	21.7	no
VC_1890_1	Unclassified	12	83030	21.8	no	12	159864.3	23.5	no	24	242894.3	22.6	no
VC_1893_0	Unclassified	10	33787.8	18.2	no	13	41217.1	25.5	no	23	75005	21.7	no
VC_1896_0	Podoviridae	17	25814	30.9	no	10	38013.9	19.6	no	27	63827.9	25.5	no
VC_1910_0	Mimiviridae	12	45771.6	21.8	no	15	68458	29.4	no	27	114229.7	25.5	no
VC_1911_0	Myoviridae	8	57086.2	14.5	no	14	35225.4	27.5	no	22	92311.6	20.8	no
VC_1927_0	Siphoviridae	4	17975.5	7.3	no	12	66069.6	23.5	no	16	84045.1	15.1	no
VC_1929_0	Unclassified	15	11221	27.3	no	9	15050.4	17.6	no	24	26271.5	22.6	no
VC_1952_0	Unclassified	14	42950.5	25.5	no	24	272434.5	47.1	no	38	315385	35.8	no
VC_1955_0	Podoviridae	14	11267	25.5	no	19	16071	37.3	no	33	27338	31.1	no
VC_1968_1	Unclassified	11	49344.3	20	no	11	24202.1	21.6	no	22	73546.3	20.8	no
VC_1978_0	Mimiviridae	6	81715.1	10.9	no	12	107347.6	23.5	no	18	189062.8	17	no
VC_1982_0	Podoviridae	25	1030075.4	45.5	no	27	96358.9	52.9	yes	52	1126434.3	49.1	no
VC_1983_0	Podoviridae	33	109444.3	60	yes	26	81713.6	51	yes	59	191158	55.7	yes
VC_1985_0	Podoviridae	4	1623.6	7.3	no	11	28334.2	21.6	no	15	29957.7	14.2	no
VC_2010_0	Unclassified	7	7043.6	12.7	no	16	101144.8	31.4	no	23	108188.4	21.7	no

VC_2033_0	Myoviridae	14	52844.5	25.5	no	12	56079.8	23.5	no	26	108924.3	24.5	no
VC_2044_0	Unclassified	7	29162.7	12.7	no	11	26221.4	21.6	no	18	55384.1	17	no
VC_2046_0	Myoviridae	16	49579.6	29.1	no	13	20443.1	25.5	no	29	70022.7	27.4	no
VC_2056_0	Siphoviridae	9	7465.2	16.4	no	19	37012	37.3	no	28	44477.2	26.4	no
VC_2072_0	Unclassified	19	280925	34.5	no	18	296978.2	35.3	no	37	577903.1	34.9	no
VC_2141_0	Podoviridae	20	36601.6	36.4	no	19	19684.5	37.3	no	39	56286.1	36.8	no
VC_2218_0	Unclassified	16	9204.2	29.1	no	17	20224.9	33.3	no	33	29429.2	31.1	no
VC_2247_0	Siphoviridae	8	1319.8	14.5	no	11	6519.2	21.6	no	19	7839	17.9	no
VC_2266_0	Unclassified	11	79976.8	20	no	11	32121.1	21.6	no	22	112098	20.8	no
VC_2266_7	Unclassified	5	57160.5	9.1	no	11	165444.9	21.6	no	16	222605.4	15.1	no
VC_2272_0	Siphoviridae	14	96298.2	25.5	no	12	139026	23.5	no	26	235324.2	24.5	no
VC_2321_0	Podoviridae	18	46613.9	32.7	no	15	37196	29.4	no	33	83809.9	31.1	no
VC_2330_1	Unclassified	14	584301.7	25.5	no	13	159440.7	25.5	no	27	743742.4	25.5	no
VC_2352_0	Unclassified	23	293449.4	41.8	no	22	201660.9	43.1	no	45	495110.2	42.5	no
VC_2366_0	Siphoviridae	23	508846.8	41.8	no	22	182227.8	43.1	no	45	691074.6	42.5	no
VC_2410_0	Unclassified	10	229764.1	18.2	no	13	169524.1	25.5	no	23	399288.2	21.7	no
VC_2420_0	Unclassified	5	6465.2	9.1	no	12	58740.7	23.5	no	17	65205.9	16	no
VC_2453_0	Myoviridae	11	4529.3	20	no	12	8683.1	23.5	no	23	13212.4	21.7	no
VC_2460_0	Phycodnaviridae	8	13804.4	14.5	no	11	16751	21.6	no	19	30555.4	17.9	no
VC_2462_0	Siphoviridae	11	6019.8	20	no	13	55535.7	25.5	no	24	61555.5	22.6	no
VC_2519_0	Unclassified	20	145686.5	36.4	no	18	49361.2	35.3	no	38	195047.7	35.8	no

VC_2542_0	Myoviridae	6	14404.5	10.9	no	15	20974.8	29.4	no	21	35379.3	19.8	no
VC_2555_0	Myoviridae	19	13697.2	34.5	no	27	20128.9	52.9	yes	46	33826.1	43.4	no
VC_2559_0	Siphoviridae	16	224924.7	29.1	no	21	72598.3	41.2	no	37	297523	34.9	no
VC_2572_0	Siphoviridae	10	1596.1	18.2	no	11	2563.1	21.6	no	21	4159.2	19.8	no
VC_2589_0	Myoviridae	14	11267.5	25.5	no	13	35790.1	25.5	no	27	47057.5	25.5	no
VC_2600_0	Podoviridae	11	12423.7	20	no	24	74023.2	47.1	no	35	86446.9	33	no
VC_2644_0	Myoviridae	14	2204.1	25.5	no	5	942.7	9.8	no	19	3146.8	17.9	no
VC_2667_0	Podoviridae	25	26791.6	45.5	no	19	63938.8	37.3	no	44	90730.4	41.5	no
VC_2678_0	Myoviridae	16	69131.8	29.1	no	16	23779.3	31.4	no	32	92911.1	30.2	no
VC_2679_0	Unclassified	6	1717.6	10.9	no	15	9129.3	29.4	no	21	10846.9	19.8	no
VC_2698_0	Mimiviridae	6	14442.5	10.9	no	13	37691	25.5	no	19	52133.5	17.9	no
VC_2751_0	Mimiviridae	5	10710.7	9.1	no	15	132000	29.4	no	20	142710.7	18.9	no
VC_2753_0	Myoviridae	10	30451.6	18.2	no	16	49791.8	31.4	no	26	80243.3	24.5	no
VC_2754_0	Myoviridae	11	18380.1	20	no	16	4174.1	31.4	no	27	22554.1	25.5	no
VC_2761_0	Unclassified	13	14526.6	23.6	no	23	76349.3	45.1	no	36	90875.8	34	no
VC_2822_0	Myoviridae	15	6869.6	27.3	no	18	12680.1	35.3	no	33	19549.7	31.1	no
VC_2825_0	Myoviridae	15	31913.9	27.3	no	16	157946.8	31.4	no	31	189860.7	29.2	no
VC_2838_0	Unclassified	17	5866.3	30.9	no	26	3161.8	51	yes	43	9028	40.6	no
VC_285_0	Siphoviridae	11	213658.4	20	no	14	54408.8	27.5	no	25	268067.3	23.6	no
VC_2872_0	Unclassified	12	341342	21.8	no	17	279808.6	33.3	no	29	621150.6	27.4	no
VC_2872_1	Unclassified	8	486497.8	14.5	no	11	608451	21.6	no	19	1094948.8	17.9	no

VC_2875_0	Unclassified	18	152681.2	32.7	no	21	129405.1	41.2	no	39	282086.3	36.8	no
VC_290_0	Siphoviridae	16	320109.4	29.1	no	21	369581.1	41.2	no	37	689690.4	34.9	no
VC_2916_0	Myoviridae	15	20665.8	27.3	no	16	170024.9	31.4	no	31	190690.8	29.2	no
VC_294_0	Siphoviridae	14	211453.4	25.5	no	14	69229.5	27.5	no	28	280683	26.4	no
VC_2967_0	Siphoviridae	6	32002.3	10.9	no	12	111761.5	23.5	no	18	143763.7	17	no
VC_297_0	Siphoviridae	20	160746	36.4	no	25	173498.7	49	no	45	334244.7	42.5	no
VC_2971_0	Unclassified	14	40594.2	25.5	no	13	21102.4	25.5	no	27	61696.6	25.5	no
VC_3018_0	Siphoviridae	14	53630.3	25.5	no	15	560861	29.4	no	29	614491.3	27.4	no
VC_3030_0	Unclassified	16	4028.7	29.1	no	23	10386	45.1	no	39	14414.7	36.8	no
VC_3035_0	Siphoviridae	10	10176	18.2	no	13	32063.6	25.5	no	23	42239.6	21.7	no
VC_3037_0	Podoviridae	15	4534.7	27.3	no	10	3586.8	19.6	no	25	8121.5	23.6	no
VC_3076_0	Podoviridae	8	2165.2	14.5	no	11	6711.3	21.6	no	19	8876.5	17.9	no
VC_3099_0	Unclassified	18	4373.1	32.7	no	12	5504.5	23.5	no	30	9877.6	28.3	no
VC_3141_0	Unclassified	6	7168.7	10.9	no	11	16902.9	21.6	no	17	24071.6	16	no
VC_3195_0	Siphoviridae	11	109874.3	20	no	12	336080.1	23.5	no	23	445954.4	21.7	no
VC_3206_0	Siphoviridae	11	4579.9	20	no	12	6185.1	23.5	no	23	10765	21.7	no
VC_3211_0	Siphoviridae	16	453176.6	29.1	no	17	462757.9	33.3	no	33	915934.5	31.1	no
VC_3316_0	Unclassified	4	2940.8	7.3	no	13	13073.3	25.5	no	17	16014.2	16	no
VC_3320_0	Unclassified	5	11279.4	9.1	no	12	9650.2	23.5	no	17	20929.7	16	no
VC_3341_0	Unclassified	6	11142.8	10.9	no	11	33023.5	21.6	no	17	44166.2	16	no
VC_3359_0	Myoviridae	9	21081.2	16.4	no	11	809550.7	21.6	no	20	830631.9	18.9	no
VC_3361_0	Unclassified	16	195512.5	29.1	no	9	17085.9	17.6	no	25	212598.4	23.6	no

VC_3377_0	Myoviridae	8	6087.4	14.5	no	16	26678.5	31.4	no	24	32765.9	22.6	no
VC_3462_0	Siphoviridae	9	8343.2	16.4	no	20	28346.5	39.2	no	29	36689.7	27.4	no
VC_3464_0	Podoviridae	15	21722.5	27.3	no	13	23414.5	25.5	no	28	45137.1	26.4	no
VC_3476_0	Myoviridae	10	2022.4	18.2	no	12	19891.1	23.5	no	22	21913.5	20.8	no
VC_3483_0	Unclassified	13	21331.9	23.6	no	14	18558	27.5	no	27	39890	25.5	no
VC_3558_0	Unclassified	5	257.5	9.1	no	11	1921.5	21.6	no	16	2179	15.1	no
VC_3564_0	Unclassified	9	24867.3	16.4	no	11	1368.5	21.6	no	20	26235.8	18.9	no
VC_3583_0	Mimiviridae	9	2050.9	16.4	no	12	16970.7	23.5	no	21	19021.6	19.8	no
VC_3599_0	Siphoviridae	8	1764.1	14.5	no	11	2087.3	21.6	no	19	3851.4	17.9	no
VC_3601_0	Siphoviridae	13	11827.1	23.6	no	7	8307	13.7	no	20	20134.1	18.9	no
VC_3613_0	Podoviridae	11	26396.5	20	no	13	11986.3	25.5	no	24	38382.8	22.6	no
VC_3641_0	Myoviridae	12	783.2	21.8	no	3	777.3	5.9	no	15	1560.5	14.2	no
VC_3688_0	Podoviridae	12	31349.9	21.8	no	19	122852.1	37.3	no	31	154202	29.2	no
VC_369_0	Myoviridae	17	67581.1	30.9	no	14	51891.1	27.5	no	31	119472.2	29.2	no
VC_370_0	Siphoviridae	10	6425.8	18.2	no	11	3023.5	21.6	no	21	9449.4	19.8	no
VC_3729_0	Myoviridae	11	19712.6	20	no	15	33508.6	29.4	no	26	53221.2	24.5	no
VC_373_0	Siphoviridae	16	7268.2	29.1	no	13	177081.3	25.5	no	29	184349.5	27.4	no
VC_3731_0	Siphoviridae	9	2061.5	16.4	no	14	4892.8	27.5	no	23	6954.3	21.7	no
VC_376_0	Siphoviridae	12	87350.4	21.8	no	12	29525.6	23.5	no	24	116876	22.6	no
VC_3769_0	Unclassified	13	5151.6	23.6	no	10	3428.2	19.6	no	23	8579.8	21.7	no
VC_386_0	Siphoviridae	19	147265	34.5	no	24	110856.6	47.1	no	43	258121.6	40.6	no
VC_387_0	Siphoviridae	6	4153.8	10.9	no	12	63595.4	23.5	no	18	67749.2	17	no
VC_3895_0	Siphoviridae	5	11128.1	9.1	no	13	12770.7	25.5	no	18	23898.8	17	no

VC_3999_0	Siphoviridae	11	787	20	no	13	1051.3	25.5	no	24	1838.3	22.6	no
VC_4001_0	Unclassified	5	12509.9	9.1	no	11	29761.1	21.6	no	16	42271	15.1	no
VC_403_0	Myoviridae	8	17704	14.5	no	17	52109.7	33.3	no	25	69813.7	23.6	no
VC_4099_0	Unclassified	14	477476.1	25.5	no	18	436751	35.3	no	32	914227.2	30.2	no
VC_4116_0	Unclassified	13	13816.8	23.6	no	6	7256.7	11.8	no	19	21073.5	17.9	no
VC_4130_0	Unclassified	20	21110.9	36.4	no	20	4445.1	39.2	no	40	25556	37.7	no
VC_4245_0	Myoviridae	8	21252.5	14.5	no	12	7248.4	23.5	no	20	28500.9	18.9	no
VC_439_0	Podoviridae	8	27925.4	14.5	no	11	20883	21.6	no	19	48808.3	17.9	no
VC_445_0	Siphoviridae	12	1373.5	21.8	no	11	10491.6	21.6	no	23	11865.1	21.7	no
VC_4499_0	Siphoviridae	10	7101.6	18.2	no	12	9941.3	23.5	no	22	17042.9	20.8	no
VC_4541_0	Unclassified	4	459.6	7.3	no	12	2434.4	23.5	no	16	2894	15.1	no
VC_455_0	Myoviridae	12	71945	21.8	no	9	7710.8	17.6	no	21	79655.8	19.8	no
VC_4576_0	Myoviridae	19	26442.7	34.5	no	14	8063.5	27.5	no	33	34506.2	31.1	no
VC_4621_0	Podoviridae	13	2633.2	23.6	no	18	35497.9	35.3	no	31	38131.1	29.2	no
VC_4643_0	Siphoviridae	4	109444.9	7.3	no	11	20802.3	21.6	no	15	130247.2	14.2	no
VC_4698_0	Unclassified	6	4442.4	10.9	no	13	63518.9	25.5	no	19	67961.3	17.9	no
VC_4709_0	Unclassified	4	28974.7	7.3	no	12	19869.9	23.5	no	16	48844.6	15.1	no
VC_500_0	Myoviridae	14	35308	25.5	no	21	282999.9	41.2	no	35	318307.9	33	no
VC_513_0	Siphoviridae	33	471616.7	60	yes	39	536480.2	76.5	yes	72	1008096.9	67.9	yes
VC_5183_0	Siphoviridae	6	942.8	10.9	no	12	3430.2	23.5	no	18	4373	17	no
VC_5443_0	Siphoviridae	5	4807.5	9.1	no	14	7234.9	27.5	no	19	12042.4	17.9	no
VC_556_0	Myoviridae	9	1476.4	16.4	no	12	12963.4	23.5	no	21	14439.8	19.8	no
VC_6018_0	Unclassified	11	2020.6	20	no	13	2529.4	25.5	no	24	4550	22.6	no

VC_6041_0	Myoviridae	10	8776.7	18.2	no	12	4643.4	23.5	no	22	13420.1	20.8	no
VC_638_0	Unclassified	9	724862.5	16.4	no	14	153963.8	27.5	no	23	878826.3	21.7	no
VC_639_0	Siphoviridae	6	15479.1	10.9	no	11	22966	21.6	no	17	38445.1	16	no
VC_640_0	Siphoviridae	4	42272	7.3	no	11	130352	21.6	no	15	172624	14.2	no
VC_650_1	Unclassified	8	16537.3	14.5	no	12	36511.3	23.5	no	20	53048.6	18.9	no
VC_6994_0	Unclassified	15	4497.5	27.3	no	15	3666.4	29.4	no	30	8163.9	28.3	no
VC_734_1	Unclassified	21	446219.9	38.2	no	20	1218835.8	39.2	no	41	1665055.6	38.7	no
VC_839_0	Siphoviridae	12	21221	21.8	no	15	56976	29.4	no	27	78197	25.5	no
VC_839_2	Siphoviridae	9	12196.4	16.4	no	11	43438.6	21.6	no	20	55635	18.9	no
VC_881_0	Siphoviridae	13	7824.5	23.6	no	11	1422.9	21.6	no	24	9247.4	22.6	no
VC_907_0	Myoviridae	3	1370.3	5.5	no	11	7166.3	21.6	no	14	8536.6	13.2	no
VC_949_0	Myoviridae	13	278355.5	23.6	no	18	64021	35.3	no	31	342376.6	29.2	no

Supplementary Table 8: VCs identified as significantly different between IBS and Control subjects using DeSeq2. Positive log2 fold changes indicate an increased abundance in IBS subjects compared with control subjects. Clusters were prefiltered before this analysis to only retain clusters containing contigs that were present in at least 10 % of all samples. The percent prevalence is the number of samples that a particular VC is present in divided by the total number of samples in the group or subgroup and expressed as a percentage. This is a truncated table. For complete table, please refer to the published article (doi: 10.1080/19490976.2021.1887719).

Cluster ID	VC_1389_0	VC_1832_0	VC_1950_0	VC_1982_0	VC_2751_0	VC_2918_0	VC_3089_0	VC_3232_0	VC_5989_0	VC_650_0
Prevalence in IBS (n = 55)	8	3	2	25	5	3	2	9	1	5
Prevalence in IBS-C (n = 21)	5	2	1	14	1	1	2	2	1	2
Prevalence in IBS-D (n = 17)	1	1	1	5	2	1	0	3	0	1
Prevalence in IBS-M (n = 17)	2	0	0	6	2	1	0	4	0	2
Prevalence in Control (n = 51)	4	9	9	27	15	8	9	4	10	8
% Prevalence in IBS (n=55)	14.55	5.45	3.64	45.45	9.09	5.45	3.64	16.36	1.82	9.09
% Prevalence in IBS-C (n=21)	23.81	9.52	4.76	66.67	4.76	4.76	9.52	9.52	4.76	9.52
% Prevalence in IBS-D (n=17)	5.88	5.88	5.88	29.41	11.76	5.88	0.00	17.65	0.00	5.88
% Prevalence in IBS-M (n =17)	11.76	0.00	0.00	35.29	11.76	5.88	0.00	23.53	0.00	11.76
% Prevalence in Control (n=51)	7.84	17.65	17.65	52.94	29.41	15.69	17.65	7.84	19.61	15.69
BaseMean	20.4372	302.2231	32.0387	3174.5670	428.0707	10.9087	41.9678	240.6148	8.6366	1575.2780
Log2 Fold Change	24.9719	-29.9704	-26.9292	3.8114	-6.2916	-23.6878	-27.2476	28.4113	-24.5406	-10.5849
lfcSE	2.6628	2.4137	2.3907	1.0993	1.8238	2.2780	2.2923	2.5626	1.9239	2.6778
Stat	9.3782	-12.4168	-11.2640	3.4672	-3.4498	-10.3985	-11.8863	11.0871	-12.7558	-3.9529
P-Value	6.71E-21	2.12E-35	1.98E-29	0.000526	0.000561	2.52E-25	1.39E-32	1.45E-28	2.90E-37	7.72E-05
BH Adjusted P-Value	5.33E-19	5.89E-33	2.75E-27	0.031195	0.031195	2.33E-23	2.58E-30	1.61E-26	1.61E-34	0.005367

Family	Mimiviridae	Unclassified	Unclassified	Podoviridae	Mimiviridae	Siphoviridae	Siphoviridae	Siphoviridae	Podoviridae	Unclassified
Family (% Votes)	62.5	100	33.33	100	76.92	50	100	75	100	100

Supplementary Table 9: VCs associated with the significant split between IBS and controls identified by Canonical Correlation Analysis with taxonomy information and predicted bacterial host species information for each VC. This is a truncated table. For complete table, please refer to the published article (doi: 10.1080/19490976.2021.1887719).

Cluster ID	SpearmanRho	t-statistic	P-value	Q-value	StandardError	Family	Family (% Votes)
VC_1156_0	-0.363	-3.97	0.00013	0.00568	0.0914	Siphoviridae	100
VC_1156_1	-0.29	-3.09	0.00259	0.0465	0.0939	Siphoviridae	100
VC_1173_0	-0.406	-4.53	1.58E-05	0.000872	0.0896	Podoviridae	100
VC_1180_0	-0.343	-3.73	0.000316	0.0104	0.0921	Unclassified	100
VC_1205_0	-0.321	-3.46	0.000797	0.0223	0.0929	Unclassified	88.89
VC_1207_12	0.302	3.23	0.00166	0.0339	0.0935	Myoviridae	100
VC_1207_4	-0.319	-3.43	0.000864	0.0236	0.0929	Unclassified	100
VC_1248_1	-0.313	-3.36	0.00108	0.0252	0.0931	Siphoviridae	100
VC_1260_0	-0.315	-3.38	0.00101	0.0249	0.0931	Siphoviridae	100
VC_1273_0	-0.302	-3.23	0.00165	0.0339	0.0935	Unclassified	100
VC_1278_0	-0.312	-3.35	0.00114	0.0262	0.0932	Unclassified	85.11
VC_1301_0	-0.317	-3.41	0.000915	0.0242	0.093	Unclassified	100
VC_1327_0	-0.305	-3.27	0.00146	0.0308	0.0934	Mimiviridae	60
VC_1348_0	-0.308	-3.3	0.00134	0.0293	0.0933	Unclassified	95.24
VC_1393_2	-0.311	-3.34	0.00117	0.0266	0.0932	Siphoviridae	100
VC_1394_0	-0.424	-4.78	5.91E-06	0.000378	0.0888	Siphoviridae	100
VC_1418_0	-0.564	-6.96	3.20E-10	7.27E-08	0.081	Mimiviridae	41.67
VC_1419_0	-0.491	-5.75	9.05E-08	1.43E-05	0.0854	Mimiviridae	46.15

VC_1421_0	-0.394	-4.37	2.96E-05	0.00148	0.0901	Mimiviridae	50
VC_1424_0	-0.461	-5.3	6.65E-07	7.17E-05	0.087	Mimiviridae	25
VC_1429_0	-0.345	-3.75	0.000287	0.00963	0.092	Mimiviridae	58.82
VC_1477_0	-0.619	-8.04	1.48E-12	1.01E-09	0.077	Siphoviridae	77.78
VC_1480_0	-0.301	-3.22	0.00173	0.0345	0.0935	Unclassified	66.67
VC_1485_0	-0.612	-7.89	3.17E-12	1.62E-09	0.0775	Siphoviridae	38.46
VC_1514_0	-0.467	-5.39	4.46E-07	5.07E-05	0.0867	Unclassified	87.5
VC_1535_0	-0.532	-6.41	4.34E-09	8.08E-07	0.083	Myoviridae	66.67
VC_1540_0	-0.511	-6.07	2.12E-08	3.62E-06	0.0843	Unclassified	100
VC_1543_0	-0.413	-4.62	1.09E-05	0.000654	0.0893	Mimiviridae	42.11
VC_1546_0	-0.459	-5.26	7.62E-07	7.80E-05	0.0871	Mimiviridae	62.5
VC_1601_0	-0.557	-6.84	5.56E-10	1.14E-07	0.0814	Mimiviridae	78.57
VC_1610_0	-0.292	-3.11	0.00239	0.0436	0.0938	Siphoviridae	100
VC_1616_0	-0.399	-4.44	2.22E-05	0.00116	0.0899	Myoviridae	71.43
VC_1641_0	-0.36	-3.94	0.000147	0.00603	0.0915	Siphoviridae	55.56
VC_1648_0	-0.476	-5.52	2.51E-07	3.21E-05	0.0862	Siphoviridae	100
VC_1680_0	-0.291	-3.1	0.00246	0.0445	0.0938	Unclassified	66.67
VC_1690_0	-0.316	-3.4	0.000971	0.0246	0.093	Myoviridae	63.64
VC_1804_1	-0.454	-5.19	1.05E-06	9.72E-05	0.0874	Myoviridae	42.86
VC_1847_0	-0.343	-3.72	0.000322	0.0104	0.0921	Unclassified	66.67
VC_1853_0	0.305	3.26	0.00149	0.0311	0.0934	Unclassified	77.78
VC_1907_0	-0.453	-5.18	1.09E-06	9.72E-05	0.0874	Unclassified	90.91
VC_1910_0	-0.625	-8.17	7.81E-13	7.99E-10	0.0765	Mimiviridae	42.86
VC_1927_0	-0.356	-3.89	0.000177	0.00684	0.0916	Siphoviridae	53.85
VC_1952_0	-0.299	-3.2	0.00185	0.0356	0.0936	Unclassified	43.75
VC_1972_0	-0.426	-4.81	5.16E-06	0.000341	0.0887	Unclassified	100
VC_1978_0	-0.627	-8.2	6.67E-13	7.99E-10	0.0764	Mimiviridae	68.42
VC_2010_0	-0.471	-5.44	3.55E-07	4.27E-05	0.0865	Unclassified	100
VC_2056_0	-0.357	-3.9	0.000173	0.00681	0.0916	Siphoviridae	50

VC_2060_0	-0.316	-3.4	0.000966	0.0246	0.093	Unclassified	100
VC_2294_0	-0.314	-3.37	0.00105	0.0252	0.0931	Unclassified	100
VC_2313_0	-0.584	-7.34	4.88E-11	2.00E-08	0.0796	Unclassified	85.71
VC_2329_0	-0.3	-3.21	0.00176	0.0346	0.0935	Unclassified	100
VC_2366_0	-0.405	-4.51	1.68E-05	0.000905	0.0897	Siphoviridae	56.41
VC_2400_0	-0.352	-3.83	0.000218	0.00795	0.0918	Myoviridae	63.64
VC_2459_2	-0.348	-3.79	0.000257	0.00877	0.0919	Myoviridae	100
VC_2534_0	-0.457	-5.25	8.25E-07	8.05E-05	0.0872	Unclassified	80
VC_2550_0	-0.449	-5.13	1.36E-06	0.000103	0.0876	Unclassified	66.67
VC_2600_0	-0.35	-3.81	0.000237	0.00835	0.0919	Podoviridae	100
VC_2636_0	-0.377	-4.15	6.74E-05	0.003	0.0908	Unclassified	60
VC_2638_0	-0.479	-5.57	2.03E-07	2.77E-05	0.0861	Unclassified	83.33
VC_2713_0	-0.306	-3.28	0.0014	0.03	0.0933	Myoviridae	71.43
VC_2751_0	-0.575	-7.16	1.20E-10	3.50E-08	0.0803	Mimiviridae	76.92
VC_2770_0	-0.316	-3.39	0.000985	0.0246	0.093	Unclassified	100
VC_2780_0	-0.396	-4.4	2.66E-05	0.00136	0.09	Siphoviridae	66.67
VC_2786_0	-0.384	-4.25	4.73E-05	0.00215	0.0905	Unclassified	62.5
VC_2825_0	-0.43	-4.85	4.29E-06	0.000292	0.0885	Myoviridae	100
VC_285_0	-0.408	-4.56	1.42E-05	0.000809	0.0895	Siphoviridae	90.91
VC_2917_0	-0.444	-5.05	1.92E-06	0.000135	0.0879	Siphoviridae	62.5
VC_2948_0	-0.314	-3.37	0.00106	0.0252	0.0931	Myoviridae	80
VC_3018_0	-0.328	-3.54	0.000605	0.018	0.0926	Siphoviridae	93.75
VC_3049_0	-0.31	-3.33	0.00122	0.0271	0.0932	Unclassified	100
VC_3106_0	-0.309	-3.32	0.00125	0.0276	0.0932	Siphoviridae	100
VC_3107_0	-0.419	-4.7	7.93E-06	0.000492	0.089	Unclassified	100
VC_3145_0	-0.451	-5.15	1.25E-06	9.81E-05	0.0875	Unclassified	100
VC_3194_0	-0.326	-3.51	0.00066	0.0193	0.0927	Siphoviridae	100
VC_3316_0	-0.385	-4.26	4.52E-05	0.00215	0.0905	Unclassified	50
VC_3321_0	-0.385	-4.25	4.68E-05	0.00215	0.0905	Unclassified	60

VC_3455_0	-0.288	-3.07	0.00275	0.049	0.0939	Siphoviridae	100
VC_3462_0	-0.452	-5.16	1.18E-06	9.72E-05	0.0875	Siphoviridae	50
VC_3523_0	-0.338	-3.67	0.000389	0.0125	0.0923	Unclassified	77.78
VC_3731_0	-0.576	-7.18	1.06E-10	3.50E-08	0.0802	Siphoviridae	50
VC_3744_0	-0.317	-3.41	0.000922	0.0242	0.093	Myoviridae	40
VC_3815_0	-0.306	-3.28	0.0014	0.03	0.0933	Siphoviridae	100
VC_3816_0	-0.287	-3.06	0.00283	0.05	0.0939	Siphoviridae	100
VC_3877_0	-0.31	-3.33	0.00121	0.0271	0.0932	Unclassified	75
VC_3903_0	-0.321	-3.46	0.000794	0.0223	0.0929	Siphoviridae	100
VC_3948_0	-0.297	-3.17	0.002	0.0375	0.0936	Unclassified	100
VC_401_0	-0.337	-3.65	0.000416	0.0129	0.0923	Myoviridae	100
VC_403_0	-0.298	-3.18	0.00195	0.0369	0.0936	Myoviridae	57.14
VC_4040_0	-0.314	-3.37	0.00106	0.0252	0.0931	Unclassified	66.67
VC_4121_0	-0.349	-3.8	0.000247	0.00857	0.0919	Siphoviridae	80
VC_4124_0	-0.36	-3.93	0.000153	0.00612	0.0915	Myoviridae	100
VC_4146_0	-0.298	-3.19	0.0019	0.0364	0.0936	Unclassified	100
VC_4172_0	-0.386	-4.26	4.43E-05	0.00215	0.0905	Mimiviridae	50
VC_4185_0	-0.408	-4.56	1.42E-05	0.000809	0.0895	Myoviridae	41.67
VC_4370_0	-0.35	-3.81	0.000236	0.00835	0.0919	Siphoviridae	100
VC_4438_0	-0.334	-3.61	0.000475	0.0145	0.0924	Siphoviridae	100
VC_4441_0	-0.318	-3.42	0.000884	0.0238	0.093	Unclassified	66.67
VC_4541_0	-0.316	-3.39	0.000979	0.0246	0.093	Unclassified	72.73
VC_4621_0	-0.483	-5.63	1.56E-07	2.28E-05	0.0859	Podoviridae	100
VC_4706_0	-0.299	-3.2	0.00184	0.0356	0.0936	Siphoviridae	100
VC_4941_0	-0.293	-3.12	0.00233	0.0434	0.0938	Unclassified	66.67
VC_4959_0	-0.449	-5.12	1.41E-06	0.000103	0.0876	Unclassified	100
VC_5183_0	-0.571	-7.1	1.58E-10	4.05E-08	0.0805	Siphoviridae	50
VC_5241_0	-0.333	-3.6	0.000485	0.0146	0.0925	Microviridae	100
VC_57_1	-0.362	-3.95	0.00014	0.00587	0.0914	Myoviridae	100

VC_5786_0	-0.337	-3.65	0.000416	0.0129	0.0923	Unclassified	66.67
VC_5943_0	-0.302	-3.23	0.00167	0.0339	0.0935	Myoviridae	80
VC_6402_0	-0.356	-3.88	0.000181	0.00687	0.0916	Unclassified	100
VC_6472_0	-0.313	-3.36	0.00107	0.0252	0.0931	Unclassified	66.67
VC_650_1	-0.354	-3.86	2.00E-04	0.00743	0.0917	Unclassified	100
VC_6629_0	-0.292	-3.11	0.00238	0.0436	0.0938	Mimiviridae	50
VC_6722_0	-0.32	-3.44	0.000835	0.0231	0.0929	Siphoviridae	100
VC_735_0	-0.324	-3.49	0.000714	0.0206	0.0928	Unclassified	100
VC_824_3	-0.452	-5.16	1.19E-06	9.72E-05	0.0875	Unclassified	100
VC_860_0	-0.361	-3.95	0.000141	0.00587	0.0914	Unclassified	100
VC_865_0	-0.301	-3.22	0.00173	0.0345	0.0935	Siphoviridae	66.67

Chapter 5

Integrative microbiome analysis shows that loss of inter-kingdom interactions is a potential driver of microbiome alterations in Irritable Bowel Syndrome.

Authors: Anubhav Das, Paul W. O'Toole, Ian B. Jeffery*

*Corresponding author.

Manuscript under preparation

Author contributions: PWOT, IBJ conceived the study idea and design. AD, IBJ, and PWOT drafted and reviewed the manuscript. AD and IBJ, performed the bioinformatics analysis.

AD Contribution:

Data processing: Bacteriome read processing and clustering into zOTUs. Other datasets were generated as part of previous chapters.

Analysis: Multivariate statistical analyses. Procrustes analysis. Inter-kingdom co-occurrence and co-exclusion network generation and analysis. Development of Machine learning models and subsequent analysis.

Others: Contribution to manuscript by drafting Introduction, Methods, Results, and Discussion sections.

Table of Contents

5.1 Abstract 330

5.2 Introduction 332

5.3 Methods 333

5.3.1 Participants and Cohort Information 333

5.3.2 Bacteriome Data Processing 335

5.3.3 Mycobiome Data Processing 336

5.3.4 Virome Data Processing 336

5.3.5 Statistical Analysis 337

5.3.6 Network Analysis 338

5.3.7 Machine Learning Analysis 338

5.4 Results 339

5.4.1 Microbiome components in combination can significantly differentiate IBS from healthy subjects, but not between the IBS clinical sub-types 339

5.4.2 Association analysis of observed synergy between microbiome components 343

5.4.3 Network analysis indicates loss of inter-kingdom connections in IBS 345

5.4.4 The predictive model performance of multi-kingdom data is similar to that of the bacteriome alone 353

5.5 Discussion 357

5.6 References 360

5.7 Supplementary information 364

5.7.1 Supplementary Figures 364

5.7.2 Supplementary Tables 367

5.1 Abstract

Microbiome science has improved our understanding of how microorganisms in the gut contribute to both healthy and diseased states. However, the use of 16S rRNA gene amplicon sequencing and, to a lesser extent, whole metagenome shotgun sequencing has resulted in a focus on the bacteriome in the microbiome field. As we have gained a better appreciation of the diversity of the gut microbiome, microbiome research has been extensively expanded beyond bacterial populations. Co-existing with bacteria and archaea, there is a less explored community consisting of eukaryotes and viruses, which not only coexist but also have a considerable influence on the health and physiology of human hosts by either modulating bacterial community, their metabolism, or contributing to host metabolome. Studies investigating the role of these components often focus on one component, and there is a need for an integrative analysis that may provide insights into cross-kingdom interactions and explain the differences in observed gut microbiomes between various populations.

Here, I report an integrative and predictive machine learning analysis of the gut bacteriome, mycobiome, and virome, and investigate the role of each component and their interactions in IBS. This work is carried out using data from a previously published cohort of samples consisting of 80 IBS patients and 65 Controls. I investigated how each of the inter-kingdom microbiome components contributed to the microbiome structure of IBS and control samples, and how these structures affected the predictive power of machine learning models based on each microbiome component individually, or in different combinations.

This research aims to improve our understanding of the microbiome and its potential use in microbiome-based therapies and diagnostic methods. The results of predictive modelling showed that both the non-bacterial microbiome and the bacteriome, either

individually or in combination, were equally effective in predicting IBS. The substantial contribution to significance of bacterial features in combined models indicates that gut bacteria are important in predicting IBS and that using measures of bacterial abundance may be a cost-effective method for diagnosing the disorder. However, it may also be worth further investigating the role of non-bacterial elements of the microbiome in understanding IBS. By analyzing the microbiome data across different kingdoms, we can understand the contribution of each microbiome component to beta-diversity and predictive model performance. The current analysis of associations and microbiome structure through co-occurrence and co-exclusion networks showed that negative associations between bacteria and other microbiome components played a significant role in determining the microbiome structure associated with health status.

5.2 Introduction

With advancements in microbiome science, the composition and function of plethora of non-bacterial components like eukaryotes (fungi), and viruses are being deeply investigated. Multiple studies have implicated the role of gut fungi (mycobiome), and viruses (virome) in various health conditions [1–10]. Typically, the mycobiome, and virome have been analysed separately [1,2,6,7,10], but studies exploring inter-kingdom relationships considering all three (bacteria, fungi, and viruses) are limited. Studies have investigated these inter-kingdom relationships in colorectal cancer [11], antibiotic response [12], Graft-versus-host disease [13], and skin microbiome [14]. These studies provided deeper insights into inter-kingdom interactions and have shown applicability of multi-kingdom markers in understanding clinically relevant microbiome interaction with the host, but similar studies in irritable bowel syndrome (IBS) are yet to be reported.

In previous studies from the laboratory, I had described the differences in gut bacteriome and metabolome associated with IBS and BAM compared to healthy controls [15], along with differences in mycobiome [1], and virome [6] of the same cohort. Here, I report an inter-kingdom integrated analysis of the bacteriome, mycobiome, and virome, to understand the relationship between these three components of the microbiome in IBS and Controls. I analyzed the bacteriome, mycobiome, and virome data separately, as well as in different combinations, to investigate the ability of models derived from the datasets to distinguish IBS from controls. Paired association, and inter-kingdom microbiome structure as co-occurrence and co-exclusion networks, were analyzed to identify relationships between the microbiome components. Finally, I report the predictive power of each individual microbiomic component's abundance

profile and their combinations using machine learning algorithms and model performance assessment.

5.3 Methods

5.3.1 Participants and Cohort Information

The cohort in the study has been reported previously [1,6,15], and includes 80 IBS patients (Rome IV criteria), along with 65 controls matched for age, gender, ethnicity, and geography. The IBS subjects were further classified into clinical subtypes, viz., IBS-C (Constipation predominant), IBS-D (Diarrhoea predominant), and IBS-M (Mixed). A descriptive statistics summary of the cohort is shown in Table 1. The availability of samples in each omic dataset varied, with detail of sample distribution provided in Table 2.

	Control (n = 65)	IBS (n = 80)
Age range, years (mean)	19-65 (45)	17-66 (39)
Sex (male/female)	16/49	15/65
BMI Class, n (%)		
Normal	25 (38)	31 (39)
Obese Class I	11 (17)	14 (18)
Obese Class II	3 (5)	5 (6)
Obese Class III	1 (2)	3 (4)
Overweight	21 (32)	22 (28)
Underweight	3 (5)	3 (4)
HADS: Anxiety, n (%)		
Normal (0-10)	59 (91)	58 (72)
Abnormal (11-21)	6 (9)	22 (28)
HADS: Depression, n (%)		
Normal (0-10)	64 (98)	70 (88)
Abnormal (11-21)	1 (2)	10 (12)

Bristol Stool Score, n (%)		
Normal	54 (83)	18 (22)
Constipated	8 (12)	22 (28)
Diarrhoea	3 (5)	40 (50)
IBS subtype, n (%)		
IBS-C		30 (38)
IBS-D	N/A	21 (26)
IBS-M		29 (36)
SeHCAT assayed, n (%)	9 (14)	46 (57)
Dietary group (FFQ), n (%)		
Omnivore	63 (97)	74 (92)
Vegetarian	1 (2)	2 (2)
Pescatarian	1 (2)	1 (1)
Gluten-free	0 (0)	4 (5)
Drinks alcohol, n (%)		
Current	54 (83)	57 (71)
Previous	0 (0)	1 (1)
Never	10 (15)	22 (28)
Smoker, n (%)		
Current	10* (15)	14* (18)
Previous	13 (20)	18 (22)
Never	42 (65)	48 (60)

* 1 subject in each group smoked e-cigarettes

BMI: Body Mass Index, FFQ: Food Frequency Questionnaire, HADS: Hospital Anxiety and Depression Scale, N/A: not applicable.

Table 1: Descriptive statistics of control and IBS subjects studied

Dataset	IBS-C	IBS-D	IBS-M	IBS	Controls	Total
Bacteriome	29	20	29	78	64	142
Mycobiome	30	21	29	80	64	144
Virome	21	17	17	55	51	106
Bacteriome_Mycobiome	29	20	29	78	63	141
Bacteriome_Virome	20	16	17	53	50	103
Mycobiome_Virome	21	17	17	55	51	106
Bacteriome_Mycobiome_Virome	20	16	17	53	50	103

Table 2: Sample distribution across omic datasets

5.3.2 Bacteriome Data Processing

Paired-end reads of bacterial 16S rRNA gene amplicon sequences, reported previously, were used for bacterial abundance estimation and were analyzed at zOTU (Zero-radius OTUs) level. Firstly, the paired-end reads were quality checked using multiqc [16], followed by merging and dereplication of the merged reads using the USEARCH pipeline. The dereplicated reads were clustered to obtain zOTUs and corresponding representative sequences using the UNOISE algorithm [17] of the pipeline. Finally, merged reads were mapped back to the zOTU representative sequences at 97% identity to obtain the zOTU counts table.

The counts data was first scalar normalized, followed by log10 transformation.

$$A(x) = log10(x + 10^{-6}) + 6 \qquad \textit{equation (1)}$$

where A(x) is the log-transformed abundance, x is the scalar normalized abundance of zOTU, and 10^{-6} is the pseudo count. The addition of six ensures that all values are positive. zOTUs present in greater than 10% of the samples were considered for further analysis.

The taxonomic classification of zOTUs was performed using SINTAX, a functionality of the USEARCH pipeline, and the RDP database (v18). A confidence threshold of 0.8 was applied for all taxa levels to consider the classification valid.

5.3.3 Mycobiome Data Processing

Sample preparation and resulting mycobiome data used in the current analysis has been reported previously [1] and in Chapter 3 of this thesis. In summary, the raw paired end reads were quality checked using multiqc [16], followed by merging, dereplication, and clustering to obtain zOTUs, using the USEARCH pipeline, and the UNOISE algorithm [17] applied therein. The merged reads were re-mapped to representative zOTU sequences at 97% identity, and counts table was generated.

The obtained counts table was then scalar normalized and was log transformed according to equation (1). zOTUs present in greater than 10% of the samples were considered for further analysis.

Representative sequences of the Fungal zOTUs were classified at different taxonomic levels using qiime2's feature-classifier [18(p2)], with default parameters, and UNITE database [19] of fungal sequences, as the training reference.

5.3.4 Virome Data Processing

The estimation of abundance of phages in the samples was based on sequences derived from faecal Virus-like particles (VLPs). The sample preparation and sequencing protocols have been reported previously [6] and in Chapter 4 of this thesis. In summary, the shotgun sequences of VLPs were quality checked, assembled, and dereplicated FastQC and multiQC v1.7 [16], metaSPAdes v3.13.0 [20], and BBMap v38.22, respectively . The obtained viral contigs were classified at different taxonomic levels using Kaiju v1.6.3. The prediction of ORFs and corresponding proteins encoded by

the viral contigs were identified and filtered using multiple methods mentioned previously [6].

The predicted proteins were then clustered using vConTACT2 v0.9.11 [21] to obtain viral clusters, whose abundance was calculated as the sum of the contig counts associated with each cluster in each sample. The viral cluster (VC) count abundance was then scalar normalized followed by log10 transformation according to equation 1. Viral clusters (VCs) present in greater than 10% of the samples were considered for further analysis.

5.3.5 Statistical Analysis

Beta diversity was calculated as Bray-Curtis dissimilarity on the log-transformed abundance using the vegan package [22] in R. The significance of the split between sample classes was tested using permutational MANOVA using the adonis function of the vegan package.

Principal Coordinate Analysis was performed using the Bray-Curtis dissimilarity as distance and visualized using the ade4 package in R. To estimate the feature loadings, canonical correlation analysis was performed, wherein the Spearman correlation of each feature with the first 2 principal axes was calculated using the psych package in R. The p-values were adjusted using the Benjamini-Hochberg method, and an adjusted p-value < 0.05 was considered as significant. The feature loadings (as correlation coefficient) were plotted using the ade4 package, and top drivers were visualized using the plotgenes function of the made4 package in R.

To investigate the association between two datasets, Procrustes analysis using the first two axes of the principal coordinate analysis was performed, and significance was

tested using the protest function of the vegan package in R. The Procrustes rotation was visualized using ggplot2.

5.3.6 Network Analysis

Microbial co-occurrence and co-exclusion networks were generated based on Spearman distance using the ccrepe [23] package in R. The function employs the principle reported previously as the ReBoot method [24]. Scalar normalized data was used, and Spearman correlation was calculated between the bacterial zOTUs and the non-bacterial features. Correlating pairs with a nominal p-value < 0.001 and correlation coefficient greater than +0.4 formed the co-occurrence networks, and pairs with nominal p-value < 0.001 and coefficient < -0.4 formed the co-exclusion networks, which were considered for network visualization and analysis. Co-occurrence and Co-Exclusion networks were generated for both IBS and Control populations.

The significance of network differences was tested using Fisher's exact test based on the number of nodes in the network and the number of input taxa used for network generation.

The visualization of the networks, as well as the calculation of network and node properties, were performed using the igraph package in R. The networks were visualized in a circular layout, with the size of each node corresponding to its degree (i.e., the number of connections for each node)

5.3.7 Machine Learning Analysis

Individual normalized microbiome datasets, i.e., bacteriome, mycobiome, and virome and their combined datasets were passed through an in-house machine learning pipeline (reported previously) [15], and model performance and feature importance were analyzed based on 10 fold cross-validation. To ensure the obtained models were

comparable, and devoid of biases due to differences in number of samples, only samples common in the three individual datasets (bacteriome, mycobiome, and virome) were considered for the analysis.

The pipeline implements a two-stage model development, where the training data is first passed through LASSO-based feature selection, followed by the development of a random forest classifier based on the selected features. Each step of the pipeline implements an internal repetitive 10-fold cross-validation to obtain the best optimized model based on the area under the curve (AUC) of the receiver operating characteristics (ROC) curve. The final random forest classifier was then used to predict the test dataset. The machine learning pipeline mainly uses the glmnet, randomforest, and caret packages and was implemented in R 4.1.

5.4 Results

5.4.1 Microbiome components in combination can significantly differentiate IBS from healthy subjects, but not between the IBS clinical sub-types

I analyzed differences in the microbiota profiles of IBS and Controls based on bacteriome, mycobiome, and virome datasets separately. In all the cases, I observed that the microbiome profiles could significantly separate the IBS from Controls but were unable to distinguish the clinical subtypes of IBS (Table 3). This observation is consistent with previously published reports [1,6,15]. Furthermore, the bacteriome demonstrated the largest separation between IBS and controls (based on R^2 value from permutational MANOVA), showing the bacteriome's importance in microbiome analysis (Table 3).

	IBS Vs Controls	IBS-Subtypes

Comparison	**R^2**	**p-value**	**R^2**	**p-value**
Bacteriome	0.0322	0.001	0.0264	0.42
Mycobiome	0.0128	0.003	0.0236	0.646
Virome	0.0155	0.001	0.0422	0.103
Bacteriome_MycoBiome	0.0304	0.001	0.0265	0.404
Bacteriome_Virome	0.0279	0.001	0.0393	0.371
MycoBiome_Virome	0.0185	0.001	0.0447	0.053
Bacteriome_MycoBiome_Virome	0.0273	0.001	0.0399	0.338

Table 3: Summary statistics of permutational MANOVA showing magnitude and significance of difference in microbiota profiles (beta diversity) between IBS and controls, or between IBS subtypes. Beta diversity was calculated as Bray-Curtis dissimilarity using log-transformed abundance.

Microbiota profiles based on each microbiome component (bacteriome, mycobiome, and virome) were visualized using principal coordinate analysis (Figure 1). Based on bacteriome and virome, the separation of IBS and controls was observed along the first two principal axes (Figure 1A and C), which relate to the two strongest trends in the dataset. The mycobiome separated the two populations along the second principal axis only (Figure 1B) with the primary axis being related to heterogeneity within the IBS and healthy populations.

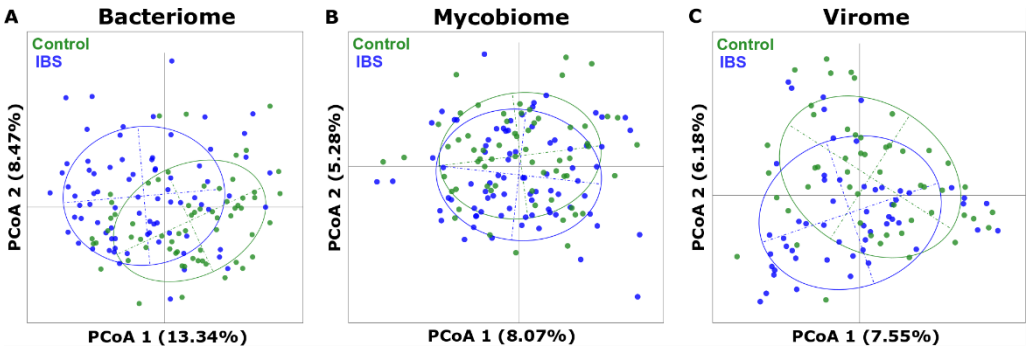


Figure 1: Compositional analysis of microbiome components. A) PCoA of beta diversity based on bacteriome data showing significant differences between IBS and controls (p=0.001; control, n=64; IBS, n=78). B) PCoA plot based on mycobiome beta diversity showing significant separation between IBS and controls (p=0.003; control, n=64; IBS, n=80). C) PCoA of virome beta diversity showing significant differences between IBS and control populations (p=0.001; control, n=51; IBS, n=55).

I also investigated the effect of integrating the relative abundances of microbiome components on the separation of IBS from controls based on beta diversity using PCoA (Figure 2) and permutational MANOVA. The combinations analyzed were bacteriome-mycobiome, bacteriome-virome, mycobiome-virome, and bacteriome-mycobiome-virome. The beta diversity of all these combinations significantly distinguished IBS from controls but did not differentiate between the IBS clinical subtypes (Table 3). The combinations where bacteriome was included demonstrated a higher degree of separation of IBS and control populations compared to the combinations where bacteriome was not included.

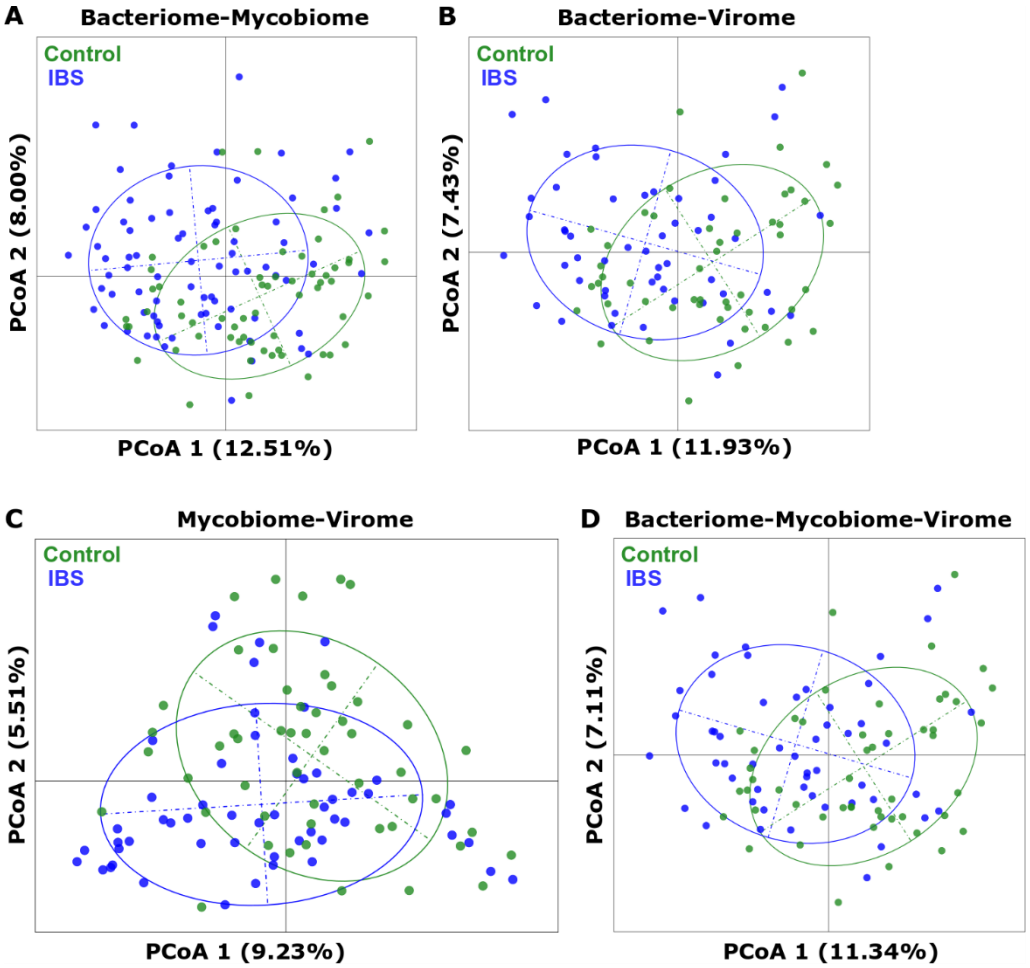


Figure 2: PCoA plots based on beta diversity (Bray-Curtis dissimilarity) using integrated datasets (log transformed abundance) show significant separation between IBS and Controls. A.) PCoA plots using Bacteriome and Mycobiome data from IBS and Control samples based on Bacteriome-Mycobiome beta diversity (p = 0.001; control, n = 63; IBS, n = 78). B.) PCoA plots based on Bacteriome and Virome beta diversity data from IBS and Control samples (p = 0.001; control, n = 50; IBS, n = 53). C.) PCoA plots based on Mycobiome and Virome beta diversity data from IBS and Control samples (p = 0.001; control, n = 51; IBS, n = 55). D.) PCoA plots based on Bacteriome, Mycobiome, and Virome data from IBS and Control samples beta diversity (p = 0.001; control, n = 50; IBS, n = 53).

Based on canonical correlation analysis, I observed many bacterial zOTUs that significantly correlated with the first two principal axes (Supplementary Figure S1 A). The loadings of mycobiome (Supplementary Figure S1 B) and virome features (Supplementary Figure S1 C) showed a smaller number of significantly associated features with the principal axes, suggesting that the Bacteriome data contained more features related to explaining the variance associated with the IBS and Control separation. The feature loadings based on canonical correlation analysis also showed that the bacterial features had higher correlation values with the principal axes compared to the non-bacterial components (Supplementary Figure S1).

Microbiome datasets, individually or in combinations, showed no significant separation of the different clinical sub-types of IBS (Supplementary Figure S2).

5.4.2 Association analysis of observed synergy between microbiome components

The degree of concordance and its statistical significance between two microbiome components was tested using pairwise Procrustes analysis. The analysis was carried out on the first two principal axes of the PCoAs, across the full dataset and separately for IBS and Control samples. The degree of association is represented by R, which is given by the equation

$R = \sqrt{1 - ss}$, where R is the degree of association and ss is the sums of squares returned by the protest function, which also tests for significance of the rotation using a permutational approach.

Inter-kingdom associations based on Procrustes analysis using beta-diversity showed significant association between bacteriome and mycobiome, and between bacteriome and phages in both IBS and Controls. No significant co-variation was observed between fungi and viruses. This observation was expected because the interactions

between bacteriome and mycobiome, as well as between bacteria and phages are known, but fungi and bacteriophages are not known to interact directly. Therefore, variables within the mycobiome and phageome are not expected to correlate with each other but with bacterial abundance in the gut (Table 4).

The degree of co-variance was calculated using samples from IBS, Controls, or both, and in all cases, the degree of co-variance between the datasets was highest for control samples (Table 4). A notable observation was that the association between datasets was highest between Bacteriome and Virome (All: 0.45, IBS: 0.425, Controls: 0.475), followed by Bacteriome and Mycobiome (All: 0.28, IBS: 0.266, Controls: 0.303), and was lowest for Virome and Mycobiome (All: 0.205, IBS: 0.211, Controls: 0.274).

Datasets	Sample (N)	R	p-value
Bacteriome-Mycobiome	All (N=141)	0.282	0.001
	Controls (N=63)	0.303	0.02
	IBS (N=78)	0.266	0.045
Bacteriome-Virome	All (N=103)	0.452	0.001
	Controls (N=50)	0.475	0.001
	IBS (N=53)	0.425	0.001
Mycobiome-Virome	All (N=106)	0.205	0.166
	Controls (N=51)	0.274	0.33
	IBS (N=55)	0.211	0.881

Table 4: Summary statistics of Procrustes analysis showing the magnitude and statistical significance of inter-kingdom covariance of beta-diversity. The first 5 principal axes of PCoA of each dataset were used for the analysis. Significance was tested based on a permutational iterative approach.

The PCoA reference and rotated Procrustes analysis axes were visualized to show the spatial arrangement of IBS and Control samples (Supplementary Figure S3). In case of Bacteriome-Mycobiome (Supplementary Figure S3 A and B), and Bacteriome-Virome (Supplementary Figure S3 C and D), the spatial arrangement of samples were observed to be in the same direction between the reference (bacteriome), and rotated matrix (mycobiome or virome). However, the sample position on the ordination plot upon rotation did not match between mycobiome and virome (Supplementary Figure S3 E and F) further indicating the lack of agreement between the two datasets.

5.4.3 Network analysis indicates loss of inter-kingdom connections in IBS

To investigate the differences in microbiome structure between IBS and Controls, I examined the inter-kingdom co-occurrence and co-exclusion patterns based on the correlation between components of the microbiome.

The co-occurrence network between Bacteriome and Mycobiome (Figure 3; A: Controls, B: IBS) was not significantly different between IBS and Controls based on node distribution, but a greater number of connections between bacterial and fungal taxa was observed in IBS (edges = 263, nodes = 291) compared to controls (edges = 230, nodes = 285). Network statistics are summarized in Table 5.

		Controls				IBS					
Comparison	**Organism**	**Samples (N)**	**Edges (N)**	**Nodes (N)**	**Input Taxa (N)**	**Samples (N)**	**Edges (N)**	**Nodes (N)**	**Input Taxa (N)**	**Odds Ratio**	**p-value**
Bacteriome-Mycobiome Co-Occurrence	Bacteriome	63	230	208	3270	78	263	220	3087	1.13	0.2297
	Mycobiome			77	146			71	144	0.872	0.6385
Bacteriome-Mycobiome Co-Exclusion	Bacteriome		107	94	3270		15	15	3087	0.165	< 0.001
	Mycobiome			45	146			8	144	0.133	< 0.001
Bacteriome-Virome Co-Occurrence	Bacteriome	50	1934	956	3162	53	1048	631	2982	0.619	< 0.001
	Virome			263	630			182	467	0.891	0.384
Bacteriome-Virome Co-Exclusion	Bacteriome		658	457	3162		250	193	2982	0.41	< 0.001
	Virome			198	630			86	467	0.493	< 0.001

Table 5: Summary statistics of Co-Occurrence and Co-Exclusion networks between Bacteriome and Mycobiome, and Bacteriome and Virome, in IBS and Control populations.

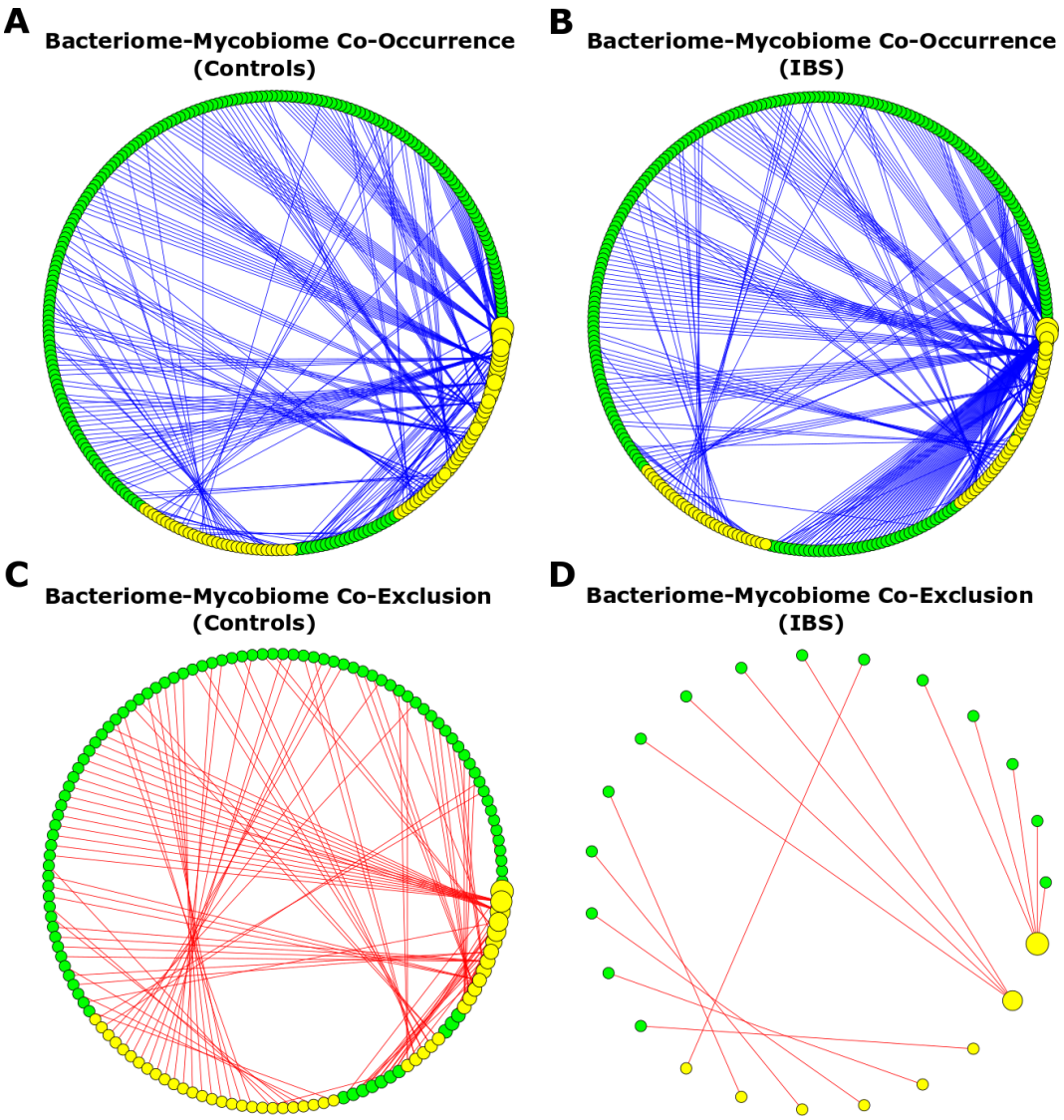


Figure 3: Inter-kingdom microbiome structure as co-occurrence and co-exclusion networks showing possible interactions between Bacteriome and Mycobiome in controls (n=63) and IBS (n=78). The networks are visualized in a circular layout with the size of each node corresponding to their degree centrality (number of connections). A.) Co-occurrence network (positive interaction) of Bacteriome and Mycobiome in controls. B.) Co-occurrence network of Bacteriome and Mycobiome in IBS showing an increase in the number of connections compared to controls. C.) Co-exclusion network (negative interaction) of Bacteriome and Mycobiome in controls. D.) Co-exclusion network of Bacteriome and Mycobiome in IBS showing a drastic reduction in the number of connections compared to controls.

