THESIS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

Mining human gut microbial metabolism through *in vitro* and *in silico* approaches

PROMI DAS

Department of Biology and Biological Engineering
CHALMERS UNIVERSITY OF TECHNOLOGY
Gothenburg, Sweden 2019

Mining human gut microbial metabolism through *in vitro* and *in silico* approaches

PROMI DAS

ISBN 978-91-7597-875-8

Doktorsavhandlingar vid Chalmers tekniska högskola
Ny serie nr 4556
ISSN 0346-718X

Division of Systems and Synthetic Biology
Department of Biology and Biological Engineering
Chalmers University of Technology
SE - 412 96 Gothenburg
Sweden
Telephone + 46 (0) 31 772 1000

Cover: Schematic representation of thesis title.

Printed by Chalmers Reproservice
Gothenburg, Sweden 2019

Mining human gut microbial metabolism through *in vitro* and *in silico* approaches

Promi Das
Department of Biology and Biological Engineering
Chalmers University of Technology

Abstract

The human gut microbiome is a consequence of mutual co-evolutionary interaction between the eukaryotic and prokaryotic parts of the mammalian holobiont. Based on the environmental and dietary inputs, there is a succession of microorganisms living inside the human colon. They have evolved to perform metabolic tasks that are not possible by the human host — for example, they breakdown complex polysaccharides and produce bioactive molecules such as short-chain fatty acids. They have the potential to transform human generated metabolites (e.g., primary bile acids) to signaling compounds such as secondary bile acids. They also produce several of B-vitamins, which otherwise human host derive through dietary means. Cognate receptors in various host cells could sense these bioactive metabolites and contribute to a wide variety of physiological function through signaling system in the host. An imbalance between the microbial activity and their effect on the host system could lead to the development of metabolic diseases.

Provided the critical role of gut microbial metabolism, this thesis presents the evaluation of metabolic genes of gut microbiota such as bile acid, vitamin, and short-chain fatty acid metabolism using metabolic reconstructions and bioinformatics analysis in different states of health. Fecal metagenomes of subjects with inflammatory bowel diseases, type 2 diabetes and malnutrition were analyzed for such functional analyses. Furthermore, abundant gut microbial species were characterized to study their growth and metabolism in *in vitro* co-cultures using network analysis. The findings explained here show the gut microbial metabolic diversity in various cohorts and conditions. It also includes a discussion on the challenges and future perspectives in a broader context of its potential application. The efforts undertaken in this work aims to inspire how the interplay between gut microbial metabolism and the host health status could contribute to the overall well-being of an individual.

Keywords: gut, metabolism, short chain fatty acids, bile acids, vitamins, metagenomics, metatranscriptomics, metabolomics, co-occurrence network, *in vitro* co-cultures.

List of publications

This thesis is based on the work in the following publications:

Paper I – *In vitro* co-cultures of human gut bacterial species as predicted from co-occurrence network analysis.
Promi Das, Boyang Ji, Petia Kovatcheva Datchary, Fredrik Bäckhed, and Jens Nielsen.
PLoS ONE (2018)
https://doi.org/10.1371/journal.pone.0195161

Paper II – Gut microbiota dysbiosis is associated with malnutrition and reduced plasma amino acid levels: Lessons from genome-scale metabolic modeling.
Manish Kumar*, Boyang Ji*, Parizad Bababei*, **Promi Das**, Dimitra Lappa, Girija Ramakrishnan, Todd E. Fox, Rashidul Haque, William A. Petri, Fredrik Bäckhed, and Jens Nielsen.
* Authors contributed equally
Metabolic Engineering (2018)
https://doi.org/10.1016/j.ymben.2018.07.018

Paper III – Uniform randomized sampling of microbial communities.
Parizad Babaei, **Promi Das**, Adil Mardinoglu, and Jens Nielsen.
In review (2019)

Paper IV – Metagenomic analysis of bile salt biotransformation in the human gut microbiome.
Promi Das, Simonas Marcišauskas, Boyang Ji, and Jens Nielsen.
In review (2019)

Paper V – Metagenomic analysis of microbe-mediated vitamin metabolism in the human gut microbiome.
Promi Das, Parizad Babaei, and Jens Nielsen.
BMC Genomics (2019)
https://doi.org/10.1186/s12864-019-5591-7

Additional research articles not included in the thesis:

Paper VI – Carbohydrate active enzymes are affected by diet transition from milk to solid food in infant gut microbiota.
Ling-Qun Ye**, Promi Das**, Peishun Li, Boyang Ji, and Jens Nielsen.
Submitted (2019)

Contributions

I. Designed and performed the *in vitro* experiments and analyzed the experimental data.
Drafted, edited and submitted the paper.

II. Designed and performed the *in vitro* experiments and analyzed the experimental data.
Co-participated in drafting the paper.

III. Designed and performed the *in vitro* experiments and analyzed the experimental data.
Co-participated in drafting the manuscript.

IV. Designed and performed the *in silico* analysis of metagenomics and metabolomics datasets.
Drafted, edited and submitted the manuscript.

V. Designed and performed the *in silico* analysis of metagenomics and metatranscriptomics datasets.
Drafted, edited and submitted the paper.

VI. Co-designed and co-participated in the *in silico* analysis of metagenomics datasets.
Co-participated in drafting the manuscript.

Preface

This dissertation serves as a partial fulfillment of the requirements to obtain the degree of Doctor of Philosophy at the Department of Biology and Biological Engineering at the Chalmers University of Technology. The Ph.D. studies were carried out between June 2015 and May 2019 at the division of Systems and Synthetic Biology, under the supervision of Jens Nielsen. The research was funded by the European Commission FP7 project METACARDIS, Knut and Alice Wallenberg Foundation, the Chalmers Foundation, and the Bill & Melinda Gates Foundation.

Promi Das

May 2019

Table of Contents

ABBREVIATIONS

aFMT	Autologous-Fecal Microbiota Transplantation
ATP	Adenosine Tri-Phosphate
BA	Bile Acid
Bai	Bile Acid Inducible
BSBGs	Bile Salt Biotransformation Genes
BSBPs	Bile Salt Biotransformation Proteins
CA	Cholic Acid
CD	Crohn's Disease
CDCA	Chenodeoxycholic Acid
CVD	Cardio Vascular Disease
EMPP	Estimated Maximal Production Potential
FODMAPs	Fermentable Oligo-, Di-, Mono-Saccharides and Polyols
FXR	Farnesoid X Receptor
GEMs	Genome-Scale Metabolic Models
GIT	Gastrointestinal Tract
HUMANN2	HMP Unified Metabolic Analysis Network 2
IBD	Inflammatory Bowel Disease
IBS	Inflammatory Bowel Syndrome
KEGG	Kyoto Encyclopedia Of Genes and Genomes
KO	KEGG Orthology
MEDUSA	Metagenomic Data Utilization and Analysis
PCR	Polymerase Chain Reaction
rRNA	Ribosomal Ribonucleic Acid
SCFAs	Short-Chain Fatty Acids
SNPs	Single Nucleotide Polymorphisms
T2D	Type 2 Diabetes
TGR5	Takeda G-Protein-Coupled Receptor 5
TMAO	Trimethylamine N-Oxide
UC	Ulcerative Colitis
UDCA	Ursodeoxycholic Acid
YCFA	Yeast Extract, Casitone, Fatty Acid
YCFAGD	Yeast Extract, Casitone, Fatty Acid Glucose Disaccharide
YCGD	Yeast Extract, Casitone, Glucose Disaccharide
YCGMS	Yeast Extract, Casitone, Glucose Maltose Starch

To my mother, Pampa.

1. BACKGROUND

1.1. EVOLUTION OF THE GUT MICROBIOME

Advances in the high-throughput sequencing technologies and bioinformatics tools and methods have propelled the field of human microbiome science. These technological resources have been crucial in systematic mining of human gut microbial ecosystem, despite low cultivation practice of most gut microorganisms. Qualitative and quantitative research of microbial communities through 16S rRNA marker-based sequencing revealed answers by addressing “Who is there?” and “How many are there?” Eventually, microbiome researchers became interested in questioning the functionality and interrelationship of microbial communities from whole genome sequencing, such as, “What do they do?” and “How do they do?” The evolution and major discoveries in microbiome research are listed in Table 1. Currently, the question has evolved into “How the microbiome could be modified to modulate our health?”

Table 1. Key advancements that set the foundation of microbiome research.