In the co-occurrence network of controls, zOTUs classified as *Cladosporium cladosporioides* were observed to have a high number of connections with multiple bacterial zOTUs belonging to species *Streptococcus salivarius, Ruminococcus bromii, Faecalibacterium prausnitzii,* and members of the *Roseburia, Lachnospiraceae,* and *Ruminococcaceae*. Other major fungal zOTUs, such as M_zOTU135 (classified as *Cladosporium sphaerospermum*), and M_zOTU28 (classified as *Tausonia pullulans*) were observed to have positive connections with multiple bacterial zOTUs belonging to *Lachnospiraceae, Ruminococcaceae,* and *Veillonellaceae*.

In IBS, two unclassified fungal zOTUs were observed to co-occur with multiple bacterial zOTUs classified as *Bacteroides stercoris*, *Bacteroides*, and *Bacteroidaceae.* I also identified positive connections between fungal zOTUs, and bacterial zOTUs classified as *B. stercoris* which was indicated to contribute to microbial dysbiosis in a previously reported study [25]. Cumulative Node degrees, i.e., the total number of connections of each classified taxa in the co-occurrence networks, of controls and IBS are listed in Supplementary Tables S1 and S2, respectively.

The co-exclusion network between Bacteriome and Mycobiome (Figure 3; C: Controls, D: IBS) was significantly different between IBS and controls (Table 5), and a dramatic reduction in negative connections was observed in IBS (edges=15, nodes=23) compared to controls (edges=107, nodes=139).

In co-exclusion networks of controls, members of *Saccharomycetales,* including *Saccharomyces cerevisiae*, formed a higher number of negative connections with bacterial zOTUs belonging to the genus *Blautia*, *Roseburia,* and families like *Lachnospiraceae*. Multiple bacterial zOTUs belonging to species *F. prausnitzii* and *Lachnospira eligens* were also observed to be negatively associated with the fungal

zOTUs. The fungal zOTU M_zOTU7 (*C. cladosporioides*) was observed to have negative edges with bacterial zOTUs belonging to *Blautia*, *Rikenellaceae,* and *Lachnospiraceae*. In IBS, the co-exclusion network was sparse, and only two fungal zOTUs, M_zOTU111 (*Sawadaea bicornis*) and M_zOTU127 (*Aureobasidium pullulans*), were observed to have a high number of negative connections with bacterial zOTUs classified as *Lachnospiraceae*, *Roseburia,* and *F. prausnitzii*. Supplementary table S3 and S4 list the cumulative node degree of bacterial and fungal zOTUs in co-exclusion networks of controls and IBS, respectively.

In summary, for the Mycobiome to Bacteriome analysis, there was a decrease in the number of negative interactions in IBS compared to controls. Within the controls, the positive and negative interactions were spread among many Bacteriome-Mycobiome pairs, while within IBS both the positive and negative interactions were primarily associated with four fungi strains that are key for the Mycobiome-bacteriome interactions within IBS. These were two unclassified variants and a *S. bicornis* and *A. pullulans* strain. The *A. pullulans* species was the most abundant of these and was present in a quarter of individuals in both the IBS and control populations. This taxon contributed 11 positive and 4 negative interactions to the structure of the microbiota within IBS. The other three strains were detected in 10% of the dataset, with all three strains being slightly more common in IBS (12.5-13.75%) compared to controls (7.8 – 9.4%). The *S. bicornis* strain contributed 5 negative associations, while the two unclassified strains contributed 39 and 40 positive interactions. These four taxa contributed 60% of all negative and 34% of all positive Mycobiome-Bacteriome interactions within IBS.

In Bacteriome-virome networks (Figure 4; A: Controls, B: IBS), I observed that the co-occurrence network of IBS (edges= 1048, nodes = 813) was significantly different from that of controls (edges= 1934, nodes = 1219), based on the distribution of bacterial nodes (Table 5). In controls, the majority of the viruses co-occurring with the bacteriome belonged to the families *Mimiviridae*, *Siphoviridae*, *Podoviridae*, and *Myoviridae* (Supplementary table S5), and major bacterial zOTUs co-occurring with such phages belonged to species *F. prausnitzii* and *Prevotella copri*, along with members of *Lachnospiraceae* and *Ruminococcaceae*. Viral clusters belonging to the *Microviridae* family were observed to be present exclusively in controls. In IBS datastes, viral clusters co-occurring with bacteria belonged to the families described above for controls, but the number of connections was considerably lower compared to controls (Supplementary table S6). These were *Mimiviridae, Siphoviridae, Podoviridae,* and *Myoviridae,* and were observed to co-occur with bacteria belonging to the family *Ruminococcaceae, Lachnospiraceae*, *Bacteroidaceae,* and species like *F. prausnitzii* and *P. copri*. While most of the viral families were common between IBS and Controls, the viral clusters belonging to the families *Inoviridae* and *Phycodnaviridae* were exclusively present in the IBS network.

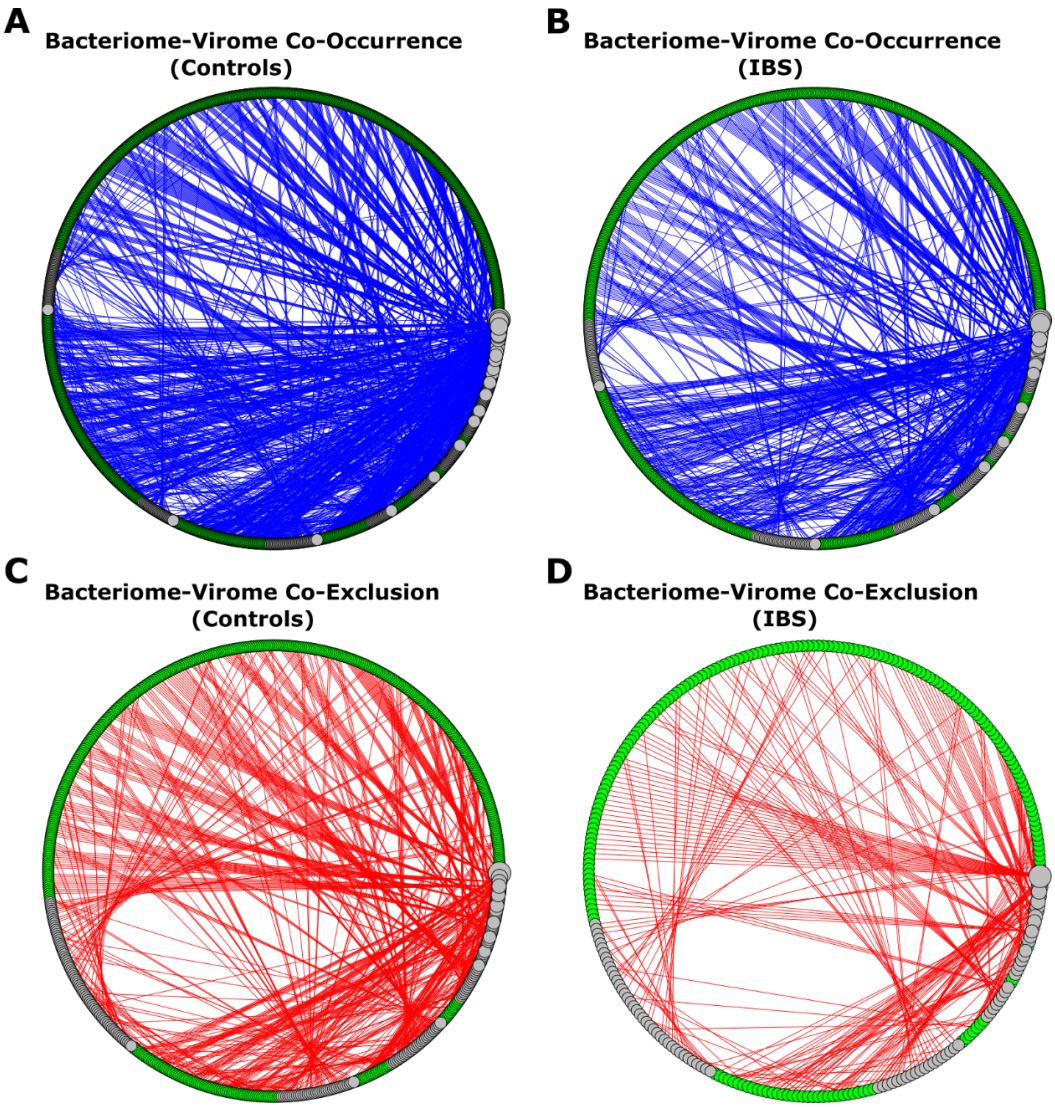


Figure 4: Inter-kingdom microbiome structure as co-occurrence and co-exclusion networks showing possible interactions between Bacteriome and Virome in Controls (n=50) and IBS(n=53). The networks are visualized in a circular layout with the size of each node corresponding to their degree centrality (number of connections). A.) Co-Occurrence network (positive interaction) of Bacteriome and Virome in Controls. B.) Co-Occurrence network of Bacteriome and Virome in IBS showing a reduction in the number of connections compared to controls. C.) Co-Exclusion network (negative interaction) of Bacteriome and Virome in Controls. D.) Co-Exclusion network of Bacteriome and Virome in IBS showing a drastic reduction in the number of connections compared to controls.

The co-exclusion network of Bacteriome-Virome (Figure 4; C: Controls, D: IBS) was significantly different between IBS (edges= 250, nodes = 279) and Controls (edges= 658, nodes = 655) based on the node distribution of both bacteria and viruses (Table 5). In controls, members of viral families *Mimiviridae*, *Siphoviridae*, and *Myoviridae* were observed to have a high number of negative connections with bacterial zOTUs (Supplementary table S7), along with other viral families including *Podoviridae*, *Phycodnaviridae*, and *Microviridae*, the last one being observed only in controls. These interacted negatively with bacterial zOTUs belonging to the family *Lachnospiraceae*, *Ruminococcaceae*, and *Bacteroidaceae*, along with species like *F. prausnitzii* that were found to have negative interactions with multiple viral clusters. In the IBS co-exclusion network of bacteria and viruses, viral clusters from *Mimiviridae*, *Siphoviridae*, *Myoviridae*, *Podoviridae*, and *Phycodnaviridae* were observed to have negative connections with bacteria, but the number of connections was comparatively fewer than that observed in the co-exclusion networks of control samples (Supplementary table S8), except *Phycodnaviridae*. These interacted negatively with bacterial zOTUs from *Lachnospiraceae*, *Bacteroides*, and species like *Agathobacter rectalis* that showed multiple negative connections with different viral clusters.

At the cluster level, the most interesting viral cluster was VC_4576_0, which had 23 negative interactions in IBS and no negative interactions in controls. However, most of the phages associated with IBS were also influential in the control microbiome profiles, although to a lesser extent (Supplementary table S9). In contrast, most of the phages associated with the non-IBS control state had no influence in the IBS microbiome profile as measured by correlation association (Supplementary table S10).

5.4.4 The predictive model performance of multi-kingdom data is similar to that of the bacteriome alone

The performance of predictive models for distinguishing IBS microbiome from controls using each dataset separately, or in combination with others, was assessed for predictive power. Machine learning analysis was carried out with 10-fold cross-validation and statistics were calculated from all samples (see Methods). The model performances of each microbiome component, on its own or in combination, did not significantly deviate from that observed based on only the bacteriome data (Figure 5 A, Table 6). Models where a combination of datasets was used favoured bacterial features compared to either fungi or viruses (Figure 5 B, Table 7), which highlights the greater importance of bacterial populations on the predictive power of the models.

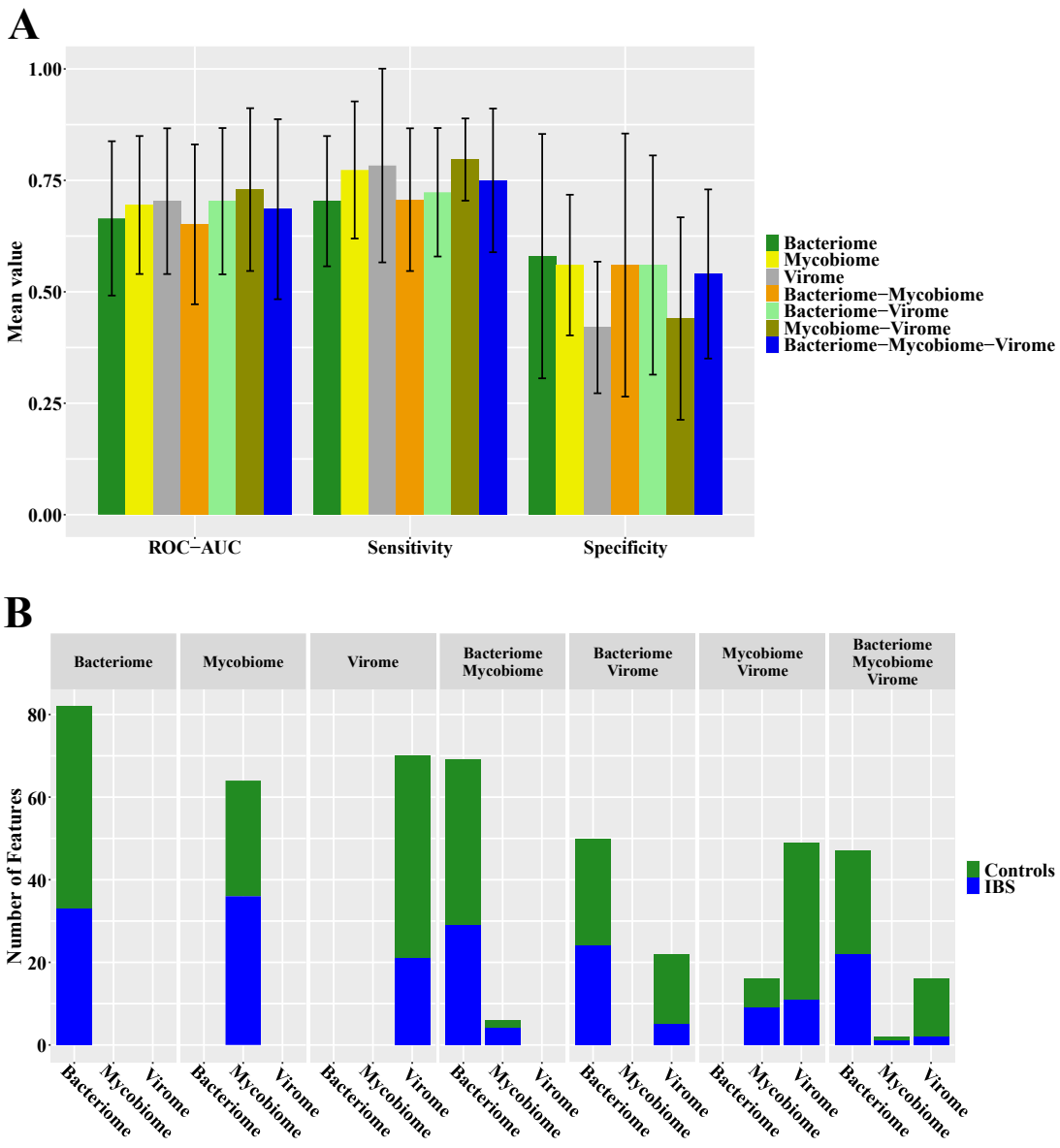


Figure 5: Summary of predictive analysis using different microbiome datasets. A.) Model performances of each dataset (individual and integrated) based on 10-fold cross-validation showing similar trends compared to the bacteriome-based analysis. B.) Number of features selected in each final model and their distribution based on association with IBS or controls.

	ROC-AUC		Sensitivity		Specificity	
Dataset	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**
Bacteriome	0.6647	0.1731	0.7033	0.1461	0.58	0.2741
Mycobiome	0.6947	0.1548	0.7733	0.1538	0.56	0.1578
Virome	0.7033	0.1635	0.7833	0.2173	0.42	0.1476
Bacteriome-Mycobiome	0.6513	0.1793	0.7067	0.1601	0.56	0.2951
Bacteriome-Virome	0.7033	0.1641	0.7233	0.1441	0.56	0.2459
Mycobiome-Virome	0.7293	0.1826	0.7967	0.0922	0.44	0.2271
Bacteriome-Mycobiome-Virome	0.6853	0.2019	0.75	0.1612	0.54	0.1897

Table 6: Predictive model performance statistics, based on 10-fold cross-validation, showing similar and comparable predictive power of different microbiome components, individually and in combinations.

	Number of Selected features					
	Bacteriome		Mycobiome		Virome	
Dataset	**IBS**	**Controls**	**IBS**	**Controls**	**IBS**	**Controls**
Bacteriome	33	49	-	-	-	-
Mycobiome	-	-	36	28	-	-
Virome	-	-	-	-	21	49
Bacteriome-Mycobiome	29	40	4	2	-	-
Bacteriome-Virome	24	26	-	-	5	17
Mycobiome-Virome	-	-	9	7	11	38
Bacteriome-Mycobiome-Virome	22	25	1	1	2	14

Table 7: Number of selected features in the final optimized models generated using the abundance profile of Bacteriome, Mycobiome, and Virome. A higher number of bacterial features as part of the predictive models shows a greater influence of the bacteriome in the prediction of IBS and controls.

In the model based on the bacteriome, members of *Lachnospiraceae*, *Ruminococcaceae*, and *Bacteroides* were associated with both IBS and controls. Species like *F. prausnitzii, A. rectalis*, and *P. copri* were associated with controls, and *Fusicatenibacter saccharivorans, Eggerthella lenta, Akkermansia muciniphila, Veillonella parvula,* and *Blautia faecis* were associated with IBS. A summary of the

selected features and their coefficients is listed in supplementary table S11. In integrated models comprising the bacteriome and mycobiome (supplementary table S12) and the bacteriome and virome (supplementary table S13), the association trends of bacterial taxa with IBS and controls were observed to be similar. In the bacteriome-virome model, zOTUs classified as *F. saccharivorans*, *Alistipes onderdonkii*, *Bacteroides eggerthii*, and *Streptococcus* were associated with IBS, while species like *P. copri*, *Alistipes ihumii* were associated with controls (supplementary table S13).

In the mycobiome-based model, species such as *Candida albicans*, *Candida dubliniensis*, and *Saccharomyces cerevisiae* were associated with irritable bowel syndrome (IBS), and members of *Saccharomycetales*, *Cladosporium*, *Aspergillus*, and a few *Penicillium* species were associated with controls (Supplementary Table S14). The mycobiome association in the integrated model with the bacteriome differed in the number of selected fungal zOTUs, and a fewer number of fungal zOTUs were included in the final optimized model (Table 7). Out of the 6 selected fungal zOTUs, 4 were associated with IBS and were classified as *C. albicans, Aspergillus ruber*, *Wickerhamomyces onychis*, and *Dipodascus geotrichum* (Supplementary Table S12). In the mycobiome-virome predictive model, fungal zOTUs classified as *C. albicans*, *A. ruber*, and *Malassezia restricta* were among the species associated with IBS. Species such as *Agaricus bisporus*, *Penicillium bialowiezense*, and a zOTU classified as *Saccharomycetales* were associated with controls (Supplementary Table S15).

In the optimized virome model, a low number of viral clusters were associated with IBS, and I observed a greater number of *Siphoviridae* and *Myoviridae* to be associated with controls (Supplementary Table S16). In the integrated bacteriome-virome model, I observed a fewer number of viral clusters being included in the final model. Viral cluster families *Podoviridae* and *Siphoviridae* were observed to be associated with

both IBS and controls, and clusters classified as *Mimiviridae* and *Myoviridae* were observed to be associated with controls (Supplementary Table S13). The association patterns of the viral clusters were observed to be different in the mycobiome-virome integrated models (Supplementary Table S15). A greater number of viral clusters classified as *Myoviridae* and *Siphoviridae* were associated with controls.

The performance of the machine learning model generated by integrating all microbiome components (bacteriome, mycobiome, and virome) was comparable to the performance of the model based solely on the bacteriome. From the feature coefficients of the final model (Supplementary Table S17), it was clear that the majority of the features comprised of bacterial zOTUs. One fungal zOTU (classified as *A. ruber*) was associated with IBS, while another fungal zOTU, classified as the genus *Aspergillus*, was associated with controls. Within the selected viral clusters, members of *Myoviridae*, *Siphoviridae*, *Podoviridae*, and *Mimiviridae* were associated with controls, while viral clusters classified as *Siphoviridae* and *Podoviridae* were associated with IBS. Among the bacterial zOTUs, those classified as *Ruminococcaceae* and *Lachnospiraceae* were associated with both IBS and controls. Bacterial species such as *F. saccharivorans*, *A. onderdonkii*, and *B. eggerthii* were found to be associated with IBS, and zOTUs classified as *P. copri*, *Bifidobacteriales*, and *Alistipes ihumii* were associated with controls.

5.5 Discussion

I report an integrated analysis of microbiome components, including bacteriome, mycobiome, and virome, to investigate differences in inter-kingdom interactions in IBS and control populations and to identify components predictive of IBS based on machine learning analysis. I observed that the microbiome and its components significantly differ between IBS and controls for all microbiome components

(individually, as well as in combinations), but these profiles failed to segregate the clinical subtypes of IBS. Inter-kingdom co-variance analysis suggested a significant synergy between the bacteriome and non-bacterial components, and no link between the mycobiome and virome.

Significant differences in inter-kingdom microbiome structure were observed between IBS and controls, based on co-occurrence and co-exclusion of the microbiome members. A large difference in co-exclusion/negative interactions between the bacteriome and non-bacterial components, as well as co-occurrence/positive interaction, along with a lower number of negative connections between bacteria and fungi, and viruses in IBS, emphasizes the reduction of potential inhibition of fungi and viruses by the bacteria in the gut, which possibly affects gut health in a detrimental way.

It is interesting to note that both the mycobiome and virome had more influence on the bacteriome than vice versa. As key species are known to provide structure to the microbiome, this points to the importance of the mycobiome and virome in maintaining the eubiotic state and suggests that the key species in IBS may not necessarily be bacterial. Therefore, the identification of the four identified fungal strains and viral clusters is of interest for shaping future research efforts.

The analysis of predictive models showed that the non-bacterial microbiome (either alone or in combination with the bacteriome) had a similar ability to predict IBS as models that only used the bacteriome. In addition, the presence of greater number of bacterial features in integrated models highlights the crucial role that gut bacteria play in predicting IBS. These findings suggest that using measures of bacterial abundance may be sufficient for creating cost-effective diagnostic tools for IBS, but it is possible

that important elements of the non-bacterial microbiome may be overlooked and are worth further investigation in understanding the disorder.

The current study indicates that microbiome structural alterations in IBS not only occur within the bacteriome, but can also be seen between the kingdoms, thereby providing insights into cross-kingdom interaction dynamics that can alter microbiome composition. The role of the gut mycobiome and virome in faecal microbiota transplantation (FMT) has recently been reviewed, discussing the impacts of fungal and viral members on the microbiome [26]. Therefore, understanding the dynamics between bacteria and other members of the microbiome can help us develop better microbiome-based therapies or mitigate risks from therapies like FMT.

This study aims to expand the current understanding of the microbiome and its potential applications in microbiome-based therapies and diagnostic methods. The analysis of predictive models showed that both the non-bacterial microbiome and the bacteriome, either alone or in combination, had a similar ability to predict IBS. The greater presence of bacterial features in integrated models suggests that gut bacteria play a crucial role in predicting IBS and using measures of bacterial abundance may be sufficient for creating cost-effective diagnostic tools for the disorder. However, important elements of the non-bacterial microbiome may also be worth further investigation in understanding IBS. By analysing microbiome data across different kingdoms, we can understand the contribution of each microbiome component to beta-diversity and predictive model performance. My analysis of associations and microbiome structure through co-occurrence and co-exclusion networks revealed that negative associations between bacteria and other microbiome components are associated with determination of the microbiome structure and health status.

5.6 References

1. Das A, O'Herlihy E, Shanahan F, et al. The fecal mycobiome in patients with Irritable Bowel Syndrome. Sci Rep 2021;11:124.

2. Mar Rodríguez M, Pérez D, Javier Chaves F, et al. Obesity changes the human gut mycobiome. Sci Rep 2015;5:14600.

3. You N, Zhuo L, Zhou J, et al. The Role of Intestinal Fungi and Its Metabolites in Chronic Liver Diseases. Gut Liver 2020;14:291–296.

4. Zhang L, Zhan H, Xu W, et al. The role of gut mycobiome in health and diseases. Therap Adv Gastroenterol 2021;14:17562848211047130.

5. Huseyin CE, O'Toole PW, Cotter PD, et al. Forgotten fungi-the gut mycobiome in human health and disease. FEMS Microbiol Rev 2017;41:479–511.

6. Coughlan S, Das A, O'Herlihy E, et al. The gut virome in Irritable Bowel Syndrome differs from that of controls. Gut Microbes 2021;13:1–15.

7. Clooney AG, Sutton TDS, Shkoporov AN, et al. Whole-Virome Analysis Sheds Light on Viral Dark Matter in Inflammatory Bowel Disease. Cell Host Microbe 2019;26:764-778.e5.

8. Ansari MH, Ebrahimi M, Fattahi MR, et al. Viral metagenomic analysis of fecal samples reveals an enteric virome signature in irritable bowel syndrome. BMC Microbiol 2020;20:123.

9. Yang K, Niu J, Zuo T, et al. Alterations in the Gut Virome in Obesity and Type 2 Diabetes Mellitus. Gastroenterology 2021;161:1257-1269.e13.

10. Nishijima S, Nagata N, Kiguchi Y, et al. Extensive gut virome variation and its associations with host and environmental factors in a population-level cohort. Nat Commun 2022;13:5252.

11. Liu N-N, Jiao N, Tan J-C, et al. Multi-kingdom microbiota analyses identify bacterial-fungal interactions and biomarkers of colorectal cancer across cohorts. Nat Microbiol 2022;7:238–250.

12. Haak BW, Argelaguet R, Kinsella CM, et al. Integrative Transkingdom Analysis of the Gut Microbiome in Antibiotic Perturbation and Critical Illness Chia N, ed. mSystems 2021;6:e01148-20.

13. Zhang F, Zuo T, Yeoh YK, et al. Longitudinal dynamics of gut bacteriome, mycobiome and virome after fecal microbiota transplantation in graft-versus-host disease. Nat Commun 2021;12:65.

14. Saheb Kashaf S, Proctor DM, Deming C, et al. Integrating cultivation and metagenomics for a multi-kingdom view of skin microbiome diversity and functions. Nat Microbiol 2022;7:169–179.

15. Jeffery IB, Das A, O'Herlihy E, et al. Differences in Fecal Microbiomes and Metabolomes of People With vs Without Irritable Bowel Syndrome and Bile Acid Malabsorption. Gastroenterology 2020;158:1016-1028.e8.

16. Ewels P, Magnusson M, Lundin S, et al. MultiQC: summarize analysis results for multiple tools and samples in a single report. Bioinformatics 2016;32:3047–3048.

17. Edgar RC, Flyvbjerg H. Error filtering, pair assembly and error correction for next-generation sequencing reads. Bioinformatics 2015;31:3476–3482.

18. Bolyen E, Rideout JR, Dillon MR, et al. Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. Nat Biotechnol 2019;37:852–857.

19. Nilsson RH, Larsson K-H, Taylor AFS, et al. The UNITE database for molecular identification of fungi: handling dark taxa and parallel taxonomic classifications. Nucleic Acids Res 2019;47:D259–D264.

20. Nurk S, Meleshko D, Korobeynikov A, et al. metaSPAdes: a new versatile metagenomic assembler. Genome Res 2017;27:824–834.

21. Bin Jang H, Bolduc B, Zablocki O, et al. Taxonomic assignment of uncultivated prokaryotic virus genomes is enabled by gene-sharing networks. Nat Biotechnol 2019;37:632–639.

22. Anderson MJ. A new method for non-parametric multivariate analysis of variance: NON-PARAMETRIC MANOVA FOR ECOLOGY. Austral Ecology 2001;26:32–46.

23. Schwager E, Bielski C, Weingart G. ccrepe: ccrepe_and_nc.score. R package 1.32.0. Available at: https://bioconductor.org/packages/release/bioc/html/ccrepe.html.

24. Faust K, Sathirapongsasuti JF, Izard J, et al. Microbial co-occurrence relationships in the human microbiome. PLoS Comput Biol 2012;8:e1002606.

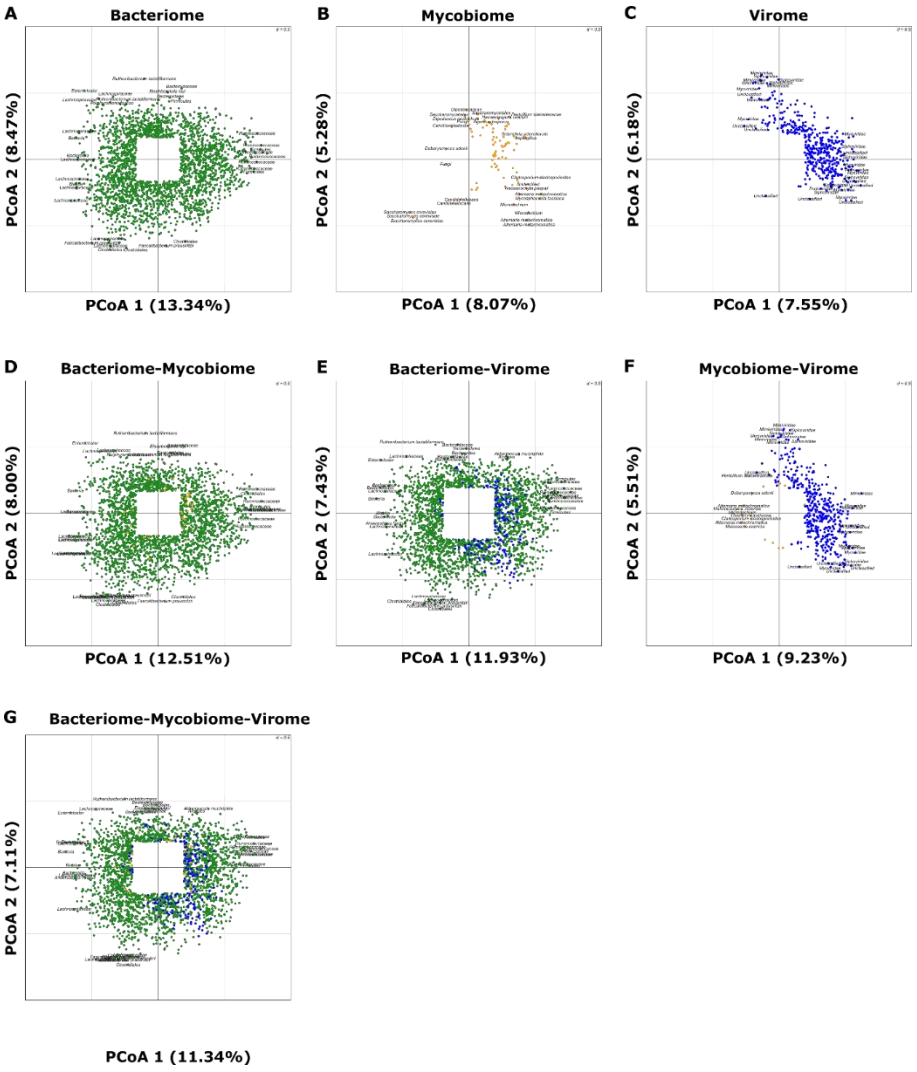
25. Casén C, Vebø HC, Sekelja M, et al. Deviations in human gut microbiota: a novel diagnostic test for determining dysbiosis in patients with IBS or IBD. Aliment Pharmacol Ther 2015;42:71–83.

26. Lam S, Bai X, Shkoporov AN, et al. Roles of the gut virome and mycobiome in faecal microbiota transplantation. The Lancet Gastroenterology & Hepatology 2022;7:472–484.

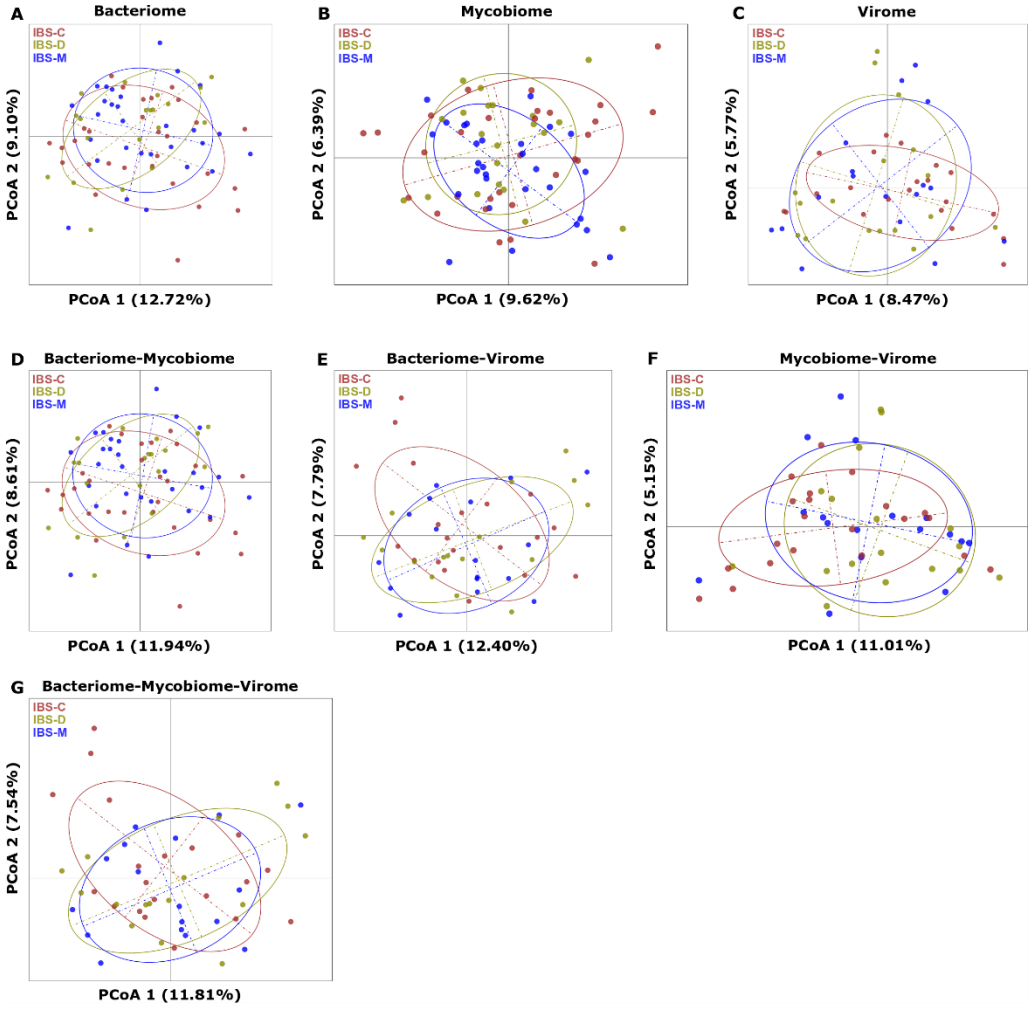
5.7 Supplementary information

5.7.1 Supplementary Figures

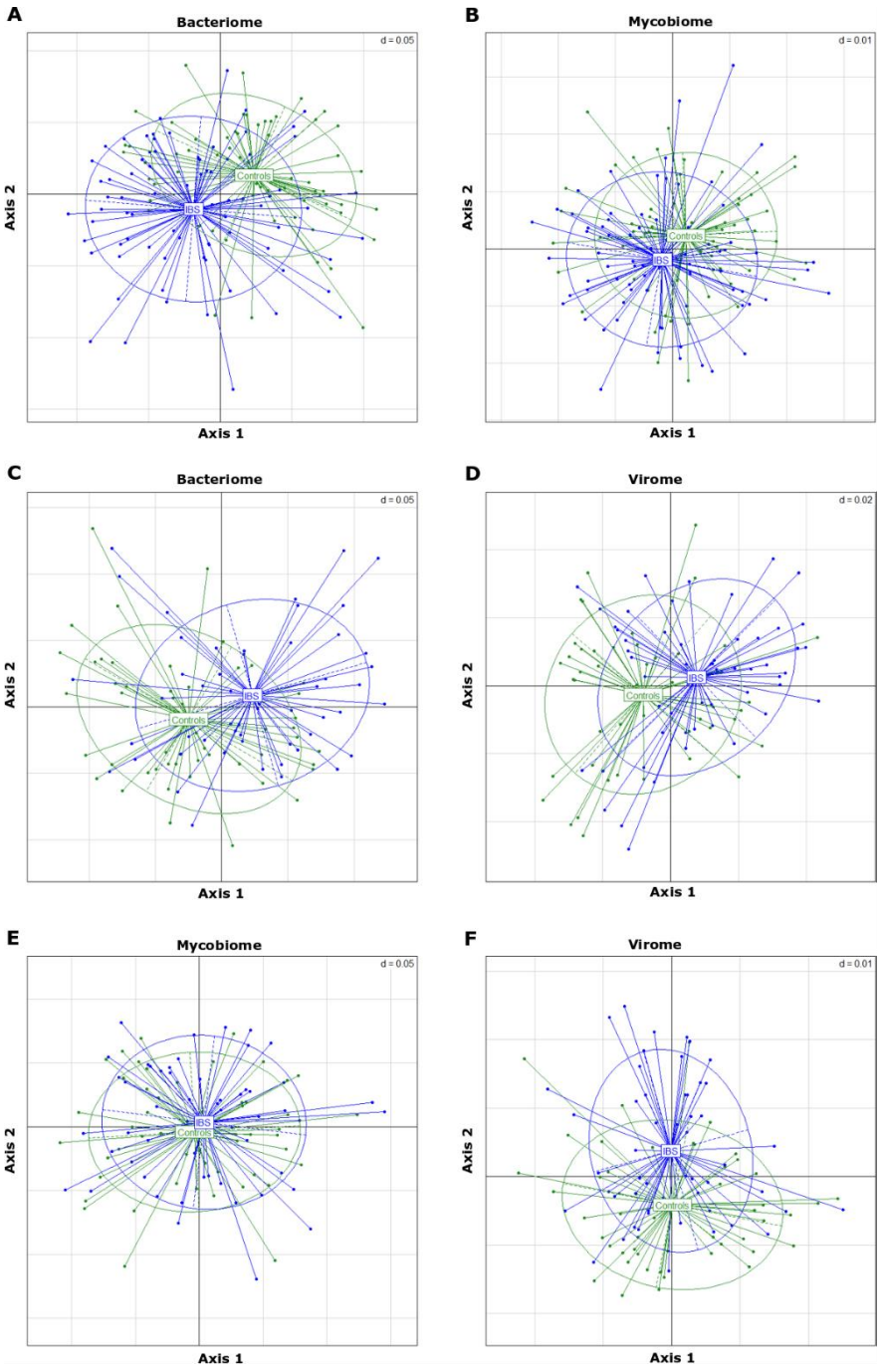
Supplementary Figure 1: Feature loadings based on the correlation between feature abundance and Principal co-ordinate axes, showing greater degree of association between Bacteriome and principal axes. Features showing significant correlation with the axes have been plotted. A.) Bacterial zOTUs (green); B.) Fungal zOTUs (yellow); C.) Viral clusters (grey); D.) Bacterial and Fungal zOTUs; E.) Bacterial zOTUs and Viral clusters; F.) Fungal zOTUs and Viral Clusters; G.) Bacterial zOTUs, Fungal zOTUs, and Viral Clusters.



Supplementary Figure 2: PCoA plots based on beta diversity (Bray-Curtis dissimilarity) using integrated datasets (log transformed abundance) showing the spatial arrangement of IBS samples based on clinical subtypes, i.e. IBS-C, IBS-D, IBS-M. None of the beta-diversity were significantly different between the clinical subtypes. A.) PCoA plots based on Bacteriome beta diversity (p = 0.42; IBS-C, n = 29; IBS-D, n = 20; IBS-M, n = 29). B.) PCoA plots based on Mycobiome beta diversity (p = 0.646; IBS-C, n = 30; IBS-D, n = 21; IBS-M, n = 29). C.) PCoA plots based on Virome beta diversity (p = 0.103; IBS-C, n = 21; IBS-D, n = 17; IBS-M, n = 17). D.) PCoA plots based on Bacteriome-Mycobiome beta diversity (p = 0.404; IBS-C, n = 29; IBS-D, n = 20; IBS-M, n = 29). E.) PCoA plots based on Bacteriome-Virome beta diversity (p = 0.371; IBS-C, n = 20; IBS-D, n = 16; IBS-M, n = 17). F.) PCoA plots based on Mycobiome-Virome beta diversity (p = 0.053; IBS-C, n = 21; IBS-D, n = 17; IBS-M, n = 17). G.) PCoA plots based on Bacteriome-Mycobiome-Virome beta diversity (p = 0.338; IBS-C, n = 20; IBS-D, n = 16; IBS-M, n = 17).



Supplementary Figure 3: Inter-kingdom association analysis based on Procrustes analysis using first 5 PCoA axes. A and B.) Procrustes plot between Bacteriome and Mycobiome (n= 141), where the arrows point towards the Bacteriome from Mycobiome (p = 0.001). C and D.) Procrustes plot between Bacteriome and Virome (n= 103), where the direction of arrows is from Virome to Bacteriome (p = 0.001). E and F.) Procrustes plot between Mycobiome and Virome (n= 106), where the direction of arrows is from Mycobiome to Virome (p = 0.166).



5.7.2 Supplementary Tables

Supplementary Table 1: Cumulative Node degrees, i.e., the total number of connections of each classified taxa in the co-occurrence network of Bacteriome and Mycobiome in controls. Nodes with degree > 1, have been reported.

Component	Taxa	Total Degree	Component	Taxa	Total Degree
Bacteriome	*Lachnospiraceae*	54	Mycobiome	*Tausonia pullulans*	10
Bacteriome	*Clostridiales*	30	Mycobiome	*Kazachstania unispora*	9
Bacteriome	*Ruminococcaceae*	27	Mycobiome	*Dipodascaceae*	7
Bacteriome	*Firmicutes*	15	Mycobiome	*Vishniacozyma victoriae*	7
Bacteriome	*Bacteroidaceae*	14	Mycobiome	*Wallemia muriae*	7
Bacteriome	*Faecalibacterium prausnitzii*	9	Mycobiome	*Bullera unica*	6
Bacteriome	*Blautia*	8	Mycobiome	*Dioszegia hungarica*	6
Bacteriome	*Roseburia*	7	Mycobiome	*Saccharomycetales*	6
Bacteriome	*Bacteria*	5	Mycobiome	*Penicillium spinulosum*	5
Bacteriome	*Bacteroides*	5	Mycobiome	*Dipodascus geotrichum*	4
Bacteriome	*Clostridia*	4	Mycobiome	*Kluyveromyces marxianus*	4
Bacteriome	*Dialister invisus*	4	Mycobiome	*Penicillium bialowiezense*	4
Bacteriome	*Bacteroidales*	3	Mycobiome	*Candida glaebosa*	3
Bacteriome	*Blautia luti*	3	Mycobiome	*Debaryomyces udenii*	3
Bacteriome	*Faecalibacterium*	2	Mycobiome	*Neoascochyta paspali*	3
Bacteriome	*Ruminococcus*	2	Mycobiome	*Penicillium lanosum*	3
Bacteriome	*Ruminococcus bromii*	2	Mycobiome	*Alternaria metachromatica*	2
Bacteriome	*Streptococcus salivarius*	2	Mycobiome	*Aspergillus*	2
Mycobiome	*Fungi*	48	Mycobiome	*Aspergillus japonicus*	2
Mycobiome	*unidentified*	25	Mycobiome	*Candida*	2
Mycobiome	*Cladosporium cladosporioides*	20	Mycobiome	*Cystofilobasidium macerans*	2
Mycobiome	*Cladosporium sphaerospermum*	12	Mycobiome	*Holtermanniella takashimae*	2
Mycobiome	*Sclerotinia sclerotiorum*	10	Mycobiome	*Saccharomyces cerevisiae*	2

Supplementary Table 2: Cumulative Node degrees, i.e., the total number of connections of each classified taxa in the co-occurrence network of Bacteriome and Mycobiome in IBS. Nodes with degree > 1, have been reported.

Component	Taxa	Total Degree	Component	Taxa	Total Degree
Bacteriome	*Bacteroides*	43	Mycobiome	*Fungi*	153
Bacteriome	*Bacteroidaceae*	35	Mycobiome	*Aureobasidium pullulans*	17
Bacteriome	*Lachnospiraceae*	35	Mycobiome	*Agaricus bisporus*	14
Bacteriome	*Bacteroides stercoris*	30	Mycobiome	*unidentified*	11
Bacteriome	*Ruminococcaceae*	22	Mycobiome	*Penicillium lanosum*	7
Bacteriome	*Faecalibacterium prausnitzii*	10	Mycobiome	*Aspergillus*	5
Bacteriome	*Clostridiales*	8	Mycobiome	*Alternaria brassicae*	4
Bacteriome	*Firmicutes*	7	Mycobiome	*Eurotiales*	4
Bacteriome	*Roseburia*	7	Mycobiome	*Saccharomycetales*	4
Bacteriome	*Blautia*	6	Mycobiome	*Candida*	3
Bacteriome	*Anaerostipes hadrus*	5	Mycobiome	*Candida tropicalis*	3
Bacteriome	*Hydrogeniiclostridium mannosilyticum*	4	Mycobiome	*Penicillium bialowiezense*	3
Bacteriome	*Bifidobacterium*	3	Mycobiome	*Vishniacozyma victoriae*	3
Bacteriome	*Clostridiales Incertae Sedis XIII*	3	Mycobiome	*Alternaria metachromatica*	2
Bacteriome	*Alistipes finegoldii*	2	Mycobiome	*Bullera unica*	2
Bacteriome	*Bacteria*	2	Mycobiome	*Candida albicans*	2
Bacteriome	*Bacteroides caccae*	2	Mycobiome	*Cladosporium cladosporioides*	2
Bacteriome	*Eisenbergiella tayi*	2	Mycobiome	*Cladosporium sphaerospermum*	2
Bacteriome	*Flintibacter*	2	Mycobiome	*Itersonilia perplexans*	2
Bacteriome	*Ihubacter massiliensis*	2	Mycobiome	*Kazachstania barnettii*	2
Bacteriome	*Parabacteroides distasonis*	2	Mycobiome	*Wallemia*	2
Bacteriome	*Ruminococcus*	2	Mycobiome	*Wickerhamomyces onychis*	2
Bacteriome	*Turicibacter sanguinis*	2	Mycobiome	*Xylariales*	2

Supplementary Table 3: Cumulative Node degrees, i.e., the total number of connections of each classified taxa in the co-exclusion network of Bacteriome and Mycobiome in Controls. Nodes with degree > 1, have been reported.

Component	Taxa	Total Degree	Component	Taxa	Total Degree
Bacteriome	*Lachnospiraceae*	38	Bacteriome	*Roseburia*	2
Bacteriome	*Ruminococcaceae*	11	Mycobiome	*Saccharomyces cerevisiae*	19
Bacteriome	*Clostridiales*	10	Mycobiome	*unidentified*	18
Bacteriome	*Lachnospira eligens*	7	Mycobiome	*Saccharomycetales*	14
Bacteriome	*Agathobacter rectalis*	4	Mycobiome	*Fungi*	12
Bacteriome	*Blautia*	4	Mycobiome	*Alternaria metachromatica*	6
Bacteriome	*Clostridium IV*	4	Mycobiome	*Hanseniaspora uvarum*	4
Bacteriome	*Bacteroidaceae*	3	Mycobiome	*Sclerotinia sclerotiorum*	4
Bacteriome	*Faecalibacterium prausnitzii*	3	Mycobiome	*Ascomycota*	3
Bacteriome	*Agathobacter*	2	Mycobiome	*Aureobasidium pullulans*	3
Bacteriome	*Anaerobutyricum*	2	Mycobiome	*Cladosporium cladosporioides*	3
Bacteriome	*Bifidobacterium longum*	2	Mycobiome	*Penicillium lanosum*	3
Bacteriome	*Clostridium XVIII*	2	Mycobiome	*Aspergillus*	2
Bacteriome	*Fusicatenibacter saccharivorans*	2			

Supplementary Table 4: Cumulative Node degrees, i.e., the total number of connections of each classified taxa in the co-exclusion network of Bacteriome and Mycobiome in IBS.

Component	Taxa	Total Degree	Component	Taxa	Total Degree
Bacteriome	*Lachnospiraceae*	5	Mycobiome	*Aureobasidium pullulans*	4
Bacteriome	*Faecalibacterium prausnitzii*	4	Mycobiome	*Alternaria metachromatica*	1
Bacteriome	*Bacteroides*	2	Mycobiome	*Cystofilobasidium macerans*	1
Bacteriome	*Roseburia*	2	Mycobiome	*Debaryomyces udenii*	1
Bacteriome	*Anaerobutyricum soehngenii*	1	Mycobiome	*Heterobasidion araucariae*	1
Bacteriome	*Clostridiales*	1	Mycobiome	*Itersonilia perplexans*	1
Mycobiome	*Sawadaea bicornis*	5	Mycobiome	*Saccharomyces cerevisiae*	1

Supplementary Table 5: Cumulative Node degrees, i.e., the total number of connections of each classified taxa in the co-occurrence network of Bacteriome and Virome in controls. Nodes with degree > 1, have been reported.

Component	Taxa	Total Degree	Component	Taxa	Total Degree
Virome	*Unclassified*	672	Bacteriome	*Clostridiales Incertae Sedis XIII*	5
Virome	*Siphoviridae*	588	Bacteriome	*Blautia faecis*	5
Virome	*Mimiviridae*	260	Bacteriome	*Alistipes onderdonkii*	5
Virome	*Myoviridae*	256	Bacteriome	*Akkermansia muciniphila*	5
Virome	*Podoviridae*	129	Bacteriome	*Verrucomicrobiales*	4
Virome	*Microviridae*	29	Bacteriome	*Ruminococcus*	4
Bacteriome	*Lachnospiraceae*	299	Bacteriome	*Rikenellaceae*	4
Bacteriome	*Ruminococcaceae*	269	Bacteriome	*Proteobacteria*	4
Bacteriome	*Clostridiales*	184	Bacteriome	*Peptostreptococcaceae*	4
Bacteriome	*Prevotella copri*	153	Bacteriome	*Holdemanella biformis*	4
Bacteriome	*Faecalibacterium prausnitzii*	152	Bacteriome	*Fusicatenibacter*	4
Bacteriome	*Firmicutes*	102	Bacteriome	*Eubacterium coprostanoligenes*	4
Bacteriome	*Bacteroides*	90	Bacteriome	*Bifidobacterium*	4
Bacteriome	*Bacteroidales*	78	Bacteriome	*Alistipes shahii*	4
Bacteriome	*Bacteroidaceae*	71	Bacteriome	*Alistipes inops*	4
Bacteriome	*Bacteria*	67	Bacteriome	*Alistipes dispar*	4
Bacteriome	*Blautia*	48	Bacteriome	*Akkermansiaceae*	4
Bacteriome	*Roseburia*	32	Bacteriome	*Turicibacter sanguinis*	3
Bacteriome	*Agathobacter rectalis*	29	Bacteriome	*Prevotellaceae*	3
Bacteriome	*Prevotella*	28	Bacteriome	*Enterobacteriaceae*	3
Bacteriome	*Clostridia*	25	Bacteriome	*Alphaproteobacteria*	3
Bacteriome	*Faecalibacterium*	24	Bacteriome	*Alistipes senegalensis*	3
Bacteriome	*Bacteroidia*	19	Bacteriome	*Alistipes finegoldii*	3
Bacteriome	*Coprococcus eutactus*	13	Bacteriome	*Agathobacter*	3
Bacteriome	*Ruminococcus bromii*	12	Bacteriome	*Romboutsia timonensis*	2
Bacteriome	*Gemmiger*	11	Bacteriome	*Phocaeicola dorei*	2
Bacteriome	*Bacteroides uniformis*	11	Bacteriome	*Odoribacter splanchnicus*	2
Bacteriome	*Alistipes*	11	Bacteriome	*Neglecta timonensis*	2
Bacteriome	*Bacteroidetes*	10	Bacteriome	*Lachnospiracea incertae sedis*	2
Bacteriome	*Anaerobutyricum*	9	Bacteriome	*Lachnospira eligens*	2
Bacteriome	*Ruminococcus callidus*	8	Bacteriome	*Intestinimonas massiliensis*	2
Bacteriome	*Phocaeicola*	8	Bacteriome	*Faecalibacillus intestinalis*	2
Bacteriome	*Fusicatenibacter saccharivorans*	7	Bacteriome	*Betaproteobacteria*	2
Bacteriome	*Streptococcus*	6	Bacteriome	*Barnesiella intestinihominis*	2

Bacteriome	*Porphyromonadaceae*	5	Bacteriome	*Anaerobutyricum soehngenii*	2
Bacteriome	*Intestinibacter bartlettii*	5			

Supplementary Table 6: Cumulative Node degrees, i.e., the total number of connections of each classified taxa in the co-occurrence network of Bacteriome and Virome in IBS. Nodes with degree > 1, have been reported.

Component	Taxa	Total Degree	Component	Taxa	Total Degree
Virome	*Unclassified*	302	Bacteriome	*Agathobacter*	4
Virome	*Siphoviridae*	270	Bacteriome	*Anaerotignum faecicola*	4
Virome	*Myoviridae*	215	Bacteriome	*Phascolarctobacterium faecium*	4
Virome	*Mimiviridae*	118	Bacteriome	*Alistipes finegoldii*	3
Virome	*Podoviridae*	113	Bacteriome	*Bacteroides fragilis*	3
Virome	*Inoviridae*	23	Bacteriome	*Barnesiella intestinihominis*	3
Virome	*Phycodnaviridae*	7	Bacteriome	*Colidextribacter massiliensis*	3
Bacteriome	*Ruminococcaceae*	186	Bacteriome	*Coprococcus*	3
Bacteriome	*Lachnospiraceae*	171	Bacteriome	*Faecalibacillus*	3
Bacteriome	*Faecalibacterium prausnitzii*	100	Bacteriome	*Odoribacter splanchnicus*	3
Bacteriome	*Bacteroides*	91	Bacteriome	*Akkermansia muciniphila*	2
Bacteriome	*Bacteroidaceae*	75	Bacteriome	*Alistipes*	2
Bacteriome	*Clostridiales*	71	Bacteriome	*Alistipes ihumii*	2
Bacteriome	*Prevotella copri*	54	Bacteriome	*Anaerotignum*	2
Bacteriome	*Roseburia*	34	Bacteriome	*Burkholderiales*	2
Bacteriome	*Bacteroidales*	29	Bacteriome	*Butyricimonas paravirosa*	2
Bacteriome	*Bacteria*	28	Bacteriome	*Faecalibacillus faecis*	2
Bacteriome	*Bacteroidia*	17	Bacteriome	*Fusicatenibacter saccharivorans*	2
Bacteriome	*Blautia*	14	Bacteriome	*Massilimicrobiota timonensis*	2
Bacteriome	*Phocaeicola*	12	Bacteriome	*Negativicutes*	2
Bacteriome	*Firmicutes*	10	Bacteriome	*Paludicola psychrotolerans*	2
Bacteriome	*Coprococcus eutactus*	9	Bacteriome	*Parabacteroides distasonis*	2
Bacteriome	*Faecalibacterium*	9	Bacteriome	*Peptostreptococcaceae*	2
Bacteriome	*Gemmiger*	9	Bacteriome	*Porphyromonadaceae*	2
Bacteriome	*Agathobacter rectalis*	8	Bacteriome	*Roseburia intestinalis*	2
Bacteriome	*Clostridium IV*	7	Bacteriome	*Roseburia inulinivorans*	2

Bacteriome	*Clostridia*	6	Bacteriome	*Streptococcus salivarius*	2
Bacteriome	*Lachnospira eligens*	6	Bacteriome	*Sutterella wadsworthensis*	2
Bacteriome	*Rikenellaceae*	5			

Supplementary Table 7: Cumulative Node degrees, i.e., the total number of connections of each classified taxa in the co-exclusion network of Bacteriome and Virome in Controls. Nodes with degree > 1, have been reported.

Component	Taxa	Total Degree	Component	Taxa	Total Degree
Virome	*Unclassified*	192	Bacteriome	*Flavonifractor plautii*	4
Virome	*Siphoviridae*	172	Bacteriome	*Oscillibacter*	4
Virome	*Myoviridae*	129	Bacteriome	*Phocaeicola*	4
Virome	*Mimiviridae*	114	Bacteriome	*Alistipes shahii*	3
Virome	*Podoviridae*	43	Bacteriome	*Bilophila wadsworthia*	3
Virome	*Microviridae*	5	Bacteriome	*Clostridia*	3
Virome	*Phycodnaviridae*	3	Bacteriome	*Clostridiales Incertae Sedis XIII*	3
Bacteriome	*Lachnospiraceae*	142	Bacteriome	*Clostridium IV*	3
Bacteriome	*Ruminococcaceae*	64	Bacteriome	*Intestinimonas butyriciproducens*	3
Bacteriome	*Bacteroides*	61	Bacteriome	*Parabacteroides distasonis*	3
Bacteriome	*Bacteroidaceae*	49	Bacteriome	*Porphyromonadaceae*	3
Bacteriome	*Faecalibacterium prausnitzii*	41	Bacteriome	*Alistipes*	2
Bacteriome	*Blautia*	31	Bacteriome	*Alistipes finegoldii*	2
Bacteriome	*Clostridiales*	26	Bacteriome	*Alistipes ihumii*	2
Bacteriome	*Bacteria*	19	Bacteriome	*Anaerobutyricum soehngenii*	2
Bacteriome	*Bacteroidales*	18	Bacteriome	*Bacteroides caccae*	2
Bacteriome	*Agathobacter rectalis*	16	Bacteriome	*Bacteroidia*	2
Bacteriome	*Bacteroides uniformis*	12	Bacteriome	*Bifidobacterium longum*	2
Bacteriome	*Firmicutes*	12	Bacteriome	*Blautia faecis*	2
Bacteriome	*Ruthenibacterium lactatiformans*	11	Bacteriome	*Blautia luti*	2
Bacteriome	*Prevotella copri*	10	Bacteriome	*Faecalibacterium*	2
Bacteriome	*Roseburia*	10	Bacteriome	*Flintibacter*	2
Bacteriome	*Agathobacter*	8	Bacteriome	*Gemmiger*	2
Bacteriome	*Anaerobutyricum*	8	Bacteriome	*Lawsonibacter asaccharolyticus*	2
Bacteriome	*Anaerostipes hadrus*	5	Bacteriome	*Mediterraneibacter faecis*	2

Bacteriome	*Bacteroides faecis*	4	Bacteriome	*Phascolarctobacterium faecium*	2
Bacteriome	*Bacteroidetes*	4	Bacteriome	*Pseudomonas*	2
Bacteriome	*Dysosmobacter welbionis*	4	Bacteriome	*Rikenellaceae*	2
Bacteriome	*Enterocloster*	4	Bacteriome	*Streptococcus*	2

Supplementary Table 8: Cumulative Node degrees, i.e., the total number of connections of each classified taxa in the co-exclusion network of Bacteriome and Virome in IBS. Nodes with degree > 1, have been reported.

Component	Taxa	Total Degree	Component	Taxa	Total Degree
Virome	*Myoviridae*	78	Bacteriome	*Anaerobutyricum*	3
Virome	*Unclassified*	69	Bacteriome	*Anaerostipes hadrus*	3
Virome	*Siphoviridae*	46	Bacteriome	*Bacteroides uniformis*	3
Virome	*Mimiviridae*	38	Bacteriome	*Blautia faecis*	3
Virome	*Podoviridae*	15	Bacteriome	*Enterocloster*	3
Virome	*Phycodnaviridae*	4	Bacteriome	*Fusicatenibacter*	3
Bacteriome	*Lachnospiraceae*	80	Bacteriome	*Hungatella*	3
Bacteriome	*Bacteroides*	23	Bacteriome	*Ruminococcaceae*	3
Bacteriome	*Agathobacter rectalis*	16	Bacteriome	*Bacteroidales*	2
Bacteriome	*Blautia*	16	Bacteriome	*Bacteroidetes*	2
Bacteriome	*Bacteroidaceae*	10	Bacteriome	*Bilophila wadsworthia*	2
Bacteriome	*Faecalibacterium prausnitzii*	10	Bacteriome	*Dialister invisus*	2
Bacteriome	*Clostridiales*	9	Bacteriome	*Flintibacter butyricus*	2
Bacteriome	*Agathobacter*	6	Bacteriome	*Lachnospiracea incertae sedis*	2
Bacteriome	*Bacteroidia*	6	Bacteriome	*Prevotella copri*	2
Bacteriome	*Bacteria*	5	Bacteriome	*Roseburia*	2
Bacteriome	*Fusicatenibacter saccharivorans*	5	Bacteriome	*Streptococcus*	2
Bacteriome	*Phocaeicola*	4			

Supplementary Table 9: Top 5 IBS-associated Phages and their network Degree property showing the number of interactions of each Phage in IBS and controls as ordered by Principal Component Analysis.

Bacteriome-Virome interactions	Groups	VC_4576_0	VC_1543_0	VC_1514_0	VC_4185_0	VC_1645_0
Co-Occurrence	Controls	1	6	5	0	0
	IBS	2	33	45	17	23
Co-Exclusion	Controls	0	3	2	0	0
	IBS	23	9	4	9	6

Supplementary Table 10: Top 12 Control associated Phages whose influence is depleted in IBS and their network Degree property showing the number of interactions of each Phage in IBS and controls as ordered by Principal Component Analysis.

Bacteriome-Virome interactions	Groups	VC_1910_0	VC_1978_0	VC_5183_0	VC_1601_0	VC_1853_0	VC_2313_0	VC_1704_1	VC_650_1	VC_1477_0	VC_2291_0	VC_1205_0	VC_2542_0
Co-Occurrence	Controls	71	57	50	33	47	36	38	39	20	30	4	14
	IBS	16	0	0	0	9	0	0	0	0	0	0	0
Co-Exclusion	Controls	36	12	10	9	7	5	3	2	10	4	17	11
	IBS	6	0	0	0	0	0	0	0	0	0	0	0

Supplementary Table 11: Lasso-coefficient, and random forest importance of selected features in the final optimized model based on bacteriome. A positive LASSO coefficient indicates association with controls, and negative indicates association with IBS.

ID	Lasso_Coefficient	RF_Importance	Taxonomy	ID	Lasso_Coefficient	RF_Importance	Taxonomy

B_zOTU527	1271.67	1.09	*Lachnospiraceae*	B_zOTU7409	5162.34	0.43	*Faecalibacterium prausnitzii*
B_zOTU413	2108	0.96	*Lachnospiraceae*	B_zOTU8068	-3491.1	0.42	*Streptococcus*
B_zOTU4928	-3893.54	0.86	*Lachnospiraceae*	B_zOTU1439	2080.18	0.41	*Lachnospiraceae*
B_zOTU170	-77.67	0.8	*Bacteroides*	B_zOTU345	-110.22	0.41	*Lachnospiraceae*
B_zOTU3628	-174.47	0.78	*Bacteroidetes*	B_zOTU4681	3163.64	0.4	*Bacteroidales*
B_zOTU826	2164.03	0.73	*Clostridiales Incertae Sedis XIII*	B_zOTU202	-19.63	0.4	*Clostridiales*
B_zOTU434	1058.78	0.71	*Lachnospiraceae*	B_zOTU6671	-2323.15	0.39	*Roseburia*
B_zOTU9017	894.02	0.69	*Lachnospiraceae*	B_zOTU3515	-74.32	0.39	*Bacteroides faecichinchillae*
B_zOTU564	523.05	0.68	*Lachnospiraceae*	B_zOTU8674	11251.58	0.39	*Ruminococcaceae*
B_zOTU8943	-783.8	0.67	*Blautia faecis*	B_zOTU242	48.44	0.39	*Clostridium XlVb*
B_zOTU6357	-948	0.66	*Fusicatenibacter saccharivorans*	B_zOTU6521	1631.73	0.39	*Agathobacter rectalis*
B_zOTU2762	231.38	0.66	*Faecalibacterium prausnitzii*	B_zOTU300	103.45	0.39	*Clostridiales*
B_zOTU5061	-142.9	0.65	*Bacteroidaceae*	B_zOTU2804	3922.54	0.39	*Lachnospiraceae*
B_zOTU4876	-3809.12	0.63	*Lachnospiraceae*	B_zOTU8168	-4987.75	0.38	*Ruminococcaceae*
B_zOTU1009	5944.85	0.62	*Blautia*	B_zOTU1447	237.51	0.38	*Ruminococcaceae*
B_zOTU8547	3441.72	0.61	*Ruminococcaceae*	B_zOTU6156	51293.24	0.37	*Agathobacter rectalis*
B_zOTU4249	171.71	0.6	*Lachnospiraceae*	B_zOTU1833	7471.36	0.36	*Ruminococcaceae*
B_zOTU6877	-3332.07	0.58	*Blautia*	B_zOTU9426	-191.21	0.36	*Lachnospiraceae*
B_zOTU9105	443.32	0.58	*Ruminococcaceae*	B_zOTU7379	-6608.16	0.36	*Ruminococcaceae*

B_zOTU813	-1089.13	0.57	*Bacteria*	B_zOTU6602	-6496.76	0.36	*Agathobacter rectalis*
B_zOTU1989	853.93	0.57	*Ruminococcaceae*	B_zOTU749	-2916.68	0.36	*Lachnospiraceae*
B_zOTU1092	-4.81	0.56	*Bacteroides*	B_zOTU9208	6665.99	0.35	*Faecalibacterium prausnitzii*
B_zOTU6547	-1641.25	0.56	*Clostridiales*	B_zOTU8285	1715.01	0.35	*Bacteroides*
B_zOTU9288	-27.69	0.56	*Fusicatenibacter saccharivorans*	B_zOTU1655	3791.1	0.35	*Bacteroidaceae*
B_zOTU5255	1568.74	0.55	*Clostridiales*	B_zOTU4667	1277.58	0.34	*Faecalibacterium*
B_zOTU2588	742.31	0.54	*Alistipes*	B_zOTU5764	20829.99	0.33	*Agathobacter*
B_zOTU113	126.51	0.54	*Holdemanella biformis*	B_zOTU1873	2380.24	0.31	*Firmicutes*
B_zOTU3313	66.06	0.53	*Agathobacter*	B_zOTU3686	1196.68	0.31	*Clostridiales*
B_zOTU7517	-1347.43	0.53	*Fusicatenibacter saccharivorans*	B_zOTU8278	-78.03	0.31	*Eggerthella lenta*
B_zOTU3544	1817.06	0.52	*Agathobacter*	B_zOTU7591	-9400.42	0.31	*Phascolarctobacterium*
B_zOTU583	588.5	0.5	*Senegalimassilia anaerobia*	B_zOTU8690	574.37	0.3	*Bacteroidaceae*
B_zOTU472	929.64	0.5	*Clostridiales*	B_zOTU8805	85.26	0.3	*Firmicutes*
B_zOTU634	-1353.75	0.5	*Hydrogeniiclostridium mannosilyticum*	B_zOTU803	509.24	0.3	*Prevotella copri*
B_zOTU5498	12710.04	0.49	*Bacteroides*	B_zOTU1244	-11472.7	0.29	*Dysosmobacter welbionis*
B_zOTU1613	7034.61	0.47	*Ruminococcaceae*	B_zOTU645	982.69	0.28	*Alistipes ihumii*
B_zOTU9537	7681.84	0.46	*Bacteroidales*	B_zOTU1442	-1517.31	0.28	*Veillonella parvula*
B_zOTU8042	-8710.29	0.46	*Lachnospiraceae*	B_zOTU5528	9108.32	0.27	*Lachnospiraceae*
B_zOTU729	363.85	0.44	*Firmicutes*	B_zOTU805	-1536.28	0.26	*Clostridiales*

B_zOTU6217	-836.41	0.43	*Lachnospiraceae*	B_zOTU7555	-425.24	0.26	*Bacteroidaceae*
B_zOTU4061	380.16	0.43	*Bacteroides*	B_zOTU52	-161.65	0.24	*Akkermansia muciniphila*
B_zOTU2144	6279.19	0.43	*Ruminococcaceae*	B_zOTU8131	3151.12	0.23	*Bacteria*

Supplementary Table 12: Lasso-coefficient, and random forest importance of selected features in the final optimized model based on bacteriome-mycobiome. A positive LASSO coefficient indicates association with controls, and negative indicates association with IBS.

ID	Lasso_Coefficient	RF_Importance	Taxonomy	ID	Lasso_Coefficient	RF_Importance	Taxonomy
B_zOTU527	920.49	1.19	*Lachnospiraceae*	B_zOTU5498	6597.9	0.46	*Bacteroides*
B_zOTU3655	-143.4	0.94	*Lachnospiraceae*	B_zOTU583	293.51	0.46	*Senegalimassilia anaerobia*
B_zOTU170	-49.78	0.93	*Bacteroides*	B_zOTU729	71.05	0.45	*Firmicutes*
B_zOTU3628	-109.83	0.89	*Bacteroidetes*	B_zOTU9537	3618.08	0.44	*Bacteroidales*
B_zOTU413	1124.68	0.85	*Lachnospiraceae*	B_zOTU345	-11.75	0.43	*Lachnospiraceae*
M_zOTU5	-0.44	0.84	*Candida albicans*	B_zOTU1613	2419.87	0.43	*Ruminococcaceae*
B_zOTU3395	-1139.83	0.81	*Lachnospiraceae*	B_zOTU7409	3630.8	0.43	*Faecalibacterium prausnitzii*
B_zOTU826	1093.13	0.8	*Clostridiales Incertae Sedis XIII*	B_zOTU8285	1913	0.42	*Bacteroides*
B_zOTU564	401.83	0.78	*Lachnospiraceae*	B_zOTU6156	24538.12	0.42	*Agathobacter rectalis*
B_zOTU6357	-203.43	0.78	*Fusicatenibacter saccharivorans*	B_zOTU2804	2404	0.42	*Lachnospiraceae*
B_zOTU4928	-3650.5	0.78	*Lachnospiraceae*	B_zOTU8674	5051.85	0.41	*Ruminococcaceae*
B_zOTU6877	-1366.59	0.77	*Blautia*	B_zOTU9208	3797.84	0.41	*Faecalibacterium prausnitzii*

B_zOTU1009	4041.47	0.76	*Blautia*	B_zOTU4681	916.93	0.4	*Bacteroidales*
B_zOTU8547	1501.62	0.75	*Ruminococcaceae*	B_zOTU1833	3658.45	0.4	*Ruminococcaceae*
B_zOTU5061	-533.4	0.74	*Bacteroidaceae*	M_zOTU233	9.96	0.39	*Fungi*
B_zOTU4876	-3292.36	0.72	*Lachnospiraceae*	B_zOTU7379	-6851.36	0.39	*Ruminococcaceae*
B_zOTU4249	90.35	0.71	*Lachnospiraceae*	B_zOTU6602	-2647.37	0.39	*Agathobacter rectalis*
B_zOTU8943	-512.29	0.69	*Blautia faecis*	M_zOTU39	-5.72	0.37	*Aspergillus ruber*
B_zOTU9017	279.1	0.66	*Lachnospiraceae*	B_zOTU5528	2987.1	0.37	*Lachnospiraceae*
B_zOTU434	541.05	0.66	*Lachnospiraceae*	B_zOTU1244	-5230.42	0.37	*Dysosmobacter welbionis*
B_zOTU6547	-214.52	0.66	*Clostridiales*	B_zOTU805	-762.91	0.36	*Clostridiales*
B_zOTU9105	161.03	0.62	*Ruminococcaceae*	B_zOTU1873	2049.76	0.35	*Firmicutes*
B_zOTU1410	-218.5	0.6	*Ruminococcaceae*	B_zOTU645	548.85	0.33	*Alistipes ihumii*
B_zOTU7517	-360.59	0.58	*Fusicatenibacter saccharivorans*	B_zOTU2827	-24.4	0.33	*Bacteroides*
B_zOTU113	81.24	0.57	*Holdemanella biformis*	B_zOTU749	-593.32	0.33	*Lachnospiraceae*
B_zOTU2144	1824.61	0.57	*Ruminococcaceae*	B_zOTU1447	0.64	0.32	*Ruminococcaceae*
B_zOTU813	-172.61	0.56	*Bacteria*	B_zOTU5764	13010.54	0.32	*Agathobacter*
B_zOTU6509	-294.23	0.56	*Clostridiales*	B_zOTU8278	-508.25	0.32	*Eggerthella lenta*
B_zOTU3544	0.95	0.54	*Agathobacter*	B_zOTU8690	484.66	0.31	*Bacteroidaceae*
B_zOTU472	365.41	0.54	*Clostridiales*	B_zOTU7591	-5346.94	0.31	*Phascolarctobacterium*
B_zOTU5255	325.68	0.53	*Clostridiales*	B_zOTU1655	659.17	0.3	*Bacteroidaceae*

B_zOTU8042	-955.43	0.51	*Lachnospiraceae*	M_zOTU84	-6.53	0.29	*Wickerhamomyces onychis*
B_zOTU634	-1298.73	0.5	*Hydrogeniiclostridium mannosilyticum*	B_zOTU803	212.18	0.27	*Prevotella copri*
B_zOTU1989	279.4	0.49	*Ruminococcaceae*	M_zOTU59	-14.11	0.26	*Dipodascus geotrichum*
B_zOTU2588	164.15	0.49	*Alistipes*	M_zOTU268	168.44	0.25	*Fungi*
B_zOTU8787	-588.57	0.49	*Gemmiger*	B_zOTU52	-95.58	0.24	*Akkermansia muciniphila*
B_zOTU1439	945.93	0.49	*Lachnospiraceae*	B_zOTU8131	1187.86	0.23	*Bacteria*
B_zOTU8068	-1640.31	0.47	*Streptococcus*				

Supplementary Table 13: Lasso-coefficient, and random forest importance of selected features in the final optimized model based on bacteriome-virome. A positive LASSO coefficient indicates association with controls, and negative indicates association with IBS.

ID	Lasso_Coefficient	RF_Importance	Taxonomy	ID	Lasso_Coefficient	RF_Importance	Taxonomy
B_zOTU527	1904.5	1.32	*Lachnospiraceae*	B_zOTU3515	-1.09	0.46	*Bacteroides faecichinchillae*
B_zOTU413	2792.37	1.04	*Lachnospiraceae*	B_zOTU8068	-27.17	0.45	*Streptococcus*
B_zOTU3655	-111.59	0.97	*Lachnospiraceae*	VC_1311_0	-52.82	0.44	*Podoviridae*
B_zOTU3628	-152.78	0.92	*Bacteroidetes*	VC_881_0	-1884.34	0.43	*Siphoviridae*
B_zOTU170	-38.99	0.87	*Bacteroides*	B_zOTU2804	889.01	0.43	*Lachnospiraceae*
B_zOTU4928	-5266.44	0.84	*Lachnospiraceae*	B_zOTU5764	12303.37	0.42	*Agathobacter*
B_zOTU564	60.23	0.84	*Lachnospiraceae*	B_zOTU483	55.18	0.42	*Paraprevotella clara*
VC_2600_0	11.07	0.82	*Podoviridae*	B_zOTU2100	-1502.74	0.41	*Lachnospiraceae*

B_zOTU9017	888.32	0.8	*Lachnospiraceae*	B_zOTU1244	-5653.4	0.41	*Dysosmobacter welbionis*
B_zOTU4876	-3268.39	0.8	*Lachnospiraceae*	VC_1418_0	72.18	0.41	*Mimiviridae*
B_zOTU8547	2803.3	0.78	*Ruminococcaceae*	B_zOTU8168	-1468.18	0.4	*Ruminococcaceae*
B_zOTU826	244.22	0.77	*Clostridiales Incertae Sedis XIII*	B_zOTU1833	461.89	0.4	*Ruminococcaceae*
B_zOTU434	659.12	0.76	*Lachnospiraceae*	B_zOTU749	-1635.7	0.39	*Lachnospiraceae*
B_zOTU1009	6959.18	0.76	*Blautia*	VC_3089_0	1237.21	0.39	*Siphoviridae*
B_zOTU4087	-828.26	0.72	*Bacteroidaceae*	VC_2330_1	-1.17	0.39	*Unclassified*
B_zOTU9105	786.75	0.68	*Ruminococcaceae*	B_zOTU7379	-10073.5	0.38	*Ruminococcaceae*
B_zOTU8665	-230.37	0.67	*Lachnospiraceae*	VC_2291_0	31.92	0.36	*Siphoviridae*
VC_1952_0	1.41	0.66	*Unclassified*	VC_958_0	5332.66	0.35	*Myoviridae*
VC_500_0	66.33	0.66	*Myoviridae*	B_zOTU8569	2391.86	0.35	*Bifidobacteriales*
B_zOTU7517	-840.59	0.66	*Fusicatenibacter saccharivorans*	VC_3303_0	0.15	0.34	*Siphoviridae*
B_zOTU5061	-561.07	0.64	*Bacteroidaceae*	B_zOTU6602	-1708.61	0.34	*Agathobacter rectalis*
B_zOTU736	-151.47	0.62	*Lachnospiraceae*	B_zOTU8131	5613.53	0.34	*Bacteria*
B_zOTU6877	-849.64	0.61	*Blautia*	VC_2178_0	15013.64	0.33	*Unclassified*
B_zOTU1989	1176.68	0.61	*Ruminococcaceae*	VC_1832_0	3.19	0.32	*Unclassified*
B_zOTU113	132.59	0.6	*Holdemanella biformis*	VC_2349_0	129.04	0.32	*Myoviridae*
B_zOTU472	602.62	0.58	*Clostridiales*	B_zOTU7591	-2868.36	0.32	*Phascolarctobacterium*
B_zOTU6339	2164.45	0.58	*Ruminococcaceae*	B_zOTU803	401.78	0.31	*Prevotella copri*
VC_3688_0	108	0.56	*Podoviridae*	B_zOTU805	-1426.11	0.31	*Clostridiales*
VC_2751_0	17.02	0.56	*Mimiviridae*	VC_287_0	-23.44	0.3	*Siphoviridae*

B_zOTU5255	626.47	0.56	*Clostridiales*	VC_739_0	22.83	0.28	*Podoviridae*
B_zOTU5498	13409.08	0.55	*Bacteroides*	B_zOTU645	496.33	0.28	*Alistipes ihumii*
VC_3377_0	223.45	0.54	*Myoviridae*	B_zOTU210	-55.01	0.28	*Bacteroides eggerthii*
B_zOTU729	88.45	0.54	*Firmicutes*	VC_2313_0	26.51	0.27	*Unclassified*
VC_1803_0	3392.33	0.52	*Myoviridae*	B_zOTU8480	-711.15	0.25	*Bacteroides*
B_zOTU9537	3579.71	0.5	*Bacteroidales*	VC_6909_0	-2205.14	0.23	*Podoviridae*
B_zOTU6156	33277.54	0.48	*Agathobacter rectalis*	B_zOTU9224	-1782.57	0.22	*Alistipes onderdonkii*

Supplementary Table 14: Lasso-coefficient, and random forest importance of selected features in the final optimized model based on mycobiome. A positive LASSO coefficient indicates association with controls, and negative indicates association with IBS.

ID	Lasso_Coefficient	RF_Importance	Taxonomy	ID	Lasso_Coefficient	RF_Importance	Taxonomy

M_zOTU7	-5.71	1.79	*Cladosporium cladosporioides*	M_zOTU84	-60.59	0.6	*Wickerhamomyces onychis*
M_zOTU5	-2.34	1.74	*Candida albicans*	M_zOTU362	0.04	0.6	*Fungi*
M_zOTU21	-4.6	1.69	*Saccharomyces cerevisiae*	M_zOTU13	-19.84	0.59	*Basidiomycota*
M_zOTU91	-0.48	1.63	*Saccharomyces cerevisiae*	M_zOTU88	-125.84	0.57	*Itersonilia perplexans*
M_zOTU10	5.67	1.6	*Fungi*	M_zOTU1474	-214.68	0.57	*Candida*
M_zOTU1	-1.08	1.49	*Saccharomycetales*	M_zOTU19	0.37	0.54	*Saccharomycetales*
M_zOTU919	0.26	1.47	*Candida*	M_zOTU114	-22.9	0.53	*Kazachstania unispora*
M_zOTU6	-3.52	1.17	*Candida dubliniensis*	M_zOTU46	15.15	0.53	*Saccharomycetales*
M_zOTU2	2.48	1.14	*unidentified*	M_zOTU59	-50.75	0.49	*Dipodascus geotrichum*
M_zOTU840	160.88	1.13	*Fungi*	M_zOTU56	19.09	0.48	*unidentified*
M_zOTU41	-46.97	0.96	*Penicillium lanosum*	M_zOTU190	-53.41	0.47	*Aspergillus penicillioides*
M_zOTU4	-4.55	0.94	*Dipodascaceae*	M_zOTU268	312.07	0.45	*Fungi*
M_zOTU11	-8.7	0.9	*Ascomycota*	M_zOTU51	-1.12	0.45	*Penicillium lanosum*
M_zOTU31	-19.1	0.86	*Vishniacozyma victoriae*	M_zOTU104	-2.01	0.45	*Fungi*
M_zOTU18	0.16	0.86	*Aspergillus*	M_zOTU135	110.82	0.44	*Cladosporium sphaerospermum*
M_zOTU1873	266.92	0.84	*Fungi*	M_zOTU100	223.09	0.42	*Dioszegia hungarica*
M_zOTU20	8.96	0.83	*Rhodotorula mucilaginosa*	M_zOTU94	-60.17	0.42	*Penicillium bialowiezense*
M_zOTU37	28.73	0.82	*Aspergillus*	M_zOTU121	22.88	0.42	*unidentified*
M_zOTU174	172.24	0.8	*Neoascochyta paspali*	M_zOTU93	6.56	0.4	*Penicillium bialowiezense*
M_zOTU86	-211.42	0.8	*Microdochium*	M_zOTU111	-255.31	0.39	*Sawadaea bicornis*
M_zOTU118	-15.53	0.8	*Alternaria metachromatica*	M_zOTU110	-29.6	0.39	*Wallemia*
M_zOTU1154	-222.43	0.76	*Fungi*	M_zOTU235	-257.5	0.38	*Cladosporium cladosporioides*
M_zOTU225	-318.26	0.76	*Neoascochyta paspali*	M_zOTU106	-15.64	0.38	*Sclerotinia sclerotiorum*

M_zOTU27	-45.28	0.74	*Candida tropicalis*	M_zOTU123	118.13	0.37	*Torulaspora delbrueckii*
M_zOTU249	43.74	0.74	*Hanseniaspora uvarum*	M_zOTU1699	24546.69	0.37	*unidentified*
M_zOTU28	40.28	0.73	*Tausonia pullulans*	M_zOTU117	-63.58	0.37	*Wickerhamomyces onychis*
M_zOTU42	6.75	0.72	*Vishniacozyma victoriae*	M_zOTU476	-769.9	0.36	*Fungi*
M_zOTU233	39.54	0.68	*Fungi*	M_zOTU1090	-281.24	0.36	*Aureobasidium pullulans*
M_zOTU874	-308.3	0.68	*Malassezia restricta*	M_zOTU209	-77.1	0.34	*Wickerhamomyces anomalus*
M_zOTU116	9.2	0.67	*Penicillium lanosum*	M_zOTU9	-25.78	0.34	*Diutina catenulata*
M_zOTU33	41.5	0.64	*Cystofilobasidium infirmominiatum*	M_zOTU311	-15.62	0.31	*Fungi*
M_zOTU216	229.69	0.62	*Saccharomycetales*	M_zOTU239	135.85	0.25	*Cladosporium cladosporioides*

Supplementary Table 15: Lasso-coefficient, and random forest importance of selected features in the final optimized model based on Mycobiome-Virome. A positive LASSO coefficient indicates association with controls, and negative indicates association with IBS.

ID	**Lasso_Coefficient**	**RF_Importance**	**Taxonomy**	**ID**	**Lasso_Coefficient**	**RF_Importance**	**Taxonomy**
M_zOTU5	-0.29	1.07	*Candida albicans*	VC_2349_0	117.3	0.49	*Myoviridae*
VC_1952_0	21.95	0.96	*Unclassified*	M_zOTU73	0.3	0.49	*Agaricus bisporus*

VC_4621_0	66.29	0.86	*Podoviridae*	M_zOTU11	-0.51	0.49	*Ascomycota*
M_zOTU10	5.65	0.85	*Fungi*	M_zOTU225	-102.01	0.48	*Neoascochyta paspali*
VC_2010_0	96.62	0.83	*Unclassified*	VC_2178_0	12171.43	0.47	*Unclassified*
VC_1632_0	20.75	0.79	*Myoviridae*	VC_464_0	1497.7	0.47	*Myoviridae*
VC_500_0	48.15	0.76	*Myoviridae*	VC_1535_0	32.49	0.47	*Myoviridae*
VC_2761_0	5.63	0.76	*Unclassified*	VC_2872_1	0	0.45	*Unclassified*
VC_2751_0	40.92	0.74	*Mimiviridae*	VC_1811_0	-895	0.45	*Siphoviridae*
M_zOTU2	2.85	0.73	*unidentified*	VC_1418_0	133.92	0.44	*Mimiviridae*
VC_3688_0	136.32	0.68	*Podoviridae*	M_zOTU874	-69.83	0.44	*Malassezia restricta*
VC_3377_0	521.7	0.65	*Myoviridae*	VC_958_0	2343.93	0.44	*Myoviridae*
VC_1803_0	700.91	0.64	*Myoviridae*	VC_881_0	-267.68	0.43	*Siphoviridae*
VC_1510_0	70.74	0.64	*Unclassified*	M_zOTU84	-12.81	0.43	*Wickerhamomyces onychis*
M_zOTU41	-3.64	0.63	*Penicillium lanosum*	VC_2291_0	280.73	0.42	*Siphoviridae*
VC_2825_0	47.66	0.62	*Myoviridae*	VC_1114_0	2321.24	0.41	*Siphoviridae*
M_zOTU37	27.15	0.6	*Aspergillus*	VC_1482_0	-7.2	0.4	*Unclassified*
VC_5989_0	5697.35	0.6	*Podoviridae*	VC_2924_0	1432.95	0.4	*Siphoviridae*
VC_4541_0	906.84	0.59	*Unclassified*	VC_5183_0	1275.3	0.39	*Siphoviridae*
M_zOTU216	399.02	0.59	*Saccharomycetales*	M_zOTU93	5.03	0.38	*Penicillium bialowiezense*
VC_3476_0	358.6	0.58	*Myoviridae*	VC_3303_0	4.59	0.38	*Siphoviridae*
VC_2967_0	10.31	0.56	*Siphoviridae*	M_zOTU59	-9.33	0.38	*Dipodascus geotrichum*
M_zOTU20	10.07	0.55	*Rhodotorula mucilaginosa*	VC_739_0	67.98	0.36	*Podoviridae*
VC_1419_0	10.72	0.54	*Mimiviridae*	VC_1710_0	19.55	0.36	*Unclassified*
VC_1205_0	195.02	0.53	*Unclassified*	VC_3125_0	-4.31	0.35	*Mimiviridae*
M_zOTU39	-2.41	0.53	*Aspergillus ruber*	VC_3978_0	-241.23	0.35	*Myoviridae*
VC_1393_2	33.86	0.52	*Siphoviridae*	VC_2701_0	582.03	0.34	*Siphoviridae*
M_zOTU1154	-40.21	0.51	*Fungi*	VC_740_0	-1.08	0.34	*Podoviridae*
VC_3089_0	1386.88	0.51	*Siphoviridae*	VC_283_3	7.36	0.32	*Siphoviridae*
VC_3641_0	-1315.7	0.51	*Myoviridae*	VC_4117_0	-325.28	0.32	*Unclassified*
VC_1311_0	-94.12	0.5	*Podoviridae*	VC_1217_0	-12.69	0.28	*Unclassified*
VC_1275_0	1472.6	0.5	*Unclassified*	VC_3655_0	-3.8	0.27	*Siphoviridae*
VC_3408_0	8818.52	0.5	*Myoviridae*				

Supplementary Table 16: Lasso-coefficient, and random forest importance of selected features in the final optimized model based on virome. A positive LASSO coefficient indicates association with controls, and negative indicates association with IBS.

ID	Lasso_Coefficient	RF_Importance	Taxonomy	ID	Lasso_Coefficient	RF_Importance	Taxonomy
VC_1632_0	363.93	0.92	*Myoviridae*	VC_2559_0	-0.65	0.45	*Siphoviridae*
VC_1952_0	61.31	0.88	*Unclassified*	VC_1311_0	-121.65	0.44	*Podoviridae*
VC_2761_0	51.26	0.85	*Unclassified*	VC_1811_0	-1036.06	0.44	*Siphoviridae*
VC_500_0	113.73	0.84	*Myoviridae*	VC_3983_0	25.63	0.43	*Siphoviridae*
VC_2751_0	56.08	0.76	*Mimiviridae*	VC_958_0	194.17	0.42	*Myoviridae*
VC_3377_0	2507.65	0.75	*Myoviridae*	VC_882_0	19.54	0.42	*Siphoviridae*
VC_4621_0	10.42	0.71	*Podoviridae*	VC_2872_1	3.32	0.42	*Unclassified*
VC_3688_0	183.84	0.66	*Podoviridae*	VC_5183_0	2421.72	0.42	*Siphoviridae*
VC_5989_0	5089.9	0.64	*Podoviridae*	VC_2701_0	2226.84	0.41	*Siphoviridae*
VC_3089_0	2918.39	0.64	*Siphoviridae*	VC_2534_0	21.79	0.4	*Unclassified*
VC_1803_0	4035.94	0.62	*Myoviridae*	VC_3303_0	16.61	0.4	*Siphoviridae*
VC_2825_0	69.87	0.61	*Myoviridae*	VC_2918_0	322.99	0.4	*Siphoviridae*
VC_640_0	20.21	0.57	*Siphoviridae*	VC_1482_0	-32.34	0.38	*Unclassified*
VC_2667_0	22.93	0.57	*Podoviridae*	VC_1217_1	-2.4	0.38	*Unclassified*
VC_2916_0	70.58	0.55	*Myoviridae*	VC_3978_0	-644.15	0.37	*Myoviridae*
VC_3476_0	723.8	0.55	*Myoviridae*	VC_650_0	11.11	0.37	*Unclassified*
VC_464_0	2597.23	0.54	*Myoviridae*	VC_3601_0	-144.45	0.36	*Siphoviridae*
VC_1205_0	986.7	0.54	*Unclassified*	VC_2400_0	93.5	0.36	*Myoviridae*
VC_1114_0	5028.02	0.53	*Siphoviridae*	VC_739_0	91.8	0.34	*Podoviridae*

VC_2967_0	60.1	0.51	*Siphoviridae*	VC_1902_0	-122.46	0.32	*Unclassified*
VC_881_0	-892.27	0.51	*Siphoviridae*	VC_4189_0	2018.11	0.32	*Unclassified*
VC_1393_2	2.15	0.51	*Siphoviridae*	VC_1290_0	-1032.88	0.32	*Inoviridae*
VC_2351_0	356.59	0.51	*Myoviridae*	VC_829_0	-40.37	0.32	*Siphoviridae*
VC_3408_0	801.4	0.51	*Myoviridae*	VC_283_3	4.8	0.32	*Siphoviridae*
VC_734_1	6.48	0.51	*Unclassified*	VC_2412_0	0.8	0.31	*Unclassified*
VC_2349_0	338.28	0.5	*Myoviridae*	VC_2146_0	-73.39	0.31	*Unclassified*
VC_3809_0	1170.57	0.49	*Unclassified*	VC_3125_0	-34.64	0.3	*Mimiviridae*
VC_1301_0	147.82	0.48	*Unclassified*	VC_1883_0	-6.09	0.3	*Mimiviridae*
VC_1419_0	125.24	0.47	*Mimiviridae*	VC_4117_0	-3047.15	0.3	*Unclassified*
VC_1177_2	74.32	0.47	*Unclassified*	VC_1217_0	-34.5	0.3	*Unclassified*
VC_1535_0	66.23	0.46	*Myoviridae*	VC_1710_0	44.19	0.29	*Unclassified*
VC_3359_0	9.28	0.46	*Myoviridae*	VC_4268_0	-677.3	0.27	*Podoviridae*
VC_3641_0	-2081.56	0.46	*Myoviridae*	VC_6909_0	-5021.57	0.27	*Podoviridae*
VC_1275_0	675.2	0.46	*Unclassified*	VC_2562_0	-45.62	0.26	*Unclassified*
VC_1418_0	221.49	0.45	*Mimiviridae*	VC_1716_0	-87.65	0.23	*Myoviridae*

Supplementary Table 17: Lasso-coefficient, and random forest importance of selected features in the final optimized model based on all microbiome components. A positive LASSO coefficient indicates association with controls, and negative indicates association with IBS.

ID	Lasso_Coefficient	RF_Importance	Taxonomy	ID	Lasso_Coefficient	RF_Importance	Taxonomy
B_zOTU527	1315.82	1.52	*Lachnospiraceae*	B_zOTU9537	1314.08	0.53	*Bacteroidales*
B_zOTU413	1866.89	1.09	*Lachnospiraceae*	VC_1803_0	2298.58	0.52	*Myoviridae*
B_zOTU3628	-111.27	0.99	*Bacteroidetes*	M_zOTU37	5.51	0.51	*Aspergillus*
B_zOTU170	-31.95	0.98	*Bacteroides*	VC_1311_0	-14.48	0.48	*Podoviridae*
B_zOTU1009	4835.65	0.9	*Blautia*	B_zOTU2804	748.18	0.48	*Lachnospiraceae*
B_zOTU9017	451.23	0.89	*Lachnospiraceae*	B_zOTU8569	1524.38	0.45	*Bifidobacteriales*

B_zOTU3655	-121.29	0.87	*Lachnospiraceae*	B_zOTU1244	-2749.08	0.44	*Dysosmobacter welbionis*
B_zOTU564	131.12	0.85	*Lachnospiraceae*	B_zOTU1833	110.22	0.43	*Ruminococcaceae*
B_zOTU4928	-3587.97	0.85	*Lachnospiraceae*	M_zOTU39	-8.94	0.43	*Aspergillus ruber*
B_zOTU4876	-2427.16	0.84	*Lachnospiraceae*	B_zOTU7379	-7394.74	0.42	*Ruminococcaceae*
VC_1952_0	5.88	0.84	*Unclassified*	VC_1832_0	2.36	0.41	*Unclassified*
B_zOTU826	259.15	0.81	*Clostridiales Incertae Sedis XIII*	VC_881_0	-1214.68	0.41	*Siphoviridae*
B_zOTU8547	2057.37	0.81	*Ruminococcaceae*	B_zOTU8168	-411.43	0.4	*Ruminococcaceae*
B_zOTU434	427.05	0.81	*Lachnospiraceae*	B_zOTU6131	0.26	0.4	*Prevotella copri*
VC_500_0	41.25	0.8	*Myoviridae*	VC_2349_0	94.16	0.39	*Myoviridae*
B_zOTU8665	-243.89	0.78	*Lachnospiraceae*	B_zOTU2100	-2662.43	0.39	*Lachnospiraceae*
B_zOTU6877	-493.41	0.77	*Blautia*	VC_2178_0	10489.93	0.39	*Unclassified*
B_zOTU5312	-2.04	0.77	*Ruminococcaceae*	VC_3089_0	728.48	0.39	*Siphoviridae*
B_zOTU5061	-439.35	0.74	*Bacteroidaceae*	VC_958_0	2879.91	0.38	*Myoviridae*
B_zOTU9105	393.35	0.73	*Ruminococcaceae*	B_zOTU7591	-1511.36	0.38	*Phascolarctobacterium*
B_zOTU7517	-151.54	0.72	*Fusicatenibacter saccharivorans*	B_zOTU749	-909.79	0.37	*Lachnospiraceae*
B_zOTU4087	-420.58	0.7	*Bacteroidaceae*	VC_739_0	5.95	0.37	*Podoviridae*
VC_3377_0	163.99	0.65	*Myoviridae*	B_zOTU805	-589.07	0.37	*Clostridiales*
B_zOTU113	71.11	0.65	*Holdemanella biformis*	B_zOTU645	302.98	0.36	*Alistipes ihumii*
VC_2751_0	4.75	0.64	*Mimiviridae*	B_zOTU803	224.11	0.35	*Prevotella copri*
B_zOTU5498	8793.43	0.64	*Bacteroides*	B_zOTU5764	8211.59	0.35	*Agathobacter*

VC_3688_0	68.2	0.61	*Podoviridae*	B_zOTU210	-11.47	0.34	*Bacteroides eggerthii*
B_zOTU1989	501.24	0.6	*Ruminococcaceae*	VC_1418_0	32.09	0.34	*Mimiviridae*
B_zOTU472	333.5	0.6	*Clostridiales*	B_zOTU8480	-120.03	0.31	*Bacteroides*
B_zOTU6339	1347.48	0.59	*Ruminococcaceae*	B_zOTU8131	2353.03	0.31	*Bacteria*
B_zOTU634	-142.54	0.58	*Hydrogeniiclostridium mannosilyticum*	VC_2313_0	13.72	0.31	*Unclassified*
B_zOTU729	2.01	0.58	*Firmicutes*	B_zOTU9224	-731.94	0.23	*Alistipes onderdonkii*
B_zOTU6156	21419.07	0.53	*Agathobacter rectalis*				

Chapter 6

Bacterial community-level metabolic modelling reveals differences in metabolic capabilities and interactions within the gut microbial ecosystem in Irritable Bowel Syndrome.

Authors: Anubhav Das, Paul W. O'Toole, Ian B. Jeffery*

*Corresponding author.

Manuscript under preparation

Author contributions: AD, and IBJ conceived the study idea and design. AD, PWOT, and IBJ drafted and reviewed the manuscript. AD and IBJ., performed the bioinformatics analysis.

AD Contribution:

Data processing: Bacterial metagenomic composition generated in Chapter 2 was re-used.

Analysis: Genome scale model simulations, and species knock-out simulations. Flux Balance analysis, hypothesis testing and multivariate statistical analyses. Co-operation and competition network generation and analysis. Elasticity analysis. Development of Machine learning models and subsequent analysis.

Others: Contribution to manuscript by drafting Introduction, Methods, Results, and Discussion sections.

Table of contents

6.1 Abstract 392
6.2 Introduction 393
6.3 Methods 394
6.3.1 Participants and Cohort information 394
6.3.2 Metagenomic sample processing 394
6.3.3 Genome scale metabolic model generation 396
6.3.4 Community growth simulations 396
6.3.5 Taxon knock-out for estimation of co-operation and competition in a community 397
6.3.6 Clustering analysis of network and node properties 398
6.3.7 Network-based predictive analysis 398
6.3.8 Classification threshold optimization 399
6.3.9 Perturbation analysis 399
6.3.10 Statistical analysis 400
6.4 Results 401
6.4.1 Metabolic model summary 401
6.4.2 Reaction flux profiles are different in IBS and Controls, but not between IBS subtypes 401
6.4.3 Inter-bacterial metabolic interactions are reduced in IBS 408
6.4.4 The co-operation network is predictive of IBS 411
6.4.5 Cluster analysis of microbiome structure 414
6.4.6 Highly interactive species strongly modulate the metabolic output of the microbiome 417
6.5 Discussion 424
6.6 References 429
6.7 Supplementary information 434
6.7.1 Supplementary Figures 434
6.7.2 Supplementary Tables 443

6.1 Abstract

With a recent shift in paradigm towards microbiome-based therapeutics and diagnostics for multiple intestinal and extra-intestinal diseases, it is imperative to first understand the underlying mechanisms of intra community interaction and metabolic output of the microbiome communities. Although understanding the biology of each bacterial species is crucial, to modulate the whole microbiome rationally, knowledge of the biological dynamics of the whole microbiome community is also essential. For this, application of computational methods and systems biology-based approaches will be pivotal.

Here, I report metabolic modelling of the gut bacteria of 80 IBS patients, and 59 matched Control subjects, based on bacterial composition and available species level Genome-scale metabolic network reconstructions (GENREs). Using Flux Balance Analysis (FBA), I have shown differences in predicted metabolic output of bacteria in IBS and Controls, and changes in uptake and secretion of metabolites which can potentially participate in inter-bacterial metabolic cross-feeding. Furthermore, I investigated differences in co-operative and competitive interaction networks, wherein I observed reduction in inter-bacterial co-operation in IBS, and provided a distinct taxonomic-metabolomic signature of dysbiosis in IBS. Finally, I investigated potential changes in metabolic output of the bacterial community upon increase in relative abundance of each species. The applied methodology, and findings reported here can help advance the research on rational microbiome modulation, and prediction and development of live biotherapeutic products in a cost-effective and time efficient manner.

6.2 Introduction

The gut microbiota performs a plethora of functions which are essential for maintenance of the human health, and alterations in these are directly implicated in pathogenesis of functional gastrointestinal disorders like IBS. Compositional and functional analysis, metabolomics, and integrated multi-omics analysis approaches have shown differentiating features in IBS studies along with changes in microbiome structure based on abundance correlation, which have strengthened our perception regarding the role of microbiota in IBS pathophysiology [1–5]. Notwithstanding the extensive microbiome studies in understanding of IBS, there is scope for further exploration of the mechanistic insights and bacterial inter-play based on metabolic interdependencies.

To understand the functional capabilities of the gut bacterial community, and the metabolic dependency based microbial community dynamics in IBS, I used a constraint-based modelling framework, MiCom [6], based on cobrapy [7] which employs COBRA (Constraint based Reconstruction and Analysis) methods [8,9]. I analysed the microbial community level genome-scale metabolic models, based on bacterial abundance, of a previously published cohort comprising of 80 IBS patients and 59 matched Controls. Based on Flux Balance Analysis (FBA), I have shown differences in the predicted rates of production/consumption of molecules relevant to IBS. Based on a minimal media-based optimization, differences in simulated exchanges (import/export) by bacterial species have also been investigated, and I observed changes in co-operative and competitive bacterial networks based on putative metabolic dependencies. I also developed a metric, based on clustering of network properties which was predictive of IBS. Finally, using simulation strategies, I investigated the effect of increase in abundance of each species on overall metabolic

output of the bacterial communities. The results suggest a global difference in metabolic capacities between microbial communities of IBS and Controls, along with significant loss of inter-microbial co-operation in IBS, which are consistent with previously published studies. Predictive analysis using information from interaction networks show high predictive power of the machine learning models. Findings from perturbation simulations indicated novel mechanistic insights in IBS, in addition to revealing species which can potentially modulate the metabolic output of the bacterial community.

6.3 Methods

6.3.1 Participants and Cohort information

The cohort analysed in this study has been reported previously [5,10,11], and consists of 80 IBS patients (Rome IV criteria), along with 65 Controls which are age, gender, ethnicity, and geography matched. Based on symptoms, the IBS subjects have been further classified as IBS-C (Constipation predominant), IBS-D (Diarrhoea predominant), and IBS-M (Mixed). A summary of the cohort is listed in Table 1. I analysed a subset of the cohort consisting of 80 IBS patients, and 59 Controls, with information of bacterial species abundance based on metagenomic sequences.

6.3.2 Metagenomic sample processing

I re-analysed the shotgun metagenomic based abundance data reported previously [5]. Reads were processed according to the standard operating procedures of Human Microbiome Project (HMP) Consortium [12]. The processed reads were then passed through HUMAnN2 pipeline [13]. The metagenomic abundance, based on MetaPhlAn2 [14], were used to generate abundance table and species level abundance was extracted.

	Control (n = 65)	IBS (n = 80)
Age range, years (mean)	19-65 (45)	17-66 (39)
Sex (male/female)	16/49	15/65
BMI Class, n (%)		
Normal	25 (38)	31 (39)
Obese Class I	11 (17)	14 (18)
Obese Class II	3 (5)	5 (6)
Obese Class III	1 (2)	3 (4)
Overweight	21 (32)	22 (28)
Underweight	3 (5)	3 (4)
HADS: Anxiety, n (%)		
Normal (0-10)	59 (91)	58 (72)
Abnormal (11-21)	6 (9)	22 (28)
HADS: Depression, n (%)		
Normal (0-10)	64 (98)	70 (88)
Abnormal (11-21)	1 (2)	10 (12)
Bristol Stool Score, n (%)		
Normal	54 (83)	18 (22)
Constipated	8 (12)	22 (28)
Diarrhoea	3 (5)	40 (50)
IBS subtype, n (%)		
IBS-C		30 (38)
IBS-D	N/A	21 (26)
IBS-M		29 (36)
SeHCAT assayed, n (%)	9 (14)	46 (57)
Dietary group (FFQ), n (%)		
Omnivore	63 (97)	74 (92)
Vegetarian	1 (2)	2 (2)
Pescatarian	1 (2)	1 (1)
Gluten-free	0 (0)	4 (5)
Drinks alcohol, n (%)		
Current	54 (83)	57 (71)
Previous	0 (0)	1 (1)
Never	10 (15)	22 (28)
Smoker, n (%)		
Current	10* (15)	14* (18)
Previous	13 (20)	18 (22)
Never	42 (65)	48 (60)

* 1 subject in each group smoked e-cigarettes

BMI: Body Mass Index, FFQ: Food Frequency Questionnaire, HADS: Hospital Anxiety and Depression Scale, N/A: not applicable.

Table 1: Descriptive statistics of Control and IBS subjects of the analysed cohort.

6.3.3 Genome scale metabolic model generation

To generate gut bacterial community level genome scale metabolic models, I used MICOM [6], a cobrapy [7] based framework, and AGORA (assembly of gut organisms through reconstruction and analysis)[15] based species level metabolic models [16]. For each metagenomic sample, species with abundance greater than or equal to 0.005% were considered for generation of community level metabolic model generation. Species-level metabolic models corresponding to the selected species along with their abundances were used to generate the community level metabolic models. To further constrain the rate of metabolite exchange rates, I set the fluxes of the community media to that of an average western diet [17]. A community trade-off value was calculated using the function available in MICOM, such that maximum growth was obtained for each community. The obtained value was 0.5.

6.3.4 Community growth simulations

For community growth simulations, I used MICOM along with OSQP [18], a quadratic programming solver, for optimization. To estimate the rates of reaction within the members of the community, I performed parsimonious flux balance analysis (pFBA) with an objective of maximizing the overall community biomass. This approach returned a unique solution based on the optimization, wherein, fluxes (reaction rates in millimoles per gram dry weight per hour, or mmol/[gDW h]) through each reaction was obtained. Cellular reactions (excluding exchange reactions) which produced or consumed at least one extra-cellular metabolite were considered for further analysis.

To investigate further, the community models were grown with a minimal media approach (minimization of total import flux from media). This yielded metabolites and their fluxes through each species compartment, which can potentially be exchanged (import/export) by other species.

All the reaction fluxes were capped at 10^{-6}, i.e., the absolute values of fluxes < 10^{-6} were set to 0. The resulting fluxes were then log10 transformed according to equation (1).

$$v_{log} = sign(v) * (log_{10}(abs(v) + 10^{-6}) + 6) \quad \text{equation (1)}$$

Where, v_{log} is the log transformed reaction flux, and v is the reaction flux.

6.3.5 Taxon knock-out for estimation of co-operation and competition in a community

To understand inter-microbial dynamics in the community, by estimation of extent of co-operation and competition, I adopted a taxon knock-out approach. In each community model I sequentially removed one species at a time and calculated the change in growth rate of other species present in the community. If the growth rate of a species was reduced upon elimination of another knocked out species, the interaction between them was considered as co-operative, while an increase in growth rate was considered as competition. Based on change in growth rate I generated co-operation and competition networks, wherein an interaction was considered strong if the absolute change was greater than 0.1. In some cases, I observed outlying values for the growth rate changes. To control the effect of this, I capped the absolute growth rate to 10. Networks were generated, and network and node properties were calculated using igraph (v. 1.2.6) in R 4.1.

To generate mean networks of IBS and Control populations, each of unfiltered co-operation and competition networks of IBS and Control samples were considered separately. For each network, edges present in greater than 10% of the samples were selected, and average of the edge weights were calculated. Edges with mean weights > 0.1 were selected, and an undirected graph was generated for co-operation/competition networks of IBS and Controls were generated and visualized

using igraph (v. 1.2.6) in R 4.1. To check for significance, the Fisher's exact test was performed, and resulting p-values were adjusted using the Benjamini-Hochberg method.

6.3.6 Clustering analysis of network and node properties

Node properties like betweenness, degree (inward and outward) of species in each sample, along with their abundance, and network properties like reciprocity of both co-operative and competitive networks were integrated and clustered using iBBiG [19]. The integrated table was run with iBBiG for 100 iterations, and modules, identified in >= 50% of the iterations, were considered for further analysis.

Predictive analysis using biologically relevant modules was performed using support vector machines (SVMs) and Leave-One-Out cross validation strategy, implemented in R4.1 using packages caret and e1071. Classification threshold of prediction probability was also optimized for stratification of samples using ROC package in R 4.1.

6.3.7 Network-based predictive analysis

For each sample, I calculated a metric called the co-operation index of each node based on their nearest neighbours in the co-operative network. For each network, sub-graphs were generated with first neighbours of each node, followed by summation of all the edge weights in the sub-graph, which I refer as the co-operation index. I generated a table based on such co-operative indices of the taxa in each sample, which was then scalar normalized across species, and used for predictive analysis. I applied our previously published machine learning pipeline [5], and analysed predictive model performance based on a 10 fold cross-validation strategy.

The machine learning pipeline mainly used glmnet, randomforest, and caret packages and was implemented in R 4.1. I also identified optimized thresholds returning maximum sensitivity at specificity of 0.8 and 0.9 along with optimized threshold based on maximum of summation of sensitivity and specificity. Threshold optimization was performed using ROC package in R 4.1.

6.3.8 Classification threshold optimization

Predictive analysis using classification algorithms provides predictive model performance metrics like sensitivity (accuracy of classifying a sample as IBS), and specificity (accuracy of classifying a sample as Control) based on default classification threshold which is mostly 0.5.

A predictive model with high sensitivity and low specificity can identify IBS patients with high accuracy but will also classify someone without the disease as a patient leading to rise in false positives. As a result, disease free subjects might be put through further unnecessary investigations. Conversely, a predictive model with low sensitivity and high specificity will misclassify the IBS patients as non-IBS controls. Therefore, a balance between the two metrics is required, and hence, classification thresholds were optimized to stratify subjects as per the defined requirements.

I optimized the classification probability thresholds based on three objectives, which were maximum of the sum of sensitivity and specificity (best), a specificity of 0.8 (SP80), and a specificity of 0.9 (SP90), using ROC package in R 4.1.

6.3.9 Perturbation analysis

The gut bacterial community perturbation simulation was performed by calculating a metric named elasticity coefficient (elasticities), using the functions provided in MICOM. It is a dimensionless normalized measure which represents rate of change of

exchange (import/export) fluxes upon change in abundance of species. For each community model, the abundance of each species was increased by 10%, and elasticities of exchange reactions (imports/exports) of the whole community was calculated. A positive elasticity coefficient indicated a stimulation of exchange rate upon increase in abundance of the species, while a negative coefficient indicated inhibition of exchange rate. For IBS and Control populations, I calculated the average elasticities of each exchanged metabolite-species pair, and features with absolute values of mean elasticities > 0.1 (corresponding to 10% change) were selected for further analysis.

To identify species with similar impact on metabolic output of the community, I generated mean elasticity profiles of species for imports and exports in IBS and Controls. These profiles were visualized as heatmap generated using Euclidean distance and ward linkage using made4 package in R 4.1.

6.3.10 Statistical analysis

Global differences in flux distributions were analysed and visualized by Principal Co-ordinate analysis (PCoA) with spearman distance, using the ade4 package in R.

Univariate differences in flux profiles were checked using two sample Welch's t-test, and nominal p-values < 0.01 were considered as significant. In other cases, Wilcoxon rank sum test was applied, and p-values were adjusted using Benjamini-Hochberg method. Adjusted p-values < 0.05 were considered as significant.

In perturbation analysis, the number of affected metabolites by species were compared between multiple groups, and statistical significance was tested using Kruskal-Wallis test, followed by post-hoc Dunn’s test between each group using dunn.test package in

R4.1. The obtained p-values were adjusted using Benjamini-Hochberg method, and an adjusted p-value < 0.05 was considered statistically significant.

6.4 Results

6.4.1 Metabolic model summary

Considering all community level metabolic models, the community comprised of 118 species, with each model consisting of a median of 23 species (IQR: 20-26). The median number of metabolites of each species was 1,434 (IQR: 1161 - 1500), which were involved in a median of 1,962 (IQR: 1381 - 2418) reactions.

6.4.2 Reaction flux profiles are different in IBS and Controls, but not between IBS subtypes

The internal reactions involving at least one extra-cellular metabolite, the predicted reaction rate profile was observed to be significantly different between IBS and Control populations (p-value = 0.001), but it failed to significantly distinguish between the clinical IBS sub-types (p-value = 0.411). The separation between IBS and Control populations, and IBS sub-types based on Principal Co-ordinate Analysis (PCoA) using spearman distance are shown in Figure 1 and Supplementary figure 1A, respectively.

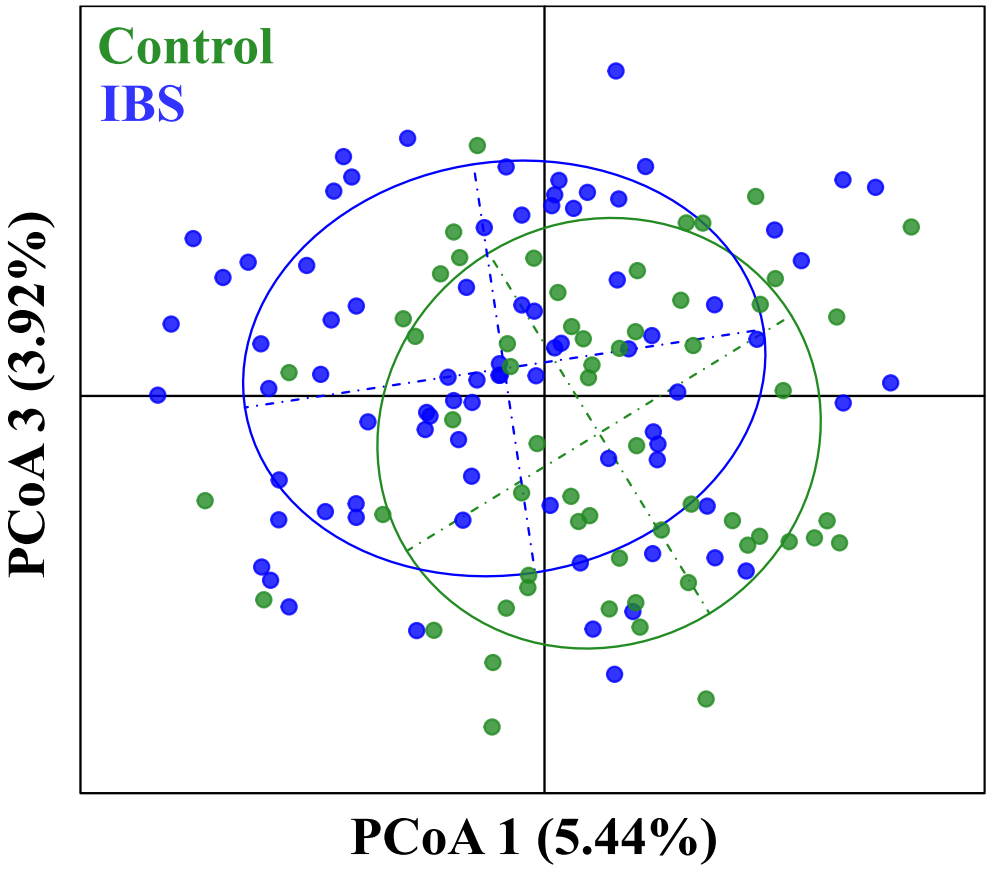


Figure 1: Principal co-ordinate analysis (PCoA) based on Spearman distance of log transformed stratified flux profiles of internal reactions showing significant differences between IBS and controls (p=0.001; control, n=59; IBS, n=80).

Univariate comparison yielded 298 predicted reactions in 26 species which were significantly different between IBS and Controls (Supplementary table 1). Out of these reactions, fluxes of 165 reactions were observed to be higher in IBS, and fluxes of 133 reactions were lower in IBS compared to controls. A summary of some reactions relevant to gut health is listed in Table 1. Multiple critical metabolites were predicted which may contribute to the pathophysiology of IBS. For example, reactions related to mucin degradation and assimilation of its catabolites, along with production of T-antigens and F1-antigens were overrepresented in IBS. Furthermore, there were noted differences in the contributing species. While in IBS, the major drivers were

Bacteroides vulgatus, Bacteroides uniformis, Roseburia inulinivorans, Bacteroides ovatus, Ruminococcus torques, Dorea formicigenerans, and *Roseburia inulinivorans,* in controls the reactions were driven by *Barnesiella intestinihominis,* and *Ruminococcus lactaris*.

Simulated production rates of short chain fatty acids (SCFAs) were observed to be similar, but the contributing species differed in IBS (*Dorea formicigenerans, Ruminococcus torques*) compared to Controls (*Odoribacter splanchnicus, Barnesiella intestinihominis*). Based on the simulations, hypoxanthine and glycine consumption by *Ruminococcus torques* were observed to be higher in IBS, and hydrogen sulphide consumption by *Barnesiella intestinihominis* and 4-amino butanoate (GABA) production by *Odoribacter splanchnicus* were lower in IBS. The number of significantly different reactions pertaining to bile acid transformation was similar in IBS and Controls, but contributed differently by *Dorea formicigenerans*, and *Alistipes putredinis*, respectively.

Reaction	Species	Reactions (N)
Elevated in IBS		
Mucin degradation	*Bacteroides vulgatus, Bacteroides uniformis, Roseburia inulinivorans, Bacteroides ovatus, Ruminococcus torques*	32
T-antigens, F1 alpha antigen	*Bacteroides vulgatus, Ruminococcus torques, Dorea formicigenerans, Roseburia inulinivorans*	6
Formate	*Dorea formicigenerans, Ruminococcus torques*	2
Acetate	*Dorea formicigenerans, Ruminococcus torques*	2
Bile acid transformation	*Dorea formicigenerans*	1
Hypoxanthine consumption	*Ruminococcus torques*	1
Glycine consumption	*Ruminococcus torques*	1
Depleted in IBS		
Mucin degradation	*Barnesiella intestinihominis, Ruminococcus lactaris*	11
T-antigens	*Barnesiella intestinihominis*	2
Formate	*Odoribacter splanchnicus*	1
Acetate	*Odoribacter splanchnicus, Barnesiella intestinihominis*	2
Bile acid transformation	*Alistipes putredinis*	1
H_2S consumption	*Barnesiella intestinihominis*	1
4-amino butanoate (GABA) production	*Odoribacter splanchnicus*	1

Table 1: A summary of selected relevant and significantly different reactions and their contributing species in IBS (N=80) and Controls (N=59).

To further investigate the predicted rate of production (import) and consumption (export) of metabolites by the bacterial species in IBS and Control microbiomes, I adopted a minimal medium approach to grow the models in the simulation environment. Overall, the predicted profiles of exported and imported metabolites could significantly differentiate between IBS and Controls (p-value = 0.001 for both

export and import profiles). This significant separation based on export and import profiles using PCoA and spearman distance are shown in Figure 2B and C, respectively. The export and import profiles were unable to distinguish between the clinical sub-types of IBS (p-value = 0.754 for exports, and p-value = 0.655 for imports; supplementary figure 1B and C, respectively).

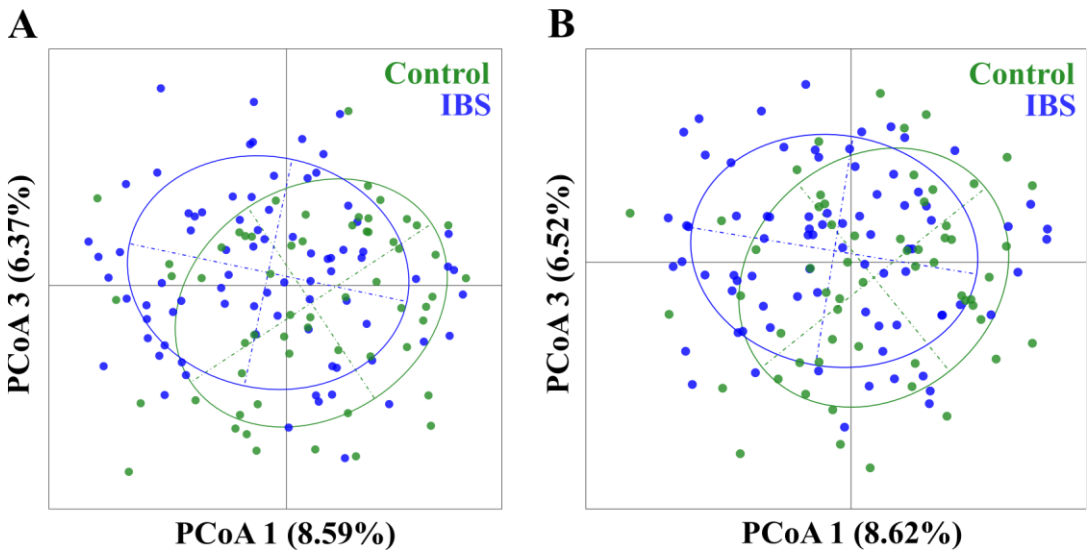


Figure 2: Analysis of differences in metabolite exchanges between IBS (N=80) and Controls (N=59). A.) Principal co-ordinate analysis (PCoA) based on spearman distance of log transformed stratified export profiles of metabolites showing significant differences between IBS and controls (p=0.001). B.) PCoA based on spearman distance of log transformed stratified import profiles of metabolites showing significant differences between IBS and controls (p=0.001).