YEAR	ADVANCEMENTS	REFERENCES
Early 1680s	Difference between the microbes found in samples from the oral cavity than that from the feces	Antoine van Leuwenhoek
1864	Germ Theory of Disease	Louis Pasteur
1876	Koch’s postulates	Robert Koch
1901	Probiotics for good health	Eli Metchnikoff
1909	Effect of diet on the composition of gut microbiota and its impact on the host’s health	Arthur I Kendall
1950	Normal flora under different states of health	Rene Dubos
1977	Ribosomal RNA genes as molecular markers for phylogenetic studies	Carl Woese and Norman Pace
1979	Shotgun sequencing method	(Gardner *et al.*, 1981)
1983	Polymerase chain reaction	Kary Mullis
1990	Start of large-scale sequencing trials	NIH
2006	Metagenomic analysis of the human gut microbiome	(Gill *et al.*, 2006)
2007	Human Microbiome Project	(Turnbaugh *et al.*, 2007)
2009	Study of the gut microbiome in lean and obese individuals	(Turnbaugh *et al.*, 2009)
2010	Microbiome gene catalog	(Qin *et al.*, 2010)
2011	Enterotypes of the human gut microbiome	(Arumugam *et al.*, 2011)
2014	Human Microbiome Project 2	(The Integrative HMP Research Network Consortium, 2014)
2015	Precision Medicine	(Zeevi *et al.*, 2015)

1.2. THE HUMAN GUT MICROBIOME

- **Human Physiology and Microbial Ecology of Gastrointestinal Tract**

The human gastrointestinal (GI) tract can be visualized into six distinct anatomical regions, extending from stomach to the rectum. These six anatomical regions are stomach, duodenum, jejunum, ileum, cecum, and rectum. On a broad level, small intestine consists of duodenum, jejunum, and ileum; while large intestine consists of cecum and rectum. The physiochemical features in each of these regions differ and thereby influences the colonization, composition, and abundance of the microbial species. These physiochemical variations are mostly attributed to host secretions (e.g., digestive enzymes, bile salts, hydrochloric acid, and mucus), availability of diet-derived metabolites, local pH, redox potential, and transit rates of the luminal content (Savage, 2003). The microbial load is low (less than 10^4 organisms per gram of stool) in the upper GI tract (i.e., stomach, duodenum, and jejunum) and gradually increases (greater than 10^7 organisms per gram of stool) towards the lower GI tract (i.e., ileum, cecum, and rectum) (Figure 1). In the upper GI tract, acid stress in the stomach, the presence of the bile acids (BAs), pancreatic enzymes in the duodenum, and the rapid wash out of the cells by the fast flow of food are the primary reasons for impeding the microbial colonization. However, in the lower GI tract, increased availability of nutrients caused by reduced transit times makes it a favorable environment for the microbes. Hence, there is a high abundance of microbial species in the colon (Savage, 2003).

The intestinal lumen is separated by the epithelial surface from the internal body environment and forms a continuum with the external environment. The epithelial surface is a single layer of intestinal cells that is covered by a mucus layer. The microbial composition present in the feces and luminal content differs significantly than the ones associated with the mucus layer (Eckburg *et al.*, 2005; Frank *et al.*, 2007).

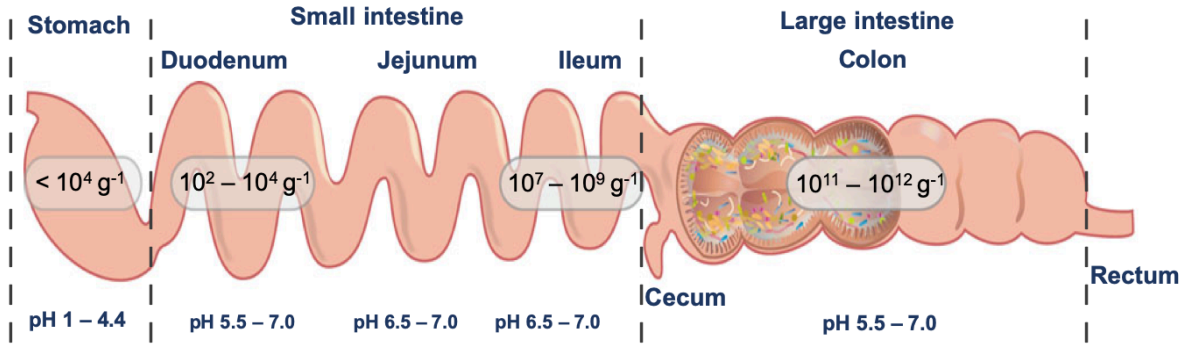

Figure 1. Schematic representation of bacterial density and the environmental pH in different anatomical regions along the length of the gastrointestinal tract. Adapted from (Kovatcheva-Datchary et al., 2013).

- **Establishment of Gut Microbiota in the GI tract**

The colonization of an infant gut microbiome begins at birth when exposed to an external environment. Microbial species that are introduced to the GI tract of the neonate are a consequence of contact between the microbes found in the skin/vagina and surrounding environment of the mother (Dominguez-Bello *et al.*, 2010). The development of the neonatal gut microbiota begins from a simple community and transforms into a climax community through a gradual and dynamic process for over two years of life. This process is determined by multiple factors, such as the type of feeding, the health status of the baby, mode of delivery, and maternal microbiota (Bäckhed *et al.*, 2015). Due to the positive redox potential in the neonatal gut, it is populated by facultative anaerobes. In fact, as the infant starts to consume solid food, the existing microbiota creates an anoxygenic condition, which becomes favorable for the proliferation of anaerobic species (Wall *et al.*, 2009). At two years of life, the succession of microbial composition reaches a state which is similar to that found in an adult intestine (Koenig *et al.*, 2011). Furthermore, the microbial structure of a matured or stabilized microbiome is shaped by diet, colonization history, type of antibiotic medication, and host immune system and other stressful factors that play a vital role in health and disease later in life (Figure 2) (Yatsunenko *et al.*, 2012).

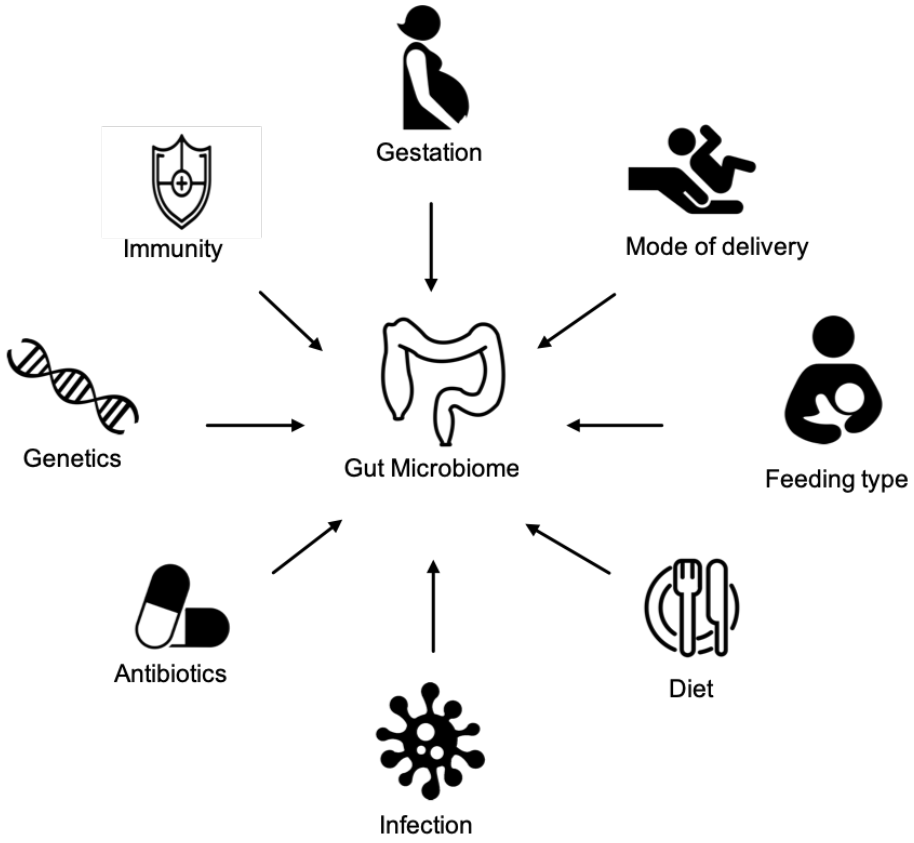

Figure 2. Involvement of diverse factors in association with the gut microbiome.

1.3. FACTORS THAT SHAPE HUMAN GUT MICROBIOTA

- **Medication**

Human gut microbiota is a reservoir of antibiotic resistance gene encoded in the genomes of microorganisms, termed resistome (Wright, 2007). In the last few decades, the use of antibiotics have posed severe threats to the public health (Laxminarayan *et al.*, 2013). These genes have been detected in the human gut microbiota of inhabitants from low-income, remote, and industrialized urban areas all over the world. Most of the studies indicate that the emergence of antibiotic resistance is driven by the rate of antibiotic consumption and misuse. The increased selective pressure of antibiotics accounts for the differences in the diversity and abundance of these genes across countries (Salyers *et al.*, 2004; Bartoloni *et al.*, 2009; Sommer *et al.*, 2009; Forslund *et al.*, 2014; Pehrsson *et al.*, 2016). They persist in the gut ecosystem for many years even after the termination of antibiotic treatment (Hu *et al.*, 2013; Clemente *et al.*, 2015).

On the other hand, non-antibiotic drugs such as anti-diabetics (e.g., metformin) (Forslund *et al.*, 2015), proton pump inhibitors (Imhann *et al.*, 2016; Jackson *et al.*, 2016), non-steroidal anti-inflammatory drugs (Paño-Pardo *et al.*, 2016), and antipsychotics (Flowers *et al.*, 2017) have been associated with changes in the composition of gut microbiome. In fact, a similar finding was evidenced in a large cohort where medication intake explained the most significant variation with other covariate-microbiota associations in altering the gut microbial composition (Darzi *et al.*, 2016). Systematic profiling of associations between human targeted drugs and gut bacteria has revealed the importance of accounting for potential medication-related confounders in future gut microbiome-disease studies. A comprehensive resource of pharmaceutical drug actions on the microbiome with further validation of the drug–microbe network facilitates more in-depth clinical and mechanistic studies, thus ultimately improving the drug development (Maier *et al.*, 2018).