Among the metabolites predicted to be exported, 75 exports consisting of 47 metabolites from 14 species were found to be significantly different between IBS and Controls, where 50 such exports were exported with a higher rate, and 24 with a lower rate in IBS (supplementary table 2). A few significantly different metabolites relevant to IBS are listed in Table 2. It is interesting to note that the predicted export of bio-transformed bile acids by *Ruminococcus torques*, and *Collinsella aerofaciens* were

among the exports which were significantly higher in IBS. Simulated exports of SCFAs were observed to be higher in both IBS and Controls, but the exporting species were observed to be different. While SCFAs significantly exported at higher rates (formate, acetate, or propionate) in IBS were exported by *Ruminococcus torques*, and *Dorea formicigenerans*, these were exported at lower rates by *Barnesiella intestinihominis, Odoribacter splanchnicus* in IBS. Furthermore, simulated export of N-acetyl-D-glucosamine (a mucin sugar) by *Ruminococcus torques* glycine and hypoxanthine by *Dorea formicigenerans* were higher in IBS compared to controls. Also, predicted glycerol export by *Ruminococcus torques*, and *Bifidobacterium longum* were higher in IBS compared to Controls, and it was lower in case of *Barnesiella intestinihominis*. I also observed a significantly lower rate of simulated export of 4-amino butyrate (GABA) by *Odoribacter splanchnicus* in IBS.

Metabolite	Species
High in IBS	
Acetate	*Ruminococcus torques, Dorea formicigenerans*
Formate	*Ruminococcus torques, Dorea formicigenerans*
Propionate	*Ruminococcus torques*
Ursocholate (bile acid)	*Ruminococcus torques*
12-Dehydrocholate (bile acid)	*Collinsella aerofaciens*
Chenodeoxycholate (bile acid)	*Collinsella aerofaciens*
Glycine	*Dorea formicigenerans*
Hypoxanthine	*Dorea formicigenerans*
N-acetyl-D-glucosamine	*Ruminococcus torques*
Glycerol	*Ruminococcus torques, Dorea formicigenerans, Bifidobacterium longum*
Low in IBS	
Hydrogen	*Odoribacter splanchnicus*
Acetate	*Barnesiella intestinihominis, Odoribacter splanchnicus*
Propionate	*Barnesiella intestinihominis*
Butyrate	*Odoribacter splanchnicus*
Glycerol	*Barnesiella intestinihominis*
4-aminobutanoate	*Odoribacter splanchnicus*

Table 2: Summary of significantly different and biologically relevant exported metabolites and the contributing species in IBS (N=80) and Controls (N=59).

Based on simulation studies, among the significantly different imported metabolites, I observed 130 imports consisting of 93 metabolites from 21 species, where 80 such imports occurred with a higher rate, and 50 with a lower rate in IBS (supplementary table 3). Among the few selected activities relevant to IBS pathophysiology, listed in Table 3, the predicted import of bile acids by *Ruminococcus torques*, *Collinsella aerofaciens*, and *Roseburia intestinalis* were overrepresented in imports happening at a higher rate in IBS. Consumption of hypoxanthine by *Ruminococcus torques* was also observed to be higher in IBS. A lower rate of consumption of N-acetylneuraminate,

and Hydrogen sulphide by *Barnesiella intestinihominis*, and glycine by *Lachnospira eligens* in IBS, was also observed. Along with the selected activities mentioned in Table 3, I also observed a predicted greater rate of consumption of amino acids, vitamins, and sugars by the microbiota in IBS compared to controls.

Metabolite	**Species**
High in IBS	
3alpha,12alpha-Dihydroxy-7-oxo-5beta-cholanate (bile acid)	*Ruminococcus torques*
N-acetylneuraminate	*Ruminococcus torques*
7-Dehydrochenodeoxycholate (bile acid)	*Ruminococcus torques*
3-Dehydrochenodeoxycholate (bile acid)	*Collinsella aerofaciens*
3-Dehydrocholate (bile acid)	*Collinsella aerofaciens*
Cholate (bile acid)	*Roseburia intestinalis*
Hypoxanthine	*Ruminococcus torques*
Low in IBS	
N-acetylneuraminate	*Barnesiella intestinihominis*
Hydrogen sulfide	*Barnesiella intestinihominis*
Glycine	*Lachnospira eligens*

Table 3: List of significantly different and biologically relevant imported metabolites and the contributing species in IBS (N=80) and Controls (N=59).

6.4.3 Inter-bacterial metabolic interactions are reduced in IBS

Network analysis constructed from taxon knockouts (reflecting potential metabolic dependencies and overlap) revealed significant differences in the microbiome structure based on putative co-operation and competition between members of the bacterial community. The mean co-operation network was observed to be sparse in IBS compared to controls (Figure 3), where the network of control samples had 30

nodes with 165 connections and IBS had 11 nodes with 12 connections. The node and edge distributions of the co-operation network of IBS was significantly different from that of Controls. (Table 4).

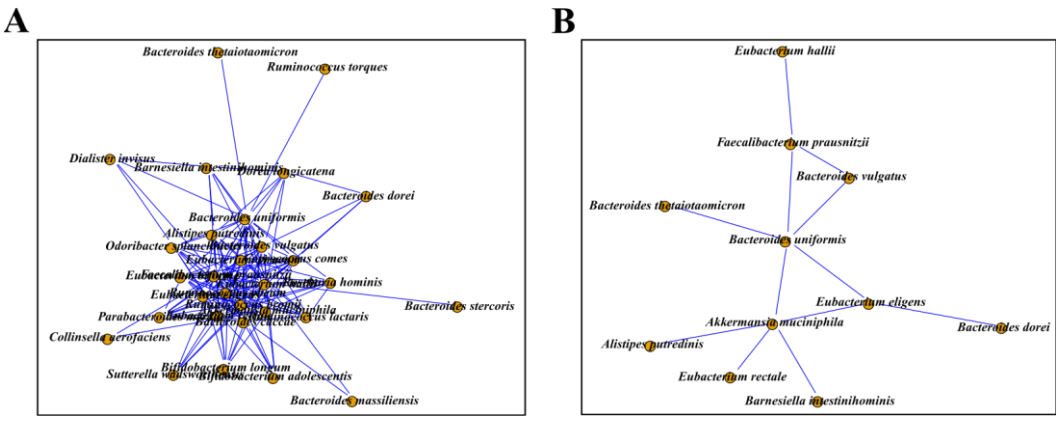


Figure 3: Inter-bacterial metabolic co-operation networks of A.) Controls (Nodes= 30, Edges= 165), and B.) IBS (Nodes= 11, Edges= 12) showing significant and drastic reduction in co-operation between bacteria in IBS. The differences were significant for node (Adjusted p-value < 0.001) as well as edge distribution (Adjusted p-value < 0.001) based on Fisher's exact test.

In contrast, the competition network of IBS and Controls were consistently sized (Figure 4). The competition network of IBS included 47 nodes and 181 connections, and the network of controls contained 47 nodes and 331 connections. The networks were not significantly different in terms of node distribution but differed significantly based on edges showing that similar taxa were associated with microbiome structures, but these contained significantly altered connections in IBS (Table 4). These results indicate a reduction in both co-operation and competition intra-community interactions between the community members in IBS.

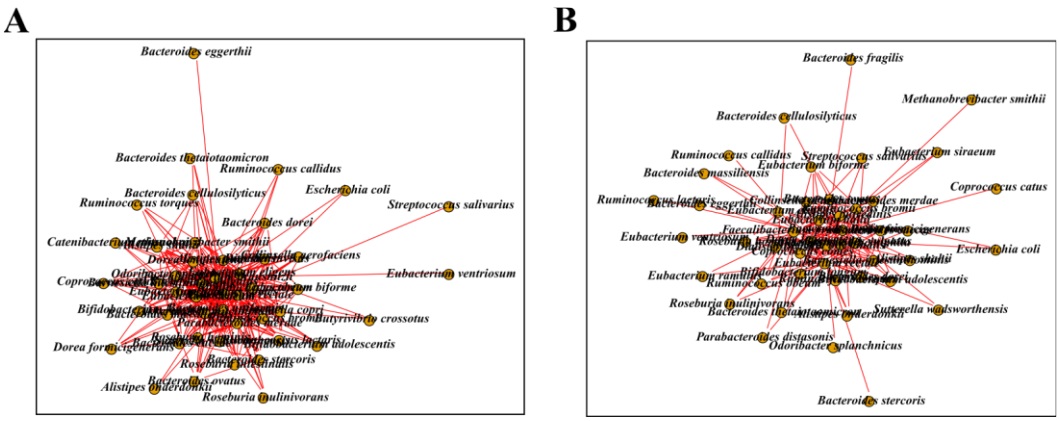


Figure 4: Inter-bacterial metabolic competition networks of A.) Controls (Nodes= 47, Edges= 331), and B.) IBS (Nodes= 47, Edges= 181) showing similar number of taxa but reduced number of connections in IBS. The differences were not significant for node distribution (Adjusted p-value = 0.202), but number of connections differed significantly (Adjusted p-value < 0.001) based on Fisher's exact test.

		Co-Operation Network	Competition Network
Input Taxa (N)	**Controls (N=59)**	91	90
Expected Edges (N)		4095	4005
Edges (N)		165	331
Nodes (N)		30	47
Input Taxa (N)	**IBS (N=80)**	110	110
Expected Edges (N)		5995	5995
Edges (N)		12	181
Nodes (N)		11	47
Odds Ratio	**Nodes**	0.228	0.684
	Edges	0.048	0.346
Adjusted p-value	**Nodes**	< 0.001	0.202
	Edges	< 0.001	< 0.001

Table 4: Summary statistics of average interaction networks of IBS and Controls.

From sample wise analysis of network properties, I observed significant differences in network edge distribution, and network density. There was also a trend (P-value < 0.1) for reciprocity between IBS and Controls, and the measure were reduced in IBS. The number of nodes between IBS and Controls was consistent (supplementary figure 2), suggesting that these results are independent of the alpha diversity. No significant differences in network properties of competition networks were observed between IBS and Controls (supplementary figure 3). This further suggests a greater importance of co-operation compared to competition in IBS microbiome structure.

6.4.4 The co-operation network is predictive of IBS

I further investigated the predictive power of the co-operation network. Based on cumulative edge weights of sub-graphs of nearest neighbours (see methods) of each node, I performed predictive analysis which indicated that the co-operation network is predictive of IBS. The mean area under the curve (AUC) of the receiver operating

characteristics (ROC) curve based on 10-fold cross validation was 0.747 (SD = 0.096), with sensitivity (accuracy of predicting IBS) of 0.838 (SD = 0.132) and specificity (accuracy of predicting Controls) of 0.39 (SD = 0.157). Optimized classification thresholds were identified, which were maximum of the sum of sensitivity and specificity (best), a specificity of 0.8 (SP80), and a specificity of 0.9 (SP90). The corresponding sensitivities and specificities of the optimized thresholds are listed in table 5.

	Threshold	**Sensitivity**	**Specificity**
Best (Max Sensitivity + Specificity)	0.575	0.738	0.678
SP80 (Specificity of 0.8)	0.663	0.55	0.814
SP90 (Specificity of 0.9)	0.771	0.275	0.915

Table 5: Optimized thresholds for classification using co-operation network indices.

The feature coefficients and random forest importance of the species in the final optimized model are listed in Table 6. I observed that a lower number of species was associated with IBS, and the random forest importance of the species was lower than that of Controls. *Dialister invisus, Parabacteroides merdae, Ruminococcus torques, Dorea formicigenerans, Bifidobacterium adolescentis*, were some of the species with high importance that were associated with IBS, while among species associated with controls, *Barnesiella intestinihominis, Coprococcus comes, Bacteroides vulgatus, Ruminococcus obeum, Odoribacter splanchnicus, Lachnospira eligens*, had higher random forest importance score.

Species	Lasso Coefficient	Feature importance
Associated with IBS		
Dialister invisus	-2.594	1.936
Parabacteroides merdae	-1.89	1.418
Ruminococcus torques	-2.089	1.391
Dorea formicigenerans	-8.591	1.356
Bifidobacterium adolescentis	-0.044	1.311
Bifidobacterium longum	-2.748	1.28
Alistipes onderdonkii	-1.891	1.196
Streptococcus salivarius	-2.057	0.919
Eubacterium ramulus	-19.908	0.804
Bifidobacterium bifidum	-14.434	0.766
Phascolarctobacterium succinatutens	-10.572	0.735
Associated with Controls		
Barnesiella intestinihominis	21.895	3.204
Coprococcus comes	13.936	2.384
Bacteroides vulgatus	7.282	2.259
Ruminococcus obeum	13.737	2.246
Alistipes shahii	2.506	1.929
Odoribacter splanchnicus	11.911	1.922
Lachnospira eligens	2.345	1.803
Eubacterium rectale	1.599	1.78
Bacteroides uniformis	2.752	1.716
Eubacterium biforme	11.917	1.656
Ruminococcus lactaris	10.357	1.424
Bacteroides cellulosilyticus	2.907	1.37
Roseburia hominis	4.47	1.264
Coprococcus catus	17.175	1.221
Akkermansia muciniphila	1.107	1.208
Ruminococcus callidus	5.549	1.207
Prevotella copri	9.11	1.064
Sutterella wadsworthensis	1.693	0.891

Table 6: Important features in the final predictive model. The table lists species associated with IBS or Controls along with their corresponding LASSO coefficients and Random Forest feature importance scores.

6.4.5 Cluster analysis of microbiome structure

Clustering analysis of node properties of taxa and their node properties in co-operation and competition networks identified 4 modules which upon further investigation suggested functional importance of the taxa involved in metabolic interactions. Among the 4 modules, Module 1 had the highest cluster size followed by modules 4, 2, and 3 (supplementary figure 4). The modules and their members are listed in supplementary table 4.

Module 1 was identified as the core microbiome as it consisted of taxa present in majority of the subjects. The core cumulative index was significantly reduced in IBS (Figure 5 A). Module 3 (the depleted core) comprised of a subset of the features present in Module 1 and was observed to be not associated with IBS (Figure 5 C). Module 2 was not significantly different between IBS and Controls (Figure 5 B), and contained node properties of species, majority of which had the potential to metabolize mucins (supplementary table 5). Module 4, cumulative index of which was significantly reduced in IBS (Figure 5 D), appeared to be dominated by dietary fibre associated species like *Bacteroides ovatus, Bacteroides uniformis, Odoribacter splanchnicus, Ruminococcus bromii, Ruminococcus lactaris*.

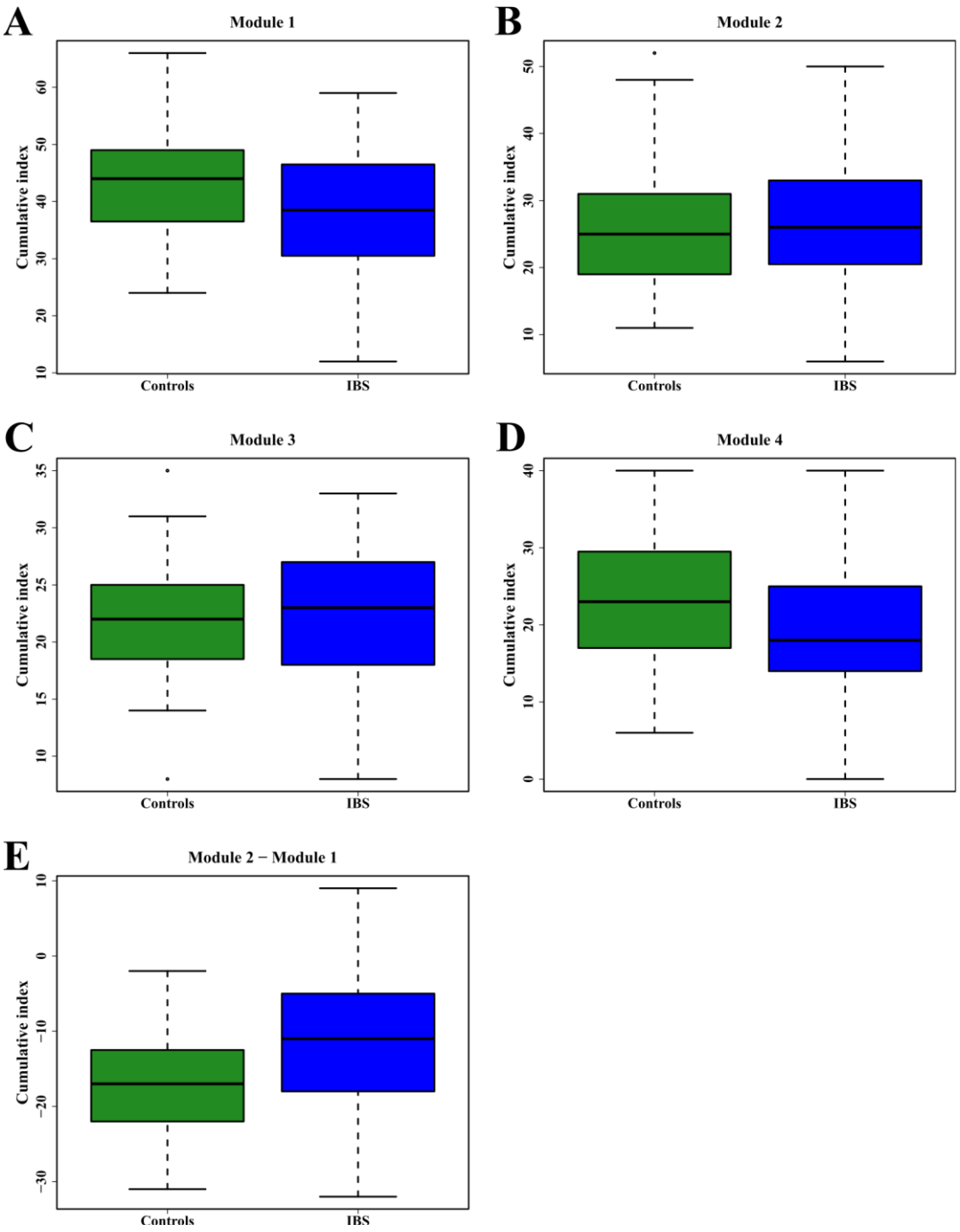


Figure 5: Differences in cumulative indices of modules (identified using iBBiG clustering methodology) between IBS (N=80) and Controls (N=59). The statistical significance of the differences was checked using Wilcoxon rank sum tests, followed by p-value adjustment using the Benjamini-Hochberg method. An adjusted p-value < 0.05 was considered as significant. A.) Differences in cumulative indices of Module 1 (core) was significant (adjusted p-value = 0.021) and was lower in IBS. B.)

Differences in cumulative indices of Module 2 (mucin degraders) was not significantly different between IBS and Controls (adjusted p-value = 0.764). C.) Differences in cumulative indices of Module 3 (reduced core) was not significantly different between IBS and Controls (adjusted p-value = 0.323). D.) Cumulative indices of Module 4 (fibre associated) was significantly less in IBS (adjusted p-value = 0.02). E.) Combined metric by negating effect of module 1 (core) from module 2 (mucin degraders) was significantly higher in IBS (adjusted p-value = 0.001).

As I observed that mucin degradation and assimilation of its catabolites are involved in the models, I formed the hypothesis that it may act within the context of structures in the microbiome. Furthermore, I wanted to investigate the interaction between the presence of mucin degraders and depletion of core in IBS. For this I calculated a metric by subtracting the cumulative index of Module 1 from that of Module 2 (Module 2 minus Module 1). The resulting score was observed to be significantly greater in IBS compared to controls showing that individuals with a greater number of Module 2 taxa and a reduced number of M1 core taxa have a greater incidence of IBS (Figure 5 E). A similar analysis looking at the fibre associated (Module 4) and Core (Module 1) did not show any relationship with IBS.

I further investigated the interaction between this score representing net effect of Module 2 (dominated by node properties of mucin degraders) and Module 4 (fibre related module). For this, I performed principal component analysis (PCA) using the cumulative index of Module 4, and net effective score of mucin degraders (Module 2 minus Module 1). The PCA biplot (Supplementary figure 5) showed significant separation of IBS samples from Controls, and the two variables were observed to occur in opposite direction along the first principal axis. Although expected, it was interesting to observe that the fibre associated module was the major driver of the

control samples, while the mucin degraders were associated with IBS. I also checked the predictive performance of the two variables using support vector machines (SVM) and Leave-one-out cross validation (LOO-CV) approach. The area under the curve (AUC) of the receiver operating characteristics (ROC) curve, based LOO-CV, was 0.682 with sensitivity (accuracy of predicting IBS) of 0.675 and specificity (accuracy of predicting Controls) of 0.559. Optimized classification thresholds were identified, which were maximum of the sum of sensitivity and specificity (best), a specificity of 0.8 (SP80), and a specificity of 0.9 (SP90). The corresponding sensitivities and specificities of the optimized thresholds are listed in table 7.

	Threshold	**Sensitivity**	**Specificity**
Best (Max Sensitivity + Specificity)	0.573	0.6125	0.746
SP80 (Specificity of 0.8)	0.632	0.488	0.814
SP90 (Specificity of 0.9)	0.71	0.325	0.915

Table 7: Optimized thresholds for classification using Fibre and mucin degraders associated variables.

6.4.6 Highly interactive species strongly modulate the metabolic output of the microbiome

I investigated the effect of microbiome perturbation on the metabolic output of the bacterial community by simulating the increase in the abundance of each species. I analysed mean stimulation or inhibition of simulated metabolite exchanges (import/export) in IBS and Control populations (Figure 6). I thus observed significant differences between modulation potential of species with different extent of co-operation.

For metabolite exports by the bacterial community, the number of metabolites predicted to be stimulated by each species showed significant differences (Kruskal Wallis test; p-value < 0.01) based on number of connections the species had in the co-operation network (Figure 6A). Within IBS or Controls, I observed that the more interactive species could potentially influence a greater number of metabolites compared to those with lower number of interactions (Supplementary Table 6), indicating the importance of keystone species in metabolic output modulation of a bacterial community. In case of metabolites, whose exports were inhibited upon increase in abundance of species (Figure 6B), the number of affected metabolites by species with only one connection was observed to be significantly lower than those with greater number of interactions (Supplementary Table 6), in both IBS and Control groups. However, I observed similar patterns of simulated impact of bacterial community perturbation on metabolic output, by members of the gut bacterial community of IBS and Control groups, at each interaction level (Supplementary Table 6).

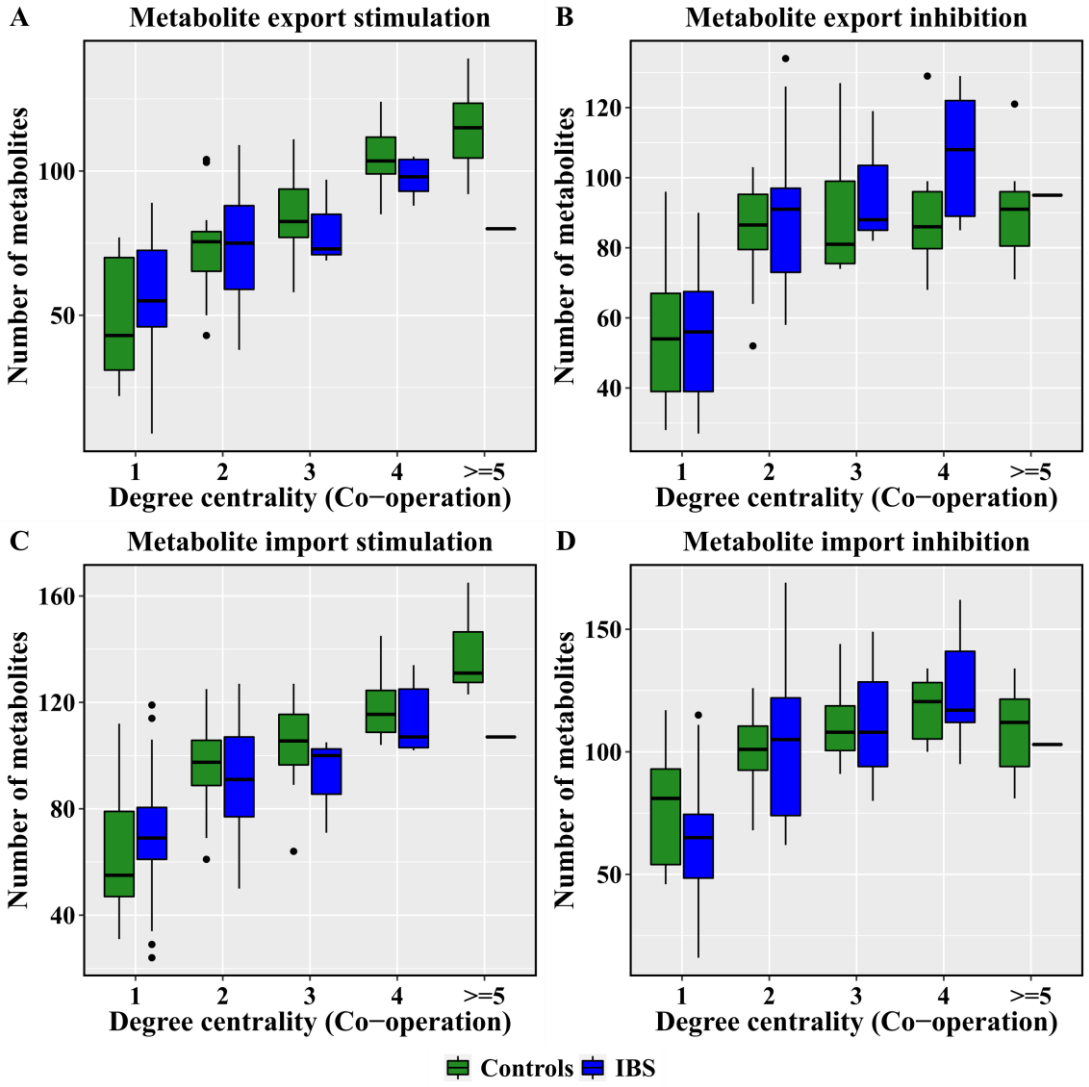


Figure 6: Comparison of modulation potential of species based on number of metabolites influenced by them and their extent of co-operation with other community members. Species were grouped based on their degree centrality in the co-operation network [1 (IBS: 27, Controls: 21), 2 (IBS: 21, Controls: 18), 3 (IBS: 3, Controls: 8), 4 (IBS: 5, Controls: 6), >=5 (IBS: 1, Controls: 7)]. Overall, all the comparisons were statistically significant (Kruskal-Wallis test, p-value < 0.001), and details of post-hoc Dunn's tests are listed in supplementary table 6. A.) Differences in number of exported metabolites stimulated by species having different extent of interactions in co-operation network. B.) Differences in number of exported metabolites inhibited by species having different extent of interactions in co-operation network. C.) Differences

in number of imported metabolites stimulated by species having different extent of interactions in co-operation network. D.) Differences in number of imported metabolites inhibited by species having different extent of interactions in co-operation network.

The predicted impact of bacterial community perturbation, by increasing the abundance of species, was also seen on consumption of metabolites by the community. For the metabolites with predicted increase in consumption upon bacterial community perturbation (Figure 6C), the species with only one connection significantly impacted a fewer number of metabolites compared to species with higher number of connections (Supplementary Table 6). Similarly, in case of predicted inhibition of metabolite consumption upon bacterial community perturbation (Figure 6D), species with higher degree of interaction affected a greater number of metabolites compared to species showing reduced co-operation with the community (Supplementary Table 6).

To predict species with similar modulation potential, I visualized average elasticity profiles of species in IBS and Controls, for metabolites imported and exported by the bacterial community. Analysis of the simulated export profile in controls (supplementary figure 6A), I observed two clusters of species where cluster 2 (blue) showed a greater potential to modulate (strong modulators) the export of metabolites by the community compared to cluster 1 (red; weak modulators). The species and their cluster memberships are listed in supplementary table 7. Some notable species in the potent clusters were *Barnesiella intestinihominis*, *Odoribacter splanchnicus*, *Faecalibacterium prausnitzii*, and *Lachnospira eligens*. I also observed clusters of metabolites which were strongly stimulated (cluster 1, red) or inhibited (cluster 2, blue). Among the affected metabolites, the cluster of speculated strong modulators potentially increased the export of SCFAs (acetate, propionate, and butyrate) along

with indole (supplementary table 8), which are known to be beneficial for maintenance of gut health [20,21]. However, an increase in the simulated export of metabolites detrimental for health like Trimethylamine (TMA), and Phenol was observed for some of the species. In case of the metabolites whose rate of exports were potentially reduced by the predicted strong modulators, I observed possible inhibition of molecules known to be detrimental for gut health, like multiple catabolites of mucin degradation (suggesting reduction in mucin degradation), along with hydrogen sulphide (supplementary table 8).

I further noticed that the network properties (outward degree, inward degree, and betweenness in co-operation and competition networks, supplementary figure 6B) of the strong modulators (cluster 2) were significantly greater than that of weaker modulators (cluster 1). This observation suggested that the highly interactive species in a community, as identified through simulations, may be impactful on metabolic output modulation.

In case of import of metabolites by the bacterial community in controls (supplementary figure 7A), three distinct clusters of species with different modulation potential were observed, wherein, members of cluster 1 (red) and 3 (green) showed stronger influence on the predicted imports compared to cluster 2 (blue), which are also summarized in supplementary table 9. The constituents of the clusters were similar to that of simulated exports. The metabolites imported by the community did not group into distinct clusters based on inhibition or stimulation. Out of the four identified cluster, metabolites in cluster 1 (red), and cluster 2 (blue) were mainly predicted to be stimulated, and inhibited, respectively. Members of cluster 1 were mainly bio-transformed bile acids, whose imports were increased by species with a higher potential to modulate, while cluster 2 (primarily inhibited imports) contained

metabolites related to mucin metabolism (indicating reduction of mucin degradation), 4-Aminobutanoate (GABA), and various amino acids and di-peptides. Cluster 3 (green) and 4 (brown) were mixed clusters with metabolite imports being reduced or increased grouped together. Imported metabolites and their cluster memberships are summarized in supplementary table 10. Network properties of the species clusters with high capability to modulate the metabolite imports also showed higher levels of interactions within the community (supplementary figure 7B). Network properties were highest for species in cluster 3, followed by those in cluster 1, with least values for species of cluster 2.

In IBS datasets, I obtained trends similar to perturbation analysis in the control population. For modulation profiles of species and their potentially exported metabolites, I observed two separate clusters of species (supplementary figure 8A), and species in cluster 2 (blue) were likely to have a stronger influence on changes in metabolite exports compared to cluster 1 (red). The composition of the clusters was similar to that observed in analysis of controls. Species and their cluster memberships are listed in supplementary table 11. The exported metabolites did not show efficient clustering. Out of the six identified clusters of simulated metabolites, members of Cluster 2 (blue), and Cluster 6 (orange) were mainly stimulated, while metabolites in Cluster 4 (brown) were predominantly inhibited. The other clusters, Cluster 1 (red), Cluster 3 (green), Cluster 5 (magenta), and were mixed clusters with no specific grouping based on increase or reduction of export rates upon increasing species abundance. Summary of exported metabolites and their cluster memberships are listed in supplementary table 12. Among the metabolites present in cluster 2 and 6 (mainly stimulated), I observed an increase in predicted export of Isovalerate (a branched SCFA), unlike in controls where an increase in the export of SCFAs was observed.

Other metabolites belonging to these clusters were metabolites related to mucin degradation, and related released antigens, which is opposite to what was observed in controls.

The differences in network properties between species clusters also showed similar trends as observed in controls. The strong modulators had significantly higher values of network properties compared to the likely weak modulators (supplementary figure 8B).

By simulating the bacterial community in the samples, I further investigated the predicted capabilities of species in IBS to modulate the metabolites imported by the IBS community (supplementary figure 9A). I obtained two broadly distinct clusters of species, out of which cluster 2 (blue) contained species with likely stronger modulation capabilities, while cluster 1 (red) contained a variety of species with possible weaker potential to change the rate of metabolite imports. Cluster 1 contained multiple species which were part of strong modulators in earlier comparisons, but the members of cluster 2 (strong modulators) were similar to previous observations. Supplementary table 13 lists the species and their cluster affiliations. The trends in network properties were similar to previous analysis, wherein, all the network properties in both co-operation and competition networks, were higher for species with greater influence on metabolites (cluster 2) compared to the weaker cluster 1 (supplementary figure 9B).

Regarding metabolites, I obtained 4 clusters, with clusters 2 (blue) and 4 (brown) containing metabolites which were mostly predicted to be inhibited by members of strongly influencing species. Cluster 3 (green) contained metabolites whose imports were mainly stimulated by the potent modulators. Cluster 1 (red) was observed to be mixed cluster and contained metabolites which were less influenced by the gut

bacterial community perturbations. Supplementary table 14 contains the metabolites and their cluster information. It was interesting to note that various predicted bile acids exchanges dominated cluster 2 (blue), whose rate of imports were mainly reduced by the strong modulators. This suggests a likely lower rate of absorption of BAs upon species perturbation, which may be detrimental to the gut. Cluster 4 (brown) which also contained metabolites with inhibited imports, comprised of additional bile acids, along with metabolites of mucin degradation, and released T antigens. Cluster 3 (green), containing metabolites with increased rate of import, included other mucin catabolites, and released antigens.

6.5 Discussion

I report a computational analysis of community level metabolic models of IBS and Controls which attempts to investigate and understand the mechanistic differences in bacterial metabolism, focusing on metabolites which can be potentially exchanged with the extracellular environment. Furthermore, I also explored inter-bacterial interactions, and differences in microbiome structure based on metabolic exchanges and competition. I also performed predictive analysis based on bacterial network properties to identify potential markers with capability to predict IBS. Moreover, through this study, I simulated the effect of increase in abundance of the species on the overall production and consumption profile of the bacterial community, to identify potent species which can modulate the gut bacterial community, and thereby, its metabolic output.

I observed global differences in intracellular reactions, exports, and imports profiles which could significantly distinguish between IBS and Controls but were incapable to significantly differentiate the clinical subtypes of IBS. These observations corroborate our previous research and further indicate inclusion of microbiome status along with

the less efficient conventional classification of IBS patients into subtypes based on differences in bowel habits [22], for better diagnosis and management of IBS.

In terms of significantly different intracellular reactions and metabolite exchanges, and their contributing species, a higher number of species, and reactions predicted to be related to mucin degradation and antigen release in IBS, indicate an increase in detrimental effect on the gut epithelium, and subsequently on gut barrier function, along with increased chances of pro-inflammatory events due to antigen release. It is interesting to note what while all the community level metabolic models were simulated in the same media (average western diet), the community members in IBS preferred more mucin degradation and related reactions, compared to controls. A previous study reported that a dietary fibre-deprived gut microbiota tends to degrade the colonic mucous barrier by degradation of mucin and utilizing the catabolites as energy source [23]. The current observations corroborate this finding, but also suggest that mucin degradation is not only driven by lack of dietary fibres, but also depends on the community composition which may enhance the effect.

Although predictive in nature, the increased rate of hypoxanthine (a source of energy for gut epithelial cells which promotes intestinal barrier development [24,25]) consumption by *Ruminococcus torques* in IBS, is in accordance with previous report linking IBS pathogenesis to lowering of availability of hypoxanthine in the gut [4]. The predicted lower rate of import of hydrogen sulphide (H_2S), a gaseous by-product of microbial metabolism associated with inflammation [26], is speculative of higher availability in the gut which may lead to increased inflammation in IBS, and is concordant with previous reports [27,28]. Furthermore, a predicted reduced rate of 4-aminobutanoate (GABA) export by *Odoribacter splanchnicus*, indicates altered modulation of gastrointestinal motility and inflammation in IBS, which has been

reported previously [29]. I also observed higher levels of potential exports as well as imports of primary and secondary bile acids in IBS, which is also suggestive of higher rates of production of such metabolites, may lead to colonic irritation and diarrhoea [30,31]. Additionally, estimated increase of simulated imports of amino acids and sugars by different species in IBS was observed which may lead to reduced availability of nutrients for the host and other members of the microbiome community. Additionally, the observed increase in the predicted rate of potential export of glycerol in IBS, may lead to increased visceral hypersensitivity and abdominal pain [32].

Based on metabolic interactions, I also observed significant alterations in the microbiome structure of IBS compared to controls, especially in the case of inter-microbial co-operation. Moreover, I observed significant reduction in number of connections, reciprocity (reflective of interdependencies), and network density in co-operation networks of IBS samples. Our previous analyses also reported alterations in microbiome structure, and reduction of co-occurrence network density in IBS [5]. Interestingly, the co-occurrence networks previously reported were based on correlations, and our current corroborating results, based on metabolic modelling adds biological context to the networks, and further strengthens the hypothesis that inter bacterial co-occurrence is driven by inter-microbial metabolic co-dependencies, as suggested by another study [33]. The co-operation network was also found to be predictive of IBS. Predictive analysis using microbiome abundance may perform better, but there is a possibility of existence of greater contribution of stochasticity in the models. However, predictive models based on our current approach using co-operation network based on potential metabolic interactions may provide a biological context to the models, which may render the model to be more robust and biologically relevant.

Additional insights obtained using clustering analysis of the node properties relate to the biology of microbiome in IBS. The significant underrepresentation of the core module in IBS reflected the alterations in microbiome diversity in IBS. The net effect of the module associated with mucin degraders, after negating the effect of the core module, was significantly higher in IBS. Conversely, the module associated with fibre associated species was significantly different, but the cumulative index was lower in IBS. Interestingly, combining the net mucin degrader effect and fibre associated module, showed significantly higher scores in IBS, and the combined metric was also observed to be predictive of IBS.

The gut bacterial community perturbation analysis, by increasing the abundance of each taxon in the community, also provided biological insights, hitherto unreported in IBS. Species having stronger influence on metabolite exchange rates were also found to be actively interactive in co-operation and competition networks. Similar trends were observed for both IBS, and Control populations. Such information is useful in screening potent bacterial species which can potentially perturb the metabolic output of the community. However, from analysis of metabolic output modulation, I noted that beneficial bacterial species as well as pathobionts can have similar and high modulation capabilities. Therefore, extreme care should be taken during selection of species/strain for clinical intervention studies.

The current findings based on computational analysis have potential clinical implications, but also require further experimental validation. I have demonstrated the utility of community level metabolic modelling in investigating differences in IBS. Findings from my differential analysis of the predicted flux profiles bear biological relevance and are also in agreement with findings reported in my previous metabolome analysis as well as in other studies. The predictive analyses in this study propose novel

approaches to utilize microbial interaction networks for diagnosing and stratifying IBS samples. The bacteriome perturbation analysis applied to IBS and Control populations, may be helpful in identification of live biotherapeutic candidates which can show desired level of interaction and metabolic output modulation.

6.6 References

1. Jeffery IB, Quigley EMM, Öhman L, et al. The microbiota link to irritable bowel syndrome: an emerging story. Gut Microbes 2012;3:572–576.

2. Jeffery IB, O'Toole PW, Öhman L, et al. An irritable bowel syndrome subtype defined by species-specific alterations in faecal microbiota. Gut 2012;61:997–1006.

3. Tap J, Derrien M, Törnblom H, et al. Identification of an Intestinal Microbiota Signature Associated With Severity of Irritable Bowel Syndrome. Gastroenterology 2017;152:111-123.e8.

4. Mars RAT, Yang Y, Ward T, et al. Longitudinal Multi-omics Reveals Subset-Specific Mechanisms Underlying Irritable Bowel Syndrome. Cell 2020;182:1460-1473.e17.

5. Jeffery IB, Das A, O'Herlihy E, et al. Differences in Fecal Microbiomes and Metabolomes of People With vs Without Irritable Bowel Syndrome and Bile Acid Malabsorption. Gastroenterology 2020;158:1016-1028.e8.

6. Diener C, Gibbons SM, Resendis-Antonio O. MICOM: Metagenome-Scale Modeling To Infer Metabolic Interactions in the Gut Microbiota. mSystems 2020;5:e00606-19.

7. Ebrahim A, Lerman JA, Palsson BO, et al. COBRApy: COnstraints-Based Reconstruction and Analysis for Python. BMC Syst Biol 2013;7:74.

8. Schellenberger J, Que R, Fleming RMT, et al. Quantitative prediction of cellular metabolism with constraint-based models: the COBRA Toolbox v2.0. Nat Protoc 2011;6:1290–1307.

9. Ng RH, Lee JW, Baloni P, et al. Constraint-Based Reconstruction and Analyses of Metabolic Models: Open-Source Python Tools and Applications to Cancer. Front Oncol 2022;12:914594.

10. Das A, O'Herlihy E, Shanahan F, et al. The fecal mycobiome in patients with Irritable Bowel Syndrome. Sci Rep 2021;11:124.

11. Coughlan S, Das A, O'Herlihy E, et al. The gut virome in Irritable Bowel Syndrome differs from that of controls. Gut Microbes 2021;13:1–15.

12. Human Microbiome Project Consortium. Structure, function and diversity of the healthy human microbiome. Nature 2012;486:207–214.

13. Franzosa EA, McIver LJ, Rahnavard G, et al. Species-level functional profiling of metagenomes and metatranscriptomes. Nat Methods 2018;15:962–968.

14. Truong DT, Franzosa EA, Tickle TL, et al. MetaPhlAn2 for enhanced metagenomic taxonomic profiling. Nat Methods 2015;12:902–903.

15. Magnúsdóttir S, Heinken A, Kutt L, et al. Generation of genome-scale metabolic reconstructions for 773 members of the human gut microbiota. Nat Biotechnol 2017;35:81–89.

16. Diener C. Artifacts for q2-micom. 2020. Available at: https://zenodo.org/record/3755182 [Accessed December 30, 2022].

17. Noronha A, Modamio J, Jarosz Y, et al. The Virtual Metabolic Human database: integrating human and gut microbiome metabolism with nutrition and disease. Nucleic Acids Res 2019;47:D614–D624.

18. Stellato B, Banjac G, Goulart P, et al. OSQP: an operator splitting solver for quadratic programs. Math Prog Comp 2020;12:637–672.

19. Gusenleitner D, Howe EA, Bentink S, et al. iBBiG: iterative binary bi-clustering of gene sets. Bioinformatics 2012;28:2484–2492.

20. Ye X, Li H, Anjum K, et al. Dual Role of Indoles Derived From Intestinal Microbiota on Human Health. Front Immunol 2022;13:903526.

21. Silva YP, Bernardi A, Frozza RL. The Role of Short-Chain Fatty Acids From Gut Microbiota in Gut-Brain Communication. Front Endocrinol 2020;11:25.

22. Drossman DA, Hasler WL. Rome IV-Functional GI Disorders: Disorders of Gut-Brain Interaction. Gastroenterology 2016;150:1257–1261.

23. Desai MS, Seekatz AM, Koropatkin NM, et al. A Dietary Fiber-Deprived Gut Microbiota Degrades the Colonic Mucus Barrier and Enhances Pathogen Susceptibility. Cell 2016;167:1339-1353.e21.

24. Lee JS, Wang RX, Alexeev EE, et al. Hypoxanthine is a checkpoint stress metabolite in colonic epithelial energy modulation and barrier function. J Biol Chem 2018;293:6039–6051.

25. Lee JS, Wang RX, Goldberg MS, et al. Microbiota-Sourced Purines Support Wound Healing and Mucous Barrier Function. iScience 2020;23:101226.

26. Guo F-F, Yu T-C, Hong J, et al. Emerging Roles of Hydrogen Sulfide in Inflammatory and Neoplastic Colonic Diseases. Front Physiol 2016;7:156.

27. Singer-Englar T, Rezaie A, Gupta K, et al. 182 - Competitive Hydrogen Gas Utilization by Methane- and Hydrogen Sulfide-Producing Microorganisms and Associated Symptoms: Results of a Novel 4-Gas Breath Test Machine. Gastroenterology 2018;154:S-47.

28. Singer-Englar T, Rezaie A, Gupta K, et al. 1089 - A Novel 4-Gas Device for Breath Testing Shows Exhaled H 2 S is Associated with Diarrhea and Abdominal Pain in a Large Scale Prospective Trial. Gastroenterology 2018;154:S-213.

29. Aggarwal S, Ahuja V, Paul J. Dysregulation of GABAergic Signalling Contributes in the Pathogenesis of Diarrhea-predominant Irritable Bowel Syndrome. J Neurogastroenterol Motil 2018;24:422–430.

30. Yang M, Gu Y, Li L, et al. Bile Acid-Gut Microbiota Axis in Inflammatory Bowel Disease: From Bench to Bedside. Nutrients 2021;13:3143.

31. Riemsma R, Al M, Corro Ramos I, et al. SeHCAT [tauroselcholic (selenium-75) acid] for the investigation of bile acid malabsorption and measurement of bile acid pool loss: a systematic review and cost-effectiveness analysis. Health Technol Assess 2013;17:1–236.

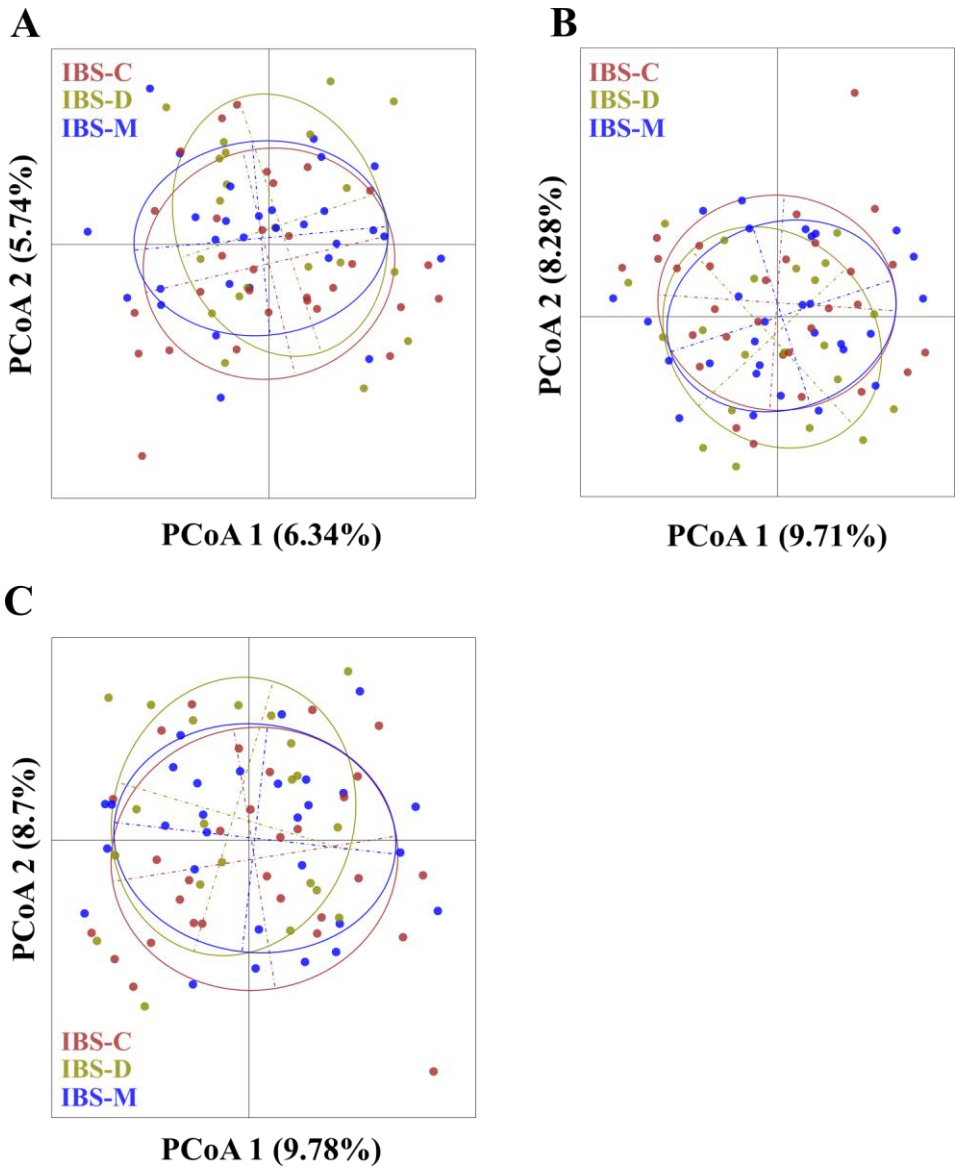
32. Louvel D, Delvaux M, Staumont G, et al. Intracolonic injection of glycerol: a model for abdominal pain in irritable bowel syndrome? Gastroenterology 1996;110:351–361.

33. Zelezniak A, Andrejev S, Ponomarova O, et al. Metabolic dependencies drive species co-occurrence in diverse microbial communities. Proc Natl Acad Sci U S A 2015;112:6449–6454.

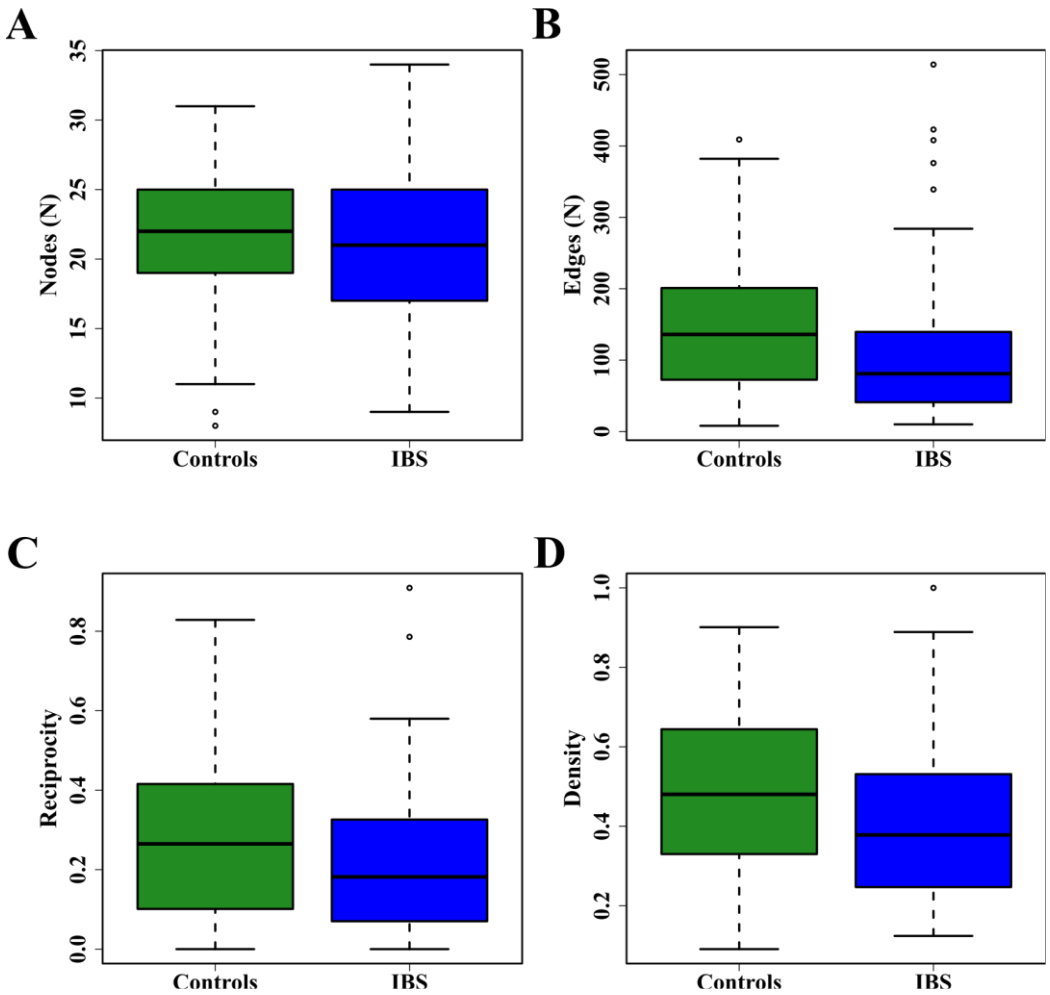
6.7 Supplementary information

6.7.1 Supplementary Figures

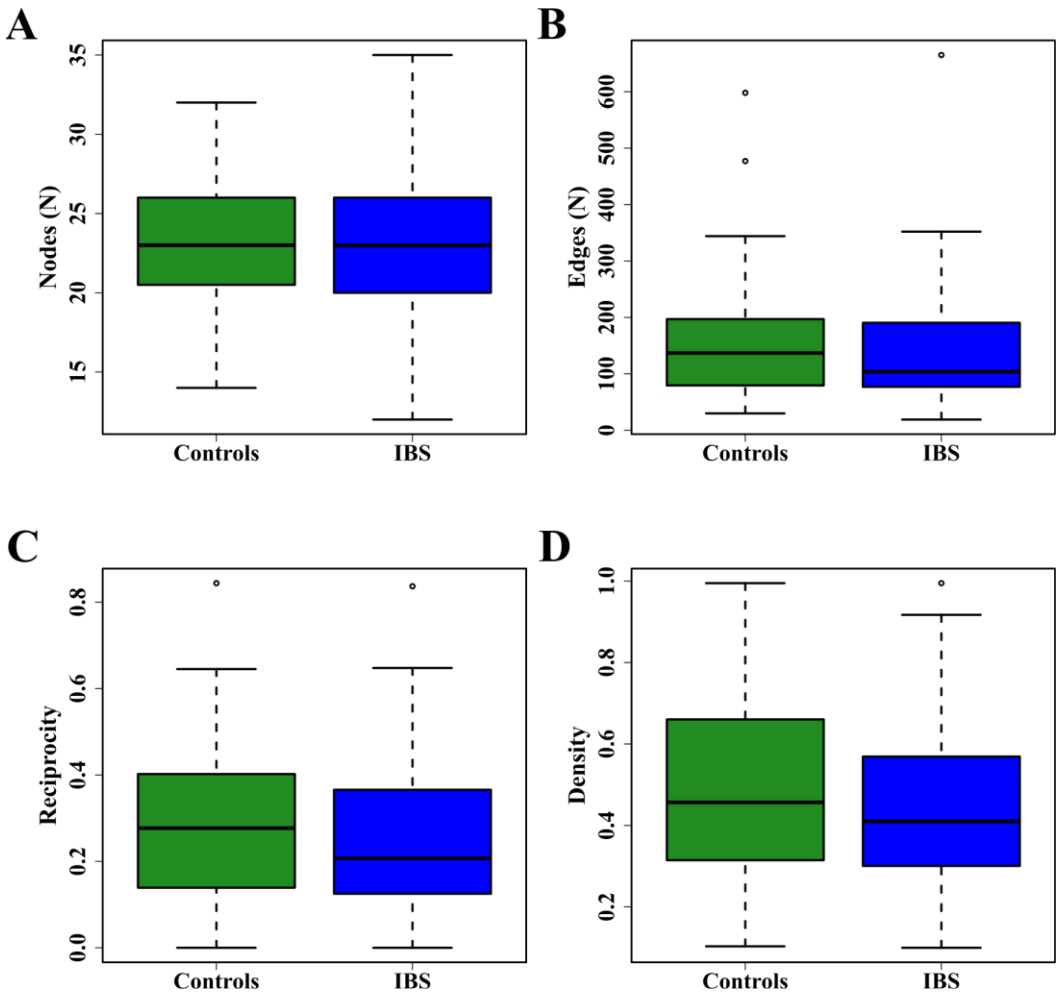
Supplementary Figure 1: PCoA plots based on spearman distance using reaction fluxes (log transformed) showing the spatial arrangement of IBS samples based on clinical subtypes, i.e. IBS-C (N=30), IBS-D (N=21), IBS-M (N=29). None of the separations were significantly different between the clinical subtypes. A.) PCoA plots based on log transformed fluxes of stratified internal reactions (p = 0.411). B.) PCoA plots based on log transformed stratified exports (p = 0.754). C.) PCoA plots based on log transformed stratified imports (p = 0.655).



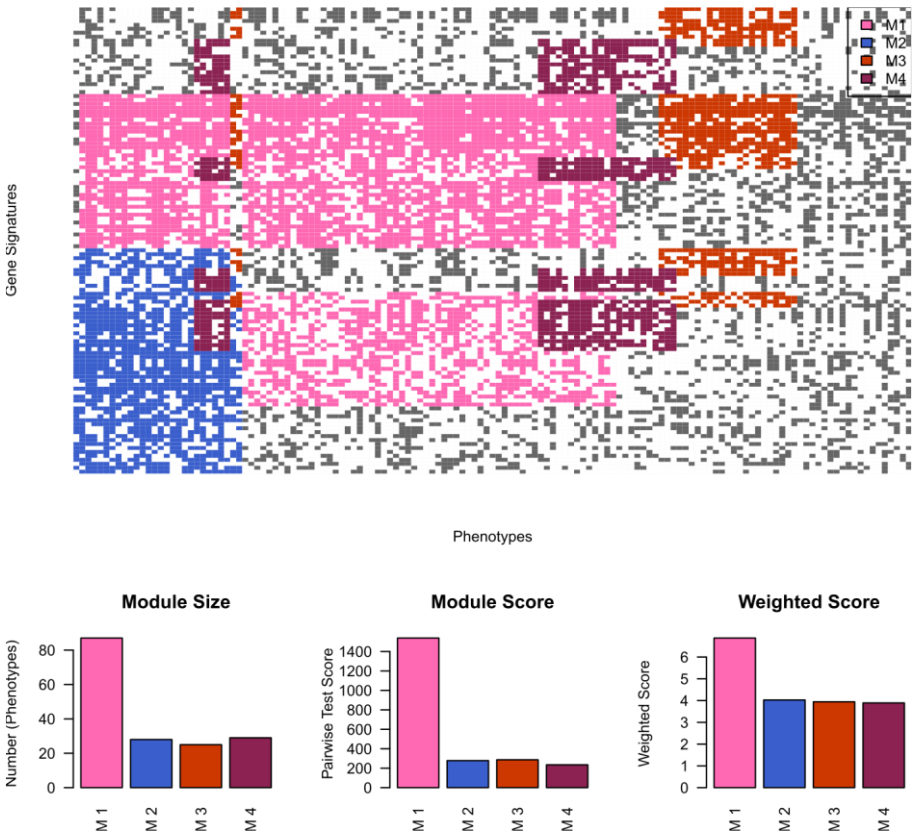
Supplementary Figure 2: Differences in network properties of co-operation networks between IBS (N=80) and Controls (N=59). Statistical significance was tested using Wilcoxon rank sum test and Benjamini-Hochberg adjusted p-value < 0.05 was considered significant. A.) Differences in number of nodes in co-operation networks of IBS and Controls samples (p = 0.283, not significant). B.) Differences in number of edges in co-operation networks of IBS and Controls samples (adjusted p-value = 0.014, significant). C.) Differences in reciprocity between nodes in co-operation networks of IBS and Controls samples (adjusted p-value = 0.056, near significant trend). D.) Differences in network density of co-operation networks of IBS and Controls samples (adjusted p-value = 0.014, significant).



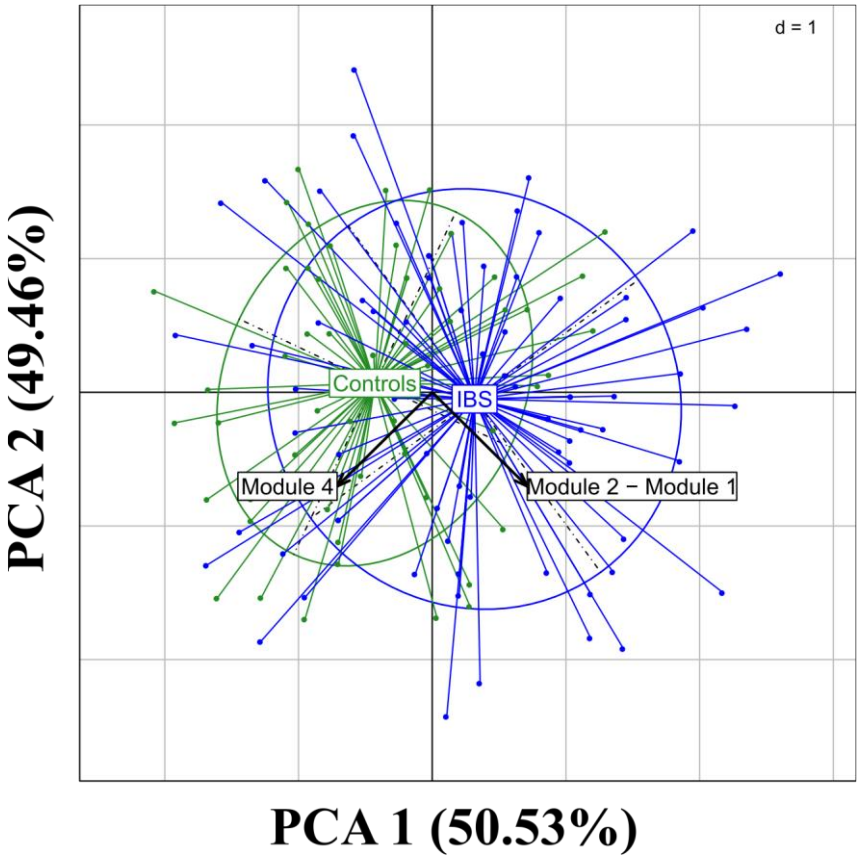
Supplementary Figure 3: Differences in network properties of competition networks between IBS (N=80) and Controls (N=59). Statistical significance was tested using Wilcoxon rank sum test and Benjamini-Hochberg adjusted p-value < 0.05 was considered significant. A.) Differences in number of nodes in competition networks of IBS and Controls samples (adjusted p-value = 0.407, not significant). B.) Differences in number of edges in competition networks of IBS and Controls samples (adjusted p-value = 0.359, not significant). C.) Differences in reciprocity between nodes in competition networks of IBS and Controls samples (adjusted p-value = 0.359, not significant). D.) Differences in network density of competition networks of IBS and Controls samples (adjusted p-value = 0.359, not significant).