Though antibiotic treatment can disrupt the commensal microbiota, reconstitution of human gut microbiome is often slow and incomplete (Jernberg *et al.*, 2007; Dethlefsen *et al.*, 2008; Dethlefsen and Relman, 2011). While probiotics have been suggested to restore the gut microbial composition from an antibiotic-induced dysbiosis, no substantial evidence has been found to support the recommendation. However, *in vivo* examination of the extent of probiotics on gut mucosal colonization was shown to delay the reconstitution of indigenous microbiome following an antibiotic exposure (Suez *et al.*, 2018). Alternately, autologous-fecal microbiota transplantation (aFMT) succeeded to induce rapid and near-complete gut mucosal microbiome re-colonization, and reversion of the human gut transcriptome as that of the pre-antibiotic configuration (Taur *et al.*, 2018).

- **Genetics**

As part of the microbiome is known to be heritable, host genetics has been an essential factor in determining the gut microbial composition (Turnbaugh *et al.*, 2009; Goodrich *et al.*, 2014, 2016; Turpin *et al.*, 2016; Xie *et al.*, 2016). Despite different phylogenetic profiles among different populations, a “core microbiome” at a functional level was identified among the various individuals (Turnbaugh *et al.*, 2009). However, the extent to which the host genetics determine the heritability remained an unsolved question. While several studies have found associations between individual bacterial taxa or pathways and host single nucleotide polymorphisms (SNPs) (Blekhman *et al.*, 2015; Bonder *et al.*, 2016; Goodrich *et al.*, 2016; Turpin *et al.*, 2016; Wang *et al.*, 2016; Xie *et al.*, 2016), these associations turned out to be statistically insignificant after multiple testing corrections (Kurilshikov *et al.*, 2017).

From re-analysis of the microbiome and genetic data of 2252 twins (TwinsUK cohort), the average heritability of the gut microbiome was estimated to be around 1.9 to 8.1%, suggesting only a small number of bacterial taxa are likely to be heritable. Furthermore, genetic ancestry or individual SNPs of the subjects were not significantly associated with their microbiome (Rothschild *et al.*, 2018). Contrarily, direct evidence supporting the notion that environmental factors predominantly shape gut microbial composition demonstrated a significant gut microbiome similarity among genetically unrelated individuals who share a household (Rothschild *et al.*, 2018). In the same study, no considerable microbiome similarity was reported among genetically related individuals who do not have a history of household sharing. It was strongly anticipated that studies with larger sample sizes could identify heritable taxa and SNP associations in the future. Nonetheless, it is unlikely to change the inference that environmental factor predominantly shapes the microbiome composition.

- **Diet**

With the help of compositional and functional tools, several studies have investigated the role of microbes in mediating diet-induced effects on host physiology. By emphasizing on diet components, the research question focused on how a particular component of each dietary pattern could influence the abundance and functional capacity of different types of bacteria in the gut, and how those gut microbiota-derived metabolites could affect human health. Based on the recent observational and interventional studies, the impact of some dietary compounds on human gut microbiome and health status is summarized in Table 2.

Table 2. Influence of dietary components on human health mediated through gut microbial metabolism.

DIET	EFFECT ON THE GUT MICROBIOME	EFFECTS ON THE HOST	REFERENCES
Fiber	Increased microbial diversity and SCFA production	Reduced CVD and T2D	(Battaglioli and Kashyap, 2018; Sasaki *et al.*, 2018; Zhao *et al.*, 2018)
FODMAP	High FODMAP diet decreased the abundance of gas-consuming bacteria, and low FODMAP diet increased Actinobacteria	Reduced symptoms of IBS	(Gibson, 2017; McIntosh *et al.*, 2017)
Vegan	Modest effects on the composition of gut microbiota, but significant differences in metabolome compared to omnivores diet	High gut microbial-mediated metabolites in the plasma compared to omnivores diet	(Orlich *et al.*, 2013; Wu *et al.*, 2016; Mihrshahi *et al.*, 2017)
Cheese	Increased Bifidobacteria and reduced Bacteroides and Clostridia	Reduced production of TMAO and increased output of SCFA	(Montel *et al.*, 2014; Zheng *et al.*, 2015)
Polyphenols	Increased Akkermansia, Bifidobacteria, Lactobacillus, and butyrate producers. Reduced LPS producers	Reduced metabolic markers of CVD risk and metabolic syndrome	(Etxeberria *et al.*, 2015; Moreno-Indias *et al.*, 2016; Ozdal *et al.*, 2016)

1.4. STRATEGIES TO STUDY GUT MICROBIOTA

The availability of high-throughput experimental techniques has allowed a more in-depth study on the composition and functional potential of a microbial community. Broadly, they are of two types (i). Compositional techniques, which analyze the number and kind of microbe present in a community and (ii). Functional techniques, which examine the community level functions executed by the bacteria present in a community. Table 3 lists the techniques used in the field of microbiome science.

Table 3. Techniques used to study the human gut microbiota from isolated samples.

COMPOSITIONAL			
TECHNIQUE	**DESCRIPTION**	**ADVANTAGES**	**DISADVANTAGES**
Culture-based	Isolation and growth of bacteria in a selective growth medium	Semi-quantitative	Labor-intensive, only a minor fraction of bacteria is cultivable
Cloned 16S rRNA gene sequencing	Cloning and sequencing of 16S rRNA gene	Quantitative	PCR-bias, cloning-bias, laborious, expensive
Temperature gradient gel electrophoresis	Gel separation of 16S rRNA amplicons using temperature	Semi-quantitative	PCR-bias
Denaturing gradient gel electrophoresis	Gel separation of 16S rRNA amplicons using denaturants	Semi-quantitative	PCR-bias
Fluorescence in-situ hybridization with flow cytometry	Hybridization of 16S rRNA regions with fluorescently labeled oligonucleotide probes	Semi-quantitative	Limited to known species as sequence information is required
Terminal restriction fragment length polymorphism	Restriction enzyme-digested fragments of 16S rRNA amplicon	Semi-quantitative	PCR-bias, low resolution
DNA microarray	Hybridization of nucleotide regions with fluorescently labeled oligonucleotide probes	Semi-quantitative	PCR bias, the possibility of cross-hybridization, lack of sensitivity
16S rRNA amplicon sequencing	Massive parallel sequencing of the16S rRNA amplicons	Quantitative	PCR-bias
Whole metagenomic sequencing	Large-scale parallel sequencing of the entire genome fragments	Quantitative	Computationally intensive analysis
FUNCTIONAL			
Metatranscriptomics	Parallel measurement of RNA levels in a mixture	Quantitative	Computationally intensive analysis
Metabolomics	Parallel measurement of metabolites levels in a mixture	Quantitative	Computationally intensive analysis
Metaproteomics	Parallel measurement of proteins levels in a mixture	Quantitative	Computationally intensive analysis

In systems biology, mathematical models such as genome-scale metabolic models are often used to capture information from complex systems in a concise way. They help in comprehending important biological reactions and enable the simulation of metabolism and other biological processes involved in cellular growth (Nielsen, 2017). Testing the metabolic capabilities of an organism under different environmental conditions and parameters has allowed robustness in simulations (O'Brien *et al.*, 2015), which is otherwise not feasible in human studies. These models provide top-down, systems biology analysis at holistic genome-scale level, with bottom-up, systems biology modeling investigation. They could be used in predicting the metabolic potential of the gut microbiota in a complex ecosystem. Reconstructing GEMs for each species of a community could enable interpretation of their metabolic interactions with other members (Sen *et al.*, 2019). One potential application of gaining knowledge from these simulations would be rational designing of synthetic microbial communities. In simplest terms, the workflow for reconstruction of a genome-scale metabolic model has been summarized in Table 4.

Table 4. General steps involved in reconstruction of a modeling framework for analysis and predictions.

ACTION	DESCRIPTION
Build	For a target organism, known metabolic reactions, genome annotation, and biochemical characterization from published articles are retrieved as raw materials.
Structure	For each metabolic reaction, stoichiometric coefficients for each of the substrate and product involved; reversibility; location in a cellular compartment; enzyme metabolic gene and its genomic location are assembled into pathways, metabolic sectors, and finally into genome-scale networks.
Model	These stoichiometric equations are converted into a numerical matrix.
Constraints	These are mathematical equations imposing mass balance constraints in a network, such as equating the consumption and production rates of every metabolite; the reaction flux within lower and upper bound at steady state.
Solution space	A solution space with allowable flux distributions that satisfy the imposed constraints is achieved.
Objective function	It is a mathematical representation of a biological process (e.g., growth, ATP consumption) which is optimized in specific environmental and physiological context during conversion of a network reconstruction to a model
Model evaluation	The accuracy, consistency, and incomplete pieces of information in a model are checked to ensure informative and qualitative outcome for e.g., growth capability.
Prediction	Next to gap filling and curation, quantitative phenotype predictions are performed. Flux state(s) that maximize the assumed objective are predicted using flux balance or flux variability analysis.
Validation	These predictions are then iteratively compared to experimental measurements until the model evaluation is satisfactory.

1.5. METABOLIC FUNCTIONS OF GUT MICROBIOTA

- **Carbohydrates and Fermentation**

Carbohydrates are a significant component of human diet, and not all of them are metabolized equally. From a microbial and human host perspective, the division in the availability of carbohydrates to both microbes and host serve different purposes. While mono- and disaccharides in the diet are readily absorbed in the small intestine, a fraction of the dietary carbohydrates is resistant or non-digestible by the human host. These carbs reach the proximal colon and serve as active substrates for metabolic breakdown by the gut microbiota (Macfarlane *et al.*, 1992). Dietary resistant oligosaccharides, undigested (resistant) starch and plant cell-wall components such as cellulose, hemicellulose, and pectin are the primary sources of microbial fermentation in the gut. These substrates are further metabolized by the microbes to short-chain fatty acids (SCFAs) primarily acetate, lactate and butyrate and thus lowering the environmental pH (Cummings *et al.*, 1987; Macfarlane *et al.*, 1992). The contribution by the gut microbe's enzymatic potential of carbohydrate degradation has been well established. For example, the human host produces only 17 carbohydrate-active enzymes, whereas gut bacterial species possess more than 200 carbohydrate-active enzymes (Cantarel *et al.*, 2012). A generalized workflow of carbohydrate degradation and fermentation is presented in Figure 3.