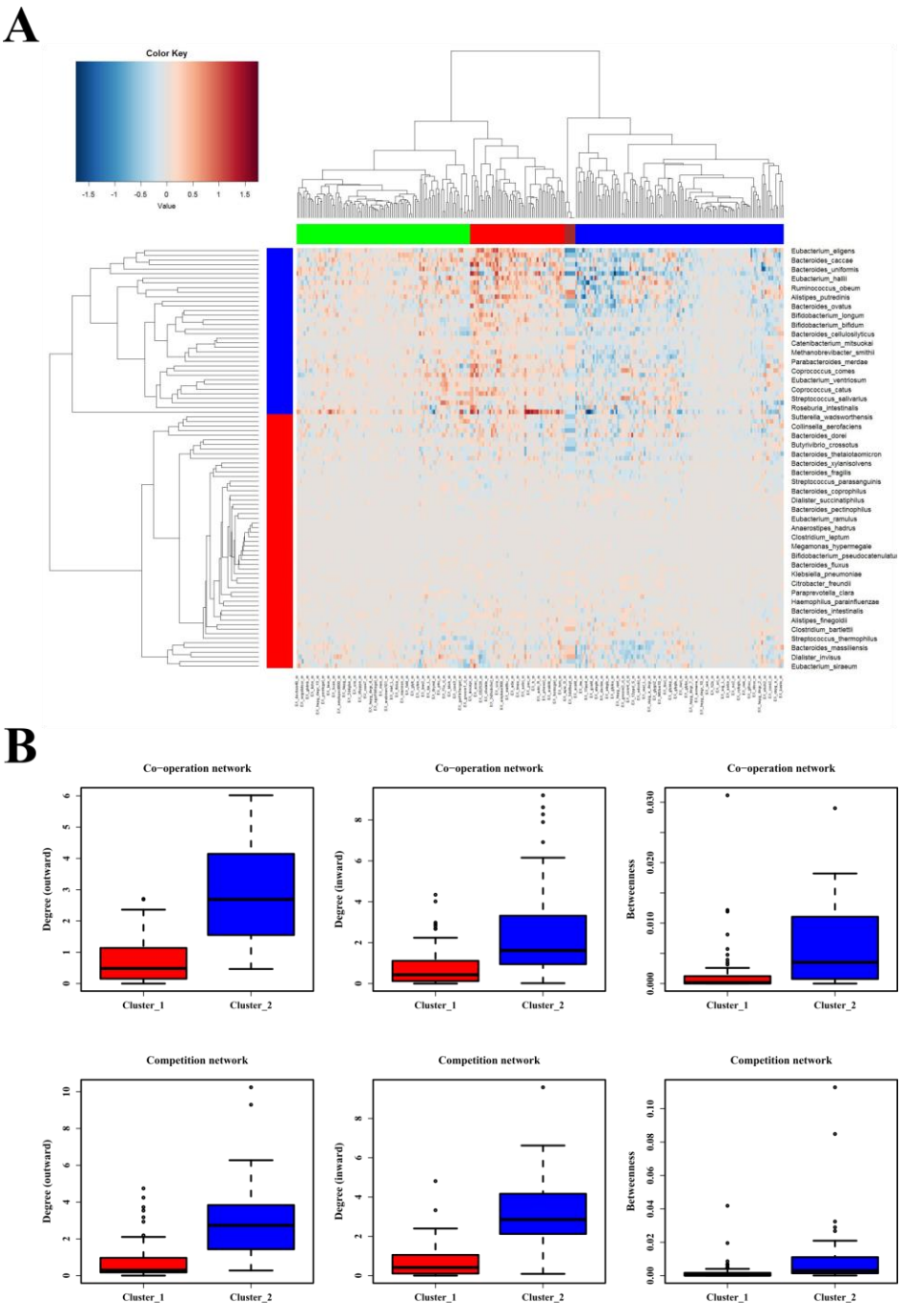
Supplementary Figure 4: Clustering and summary of modules of node properties of co-operation and competition networks identified using iBBiG methodology.



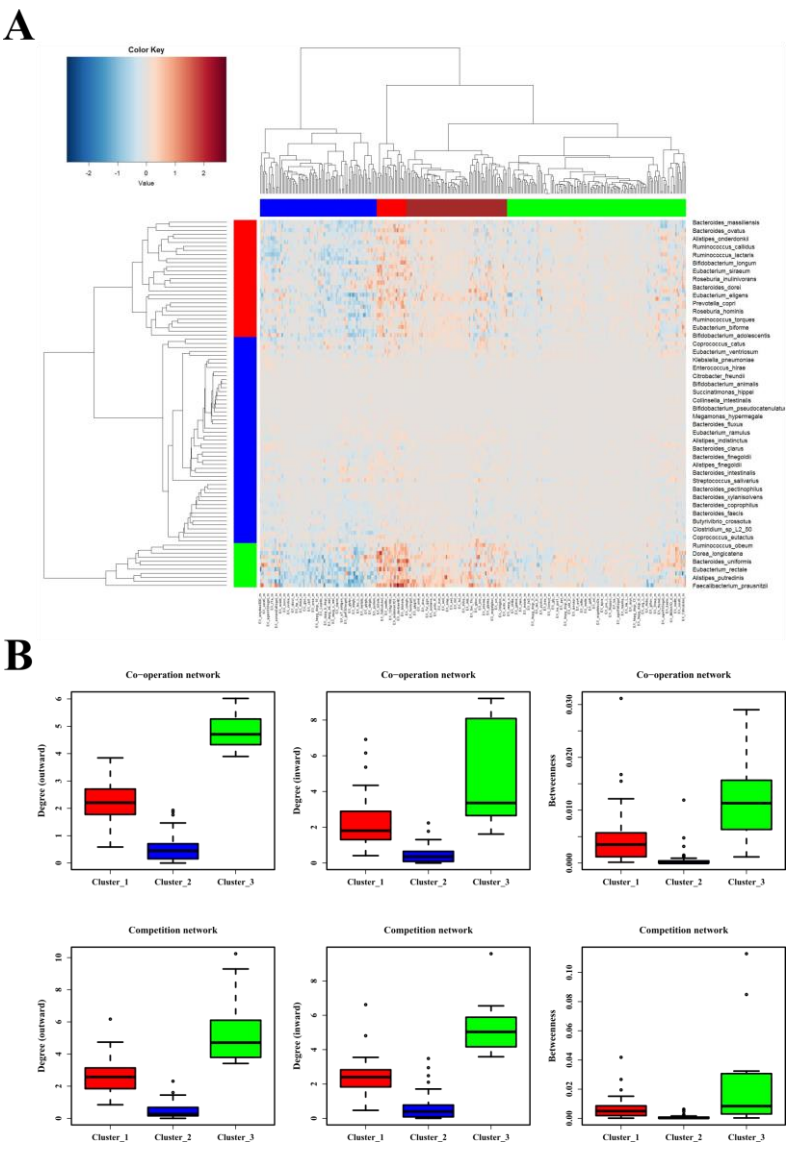
Supplementary Figure 5: Principal component analysis (PCA) of cumulative indices representing fibre associated taxa (Module 4) and net effective index of mucin degrading species (Module 2- Module 1) showing significant spatial separation between IBS and controls (p=0.001; control, n=59; IBS, n=80).



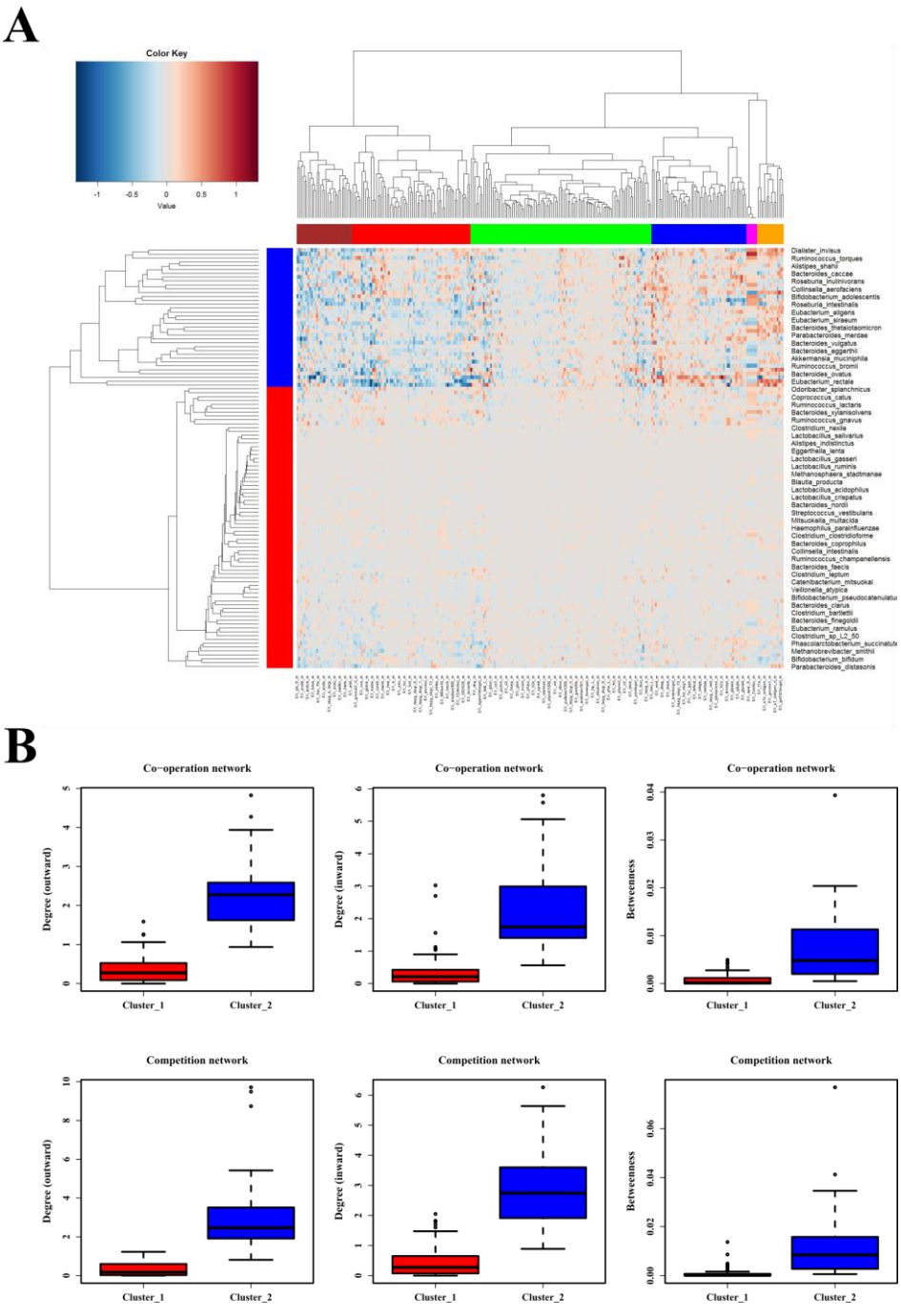
Supplementary Figure 6: A.) Heatmap of elasticities (change in rate of exchange) of metabolite exports upon increasing the abundance of species in Controls. B.) Differences in network properties of species belonging to cluster 1 (red; $N_{species}$ = 55) and cluster 2 (blue; $N_{species}$ = 36). Top panel (L to R) shows differences in outward degree, inward degree, and betweenness of co-operation network, respectively. Bottom panel (L to R) shows differences in outward degree, inward degree, and betweenness of competition network, respectively. All comparisons were statistically significant (Wilcoxon rank sum test; Adjusted p-value < 0.001).



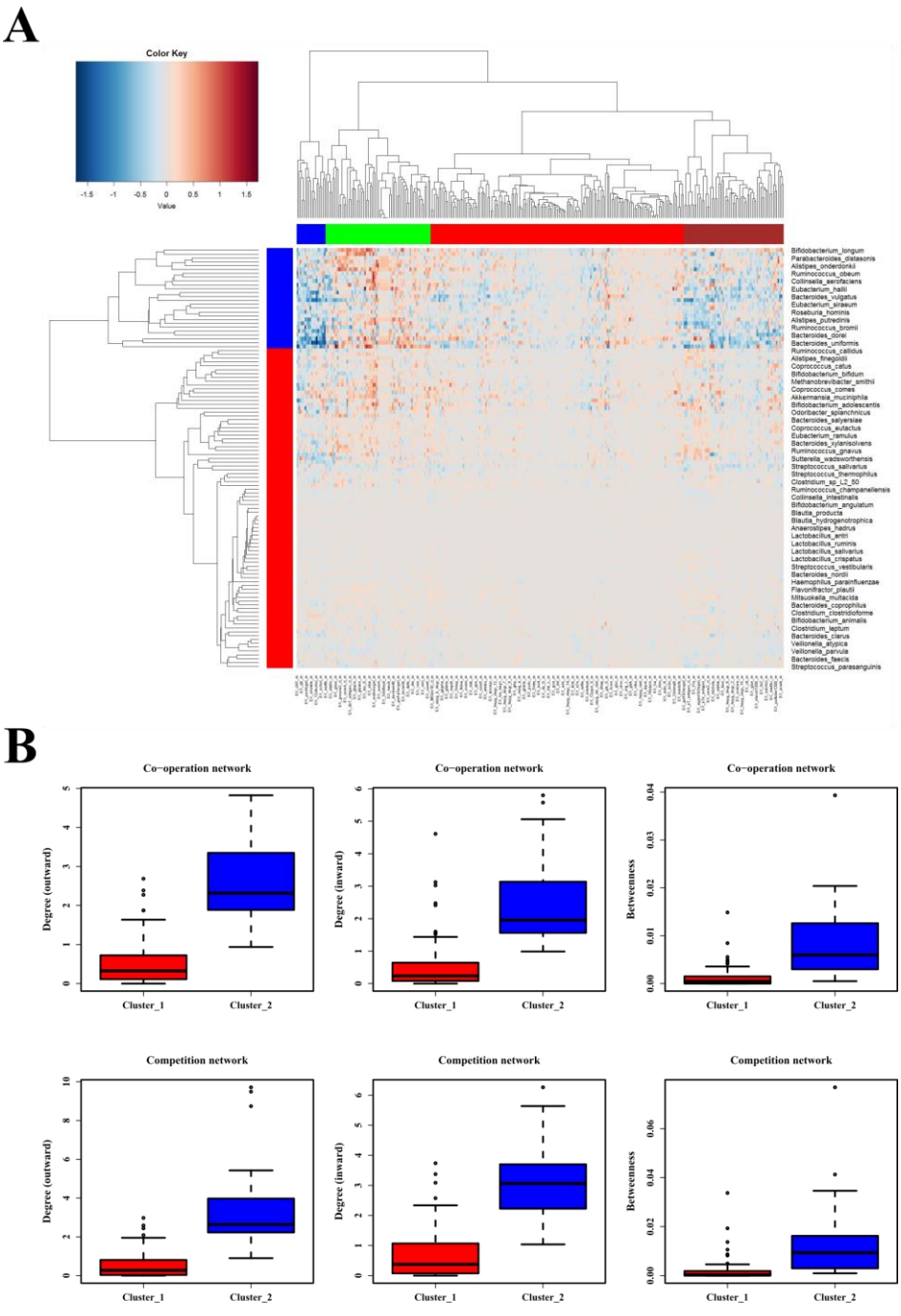
Supplementary Figure 7: A.) Heatmap of elasticities (change in rate of exchange) of metabolite imports upon increasing the abundance of species in Controls. B.) Differences in network properties of species belonging to cluster 1 (red; $N_{species}$ = 29), cluster 2 (blue; $N_{species}$ = 51), and cluster 3 (green; $N_{species}$ = 11). Top panel (L to R) shows differences in outward degree, inward degree, and betweenness of co-operation network, respectively. Bottom panel (L to R) shows differences in outward degree, inward degree, and betweenness of competition network, respectively. All comparisons were statistically significant (Kruskal-Wallis test; Adjusted p-value < 0.001).



Supplementary Figure 8: A.) Heatmap of elasticities (change in rate of exchange) of metabolite exports upon increasing the abundance of species in IBS. B.) Differences in network properties of species belonging to cluster 1 (red; $N_{species}$ = 73) and cluster 2 (blue; $N_{species}$ = 36). Top panel (L to R) shows differences in outward degree, inward degree, and betweenness of co-operation network, respectively. Bottom panel (L to R) shows differences in outward degree, inward degree, and betweenness of competition network, respectively. All comparisons were statistically significant (Wilcoxon rank sum test; Adjusted p-value < 0.001).



Supplementary Figure 9: A.) Heatmap of elasticities (change in rate of exchange) of metabolite imports upon increasing the abundance of species in IBS. B.) Differences in network properties of species belonging to cluster 1 (red; $N_{species}$ = 83) and cluster 2 (blue; $N_{species}$ = 26). Top panel (L to R) shows differences in outward degree, inward degree, and betweenness of co-operation network, respectively. Bottom panel (L to R) shows differences in outward degree, inward degree, and betweenness of competition network, respectively. All comparisons were statistically significant (Wilcoxon rank sum test; Adjusted p-value < 0.001).



6.7.2 Supplementary Tables

Supplementary Table 1: Summary statistics of significantly different intracellular reactions producing at least one extracellular metabolite, along with the contributing species based on Welch's t-test between IBS (N=80) and Controls (N=59). Effect size was calculated using Cohen's D formula.

Reaction (ID)	Reaction (name)	Species	IBS (mean)	IBS (SD)	Control (mean)	Control (SD)	Effect size	pvalue	Reactants (name)	Products (name)
PROt4r	proline transporter	*Ruminococcus torques*	1.618	1.842	0.424	1.194	0.769	< 0.001	Sodium;L-proline	Sodium;L-proline
GLCNACASE_OGLYCAN6e	N-acetylglucosaminidase, mucin-type O-glycans, extracellular	*Ruminococcus torques*	1.301	1.66	0.317	0.806	0.754	< 0.001	Core 3;Water	Tn antigen;N-acetyl-D-glucosamine
BGALASE_OGLYCAN14e	Beta-galactosidase, mucin-type O-glycans, extracellular	*Ruminococcus torques*	1.227	1.617	0.284	0.788	0.741	< 0.001	mucin-type O-glycan No 12;Water	Core 3;D-Galactose
BHEXNACASE_OGLYCAN31e	Beta-N-acetylhexosaminidase, mucin-type O-glycans, extracellular	*Ruminococcus torques*	1.003	1.392	0.237	0.618	0.712	< 0.001	mucin-type O-glycan No 2;Water	N-acetyl-D-glucosamine;sialyl-Tn antigen

BHEXNACASE_OGLYCAN10e	Beta-N-acetylhexosaminidase, mucin-type O-glycans, extracellular	*Ruminococcus torques*	1.667	1.936	0.535	1.342	0.679	< 0.001	mucin-type O-glycan No 164;Water	mucin-type O-glycan No 186;N-acetyl-D-glucosamine
GALt4	Galactose Transport via Sodium Symport	*Ruminococcus torques*	1.865	1.911	0.723	1.492	0.666	< 0.001	D-Galactose;Sodium	D-Galactose;Sodium
HISt2r	L-histidine reversible transport via proton symport	*Ruminococcus torques*	1.854	1.771	0.79	1.456	0.656	< 0.001	proton;L-histidine	proton;L-histidine
GALt1r	Galactose Transport (Uniport)	*Ruminococcus torques*	0.94	1.697	0.143	0.526	0.634	< 0.001	D-Galactose	D-Galactose
CGLYabc	L-cysteinylglycine transport via ABC system	*Ruminococcus torques*	1.239	1.67	0.383	0.942	0.631	< 0.001	ATP;L-cysteinylglycine;Water	ADP;L-cysteinylglycine;proton;hydrogenphosphate
MGt5	magnesium transport in/out via permease (no H+)	*Ruminococcus torques*	1.334	1.266	0.601	1.047	0.631	< 0.001	magnesium	magnesium
RIBFLVt2r	riboflavin transport in via proton symport reversible	*Ruminococcus torques*	1.495	1.414	0.681	1.182	0.624	< 0.001	proton;Riboflavin	proton;Riboflavin

COBALTt5	cobalt transport in/out via permease (no H+)	*Ruminococcus torques*	1.318	1.263	0.597	1.041	0.624	< 0.001	Co2+	Co2+
ASNt2r	L-asparagine reversible transport via proton symport	*Ruminococcus torques*	2.024	1.962	0.893	1.668	0.621	< 0.001	L-asparagine;proton	L-asparagine;proton
NACt2r	Nicotinic acid transport by proton symport reversible	*Ruminococcus torques*	1.485	1.413	0.681	1.182	0.618	< 0.001	proton;Nicotinate	proton;Nicotinate
ILEt2r	L-isoleucine reversible transport via proton symport	*Ruminococcus torques*	2.152	2.021	0.998	1.728	0.614	< 0.001	proton;L-isoleucine	proton;L-isoleucine
URIt2r	Uridine Transport in via Proton Symport Reversible	*Ruminococcus torques*	2.158	2.03	1.003	1.737	0.611	< 0.001	proton;Uridine	proton;Uridine
PHEt2r	L-phenylalanine reversible transport via proton symport	*Ruminococcus torques*	2.035	1.921	0.949	1.644	0.608	< 0.001	proton;L-phenylalanine	proton;L-phenylalanine

FORt2r	formate transport in via proton symport	*Ruminococcus torques*	2.289	2.225	1.053	1.826	0.607	< 0.001	Formate;proton	Formate;proton
LYSt3r	L-lysine transport out via proton antiport reversible	*Ruminococcus torques*	3.493	3.266	1.642	2.84	0.605	< 0.001	proton;L-lysinium(1+)	proton;L-lysinium(1+)
TYRt2r	L-tyrosine reversible transport via proton symport	*Ruminococcus torques*	1.959	1.856	0.916	1.587	0.604	< 0.001	proton;L-tyrosine	proton;L-tyrosine
ATPS4	ATP synthase (four protons for one ATP)	*Ruminococcus torques*	3.323	3.116	1.564	2.705	0.603	< 0.001	ADP;proton;hydrogenphosphate	ATP;Water;proton
DALAt2r	D-Alanine Transport via Proton Symport	*Ruminococcus torques*	2.058	1.964	0.961	1.67	0.602	< 0.001	D-alanine;proton	D-alanine;proton
LYSt2r	L-lysine reversible transport via proton symport	*Ruminococcus torques*	3.485	3.268	1.642	2.84	0.602	< 0.001	proton;L-lysinium(1+)	proton;L-lysinium(1+)
CA2abc	calcium transport via ABC system	*Ruminococcus torques*	1.297	1.278	0.602	1.047	0.596	0.001	ATP;calcium(2+);Water	ADP;calcium(2+);proton;hydrogenphosphate
FE2abc	iron (II) transport via ABC system	*Ruminococcus torques*	1.297	1.278	0.602	1.047	0.596	0.001	ATP;Fe2+;Water	ADP;Fe2+;proton;hydrogenphosphate

FE3abc	iron (III) transport via ABC system	*Ruminococcus torques*	1.297	1.278	0.602	1.047	0.596	0.001	ATP;Fe3+;Water	ADP;Fe3+;proton;hydrogenphosphate
MNabc	manganese transport via ABC system	*Ruminococcus torques*	1.297	1.278	0.602	1.047	0.596	0.001	ATP;Water;Mn2+	ADP;proton;Mn2+;hydrogenphosphate
r2137	Major Facilitator(MFS) TCDB:2.A.1.14.6	*Ruminococcus torques*	1.297	1.277	0.602	1.047	0.596	0.001	Chloride	Chloride
SO4t2	sulfate transport in via proton symport	*Ruminococcus torques*	1.297	1.277	0.602	1.047	0.596	0.001	proton;sulfate	proton;sulfate
ZNabc	zinc transport via ABC system	*Ruminococcus torques*	1.297	1.278	0.602	1.047	0.596	0.001	ATP;Water;Zinc	ADP;proton;hydrogenphosphate;Zinc
H2Ot	H2O Transport via Diffusion	*Ruminococcus torques*	2.904	2.767	1.374	2.378	0.593	0.001	Water	Water
PNTOt4	Pantothenate sodium symporter	*Ruminococcus torques*	1.462	1.448	0.683	1.184	0.589	0.001	Sodium;(R)-Pantothenate	Sodium;(R)-Pantothenate
SIAASE_OGLYCAN1e	Sialidase, mucin-type O-glycans, extracellular	*Ruminococcus torques*	0.804	1.352	0.186	0.607	0.589	< 0.001	Water;sialyl-Tn antigen	Tn antigen;N-acetylneuraminate

GLNt2r	L-glutamine reversible transport via proton symport	*Ruminococcus torques*	2.309	2.214	1.099	1.908	0.586	0.001	L-glutamine;proton	L-glutamine;proton
GLYCt	Glycerol Transport via Channel	*Dorea formicigenerans*	1.27	1.425	0.521	1.115	0.585	0.001	Glycerol	Glycerol
SERt2r	L-serine reversible transport via proton symport	*Ruminococcus torques*	2.092	2.035	0.993	1.721	0.583	0.001	proton;L-serine	proton;L-serine
BGALASE_OGLYCAN10e	Beta-galactosidase, mucin-type O-glycans, extracellular	*Ruminococcus torques*	0.951	1.792	0.116	0.955	0.581	0.001	mucin-type O-glycan No 113;Water	mucin-type O-glycan No 163;D-Galactose
O2t	O2 Transport (Diffusion)	*Ruminococcus torques*	0.768	0.933	0.311	0.623	0.577	0.001	O2	O2
ACGApts	N-Acetyl-D-glucosamine transport via PEP:Pyr PTS	*Ruminococcus torques*	2.212	2.232	1.047	1.816	0.573	0.001	N-acetyl-D-glucosamine;Phosphoenolpyruvate	N-acetyl-D-glucosamine 6-phosphate(2-);pyruvate
Kt1r	potassium transport via uniport reversible (facilitated diffusion)	*Ruminococcus torques*	1.27	1.293	0.599	1.044	0.571	0.001	potassium	potassium

SIAASE_OGLYCAN59e	Sialidase, mucin-type O-glycans, extracellular	*Ruminococcus torques*	0.592	1.073	0.119	0.496	0.566	0.001	mucin-type O-glycan No 2;Water	N-acetylneuraminate;Core 3
ESHCYS_FPe	Extracellular electron shuttle, cysteine to oxygen, F. prausnitzii specific	*Ruminococcus torques*	0.692	0.878	0.28	0.54	0.565	0.001	L-cystine;Water	L-cysteine;O2
CYSt2r	L-cysteine reversible transport via proton symport	*Dorea formicigenerans*	1.369	1.589	0.564	1.249	0.564	0.001	L-cysteine;proton	L-cysteine;proton
PYDXtr	Transport of Pyridoxal	*Ruminococcus torques*	1.276	1.304	0.608	1.055	0.563	0.001	Pyridoxal	Pyridoxal
DAPabc	M-diaminopimelic acid ABC transport	*Ruminococcus torques*	1.825	1.852	0.874	1.515	0.562	0.001	meso-2,6-Diaminoheptanedioate;ATP;Water	meso-2,6-Diaminoheptanedioate;ADP;proton;hydrogenphosphate
GUAt2	guanine transport in via proton symport	*Dorea formicigenerans*	1.35	1.497	0.591	1.207	0.558	0.001	Guanine;proton	Guanine;proton
HYXNt	Hypoxanthine Transport	*Ruminococcus torques*	2.149	2.168	1.039	1.802	0.557	0.001	Hypoxanthine	Hypoxanthine

BHEXNACASE_OGLYCAN31e	Beta-N-acetylhexosaminidase, mucin-type O-glycans, extracellular	*Dorea formicigenerans*	0.699	1.328	0.152	0.424	0.555	0.001	mucin-type O-glycan No 2;Water	N-acetyl-D-glucosamine;sialyl-Tn antigen
GLYLEUPEPT1tc	Transport of Glycylleucine by the Apical Pept1 Amino Acid Transporters	*Dorea formicigenerans*	1.726	1.919	0.765	1.524	0.554	0.001	Glycylleucine;proton	Glycylleucine;proton
ACt2r	Acetate Reversible Transport via Proton Symport	*Ruminococcus torques*	2.433	2.473	1.177	2.039	0.554	0.001	acetate;proton	acetate;proton
PROt4r	proline transporter	*Dorea formicigenerans*	1.6	1.754	0.729	1.373	0.553	0.001	Sodium;L-proline	Sodium;L-proline
FE2abc	iron (II) transport via ABC system	*Dorea formicigenerans*	0.964	1.131	0.398	0.906	0.551	0.001	ATP;Fe2+;Water	ADP;Fe2+;proton;hydrogenphosphate
DAPabc	M-diaminopimelic acid ABC transport	*Dorea formicigenerans*	0.486	0.665	0.146	0.567	0.55	0.002	meso-2,6-Diaminoheptanedioate;ATP;Water	meso-2,6-Diaminoheptanedioate;ADP;proton;hydrogenphosphate
NAt3_1	Sodium Proton Antiporter (H:Na Is 1:1)	*Dorea formicigenerans*	1.519	1.82	0.629	1.424	0.545	0.002	proton;Sodium	proton;Sodium

PNTOt4	Pantothenate sodium symporter	*Dorea formicigenerans*	1.131	1.256	0.509	1.019	0.544	0.002	Sodium;(R)-Pantothenate	Sodium;(R)-Pantothenate
ACGApts	N-Acetyl-D-glucosamine transport via PEP:Pyr PTS	*Dorea formicigenerans*	1.797	2.027	0.821	1.537	0.542	0.002	N-acetyl-D-glucosamine;Phosphoenolpyruvate	N-acetyl-D-glucosamine 6-phosphate(2-);pyruvate
NH4tb	Transport of Ammonium ion	*Dorea formicigenerans*	2.049	2.256	0.951	1.786	0.54	0.002	Ammonium	Ammonium
METt2r	L-methionine reversible transport via proton symport	*Ruminococcus torques*	1.884	1.915	0.929	1.611	0.539	0.002	proton;L-methionine	proton;L-methionine
MK7te	Menaquinone secretion (menaquinone 7)	*Dorea formicigenerans*	1.229	1.354	0.564	1.111	0.537	0.002	Menaquinone 7	Menaquinone 7
GLYC3Pabc	sn-Glycerol 3-phosphate transport via ABC system	*Ruminococcus torques*	2.317	2.332	1.154	1.998	0.536	0.002	ATP;Glycerol 3-phosphate;Water	ADP;Glycerol 3-phosphate;proton;hydrogenphosphate
NACt2r	Nicotinic acid transport by proton symport reversible	*Dorea formicigenerans*	1.957	2.146	0.919	1.734	0.532	0.002	proton;Nicotinate	proton;Nicotinate

Coabc	Cobalt transport via ABC system	*Dorea formicigenerans*	0.994	1.109	0.464	0.881	0.528	0.002	ATP;Co2+;Water	ADP;Co2+;proton;hydrogenphosphate
CAt4i	Calcium transport via proton antiport, irreversible	*Dorea formicigenerans*	0.996	1.11	0.467	0.882	0.528	0.002	calcium(2+);proton	calcium(2+);proton
FE3abc	iron (III) transport via ABC system	*Dorea formicigenerans*	0.995	1.11	0.466	0.882	0.528	0.002	ATP;Fe3+;Water	ADP;Fe3+;proton;hydrogenphosphate
MG2abc	magnesium transport via ABC system	*Dorea formicigenerans*	0.995	1.11	0.466	0.882	0.528	0.002	ATP;Water;magnesium	magnesium;ADP;proton;hydrogenphosphate
MNt2	manganese transport in via proton symport	*Dorea formicigenerans*	0.995	1.11	0.466	0.882	0.528	0.002	proton;Mn2+	proton;Mn2+
r2137	Major Facilitator(MFS) TCDB:2.A.1.14.6	*Dorea formicigenerans*	0.996	1.11	0.466	0.882	0.528	0.002	Chloride	Chloride
THMt3	Thiamine Transport in via Proton Antiport	*Dorea formicigenerans*	0.997	1.11	0.467	0.883	0.528	0.002	proton;Thiamin	proton;Thiamin
Cut1	Copper export via ATPase	*Dorea formicigenerans*	0.618	0.898	0.2	0.669	0.528	0.002	ADP;Cu2+;proton;hydrogenphosphate	ATP;Cu2+;Water

ESHGLU_FPe	Extracellular electron shuttle, glutathione to oxygen, F. prausnitzii specific	*Ruminococcus torques*	0.682	0.86	0.3	0.552	0.528	0.002	Oxidized glutathione;Water	Reduced glutathione;O2
ZN2t4	ZN2t4	*Dorea formicigenerans*	0.996	1.11	0.467	0.882	0.528	0.002	proton;potassium;Zinc	proton;potassium;Zinc
ACNAMtNa	N-acetylneuraminate sodium symport	*Ruminococcus torques*	1.604	1.8	0.739	1.468	0.527	0.002	N-acetylneuraminate;Sodium	N-acetylneuraminate;Sodium
SERt2r	L-serine reversible transport via proton symport	*Dorea formicigenerans*	1.483	1.714	0.66	1.392	0.527	0.002	proton;L-serine	proton;L-serine
ACt2r	Acetate Reversible Transport via Proton Symport	*Dorea formicigenerans*	2.443	2.668	1.166	2.173	0.525	0.002	acetate;proton	acetate;proton
FORt2r	formate transport in via proton symport	*Dorea formicigenerans*	2.446	2.671	1.167	2.18	0.525	0.002	Formate;proton	Formate;proton
Kt3r	potassium proton antiporter	*Dorea formicigenerans*	1.134	1.256	0.536	1.005	0.525	0.002	proton;potassium	proton;potassium
AMPt2	AMP transport in via proton symport, irreversible	*Dorea formicigenerans*	1.453	1.684	0.654	1.353	0.523	0.002	AMP;proton	AMP;proton

PYDXabc	Pyridoxal transport via ECF transporter	*Dorea formicigenerans*	1.009	1.116	0.465	0.96	0.523	0.002	ATP;Water;Pyridoxal	ADP;proton;hydrogenphosphate;Pyridoxal
NMNP	NMN permease	*Dorea formicigenerans*	1.961	2.15	0.94	1.741	0.522	0.002	NMN	NMN
SHEMEabc	Siroheme transport via ABC system	*Dorea formicigenerans*	1.04	1.15	0.497	0.917	0.522	0.002	ATP;Water;Siroheme	ADP;proton;hydrogenphosphate;Siroheme
VALt2r	L-valine reversible transport via proton symport	*Ruminococcus torques*	2.086	2.194	1.039	1.799	0.522	0.002	proton;L-valine	proton;L-valine
ASNt2r	L-asparagine reversible transport via proton symport	*Dorea formicigenerans*	1.583	1.735	0.76	1.41	0.521	0.002	L-asparagine;proton	L-asparagine;proton
DCYTt2r	deoxycytidine transport in via proton symport reversible	*Dorea formicigenerans*	1.35	1.472	0.654	1.196	0.519	0.003	Deoxycytidine;proton	Deoxycytidine;proton
12DGR180ti	1,2 diacylglycerol transport (extracellular to cytoplasm, n-C18:0)	*Dorea formicigenerans*	1.339	1.472	0.644	1.196	0.518	0.003	1,2-Diacyl-sn-glycerol (dioctadecanoyl, n-C18:0)	1,2-Diacyl-sn-glycerol (dioctadecanoyl, n-C18:0)

GLNt2r	L-glutamine reversible transport via proton symport	*Dorea formicigenerans*	1.682	1.86	0.807	1.497	0.518	0.003	L-glutamine;proton	L-glutamine;proton
GLYt2r	Glycine Reversible Transport via Proton Symport	*Ruminococcus torques*	0.88	1.544	0.249	0.777	0.517	0.002	Glycine;proton	Glycine;proton
TRPt2r	L-tryptophan reversible transport via proton symport	*Dorea formicigenerans*	1.385	1.523	0.668	1.234	0.517	0.003	proton;L-tryptophan	proton;L-tryptophan
ALAHISabc	L-alanyl-L-histidine transport via ABC system	*Dorea formicigenerans*	1.47	1.611	0.712	1.31	0.516	0.003	L-alanyl-L-histidine;ATP;Water	ADP;L-alanyl-L-histidine;proton;hydrogenphosphate
PHEt2r	L-phenylalanine reversible transport via proton symport	*Dorea formicigenerans*	1.614	1.771	0.782	1.44	0.516	0.003	proton;L-phenylalanine	proton;L-phenylalanine
ARGt2	L arginine transport in via proton symport	*Dorea formicigenerans*	1.643	1.8	0.798	1.466	0.515	0.003	L-argininium(1+);proton	L-argininium(1+);proton
HEMEti	Heme transport via ABC system	*Dorea formicigenerans*	0.942	1.07	0.439	0.87	0.515	0.003	ATP;Water;Protoheme	ADP;proton;Protoheme;hydrogenphosphate

LYSt2r	L-lysine reversible transport via proton symport	*Dorea formicigenerans*	1.739	1.904	0.845	1.553	0.515	0.003	proton;L-lysinium(1+)	proton;L-lysinium(1+)
TYRt2r	L-tyrosine reversible transport via proton symport	*Dorea formicigenerans*	1.517	1.667	0.735	1.353	0.515	0.003	proton;L-tyrosine	proton;L-tyrosine
ILEt2r	L-isoleucine reversible transport via proton symport	*Dorea formicigenerans*	1.756	1.922	0.855	1.568	0.514	0.003	proton;L-isoleucine	proton;L-isoleucine
VALt2r	L-valine reversible transport via proton symport	*Dorea formicigenerans*	1.729	1.893	0.843	1.546	0.513	0.003	proton;L-valine	proton;L-valine
GCALDt	Glycoaldehydye reversible transport	*Dorea formicigenerans*	1.525	1.681	0.743	1.361	0.512	0.003	glycolaldehyde	glycolaldehyde
Cuabc	Copper transport via ABC system	*Dorea formicigenerans*	0.853	1.073	0.337	0.949	0.51	0.003	Cu2+;ATP;Water	Cu2+;ADP;proton;hydrogenphosphate
THRt2r	L-threonine reversible transport via proton symport	*Dorea formicigenerans*	1.631	1.786	0.798	1.469	0.509	0.003	proton;L-threonine	proton;L-threonine

CYTBD	cytochrome oxidase bd (ubiquinol-8: 2 protons)	*Dorea formicigenerans*	0.919	1.338	0.322	0.989	0.507	0.003	proton;O2;Ubiquinol-8	Water;proton;Ubiquinone-8
RIBFLVt2r	riboflavin transport in via proton symport reversible	*Dorea formicigenerans*	1.148	1.264	0.563	1.032	0.507	0.003	proton;Riboflavin	proton;Riboflavin
ALAt2r	L-Alanine Reversible Transport via Proton Symport	*Dorea formicigenerans*	1.7	1.947	0.809	1.549	0.506	0.003	L-alanine;proton	L-alanine;proton
ATPS4	ATP synthase (four protons for one ATP)	*Dorea formicigenerans*	2.571	2.801	1.272	2.315	0.506	0.003	ADP;proton;hydrogenphosphate	ATP;Water;proton
FOLt	FOLt	*Dorea formicigenerans*	1.232	1.353	0.606	1.107	0.506	0.003	Folate;proton	Folate;proton
H2Ot	H2O Transport via Diffusion	*Dorea formicigenerans*	2.831	3.082	1.407	2.558	0.503	0.004	Water	Water
BHEXNACASE_OGLYCAN10e	Beta-N-acetylhexosaminidase, mucin-type O-glycans, extracellular	*Dorea formicigenerans*	1.149	1.761	0.409	1.118	0.502	0.003	mucin-type O-glycan No 164;Water	mucin-type O-glycan No 186;N-acetyl-D-glucosamine

METt2r	L-methionine reversible transport via proton symport	*Dorea formicigenerans*	1.434	1.728	0.668	1.32	0.498	0.004	proton;L-methionine	proton;L-methionine
CO2t	CO2 Transporter via Diffusion	*Ruminococcus torques*	2.264	2.494	1.138	2	0.498	0.004	carbon dioxide	carbon dioxide
THRt2r	L-threonine reversible transport via proton symport	*Ruminococcus torques*	1.764	2.007	0.849	1.66	0.497	0.004	proton;L-threonine	proton;L-threonine
BGALASE_OGLYCAN90e	Beta-galactosidase, mucin-type O-glycans, extracellular	*Ruminococcus torques*	1.199	1.576	0.518	1.134	0.496	0.004	released mucin-type O-glycan No 12;Water	Beta-1,3-N-acetylglucosamine-N-acetylgalactosamine;D-Galactose
Cuabc	Copper transport via ABC system	*Ruminococcus torques*	1.186	1.342	0.593	1.037	0.495	0.004	Cu2+;ATP;Water	Cu2+;ADP;proton;hydrogenphosphate
SPMDabc	spermidine transport via ABC system	*Dorea formicigenerans*	1.013	1.16	0.498	0.928	0.49	0.004	ATP;Water;Spermidine	ADP;proton;hydrogenphosphate;Spermidine
URAt2	uracil transport in via proton symport	*Dorea formicigenerans*	1.084	1.449	0.448	1.139	0.489	0.004	proton;uracil	proton;uracil
ACGAMtr2	acgam[c] transport via proton symport	*Ruminococcus torques*	0.988	1.724	0.302	0.992	0.488	0.004	N-acetyl-D-glucosamine;proton	N-acetyl-D-glucosamine;proton

HDCAtr	Fatty Acid Transport via Diffusion	*Parabacteroides distasonis*	0.309	0.49	0.115	0.294	0.48	0.004	Hexadecanoate (n-C16:0)	Hexadecanoate (n-C16:0)
ESHGLU2_FPe	Extracellular electron shuttle, riboflavin to glutathione, F. prausnitzii specific	*Ruminococcus torques*	0.709	0.997	0.309	0.649	0.475	0.005	Reduced glutathione;Riboflavin	Oxidized glutathione;Reduced riboflavin
FUMt2r	fumarate reversible transport via symport	*Dorea formicigenerans*	1.287	1.422	0.665	1.221	0.47	0.006	Fumarate;proton	Fumarate;proton
GLUt2r	L-glutamate transport via proton symport, reversible	*Dorea formicigenerans*	1.542	1.822	0.762	1.479	0.47	0.006	L-glutamate(1-);proton	L-glutamate(1-);proton
SIAASE_OGLYCAN146e	Sialidase, mucin-type O-glycans, extracellular	*Ruminococcus torques*	0.708	1.059	0.271	0.777	0.47	0.006	released mucin-type O-glycan No 2;Water	N-acetylneuraminate;Beta-1,3-N-acetylglucosamine-N-acetylgalactosamine
BHEXNACASE_OGLYCAN45e	Beta-N-acetylhexosaminidase, mucin-type O-glycans, extracellular	*Ruminococcus torques*	1.502	2.019	0.682	1.452	0.467	0.006	released mucin-type O-glycan No 137;Water	released mucin-type O-glycan No 195;N-acetyl-D-glucosamine

BHEXNACASE_OGLYCAN77e	Beta-N-acetylhexosaminidase, mucin-type O-glycans, extracellular	*Ruminococcus torques*	0.972	1.42	0.41	0.948	0.465	0.006	released mucin-type O-glycan No 2;Water	N-acetyl-D-glucosamine;Alpha-2,6-N-acetylneuraminate-N-acetylgalactosamine
12aHSDHe	12-alpha hydroxysteroid dehydrogenase, NAD+ dependent, cholate specific	*Dorea formicigenerans*	2.111	2.55	1.03	2.11	0.462	0.007	12-Dehydrocholate;proton;Nicotinamide adenine dinucleotide - reduced	Cholate;Nicotinamide adenine dinucleotide
GAM2STASE_HS3	N-sulfoglucosamine sulfohydrolase, heparin degradation	*Bacteroides thetaiotaomicron*	0.153	0.272	0.033	0.247	0.461	0.008	Water;Product AS3TASE_HS2	proton;Product GAM2STASE_HS3;sulfate
SO4t2	sulfate transport in via proton symport	*Dorea formicigenerans*	0.795	1.201	0.312	0.875	0.46	0.007	proton;sulfate	proton;sulfate
FUCASE_OGLYCAN153e	Alpha-L-fucosidase, mucin-type O-glycans, extracellular	*Bacteroides uniformis*	0.4	0.946	0.017	0.71	0.458	0.007	released mucin-type O-glycan No 182;Water	Beta-1,6-N-acetylglucosamine-N-acetylgalactosamine;L-fucose

LEUt2r	L-leucine reversible transport via proton symport	*Dorea formicigenerans*	1.062	1.456	0.458	1.173	0.457	0.008	proton;L-leucine	proton;L-leucine
THMabc	thiamine transport via ABC system	*Bifidobacterium longum*	1.13	1.03	0.67	0.989	0.456	0.009	ATP;Water;Thiamin	ADP;proton;hydrogenphosphate;Thiamin
ALAt2r	L-Alanine Reversible Transport via Proton Symport	*Ruminococcus torques*	1.388	2.026	0.602	1.367	0.454	0.007	L-alanine;proton	L-alanine;proton
BGALASE_OGLYCAN90e	Beta-galactosidase, mucin-type O-glycans, extracellular	*Bacteroides vulgatus*	0.446	1.195	0.043	0.385	0.454	0.006	released mucin-type O-glycan No 12;Water	Beta-1,3-N-acetylglucosamine-N-acetylgalactosamine;D-Galactose
BHEXNACASE_OGLYCAN89e	Beta-N-acetylhexosaminidase, mucin-type O-glycans, extracellular	*Ruminococcus torques*	1.193	1.679	0.542	1.165	0.451	0.008	Beta-1,3-N-acetylglucosamine-N-acetylgalactosamine;Water	N-acetyl-D-galactosamine;N-acetyl-D-glucosamine
SUCCt2r	succinate transport via proton symport	*Dorea formicigenerans*	1.077	1.36	0.516	1.139	0.447	0.009	proton;Succinate	proton;Succinate
BTNabc	Biotin transport via ABC system	*Ruminococcus torques*	0.838	0.924	0.454	0.79	0.446	0.01	Biotin;ATP;Water	ADP;Biotin;proton;hydrogenphosphate

FRUpts	D-fructose transport via PEP:Pyr PTS	*Dorea formicigenerans*	1.981	2.366	1.004	2.004	0.446	0.01	D-Fructose;Phosphoenolpyruvate	D-Fructose 1-phosphate;pyruvate
PIMEtr	pimelate transport via diffusion, reversible	*Ruminococcus torques*	0.837	0.925	0.454	0.79	0.446	0.01	Pimelate	Pimelate
CS4TASE	chondro-4-sulfatase	*Bacteroides thetaiotaomicron*	0.152	0.291	0.037	0.223	0.442	0.01	Dimer GalNAc4S-GlcA;Water	Dimer GalNAc-GlcA;proton;sulfate
CS6TASE	chondro-6-sulfatase	*Bacteroides thetaiotaomicron*	0.15	0.288	0.037	0.216	0.441	0.01	Dimer GalNAc6S-GlcA;Water	Dimer GalNAc-GlcA;proton;sulfate
AS3TASE_HS1	N-acetylglucosamine-3-sulfatase, heparin degradation	*Bacteroides thetaiotaomicron*	0.265	0.502	0.08	0.327	0.437	0.01	Product AS6TASE_HS1, GlcN2S-GlcA-GlcN2,3diS;Water	Product AS3TASE_HS1, GlcN2S-GlcA-GlcN2S;sulfate;proton
GLCNACASE_HS1	Alpha-N-acetylglucosaminidase, heparin degradation	*Bacteroides thetaiotaomicron*	0.241	0.677	0.013	0.314	0.432	0.009	Water;proton;Product GAM2STASE_HS2	D-Glucosamine;D-glucuronate
SO4t2	sulfate transport in via proton symport	*Faecalibacterium prausnitzii*	2.53	1.43	3.018	0.739	-0.43	0.01	proton;sulfate	proton;sulfate

ESHGLU_FPe	Extracellular electron shuttle, glutathione to oxygen, F. prausnitzii specific	*Faecalibacterium prausnitzii*	1.681	1.093	2.066	0.612	-0.43	0.01	Oxidized glutathione;Water	Reduced glutathione;O2
ALAt2r	L-Alanine Reversible Transport via Proton Symport	*Faecalibacterium prausnitzii*	3.914	2.09	4.639	1.088	-0.44	0.009	L-alanine;proton	L-alanine;proton
PIt7ir	PIt7ir	*Ruminococcus lactaris*	1.075	1.958	2.032	2.244	-0.45	0.01	Sodium;hydrogenphosphate	Sodium;hydrogenphosphate
ADNCNT3tc	Adenosine Transport Cnt3	*Odoribacter splanchnicus*	1.383	1.866	2.231	1.864	-0.46	0.009	Adenosine;proton	Adenosine;proton
CYSt2r	L-cysteine reversible transport via proton symport	*Odoribacter splanchnicus*	1.182	1.582	1.924	1.68	-0.46	0.009	L-cysteine;proton	L-cysteine;proton
ACGApts	N-Acetyl-D-glucosamine transport via PEP:Pyr PTS	*Ruminococcus lactaris*	1.128	2.086	2.158	2.396	-0.46	0.009	N-acetyl-D-glucosamine;Phosphoenolpyruvate	N-acetyl-D-glucosamine 6-phosphate(2-);pyruvate
GLUt2r	L-glutamate transport via proton symport, reversible	*Odoribacter splanchnicus*	1.628	2.212	2.678	2.371	-0.46	0.009	L-glutamate(1-);proton	L-glutamate(1-);proton

GCALDt	Glycoaldehydye reversible transport	*Odoribacter splanchnicus*	0.851	1.145	1.389	1.198	-0.46	0.009	glycolaldehyde	glycolaldehyde
SERt2r	L-serine reversible transport via proton symport	*Ruminococcus lactaris*	1.041	1.886	1.984	2.199	-0.46	0.009	proton;L-serine	proton;L-serine
4HPRO_LTt	trans-4-hydroxy-L-proline transport	*Odoribacter splanchnicus*	1.302	1.742	2.126	1.833	-0.46	0.009	trans-4-hydroxy-L-proline;proton	trans-4-hydroxy-L-proline;proton
GUAt2	guanine transport in via proton symport	*Faecalibacterium prausnitzii*	3.581	1.507	4.126	0.7	-0.46	0.005	Guanine;proton	Guanine;proton
COBALTt5	cobalt transport in/out via permease (no H+)	*Odoribacter splanchnicus*	0.661	0.891	1.09	0.957	-0.46	0.008	Co2+	Co2+
PHEt2r	L-phenylalanine reversible transport via proton symport	*Barnesiella intestinihominis*	1.877	2.062	2.792	1.87	-0.47	0.007	proton;L-phenylalanine	proton;L-phenylalanine
Cuabc	Copper transport via ABC system	*Barnesiella intestinihominis*	1.009	1.075	1.478	0.931	-0.47	0.007	Cu2+;ATP;Water	Cu2+;ADP;proton;hydrogenphosphate
GLUABUTt7	4-aminobutyrate/glutamate antiport	*Odoribacter splanchnicus*	1.152	2.135	2.269	2.631	-0.47	0.009	4-Aminobutanoate;L-glutamate(1-)	4-Aminobutanoate;L-glutamate(1-)

SO4t2	sulfate transport in via proton symport	*Odoribacter splanchnicus*	0.661	0.892	1.092	0.957	-0.47	0.008	proton;sulfate	proton;sulfate
CAt4i	Calcium transport via proton antiport, irreversible	*Odoribacter splanchnicus*	0.662	0.892	1.095	0.958	-0.47	0.008	calcium(2+);proton	calcium(2+);proton
HEMEti	Heme transport via ABC system	*Odoribacter splanchnicus*	0.662	0.892	1.095	0.958	-0.47	0.008	ATP;Water;Protoheme	ADP;proton;Protoheme;hydrogenphosphate
MNt2	manganese transport in via proton symport	*Odoribacter splanchnicus*	0.662	0.892	1.095	0.958	-0.47	0.008	proton;Mn2+	proton;Mn2+
r2137	Major Facilitator(MFS) TCDB:2.A.1.14.6	*Odoribacter splanchnicus*	0.662	0.892	1.095	0.958	-0.47	0.008	Chloride	Chloride
SHEMEabc	Siroheme transport via ABC system	*Odoribacter splanchnicus*	0.662	0.892	1.095	0.958	-0.47	0.008	ATP;Water;Siroheme	ADP;proton;hydrogenphosphate;Siroheme
SPMDabc	spermidine transport via ABC system	*Odoribacter splanchnicus*	0.662	0.892	1.095	0.958	-0.47	0.008	ATP;Water;Spermidine	ADP;proton;hydrogenphosphate;Spermidine
CGLYabc	L-cysteinylglycine transport via ABC system	*Odoribacter splanchnicus*	0.634	0.862	1.052	0.924	-0.47	0.008	ATP;L-cysteinylglycine;Water	ADP;L-cysteinylglycine;proton;hydrogenphosphate

CYFO	Aerobic reductase, extracellular (Riboflavin)	*Faecalibacterium prausnitzii*	1.665	1.144	2.089	0.576	-0.47	0.005	Riboflavin;Water	O2;Reduced riboflavin
ASNt2r	L-asparagine reversible transport via proton symport	*Odoribacter splanchnicus*	1.327	1.776	2.191	1.911	-0.47	0.008	L-asparagine;proton	L-asparagine;proton
Kt1r	potassium transport via uniport reversible (facilitated diffusion)	*Faecalibacterium prausnitzii*	2.914	1.169	3.348	0.581	-0.47	0.005	potassium	potassium
OCDCAtr	Fatty Acid Transport via Diffusion	*Barnesiella intestinihominis*	1.604	1.695	2.357	1.502	-0.47	0.006	octadecanoate (n-C18:0)	octadecanoate (n-C18:0)
ZN2t4	ZN2t4	*Barnesiella intestinihominis*	1.058	1.171	1.576	1.03	-0.47	0.006	proton;potassium;Zinc	proton;potassium;Zinc
4HBZt2	4-Hydroxybenzoate transport via proton symport	*Barnesiella intestinihominis*	1.074	1.16	1.592	1.034	-0.47	0.006	4-hydroxybenzoate;proton	4-hydroxybenzoate;proton
Cuabc	Copper transport via ABC system	*Faecalibacterium prausnitzii*	2.55	1.342	3.043	0.603	-0.47	0.004	Cu2+;ATP;Water	Cu2+;ADP;proton;hydrogenphosphate

NADH6	NADH dehydrogenase (ubiquinone-8 & 3.5 protons)	*Odoribacter splanchnicus*	1.973	2.636	3.222	2.617	-0.48	0.006	proton;Nicotinamide adenine dinucleotide - reduced;Ubiquinone-8	proton;Nicotinamide adenine dinucleotide;Ubiquinol-8
SERt2r	L-serine reversible transport via proton symport	*Odoribacter splanchnicus*	1.352	1.828	2.255	1.966	-0.48	0.007	proton;L-serine	proton;L-serine
GLCNACASE_OGLYCAN3e	N-acetylglucosaminidase, mucin-type O-glycans, extracellular	*Barnesiella intestinihominis*	0.653	1.118	1.239	1.331	-0.48	0.007	Core 2;Water	T antigen (core 1);N-acetyl-D-glucosamine
12DGR180ti	1,2 diacylglycerol transport (extracellular to cytoplasm, n-C18:0)	*Faecalibacterium prausnitzii*	2.535	1.105	2.946	0.507	-0.48	0.004	1,2-Diacyl-sn-glycerol (dioctadecanoyl, n-C18:0)	1,2-Diacyl-sn-glycerol (dioctadecanoyl, n-C18:0)
Coabc	Cobalt transport via ABC system	*Faecalibacterium prausnitzii*	2.636	1.148	3.068	0.554	-0.48	0.004	ATP;Co2+;Water	ADP;Co2+;proton;hydrogenphosphate
ZN2t4	ZN2t4	*Odoribacter splanchnicus*	0.675	0.909	1.12	0.952	-0.48	0.006	proton;potassium;Zinc	proton;potassium;Zinc
LYSt2r	L-lysine reversible transport via proton symport	*Odoribacter splanchnicus*	1.387	1.861	2.285	1.889	-0.48	0.006	proton;L-lysinium(1+)	proton;L-lysinium(1+)

RIBFLVt2r	riboflavin transport in via proton symport reversible	*Faecalibacterium prausnitzii*	2.901	1.23	3.363	0.586	-0.48	0.004	proton;Riboflavin	proton;Riboflavin
TRPt2r	L-tryptophan reversible transport via proton symport	*Faecalibacterium prausnitzii*	3.378	1.377	3.892	0.64	-0.48	0.004	proton;L-tryptophan	proton;L-tryptophan
CYTDt2r	Cytidine Transport in via Proton Symport, Reversible	*Odoribacter splanchnicus*	1.301	1.74	2.143	1.766	-0.48	0.006	Cytidine;proton	Cytidine;proton
FE3abc	iron (III) transport via ABC system	*Faecalibacterium prausnitzii*	2.629	1.163	3.067	0.556	-0.48	0.004	ATP;Fe3+;Water	ADP;Fe3+;proton;hydrogenphosphate
r2137	Major Facilitator(MFS) TCDB:2.A.1.14.6	*Faecalibacterium prausnitzii*	2.629	1.163	3.067	0.556	-0.48	0.004	Chloride	Chloride
DGSNt2	deoxyguanosine transport in via proton symport	*Odoribacter splanchnicus*	1.288	1.724	2.128	1.766	-0.48	0.006	Deoxyguanosine;proton	Deoxyguanosine;proton
SPMDabc	spermidine transport via ABC system	*Faecalibacterium prausnitzii*	2.553	1.065	2.959	0.532	-0.48	0.004	ATP;Water;Spermidine	ADP;proton;hydrogenphosphate;Spermidine

THMt3	Thiamine Transport in via Proton Antiport	*Faecalibacterium prausnitzii*	2.521	1.124	2.947	0.544	-0.48	0.004	proton;Thiamin	proton;Thiamin
CSNt6	cytosine transport inout via proton symporter	*Lachnospira eligens*	0.054	0.488	0.637	1.629	-0.49	0.01	cytosine;proton	cytosine;proton
THRt2r	L-threonine reversible transport via proton symport	*Odoribacter splanchnicus*	1.535	2.054	2.542	2.097	-0.49	0.006	proton;L-threonine	proton;L-threonine
ACt2r	Acetate Reversible Transport via Proton Symport	*Odoribacter splanchnicus*	1.861	2.488	3.076	2.509	-0.49	0.005	acetate;proton	acetate;proton
FE2abc	iron (II) transport via ABC system	*Faecalibacterium prausnitzii*	2.6	1.23	3.065	0.561	-0.49	0.003	ATP;Fe2+;Water	ADP;Fe2+;proton;hydrogenphosphate
FOLt	FOLt	*Faecalibacterium prausnitzii*	2.932	1.275	3.416	0.59	-0.49	0.003	Folate;proton	Folate;proton
ATPS4	ATP synthase (four protons for one ATP)	*Odoribacter splanchnicus*	2.054	2.742	3.394	2.755	-0.49	0.005	ADP;proton;hydrogenphosphate	ATP;Water;proton
URAt2r	uracil transport in via proton symport reversible	*Faecalibacterium prausnitzii*	4.798	2.25	5.635	0.914	-0.49	0.003	proton;uracil	proton;uracil

METt2r	L-methionine reversible transport via proton symport	*Odoribacter splanchnicus*	1.249	1.672	2.08	1.72	-0.49	0.005	proton;L-methionine	proton;L-methionine
PNTOt4	Pantothenate sodium symporter	*Faecalibacterium prausnitzii*	2.873	1.282	3.362	0.589	-0.49	0.003	Sodium;(R)-Pantothenate	Sodium;(R)-Pantothenate
DGSNt2	deoxyguanosine transport in via proton symport	*Prevotella copri*	0.088	0.663	0.579	1.24	-0.49	0.007	Deoxyguanosine;proton	Deoxyguanosine;proton
ARGt2r	L-arganine reversible transport via proton symport	*Odoribacter splanchnicus*	1.351	1.808	2.256	1.855	-0.49	0.005	L-argininium(1+);proton	L-argininium(1+);proton
TTDCAtr	Tetradecanoate (Transport)	*Barnesiella intestinihominis*	1.561	1.561	2.278	1.338	-0.49	0.004	tetradecanoate (n-C14:0)	tetradecanoate (n-C14:0)
GLYC3Pt	glycerol-3-phosphate transport in via proton symport	*Odoribacter splanchnicus*	1.483	1.983	2.48	2.035	-0.5	0.005	Glycerol 3-phosphate;proton	Glycerol 3-phosphate;proton
DEXTRINabc	Dextrin uptake via ABC transporter	*Ruminococcus lactaris*	0.189	0.599	0.563	0.884	-0.5	0.006	ADP;Dextrin;proton;hydrogenphosphate	Dextrin;ATP;Water
MGt5	magnesium transport in/out via permease (no H+)	*Odoribacter splanchnicus*	0.676	0.909	1.14	0.957	-0.5	0.005	magnesium	magnesium

GLYTYRabc	Glycyl-L-tyrosine transport via ABC system	*Bacteroides massiliensis*	0.046	0.606	0.56	1.33	-0.5	0.007	ATP;Glycyl-L-tyrosine;Water	ADP;Glycyl-L-tyrosine;proton;hydrogenphosphate
FDNADOX_H	Ferredoxin:NAD oxidoreductase, generating a proton motive force	*Odoribacter splanchnicus*	1.96	2.622	3.286	2.692	-0.5	0.004	Reduced ferredoxin;Nicotinamide adenine dinucleotide	Oxidized ferredoxin;proton;Nicotinamide adenine dinucleotide - reduced
GTHOXti3	oxidized glutathione irreversible uniport	*Barnesiella intestinihominis*	0.953	0.977	1.408	0.833	-0.5	0.004	Oxidized glutathione	Oxidized glutathione
TYRt2r	L-tyrosine reversible transport via proton symport	*Barnesiella intestinihominis*	1.896	1.903	2.805	1.672	-0.51	0.003	proton;L-tyrosine	proton;L-tyrosine
BGALASE_OGLYCAN90e	Beta-galactosidase, mucin-type O-glycans, extracellular	*Barnesiella intestinihominis*	0.766	1.395	1.556	1.697	-0.51	0.004	released mucin-type O-glycan No 12;Water	Beta-1,3-N-acetylglucosamine-N-acetylgalactosamine;D-Galactose
CBL1abc	Cob(1)alamin transport via ABC system	*Barnesiella intestinihominis*	0.675	1.372	1.408	1.503	-0.51	0.004	ADP;Cob(I)alamin;proton;hydrogen phosphate	Cob(I)alamin;ATP;Water
ADOCBLabc	Adenosylcobalamin transport via ABC system	*Barnesiella intestinihominis*	0.676	1.372	1.409	1.502	-0.51	0.004	Adenosylcobalamin;ATP;Water	Adenosylcobalamin;ADP;proton;hydrogenphosphate