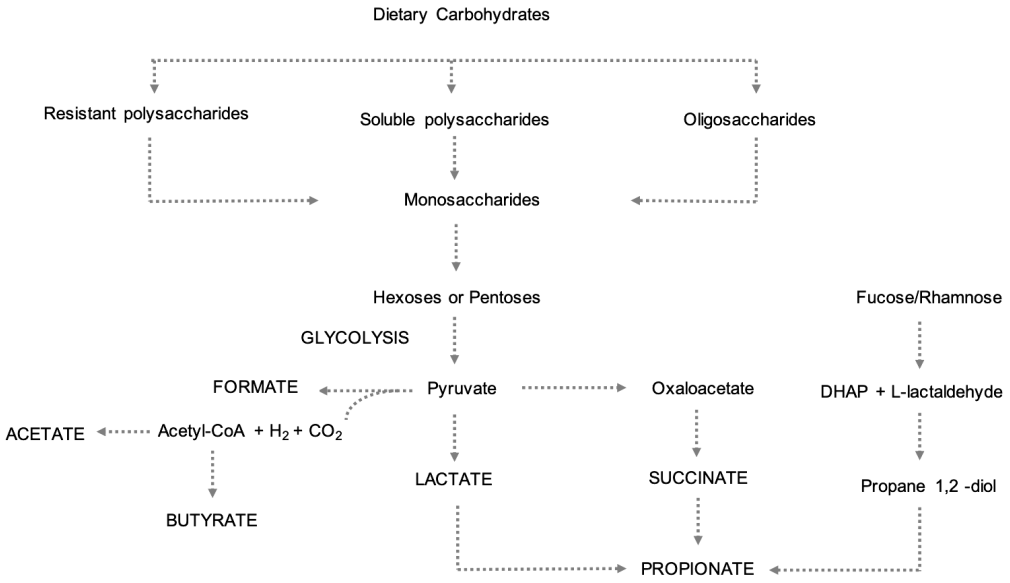

Figure 3. Microbial fermentation of dietary carbohydrates to short-chain fatty acids in the human colon.

The fate of the major SCFAs produced is dependent on their location (Koh *et al.*, 2016). Acetate produced in the gut enters the liver through the portal circulation and eventually is used for cholesterol synthesis. However, a significant fraction passes the liver and becomes available to peripheral tissues as an energy source. Propionate is used for gluconeogenesis in the liver and inhibits cholesterol synthesis from acetate (Wolever *et al.*, 1991). Propionate thus represents a

source of glucose for the host. It is also used for intestinal gluconeogenesis (De Vadder *et al.*, 2014). Most of the butyrate that is absorbed is utilized by colonocytes (Cummings *et al.*, 1987), acting as their primary energy source. Relatively little butyrate reaches the circulatory system (Cummings *et al.*, 1987).

- **Proteins and Fermentation**

Proteins are the second major component of human diet and are metabolized to different extent (e.g., animal and plant proteins) by human digestive enzymes and colonic gut microbiota. Dietary protein assimilation is impeded by gastric acid in the upper GIT, which results in a considerable amount of protein digest traversing from proximal to distal colon (Evenepoel *et al.*, 1998). As the availability of carbohydrate decreases in the distal colon, protein degradation takes place near neutral pH, where proteins and amino acids become the primary source of energetic inputs for the gut microbiota (Macfarlane *et al.*, 1992). Besides dietary proteins, endogenous material such as pancreatic enzymes, mucus, and exfoliated epithelial cells also serve as protein sources for microbial fermentation. A variety of secondary metabolites, with a range of host effects, are produced from the breakdown of amino acids in the gut (Table 5). Besides SCFAs, hydrogen, carbon dioxide, indoles, decarboxylated amino acids, branched chain fatty acids, ammonia, amines, N-nitroso compounds, phenolic compounds, and sulfides are also produced in the colon (Cummings and Macfarlane, 1991).

Table 5. Influence of dietary-protein derived microbial-mediated metabolic end products on human health.

PRODUCT	SOURCE	EFFECT	RESULT	REFERENCES
p-cresol	Fermentation of aromatic amino acids	Damages DNA of colonic epithelial cells	Cancer development, Kidney failure	(Vanholder *et al.*, 1999; Toden *et al.*, 2005)
Hydrogen sulfide	Fermentation of sulfur-containing amino acids, by the action of sulfate-reducing bacteria that utilize hydrogen and sulfate; degradation of taurine	Blocks the oxidation of butyrate in colonic epithelial cells by inhibiting the activity of acyl-CoA-dehydrogenase goblet cell depletion; superficial ulceration, and causes genomic DNA damage	Ulcerative colitis and colon cancer development	(Pitcher and Cummings, 1996; Magee *et al.*, 2000; Attene-Ramos *et al.*, 2006)
Indole acetic acid and indole-3-propionic acid	Metabolism of tryptophan generates indole derivatives	Ligands for the aryl hydrocarbon receptor	Protective against colitis, and involved in reducing inflammation in the central nervous system	(Lamas *et al.*, 2016; Rothhammer *et al.*, 2016)

Tryptamine	Decarboxylation of dietary tryptophan	A β-arylamine neurotransmitter, which induces ion secretion in gut epithelial cells	Alters overall serotonin dynamics in the host affect gastrointestinal transit time.	(Lamas *et al.*, 2016; Rothhammer *et al.*, 2016)

- **Bile acids and Biotransformation**

Human host-mediated primary bile acids, e.g., chenodeoxycholic acid (CDCA) and cholic acid (CA), are produced in the liver. After meal consumption, these BAs are released into the duodenum to facilitate the emulsification of dietary lipids. Then, they are re-absorbed through the ileal active transporter (generally known as IBAT or ASBT) and re-circulated to the liver through portal vein blood. This entire process is called enterohepatic circulation and preserves around 95% of the bile acid pool (Hofmann, 2011).

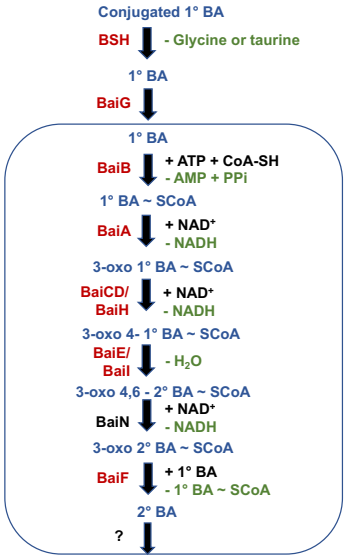

Figure 4. Schematic representation of bile salt biotransformation in the human colon. Co-substrates and by-products have been highlighted in black and green color respectively. Conjugated primary bile salts are first deconjugated to primary bile acids by BSH enzyme. Then they enter the cell via BaiG membrane transporter, where the blue outline denotes the cell membrane of a bacterium. The primary bile acid gets transformed to secondary bile acid through series of reactions. The transporter required to export the secondary bile acid outside of the cell has not been determined. Enzymatic proteins that were studied have been highlighted in red color

The remaining 5% of the bile acid pool is prevented from active reuptake due to a change in chemical structure (e.g., removal of glycine or taurine from the primary bile acids). Microbial metabolism of BA begins with the deconjugation of the conjugated primary BA, followed by dehydroxylation into secondary BAs in the large intestine. This multi-step pathway of dehydroxylation (Figure 4) is encoded in a bile acid-inducible (bai) operon as identified in

Clostridium clusters XIVa and XI, and Eubacterium from Firmicutes (Ridlon *et al.*, 2006). These secondary BAs are a critical class of bioactive compounds that can regulate gut microbial composition, and several host metabolic processes via activation of cognate receptors (like farnesoid X receptor (FXR) (Wahlström *et al.*, 2016; Molinaro *et al.*, 2018), Takeda G-protein-coupled receptor 5 (TGR5)) present in the gut and the liver (Holmes *et al.*, 2012; De Aguiar Vallim *et al.*, 2013). Secondary bile acids such as deoxycholic acid (DCA), lithocholic acid (LCA), ursodeoxycholic acid (UCDA) are known to bind BA receptors such as TGR5, VDR, and exert anti-inflammatory effects in the colon (Hylemon *et al.*, 2009; Lefebvre *et al.*, 2009; Ogilvie and Jones, 2012; Ridlon *et al.*, 2014; Zhou and Hylemon, 2014; Martinot *et al.*, 2017; Pols *et al.*, 2017). Their functional role in signaling could result in a beneficial outcome by reducing inflammation in inflammatory bowel diseases.

- **Vitamin Biosynthesis and Utilization**

Beside the role of gut microbiota in nutrient digestion and energy recovery, they are a potential source of vitamins. Microbes in the gut are known to synthesize K_2 and B-vitamins such as biotin, cobalamin, folate, nicotinic acid, pantothenic acid, pyridoxine, riboflavin, and thiamine (Hill, 1997). While dietary vitamins are absorbed in the small intestine, the majority of the microbial-produced vitamin uptake takes place in the colon (Ichihashi *et al.*, 1992; Said and Mohammed, 2006). Several of the B-vitamin synthetic pathways require the presence of other B-vitamins as "coenzyme requirement" in a process called "vitamin cannibalism" as shown in Figure 5.

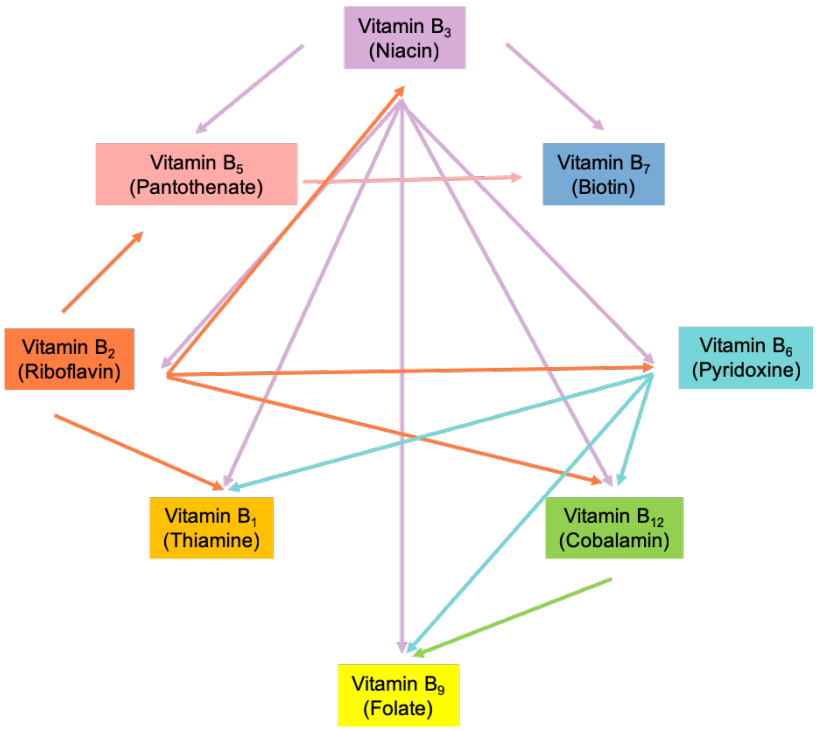

Figure 5. Schematic representation of the vitamin usage for synthesis of other vitamins. The direction of an arrow denotes the direction of vitamin synthesis.

1.6. CURRENT AND FUTURE SOLUTIONS TO MANIPULATE THE GUT MICROBIOME

Prebiotics

Changing diets have led to remarkable changes in the structure and composition of the gut microbiome within a definite period. For example, swapping food for two weeks between African Americans and rural Africans have led to significant production of butyrate in the African Americans consuming a rural diet (O'Keefe *et al.*, 2015). The critical element that has led to selective growth of beneficial colonic microorganisms is dietary fiber (i.e., prebiotics). The development of a community in the presence of these substrates depends on the carbohydrate structure, the enzymatic potential of the microbes in this community to degrade and utilize them, and their ability to tolerate the change in the environmental conditions (such as low pH resulting from fermentation's acidic products). However, the selective growth of these species is continued as long as they are fed with their favorable substrates (Bindels *et al.*, 2015).

Probiotics and Fecal Microbiota Transplantation

In a wide range of pathologies and a substantial number of observational studies, the benefits of probiotics consumption has been shown by two ways, either through significant statistical results or statistically insignificant but with improved clinical phenotypes (Thompson *et al.*, 2017; Veronese *et al.*, 2018). There have been debatable concerns with the administration of live microorganisms for not being able to establish and exert an effect on the commensal community (Kristensen *et al.*, 2016; Walter *et al.*, 2018). Conversely, probiotics are resorted to post-antibiotic induced dysbiosis to restore the gut microbial balance (Ekmekciu *et al.*, 2017). However, experimental evidence in both mice and humans show that probiotics delay the extent of reconstitution of the indigenous microbiome. Alternatively, auto-FMT is a promising approach to achieve a rapid and near complete microbiome reconstitution. Furthermore, the human gut transcriptome is also reversed towards homeostatic reconfiguration (Suez *et al.*, 2018; Taur *et al.*, 2018).

Precision Medicine

Following the mechanistic validations for the partial correlations between the mucosal and stool microbiome, human subjects have exhibited variability in probiotic colonization in two ways, with either a permissive (responsive) or resistant (non-responsive) pattern (Zmora *et al.*, 2018). Given the diversity in the gut microbiota of subjects with person-specific, country-specific, precision medicine could provide promising results (Shapiro *et al.*, 2017). Based on the clinical and microbiome profile of an individual, personalized diets could be designed through data-driven machine learning algorithms. In connection to this, tailoring of personalized dietary intervention has successfully modified high levels of post-prandial blood glucose and further metabolic consequences in T2D subjects (Zeevi *et al.*, 2015). A similar approach of personalized dietary interventions for a variety of metabolic, inflammatory and complex diseases could be valuable to evaluate the nutritional effects of diet on disease development and progression. This could be of practical value as it caters sufficient evidence to the scheme of a clinical decision process.

2. RESEARCH CHAPTERS

The scientific evolution of the thesis:

Chapter I

- Analysis of the system – microbe-microbe interaction using network analysis and *in vitro* fermentation.

Chapter II

- Analysis of the system – metabolic potential of microbial species in a complex community using genome-scale models (GEMs) in healthy and diseased subjects.
- Analysis of the system – microbial interactions in a simplified community using GEMs.

Chapter III

- Analysis of the system – Functional analysis of microbe-mediated bile acid metabolism with bioinformatic tools and omics-data in healthy and diseased subjects.

Chapter IV

- Analysis of the system – Functional analysis of microbe-mediated vitamin metabolism with bioinformatic tools and omics-data in diverse cohorts at a global level.

Chapter I – Co-occurrence Network Analysis and *in vitro* Cultivation

PAPER I – *In vitro* co-cultures of human gut bacterial species as predicted from co-occurrence network analysis

OBJECTIVES

This study aimed:

- To perform co-occurrence network analysis of microbial community data and identify non-random co-occurrence patterns
- To conduct *in vitro* co-culture experiments between the human gut bacterial species to investigate the growth and metabolic outcome of pairwise cultures

MOTIVATION

Ecological studies using interaction networks and their topologies have revealed that pairs of interacting partners determine functionality and robustness in a community. Thereby, it serves as a fundamental unit for understanding the productivity and dynamics of a community. However, these networks have not been validated adequately to reveal potential interactions (positive or negative) between gut bacterial species due to technical limitations and complex interpretation of the data (Berry and Widder, 2014). Hence, co-occurrence networks were created from microbial sequencing data to identify non-random co-occurrence patterns between bacterial species of a microbial community.

ANALYSIS, RESULTS, AND DISCUSSION

I. Co-occurrence network analysis of microbial community data

Based on the analysis of 782 human shotgun metagenomes from four different studies, Sweden (Karlsson *et al.*, 2013), Europe (Qin *et al.*, 2010), USA (Turnbaugh *et al.*, 2007), and China (Qin *et al.*, 2012), species abundances were computed through MEDUSA (Karlsson *et al.*, 2014). For the top four abundant genera, co-occurrence networks were created and compared with two existing methods: i) Spearman correlation (Spearman, 1904), which has been used for non-parametric statistical testing to measure correlation, and ii) SparCC method (Kurtz *et al.*, 2015), which has been known to offer high-precision detection of linear relationships in a compositional dataset. Both the tested methods yielded similar results, and the network generated by Spearman correlation is shown in Figure 6.