GCHOLBHSe	Glycocholate hydrolysis by bile salt hydrolase (EC 3.5.1.24)	*Alistipes putredinis*	0.061	0.593	0.385	0.668	-0.51	0.004	Cholate;Glycine;proton	glycocholate;Water
TRPt2r	L-tryptophan reversible transport via proton symport	*Odoribacter splanchnicus*	1.135	1.552	1.975	1.718	-0.51	0.004	proton;L-tryptophan	proton;L-tryptophan
CAt4i	Calcium transport via proton antiport, irreversible	*Barnesiella intestinihominis*	1.105	1.121	1.639	0.948	-0.51	0.003	calcium(2+);proton	calcium(2+);proton
FE2abc	iron (II) transport via ABC system	*Barnesiella intestinihominis*	1.105	1.121	1.639	0.948	-0.51	0.003	ATP;Fe2+;Water	ADP;Fe2+;proton;hydrogenphosphate
FE3abc	iron (III) transport via ABC system	*Barnesiella intestinihominis*	1.105	1.121	1.639	0.948	-0.51	0.003	ATP;Fe3+;Water	ADP;Fe3+;proton;hydrogenphosphate
HEMEti	Heme transport via ABC system	*Barnesiella intestinihominis*	1.105	1.121	1.639	0.948	-0.51	0.003	ATP;Water;Protoheme	ADP;proton;Protoheme;hydrogenphosphate
MALTpts	maltose transport via PEP:Pyr PTS	*Coprococcus comes*	0.683	1.276	1.333	1.248	-0.51	0.003	Maltose;Phosphoenolpyruvate	Maltose 6-phosphate;pyruvate
MNabc	manganese transport via ABC system	*Barnesiella intestinihominis*	1.105	1.121	1.639	0.948	-0.51	0.003	ATP;Water;Mn2+	ADP;proton;Mn2+;hydrogenphosphate

r2137	Major Facilitator(MFS) TCDB:2.A.1.14.6	*Barnesiella intestinihominis*	1.105	1.121	1.639	0.948	-0.51	0.003	Chloride	Chloride
SHEMEabc	Siroheme transport via ABC system	*Barnesiella intestinihominis*	1.105	1.121	1.639	0.948	-0.51	0.003	ATP;Water;Siroheme	ADP;proton;hydrogenphosphate;Siroheme
SO4t2	sulfate transport in via proton symport	*Barnesiella intestinihominis*	1.105	1.121	1.639	0.948	-0.51	0.003	proton;sulfate	proton;sulfate
SPMDabc	spermidine transport via ABC system	*Barnesiella intestinihominis*	1.105	1.121	1.639	0.948	-0.51	0.003	ATP;Water;Spermidine	ADP;proton;hydrogenphosphate;Spermidine
THMabc	thiamine transport via ABC system	*Barnesiella intestinihominis*	1.105	1.121	1.639	0.948	-0.51	0.003	ATP;Water;Thiamin	ADP;proton;hydrogenphosphate;Thiamin
COBALTt5	cobalt transport in/out via permease (no H+)	*Barnesiella intestinihominis*	1.105	1.121	1.639	0.948	-0.51	0.003	Co2+	Co2+
PIt7ir	PIt7ir	*Roseburia hominis*	0.032	0.479	0.312	0.599	-0.52	0.004	Sodium;hydrogenphosphate	Sodium;hydrogenphosphate
HDCAt2	Hexadecanoate transport via proton symport	*Odoribacter splanchnicus*	0.281	0.427	0.547	0.588	-0.52	0.004	proton;Hexadecanoate (n-C16:0)	proton;Hexadecanoate (n-C16:0)

BHEXNACASE_OGLYCAN44e	Beta-N-acetylhexosaminidase, mucin-type O-glycans, extracellular	*Barnesiella intestinihominis*	0.56	1.113	1.196	1.331	-0.52	0.004	GlcNAc-alpha-1,4-Core 2;Water	N-acetyl-D-glucosamine;GlcNAc-alpha-1,4-Core 1
OCDCAt2	Octadecanoate transport via proton symport	*Odoribacter splanchnicus*	0.359	0.494	0.663	0.661	-0.52	0.004	proton;octadecanoate (n-C18:0)	proton;octadecanoate (n-C18:0)
RIBFLVt2r	riboflavin transport in via proton symport reversible	*Barnesiella intestinihominis*	1.257	1.267	1.869	1.065	-0.52	0.002	proton;Riboflavin	proton;Riboflavin
FORt2r	formate transport in via proton symport	*Odoribacter splanchnicus*	1.166	1.655	2.172	2.156	-0.52	0.003	Formate;proton	Formate;proton
26DAPt2r	meso-2,6-Diaminoheptanedioate transport via proton symport, reversible	*Barnesiella intestinihominis*	2.143	2.172	3.197	1.833	-0.53	0.002	meso-2,6-Diaminoheptanedioate;proton	meso-2,6-Diaminoheptanedioate;proton
MGt5	magnesium transport in/out via permease (no H+)	*Barnesiella intestinihominis*	1.092	1.118	1.639	0.947	-0.53	0.002	magnesium	magnesium
FOLt	FOLt	*Barnesiella intestinihominis*	1.348	1.354	2.008	1.135	-0.53	0.002	Folate;proton	Folate;proton

NACt2r	Nicotinic acid transport by proton symport reversible	*Odoribacter splanchnicus*	0.763	1.025	1.328	1.1	-0.53	0.003	proton;Nicotinate	proton;Nicotinate
ATPS4	ATP synthase (four protons for one ATP)	*Barnesiella intestinihominis*	2.852	2.935	4.309	2.5	-0.54	0.002	ADP;proton;hydrogenphosphate	ATP;Water;proton
BHEXNACASE_OGLYCAN45e	Beta-N-acetylhexosaminidase, mucin-type O-glycans, extracellular	*Barnesiella intestinihominis*	0.981	1.752	2	2.036	-0.54	0.002	released mucin-type O-glycan No 137;Water	released mucin-type O-glycan No 195;N-acetyl-D-glucosamine
ACGAMtr2	acgam[c] transport via proton symport	*Barnesiella intestinihominis*	1.964	2.32	3.147	2.042	-0.54	0.002	N-acetyl-D-glucosamine;proton	N-acetyl-D-glucosamine;proton
TRPt2r	L-tryptophan reversible transport via proton symport	*Barnesiella intestinihominis*	1.709	1.702	2.557	1.418	-0.54	0.002	proton;L-tryptophan	proton;L-tryptophan
CYSt2r	L-cysteine reversible transport via proton symport	*Barnesiella intestinihominis*	1.832	1.821	2.744	1.516	-0.55	0.002	L-cysteine;proton	L-cysteine;proton

HISt2r	L-histidine reversible transport via proton symport	*Barnesiella intestinihominis*	1.827	1.815	2.737	1.509	-0.55	0.002	proton;L-histidine	proton;L-histidine
H2St	h2s transport (diffusion)	*Barnesiella intestinihominis*	1.936	1.922	2.902	1.597	-0.55	0.002	Hydrogen sulfide	Hydrogen sulfide
Kt1r	potassium transport via uniport reversible (facilitated diffusion)	*Barnesiella intestinihominis*	1.214	1.298	1.865	1.062	-0.55	0.002	potassium	potassium
ACt2r	Acetate Reversible Transport via Proton Symport	*Barnesiella intestinihominis*	2.561	2.546	3.85	2.123	-0.55	0.002	acetate;proton	acetate;proton
SUCCt2r	succinate transport via proton symport	*Barnesiella intestinihominis*	2.944	2.923	4.435	2.436	-0.55	0.001	proton;Succinate	proton;Succinate
GSNt2r	Guanosine Transport in via Proton Symport Reversible	*Barnesiella intestinihominis*	1.272	1.658	2.168	1.556	-0.56	0.001	Guanosine;proton	Guanosine;proton
NADH6	NADH dehydrogenase (ubiquinone-8 & 3.5 protons)	*Barnesiella intestinihominis*	3.027	2.987	4.555	2.472	-0.56	0.001	proton;Nicotinamide adenine dinucleotide - reduced;Ubiquinone-8	proton;Nicotinamide adenine dinucleotide;Ubiquinol-8

NACt2r	Nicotinic acid transport by proton symport reversible	*Barnesiella intestinihominis*	1.272	1.269	1.93	1.073	-0.56	0.001	proton;Nicotinate	proton;Nicotinate
GLYC3Pt	glycerol-3-phosphate transport in via proton symport	*Barnesiella intestinihominis*	2.981	2.943	4.505	2.447	-0.56	0.001	Glycerol 3-phosphate;proton	Glycerol 3-phosphate;proton
PIt6b	Transport of Orthophosphate	*Barnesiella intestinihominis*	2.927	2.893	4.425	2.409	-0.56	0.001	proton;hydrogenphosphate	proton;hydrogenphosphate
CYTDt2r	Cytidine Transport in via Proton Symport, Reversible	*Barnesiella intestinihominis*	2.083	2.254	3.247	1.827	-0.57	0.001	Cytidine;proton	Cytidine;proton
ASNt2r	L-asparagine reversible transport via proton symport	*Barnesiella intestinihominis*	2.204	2.338	3.433	1.94	-0.57	0.001	L-asparagine;proton	L-asparagine;proton
MANTRabc	mannotriose transport via ABC transporter	*Barnesiella intestinihominis*	1.489	1.644	2.379	1.454	-0.57	0.001	ATP;Water;mannotriose (beta-1,4)	ADP;proton;mannotriose (beta-1,4);hydrogenphosphate

GLCNACASE_OGLYCAN2e	N-acetylglucosaminidase, mucin-type O-glycans, extracellular	*Barnesiella intestinihominis*	0.652	1.135	1.36	1.323	-0.57	0.001	GlcNAc-alpha-1,4-Core 2;Water	N-acetyl-D-glucosamine;Core 2
NACUP	Uptake of Nicotinic Acid	*Faecalibacterium prausnitzii*	2.769	1.326	3.363	0.586	-0.58	< 0.001	Nicotinate	Nicotinate
ADOCBLabc	Adenosylcobalamin transport via ABC system	*Odoribacter splanchnicus*	0.517	0.862	1.057	0.978	-0.59	0.001	Adenosylcobalamin;ATP;Water	Adenosylcobalamin;ADP;proton;hydrogenphosphate
CBL1abc	Cob(1)alamin transport via ABC system	*Odoribacter splanchnicus*	0.517	0.861	1.057	0.978	-0.59	0.001	ADP;Cob(I)alamin;proton;hydrogen phosphate	Cob(I)alamin;ATP;Water
H2Ot	H2O Transport via Diffusion	*Barnesiella intestinihominis*	2.791	2.867	4.336	2.349	-0.59	0.001	Water	Water
CAt4i	Calcium transport via proton antiport, irreversible	*Faecalibacterium prausnitzii*	2.5	1.233	3.067	0.557	-0.59	< 0.001	calcium(2+);proton	calcium(2+);proton
GLCNACASE_OGLYCAN4e	N-acetylglucosaminidase, mucin-type O-glycans, extracellular	*Barnesiella intestinihominis*	0.594	1.124	1.327	1.285	-0.61	0.001	GlcNAc-alpha-1,4-Core 1;Water	T antigen (core 1);N-acetyl-D-glucosamine

DGSNt2	deoxyguanosine transport in via proton symport	*Barnesiella intestinihominis*	1.487	1.603	2.39	1.309	-0.62	< 0.001	Deoxyguanosine;proton	Deoxyguanosine;proton
AGLCNACASE_OGLYCAN8e	Alpha-N-acetylglucosaminidase, mucin-type O-glycans, extracellular	*Barnesiella intestinihominis*	0.732	1.214	1.552	1.348	-0.64	< 0.001	released GlcNAc-alpha-1,4-Core 2;Water	N-acetyl-D-glucosamine;released Core 2
r1162	Active Transport	*Roseburia hominis*	0.015	0.511	0.386	0.635	-0.65	< 0.001	ADP;proton;hydrogenphosphate;sulfate	ATP;Water;sulfate
BHEXNACASE_OGLYCAN79e	Beta-N-acetylhexosaminidase, mucin-type O-glycans, extracellular	*Barnesiella intestinihominis*	0.666	1.242	1.541	1.44	-0.65	< 0.001	released mucin-type O-glycan No 39;Water	N-acetyl-D-glucosamine;Beta-1,3-Galactose-N-acetylgalactosamine
BHEXNACASE_OGLYCAN94e	Beta-N-acetylhexosaminidase, mucin-type O-glycans, extracellular	*Barnesiella intestinihominis*	0.582	1.142	1.412	1.382	-0.66	< 0.001	released GlcNAc-alpha-1,4-Core 2;Water	N-acetyl-D-glucosamine;released GlcNAc-alpha-1,4-Core 1

AGLCNACASE_OGLYCAN7e	Alpha-N-acetylglucosaminidase, mucin-type O-glycans, extracellular	*Barnesiella intestinihominis*	0.663	1.226	1.534	1.33	-0.68	< 0.001	released GlcNAc-alpha-1,4-Core 1;Water	N-acetyl-D-glucosamine;Beta-1,3-Galactose-N-acetylgalactosamine
KESTOTTRASEe	Beta-2-1 fructan degradation, extracellular	*Bacteroides caccae*	0.19	0.504	-0.197	0.647	0.667	< 0.001	Water;Kestotetraose (3 fru, 1 glc inulin-type fructo-oligosaccharide)	D-Fructose;D-glucose
HDCAt2	Hexadecanoate transport via proton symport	*Roseburia hominis*	0.07	0.356	-0.204	0.503	0.629	< 0.001	proton;Hexadecanoate (n-C16:0)	proton;Hexadecanoate (n-C16:0)
KESTOPTASEe	Beta-2-1 fructan degradation, extracellular	*Bacteroides caccae*	0.399	1.159	-0.236	1.13	0.555	0.002	Water;Kestopentaose (4 fru, 1 glc inulin-type fructo-oligosaccharide)	D-Fructose;D-glucose
ALALEUabc	L-alanyl-L-leucine transport via ABC system	*Methanobrevibacter smithii*	0.09	0.648	-0.254	0.594	0.553	0.002	L-alanyl-L-leucine;ATP;Water	ADP;L-alanyl-L-leucine;proton;hydrogenphosphate
CHTBSabc	Chitobiose transport via ABC system	*Bacteroides stercoris*	0.134	0.367	-0.071	0.397	0.534	0.002	N,N-diacetylchitobiose;ATP;Water	ADP;N,N-diacetylchitobiose;proton;hydrogenphosphate

INULINASEe	Inulin degradation, extracellular	*Bacteroides caccae*	0.034	1.336	-0.662	1.289	0.53	0.002	Water;Inulin	D-Fructose;Kestopentaose
GLUt2r	L-glutamate transport via proton symport, reversible	*Ruminococcus torques*	0.587	1.579	-0.012	0.407	0.52	0.002	L-glutamate(1-);proton	L-glutamate(1-);proton
RIBABCte	D-ribose transport out via ABC system	*Escherichia coli*	0.111	0.396	-0.065	0.324	0.487	0.005	ATP;Water;D-ribose	ADP;proton;hydrogenphosphate;D-ribose
BHEXNACASE_OGLYCAN89e	Beta-N-acetylhexosaminidase, mucin-type O-glycans, extracellular	*Roseburia inulinivorans*	0.253	0.86	-0.075	0.415	0.486	0.004	Beta-1,3-N-acetylglucosamine-N-acetylgalactosamine;Water	N-acetyl-D-galactosamine;N-acetyl-D-glucosamine
FUCASE_OGLYCAN135e	Alpha-L-fucosidase, mucin-type O-glycans, extracellular	*Bacteroides uniformis*	0.34	0.857	-0.039	0.697	0.485	0.005	released mucin-type O-glycan No 144;Water	released Core 2;L-fucose
BHEXNACASE_OGLYCAN79e	Beta-N-acetylhexosaminidase, mucin-type O-glycans, extracellular	*Roseburia inulinivorans*	0.201	0.718	-0.07	0.334	0.484	0.004	released mucin-type O-glycan No 39;Water	N-acetyl-D-glucosamine;Beta-1,3-Galactose-N-acetylgalactosamine

PYDXabc	Pyridoxal transport via ECF transporter	*Bacteroides stercoris*	0.048	0.229	-0.071	0.272	0.474	0.007	ATP;Water;Pyridoxal	ADP;proton;hydrogenphosphate;Pyridoxal
GLCAASE9e	beta-glucuronidase, hyaluronan degradation	*Bacteroides thetaiotaomicron*	0.158	0.484	-0.027	0.27	0.472	0.005	Water;hyaluronan biosynthesis, precursor 1	N-acetyl-D-glucosamine;D-glucuronate
GLCAASE8e	beta-glucuronidase, hyaluronan degradation	*Bacteroides thetaiotaomicron*	0.105	0.362	-0.023	0.136	0.469	0.004	Water;hyaluronan	D-glucuronate;hyaluronan degradation product 1
BHEXNACASE_OGLYCAN77e	Beta-N-acetylhexosaminidase, mucin-type O-glycans, extracellular	*Bacteroides vulgatus*	0.402	1.091	-0.007	0.602	0.465	0.005	released mucin-type O-glycan No 2;Water	N-acetyl-D-glucosamine;Alpha-2,6-N-acetylneuraminate-N-acetylgalactosamine
BGALASE_OGLYCAN100e	Beta-galactosidase, mucin-type O-glycans, extracellular	*Roseburia inulinivorans*	0.265	0.84	-0.101	0.747	0.461	0.008	released mucin-type O-glycan No 18;Water	released mucin-type O-glycan No 175;D-Galactose
GALNACASE_OGLYCAN1e	N-acetylgalactosaminidase, mucin-type O-glycans, extracellular	*Bacteroides vulgatus*	0.24	0.607	-0.009	0.465	0.459	0.007	Tn antigen;Water	protein-linked serine or threonine residue (O-glycosylation site);N-acetyl-D-galactosamine

BHEXNACASE_OGLYCAN90e	Beta-N-acetylhexosaminidase, mucin-type O-glycans, extracellular	*Bacteroides uniformis*	0.223	0.68	-0.06	0.552	0.458	0.008	released Core 4;Water	N-acetyl-D-glucosamine;Beta-1,3-N-acetylglucosamine-N-acetylgalactosamine
FUCASE_OGLYCAN40e	Alpha-L-fucosidase, mucin-type O-glycans, extracellular	*Bacteroides uniformis*	0.144	0.443	-0.058	0.448	0.454	0.009	mucin-type O-glycan No 182;Water	Core 6;L-fucose
BHEXNACASE_OGLYCAN79e	Beta-N-acetylhexosaminidase, mucin-type O-glycans, extracellular	*Bacteroides vulgatus*	0.381	1.134	-0.032	0.615	0.453	0.007	released mucin-type O-glycan No 39;Water	N-acetyl-D-glucosamine;Beta-1,3-Galactose-N-acetylgalactosamine
BHEXNACASE_OGLYCAN35e	Beta-N-acetylhexosaminidase, mucin-type O-glycans, extracellular	*Bacteroides vulgatus*	0.212	0.657	-0.039	0.448	0.446	0.008	mucin-type O-glycan No 39;Water	T antigen (core 1);N-acetyl-D-glucosamine
ACNAMt2r	N-acetylneuraminate proton symport, reversible	*Bacteroides eggerthii*	0.363	1.262	-0.096	0.734	0.444	0.008	N-acetylneuraminate;proton	N-acetylneuraminate;proton

ALAASPabc	L-alanyl-L-aspartate transport via ABC system	*Bacteroides stercoris*	0.336	1.026	-0.032	0.566	0.444	0.008	L-alanyl-L-aspartate;ATP;Water	ADP;L-alanyl-L-aspartate;proton;hydrogenphosphate
FUCASE_OGLYCAN114e	Alpha-L-fucosidase, mucin-type O-glycans, extracellular	*Bacteroides ovatus*	0.124	1.361	-0.391	0.943	0.439	0.01	released mucin-type O-glycan No 104;Water	released mucin-type O-glycan No 114;L-fucose
CSNt2	cytosine transport in via proton symport	*Escherichia coli*	0.288	1.067	-0.059	0.352	0.437	0.008	cytosine;proton	cytosine;proton
BGALASE_OGLYCAN110e	Beta-galactosidase, mucin-type O-glycans, extracellular	*Roseburia inulinivorans*	0.158	0.651	-0.07	0.355	0.435	0.009	released mucin-type O-glycan No 46;Water	released Core 2;D-Galactose
BHEXNACASE_OGLYCAN51e	Beta-N-acetylhexosaminidase, mucin-type O-glycans, extracellular	*Roseburia inulinivorans*	0.145	0.661	-0.074	0.273	0.434	0.009	released mucin-type O-glycan No 154;Water	N-acetyl-D-glucosamine;released F1alpha antigen
THRt2r	L-threonine reversible transport via proton symport	*Escherichia coli*	0.331	1.261	-0.089	0.591	0.427	0.01	proton;L-threonine	proton;L-threonine

RIBabc	D-ribose transport via ABC system	*Escherichia coli*	-0.093	0.398	0.061	0.288	-0.44	0.009	ATP;Water;D-ribose	ADP;proton;hydrogenphosphate;D-ribose
GLYASNabc	Glycyl-L-asparagine transport via ABC system	*Sutterella wadsworthensis*	-0.128	0.375	0.032	0.326	-0.46	0.008	ATP;Glycyl-L-asparagine;Water	ADP;Glycyl-L-asparagine;proton;hydrogenphosphate
GLYTYRabc	Glycyl-L-tyrosine transport via ABC system	*Sutterella wadsworthensis*	-0.099	0.341	0.046	0.296	-0.46	0.008	ATP;Glycyl-L-tyrosine;Water	ADP;Glycyl-L-tyrosine;proton;hydrogenphosphate
GLYPHEabc	Glycyl-L-phenylalanine transport via ABC system	*Sutterella wadsworthensis*	-0.092	0.339	0.058	0.316	-0.46	0.008	ATP;Glycylphenylalanine;Water	ADP;Glycylphenylalanine;proton;hydrogenphosphate
ALAGLYabc	L-alanylglycine transport via ABC system	*Sutterella wadsworthensis*	-0.073	0.406	0.1	0.339	-0.46	0.007	L-alanylglycine;ATP;Water	ADP;L-alanylglycine;proton;hydrogenphosphate
LCTNBabc	Lacto-N-biose transport via ABC system	*Bifidobacterium bifidum*	-0.029	0.268	0.11	0.317	-0.48	0.007	ATP;Water;Lacto-N-biose	ADP;proton;Lacto-N-biose;hydrogenphosphate
GLYPROabc	Glycyl-L-proline transport via ABC system	*Sutterella wadsworthensis*	-0.109	0.359	0.05	0.305	-0.48	0.006	ATP;Glycylproline;Water	ADP;Glycylproline;proton;hydrogenphosphate
FE2abc	iron (II) transport via ABC system	*Odoribacter splanchnicus*	-0.071	0.442	0.24	0.727	-0.52	0.004	ATP;Fe2+;Water	ADP;Fe2+;proton;hydrogenphosphate

CYSt2r	L-cysteine reversible transport via proton symport	*Bacteroides thetaiotaomicron*	-0.335	1.189	0.199	0.726	-0.54	0.001	L-cysteine;proton	L-cysteine;proton
PPIt2e	Diphosphate export via proton symport	*Bacteroides thetaiotaomicron*	-0.242	0.63	0.101	0.49	-0.61	< 0.001	proton;Diphosphate	proton;Diphosphate

Supplementary Table 2: Summary statistics of significantly different exported metabolites and their contributing species based on Welch's t-test between IBS (N=80) and Controls (N=59). Effect size was calculated using Cohen's D formula.

ID	Species	IBS (mean)	IBS (sd)	Control (mean)	Control (sd)	Effect size	pvalue	Metabolite name
uchol_e_	*Ruminococcus torques*	1.333	1.629	0.2	0.757	0.892	0	Ursocholate
HC02194_e_	*Ruminococcus torques*	1.053	1.493	0.201	0.757	0.719	0	Ursodiol
nh4_e_	*Dorea formicigenerans*	1.527	1.886	0.457	1.173	0.681	0	Ammonium
h_e_	*Ruminococcus torques*	2.581	2.418	1.141	1.986	0.651	0	proton
ac_e_	*Ruminococcus torques*	2.373	2.24	1.05	1.833	0.646	0	acetate
lac_D_e_	*Ruminococcus torques*	2.189	2.086	0.974	1.699	0.639	0	(R)-lactate
leu_L_e_	*Dorea formicigenerans*	0.427	0.957	0	0	0.63	0	L-leucine
ppa_e_	*Ruminococcus torques*	2.193	2.067	0.996	1.742	0.626	0	propionate
for_e_	*Ruminococcus torques*	2.262	2.184	1.025	1.787	0.62	0	Formate
ala_L_e_	*Ruminococcus torques*	1.774	2.003	0.697	1.427	0.619	0	L-alanine
gln_L_e_	*Ruminococcus torques*	1.193	1.775	0.298	1.022	0.618	0	L-glutamine
succ_e_	*Ruminococcus torques*	1.816	1.744	0.844	1.482	0.6	0.001	Succinate
urea_e_	*Ruminococcus torques*	1.39	1.425	0.62	1.143	0.596	0.001	Urea
val_L_e_	*Ruminococcus torques*	1.409	1.822	0.497	1.299	0.577	0.001	L-valine
met_L_e_	*Dorea formicigenerans*	1.028	1.455	0.336	0.903	0.572	0.001	L-methionine
lac_L_e_	*Ruminococcus torques*	1.638	1.947	0.671	1.43	0.566	0.001	(S)-lactate
C02528_e_	*Collinsella aerofaciens*	1.561	1.789	0.714	1.267	0.547	0.001	chenodeoxycholate
glyc_e_	*Ruminococcus torques*	1.411	1.746	0.588	1.263	0.54	0.002	Glycerol
12dhchol_e_	*Collinsella aerofaciens*	1.452	1.865	0.592	1.274	0.538	0.002	12-Dehydrocholate

pro_L_e_	*Bifidobacterium longum*	1.295	1.738	0.508	1.155	0.534	0.002	L-proline
lac_L_e_	*Dorea formicigenerans*	1.619	1.815	0.746	1.441	0.533	0.002	(S)-lactate
ac_e_	*Dorea formicigenerans*	1.883	2.089	0.876	1.691	0.53	0.002	acetate
ppi_e_	*Bifidobacterium longum*	0.384	0.923	0.032	0.181	0.53	0.001	Diphosphate
arg_L_e_	*Bacteroides ovatus*	1.162	1.701	0.423	1.049	0.523	0.002	L-argininium(1+)
gthrd_e_	*Ruminococcus torques*	1.15	1.537	0.46	1.067	0.521	0.002	Reduced glutathione
etoh_e_	*Ruminococcus torques*	1.631	1.908	0.751	1.443	0.52	0.002	ethanol
Lcystin_e_	*Ruminococcus torques*	1.276	1.545	0.566	1.164	0.519	0.002	L-cystine
nac_e_	*Ruminococcus torques*	0.85	1.128	0.341	0.811	0.518	0.002	Nicotinate
glyc_e_	*Bifidobacterium longum*	1.429	1.88	0.607	1.245	0.515	0.002	Glycerol
gly_e_	*Dorea formicigenerans*	1.405	1.582	0.682	1.263	0.505	0.003	Glycine
for_e_	*Dorea formicigenerans*	1.788	1.972	0.884	1.616	0.502	0.004	Formate
5mthf_e_	*Bifidobacterium longum*	1.27	1.505	0.607	1.135	0.497	0.004	5-Methyltetrahydrofolate
h2o_e_	*Dorea formicigenerans*	1.888	2.075	0.95	1.745	0.489	0.005	Water
leu_L_e_	*Bifidobacterium longum*	1.005	1.569	0.375	0.937	0.487	0.004	L-leucine
leu_L_e_	*Ruminococcus torques*	0.877	1.389	0.311	0.883	0.486	0.004	L-leucine
ptrc_e_	*Ruminococcus torques*	1.309	1.385	0.677	1.241	0.481	0.005	Putrescine
nac_e_	*Dorea formicigenerans*	0.779	1.109	0.323	0.794	0.472	0.005	Nicotinate
thymd_e_	*Bifidobacterium longum*	1.245	1.311	0.688	1.034	0.472	0.006	Thymidine
co2_e_	*Ruminococcus torques*	1.506	2.01	0.671	1.52	0.469	0.006	carbon dioxide
acgam_e_	*Ruminococcus torques*	1.064	1.657	0.404	1.132	0.465	0.006	N-acetyl-D-glucosamine
fru_e_	*Bifidobacterium longum*	1.924	2.108	1.026	1.769	0.462	0.007	D-Fructose
glyc_e_	*Dorea formicigenerans*	1.134	1.481	0.521	1.157	0.462	0.007	Glycerol
acald_e_	*Ruminococcus torques*	1.565	1.984	0.746	1.536	0.461	0.007	acetaldehyde
hxan_e_	*Dorea formicigenerans*	1.427	1.704	0.715	1.375	0.46	0.007	Hypoxanthine
cytd_e_	*Bifidobacterium longum*	1.444	1.944	0.672	1.441	0.451	0.008	Cytidine
succ_e_	*Bifidobacterium longum*	0.461	1.248	0.06	0.349	0.438	0.007	Succinate
dcyt_e_	*Dorea formicigenerans*	0.607	1.128	0.204	0.658	0.436	0.009	Deoxycytidine
gly_e_	*Bifidobacterium adolescentis*	0.678	1.445	0.193	0.644	0.434	0.009	Glycine
thymd_e_	*Parabacteroides distasonis*	0.434	0.908	0.118	0.485	0.434	0.009	Thymidine
ac_e_	*Barnesiella intestinihominis*	1.941	2.061	2.849	1.956	-0.452	0.009	acetate

but_e_	*Odoribacter splanchnicus*	1.347	1.83	2.206	1.929	-0.456	0.009	butyrate
h2_e_	*Odoribacter splanchnicus*	0.902	1.349	1.565	1.527	-0.46	0.009	Hydrogen
ac_e_	*Odoribacter splanchnicus*	1.437	1.937	2.357	2.034	-0.463	0.008	acetate
ser_L_e_	*Faecalibacterium prausnitzii*	3.28	2.2	4.215	1.784	-0.467	0.006	L-serine
ser_L_e_	*Barnesiella intestinihominis*	1.826	1.933	2.702	1.785	-0.471	0.007	L-serine
nh4_e_	*Odoribacter splanchnicus*	1.402	1.905	2.335	2.025	-0.475	0.007	Ammonium
h2o_e_	*Odoribacter splanchnicus*	0.891	1.534	1.697	1.837	-0.476	0.007	Water
lac_D_e_	*Faecalibacterium prausnitzii*	3.715	2.174	4.615	1.541	-0.477	0.005	(R)-lactate
nh4_e_	*Ruminococcus lactaris*	0.387	1.125	1.022	1.506	-0.477	0.008	Ammonium
ppa_e_	*Barnesiella intestinihominis*	2.033	2.093	3.011	1.864	-0.493	0.004	propionate
amylose300_e_	*Eubacterium rectale*	0.243	0.638	0.645	0.939	-0.501	0.005	amylose (n=300 repeat units, alpha-1,4-glc)
h_e_	*Odoribacter splanchnicus*	1.446	1.957	2.474	2.092	-0.508	0.004	proton
acald_e_	*Ruminococcus lactaris*	0.228	0.902	0.922	1.7	-0.51	0.005	acetaldehyde
glyc_e_	*Barnesiella intestinihominis*	1.323	1.444	2.042	1.36	-0.512	0.003	Glycerol
dad_2_e_	*Odoribacter splanchnicus*	0.581	0.851	1.053	0.983	-0.513	0.004	2-deoxyadenosine
pi_e_	*Barnesiella intestinihominis*	2.216	2.272	3.327	2.019	-0.517	0.003	hydrogenphosphate
thr_L_e_	*Lachnospira eligens*	0.616	1.202	1.487	1.922	-0.544	0.003	L-threonine
asp_L_e_	*Faecalibacterium prausnitzii*	4.554	1.961	5.405	0.919	-0.556	0.001	L-aspartate(1-)
h_e_	*Barnesiella intestinihominis*	2.497	2.497	3.83	2.162	-0.571	0.001	proton
asp_L_e_	*Coprococcus comes*	0.879	1.561	1.904	1.957	-0.579	0.001	L-aspartate(1-)
succ_e_	*Barnesiella intestinihominis*	2.272	2.285	3.513	1.993	-0.579	0.001	Succinate
arg_L_e_	*Lachnospira eligens*	1.364	1.776	2.538	2.068	-0.609	0.001	L-argininium(1+)
4abut_e_	*Odoribacter splanchnicus*	1.035	1.613	2.128	1.889	-0.622	0	4-Aminobutanoate

Supplementary Table 3: Summary statistics of significantly different imported metabolites and their contributing species based on Welch's t-test between IBS (N=80) and Controls (N=59). Effect size was calculated using Cohen's D formula.

ID	Species	IBS (mean)	IBS (sd)	Control (mean)	Control (sd)	Effect size	pvalue	Metabolite name
7ocholate_e_	*Ruminococcus torques*	1.333	1.629	0.2	0.757	0.892	0	3alpha,12alpha-Dihydroxy-7-oxo-5beta-cholanate
glyc3p_e_	*Ruminococcus torques*	1.368	1.457	0.421	0.995	0.759	0	Glycerol 3-phosphate
7dhcdchol_e_	*Ruminococcus torques*	1.053	1.493	0.201	0.757	0.719	0	7-Dehydrochenodeoxycholate
phe_L_e_	*Ruminococcus torques*	1.023	1.216	0.3	0.744	0.717	0	L-phenylalanine
fru_e_	*Ruminococcus torques*	1.887	1.86	0.764	1.413	0.68	0	D-Fructose
glu_L_e_	*Ruminococcus torques*	1.62	1.941	0.535	1.312	0.655	0	L-glutamate(1-)
ala_D_e_	*Ruminococcus torques*	1.887	2.035	0.734	1.462	0.651	0	D-alanine
glc_D_e_	*Ruminococcus torques*	1.807	1.954	0.754	1.398	0.62	0	D-glucose
so4_e_	*Dorea formicigenerans*	0.668	0.98	0.192	0.518	0.607	0	sulfate
26dap_M_e_	*Ruminococcus torques*	1.639	1.617	0.739	1.35	0.604	0.001	meso-2,6-Diaminoheptanedioate
pi_e_	*Ruminococcus torques*	1.335	1.426	0.588	1.078	0.591	0.001	hydrogenphosphate
tre_e_	*Dorea formicigenerans*	0.98	1.263	0.353	0.835	0.585	0.001	Trehalose
acnam_e_	*Ruminococcus torques*	2.009	1.946	0.956	1.67	0.581	0.001	N-acetylneuraminate
ca2_e_	*Ruminococcus torques*	0.487	0.667	0.181	0.355	0.574	0.001	calcium(2+)
cl_e_	*Ruminococcus torques*	0.487	0.667	0.181	0.355	0.574	0.001	Chloride
cu2_e_	*Ruminococcus torques*	0.487	0.667	0.181	0.355	0.574	0.001	Cu2+
fe2_e_	*Ruminococcus torques*	0.487	0.667	0.181	0.355	0.574	0.001	Fe2+
fe3_e_	*Ruminococcus torques*	0.487	0.667	0.181	0.355	0.574	0.001	Fe3+
gam_e_	*Dorea formicigenerans*	1.004	1.25	0.38	0.891	0.574	0.001	D-Glucosamine
k_e_	*Ruminococcus torques*	0.487	0.667	0.181	0.355	0.574	0.001	potassium
mg2_e_	*Ruminococcus torques*	0.487	0.667	0.181	0.355	0.574	0.001	magnesium
mn2_e_	*Ruminococcus torques*	0.487	0.667	0.181	0.355	0.574	0.001	Mn2+
so4_e_	*Ruminococcus torques*	0.487	0.667	0.181	0.355	0.574	0.001	sulfate
zn2_e_	*Ruminococcus torques*	0.487	0.667	0.181	0.355	0.574	0.001	Zinc
cobalt2_e_	*Ruminococcus torques*	0.487	0.667	0.181	0.355	0.573	0.001	Co2+
asn_L_e_	*Ruminococcus torques*	0.32	0.732	0.018	0.14	0.572	0.001	L-asparagine
man_e_	*Ruminococcus torques*	1.488	1.649	0.656	1.252	0.569	0.001	D-Mannose
cgly_e_	*Ruminococcus torques*	1.749	1.748	0.84	1.476	0.562	0.001	L-cysteinylglycine

fru_e_	*Dorea formicigenerans*	1.362	1.635	0.557	1.219	0.558	0.001	D-Fructose
pnto_R_e_	*Ruminococcus torques*	0.544	0.768	0.204	0.418	0.551	0.001	(R)-Pantothenate
3dhcdchol_e_	*Collinsella aerofaciens*	1.561	1.789	0.714	1.268	0.547	0.001	3-Dehydrochenodeoxycholate
nh4_e_	*Ruminococcus torques*	1.619	2.032	0.662	1.43	0.545	0.001	Ammonium
ile_L_e_	*Ruminococcus torques*	0.833	1.203	0.285	0.765	0.544	0.001	L-isoleucine
q8_e_	*Bacteroides thetaiotaomicron*	0.173	0.432	0.006	0.043	0.544	0.001	Ubiquinone-8
thr_L_e_	*Dorea formicigenerans*	1.07	1.581	0.378	0.927	0.534	0.002	L-threonine
cellb_e_	*Ruminococcus torques*	1.151	1.577	0.449	1.03	0.527	0.002	cellobiose
malttr_e_	*Ruminococcus torques*	0.889	1.463	0.264	0.837	0.524	0.002	maltotriose
hxan_e_	*Ruminococcus torques*	1.171	1.323	0.559	0.999	0.522	0.002	Hypoxanthine
gthox_e_	*Ruminococcus torques*	1.034	1.399	0.41	0.958	0.521	0.002	Oxidized glutathione
glcur_e_	*Bacteroides ovatus*	0.484	0.959	0.105	0.401	0.516	0.002	D-glucuronate
val_L_e_	*Dorea formicigenerans*	0.789	1.107	0.317	0.693	0.511	0.003	L-valine
gal_e_	*Ruminococcus torques*	0.661	1.293	0.139	0.663	0.509	0.002	D-Galactose
glc_D_e_	*Dorea formicigenerans*	1.132	1.445	0.48	1.113	0.506	0.003	D-glucose
pydx_e_	*Ruminococcus torques*	0.327	0.619	0.09	0.249	0.502	0.002	Pyridoxal
cit_e_	*Bacteroides thetaiotaomicron*	0.284	0.603	0.059	0.195	0.501	0.002	Citrate
pnto_R_e_	*Dorea formicigenerans*	0.319	0.432	0.135	0.292	0.499	0.003	(R)-Pantothenate
acgam_e_	*Dorea formicigenerans*	1.379	1.601	0.656	1.298	0.496	0.004	N-acetyl-D-glucosamine
amp_e_	*Dorea formicigenerans*	0.709	1.054	0.266	0.702	0.495	0.004	AMP
fe2_e_	*Dorea formicigenerans*	0.258	0.362	0.106	0.24	0.495	0.003	Fe2+
thr_L_e_	*Ruminococcus torques*	1.466	1.79	0.671	1.392	0.495	0.004	L-threonine
pro_L_e_	*Ruminococcus torques*	0.449	0.935	0.085	0.461	0.493	0.003	L-proline
mqn7_e_	*Dorea formicigenerans*	0.308	0.433	0.128	0.286	0.491	0.004	Menaquinone 7
12ppd_S_e_	*Ruminococcus torques*	1.789	1.905	0.924	1.621	0.489	0.005	(S)-propane-1,2-diol
met_L_e_	*Ruminococcus torques*	0.405	0.892	0.07	0.379	0.489	0.003	L-methionine
trp_L_e_	*Dorea formicigenerans*	0.47	0.612	0.208	0.448	0.488	0.004	L-tryptophan
ura_e_	*Ruminococcus torques*	1.183	1.604	0.514	1.085	0.488	0.004	uracil
xyl_D_e_	*Roseburia intestinalis*	0.872	1.567	0.261	0.827	0.488	0.004	aldehydo-D-xylose
asp_L_e_	*Coprococcus comes*	0.803	1.54	0.221	0.701	0.486	0.003	L-aspartate(1-)
tre_e_	*Bifidobacterium longum*	1.587	1.748	0.834	1.348	0.482	0.005	Trehalose

ura_e_	*Dorea formicigenerans*	0.803	1.11	0.35	0.738	0.481	0.005	uracil
ribflv_e_	*Dorea formicigenerans*	0.281	0.393	0.118	0.28	0.479	0.005	Riboflavin
orn_e_	*Bacteroides ovatus*	1.248	1.667	0.564	1.147	0.478	0.005	Ornithine
fuc_L_e_	*Ruminococcus torques*	1.425	1.575	0.738	1.313	0.474	0.006	L-fucose
gcald_e_	*Dorea formicigenerans*	0.537	1.023	0.142	0.583	0.474	0.005	glycolaldehyde
malt_e_	*Bifidobacterium longum*	1.992	2.037	1.113	1.659	0.473	0.006	Maltose
gua_e_	*Dorea formicigenerans*	1.436	1.706	0.717	1.367	0.466	0.007	Guanine
h2o_e_	*Ruminococcus torques*	0.71	1.562	0.142	0.764	0.462	0.005	Water
pro_L_e_	*Roseburia inulinivorans*	0.813	1.197	0.351	0.787	0.457	0.007	L-proline
cholate_e_	*Roseburia intestinalis*	1.069	1.737	0.414	1.067	0.454	0.007	Cholate
lcts_e_	*Bacteroides ovatus*	0.922	1.245	0.412	0.984	0.454	0.008	Lactose
arab_L_e_	*Dorea formicigenerans*	0.543	1.101	0.149	0.551	0.453	0.007	L-arabinose
his_L_e_	*Dorea formicigenerans*	0.53	0.689	0.256	0.506	0.453	0.008	L-histidine
malt_e_	*Dorea formicigenerans*	0.808	1.172	0.349	0.866	0.445	0.009	Maltose
3dhchol_e_	*Collinsella aerofaciens*	1.568	1.847	0.841	1.411	0.443	0.009	3-Dehydrocholate
tyr_L_e_	*Ruminococcus torques*	0.549	1.041	0.179	0.56	0.442	0.008	L-tyrosine
fol_e_	*Dorea formicigenerans*	0.317	0.453	0.143	0.335	0.438	0.01	Folate
csn_e_	*Bifidobacterium longum*	1.243	1.778	0.584	1.186	0.436	0.01	cytosine
amylopect900_e_	*Bifidobacterium adolescentis*	0.266	0.685	0.045	0.225	0.434	0.008	Amylopectin, 900 repeat units
fe3_e_	*Bacteroides thetaiotaomicron*	0.241	0.52	0.066	0.24	0.433	0.009	Fe3+
rbflvrd_e_	*Ruminococcus torques*	0.549	1.18	0.149	0.569	0.432	0.009	Reduced riboflavin
ddca_e_	*Barnesiella intestinihominis*	1.106	1.183	1.609	1.052	-0.45	0.009	laurate
phe_L_e_	*Lachnospira eligens*	0.76	1.062	1.302	1.289	-0.46	0.01	L-phenylalanine
leu_L_e_	*Odoribacter splanchnicus*	0.841	1.176	1.401	1.253	-0.46	0.009	L-leucine
hdca_e_	*Alistipes putredinis*	1.511	1.844	2.386	1.94	-0.46	0.008	Hexadecanoate (n-C16:0)
h2o_e_	*Ruminococcus lactaris*	0.546	1.344	1.267	1.733	-0.47	0.009	Water
nac_e_	*Lachnospira eligens*	0.513	0.662	0.86	0.823	-0.47	0.009	Nicotinate
glu_L_e_	*Odoribacter splanchnicus*	1.07	1.628	1.91	1.952	-0.47	0.008	L-glutamate(1-)
pydx_e_	*Lachnospira eligens*	0.423	0.54	0.717	0.709	-0.47	0.009	Pyridoxal
met_L_e_	*Odoribacter splanchnicus*	0.595	0.827	1.017	0.964	-0.47	0.008	L-methionine
his_L_e_	*Bacteroides caccae*	0.462	0.995	0.997	1.265	-0.47	0.008	L-histidine

ocdca_e_	*Barnesiella intestinihominis*	1.027	1.108	1.52	0.984	-0.47	0.006	octadecanoate (n-C18:0)
4hbz_e_	*Odoribacter splanchnicus*	0.069	0.214	0.215	0.379	-0.47	0.01	4-hydroxybenzoate
ca2_e_	*Odoribacter splanchnicus*	0.069	0.214	0.215	0.379	-0.47	0.01	calcium(2+)
cl_e_	*Odoribacter splanchnicus*	0.069	0.214	0.215	0.379	-0.47	0.01	Chloride
cobalt2_e_	*Odoribacter splanchnicus*	0.069	0.214	0.215	0.379	-0.47	0.01	Co2+
cu2_e_	*Odoribacter splanchnicus*	0.069	0.214	0.215	0.379	-0.47	0.01	Cu2+
k_e_	*Odoribacter splanchnicus*	0.069	0.214	0.215	0.379	-0.47	0.01	potassium
mg2_e_	*Odoribacter splanchnicus*	0.069	0.214	0.215	0.379	-0.47	0.01	magnesium
mn2_e_	*Odoribacter splanchnicus*	0.069	0.214	0.215	0.379	-0.47	0.01	Mn2+
pheme_e_	*Odoribacter splanchnicus*	0.069	0.214	0.215	0.379	-0.47	0.01	Protoheme
sheme_e_	*Odoribacter splanchnicus*	0.069	0.214	0.215	0.379	-0.47	0.01	Siroheme
so4_e_	*Odoribacter splanchnicus*	0.069	0.214	0.215	0.379	-0.47	0.01	sulfate
spmd_e_	*Odoribacter splanchnicus*	0.069	0.214	0.215	0.379	-0.47	0.01	Spermidine
zn2_e_	*Odoribacter splanchnicus*	0.069	0.214	0.215	0.379	-0.47	0.01	Zinc
cgly_e_	*Odoribacter splanchnicus*	0.336	0.731	0.731	0.922	-0.47	0.008	L-cysteinylglycine
12dgr180_e_	*Odoribacter splanchnicus*	0.43	0.615	0.764	0.775	-0.48	0.007	1,2-Diacyl-sn-glycerol (dioctadecanoyl, n-C18:0)
gly_e_	*Lachnospira eligens*	0.625	1.389	1.441	1.963	-0.48	0.008	Glycine
lys_L_e_	*Odoribacter splanchnicus*	1.212	1.663	2.053	1.834	-0.48	0.006	L-lysinium(1+)
h2s_e_	*Barnesiella intestinihominis*	1.165	1.25	1.74	1.136	-0.48	0.005	Hydrogen sulfide
glc_D_e_	*Odoribacter splanchnicus*	1.118	1.581	1.928	1.769	-0.48	0.006	D-glucose
xan_e_	*Alistipes shahii*	0.124	0.518	0.589	1.25	-0.49	0.009	Xanthine
nh4_e_	*Lachnospira eligens*	0.483	1.234	1.356	2.217	-0.49	0.008	Ammonium
pyr_e_	*Roseburia hominis*	0.255	0.865	0.915	1.693	-0.49	0.007	pyruvate
3mop_e_	*Faecalibacterium prausnitzii*	1.672	2.078	2.687	2.046	-0.49	0.005	3-methyl-2-oxopentanoate
adn_e_	*Odoribacter splanchnicus*	1.021	1.518	1.816	1.66	-0.5	0.005	Adenosine
man_e_	*Odoribacter splanchnicus*	0.773	1.23	1.45	1.471	-0.5	0.005	D-Mannose
pydx_e_	*Odoribacter splanchnicus*	0.062	0.207	0.215	0.38	-0.5	0.006	Pyridoxal
lys_L_e_	*Eubacterium siraeum*	0.104	0.567	0.667	1.482	-0.5	0.007	L-lysinium(1+)
ocdcea_e_	*Barnesiella intestinihominis*	1.157	1.228	1.741	1.079	-0.51	0.004	octadecenoate (n-C18:1)
dgsn_e_	*Odoribacter splanchnicus*	0.579	0.848	1.045	0.98	-0.51	0.004	Deoxyguanosine
drib_e_	*Roseburia hominis*	0.255	0.774	0.825	1.366	-0.51	0.005	2-deoxy-D-ribose

co2_e_	*Faecalibacterium prausnitzii*	3.026	2.481	4.224	2.165	-0.52	0.003	carbon dioxide
glyc3p_e_	*Barnesiella intestinihominis*	2.227	2.282	3.342	2.027	-0.52	0.003	Glycerol 3-phosphate
glyc_e_	*Odoribacter splanchnicus*	1.276	1.733	2.218	1.885	-0.52	0.003	Glycerol
acnam_e_	*Barnesiella intestinihominis*	0.173	0.478	0.518	0.789	-0.53	0.004	N-acetylneuraminate
taur_e_	*Eubacterium siraeum*	0.028	0.18	0.266	0.584	-0.55	0.003	Taurine
co2_e_	*Barnesiella intestinihominis*	1.979	2.105	3.122	1.97	-0.56	0.001	carbon dioxide
ser_L_e_	*Odoribacter splanchnicus*	0.496	0.977	1.289	1.516	-0.62	0.001	L-serine
orn_e_	*Lachnospira eligens*	1.389	1.8	2.637	2.026	-0.65	0	Ornithine
co2_e_	*Eubacterium hallii*	0.12	0.757	1.193	1.999	-0.71	0	carbon dioxide

Supplementary Table 4: List of species, their node properties along with their module memberships based on cluster analysis of node properties using iBBiG.

Node Name	Property	Module 1	Module 2	Module 3	Module 4	Node Name	Property	Module 1	Module 2	Module 3	Module 4
Akkermansia muciniphila	abundance		1			*Dialister invisus*	competition degree (inwards)			1	
Akkermansia muciniphila	competition degree (inwards)		1			*Dialister invisus*	competition degree (outwards)			1	
Alistipes onderdonkii	abundance		1		1	*Dialister invisus*	cooperation degree (inwards)			1	
Alistipes onderdonkii	competition degree (inwards)		1		1	*Dialister invisus*	cooperation degree (outwards)			1	
Alistipes onderdonkii	competition degree (outwards)		1			*Dorea formicigenerans*	abundance		1		
Alistipes onderdonkii	cooperation degree (inwards)		1		1	*Dorea formicigenerans*	competition degree (inwards)		1		
Alistipes putredinis	abundance	1			1	*Dorea formicigenerans*	competition degree (outwards)		1		
Alistipes putredinis	competition degree (inwards)	1	1		1	*Dorea longicatena*	abundance	1			
Alistipes putredinis	competition degree (outwards)	1	1		1	*Dorea longicatena*	competition degree (inwards)	1			
Alistipes putredinis	cooperation degree (inwards)	1			1	*Dorea longicatena*	competition degree (outwards)	1	1		

Alistipes putredinis	cooperation degree (outwards)	1			1	*Dorea longicatena*	cooperation degree (inwards)		1		
Alistipes shahii	abundance		1		1	*Dorea longicatena*	cooperation degree (outwards)	1	1		
Alistipes shahii	competition degree (inwards)		1		1	*Lachnospira eligens*	abundance	1		1	1
Alistipes shahii	competition degree (outwards)		1		1	*Lachnospira eligens*	competition degree (inwards)	1		1	1
Alistipes shahii	cooperation degree (inwards)		1		1	*Lachnospira eligens*	competition degree (outwards)	1	1		1
Alistipes shahii	cooperation degree (outwards)		1		1	*Lachnospira eligens*	cooperation degree (outwards)				1
Bacteroides caccae	abundance	1			1	*Eubacterium hallii*	abundance	1		1	
Bacteroides caccae	competition degree (inwards)	1			1	*Eubacterium hallii*	competition degree (inwards)	1		1	
Bacteroides caccae	competition degree (outwards)	1			1	*Eubacterium hallii*	competition degree (outwards)	1			
Bacteroides caccae	cooperation degree (outwards)	1			1	*Eubacterium hallii*	cooperation degree (inwards)	1			
Bacteroides dorei	abundance		1		1	*Eubacterium hallii*	cooperation degree (outwards)	1			
Bacteroides dorei	competition degree (inwards)		1		1	*Eubacterium rectale*	abundance	1		1	
Bacteroides dorei	competition degree (outwards)		1		1	*Eubacterium rectale*	betweenness (competition)		1		
Bacteroides dorei	cooperation degree (outwards)		1			*Eubacterium rectale*	competition degree (inwards)	1		1	
Bacteroides ovatus	abundance		1	1	1	*Eubacterium rectale*	competition degree (outwards)	1		1	
Bacteroides ovatus	competition degree (inwards)		1	1	1	*Eubacterium rectale*	cooperation degree (inwards)	1		1	
Bacteroides ovatus	competition degree (outwards)		1	1	1	*Eubacterium rectale*	cooperation degree (outwards)	1			
Bacteroides ovatus	cooperation degree (inwards)			1		*Faecalibacterium prausnitzii*	abundance	1		1	
Bacteroides uniformis	abundance	1		1		*Faecalibacterium prausnitzii*	betweenness (competition)	1		1	

Bacteroides uniformis	competition degree (inwards)	1			1	1	*Faecalibacterium prausnitzii*	competition degree (inwards)	1			1	
Bacteroides uniformis	competition degree (outwards)	1			1		*Faecalibacterium prausnitzii*	competition degree (outwards)	1			1	
Bacteroides uniformis	cooperation degree (inwards)	1			1	1	*Faecalibacterium prausnitzii*	cooperation degree (inwards)	1			1	
Bacteroides uniformis	cooperation degree (outwards)	1					*Faecalibacterium prausnitzii*	cooperation degree (outwards)	1			1	
Bacteroides vulgatus	abundance	1					*Odoribacter splanchnicus*	abundance	1				1
Bacteroides vulgatus	competition degree (inwards)	1					*Odoribacter splanchnicus*	competition degree (inwards)					1
Bacteroides vulgatus	competition degree (outwards)	1					*Odoribacter splanchnicus*	competition degree (outwards)					1
Bacteroides vulgatus	cooperation degree (inwards)	1					*Odoribacter splanchnicus*	cooperation degree (inwards)					1
Bacteroides vulgatus	cooperation degree (outwards)	1					*Odoribacter splanchnicus*	cooperation degree (outwards)					1
Barnesiella intestinihominis	abundance	1	1		1		*Parabacteroides merdae*	abundance	1				
Barnesiella intestinihominis	competition degree (inwards)	1	1		1		*Parabacteroides merdae*	competition degree (inwards)	1				
Barnesiella intestinihominis	competition degree (outwards)	1	1		1		*Parabacteroides merdae*	competition degree (outwards)		1			
Barnesiella intestinihominis	cooperation degree (inwards)	1	1		1		*Reciprocity*	competition	1			1	
Barnesiella intestinihominis	cooperation degree (outwards)	1	1		1		*Reciprocity*	cooperation	1			1	
Bifidobacterium adolescentis	abundance	1	1				*Roseburia intestinalis*	abundance				1	
Bifidobacterium adolescentis	competition degree (inwards)		1				*Roseburia intestinalis*	competition degree (inwards)				1	
Bifidobacterium adolescentis	competition degree (outwards)	1	1				*Roseburia inulinivorans*	abundance				1	
Bifidobacterium adolescentis	cooperation degree (inwards)	1	1				*Roseburia inulinivorans*	competition degree (inwards)				1	
Bifidobacterium adolescentis	cooperation degree (outwards)		1				*Ruminococcus bromii*	abundance	1	1	1	1	

Bifidobacterium longum	abundance	1	1	1	*Ruminococcus bromii*	competition degree (inwards)	1	1	1	1
Bifidobacterium longum	competition degree (inwards)		1	1	*Ruminococcus bromii*	competition degree (outwards)	1	1		1
Bifidobacterium longum	competition degree (outwards)		1		*Ruminococcus bromii*	cooperation degree (inwards)		1		1
Bifidobacterium longum	cooperation degree (inwards)		1	1	*Ruminococcus bromii*	cooperation degree (outwards)				1
Bifidobacterium longum	cooperation degree (outwards)		1		*Ruminococcus lactaris*	abundance				1
Collinsella aerofaciens	abundance	1	1		*Ruminococcus obeum*	abundance	1	1		
Collinsella aerofaciens	competition degree (inwards)	1	1		*Ruminococcus obeum*	competition degree (inwards)	1	1		
Collinsella aerofaciens	competition degree (outwards)		1		*Ruminococcus obeum*	competition degree (outwards)	1	1		
Collinsella aerofaciens	cooperation degree (inwards)		1		*Ruminococcus obeum*	cooperation degree (outwards)	1			
Collinsella aerofaciens	cooperation degree (outwards)		1		*Ruminococcus torques*	abundance		1		
Coprococcus comes	abundance	1	1		*Ruminococcus torques*	competition degree (inwards)		1		
Coprococcus comes	competition degree (inwards)	1	1		*Ruminococcus torques*	competition degree (outwards)		1		
Coprococcus comes	competition degree (outwards)	1	1		*Ruminococcus torques*	cooperation degree (outwards)		1		
Dialister invisus	abundance			1						

Supplementary Table 5: List of species, belonging to iBBiG Module 2 (mucin degraders), and their potential to degrade mucin based on number of O-glycan mediated reactions.

Species	Number of reactions	Species	Number of reactions
Alistipes onderdonkii	0	*Dorea formicigenerans*	192
Alistipes putredinis	0	*Parabacteroides merdae*	197
Bifidobacterium adolescentis	0	*Eubacterium rectale*	278
Collinsella aerofaciens	0	*Bifidobacterium longum*	427
Lachnospira eligens	0	*Ruminococcus torques*	554
Coprococcus comes	12	*Barnesiella intestinihominis*	557
Ruminococcus obeum	21	*Bacteroides ovatus*	572
Alistipes shahii	26	*Akkermansia muciniphila*	665
Dorea longicatena	38	*Bacteroides dorei*	665
Ruminococcus bromii	78		

Supplementary Table 6: Statistical summary of Dunn's tests performed to compare metabolic modulation potential of species showing different degrees of co-operation within a microbiome community.

Group 1	Group 2	Difference Z-score (Group 1- Group 2)	P.adjusted	Comparison
Controls;1	Controls;>=5	-5.375	0	Metabolite Export stimulation
Controls;2	Controls;>=5	-3.149	0.004	Metabolite Export stimulation
Controls;2	Controls;1	2.935	0.007	Metabolite Export stimulation
Controls;3	Controls;>=5	-1.693	0.097	Metabolite Export stimulation
Controls;3	Controls;1	3.537	0.002	Metabolite Export stimulation
Controls;3	Controls;2	1.239	0.186	Metabolite Export stimulation
Controls;4	Controls;>=5	-0.351	0.418	Metabolite Export stimulation
Controls;4	Controls;1	4.645	0	Metabolite Export stimulation
Controls;4	Controls;2	2.561	0.018	Metabolite Export stimulation
Controls;4	Controls;3	1.26	0.187	Metabolite Export stimulation

IBS;1	IBS;>=5	-1.069	0.229	Metabolite Export stimulation
IBS;2	IBS;>=5	-0.353	0.429	Metabolite Export stimulation
IBS;2	IBS;1	2.502	0.02	Metabolite Export stimulation
IBS;3	IBS;>=5	-0.217	0.433	Metabolite Export stimulation
IBS;3	IBS;1	1.377	0.165	Metabolite Export stimulation
IBS;3	IBS;2	0.179	0.439	Metabolite Export stimulation
IBS;4	IBS;>=5	0.557	0.371	Metabolite Export stimulation
IBS;4	IBS;1	3.49	0.002	Metabolite Export stimulation
IBS;4	IBS;2	1.952	0.064	Metabolite Export stimulation
IBS;4	IBS;3	1.179	0.199	Metabolite Export stimulation
IBS;>=5	Controls;>=5	-0.889	0.263	Metabolite Export stimulation
IBS;1	Controls;1	1.055	0.226	Metabolite Export stimulation
IBS;2	Controls;2	0.287	0.425	Metabolite Export stimulation
IBS;3	Controls;3	-0.479	0.395	Metabolite Export stimulation
IBS;4	Controls;4	-0.238	0.435	Metabolite Export stimulation
Controls;1	Controls;>=5	-3.287	0.0029	Metabolite Export inhibition
Controls;2	Controls;>=5	-0.302	0.5362	Metabolite Export inhibition
Controls;2	Controls;1	4.047	0.0003	Metabolite Export inhibition
Controls;3	Controls;>=5	-0.216	0.5488	Metabolite Export inhibition
Controls;3	Controls;1	3.184	0.0033	Metabolite Export inhibition
Controls;3	Controls;2	0.054	0.4785	Metabolite Export inhibition
Controls;4	Controls;>=5	-0.099	0.5055	Metabolite Export inhibition
Controls;4	Controls;1	2.98	0.0054	Metabolite Export inhibition
Controls;4	Controls;2	0.169	0.5267	Metabolite Export inhibition
Controls;4	Controls;3	0.105	0.5156	Metabolite Export inhibition
IBS;1	IBS;>=5	-1.723	0.1194	Metabolite Export inhibition
IBS;2	IBS;>=5	-0.418	0.5847	Metabolite Export inhibition
IBS;2	IBS;1	4.558	0.0001	Metabolite Export inhibition
IBS;3	IBS;>=5	-0.153	0.5067	Metabolite Export inhibition
IBS;3	IBS;1	2.592	0.0153	Metabolite Export inhibition

IBS;3	IBS;2	0.407	0.5495	Metabolite Export inhibition
IBS;4	IBS;>=5	0.18	0.5355	Metabolite Export inhibition
IBS;4	IBS;1	4.01	0.0003	Metabolite Export inhibition
IBS;4	IBS;2	1.258	0.2606	Metabolite Export inhibition
IBS;4	IBS;3	0.513	0.6219	Metabolite Export inhibition
IBS;>=5	Controls;>=5	0.323	0.56	Metabolite Export inhibition
IBS;1	Controls;1	0.087	0.4986	Metabolite Export inhibition
IBS;2	Controls;2	0.161	0.5164	Metabolite Export inhibition
IBS;3	Controls;3	0.414	0.5659	Metabolite Export inhibition
IBS;4	Controls;4	0.988	0.3637	Metabolite Export inhibition
Controls;1	Controls;>=5	-5.5	0	Metabolite Import stimulation
Controls;2	Controls;>=5	-2.934	0.0063	Metabolite Import stimulation
Controls;2	Controls;1	3.404	0.0021	Metabolite Import stimulation
Controls;3	Controls;>=5	-1.895	0.0726	Metabolite Import stimulation
Controls;3	Controls;1	3.417	0.0024	Metabolite Import stimulation
Controls;3	Controls;2	0.768	0.332	Metabolite Import stimulation
Controls;4	Controls;>=5	-0.728	0.328	Metabolite Import stimulation
Controls;4	Controls;1	4.31	0.0001	Metabolite Import stimulation
Controls;4	Controls;2	1.913	0.0738	Metabolite Import stimulation
Controls;4	Controls;3	1.066	0.2302	Metabolite Import stimulation
IBS;1	IBS;>=5	-1.387	0.155	Metabolite Import stimulation
IBS;2	IBS;>=5	-0.656	0.32	Metabolite Import stimulation
IBS;2	IBS;1	2.548	0.0174	Metabolite Import stimulation
IBS;3	IBS;>=5	-0.583	0.3404	Metabolite Import stimulation
IBS;3	IBS;1	1.215	0.202	Metabolite Import stimulation
IBS;3	IBS;2	-0.003	0.4986	Metabolite Import stimulation
IBS;4	IBS;>=5	0.167	0.4436	Metabolite Import stimulation
IBS;4	IBS;1	3.277	0.0026	Metabolite Import stimulation
IBS;4	IBS;2	1.716	0.102	Metabolite Import stimulation
IBS;4	IBS;3	1.172	0.2086	Metabolite Import stimulation

IBS;>=5	Controls;>=5	-0.719	0.3219	Metabolite Import stimulation
IBS;1	Controls;1	0.753	0.3278	Metabolite Import stimulation
IBS;2	Controls;2	-0.414	0.3916	Metabolite Import stimulation
IBS;3	Controls;3	-0.681	0.3186	Metabolite Import stimulation
IBS;4	Controls;4	-0.299	0.4199	Metabolite Import stimulation
Controls;1	Controls;>=5	-2.484	0.0244	Metabolite Import Inhibition
Controls;2	Controls;>=5	-0.585	0.4054	Metabolite Import Inhibition
Controls;2	Controls;1	2.563	0.0212	Metabolite Import Inhibition
Controls;3	Controls;>=5	0.227	0.499	Metabolite Import Inhibition
Controls;3	Controls;1	2.892	0.0096	Metabolite Import Inhibition
Controls;3	Controls;2	0.89	0.3822	Metabolite Import Inhibition
Controls;4	Controls;>=5	0.748	0.4092	Metabolite Import Inhibition
Controls;4	Controls;1	3.24	0.0038	Metabolite Import Inhibition
Controls;4	Controls;2	1.435	0.2002	Metabolite Import Inhibition
Controls;4	Controls;3	0.553	0.3957	Metabolite Import Inhibition
IBS;1	IBS;>=5	-1.29	0.2219	Metabolite Import Inhibition
IBS;2	IBS;>=5	-0.08	0.514	Metabolite Import Inhibition
IBS;2	IBS;1	4.234	0.0005	Metabolite Import Inhibition
IBS;3	IBS;>=5	0.153	0.494	Metabolite Import Inhibition
IBS;3	IBS;1	2.449	0.0248	Metabolite Import Inhibition
IBS;3	IBS;2	0.419	0.447	Metabolite Import Inhibition
IBS;4	IBS;>=5	0.587	0.418	Metabolite Import Inhibition
IBS;4	IBS;1	4.018	0.0003	Metabolite Import Inhibition
IBS;4	IBS;2	1.455	0.2047	Metabolite Import Inhibition
IBS;4	IBS;3	0.638	0.453	Metabolite Import Inhibition
IBS;>=5	Controls;>=5	-0.169	0.4993	Metabolite Import Inhibition
IBS;1	Controls;1	-1.41	0.1876	Metabolite Import Inhibition
IBS;2	Controls;2	-0.006	0.5088	Metabolite Import Inhibition
IBS;3	Controls;3	-0.18	0.5077	Metabolite Import Inhibition
IBS;4	Controls;4	0.075	0.4918	Metabolite Import Inhibition

Supplementary Table 7: List of species and their cluster affiliation (corresponding to supplementary figure S6) along with the number of exported metabolites influenced (stimulated/inhibited) in Controls. The features are listed according to the order in the dendrogram shown in the corresponding heatmap.