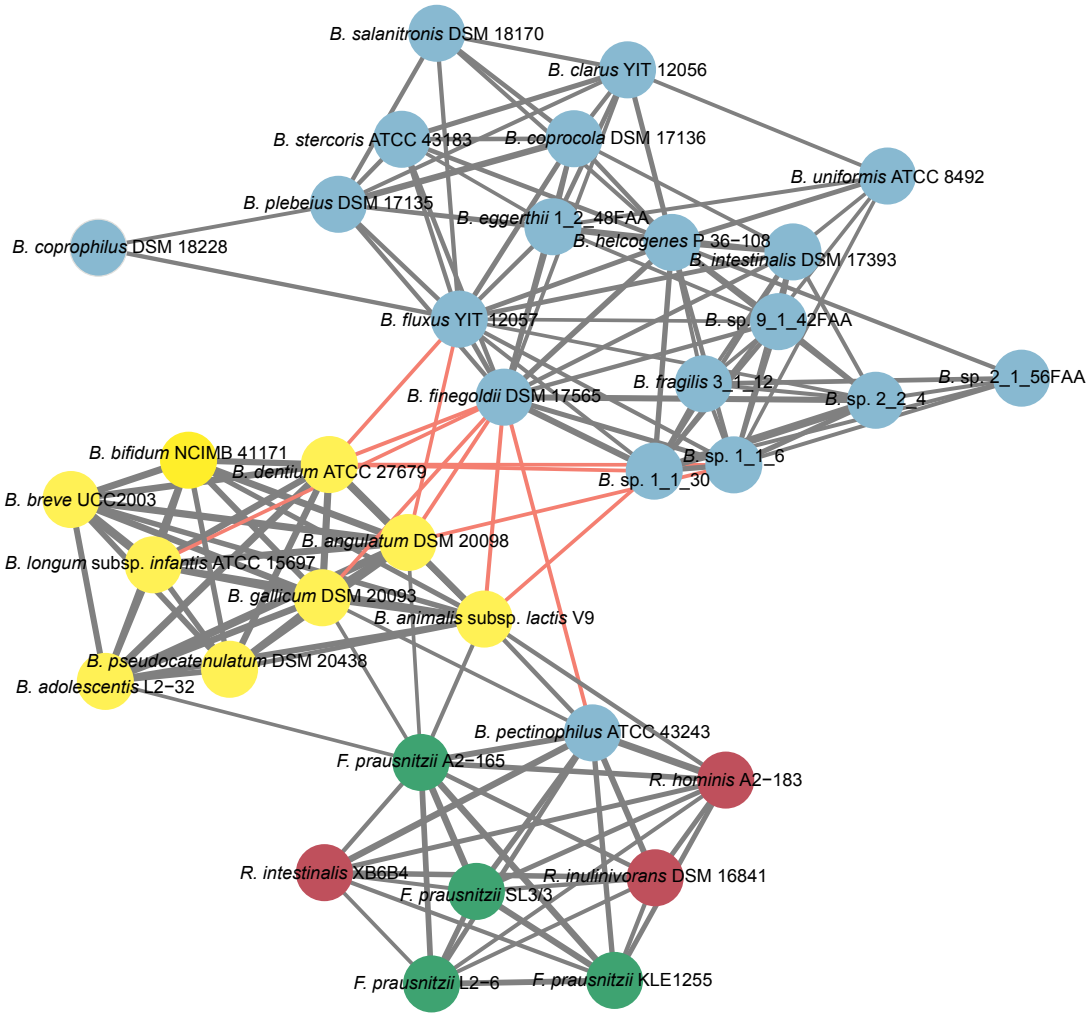

Figure 6. Co-occurrence network of abundant species. The nodes represent microbial species from the genera: Bacteroides (blue), Bifidobacterium (yellow), Faecalibacterium (green) and Roseburia (maroon). Positive and negative correlations are shown in gray and red lines, respectively. The thickness of the edges denotes the level of association based on the value of the corresponding Spearman correlation coefficient.

Interpretations of the *in silico* predictions

Species of Faecalibacterium and Roseburia exhibit a positive correlation in occurrence as indicated in gray lines. Species of Bacteroides except for (*Bacteroides pectinophilus*) and Bifidobacterium exhibit negative occurrence of correlation, as indicated in red lines. As these networks are undirected and based on correlation, it provides no information regarding the existence of an interaction, or the direction of causality if such an interaction does exist. Thereby, it is presumed that the positive and negative correlations describe only the tendency of co-occurrence between species. A species pair that exhibits a statistically significant positive correlation could indicate either a true positive ecological interaction (such as mutualism or commensalism) or a preference for shared environmental conditions.

On the other hand, a species pair that exhibits a statistically significant negative correlation could indicate either a true negative ecological interaction (such as competition or amensalism) or preference for different environmental conditions. The environmental conditions surrounding the pair of species could influence to an extent such that one could never proliferate until its favorable conditions are met. Ecological interactions could affect the outcome of co-occurrence through metabolite exchanges either one way or in both the directions. To identify and understand the environmental driver of the co-occurrence patterns, pairs of species were co-cultured *in vitro* to evaluate their growth and metabolic behavior. Based on the predictions, i.e., *Bifidobacterium adolescentis* and *Bacteroides thetaiotaomicron* were selected to inspect negative correlation, while *Faecalibacterium prausnitzii* and *Roseburia inulinivorans* were chosen to study positive correlation.

II. Evaluation of predicted co-occurrence patterns *in vitro*

a. *In vitro* co-culture between *Bifidobacterium adolescentis* and *Bacteroides thetaiotaomicron*

Both the species were cultured in YCGMS (Yeast Casitone Glucose Maltose Starch) growth medium in an anaerobic chamber at 37° C Figure 7A and 7B show the growth characteristics of each species in mono- and co-cultures.

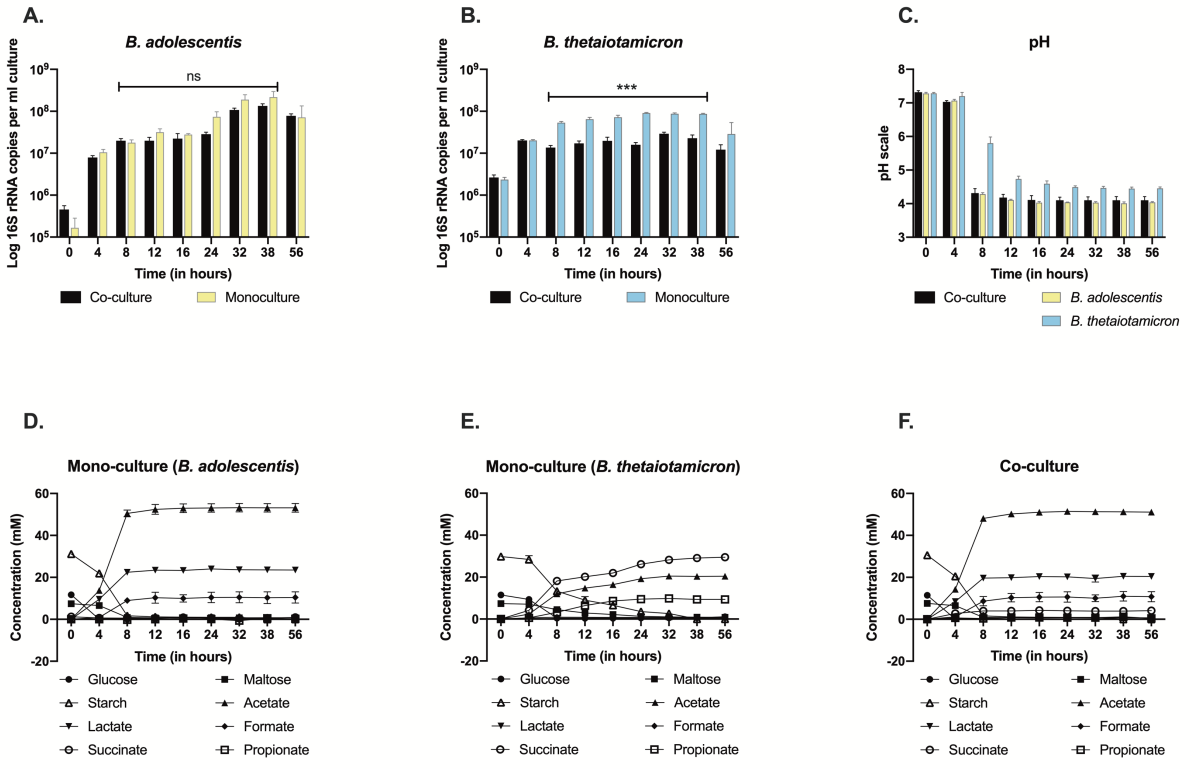

Figure 7. Growth kinetics, extracellular metabolite concentrations, and pH profile of mono- and co-cultures of *Bifidobacterium adolescentis* and *Bacteroides thetaiotaomicron* cultured in YCGMS (Yeast Casitone Glucose Maltose Starch) medium over 56 hours. P-values of less than 0.01 are indicated with asterisks.

For *B. adolescentis*, no significant difference was observed in the copy number of the 16S rRNA gene between the mono- and co-culture (Figure 7A). However, for *B. thetaiotaomicron*, the abundance of the 16S rRNA gene copies was significantly different from 8 to 32 hours of growth between the co-culture and mono-culture (Figure 7B, Student's t-test, $p < 0.01$). In Figure 7C, the extracellular pH of the co-culture was similar to that of the *B. adolescentis* mono-culture, and different from that of the *B. thetaiotamicron* mono-culture. Furthermore, the extracellular metabolite profile and their concentration levels in the co-culture were more similar to that of the *B. adolescentis* mono-culture than that of the *B. thetaiotamicron* mono-culture (Figure 7D - 7F). In the co-culture and the *B. adolescentis* mono-culture, acetate, lactate, and formate were the metabolites produced in the highest quantities. However, the levels of succinate and propionate in the *B. thetaiotaomicron* mono-culture were higher compared to that in the co-culture.

Around 8 hours of fermentation in the co-culture, there was a dramatic pH drop from 7.2 to 4.0, which corresponded to a parallel increase in the concentration of acidic metabolites such as acetate, lactate, and formate. In contrast, due to a different route of fermentative metabolism by *B. thetaiotaomicron*, the extracellular pH of the medium reached only 5.8 in its mono-culture around 8 hours of fermentation. This observation suggests that the low pH attained in the co-culture could potentially render an unfavorable environment and thereby affect the growth and metabolism of *B. thetaiotaomicron* in the presence of *B. adolescentis* irrespective of available carbon substrate in the medium. To confirm if pH was a potential cause of the growth inhibition, a preliminary pH perturbation experiment was performed. In this experiment, the environmental pH was changed from its initial state in two different ways: (i). it was changed from an acidic to basic pH, and (ii). it was modified from basic to an acidic pH. Upon doing so, it was observed that the condition with acidic pH reduced the abundance of 16S rRNA copies of *B. thetaiotaomicron,* whereas the condition with basic pH resumed its growth by a modest level. In short, the inability of one acid-intolerant bacterium growing in the presence of another acid-tolerant bacterium resulted in a negative relationship regarding their growth and metabolism towards each other.

b. *In vitro* co-culture between *Faecalibacterium prausnitzii* and *Roseburia inulinivorans*

Both the species have been reported to use acetate as a co-substrate in the growth medium (Rivière *et al.*, 2016). Hence, to evaluate their co-occurrence pattern, two different media conditions were designed, i.e., one with the acetate named as YCFAGD (Yeast Casitone Fatty Acid Glucose Disaccharide) medium and the other one without acetate named as YCGD medium. They were cultured in an anaerobic chamber at 37° C.