Species	Cluster ID	Color	Metabolites stimulated	Metabolites inhibited	Species	Cluster ID	Color	Metabolites stimulated	Metabolites inhibited
Lachnospira eligens	2	blue	105	99	*Bacteroides xylanisolvens*	1	red	38	48
Bacteroides vulgatus	2	blue	115	99	*Bacteroides coprocola*	1	red	22	51
Bacteroides caccae	2	blue	126	81	*Bacteroides fragilis*	1	red	24	59
Alistipes shahii	2	blue	111	85	*Bacteroides faecis*	1	red	36	65
Bacteroides uniformis	2	blue	103	121	*Streptococcus parasanguinis*	1	red	4	50
Eubacterium rectale	2	blue	98	129	*Bacteroides finegoldii*	1	red	31	33
Eubacterium hallii	2	blue	92	91	*Bacteroides coprophilus*	1	red	24	35
Barnesiella intestinihominis	2	blue	102	85	*Prevotella stercorea*	1	red	31	26
Ruminococcus obeum	2	blue	121	71	*Dialister succinatiphilus*	1	red	33	22
Dorea longicatena	2	blue	124	68	*Parabacteroides johnsonii*	1	red	19	23
Alistipes putredinis	2	blue	114	87	*Bacteroides pectinophilus*	1	red	18	25
Bifidobacterium adolescentis	2	blue	80	102	*Ruminococcus champanellensis*	1	red	16	17
Bacteroides ovatus	2	blue	58	127	*Eubacterium ramulus*	1	red	14	22
Ruminococcus torques	2	blue	65	103	*Bifidobacterium animalis*	1	red	1	1
Bifidobacterium longum	2	blue	68	74	*Anaerostipes hadrus*	1	red	3	2
Escherichia coli	2	blue	76	76	*Methanosphaera stadtmanae*	1	red	3	2
Bifidobacterium bifidum	2	blue	77	58	*Clostridium leptum*	1	red	3	5
Bacteroides eggerthii	2	blue	78	96	*Flavonifractor plautii*	1	red	8	10
Bacteroides cellulosilyticus	2	blue	51	101	*Megamonas hypermegale*	1	red	10	8
Parabacteroides goldsteinii	2	blue	53	45	*Succinatimonas hippei*	1	red	5	12
Catenibacterium mitsuokai	2	blue	43	67	*Bifidobacterium pseudocatenulatum*	1	red	6	12
Clostridium sp L2 50	2	blue	42	68	*Desulfovibrio piger*	1	red	19	5
Methanobrevibacter smithii	2	blue	49	85	*Bacteroides fluxus*	1	red	22	8
Eubacterium biforme	2	blue	75	81	*Collinsella intestinalis*	1	red	25	6
Parabacteroides merdae	2	blue	85	78	*Klebsiella pneumoniae*	1	red	27	16
Odoribacter splanchnicus	2	blue	90	77	*Clostridium hathewayi*	1	red	28	11
Coprococcus comes	2	blue	106	80	*Citrobacter freundii*	1	red	37	14
Ruminococcus bromii	2	blue	105	74	*Enterococcus hirae*	1	red	36	29
Eubacterium ventriosum	2	blue	73	40	*Paraprevotella clara*	1	red	30	22
Dorea formicigenerans	2	blue	75	43	*Bacteroides clarus*	1	red	40	21

Coprococcus catus	2	blue	103	52	*Haemophilus parainfluenzae*	1	red	31	39
Roseburia inulinivorans	2	blue	104	64	*Bacteroides salyersiae*	1	red	30	38
Streptococcus salivarius	2	blue	74	54	*Bacteroides intestinalis*	1	red	50	22
Ruminococcus lactaris	2	blue	71	85	*Alistipes indistinctus*	1	red	61	40
Roseburia intestinalis	2	blue	83	88	*Alistipes finegoldii*	1	red	31	36
Faecalibacterium prausnitzii	2	blue	139	93	*Phascolarctobacterium succinatutens*	1	red	70	34
Sutterella wadsworthensis	1	red	60	81	*Clostridium bartlettii*	1	red	43	28
Alistipes onderdonkii	1	red	78	79	*Bacteroides plebeius*	1	red	60	40
Collinsella aerofaciens	1	red	85	76	*Streptococcus thermophilus*	1	red	32	74
Akkermansia muciniphila	1	red	79	100	*Prevotella copri*	1	red	64	96
Bacteroides dorei	1	red	83	79	*Bacteroides massiliensis*	1	red	80	98
Ruminococcus callidus	1	red	39	78	*Parabacteroides distasonis*	1	red	50	89
Butyrivibrio crossotus	1	red	56	54	*Dialister invisus*	1	red	79	92
Coprococcus eutactus	1	red	42	84	*Roseburia hominis*	1	red	74	84
Bacteroides thetaiotaomicron	1	red	43	93	*Eubacterium siraeum*	1	red	66	100
Bacteroides stercoris	1	red	72	86					

Supplementary Table 8: List of exchanged metabolites by the community (exports) and their cluster memberships (corresponding to supplementary figure 6) in Controls. The features are listed according to the order in the dendrogram shown in the corresponding heatmap.

Reaction ID	Cluster ID	Color	Metabolite	Reaction ID	Cluster ID	Color	Metabolite
EX_dextran40_m	3	green	Dextran 40, 1,6-alpha-D-Glucan	EX_indole_m	1	red	Indole
EX_ispre_m	3	green	Isoprene	EX_alathr_m	1	red	L-alanyl-L-threonine
EX_meoh_m	3	green	methanol	EX_phenol_m	1	red	Phenol
EX_acgalidour_m	3	green	Dimer GalNAc-IdoA	EX_inulin_m	1	red	Chicory inulin
EX_levanttr_m	3	green	Levantriose	EX_mantr_m	1	red	mannotriose (beta-1,4)
EX_pydxn_m	3	green	Pyridoxine	EX_arabttr_m	1	red	Arabinotriose
EX_glctnb_m	3	green	Galacto-N-biose	EX_so3_m	1	red	Sulfite
EX_starch1200_m	3	green	starch n=1200 repeat units	EX_bhb_m	1	red	(R)-3-hydroxybutyrate
EX_glcmannan_m	3	green	Konjac glucomannan	EX_homogal_m	1	red	Homogalacturonan
EX_lichn_m	3	green	Lichenin from Icelandic moss	EX_stys_m	1	red	Stachyose
EX_hista_m	3	green	histamium	EX_thf_m	1	red	5,6,7,8-Tetrahydrofolate
EX_arabinogal_m	3	green	Larch arabinogalactan	EX_tym_m	1	red	tyraminium
EX_hspg_degr_13_m	3	green	Product GLCAASE_HS	EX_5mthf_m	1	red	5-Methyltetrahydrofolate
EX_Lcyst_m	3	green	L-cysteate(1-)	EX_isochol_m	1	red	Isocholate
EX_pro_L_m	3	green	L-proline	EX_apio_D_m	4	brown	D-Apiose (rare sugar in rhamnogalacturonan II)

EX_pecticgal_m	3	green	Pectic galactan (potato)	EX_acerA_m	4	brown	Aceric acid (3-C-carboxy-5-deoxy-l-xylose) (rare sugar in rhamnogalacturonan II)
EX_uchol_m	3	green	Ursocholate	EX_kdo_m	4	brown	3-Deoxy-D-manno-2-octulosonate
EX_cbl1_m	3	green	Cob(I)alamin	EX_3ddlhept_m	4	brown	3-deoxy-d-lyxo-heptulosaric acid (rare sugar in rhamnogalacturonan II)
EX_taur_m	3	green	Taurine	EX_2omfuc_m	4	brown	2-o-methylfucose (rare sugar in rhamnogalacturonan II)
EX_h2o_m	3	green	Water	EX_2omxyl_m	4	brown	2-O-methylxylose (rare sugar in rhamnogalacturonan II)
EX_for_m	3	green	Formate	EX_acald_m	2	blue	acetaldehyde
EX_hxan_m	3	green	Hypoxanthine	EX_succ_m	2	blue	Succinate
EX_Ser_Thr_m	3	green	protein-linked serine or threonine residue (O-glycosylation site)	EX_4hbz_m	2	blue	4-hydroxybenzoate
EX_icdchol_m	3	green	Isochenodeoxycholate	EX_ins_m	2	blue	Inosine
EX_amylopect900_m	3	green	Amylopectin, 900 repeat units	EX_galur_m	2	blue	D-Galacturonate
EX_hspg_degr_6_m	3	green	Product GAM2STASE_HS2, GlcN-GlcA-GlcN	EX_lac_D_m	2	blue	(R)-lactate
EX_hspg_degr_5_m	3	green	Product AS3TASE_HS1, GlcN2S-GlcA-GlcN2S	EX_13ppd_m	2	blue	Propane-1,3-diol
EX_hspg_m	3	green	heparan sulfate proteoglycan	EX_ppi_m	2	blue	Diphosphate
EX_lmn30_m	3	green	laminarin (n=30 repeat units, beta -1,3 glc)	EX_4abut_m	2	blue	4-Aminobutanoate
EX_xyl_D_m	3	green	aldehydo-D-xylose	EX_gcald_m	2	blue	glycolaldehyde
EX_2ddglcn_m	3	green	2-Dehydro-3-deoxy-D-gluconate	EX_glyclt_m	2	blue	glycolate
EX_phe_L_m	3	green	L-phenylalanine	EX_alahis_m	2	blue	L-alanyl-L-histidine
EX_lctnb_m	3	green	Lacto-N-biose	EX_alagln_m	2	blue	L-alanyl-L-glutamine
EX_ura_m	3	green	uracil	EX_alaasp_m	2	blue	L-alanyl-L-aspartate
EX_gbbtn_m	3	green	gamma-butyrobetaine	EX_alaglu_m	2	blue	L-alanyl-L-glutamate
EX_4hoxpacd_m	3	green	(4-hydroxyphenyl)acetaldehyde	EX_dextrin_m	2	blue	Dextrin
EX_dhpppn_m	3	green	3-(2,3-Dihydroxyphenyl)propanoate	EX_din_m	2	blue	Deoxyinosine
EX_idon_L_m	3	green	L-Idonate	EX_glypro_m	2	blue	Glycylproline
EX_pullulan1200_m	3	green	pullulan	EX_alagly_m	2	blue	L-alanylglycine
EX_core7_m	3	green	Core 7	EX_metala_m	2	blue	L-methionyl-L-alanine
EX_ha_deg1_m	3	green	hyaluronan degradation product 1	EX_alaleu_m	2	blue	L-alanyl-L-leucine
EX_arab_L_m	3	green	L-arabinose	EX_glyleu_m	2	blue	Glycylleucine
EX_hspg_degr_4_m	3	green	Product AS6TASE_HS1	EX_malttr_m	2	blue	maltotriose
EX_man6p_m	3	green	D-Mannose 6-phosphate	EX_rbflvrd_m	2	blue	Reduced riboflavin
EX_o2_m	3	green	O2	EX_hspg_rest_m	2	blue	Undegradable product GALASE_HS
EX_agal1b6agal_m	3	green	Beta-1,6-N,N-diacetylgalactosamine	EX_levanb_m	2	blue	Levanbiose
EX_crn_m	3	green	L-Carnitine	EX_gncore2_m	2	blue	GlcNAc-alpha-1,4-Core 2
EX_hdca_m	3	green	Hexadecanoate (n-C16:0)	EX_gncore2_rl_m	2	blue	released GlcNAc-alpha-1,4-Core 2
EX_man_m	3	green	D-Mannose	EX_agal1b3agal_m	2	blue	Beta-1,3-N,N-diacetylgalactosamine
EX_ncam_m	3	green	Nicotinamide	EX_core4_m	2	blue	Core 4
EX_sbt_D_m	3	green	D-glucitol	EX_core4_rl_m	2	blue	released Core 4
EX_arabinan101_m	3	green	Arabinan	EX_lald_L_m	2	blue	(S)-lactaldehyde

EX_hspg_degr_8_m	3	green	Product IDOURASE_HS2	EX_glc_D_m	2	blue	D-glucose
EX_malt_m	3	green	Maltose	EX_12ppd_S_m	2	blue	(S)-propane-1,2-diol
EX_met_L_m	3	green	L-methionine	EX_dsT_antigen_m	2	blue	disialyl-T antigen
EX_fuc_L_m	3	green	L-fucose	EX_dsT_antigen_rl_m	2	blue	released disialyl-T antigen
EX_hspg_degr_12_m	3	green	Product GLCNACASE_HS2	EX_adocbl_m	2	blue	Adenosylcobalamin
EX_ttdca_m	3	green	tetradecanoate (n-C14:0)	EX_acgalglcur_m	2	blue	Dimer GalNAc-GlcA
EX_xan_m	3	green	Xanthine	EX_glcur_m	2	blue	D-glucuronate
EX_glyc_m	3	green	Glycerol	EX_mal_L_m	2	blue	(S)-malate(2-)
EX_mannan_m	3	green	Alpha-mannan, yeast	EX_cspg_b_degr_m	2	blue	Dimer GalNAc4S-IdoA2S
EX_ile_L_m	3	green	L-isoleucine	EX_cspg_c_degr_m	2	blue	Dimer GalNAc6S-GlcA
EX_nac_m	3	green	Nicotinate	EX_cspg_a_degr_m	2	blue	Dimer GalNAc4S-GlcA
EX_fald_m	3	green	formaldehyde	EX_cspg_ab_rest_m	2	blue	Not degradable product CSABCPASE_A/B
EX_ha_m	3	green	hyaluronan	EX_cspg_c_rest_m	2	blue	Not degradable product CSABCPASE_C
EX_ha_pre1_m	3	green	hyaluronan biosynthesis, precursor 1	EX_glygn2_m	2	blue	glycogen, structure 2 (glycogenin-1,6-{7[1,4-Glc], 4[1,4-Glc]})
EX_glyb_m	3	green	glycine betaine	EX_chtbs_m	2	blue	N,N-diacetylchitobiose
EX_HC02194_m	3	green	Ursodiol	EX_glycys_m	2	blue	Gly-Cys
EX_fol_m	3	green	Folate	EX_MGlcn10_m	2	blue	mucin-type O-glycan No 10
EX_xylan_m	3	green	Oat spelt xylan	EX_MGlcn10_rl_m	2	blue	released mucin-type O-glycan No 10
EX_lac_L_m	3	green	(S)-lactate	EX_glymet_m	2	blue	Glycyl-L-methionine
EX_pyr_m	3	green	pyruvate	EX_tsul_m	2	blue	thiosulfate(2-)
EX_lmn2_m	3	green	Laminaribiose	EX_glygn4_m	2	blue	glycogen, structure 4 (glycogenin-1,6-{2[1,4-Glc], [1,4-Glc]})
EX_core8_m	3	green	Core 8	EX_glygn5_m	2	blue	glycogen, structure 5 (glycogenin-2[1,4-Glc])
EX_gal1a3acgal_m	3	green	Alpha-1,3-Galactose-N-acetylgalactosamine	EX_glyasp_m	2	blue	Glycyl-L-aspartate
EX_lys_L_m	3	green	L-lysinium(1+)	EX_glyasn_m	2	blue	Glycyl-L-asparagine
EX_trypta_m	3	green	tryptaminium	EX_glygln_m	2	blue	Glycyl-L-glutamine
EX_gthrd_m	3	green	Reduced glutathione	EX_glyglu_m	2	blue	Glycyl-L-glutamate
EX_thymd_m	3	green	Thymidine	EX_nh4_m	2	blue	Ammonium
EX_diact_m	3	green	Diacetyl	EX_etoh_m	2	blue	ethanol
EX_mqn8_m	3	green	Menaquinone 8	EX_duri_m	2	blue	Deoxyuridine
EX_ptrc_m	3	green	Putrescine	EX_levantttr_m	2	blue	Levantetraose
EX_Tn_antigen_m	3	green	Tn antigen	EX_glyphe_m	2	blue	Glycylphenylalanine
EX_f1a_m	3	green	F1alpha	EX_glytyr_m	2	blue	Glycyl-L-tyrosine
EX_f1a_rl_m	3	green	released F1alpha antigen	EX_dpcoa_m	2	blue	Dephospho-CoA
EX_acnam2a6agal_m	3	green	Alpha-2,6-N-acetylneuraminate-N-acetylgalactosamine	EX_hspg_degr_11_m	2	blue	Product GAM2STASE_HS3
EX_sTn_antigen_m	3	green	sialyl-Tn antigen	EX_hspg_degr_7_m	2	blue	Product IS2TASE_HS1
EX_btoh_m	3	green	butanol	EX_2hyoxplac_m	2	blue	2-Hydroxyphenylacetate
EX_Tyr_ggn_m	3	green	Tyr-194 of apo-glycogenin protein (primer for glycogen synthesis)	EX_hspg_degr_3_m	2	blue	Product IDOURASE_HS1

EX_agam1b3agal_m	3	green	Beta-1,3-N-acetylglucosamine-N-acetylgalactosamine	EX_acmana_m	2	blue	N-acetyl-D-mannosamine
EX_core3_m	3	green	Core 3	EX_val_L_m	2	blue	L-valine
EX_acgal_m	3	green	N-acetyl-D-galactosamine	EX_3mop_m	2	blue	3-methyl-2-oxopentanoate
EX_core5_m	3	green	Core 5	EX_hspg_degr_10_m	2	blue	Product AS3TASE_HS2
EX_gal1b3acgal_m	3	green	Beta-1,3-Galactose-N-acetylgalactosamine	EX_glcn_m	2	blue	D-gluconate
EX_T_antigen_m	3	green	T antigen (core 1)	EX_pi_m	2	blue	hydrogenphosphate
EX_gncore1_m	3	green	GlcNAc-alpha-1,4-Core 1	EX_csn_m	2	blue	cytosine
EX_gncore1_rl_m	3	green	released GlcNAc-alpha-1,4-Core 1	EX_pime_m	2	blue	Pimelate
EX_sT_antigen_m	3	green	sialyl-T antigen	EX_pydam_m	2	blue	Pyridoxamine
EX_sT_antigen_rl_m	3	green	released sialyl-Tn antigen	EX_h2s_m	2	blue	Hydrogen sulfide
EX_isocapr_m	1	red	Isocaproate	EX_mnl_m	2	blue	D-Mannitol
EX_isoval_m	1	red	Isovalerate, 3-Methylbutanoate	EX_spmd_m	2	blue	Spermidine
EX_fuc1p_L_m	1	red	L-Fucose 1-phosphate	EX_uri_m	2	blue	Uridine
EX_acac_m	1	red	acetoacetate	EX_acgalidour2s_m	2	blue	Dimer GalNAc-IdoA2S
EX_isobut_m	1	red	Isobutyrate, 2-Methylpropanoate	EX_cit_m	2	blue	Citrate
EX_dgchol_m	1	red	glycochenodeoxycholate	EX_arg_L_m	2	blue	L-argininium(1+)
EX_C02528_m	1	red	chenodeoxycholate	EX_gln_L_m	2	blue	L-glutamine
EX_tdchola_m	1	red	taurochenodeoxycholate	EX_Lkynr_m	2	blue	L-kynurenine
EX_12dhchol_m	1	red	12-Dehydrocholate	EX_salcn_m	2	blue	Salicin
EX_cholate_m	1	red	Cholate	EX_bglc_m	2	blue	beta-glucans
EX_gchola_m	1	red	glycocholate	EX_chor_m	2	blue	chorismate
EX_tchola_m	1	red	taurocholate	EX_no2_m	2	blue	Nitrite
EX_3dhcdchol_m	1	red	3-Dehydrochenodeoxycholate	EX_4hpro_LT_m	2	blue	trans-4-hydroxy-L-proline
EX_3dhchol_m	1	red	3-Dehydrocholate	EX_galmannan_m	2	blue	Carob galactomannan
EX_akg_m	1	red	2-Oxoglutarate	EX_xyluglc_m	2	blue	Xyluglucan
EX_oxa_m	1	red	oxalate(2-)	EX_fru_m	2	blue	D-Fructose
EX_7dhcdchol_m	1	red	7-Dehydrochenodeoxycholate	EX_cys_L_m	2	blue	L-cysteine
EX_7ocholate_m	1	red	3alpha,12alpha-Dihydroxy-7-oxo-5beta-cholanate	EX_rmn_m	2	blue	L-Rhamnose
EX_amylose300_m	1	red	amylose (n=300 repeat units, alpha-1,4-glc)	EX_no3_m	2	blue	Nitrate
EX_arabinoxyl_m	1	red	Arabinoxylan, 30 xyl_D, 18 arab_L, 3 gal, 3 glc_D	EX_gal_m	2	blue	D-Galactose
EX_arbt_m	1	red	Arbutin	EX_gthox_m	2	blue	Oxidized glutathione
EX_malthx_m	1	red	Maltohexaose	EX_n2_m	2	blue	Nitrogen
EX_h2_m	1	red	Hydrogen	EX_pect_m	2	blue	pectins
EX_tma_m	1	red	trimethylamine	EX_idour_m	2	blue	L-iduronate
EX_ade_m	1	red	Adenine	EX_hspg_degr_1_m	2	blue	Product 1 HEPARL1_e
EX_gua_m	1	red	Guanine	EX_gam26s_m	2	blue	Glucosamine 2,6-sulfate
EX_15dap_m	1	red	1,5-Diaminopentane	EX_hspg_degr_2_m	2	blue	Product 2 HEPARL1_e
EX_tmao_m	1	red	Trimethylamine N-oxide	EX_Lcystin_m	2	blue	L-cystine
EX_xtsn_m	1	red	xanthosine	EX_dtmp_m	2	blue	dTMP

EX_5mta_m	1	red	5-Methylthioadenosine	EX_strch2_m	2	blue	starch
EX_selni_m	1	red	selenite	EX_core6_m	2	blue	Core 6
EX_urea_m	1	red	Urea	EX_agam1b6agal_m	2	blue	Beta-1,6-N-acetylglucosamine-N-acetylgalactosamine
EX_ocdca_m	1	red	octadecanoate (n-C18:0)	EX_core2_m	2	blue	Core 2
EX_pac_m	1	red	phenylacetate	EX_core2_rl_m	2	blue	released Core 2
EX_ppa_m	1	red	propionate	EX_cspg_c_m	2	blue	chondroitin sulfate C (GalNAc6S-GlcA) proteoglycan
EX_ac_m	1	red	acetate	EX_cspg_a_m	2	blue	chondroitin sulfate A (GalNAc4S-GlcA) proteoglycan
EX_h_m	1	red	proton	EX_cspg_b_m	2	blue	chondroitin sulfate B / dermatan sulfate (IdoA2S-GalNAc4S) proteoglycan
EX_but_m	1	red	butyrate	EX_g6p_m	2	blue	D-Glucose 6-phosphate
EX_co2_m	1	red	carbon dioxide	EX_kesto_m	2	blue	Kestose (2 fru, 1 glc inulin-type fructo-oligosaccharide)
EX_actn_R_m	1	red	(R)-Acetoin	EX_kestottr_m	2	blue	Kestotetraose (3 fru, 1 glc inulin-type fructo-oligosaccharide)

Supplementary Table 9: List of species and their cluster affiliation (corresponding to supplementary figure S7) along with the number of imported metabolites influenced (stimulated/inhibited) in Controls. The features are listed according to the order in the dendrogram shown in the corresponding heatmap.

Species	Cluster ID	Color	Metabolites stimulated	Metabolites inhibited	Species	Cluster ID	Color	Metabolites stimulated	Metabolites inhibited
Bacteroides massiliensis	1	red	89	127	*Bifidobacterium pseudocatenulatum*	2	blue	8	20
Bacteroides cellulosilyticus	1	red	61	126	*Desulfovibrio piger*	2	blue	20	20
Bacteroides ovatus	1	red	113	116	*Megamonas hypermegale*	2	blue	11	34
Sutterella wadsworthensis	1	red	88	90	*Clostridium hathewayi*	2	blue	31	12
Alistipes onderdonkii	1	red	95	106	*Bacteroides fluxus*	2	blue	33	14
Akkermansia muciniphila	1	red	91	113	*Prevotella stercorea*	2	blue	40	49
Ruminococcus callidus	1	red	58	97	*Eubacterium ramulus*	2	blue	50	28
Bacteroides thetaiotaomicron	1	red	79	112	*Parabacteroides goldsteinii*	2	blue	56	44
Ruminococcus lactaris	1	red	72	119	*Alistipes indistinctus*	2	blue	72	52
Parabacteroides merdae	1	red	104	101	*Paraprevotella clara*	2	blue	52	41
Bifidobacterium longum	1	red	64	144	*Bacteroides clarus*	2	blue	39	52
Odoribacter splanchnicus	1	red	123	91	*Bacteroides salyersiae*	2	blue	53	53
Eubacterium siraeum	1	red	97	103	*Bacteroides finegoldii*	2	blue	31	54
Dialister invisus	1	red	125	99	*Clostridium bartlettii*	2	blue	48	46
Roseburia inulinivorans	1	red	100	92	*Alistipes finegoldii*	2	blue	47	48
Bacteroides stercoris	1	red	112	70	*Phascolarctobacterium succinatutens*	2	blue	79	52
Bacteroides dorei	1	red	113	92	*Bacteroides intestinalis*	2	blue	68	53

Roseburia intestinalis	1	red	108	104	*Catenibacterium mitsuokai*	2	blue	75	53
Lachnospira eligens	1	red	114	123	*Streptococcus salivarius*	2	blue	86	106
Alistipes shahii	1	red	127	105	*Parabacteroides johnsonii*	2	blue	25	47
Prevotella copri	1	red	109	117	*Bacteroides pectinophilus*	2	blue	26	51
Collinsella aerofaciens	1	red	100	99	*Streptococcus parasanguinis*	2	blue	24	55
Roseburia hominis	1	red	93	97	*Bacteroides xylanisolvens*	2	blue	45	70
Ruminococcus bromii	1	red	111	101	*Haemophilus parainfluenzae*	2	blue	45	55
Ruminococcus torques	1	red	98	95	*Bacteroides coprophilus*	2	blue	34	92
Bifidobacterium bifidum	1	red	86	81	*Bacteroides coprocola*	2	blue	36	91
Eubacterium biforme	1	red	98	105	*Bacteroides faecis*	2	blue	47	99
Escherichia coli	1	red	105	116	*Bacteroides fragilis*	2	blue	49	93
Bifidobacterium adolescentis	1	red	99	111	*Butyrivibrio crossotus*	2	blue	48	79
Methanobrevibacter smithii	2	blue	77	108	*Bacteroides plebeius*	2	blue	45	80
Coprococcus catus	2	blue	113	68	*Clostridium sp L2 50*	2	blue	55	90
Bacteroides eggerthii	2	blue	106	89	*Streptococcus thermophilus*	2	blue	40	109
Eubacterium ventriosum	2	blue	68	54	*Coprococcus eutactus*	2	blue	79	59
Dorea formicigenerans	2	blue	90	82	*Parabacteroides distasonis*	2	blue	69	94
Klebsiella pneumoniae	2	blue	50	31	*Ruminococcus obeum*	3	green	131	89
Dialister succinatiphilus	2	blue	29	34	*Coprococcus comes*	3	green	127	81
Enterococcus hirae	2	blue	51	26	*Dorea longicatena*	3	green	117	118
Ruminococcus champanellensis	2	blue	14	22	*Barnesiella intestinihominis*	3	green	145	100
Citrobacter freundii	2	blue	29	12	*Bacteroides uniformis*	3	green	123	134
Clostridium leptum	2	blue	4	1	*Bacteroides caccae*	3	green	165	99
Bifidobacterium animalis	2	blue	2	3	*Eubacterium rectale*	3	green	127	130
Anaerostipes hadrus	2	blue	3	4	*Eubacterium hallii*	3	green	128	116
Succinatimonas hippei	2	blue	7	9	*Alistipes putredinis*	3	green	107	134
Methanosphaera stadtmanae	2	blue	8	13	*Bacteroides vulgatus*	3	green	133	127
Collinsella intestinalis	2	blue	18	4	*Faecalibacterium prausnitzii*	3	green	160	112
Flavonifractor plautii	2	blue	16	9					

Supplementary Table 10: List of exchanged metabolites by the community (imports) and their cluster memberships (corresponding to supplementary figure 7) in Controls. The features are listed according to the order in the dendrogram shown in the corresponding heatmap.

Reaction ID	Cluster ID	Color	Metabolite	Reaction ID	Cluster ID	Color	Metabolite
EX_amylose300_m	2	blue	amylose	EX_glcmannan_m	4	brown	Konjac glucomannan
EX_arabinogal_m	2	blue	Larch arabinogalactan	EX_pecticgal_m	4	brown	Pectic galactan (potato)
EX_agam1b6agal_m	2	blue	Beta-1,6-N-acetylglucosamine-N-acetylgalactosamine	EX_glcur_m	4	brown	D-glucuronate
EX_core6_m	2	blue	Core 6	EX_glygn2_m	4	brown	glycogen, structure 2 (glycogenin-1,6-{7[1,4-Glc], 4[1,4-Glc]})
EX_core2_m	2	blue	Core 2	EX_glygn4_m	4	brown	glycogen, structure 4 (glycogenin-1,6-{2[1,4-Glc], [1,4-Glc]})
EX_core2_rl_m	2	blue	released Core 2	EX_btn_m	4	brown	Biotin
EX_agam1b3agal_m	2	blue	Beta-1,3-N-acetylglucosamine-N-acetylgalactosamine	EX_pheme_m	4	brown	Protoheme
EX_core3_m	2	blue	Core 3	EX_q8_m	4	brown	Ubiquinone-8
EX_Tn_antigen_m	2	blue	Tn antigen	EX_sheme_m	4	brown	Siroheme
EX_f1a_m	2	blue	F1alpha	EX_glyclt_m	4	brown	glycolate
EX_f1a_rl_m	2	blue	released F1alpha antigen	EX_pyr_m	4	brown	pyruvate
EX_acgal_m	2	blue	N-acetyl-D-galactosamine	EX_glyasp_m	4	brown	Glycyl-L-aspartate
EX_acnam2a6agal_m	2	blue	Alpha-2,6-N-acetylneuraminate-N-acetylgalactosamine	EX_glymet_m	4	brown	Glycyl-L-methionine
EX_sTn_antigen_m	2	blue	sialyl-Tn antigen	EX_pect_m	4	brown	pectins
EX_arabinoxyl_m	2	blue	Arabinoxylan	EX_acgalglcur_m	4	brown	Dimer GalNAc-GlcA
EX_acald_m	2	blue	acetaldehyde	EX_acgalidour_m	4	brown	Dimer GalNAc-IdoA
EX_gcald_m	2	blue	glycolaldehyde	EX_mqn7_m	4	brown	Menaquinone 7
EX_glc_D_m	2	blue	D-glucose	EX_crn_m	4	brown	L-Carnitine
EX_succ_m	2	blue	Succinate	EX_gbbtn_m	4	brown	gamma-butyrobetaine
EX_adocbl_m	2	blue	Adenosylcobalamin	EX_isobut_m	4	brown	Isobutyrate, 2-Methylpropanoate
EX_levanttr_m	2	blue	Levantriose	EX_2ddglcn_m	4	brown	2-Dehydro-3-deoxy-D-gluconate
EX_ocdca_m	2	blue	octadecanoate (n-C18:0)	EX_glyb_m	4	brown	glycine betaine
EX_fe2_m	2	blue	Fe2+	EX_dtmp_m	4	brown	dTMP
EX_levantttr_m	2	blue	Levantetraose	EX_actn_R_m	4	brown	(R)-Acetoin
EX_idour_m	2	blue	L-iduronate	EX_mal_L_m	4	brown	(S)-malate(2-)
EX_fru_m	2	blue	D-Fructose	EX_rhamnogalurII_m	3	green	Wine rhamnogalacturonan II
EX_phe_L_m	2	blue	L-phenylalanine	EX_cspg_c_m	3	green	chondroitin sulfate C (GalNAc6S-GlcA) proteoglycan
EX_trp_L_m	2	blue	L-tryptophan	EX_cspg_a_m	3	green	chondroitin sulfate A (GalNAc4S-GlcA) proteoglycan
EX_duri_m	2	blue	Deoxyuridine	EX_cspg_b_m	3	green	chondroitin sulfate B / dermatan sulfate (IdoA2S-GalNAc4S) proteoglycan
EX_met_L_m	2	blue	L-methionine	EX_chtbs_m	3	green	N,N-diacetylchitobiose

EX_ttdca_m	2	blue	tetradecanoate (n-C14:0)	EX_fuc_L_m	3	green	L-fucose
EX_cys_L_m	2	blue	L-cysteine	EX_xan_m	3	green	Xanthine
EX_lald_L_m	2	blue	(S)-lactaldehyde	EX_glctnb_m	3	green	Galacto-N-biose
EX_glyc_m	2	blue	Glycerol	EX_btoh_m	3	green	butanol
EX_hspg_m	2	blue	heparan sulfate proteoglycan	EX_lctnb_m	3	green	Lacto-N-biose
EX_val_L_m	2	blue	L-valine	EX_man_m	3	green	D-Mannose
EX_cbl1_m	2	blue	Cob(I)alamin	EX_bhb_m	3	green	(R)-3-hydroxybutyrate
EX_4abut_m	2	blue	4-Aminobutanoate	EX_levan1000_m	3	green	Levan, 1000 fructose units
EX_dpcoa_m	2	blue	Dephospho-CoA	EX_indole_m	3	green	Indole
EX_urea_m	2	blue	Urea	EX_gthrd_m	3	green	Reduced glutathione
EX_etoh_m	2	blue	ethanol	EX_xyl_D_m	3	green	aldehydo-D-xylose
EX_12ppd_S_m	2	blue	(S)-propane-1,2-diol	EX_fe3_m	3	green	Fe3+
EX_hspg_degr_15_m	2	blue	Product IDOURASE_HS3	EX_hspg_degr_14_m	3	green	Product GLCNACASE_HS3
EX_amylopect900_m	2	blue	Amylopectin	EX_hspg_degr_2_m	3	green	Product 2 HEPARL1_e
EX_pullulan1200_m	2	blue	pullulan	EX_hspg_degr_3_m	3	green	Product IDOURASE_HS1
EX_Lcystin_m	2	blue	L-cystine	EX_hspg_degr_8_m	3	green	Product IDOURASE_HS2
EX_o2_m	2	blue	O2	EX_pi_m	3	green	hydrogenphosphate
EX_strch2_m	2	blue	starch, structure 2 (1,6-{2[1,4-Glc], [1,4-Glc]})	EX_lac_D_m	3	green	(R)-lactate
EX_cspg_b_degr_m	2	blue	Dimer GalNAc4S-IdoA2S	EX_kestopt_m	3	green	Kestopentaose
EX_cspg_c_degr_m	2	blue	Dimer GalNAc6S-GlcA	EX_ins_m	3	green	Inosine
EX_cspg_a_degr_m	2	blue	Dimer GalNAc4S-GlcA	EX_pydxn_m	3	green	Pyridoxine
EX_cspg_ab_rest_m	2	blue	Not degradable product CSABCPASE_A/B	EX_apio_D_m	3	green	D-Apiose (rare sugar in rhamnogalacturonan II)
EX_gal_m	2	blue	D-Galactose	EX_acerA_m	3	green	Aceric acid (3-C-carboxy-5-deoxy-l-xylose) (rare sugar in rhamnogalacturonan II)
EX_meoh_m	2	blue	methanol	EX_kdo_m	3	green	3-Deoxy-D-manno-2-octulosonate
EX_cspg_c_rest_m	2	blue	Not degradable product CSABCPASE_C	EX_3ddlhept_m	3	green	3-deoxy-d-lyxo-heptulosaric acid (rare sugar in rhamnogalacturonan II)
EX_na1_m	2	blue	Sodium	EX_2omfuc_m	3	green	2-o-methylfucose (rare sugar in rhamnogalacturonan II)
EX_din_m	2	blue	Deoxyinosine	EX_2omxyl_m	3	green	2-O-methylxylose (rare sugar in rhamnogalacturonan II)
EX_mantr_m	2	blue	mannotriose (beta-1,4)	EX_levanb_m	3	green	Levanbiose
EX_gncore1_m	2	blue	GlcNAc-alpha-1,4-Core 1	EX_ade_m	3	green	Adenine
EX_gncore1_rl_m	2	blue	released GlcNAc-alpha-1,4-Core 1	EX_gly_m	3	green	Glycine
EX_sT_antigen_m	2	blue	sialyl-T antigen	EX_hspg_degr_4_m	3	green	Product AS6TASE_HS1
EX_sT_antigen_rl_m	2	blue	released sialyl-Tn antigen	EX_ha_m	3	green	hyaluronan
EX_galur_m	2	blue	D-Galacturonate	EX_ha_pre1_m	3	green	hyaluronan biosynthesis, precursor 1
EX_gal1b3acgal_m	2	blue	Beta-1,3-Galactose-N-acetylgalactosamine	EX_4hpro_LT_m	3	green	trans-4-hydroxy-L-proline
EX_T_antigen_m	2	blue	T antigen (core 1)	EX_gln_L_m	3	green	L-glutamine
EX_glyphe_m	2	blue	Glycylphenylalanine	EX_ppi_m	3	green	Diphosphate
EX_glytyr_m	2	blue	Glycyl-L-tyrosine	EX_hspg_degr_1_m	3	green	Product 1 HEPARL1_e
EX_alaleu_m	2	blue	L-alanyl-L-leucine	EX_hspg_degr_11_m	3	green	Product GAM2STASE_HS3

EX_glyleu_m	2	blue	Glycylleucine	EX_hspg_degr_6_m	3	green	Product GAM2STASE_HS2
EX_alagly_m	2	blue	L-alanylglycine	EX_idon_L_m	3	green	L-Idonate
EX_alathr_m	2	blue	L-alanyl-L-threonine	EX_ala_L_m	3	green	L-alanine
EX_n2_m	2	blue	Nitrogen	EX_sbt_D_m	3	green	D-glucitol
EX_lac_L_m	2	blue	(S)-lactate	EX_hspg_degr_12_m	3	green	Product GLCNACASE_HS2
EX_core5_m	2	blue	Core 5	EX_xtsn_m	3	green	xanthosine
EX_glycys_m	2	blue	Gly-Cys	EX_isocapr_m	3	green	Isocaproate
EX_metala_m	2	blue	L-methionyl-L-alanine	EX_isoval_m	3	green	Isovalerate, 3-Methylbutanoate
EX_glyasn_m	2	blue	Glycyl-L-asparagine	EX_mqn8_m	3	green	Menaquinone 8
EX_glygln_m	2	blue	Glycyl-L-glutamine	EX_so4_m	3	green	sulfate
EX_glyglu_m	2	blue	Glycyl-L-glutamate	EX_arab_L_m	3	green	L-arabinose
EX_alaasp_m	2	blue	L-alanyl-L-aspartate	EX_arbt_m	3	green	Arbutin
EX_alagln_m	2	blue	L-alanyl-L-glutamine	EX_raffin_m	3	green	Raffinose
EX_alaglu_m	2	blue	L-alanyl-L-glutamate	EX_sucr_m	3	green	Sucrose
EX_alahis_m	2	blue	L-alanyl-L-histidine	EX_glu_L_m	3	green	L-glutamate(1-)
EX_glypro_m	2	blue	Glycylproline	EX_melib_m	3	green	Melibiose
EX_pydam_m	2	blue	Pyridoxamine	EX_fum_m	3	green	Fumarate
EX_xylan_m	2	blue	Oat spelt xylan	EX_Lcyst_m	3	green	L-cysteate(1-)
EX_gchola_m	1	red	glycocholate	EX_galt_m	3	green	galactitol
EX_tchola_m	1	red	taurocholate	EX_tym_m	3	green	tyraminium
EX_12dhchol_m	1	red	12-Dehydrocholate	EX_pime_m	3	green	Pimelate
EX_cholate_m	1	red	Cholate	EX_uri_m	3	green	Uridine
EX_3dhcdchol_m	1	red	3-Dehydrochenodeoxycholate	EX_bglc_m	3	green	beta-glucans
EX_3dhchol_m	1	red	3-Dehydrocholate	EX_orn_m	3	green	Ornithine
EX_dextrin_m	1	red	Dextrin	EX_acgalidour2s_m	3	green	Dimer GalNAc-IdoA2S
EX_malthx_m	1	red	Maltohexaose	EX_hspg_degr_10_m	3	green	Product AS3TASE_HS2
EX_arabttr_m	1	red	Arabinotriose	EX_no3_m	3	green	Nitrate
EX_tsul_m	1	red	thiosulfate(2-)	EX_ispre_m	3	green	Isoprene
EX_12dgr180_m	1	red	1,2-Diacyl-sn-glycerol (dioctadecanoyl, n-C18:0)	EX_no2_m	3	green	Nitrite
EX_lys_L_m	1	red	L-lysinium(1+)	EX_core7_m	3	green	Core 7
EX_lmn2_m	1	red	Laminaribiose	EX_pro_L_m	3	green	L-proline
EX_arabinan101_m	1	red	Arabinan	EX_hxan_m	3	green	Hypoxanthine
EX_lmn30_m	1	red	laminarin (n=30 repeat units, beta -1,3 glc)	EX_ura_m	3	green	uracil
EX_lichn_m	1	red	Lichenin from Icelandic moss	EX_3hpppn_m	3	green	3-(3-hydroxy-phenyl)propionate
EX_dextran40_m	1	red	Dextran 40, 1,6-alpha-D-Glucan	EX_ethso3_m	3	green	ethanesulfonate
EX_dgchol_m	1	red	glycochenodeoxycholate	EX_2hyoxplac_m	3	green	2-Hydroxyphenylacetate
EX_C02528_m	1	red	chenodeoxycholate	EX_tre_m	3	green	Trehalose
EX_tdchola_m	1	red	taurochenodeoxycholate	EX_lcts_m	3	green	Lactose
EX_HC02194_m	1	red	Ursodiol	EX_salcn_m	3	green	Salicin

EX_uchol_m	1	red	Ursocholate	EX_agal1b6agal_m	3	green	Beta-1,6-N,N-diacetylgalactosamine
EX_xyluglc_m	4	brown	Xyluglucan	EX_asp_L_m	3	green	L-aspartate(1-)
EX_galmannan_m	4	brown	Carob galactomannan	EX_fuc1p_L_m	3	green	L-Fucose 1-phosphate
EX_rhamnogalurI_m	4	brown	Potato rhamnogalacturonan I	EX_man6p_m	3	green	D-Mannose 6-phosphate
EX_homogal_m	4	brown	Homogalacturonan	EX_thm_m	3	green	Thiamin
EX_g6p_m	4	brown	D-Glucose 6-phosphate	EX_for_m	3	green	Formate
EX_ile_L_m	4	brown	L-isoleucine	EX_his_L_m	3	green	L-histidine
EX_glygn5_m	4	brown	glycogen, structure 5 (glycogenin-2[1,4-Glc])	EX_so3_m	3	green	Sulfite
EX_h2o_m	4	brown	Water	EX_15dap_m	3	green	1,5-Diaminopentane
EX_hspg_degr_7_m	4	brown	Product IS2TASE_HS1	EX_dcyt_m	3	green	Deoxycytidine
EX_nh4_m	4	brown	Ammonium	EX_tyr_L_m	3	green	L-tyrosine
EX_icdchol_m	4	brown	Isochenodeoxycholate	EX_ha_deg1_m	3	green	hyaluronan degradation product 1
EX_isochol_m	4	brown	Isocholate	EX_hspg_degr_13_m	3	green	Product GLCAASE_HS
EX_chor_m	4	brown	chorismate	EX_glcn_m	3	green	D-gluconate
EX_cit_m	4	brown	Citrate	EX_metsox_S_L_m	3	green	L-Methionine Sulfoxide
EX_gua_m	4	brown	Guanine	EX_hspg_degr_9_m	3	green	Product AS6TASE_HS2
EX_Tyr_ggn_m	4	brown	Tyr-194 of apo-glycogenin protein (primer for glycogen synthesis)	EX_gam26s_m	3	green	Glucosamine 2,6-sulfate
EX_chol_m	4	brown	Choline	EX_5mta_m	3	green	5-Methylthioadenosine
EX_spmd_m	4	brown	Spermidine	EX_arg_L_m	3	green	L-argininium(1+)
EX_acmana_m	4	brown	N-acetyl-D-mannosamine	EX_hspg_degr_5_m	3	green	Product AS3TASE_HS1
EX_pydx_m	4	brown	Pyridoxal	EX_leu_L_m	3	green	L-leucine
EX_2dmmq8_m	4	brown	2-Demethylmenaquinone 8	EX_4abz_m	3	green	4-Aminobenzoate
EX_pnto_R_m	4	brown	(R)-Pantothenate	EX_ribflv_m	3	green	Riboflavin
EX_acac_m	4	brown	acetoacetate	EX_malt_m	3	green	Maltose
EX_2obut_m	4	brown	2-Oxobutanoate	EX_gthox_m	3	green	Oxidized glutathione
EX_but_m	4	brown	butyrate	EX_ptrc_m	3	green	Putrescine
EX_akg_m	4	brown	2-Oxoglutarate	EX_mannan_m	3	green	Alpha-mannan, yeast
EX_dgsn_m	4	brown	Deoxyguanosine	EX_3mop_m	3	green	3-methyl-2-oxopentanoate
EX_mnl_m	4	brown	D-Mannitol	EX_4hbz_m	3	green	4-hydroxybenzoate
EX_lanost_m	4	brown	lanosterol	EX_rmn_m	3	green	L-Rhamnose
EX_ocdcea_m	4	brown	octadecenoate (n-C18:1)	EX_thymd_m	3	green	Thymidine
EX_amp_m	4	brown	AMP	EX_gncore2_m	3	green	GlcNAc-alpha-1,4-Core 2
EX_rib_D_m	4	brown	D-ribose	EX_gncore2_rl_m	3	green	released GlcNAc-alpha-1,4-Core 2
EX_gsn_m	4	brown	Guanosine	EX_agal1b3agal_m	3	green	Beta-1,3-N,N-diacetylgalactosamine
EX_sel_m	4	brown	selenate	EX_core4_m	3	green	Core 4
EX_strch1_m	4	brown	starch, structure 1 (1,6-{7[1,4-Glc], 4[1,4-Glc]})	EX_core4_rl_m	3	green	released Core 4
EX_26dap_M_m	4	brown	meso-2,6-Diaminoheptanedioate	EX_kesto_m	3	green	Kestose
EX_thf_m	4	brown	5,6,7,8-Tetrahydrofolate	EX_kestottr_m	3	green	Kestotetraose
EX_ala_D_m	4	brown	D-alanine	EX_MGlcn10_m	3	green	mucin-type O-glycan No 10

EX_csn_m	4	brown	cytosine	EX_MGlcn10_rl_m	3	green	released mucin-type O-glycan No 10
EX_fol_m	4	brown	Folate	EX_dsT_antigen_m	3	green	disialyl-T antigen
EX_ncam_m	4	brown	Nicotinamide	EX_dsT_antigen_rl_m	3	green	released disialyl-T antigen
EX_arab_D_m	4	brown	D-Arabinose	EX_malttr_m	3	green	maltotriose
EX_ddca_m	4	brown	laurate	EX_inulin_m	3	green	Chicory inulin
EX_nmn_m	4	brown	NMN	EX_stys_m	3	green	Stachyose
EX_dad_2_m	4	brown	2-deoxyadenosine	EX_core8_m	3	green	Core 8
EX_drib_m	4	brown	2-deoxy-D-ribose	EX_gal1a3acgal_m	3	green	Alpha-1,3-Galactose-N-acetylgalactosamine
EX_starch1200_m	4	brown	starch	EX_7ocholate_m	3	green	3alpha,12alpha-Dihydroxy-7-oxo-5beta-cholanate
EX_hdca_m	4	brown	Hexadecanoate (n-C16:0)	EX_7dhcdchol_m	3	green	7-Dehydrochenodeoxycholate
EX_Ser_Thr_m	4	brown	protein-linked serine or threonine residue (O-glycosylation site)	EX_taur_m	3	green	Taurine

Supplementary Table 11: List of species and their cluster affiliation (corresponding to supplementary figure 8) along with the number of exported metabolites influenced (stimulated/inhibited) in IBS. The features are listed according to the order in the dendrogram shown in the corresponding heatmap.

Species	Cluster ID	Color	Metabolites stimulated	Metabolites inhibited	Species	Cluster ID	Color	Metabolites stimulated	Metabolites inhibited
Dialister invisus	2	blue	93	85	*Lactobacillus antri*	1	red	2	1
Alistipes onderdonkii	2	blue	106	69	*Lactobacillus ruminis*	1	red	6	4
Ruminococcus torques	2	blue	100	91	*Klebsiella pneumoniae*	1	red	7	3
Barnesiella intestinihominis	2	blue	65	87	*Methanosphaera stadtmanae*	1	red	8	6
Alistipes shahii	2	blue	89	48	*Bifidobacterium catenulatum*	1	red	11	4
Coprococcus comes	2	blue	88	58	*Blautia producta*	1	red	8	8
Bacteroides caccae	2	blue	80	95	*Leuconostoc lactis*	1	red	8	12
Ruminococcus obeum	2	blue	109	60	*Lactobacillus acidophilus*	1	red	5	17
Roseburia inulinivorans	2	blue	78	82	*Megamonas hypermegale*	1	red	9	11
Dorea longicatena	2	blue	97	82	*Lactobacillus crispatus*	1	red	11	10
Collinsella aerofaciens	2	blue	98	85	*Prevotella stercorea*	1	red	4	12
Bifidobacterium longum	2	blue	104	73	*Bacteroides nordii*	1	red	10	13
Bifidobacterium adolescentis	2	blue	79	101	*Citrobacter freundii*	1	red	17	9
Bacteroides dorei	2	blue	69	126	*Streptococcus vestibularis*	1	red	20	13
Roseburia intestinalis	2	blue	52	134	*Bifidobacterium dentium*	1	red	9	19
Roseburia hominis	2	blue	80	83	*Mitsuokella multacida*	1	red	34	5

Lachnospira eligens	2	blue	66	97	*Bifidobacterium animalis*	1	red	24	11
Streptococcus salivarius	2	blue	55	90	*Haemophilus parainfluenzae*	1	red	29	27
Eubacterium siraeum	2	blue	86	74	*Clostridium symbiosum*	1	red	31	19
Dorea formicigenerans	2	blue	64	97	*Clostridium clostridioforme*	1	red	30	15
Bacteroides thetaiotaomicron	2	blue	74	66	*Parabacteroides goldsteinii*	1	red	26	17
Bacteroides fragilis	2	blue	75	67	*Bacteroides coprophilus*	1	red	23	17
Parabacteroides merdae	2	blue	73	88	*Lactobacillus vaginalis*	1	red	52	25
Alistipes putredinis	2	blue	75	94	*Collinsella intestinalis*	1	red	3	29
Bacteroides vulgatus	2	blue	69	119	*Acidaminococcus fermentans*	1	red	1	36
Sutterella wadsworthensis	2	blue	38	65	*Ruminococcus champanellensis*	1	red	9	36
Bacteroides eggerthii	2	blue	68	87	*Bifidobacterium angulatum*	1	red	10	35
Bacteroides cellulosilyticus	2	blue	57	100	*Bacteroides faecis*	1	red	9	39
Akkermansia muciniphila	2	blue	88	73	*Bacteroides salyersiae*	1	red	33	19
Escherichia coli	2	blue	59	96	*Clostridium leptum*	1	red	47	27
Ruminococcus bromii	2	blue	55	125	*Bacteroides plebeius*	1	red	39	23
Eubacterium hallii	2	blue	88	89	*Catenibacterium mitsuokai*	1	red	33	49
Bacteroides ovatus	2	blue	83	96	*Veillonella dispar*	1	red	3	39
Faecalibacterium prausnitzii	2	blue	104	108	*Veillonella atypica*	1	red	2	40
Eubacterium rectale	2	blue	98	129	*Veillonella parvula*	1	red	6	41
Bacteroides uniformis	2	blue	105	122	*Bifidobacterium pseudocatenulatum*	1	red	7	47
Odoribacter splanchnicus	1	red	58	69	*Streptococcus thermophilus*	1	red	39	35
Bacteroides stercoris	1	red	84	63	*Bacteroides clarus*	1	red	48	50
Coprococcus catus	1	red	70	50	*Parabacteroides johnsonii*	1	red	37	35
Alistipes finegoldii	1	red	58	57	*Clostridium bartlettii*	1	red	38	32
Ruminococcus lactaris	1	red	54	35	*Streptococcus parasanguinis*	1	red	16	39
Coprococcus eutactus	1	red	68	32	*Bacteroides finegoldii*	1	red	29	49
Bacteroides xylanisolvens	1	red	75	25	*Bacteroides coprocola*	1	red	49	45
Eubacterium ventriosum	1	red	58	50	*Eubacterium ramulus*	1	red	43	31
Ruminococcus gnavus	1	red	84	39	*Bacteroides intestinalis*	1	red	54	41
Ruminococcus callidus	1	red	59	56	*Clostridium sp L2 50*	1	red	28	52
Clostridium nexile	1	red	32	19	*Clostridium bolteae*	1	red	71	36
Clostridium hathewayi	1	red	34	23	*Phascolarctobacterium succinatutens*	1	red	55	68
Lactobacillus salivarius	1	red	11	10	*Eubacterium biforme*	1	red	47	63
Flavonifractor plautii	1	red	23	4	*Methanobrevibacter smithii*	1	red	45	65
Alistipes indistinctus	1	red	19	11	*Prevotella copri*	1	red	28	89
Eubacterium dolichum	1	red	3	1	*Bifidobacterium bifidum*	1	red	49	59

Eggerthella lenta	1	red	2	3	*Bacteroides massiliensis*	1	red	56	71
Anaerostipes hadrus	1	red	2	3	*Parabacteroides distasonis*	1	red	53	87
Lactobacillus gasseri	1	red	1	2					

Supplementary Table 12: List of exchanged metabolites by the community (exports) and their cluster memberships (corresponding to supplementary figure 8) in IBS. The features are listed according to the order in the dendrogram shown in the corresponding heatmap.

Reaction ID	Cluster ID	Color	Metabolite	Reaction ID	Cluster ID	Color	Metabolite
EX_glc_D_m	4	brown	D-glucose	EX_rmn_m	3	green	L-Rhamnose
EX_Lcystin_m	4	brown	L-cystine	EX_ch4s_m	3	green	Methanethiol
EX_malttr_m	4	brown	maltotriose	EX_mannan_m	3	green	Alpha-mannan, yeast
EX_acald_m	4	brown	acetaldehyde	EX_arab_L_m	3	green	L-arabinose
EX_ins_m	4	brown	Inosine	EX_Lcyst_m	3	green	L-cysteate(1-)
EX_lac_D_m	4	brown	(R)-lactate	EX_starch1200_m	3	green	starch
EX_isobut_m	4	brown	Isobutyrate, 2-Methylpropanoate	EX_xyl_D_m	3	green	aldehydo-D-xylose
EX_lac_L_m	4	brown	(S)-lactate	EX_arabinoxyl_m	3	green	Arabinoxylan
EX_succ_m	4	brown	Succinate	EX_crn_m	3	green	L-Carnitine
EX_hdca_m	4	brown	Hexadecanoate (n-C16:0)	EX_tmao_m	3	green	Trimethylamine N-oxide
EX_thymd_m	4	brown	Thymidine	EX_idon_L_m	3	green	L-Idonate
EX_4abut_m	4	brown	4-Aminobutanoate	EX_idour_m	3	green	L-iduronate
EX_Ser_Thr_m	4	brown	protein-linked serine or threonine residue (O-glycosylation site)	EX_xan_m	3	green	Xanthine
EX_3dhchol_m	4	brown	3-Dehydrocholate	EX_hspg_degr_8_m	3	green	Product IDOURASE_HS2
EX_7ocholate_m	4	brown	3alpha,12alpha-Dihydroxy-7-oxo-5beta-cholanate	EX_pullulan1200_m	3	green	pullulan
EX_inulin_m	4	brown	Chicory inulin	EX_hspg_degr_2_m	3	green	Product 2 HEPARL1_e
EX_3dhcdchol_m	4	brown	3-Dehydrochenodeoxycholate	EX_hspg_degr_3_m	3	green	Product IDOURASE_HS1
EX_glycys_m	4	brown	Gly-Cys	EX_hspg_degr_5_m	3	green	Product AS3TASE_HS1
EX_cspg_a_degr_m	4	brown	Dimer GalNAc4S-GlcA	EX_hspg_m	3	green	heparan sulfate proteoglycan
EX_cspg_c_degr_m	4	brown	Dimer GalNAc6S-GlcA	EX_cspg_b_degr_m	3	green	Dimer GalNAc4S-IdoA2S
EX_acgalidour_m	4	brown	Dimer GalNAc-IdoA	EX_gam26s_m	3	green	Glucosamine 2,6-sulfate
EX_acac_m	4	brown	acetoacetate	EX_gbbtn_m	3	green	gamma-butyrobetaine
EX_cbl1_m	4	brown	Cob(I)alamin	EX_man_m	3	green	D-Mannose
EX_dextrin_m	4	brown	Dextrin	EX_arabinan101_m	3	green	Arabinan
EX_malthx_m	4	brown	Maltohexaose	EX_xyluglc_m	3	green	Xyluglucan
EX_core5_m	4	brown	Core 5	EX_malt_m	3	green	Maltose
EX_indole_m	4	brown	Indole	EX_homogal_m	3	green	Homogalacturonan
EX_hista_m	4	brown	histamium	EX_arg_L_m	3	green	L-argininium(1+)

EX_uchol_m	4	brown	Ursocholate	EX_rhamnogalurI_m	3	green	Potato rhamnogalacturonan I
EX_15dap_m	4	brown	1,5-Diaminopentane	EX_trp_L_m	3	green	L-tryptophan
EX_arbt_m	4	brown	Arbutin	EX_4hpro_LT_m	3	green	trans-4-hydroxy-L-proline
EX_agal1b3agal_m	1	red	Beta-1,3-N,N-diacetylgalactosamine	EX_2hyoxplac_m	3	green	2-Hydroxyphenylacetate
EX_core4_rl_m	1	red	released Core 4	EX_dhcinnm_m	3	green	2,3-dihydroxicinnamic acid
EX_gncore2_rl_m	1	red	released GlcNAc-alpha-1,4-Core 2	EX_dhpppn_m	3	green	3-(2,3-Dihydroxyphenyl)propanoate
EX_core4_m	1	red	Core 4	EX_cys_L_m	3	green	L-cysteine
EX_gncore2_m	1	red	GlcNAc-alpha-1,4-Core 2	EX_hspg_degr_9_m	3	green	Product AS6TASE_HS2
EX_oxa_m	1	red	oxalate(2-)	EX_rhamnogalurII_m	3	green	Wine rhamnogalacturonan II
EX_thf_m	1	red	5,6,7,8-Tetrahydrofolate	EX_no3_m	3	green	Nitrate
EX_levantttr_m	1	red	Levantetraose	EX_phe_L_m	3	green	L-phenylalanine
EX_glctnb_m	1	red	Galacto-N-biose	EX_salcn_m	3	green	Salicin
EX_lctnb_m	1	red	Lacto-N-biose	EX_pro_L_m	3	green	L-proline
EX_strch2_m	1	red	starch, structure 2 (1,6-{2[1,4-Glc], [1,4-Glc]})	EX_fol_m	3	green	Folate
EX_kesto_m	1	red	Kestose	EX_ade_m	3	green	Adenine
EX_kestottr_m	1	red	Kestotetraose	EX_bhb_m	3	green	(R)-3-hydroxybutyrate
EX_g6p_m	1	red	D-Glucose 6-phosphate	EX_phenol_m	3	green	Phenol
EX_core8_m	1	red	Core 8	EX_lichn_m	3	green	Lichenin from Icelandic moss
EX_gal1a3acgal_m	1	red	Alpha-1,3-Galactose-N-acetylgalactosamine	EX_pyr_m	3	green	pyruvate
EX_glygn4_m	1	red	glycogen, structure 4 (glycogenin-1,6-{2[1,4-Glc], [1,4-Glc]})	EX_n2_m	3	green	Nitrogen
EX_mantr_m	1	red	mannotriose (beta-1,4)	EX_so3_m	3	green	Sulfite
EX_stys_m	1	red	Stachyose	EX_mqn8_m	3	green	Menaquinone 8
EX_ispre_m	1	red	Isoprene	EX_btoh_m	3	green	butanol
EX_tma_m	1	red	trimethylamine	EX_dpcoa_m	3	green	Dephospho-CoA
EX_nac_m	1	red	Nicotinate	EX_actn_R_m	3	green	(R)-Acetoin
EX_ppa_m	1	red	propionate	EX_pecticgal_m	3	green	Pectic galactan (potato)
EX_h_m	1	red	proton	EX_glyclt_m	3	green	glycolate
EX_h2_m	1	red	Hydrogen	EX_glygn2_m	3	green	glycogen, structure 2 (glycogenin-1,6-{7[1,4-Glc], 4[1,4-Glc]})
EX_ac_m	1	red	acetate	EX_tsul_m	3	green	thiosulfate(2-)
EX_pac_m	1	red	phenylacetate	EX_fuc1p_L_m	3	green	L-Fucose 1-phosphate
EX_2ddglcn_m	1	red	2-Dehydro-3-deoxy-D-gluconate	EX_trypta_m	3	green	tryptaminium
EX_icdchol_m	1	red	Isochenodeoxycholate	EX_cspg_a_m	3	green	chondroitin sulfate A (GalNAc4S-GlcA) proteoglycan
EX_for_m	1	red	Formate	EX_cspg_b_m	3	green	chondroitin sulfate B / dermatan sulfate (IdoA2S-GalNAc4S) proteoglycan
EX_man6p_m	1	red	D-Mannose 6-phosphate	EX_cspg_c_m	3	green	chondroitin sulfate C (GalNAc6S-GlcA) proteoglycan
EX_taur_m	1	red	Taurine	EX_mal_L_m	2	blue	(S)-malate(2-)
EX_but_m	1	red	butyrate	EX_tym_m	2	blue	tyraminium
EX_HC02194_m	1	red	Ursodiol	EX_gcald_m	2	blue	glycolaldehyde

EX_adocbl_m	1	red	Adenosylcobalamin	EX_nh4_m	2	blue	Ammonium
EX_hspg_degr_6_m	1	red	Product GAM2STASE_HS2	EX_dsT_antigen_m	2	blue	disialyl-T antigen
EX_hspg_degr_11_m	1	red	Product GAM2STASE_HS3	EX_dsT_antigen_rl_m	2	blue	released disialyl-T antigen
EX_hspg_degr_4_m	1	red	Product AS6TASE_HS1	EX_dtmp_m	2	blue	dTMP
EX_hspg_degr_1_m	1	red	Product 1 HEPARL1_e	EX_12ppd_S_m	2	blue	(S)-propane-1,2-diol
EX_hspg_degr_7_m	1	red	Product IS2TASE_HS1	EX_arabttr_m	2	blue	Arabinotriose
EX_hspg_rest_m	1	red	Undegradable product GALASE_HS	EX_diact_m	2	blue	Diacetyl
EX_isochol_m	1	red	Isocholate	EX_glyb_m	2	blue	glycine betaine
EX_amylopect900_m	1	red	Amylopectin, 900 repeat units	EX_fuc_L_m	2	blue	L-fucose
EX_hspg_degr_10_m	1	red	Product AS3TASE_HS2	EX_arabinogal_m	2	blue	Larch arabinogalactan
EX_hspg_degr_12_m	1	red	Product GLCNACASE_HS2	EX_core7_m	2	blue	Core 7
EX_ha_m	1	red	hyaluronan	EX_xtsn_m	2	blue	xanthosine
EX_ha_pre1_m	1	red	hyaluronan biosynthesis, precursor 1	EX_hspg_degr_13_m	2	blue	Product GLCAASE_HS
EX_chtbs_m	1	red	N,N-diacetylchitobiose	EX_ile_L_m	2	blue	L-isoleucine
EX_alagln_m	1	red	L-alanyl-L-glutamine	EX_gthrd_m	2	blue	Reduced glutathione
EX_alaglu_m	1	red	L-alanyl-L-glutamate	EX_ha_deg1_m	2	blue	hyaluronan degradation product 1
EX_MGlcn10_m	1	red	mucin-type O-glycan No 10	EX_ura_m	2	blue	uracil
EX_MGlcn10_rl_m	1	red	released mucin-type O-glycan No 10	EX_fru_m	2	blue	D-Fructose
EX_glygn5_m	1	red	glycogen, structure 5 (glycogenin-2[1,4-Glc])	EX_Tyr_ggn_m	2	blue	Tyr-194 of apo-glycogenin protein (primer for glycogen synthesis)
EX_meoh_m	1	red	methanol	EX_etoh_m	2	blue	ethanol
EX_din_m	1	red	Deoxyinosine	EX_alahis_m	2	blue	L-alanyl-L-histidine
EX_acgalglcur_m	1	red	Dimer GalNAc-GlcA	EX_rbflvrd_m	2	blue	Reduced riboflavin
EX_amylose300_m	1	red	amylose	EX_lys_L_m	2	blue	L-lysinium(1+)
EX_gchola_m	1	red	glycocholate	EX_pect_m	2	blue	pectins
EX_tchola_m	1	red	taurocholate	EX_alagly_m	2	blue	L-alanylglycine
EX_12dhchol_m	1	red	12-Dehydrocholate	EX_alathr_m	2	blue	L-alanyl-L-threonine
EX_cholate_m	1	red	Cholate	EX_alaleu_m	2	blue	L-alanyl-L-leucine
EX_dgchol_m	1	red	glycochenodeoxycholate	EX_metala_m	2	blue	L-methionyl-L-alanine
EX_C02528_m	1	red	chenodeoxycholate	EX_alaasp_m	2	blue	L-alanyl-L-aspartate
EX_tdchola_m	1	red	taurochenodeoxycholate	EX_cspg_ab_rest_m	2	blue	Not degradable product CSABCPASE_A/B
EX_levanb_m	1	red	Levanbiose	EX_cspg_c_rest_m	2	blue	Not degradable product CSABCPASE_C
EX_levanttr_m	1	red	Levantriose	EX_dextran40_m	2	blue	Dextran 40, 1,6-alpha-D-Glucan
EX_13ppd_m	3	green	Propane-1,3-diol	EX_duri_m	2	blue	Deoxyuridine
EX_urea_m	3	green	Urea	EX_glcmannan_m	2	blue	Konjac glucomannan
EX_akg_m	3	green	2-Oxoglutarate	EX_hxan_m	2	blue	Hypoxanthine
EX_core2_m	3	green	Core 2	EX_glyc_m	2	blue	Glycerol
EX_core2_rl_m	3	green	released Core 2	EX_h2o_m	2	blue	Water
EX_agam1b6agal_m	3	green	Beta-1,6-N-acetylglucosamine-N-acetylgalactosamine	EX_4hbz_m	2	blue	4-hydroxybenzoate

EX_core6_m	3	green	Core 6	EX_ptrc_m	2	blue	Putrescine
EX_galur_m	3	green	D-Galacturonate	EX_isocapr_m	2	blue	Isocaproate
EX_lald_L_m	3	green	(S)-lactaldehyde	EX_isoval_m	2	blue	Isovalerate, 3-Methylbutanoate
EX_ppi_m	3	green	Diphosphate	EX_glyleu_m	2	blue	Glycylleucine
EX_5mthf_m	3	green	5-Methyltetrahydrofolate	EX_glyasn_m	2	blue	Glycyl-L-asparagine
EX_glcur_m	3	green	D-glucuronate	EX_glyasp_m	2	blue	Glycyl-L-aspartate
EX_7dhcdchol_m	3	green	7-Dehydrochenodeoxycholate	EX_glygln_m	2	blue	Glycyl-L-glutamine
EX_xylan_m	3	green	Oat spelt xylan	EX_glyglu_m	2	blue	Glycyl-L-glutamate
EX_co2_m	3	green	carbon dioxide	EX_glypro_m	2	blue	Glycylproline
EX_ocdca_m	3	green	octadecanoate (n-C18:0)	EX_glytyr_m	2	blue	Glycyl-L-tyrosine
EX_gua_m	3	green	Guanine	EX_glymet_m	2	blue	Glycyl-L-methionine
EX_pydxn_m	3	green	Pyridoxine	EX_glyphe_m	2	blue	Glycylphenylalanine
EX_agal1b6agal_m	3	green	Beta-1,6-N,N-diacetylgalactosamine	EX_acerA_m	5	magenta	Aceric acid (3-C-carboxy-5-deoxy-l-xylose) (rare sugar in rhamnogalacturonan II)
EX_galmannan_m	3	green	Carob galactomannan	EX_apio_D_m	5	magenta	D-Apiose (rare sugar in rhamnogalacturonan II)
EX_o2_m	3	green	O2	EX_kdo_m	5	magenta	3-Deoxy-D-manno-2-octulosonate
EX_34dhpha_m	3	green	(3,4-dihydroxyphenyl)acetate	EX_3ddlhept_m	5	magenta	3-deoxy-d-lyxo-heptulosaric acid (rare sugar in rhamnogalacturonan II)
EX_oaa_m	3	green	Oxaloacetate	EX_2omfuc_m	5	magenta	2-o-methylfucose (rare sugar in rhamnogalacturonan II)
EX_3mop_m	3	green	3-methyl-2-oxopentanoate	EX_2omxyl_m	5	magenta	2-O-methylxylose (rare sugar in rhamnogalacturonan II)
EX_ttdca_m	3	green	tetradecanoate (n-C14:0)	EX_Tn_antigen_m	6	orange	Tn antigen
EX_sbt_D_m	3	green	D-glucitol	EX_f1a_m	6	orange	F1alpha
EX_glcn_m	3	green	D-gluconate	EX_f1a_rl_m	6	orange	released F1alpha antigen
EX_csn_m	3	green	cytosine	EX_acnam2a6agal_m	6	orange	Alpha-2,6-N-acetylneuraminate-N-acetylgalactosamine
EX_dcyt_m	3	green	Deoxycytidine	EX_sTn_antigen_m	6	orange	sialyl-Tn antigen
EX_ncam_m	3	green	Nicotinamide	EX_agam1b3agal_m	6	orange	Beta-1,3-N-acetylglucosamine-N-acetylgalactosamine
EX_met_L_m	3	green	L-methionine	EX_core3_m	6	orange	Core 3
EX_hspg_degr_15_m	3	green	Product IDOURASE_HS3	EX_gncore1_m	6	orange	GlcNAc-alpha-1,4-Core 1
EX_phpyr_m	3	green	keto-phenylpyruvate	EX_gncore1_rl_m	6	orange	released GlcNAc-alpha-1,4-Core 1
EX_hspg_degr_14_m	3	green	Product GLCNACASE_HS3	EX_sT_antigen_m	6	orange	sialyl-T antigen
EX_pydam_m	3	green	Pyridoxamine	EX_sT_antigen_rl_m	6	orange	released sialyl-Tn antigen
EX_h2s_m	3	green	Hydrogen sulfide	EX_lmn2_m	6	orange	Laminaribiose
EX_pi_m	3	green	hydrogenphosphate	EX_acgal_m	6	orange	N-acetyl-D-galactosamine
EX_5mta_m	3	green	5-Methylthioadenosine	EX_gal1b3acgal_m	6	orange	Beta-1,3-Galactose-N-acetylgalactosamine
EX_spmd_m	3	green	Spermidine	EX_T_antigen_m	6	orange	T antigen (core 1)

Supplementary Table 13: List of species and their cluster affiliation (corresponding to supplementary figure 9) along with the number of imported metabolites influenced (stimulated/inhibited) in IBS. The features are listed according to the order in the dendrogram shown in the corresponding heatmap.

Species	Cluster ID	Color	Metabolites stimulated	Metabolites inhibited	Species	Cluster ID	Color	Metabolites stimulated	Metabolites inhibited
Bifidobacterium longum	2	blue	109	80	*Bacteroides massiliensis*	1	red	69	52
Bacteroides caccae	2	blue	107	103	*Streptococcus salivarius*	1	red	49	115
Parabacteroides distasonis	2	blue	80	85	*Prevotella copri*	1	red	27	114
Bacteroides thetaiotaomicron	2	blue	106	63	*Streptococcus thermophilus*	1	red	47	32
Alistipes onderdonkii	2	blue	127	73	*Clostridium bartlettii*	1	red	44	48
Dialister invisus	2	blue	125	95	*Clostridium sp L2 50*	1	red	56	46
Ruminococcus obeum	2	blue	98	96	*Bacteroides coprocola*	1	red	88	34
Barnesiella intestinihominis	2	blue	79	116	*Ruminococcus champanellensis*	1	red	5	27
Collinsella aerofaciens	2	blue	96	118	*Lactobacillus acidophilus*	1	red	10	34
Dorea longicatena	2	blue	105	108	*Collinsella intestinalis*	1	red	7	38
Eubacterium hallii	2	blue	103	117	*Acidaminococcus fermentans*	1	red	11	32
Lachnospira eligens	2	blue	55	131	*Bifidobacterium angulatum*	1	red	21	33
Bacteroides vulgatus	2	blue	71	149	*Citrobacter freundii*	1	red	10	12
Roseburia intestinalis	2	blue	50	169	*Blautia producta*	1	red	9	11
Eubacterium siraeum	2	blue	83	107	*Lactobacillus gasseri*	1	red	1	0
Bacteroides cellulosilyticus	2	blue	71	115	*Blautia hydrogenotrophica*	1	red	1	0
Roseburia hominis	2	blue	69	111	*Klebsiella pneumoniae*	1	red	3	0
Escherichia coli	2	blue	86	105	*Anaerostipes hadrus*	1	red	4	3
Alistipes putredinis	2	blue	83	129	*Eggerthella lenta*	1	red	0	6
Bacteroides ovatus	2	blue	101	122	*Lactobacillus antri*	1	red	0	6
Ruminococcus bromii	2	blue	74	140	*Bifidobacterium catenulatum*	1	red	6	4
Ruminococcus torques	2	blue	91	115	*Lactobacillus ruminis*	1	red	5	6
Bacteroides dorei	2	blue	81	141	*Prevotella stercorea*	1	red	5	9
Faecalibacterium prausnitzii	2	blue	134	112	*Lactobacillus salivarius*	1	red	10	7
Bacteroides uniformis	2	blue	107	162	*Methanosphaera stadtmanae*	1	red	20	8
Eubacterium rectale	2	blue	102	141	*Lactobacillus crispatus*	1	red	15	14
Ruminococcus callidus	1	red	61	75	*Leuconostoc lactis*	1	red	23	11
Catenibacterium mitsuokai	1	red	61	65	*Streptococcus vestibularis*	1	red	13	14
Alistipes finegoldii	1	red	75	42	*Bifidobacterium dentium*	1	red	12	25
Bacteroides stercoris	1	red	85	70	*Bacteroides nordii*	1	red	26	17
Coprococcus catus	1	red	83	81	*Megamonas hypermegale*	1	red	18	33
Eubacterium biforme	1	red	60	73	*Haemophilus parainfluenzae*	1	red	29	46

Bifidobacterium bifidum	1	red	68	65	*Parabacteroides goldsteinii*	1	red	37	13
Phascolarctobacterium succinatutens	1	red	64	74	*Flavonifractor plautii*	1	red	48	13
Methanobrevibacter smithii	1	red	81	58	*Alistipes indistinctus*	1	red	39	21
Parabacteroides merdae	1	red	100	80	*Mitsuokella multacida*	1	red	53	12
Coprococcus comes	1	red	103	65	*Lactobacillus vaginalis*	1	red	48	33
Alistipes shahii	1	red	119	57	*Bacteroides coprophilus*	1	red	33	26
Akkermansia muciniphila	1	red	114	72	*Clostridium symbiosum*	1	red	29	34
Dorea formicigenerans	1	red	113	74	*Clostridium clostridioforme*	1	red	37	22
Bifidobacterium adolescentis	1	red	116	69	*Clostridium hathewayi*	1	red	33	21
Roseburia inulinivorans	1	red	107	85	*Bifidobacterium animalis*	1	red	44	20
Odoribacter splanchnicus	1	red	77	62	*Clostridium nexile*	1	red	36	31
Bacteroides fragilis	1	red	77	68	*Clostridium leptum*	1	red	72	16
Bacteroides salyersiae	1	red	46	38	*Parabacteroides johnsonii*	1	red	34	59
Bacteroides intestinalis	1	red	57	50	*Bacteroides clarus*	1	red	36	56
Coprococcus eutactus	1	red	100	55	*Veillonella dispar*	1	red	11	36
Eubacterium ventriosum	1	red	65	67	*Veillonella atypica*	1	red	7	42
Eubacterium ramulus	1	red	79	35	*Bifidobacterium pseudocatenulatum*	1	red	11	41
Clostridium bolteae	1	red	77	51	*Veillonella parvula*	1	red	8	80
Bacteroides xylanisolvens	1	red	68	34	*Bacteroides finegoldii*	1	red	34	52
Ruminococcus lactaris	1	red	64	33	*Bacteroides faecis*	1	red	24	40
Ruminococcus gnavus	1	red	114	39	*Bacteroides plebeius*	1	red	49	39
Bacteroides eggerthii	1	red	70	93	*Streptococcus parasanguinis*	1	red	38	60
Sutterella wadsworthensis	1	red	72	82					

Supplementary Table 14: List of exchanged metabolites by the community (imports) and their cluster memberships (corresponding to supplementary figure 9) in IBS. The features are listed according to the order in the dendrogram shown in the corresponding heatmap.

Reaction ID	Cluster ID	Color	Metabolite	Reaction ID	Cluster ID	Color	Metabolite
EX_cbl1_m	2	blue	Cob(I)alamin	EX_bglc_m	1	red	beta-glucans
EX_lichn_m	2	blue	Lichenin from Icelandic moss	EX_dad_2_m	1	red	2-deoxyadenosine
EX_galur_m	2	blue	D-Galacturonate	EX_rib_D_m	1	red	D-ribose
EX_q8_m	2	blue	Ubiquinone-8	EX_so3_m	1	red	Sulfite
EX_sheme_m	2	blue	Siroheme	EX_ade_m	1	red	Adenine
EX_strch2_m	2	blue	starch, structure 2 (1,6-{2[1,4-Glc], [1,4-Glc]})	EX_his_L_m	1	red	L-histidine
EX_pecticgal_m	2	blue	Pectic galactan (potato)	EX_Lcyst_m	1	red	L-cysteate(1-)
EX_3dhchol_m	2	blue	3-Dehydrocholate	EX_hspg_degr_10_m	1	red	Product AS3TASE_HS2
EX_levantttr_m	2	blue	Levantetraose	EX_gua_m	1	red	Guanine
EX_cholate_m	2	blue	Cholate	EX_tmao_m	1	red	Trimethylamine N-oxide
EX_gchola_m	2	blue	glycocholate	EX_ura_m	1	red	uracil
EX_tchola_m	2	blue	taurocholate	EX_thf_m	1	red	5,6,7,8-Tetrahydrofolate
EX_12dhchol_m	2	blue	12-Dehydrocholate	EX_thr_L_m	1	red	L-threonine
EX_7ocholate_m	2	blue	3alpha,12alpha-Dihydroxy-7-oxo-5beta-cholanate	EX_dcyt_m	1	red	Deoxycytidine
EX_HC02194_m	2	blue	Ursodiol	EX_adn_m	1	red	Adenosine
EX_uchol_m	2	blue	Ursocholate	EX_orn_m	1	red	Ornithine
EX_3dhcdchol_m	2	blue	3-Dehydrochenodeoxycholate	EX_btoh_m	1	red	butanol
EX_7dhcdchol_m	2	blue	7-Dehydrochenodeoxycholate	EX_hspg_degr_14_m	1	red	Product GLCNACASE_HS3
EX_malttr_m	3	green	maltotriose	EX_sbt_D_m	1	red	D-glucitol
EX_metala_m	3	green	L-methionyl-L-alanine	EX_trp_L_m	1	red	L-tryptophan
EX_alagly_m	3	green	L-alanylglycine	EX_bhb_m	1	red	(R)-3-hydroxybutyrate
EX_alathr_m	3	green	L-alanyl-L-threonine	EX_3hphac_m	1	red	3-Hydroxyphenylacetate
EX_glcmannan_m	3	green	Konjac glucomannan	EX_oaa_m	1	red	Oxaloacetate
EX_xylan_m	3	green	Oat spelt xylan	EX_n2o_m	1	red	Nitrous oxide
EX_glcur_m	3	green	D-glucuronate	EX_no_m	1	red	Nitric oxide
EX_inulin_m	3	green	Chicory inulin	EX_succ_m	1	red	Succinate
EX_gncore2_m	3	green	GlcNAc-alpha-1,4-Core 2	EX_raffin_m	1	red	Raffinose
EX_gncore2_rl_m	3	green	released GlcNAc-alpha-1,4-Core 2	EX_sucr_m	1	red	Sucrose
EX_agal1b3agal_m	3	green	Beta-1,3-N,N-diacetylgalactosamine	EX_fru_m	1	red	D-Fructose
EX_core4_m	3	green	Core 4	EX_kestopt_m	1	red	Kestopentaose
EX_core4_rl_m	3	green	released Core 4	EX_etoh_m	1	red	ethanol
EX_arabinoxyl_m	3	green	Arabinoxylan, 30 xyl_D, 18 arab_L, 3 gal, 3 glc_D	EX_lald_L_m	1	red	(S)-lactaldehyde
EX_dsT_antigen_m	3	green	disialyl-T antigen	EX_12ppd_S_m	1	red	(S)-propane-1,2-diol
EX_dsT_antigen_rl_m	3	green	released disialyl-T antigen	EX_nh4_m	1	red	Ammonium

EX_mantr_m	3	green	mannotriose (beta-1,4)	EX_pydam_m	1	red	Pyridoxamine
EX_chtbs_m	3	green	N,N-diacetylchitobiose	EX_cspg_ab_rest_m	1	red	Not degradable product CSABCPASE_A/B
EX_glypro_m	3	green	Glycylproline	EX_cspg_c_rest_m	1	red	Not degradable product CSABCPASE_C
EX_glygln_m	3	green	Glycyl-L-glutamine	EX_ins_m	1	red	Inosine
EX_glyglu_m	3	green	Glycyl-L-glutamate	EX_dtmp_m	1	red	dTMP
EX_glyasn_m	3	green	Glycyl-L-asparagine	EX_pi_m	1	red	hydrogenphosphate
EX_glyasp_m	3	green	Glycyl-L-aspartate	EX_lmn2_m	1	red	Laminaribiose
EX_glyleu_m	3	green	Glycylleucine	EX_glc_D_m	1	red	D-glucose
EX_lac_D_m	3	green	(R)-lactate	EX_lmn30_m	1	red	laminarin
EX_alagln_m	3	green	L-alanyl-L-glutamine	EX_isobut_m	1	red	Isobutyrate, 2-Methylpropanoate
EX_alaglu_m	3	green	L-alanyl-L-glutamate	EX_hxan_m	1	red	Hypoxanthine
EX_stys_m	3	green	Stachyose	EX_pro_L_m	1	red	L-proline
EX_xyluglc_m	3	green	Xyluglucan	EX_thymd_m	1	red	Thymidine
EX_acgal_m	3	green	N-acetyl-D-galactosamine	EX_glyc_m	1	red	Glycerol
EX_arabinogal_m	3	green	Larch arabinogalactan	EX_gly_m	1	red	Glycine
EX_amylose300_m	3	green	amylose	EX_phe_L_m	1	red	L-phenylalanine
EX_rhamnogalurII_m	3	green	Wine rhamnogalacturonan II	EX_chor_m	1	red	chorismate
EX_acerA_m	3	green	Aceric acid (3-C-carboxy-5-deoxy-l-xylose) (rare sugar in rhamnogalacturonan II)	EX_val_L_m	1	red	L-valine
EX_apio_D_m	3	green	D-Apiose (rare sugar in rhamnogalacturonan II)	EX_ile_L_m	1	red	L-isoleucine
EX_kdo_m	3	green	3-Deoxy-D-manno-2-octulosonate	EX_arg_L_m	1	red	L-argininium(1+)
EX_3ddlhept_m	3	green	3-deoxy-d-lyxo-heptulosaric acid (rare sugar in rhamnogalacturonan II)	EX_cys_L_m	1	red	L-cysteine
EX_2omfuc_m	3	green	2-o-methylfucose (rare sugar in rhamnogalacturonan II)	EX_lanost_m	1	red	lanosterol
EX_2omxyl_m	3	green	2-O-methylxylose (rare sugar in rhamnogalacturonan II)	EX_glyb_m	1	red	glycine betaine
EX_meoh_m	3	green	methanol	EX_leu_L_m	1	red	L-leucine
EX_rmn_m	3	green	L-Rhamnose	EX_no3_m	1	red	Nitrate
EX_pect_m	3	green	pectins	EX_xtsn_m	1	red	xanthosine
EX_dextran40_m	3	green	Dextran 40, 1,6-alpha-D-Glucan	EX_no2_m	1	red	Nitrite
EX_pydxn_m	3	green	Pyridoxine	EX_4hpro_LT_m	1	red	trans-4-hydroxy-L-proline
EX_core8_m	3	green	Core 8	EX_hspg_rest_m	1	red	Undegradable product GALASE_HS
EX_gal1a3acgal_m	3	green	Alpha-1,3-Galactose-N-acetylgalactosamine	EX_gthrd_m	1	red	Reduced glutathione
EX_homogal_m	3	green	Homogalacturonan	EX_agal1b6agal_m	1	red	Beta-1,6-N,N-diacetylgalactosamine
EX_galmannan_m	3	green	Carob galactomannan	EX_ispre_m	1	red	Isoprene
EX_levanttr_m	3	green	Levantriose	EX_salcn_m	1	red	Salicin
EX_levanb_m	3	green	Levanbiose	EX_galt_m	1	red	galactitol
EX_gthox_m	3	green	Oxidized glutathione	EX_3hpppn_m	1	red	3-(3-hydroxy-phenyl)propionate
EX_ribflv_m	3	green	Riboflavin	EX_cinnm_m	1	red	trans-Cinnamate
EX_arabinan101_m	3	green	Arabinan	EX_pppn_m	1	red	Phenylpropanoate
EX_o2_m	3	green	O2	EX_ha_m	1	red	hyaluronan

EX_din_m	3	green	Deoxyinosine	EX_ha_pre1_m	1	red	hyaluronan biosynthesis, precursor 1
EX_gal_m	3	green	D-Galactose	EX_indole_m	1	red	Indole
EX_glymet_m	3	green	Glycyl-L-methionine	EX_thm_m	1	red	Thiamin
EX_xan_m	3	green	Xanthine	EX_arbt_m	1	red	Arbutin
EX_spmd_m	3	green	Spermidine	EX_pnto_R_m	1	red	(R)-Pantothenate
EX_ala_L_m	3	green	L-alanine	EX_actn_R_m	1	red	(R)-Acetoin
EX_mnl_m	3	green	D-Mannitol	EX_fuc_L_m	1	red	L-fucose
EX_alaasp_m	3	green	L-alanyl-L-aspartate	EX_core7_m	1	red	Core 7
EX_mal_L_m	3	green	(S)-malate(2-)	EX_acac_m	1	red	acetoacetate
EX_core6_m	3	green	Core 6	EX_uri_m	1	red	Uridine
EX_glygn2_m	3	green	glycogen, structure 2 (glycogenin-1,6-{7[1,4-Glc], 4[1,4-Glc]})	EX_arab_D_m	1	red	D-Arabinose
EX_MGlcn10_m	1	red	mucin-type O-glycan No 10	EX_2dmmq8_m	1	red	2-Demethylmenaquinone 8
EX_MGlcn10_rl_m	1	red	released mucin-type O-glycan No 10	EX_4hbz_m	1	red	4-hydroxybenzoate
EX_acgalglcur_m	1	red	Dimer GalNAc-GlcA	EX_kesto_m	1	red	Kestose
EX_acgalidour_m	1	red	Dimer GalNAc-IdoA	EX_kestottr_m	1	red	Kestotetraose
EX_cspg_b_degr_m	1	red	Dimer GalNAc4S-IdoA2S	EX_mqn7_m	1	red	Menaquinone 7
EX_4abut_m	1	red	4-Aminobutanoate	EX_Ser_Thr_m	1	red	protein-linked serine or threonine residue (O-glycosylation site)
EX_n2_m	1	red	Nitrogen	EX_gal1b3acgal_m	4	brown	Beta-1,3-Galactose-N-acetylgalactosamine
EX_glyphe_m	1	red	Glycylphenylalanine	EX_T_antigen_m	4	brown	T antigen (core 1)
EX_glytyr_m	1	red	Glycyl-L-tyrosine	EX_sT_antigen_m	4	brown	sialyl-T antigen
EX_crn_m	1	red	L-Carnitine	EX_sT_antigen_rl_m	4	brown	released sialyl-Tn antigen
EX_gbbtn_m	1	red	gamma-butyrobetaine	EX_gncore1_m	4	brown	GlcNAc-alpha-1,4-Core 1
EX_fol_m	1	red	Folate	EX_gncore1_rl_m	4	brown	released GlcNAc-alpha-1,4-Core 1
EX_acmana_m	1	red	N-acetyl-D-mannosamine	EX_f1a_m	4	brown	F1alpha
EX_mqn8_m	1	red	Menaquinone 8	EX_f1a_rl_m	4	brown	released F1alpha antigen
EX_metsox_S_L_m	1	red	L-Methionine Sulfoxide	EX_Tn_antigen_m	4	brown	Tn antigen
EX_strch1_m	1	red	starch, structure 1 (1,6-{7[1,4-Glc], 4[1,4-Glc]})	EX_agam1b3agal_m	4	brown	Beta-1,3-N-acetylglucosamine-N-acetylgalactosamine
EX_3mop_m	1	red	3-methyl-2-oxopentanoate	EX_core3_m	4	brown	Core 3
EX_arab_L_m	1	red	L-arabinose	EX_acnam2a6agal_m	4	brown	Alpha-2,6-N-acetylneuraminate-N-acetylgalactosamine
EX_12dgr180_m	1	red	1,2-Diacyl-sn-glycerol (dioctadecanoyl, n-C18:0)	EX_sTn_antigen_m	4	brown	sialyl-Tn antigen
EX_ttdca_m	1	red	tetradecanoate (n-C14:0)	EX_agam1b6agal_m	4	brown	Beta-1,6-N-acetylglucosamine-N-acetylgalactosamine
EX_hdca_m	1	red	Hexadecanoate (n-C16:0)	EX_core2_m	4	brown	Core 2
EX_csn_m	1	red	cytosine	EX_core2_rl_m	4	brown	released Core 2
EX_ncam_m	1	red	Nicotinamide	EX_C02528_m	4	brown	chenodeoxycholate
EX_ala_D_m	1	red	D-alanine	EX_dgchol_m	4	brown	glycochenodeoxycholate
EX_dgsn_m	1	red	Deoxyguanosine	EX_tdchola_m	4	brown	taurochenodeoxycholate

EX_drib_m	1	red	2-deoxy-D-ribose	EX_cspg_a_degr_m	4	brown	Dimer GalNAc4S-GlcA
EX_chol_m	1	red	Choline	EX_cspg_c_degr_m	4	brown	Dimer GalNAc6S-GlcA
EX_fe3_m	1	red	Fe3+	EX_alahis_m	4	brown	L-alanyl-L-histidine
EX_gsn_m	1	red	Guanosine	EX_pyr_m	4	brown	pyruvate
EX_glygn5_m	1	red	glycogen, structure 5 (glycogenin-2[1,4-Glc])	EX_Lcystin_m	4	brown	L-cystine
EX_tsul_m	1	red	thiosulfate(2-)	EX_taur_m	4	brown	Taurine
EX_core5_m	1	red	Core 5	EX_na1_m	4	brown	Sodium
EX_adocbl_m	1	red	Adenosylcobalamin	EX_gam26s_m	4	brown	Glucosamine 2,6-sulfate
EX_glycys_m	1	red	Gly-Cys	EX_hspg_degr_8_m	4	brown	Product IDOURASE_HS2
EX_alaleu_m	1	red	L-alanyl-L-leucine	EX_h2o_m	4	brown	Water
EX_btn_m	1	red	Biotin	EX_Tyr_ggn_m	4	brown	Tyr-194 of apo-glycogenin protein (primer for glycogen synthesis)
EX_pheme_m	1	red	Protoheme	EX_hspg_degr_2_m	4	brown	Product 2 HEPARL1_e
EX_isocapr_m	1	red	Isocaproate	EX_hspg_degr_3_m	4	brown	Product IDOURASE_HS1
EX_isoval_m	1	red	Isovalerate, 3-Methylbutanoate	EX_hspg_degr_5_m	4	brown	Product AS3TASE_HS1
EX_malt_m	1	red	Maltose	EX_ocdcea_m	4	brown	octadecenoate (n-C18:1)
EX_hspg_degr_12_m	1	red	Product GLCNACASE_HS2	EX_glyclt_m	4	brown	glycolate
EX_hspg_degr_6_m	1	red	Product GAM2STASE_HS2	EX_ocdca_m	4	brown	octadecanoate (n-C18:0)
EX_duri_m	1	red	Deoxyuridine	EX_hspg_degr_15_m	4	brown	Product IDOURASE_HS3
EX_ha_deg1_m	1	red	hyaluronan degradation product 1	EX_idon_L_m	4	brown	L-Idonate
EX_hspg_degr_13_m	1	red	Product GLCAASE_HS	EX_met_L_m	4	brown	L-methionine
EX_hspg_degr_1_m	1	red	Product 1 HEPARL1_e	EX_cit_m	4	brown	Citrate
EX_hspg_degr_7_m	1	red	Product IS2TASE_HS1	EX_ddca_m	4	brown	laurate
EX_hspg_degr_11_m	1	red	Product GAM2STASE_HS3	EX_urea_m	4	brown	Urea
EX_hspg_degr_4_m	1	red	Product AS6TASE_HS1	EX_glcn_m	4	brown	D-gluconate
EX_hspg_degr_9_m	1	red	Product AS6TASE_HS2	EX_hspg_m	4	brown	heparan sulfate proteoglycan
EX_levan1000_m	1	red	Levan	EX_xyl_D_m	4	brown	aldehydo-D-xylose
EX_mannan_m	1	red	Alpha-mannan, yeast	EX_starch1200_m	4	brown	starch
EX_g6p_m	1	red	D-Glucose 6-phosphate	EX_glygn4_m	4	brown	glycogen, structure 4 (glycogenin-1,6-{2[1,4-Glc], [1,4-Glc]})
EX_man_m	1	red	D-Mannose	EX_4abz_m	4	brown	4-Aminobenzoate
EX_cspg_c_m	1	red	chondroitin sulfate C (GalNAc6S-GlcA) proteoglycan	EX_fe2_m	4	brown	Fe2+
EX_cspg_a_m	1	red	chondroitin sulfate A (GalNAc4S-GlcA) proteoglycan	EX_isochol_m	4	brown	Isocholate
EX_cspg_b_m	1	red	chondroitin sulfate B / dermatan sulfate (IdoA2S-GalNAc4S) proteoglycan	EX_dpcoa_m	4	brown	Dephospho-CoA
EX_idour_m	1	red	L-iduronate	EX_icdchol_m	4	brown	Isochenodeoxycholate
EX_glctnb_m	1	red	Galacto-N-biose	EX_arabttr_m	4	brown	Arabinotriose
EX_lctnb_m	1	red	Lacto-N-biose	EX_dextrin_m	4	brown	Dextrin
EX_2ddglcn_m	1	red	2-Dehydro-3-deoxy-D-gluconate	EX_malthx_m	4	brown	Maltohexaose
EX_pydx_m	1	red	Pyridoxal	EX_rhamnogalurI_m	4	brown	Potato rhamnogalacturonan I

EX_ptrc_m	1	red	Putrescine	EX_amylopect900_m	4	brown	Amylopectin
EX_lys_L_m	1	red	L-lysinium(1+)	EX_pullulan1200_m	4	brown	pullulan
EX_5mta_m	1	red	5-Methylthioadenosine	EX_gcald_m	4	brown	glycolaldehyde
EX_5mthf_m	1	red	5-Methyltetrahydrofolate	EX_lac_L_m	4	brown	(S)-lactate
EX_nmn_m	1	red	NMN	EX_acald_m	4	brown	acetaldehyde
EX_so4_m	1	red	sulfate	EX_akg_m	4	brown	2-Oxoglutarate

Chapter 7

Discussion

Table of contents

7.1 General summary 527

7.2 The gut microbiome is significantly altered in IBS but cannot differentiate IBS subtypes.....529

7.3 Metabolic analysis revealed alterations in IBS and BAM 530

7.4 Ecological alterations were observed in IBS. 533

7.5 Computational metabolic modelling can identify potential microbiome modulators.....535

7.6 Study limitations and scope for improvement 535

7.7 Future directions.....537

7.8 Conclusion 538

7.9 References.....539

7.1 General summary

My research thesis focussed on Irritable Bowel Syndrome (IBS), which is a chronic functional gastrointestinal disorder in which patients suffer from no functional anomalies but are symptomatic for abdominal pain and distension accompanied by alterations in bowel habits. Due to its high global prevalence of ~11.2%[1], further research to understand IBS aetiology is warranted. Current clinical diagnosis and classification of IBS are based on symptoms such as abdominal pain, changes in bowel habits, and appearance of stool, as directed by the Rome foundation[2].

My research aimed at furthering the existing knowledge base of microbiome research, through the development and implementation of methods which may aid in therapeutic options through rational microbiome modulation and personalised medicine exploiting the stratification of IBS patients. For this, a cohort of 80 IBS patients and 65 matched Control subjects was established by our group. I analysed the faecal microbiome, metabolome, habitual diet and investigated their differences in IBS and Bile acid malabsorption (BAM). Predictive analysis of the multi-omics datasets based on a novel machine learning approach provided microbial and metabolomic signatures which were predictive of IBS and BAM. Microbiome structure analysis based on correlation and metabolic interaction networks provided additional insights into differences in interaction patterns within the microbiome between IBS and Controls. Finally, my attempt at computational estimation of the metabolic events of the microbiome using genome scale metabolic models also provided distinct taxonomic-metabolomic signature of dysregulation in IBS, which can be used for LBP development, as well as development of microbiome-based diagnostics for IBS.

There have been numerous attempts to predict IBS and other disorders using machine learning algorithms and microbiome data [3–7]. These machine learning models may aid

in better understanding of the variance within the data; however, lack of standardization, transparency, and validation, along with disparity in methods, lead to generation of predictive models with overestimated performance. The majority of the machine learning algorithms are data driven and are sensitive to trends in the training data (in some cases, stochastic), which may lead to overfitting where the model performs much better during training than testing on a held-out dataset [7]. Additionally, due to drifts in data, the transferability of developed models to new datasets is limited. For example, microbiome strain-based models trained on samples from one demographic parameter like geography or age, often cannot be used to predict samples belonging to a different age group or place. A similar issue had been addressed previously wherein the identification of microbiome alterations were improved upon age adjustment during model training [8].

A biologically relevant context-dependent predictive model development may mitigate some of these discussed limitations. The advantage of including biologically meaningful context on applicability and transferability of predictive models have been demonstrated in two recent studies, identified competing guilds of microbiome, application of which could be extended to various diseases including Type 2 Diabetes, Atherosclerotic Cardiovascular disease, and liver cirrhosis, along with association with COVID-19 severity and prognosis [9,10]. In my analysis, I have applied known biological concepts relevant to IBS, like loss of core microbiome or fibre associated taxa, and gain in mucin degradation, in developing predictive models for IBS. Additionally, the predictive models developed in this study, based on inter-microbial co-operation, also indirectly includes biological context like extent of potential metabolic interactions within a microbiome community.

7.2 The gut microbiome is significantly altered in IBS but cannot differentiate IBS subtypes

I have shown the faecal bacteriome, mycobiome, and virome to be significantly different between IBS patients and Control subjects but that it could not be used to differentiate between clinical sub-types of IBS. This observation was consistent across analyses using 16S rRNA gene amplicons, shotgun metagenomic sequences, and Co-abundant gene groups (CAGs) representing strain-level variables, which are commonly referred as metagenomic species. A recent integrative study of multiple datasets also showed the microbiome's inability to segregate the clinical subtypes of IBS [11], which corroborated my observation.

Inter-kingdom analysis also showed that the integrated microbiome components could be used to significantly differentiate IBS from Controls, but not the IBS sub-types. It was interesting to note that a higher degree of separation was observed between IBS and Controls when bacteriome abundance was part of the integrated abundance, suggesting bacteriome to be the major driver of variance in the microbiome. Additionally, I observed a greater and significant agreement and co-variation between bacteriome and mycobiome, and bacteriome and virome, but the mycobiome did not show a substantial agreement with the virome. Since the virome analysis was focussed on bacteriophages, the significant agreement of virome with bacteriome, but not with mycobiome was intuitive.

Microbiome alterations were also reflected in predictive analysis. I identified bacterial strain level metagenomic variables (based on CAGs) and strain level bacteriome features (zOTUs) which were predictive of IBS. Among the non-bacterial components of the microbiome, the mycobiome showed marginal but significant predictive capabilities, while the gut virome, and inter-kingdom integrated profiles showed

prediction performances similar to that of bacteriome, therefore, suggesting the sufficiency of bacteriome abundance in prediction and stratification of IBS patients.

7.3 Metabolic analysis revealed alterations in IBS and BAM

Compositional and functional alterations in the microbiome are reflected as metabolic alterations, which subsequently affects the host's physiology. Therefore, investigation of the host metabolome, and contribution of the microbiome to the metabolomic repertoire of the host is crucial for the understanding of disease aetiology. Additionally, for rational microbiome modulation and identification of robust candidates for diagnostic and therapeutic applications, requirement of iterative experimentation and metabolome quantification arises, which is not cost effective. Therefore, computational approaches like use of genome scale metabolic models can improve the time and cost requirements of the screening process. In this thesis, I have investigated the metabolic alterations through analysis of untargeted metabolomics of urine and faecal samples as well as computational metabolic modelling of the microbiome communities. The obtained differences were concurrent which further validated the application of computational methods to understand metabolic dynamics within a microbiome.

Differences in urine and faecal metabolites between IBS and Controls were found to be significant, and like microbiota composition, the metabolome profiles were not significantly different between IBS sub-types. Community level genome scale metabolic model analysis also showed that overall, internal reactions, metabolite exports, and imports profiles were significantly different between IBS and Controls, but these profiles were similar between IBS subtypes.

In metabolic modelling analysis, I observed increased imports of amino acids by the gut microbiota in IBS, which may reduce the availability of this nutrient for the host related functions. I also observed that there were increased levels of predicted bile acid exchanges between microbes in IBS, based on metabolic model analysis, suggesting greater availability of BAs in the extracellular environment, which may be linked to observed higher levels of BAs in the faecal metabolomes of IBS patients compared to controls in the cohort reported in this thesis, which has also been reported previously [12–15].

Other major alterations observed in the urine metabolome of IBS were higher levels of metabolites with known associated with fatty acid oxidation/metabolism in other diseases [16–18].

Another notable difference in faecal metabolome was significantly lower levels of adenosine and inosine in IBS samples. Adenosine and its downstream metabolite inosine are part of the human-microbiome nucleotide degradation pathway, which may be redirected to purine salvage pathway through hypoxanthine, thereby producing nucleotides which act as energy source for gut epithelial cells. Lower levels of adenosine and inosine in IBS are suggestive of lower levels of purine salvage in IBS, leading to possible impediment of usage of such nucleotides as energy source by gut epithelial cells. These observation corroborate previous studies which indicate adenosine, inosine, and hypoxanthine to be biomarkers of IBS [19,20].

From computational analysis of community level metabolic models, I observed significant and higher levels of mucin degradation and release of mucin carbohydrate antigens like mucin T-antigens, Tn-antigens, and F1 alpha antigens, by the microbiota in IBS. Additionally, an increased rate of export of glycerol (linked to visceral

hypersensitivity and abdominal pain [21]) was observed in IBS. A significantly reduced rate of import of hydrogen sulphide (a by-product of microbial metabolism linked to inflammation [22]), and a reduced rate of export of 4-aminobutanoate (GABA; a metabolite linked to modulation of gastrointestinal motility and inflammation [23]) by *Odoribacter splanchnicus* , were observed in IBS, suggesting possible induction of inflammation and dysregulation of gastrointestinal motility, respectively, to be higher in IBS. Previous reports also postulate the association of IBS symptom severity to the transit time of the stool which is directly linked to gastrointestinal motility [24], as well as impact of hydrogen sulphide in patients with IBS-D [22,25].

In agreement with previous observations[26], dietary components representing a high fibre diet were observed to be significantly associated with Controls but not IBS. These components showed concordance with faecal and urine metabolites associated with Control subject derived samples. Dietary fibre deprivation has been shown to encourage degradation of mucins which are utilized along with the associated catabolites as an energy source [27]. Metabolic model analysis revealed that when microbial communities are modelled with an average western diet (not rich in fibres), rates of mucin degradation are elevated. This effect is more pronounced in IBS indicating that the lack of dietary fibres, driving the mucin degradation, is amplified by the composition of the microbial community associated with IBS, specifically due to a reduction/loss of core species in IBS.

The urine metabolome was able to predict IBS with high accuracy, and contained signatures related to IBS. Findings from predictive analysis of the faecal metabolome showed that the faecal metabolome was predictive of IBS. A high precision was also seen in the classification of BAM status based on % SeHCAT retention. Such an

approach can be useful in distinguishing patients with BAM and non-BAM associated IBS.

7.4 Ecological alterations were observed in IBS.

Determination and interpretation of ecological structure of a population can be achieved by employing multiple analytical approaches. One such approach is to identify agreement between omics datasets. It was hypothesized that a healthy state should show a higher concordance between microbiome composition and function, diet, and metabolome. Findings from my inter-omics association analysis indicated the same, wherein, a greater number of microbiome, metabolome, and dietary components were associated with controls and with a higher degree of agreement between them, while in IBS subjects, inter-dataset agreement significantly reduced. This further suggested an increased level of discordance in microbiome-host co-metabolism in IBS. A previous study also reported weaker microbe-metabolome associations in IBS compared to controls, further indicating loss of agreement between microbiome-host co-metabolism in altered microbiome state of IBS [28]. The host-microbiota co-metabolism involve metabolites produced by the host and its microbiome which work as molecular information shuttles [29]. Therefore, a dysregulation of such events may be related to manifestation of IBS symptoms [30].

I also investigated the microbiome in terms of bacterial co-occurrence networks and observed a lack of number of inter connections between bacterial species in IBS, suggestive of reduced co-operation between the community members. Similar observations were made in a previous study [31], which further indicates the potential role of inter-bacterial co-operation in maintenance of a healthy microbiota. A similar lack of co-operation between bacterial species in IBS was observed based on my metabolic interaction-based network analysis using genome scale metabolic models,

which further suggests that the observed network differences are not statistical artifacts but are biologically driven. Interestingly, the co-operation network-based metrics were also observed to be predictive of IBS.

This lack of co-operation was not limited to the bacterial community of microbiome in IBS. I analysed inter-kingdom structural differences between bacteriome and mycobiome, as well as between bacteriome and virome of the same subjects. I observed significant differences in the inter-kingdom microbiome structure between IBS and controls. Besides observing reduced co-occurrence/positive interactions between bacteriome and non-bacterial components in IBS, I observed a large difference in inter-kingdom negative interactions (co-exclusion) between IBS and Control populations. Reduced levels of negative connections between bacteriome and other kingdoms in IBS directs towards reduction in inhibition of fungal and viral growths by the gut bacteria, which might affect the gut health in a detrimental way. Hence, the results indicate that co-occurrence is not the only factor to be considered for microbiome structure analysis, and there exists a significant role of the negative associations between bacteria and other microbiome components in determining a healthy microbiome structure.

Two recent studies investigating microbiome structure in health and diseases, have identified self-competing guilds of microbial species, proposed to be part of the core microbiome which can serve as robust biomarkers for microbiome-based diagnosis of diseases like Type 2 Diabetes, Atherosclerotic Cardiovascular disease, and liver cirrhosis, along with association with COVID-19 severity and prognosis [9,10]. These current results also indicate the consideration of inter-microbial negative interactions for diagnostic purposes.

7.5 Computational metabolic modelling can identify potential microbiome modulators.

Computational metabolic modelling of the microbiome community of IBS and Control samples revealed biologically relevant changes in the metabolic output of the microbiome. One may argue and question the validity of such computational predictions, but occurrence of similar alterations in experimentally derived metabolomic data along with similar observations reported in independent studies strengthens the scope of utilizing computational methods in understanding metabolic dynamics of the microbiota.

I proceeded with microbiome perturbation studies to identify effect of increase of different species on the global metabolic output of the community. The results revealed some interesting and novel insights which can further our understanding of IBS pathophysiology. I observed that the metabolic output modulation capabilities of a species are related to its level of co-operation with the other members of the microbiome community. Moreover, the analysis predicted species which are capable of strongly impacting the microbiome metabolism upon increase in their abundance. These strong modulators were also found to be the most interactive ones in the co-operation and competitive networks.

7.6 Study limitations and scope for improvement

Although my research achieved a comprehensive analysis of microbiome and its metabolic dynamics in IBS, it is not devoid of limitations and requires further investigation. One major limitation of my study was analysis of a cohort with only 80 IBS subjects and 65 Controls from a single European city with subjects belonging to

one ethnic group. The findings of this thesis should be replicated on larger populations spanning across geographies, and ethnicities.

I have analysed the microbiome mainly at species or higher taxonomic levels. In recent years, we have learnt that strains matter in defining the gut microbiome. Although there are common genes present in all the strains of a species, there is a potentially relevant genomic variability at strain level [32]. For example, genomic comparison of strains of *Faecalibacterium prausnitzii*, a commonly found species in the gut, showed a low level of average nucleotide identity, and a high level of genome plasticity between the genomes [33].

As stated previously, a large-scale analysis is required to strengthen the claims, including the trained predictive models. A machine learning model, stratifying patients with IBS or patients with BAM, if trained on a larger set of samples, is expected to show increased robustness in predictive performance.

The metabolic models used in this study were species based, and the AGORA V1 models used, included a limited number of species in their database. With release of AGORA 2 [34], the list of strain level metabolic models has increased. Therefore, there is a scope of improving the metabolic models and their outcomes, as a greater level of diversity can now be captured during community level metabolic model development. Metabolic models of the microbiome community, based on strain-level information, can provide additional insight which might have been overlooked in the current study. Additionally, a major limitation of such computational approach is the validity of signatures in the real world. It is imperative to experimentally validate any proposed mechanism derived from such computational analysis.

7.7 Future directions

Future work may include expansion of the current analysis across geographies and ethnicities using large population sizes. Additionally, quantitative estimation reported metabolic markers using targeted metabolomics can aid in development of diagnostic and therapeutic products.

As discussed previously, metabolic modelling using strain level metabolic models can provide insights that might have been missed in my current analysis. Analysis based on metagenome-assembled genomes (MAGs), and development of metabolic models thereof, can prove to be biologically more relevant, and may capture strain level variability in metabolism. Furthermore, the identified modulators of the microbiome using computational modelling need to be validated experimentally for clinical applications.

Although more complex, inclusion of metabolic models of other non-bacterial components of the microbiome along with metabolic model of host will, first, capture the metabolic diversity within the gut, and second, may provide further information related to host-microbiome co-metabolism and crosstalk.

With increasing interest in microbiome-based diagnostics, complex machine learning methods including deep learning, and neural networks, in microbiome research have been applied [35,36]. These predictive algorithms have shown promise, however, the majority of the machine learning methods are data-driven (learning performance is influenced by variance in the training data), and lack biological contexts [36]. Providing biological context to these algorithms based on multi-omics analyses may help in development of robust predictive models, which may have better biological interpretability. Moreover, models based on latent variables based on a fixed external

knowledge base, may reduce the impact of data drift, and increase the transferability of the models to other datasets.

With increase in the data quality as well as quantity, along with development of variety of databases containing biological information relevant to microbiome, predictive modelling can help.

7.8 Conclusion

The findings of my research have clinical implications. The identified microbiome and metabolic markers can be translated to clinical practice. My findings indicated that the conventional Rome subtyping of IBS is not reflected by the microbiome and metabolomic alterations, and an alternative microbiome-based stratification strategy of IBS patients should be adopted for microbiome-based treatments such as LBPs. The predictive model using metabolome could predict BAM, therefore can aid in focussing on patients with non-BAM associated IBS and allow for the treatment of BAM with sequestrants within the part of the IBS population that need them. Other microbiome based predictive models of IBS can be applied for an informed stratification of patients, which may further aid in patient screening for trials of microbiome directed therapies.

Owing to the low clinical efficacy of probiotics in IBS, and lack of specificity of other available treatment modalities (like rifaximin, and Faecal Microbiota Transplantation), development of rationally designed live biotherapeutic products is required, which will not only interact with the recipient's microbiome, but will impact the host-microbiome co-metabolism in a beneficial way. My analysis of metabolic models provides the meaningful findings, and a screening strategy of potent modulators, which can help in identifying potential candidates for LBP development.

7.9 References

1. Mearin F, Lacy BE, Chang L, et al. Bowel Disorders. Gastroenterology 2016:S0016-5085(16)00222–5.

2. Drossman DA, Hasler WL. Rome IV-Functional GI Disorders: Disorders of Gut-Brain Interaction. Gastroenterology 2016;150:1257–1261.

3. Topçuoğlu BD, Lesniak NA, Ruffin MT, et al. A Framework for Effective Application of Machine Learning to Microbiome-Based Classification Problems Blaser MJ, ed. mBio 2020;11:e00434-20.

4. Botschuijver S, Roeselers G, Levin E, et al. Intestinal Fungal Dysbiosis Is Associated With Visceral Hypersensitivity in Patients With Irritable Bowel Syndrome and Rats. Gastroenterology 2017;153:1026–1039.

5. Hong G, Li Y, Yang M, et al. Gut fungal dysbiosis and altered bacterial-fungal interaction in patients with diarrhea-predominant irritable bowel syndrome: An explorative study. Neurogastroenterol Motil 2020:e13891.

6. Li Y, Hong G, Yang M, et al. Fecal bacteria can predict the efficacy of rifaximin in patients with diarrhea-predominant irritable bowel syndrome. Pharmacol Res 2020;159:104936.

7. McCoubrey LE, Elbadawi M, Orlu M, et al. Harnessing machine learning for development of microbiome therapeutics. Gut Microbes 2021;13:1–20.

8. Ghosh TS, Das M, Jeffery IB, et al. Adjusting for age improves identification of gut microbiome alterations in multiple diseases. eLife 2020;9:e50240.

9. Wu G, Xu T, Zhao N, et al. *Two Competing Guilds as a Core Microbiome Signature for Health Recovery*. Microbiology; 2022. Available at: http://biorxiv.org/lookup/doi/10.1101/2022.05.02.490290 [Accessed March 8, 2023].

10. Guo M, Wu G, Tan Y, et al. Guild-Level Microbiome Signature Associated with COVID-19 Severity and Prognosis Huffnagle GB, ed. mBio 2023;14:e03519-22.

11. Mujagic Z, Kasapi M, Jonkers DM, et al. Integrated fecal microbiome-metabolome signatures reflect stress and serotonin metabolism in irritable bowel syndrome. Gut Microbes 2022;14:2063016.

12. Dior M, Delagrèverie H, Duboc H, et al. Interplay between bile acid metabolism and microbiota in irritable bowel syndrome. Neurogastroenterol Motil 2016;28:1330–1340.

13. Shin A, Camilleri M, Vijayvargiya P, et al. Bowel functions, fecal unconjugated primary and secondary bile acids, and colonic transit in patients with irritable bowel syndrome. Clin Gastroenterol Hepatol 2013;11:1270-1275.e1.

14. Wong BS, Camilleri M, Carlson P, et al. Increased bile acid biosynthesis is associated with irritable bowel syndrome with diarrhea. Clin Gastroenterol Hepatol 2012;10:1009-1015.e3.

15. Min YW, Rezaie A, Pimentel M. Bile Acid and Gut Microbiota in Irritable Bowel Syndrome. J Neurogastroenterol Motil 2022;28:549–561.

16. Hoppel C. The role of carnitine in normal and altered fatty acid metabolism. Am J Kidney Dis 2003;41:S4-12.

17. Liu X, Liu Y, Cheng M, et al. Metabolomic Responses of Human Hepatocytes to Emodin, Aristolochic Acid, and Triptolide: Chemicals Purified from Traditional Chinese Medicines. J Biochem Mol Toxicol 2015;29:533–543.

18. Vishwanath VA. Fatty Acid Beta-Oxidation Disorders: A Brief Review. Ann Neurosci 2016;23:51–55.

19. Ponnusamy K, Choi JN, Kim J, et al. Microbial community and metabolomic comparison of irritable bowel syndrome faeces. J Med Microbiol 2011;60:817–827.

20. Mars RAT, Yang Y, Ward T, et al. Longitudinal Multi-omics Reveals Subset-Specific Mechanisms Underlying Irritable Bowel Syndrome. Cell 2020;182:1460-1473.e17.

21. Louvel D, Delvaux M, Staumont G, et al. Intracolonic injection of glycerol: a model for abdominal pain in irritable bowel syndrome? Gastroenterology 1996;110:351–361.

22. Guo F-F, Yu T-C, Hong J, et al. Emerging Roles of Hydrogen Sulfide in Inflammatory and Neoplastic Colonic Diseases. Front Physiol 2016;7:156.

23. Aggarwal S, Ahuja V, Paul J. Dysregulation of GABAergic Signalling Contributes in the Pathogenesis of Diarrhea-predominant Irritable Bowel Syndrome. J Neurogastroenterol Motil 2018;24:422–430.

24. Tap J, Derrien M, Törnblom H, et al. Identification of an Intestinal Microbiota Signature Associated With Severity of Irritable Bowel Syndrome. Gastroenterology 2017;152:111-123.e8.

25. Singer-Englar T, Rezaie A, Gupta K, et al. 1089 - A Novel 4-Gas Device for Breath Testing Shows Exhaled H 2 S is Associated with Diarrhea and Abdominal Pain in a Large Scale Prospective Trial. Gastroenterology 2018;154:S-213.

26. Moayyedi P, Quigley EMM, Lacy BE, et al. The effect of fiber supplementation on irritable bowel syndrome: a systematic review and meta-analysis. Am J Gastroenterol 2014;109:1367–1374.

27. Desai MS, Seekatz AM, Koropatkin NM, et al. A Dietary Fiber-Deprived Gut Microbiota Degrades the Colonic Mucus Barrier and Enhances Pathogen Susceptibility. Cell 2016;167:1339-1353.e21.

28. Shankar V, Homer D, Rigsbee L, et al. The networks of human gut microbe-metabolite associations are different between health and irritable bowel syndrome. ISME J 2015;9:1899–1903.

29. Nicholson JK, Holmes E, Kinross J, et al. Host-Gut Microbiota Metabolic Interactions. Science 2012;336:1262–1267.

30. Xiao L, Liu Q, Luo M, et al. Gut Microbiota-Derived Metabolites in Irritable Bowel Syndrome. Front Cell Infect Microbiol 2021;11:729346.

31. Shankar V, Agans R, Holmes B, et al. Do gut microbial communities differ in pediatric IBS and health? Gut Microbes 2013;4:347–352.

32. Shanahan F, Ghosh TS, O'Toole PW. The Healthy Microbiome—What Is the Definition of a Healthy Gut Microbiome? Gastroenterology 2021;160:483–494.

33. Fitzgerald CB, Shkoporov AN, Sutton TDS, et al. Comparative analysis of Faecalibacterium prausnitzii genomes shows a high level of genome plasticity and warrants separation into new species-level taxa. BMC Genomics 2018;19:931.

34. Heinken A, Hertel J, Acharya G, et al. Genome-scale metabolic reconstruction of 7,302 human microorganisms for personalized medicine. Nat Biotechnol 2023. Available at: https://www.nature.com/articles/s41587-022-01628-0 [Accessed January 23, 2023].

35. Morton JT, Aksenov AA, Nothias LF, et al. Learning representations of microbe-metabolite interactions. Nat Methods 2019;16:1306–1314.

36. Baranwal M, Clark RL, Thompson J, et al. Recurrent neural networks enable design of multifunctional synthetic human gut microbiome dynamics. eLife 2022;11:e73870.