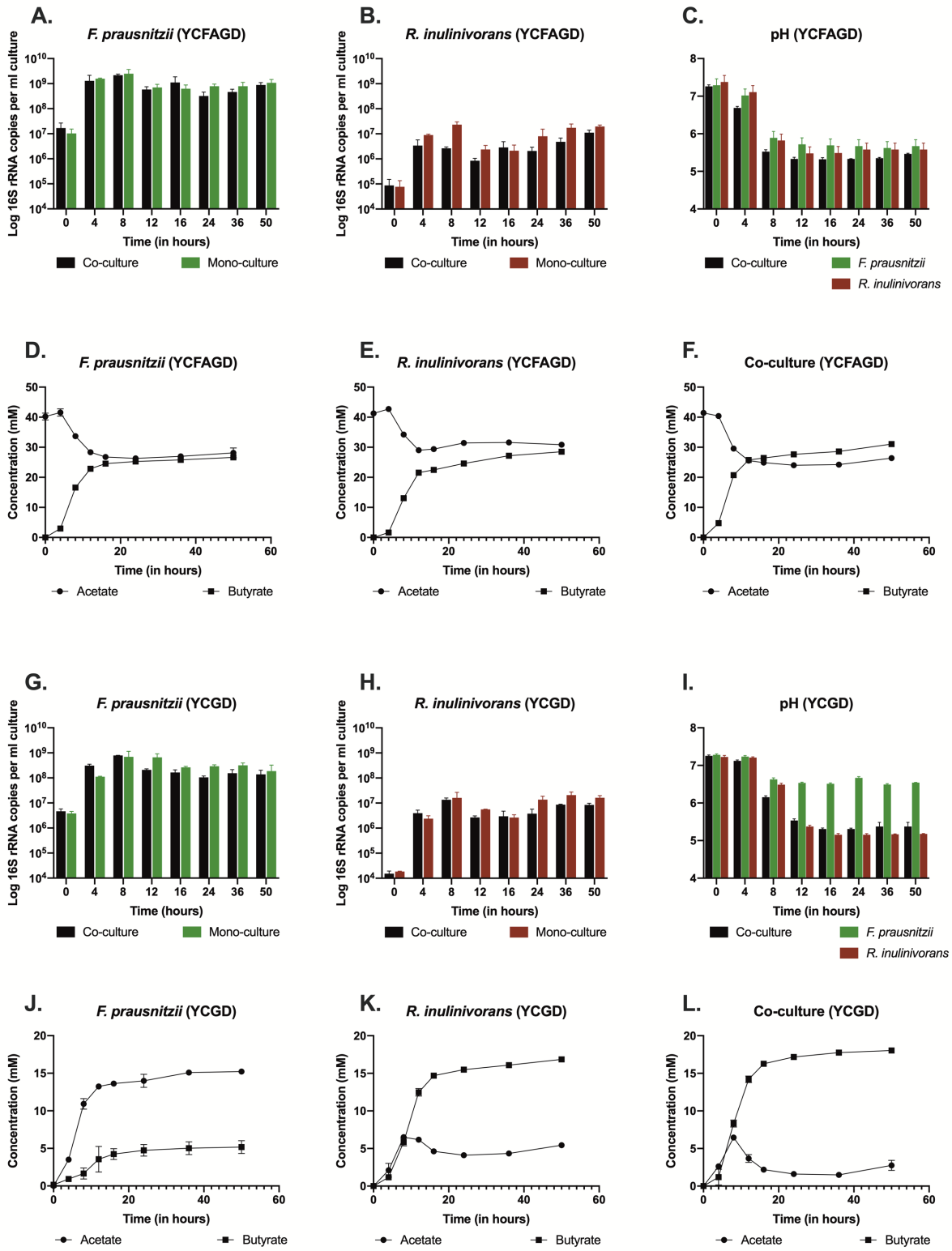

Figure 8. Growth kinetics, extracellular metabolite and pH profile of mono- and co-cultures of *F. prausnitzii* and *R. inulinivorans* cultured in YCFAGD (Yeast Casitone Free Acetate Glucose Disaccharide) and YCGD (Yeast Casitone Glucose Disaccharide) medium for 50 hours. **A, B, G, H.** Log_{10} of 16S rRNA gene copies per ml culture. **C, I.** pH profile. **D-F, J-L.** Acetate and butyrate profile in two media conditions.

Figure 8A, 8B, and 8C show the growth and extracellular pH characteristics of each species in mono- and co-culture in YCFAGD medium. The 16S rRNA gene abundance profile and the pH profile of *F. prausnitzii* and *R. inulinivorans* were similar to that found in their co-culture. Figure 8D, 8E, and 8F show a similar trend for the acetate consumption and butyrate production over time in each of the respective cultures. However, on similar comparison for metabolites as described before, in YCGD medium, the growth kinetics in Figure 8G and 8H were identical to Figure 8A and 8B. Due to the partial consumption of disaccharides, *F. prausnitzii* could not achieve complete fermentation and hence the YCGD medium (Figure 8C) reached to a final pH of around 6.5 compared to 5.8 in YCFAGD medium (Figure 8I). Also, *F. prausnitzii* mono-culture produced acetate (Figure 8J), rather than consuming its own secreted acetate, unlike *R. inulinivorans* in the YCGD medium (Figure 8K) post 8 hours of fermentation.

Under both the media conditions, the species grew together until exhaustion of carbon substrates, probably due to an overlap in their preferred pH range. However, the absence of acetate did influence the level of butyrate production and consumption of self-secreted acetate. To assess if the inoculum density had an impact on the growth outcome of mono- and co-culture, both the species were cultured in YCGD medium, with a higher inoculum density of *R. inulinivorans* than *F. prausnitzii*. The outcome remained the same. These observations implied the presence of a positive relationship between this pair of species who share a common pH in their extracellular environment.

III. Preliminary assessment of tri-culture for the formulation of synthetic microbial consortia

To verify if these co-occurrence patterns in co-culture could be observed in a tri-culture microbial community, three species (i.e., *B. adolescentis, F. prausnitzii* and *B. thetaiotamicron*) were cultured in YCFAGD medium in an anaerobic chamber at 37°C. The growth performance of each species was estimated through the production of a unique or major metabolic product. For instance, acetate, butyrate, and succinate were used as a marker to assess the growth *B. adolescentis, F. prausnitzii* and *B. thetaiotamicron* respectively.

From Figure 9A, 9D, and 9G, the growth characteristics of each species cultured in co- and tri-cultures are summarized as follows: *B. adolescentis* began its growth and reached its saturation phase around 11-13 hours of fermentation, while *F. prausnitzii* grew after 13 hours of fermentation. *B. thetaiotamicron* grew moderately only in co-culture with *F. prausnitzii*. Regarding the production of SCFAs as presented in Figure 9B, 9E, and 9H, *B. adolescentis* showed no difference in the concentration of acetate production in its bi- and tri-member co-cultures, while *F. prausnitzii* and *B. thetaiotamicron* produced a significantly higher concentration of butyrate and succinate in the bi-member co-culture as compared to the tri-culture.

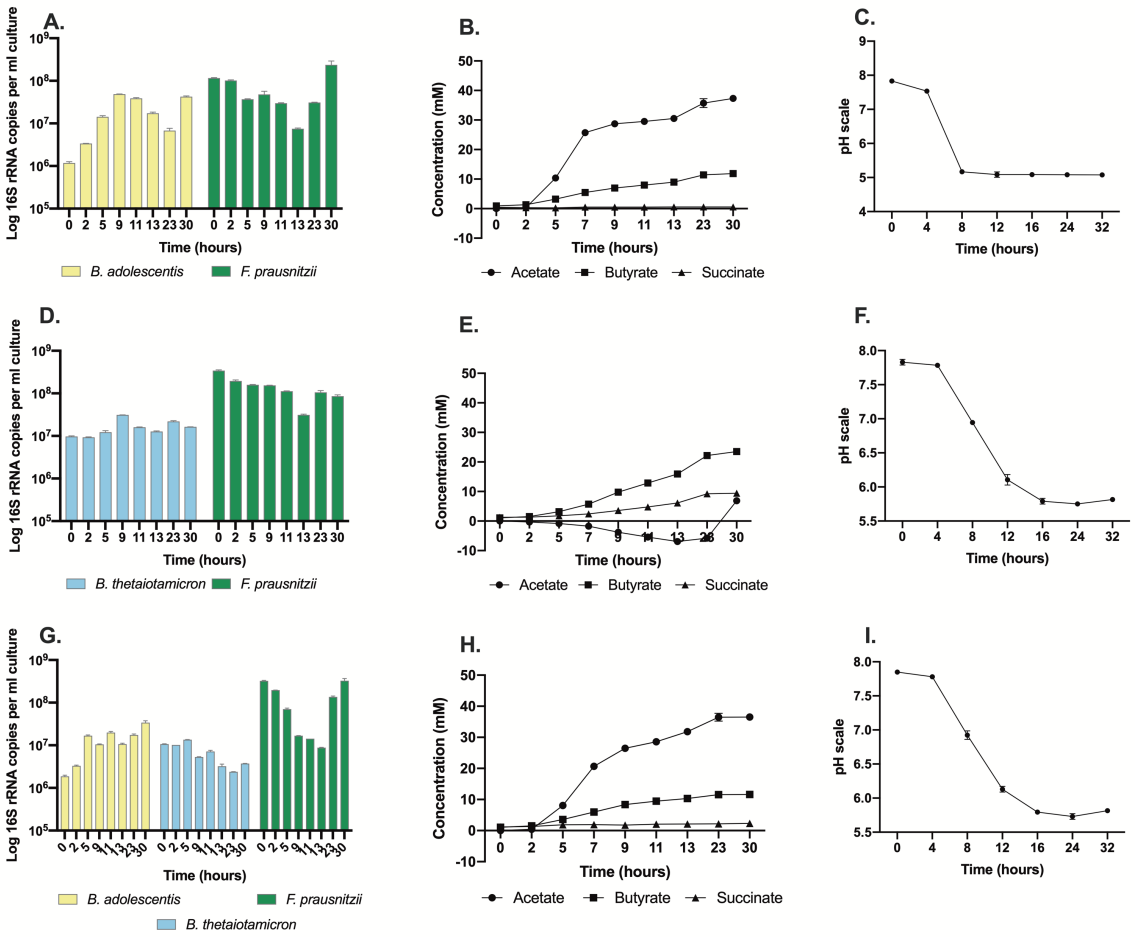

Figure 9. Growth kinetics, extracellular metabolite, and pH profile of *B. adolescentis*, *B. thetaiotamicron*, and *F. prausnitzii* co-cultured in YCFAGD (Yeast Casitone Free Acetate Glucose Disaccharide) medium over 30 hours. **A, D, G.** Log_{10} of 16S rRNA gene copies per ml culture. **B, E, H.** Acetate, butyrate, and succinate concentration. **C, F, I**. pH profile of co-cultures in YCFAGD medium.

For the pH profiles from Figure 9C, 9F, and 9I, co-cultures with the combination of *F. prausnitzii* and *B. thetaiotamicron* reached a final pH of 5.5 – 6, whereas the co-culture with *B. adolescentis* and *F. prausnitzii* reached a final pH of 5.1. Once again, it was observed that co-cultures with species of similar pH optima produced metabolites higher than that of their other combination of co-cultures. The overall growth and metabolic behavior of the tri-culture was predictable based on their respective co-cultures.

This study generated a co-occurrence network based on the species abundance in a microbial community. Experimental observations suggest that (i). the growth performance of a species in co-culture could be dependent on the type of its partner species present, as their metabolic activity might result in metabolite-mediated change in the environmental culture conditions, (ii). change in metabolic behavior under varied media composition resulted in different level of end-products, which underscores the influence of nutrient consumption by the species, and (iii). step-

by step approach of co-culturing species from bi-member to multi-member could help understand and predict the assembly of synthetic microbial communities. Therefore, investigation of co-occurrence patterns through *in vitro* pairwise cultivation further verified and recognized the importance of these network analyses.

Chapter II – Genome-scale Modeling of Microbial Communities

PAPER II – Gut microbiota dysbiosis is associated with malnutrition and reduced plasma amino acid levels: Lessons from genome-scale metabolic modeling.

OBJECTIVES

This study aimed:

- To reconstruct well-curated GEMs for representative bacterial species
- To evaluate the growth performance of GEMs through experimental validations
- To investigate the metabolic capabilities of GEMs for metabolites such as SCFAs and AAs

MOTIVATION

Different types of malnutrition (severe or acute) have exhibited gut microbial immaturity. Comparative 16S taxonomic profiling have revealed significant differences in the gut microbiota of malnourished subjects from that of healthy ones (Smith *et al.*, 2013; Subramanian *et al.*, 2014; Blanton *et al.*, 2016). However, knowledge related to metabolic differences leading to the functional variations in gut microbiota of malnourished was limited. Hence, using genome-scale metabolic models (GEMs), this study proposed to examine the metabolic variations in the gut microbiota of healthy and malnourished subjects. In this modeling framework, each of the abundant gut microbial species was modeled, and examined for their short chain fatty acid and amino acid production potential. Through implementation of GEMs and metagenomics, this approach presented a prospective way to achieve functional insights into the malnutrition-associated gut microbiota.

ANALYSIS, RESULTS, AND DISCUSSION

I. GEMs of representative microbial members in healthy and malnourished children

Gut metagenomes of children from three different countries (Malawi, Bangladesh, and Sweden) (Smith *et al.*, 2013; Subramanian *et al.*, 2014; Bäckhed *et al.*, 2015; Blanton *et al.*, 2016) were examined for this study. To simplify the modeling framework and understand the metabolic potential of bacterial species in a community, species with relative abundance of greater than 90% were identified from taxonomic studies, from all groups of children at different time-points, with a total of 68 species were selected (Figure 10). Among the 68 species, 18, 9, and 22 species

were unique in Malawian, Bangladeshi, and Swedish children, respectively. The remaining seven species were common among all groups. However, GEMs were reconstructed for 58 of the 68 species. Ten out of the 68 species were the least abundant. Also, the complete metabolic maps could not be built due to the lack of available annotation for related genomes in the public databases. Hence, they were excluded from further analysis. Once the draft models for 58 species were reconstructed, they were manually curated. These curations involved removal of unnecessary gap-filled reactions, improving the metabolic annotations for anaerobic growth, and enabling the key tasks centering around production of valuable metabolites such as short-chain fatty acids and amino acids.

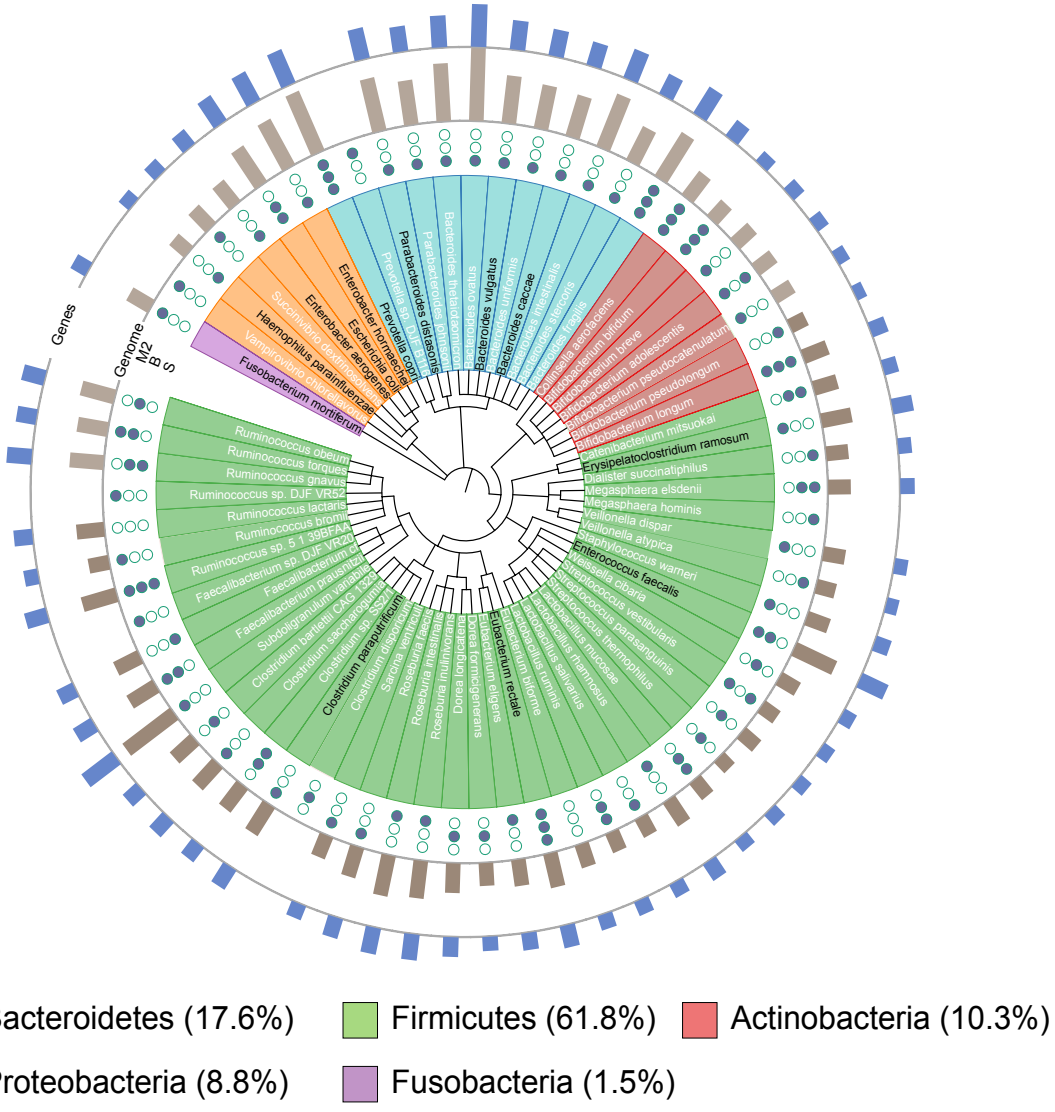

Figure 10. Taxonomic relationship among the 68 selected abundant gut bacterial species in healthy and malnourished groups. Bacterial species colored in black are opportunists. Colored solid points and circles denote the presence and absence, respectively, of each species in the children group (Bangladesh (B), Malawi (M2) and Sweden (S)). In outer two layers, brown and blue bars represent the genome size and number of genes, respectively.

II. Evaluation of predicted growth rates for six microbial species *in vitro*

To examine the accuracy of manual curation to the draft GEMs, the growth rate of bacterial species was predicted and verified *in vitro*. Based on the feasibility of culturing gut bacterial species, six species were chosen to validate the *in silico* predictions. Experimental growth rates for *Bacteroides thetaiotaomicron* (M6), *Bifidobacterium adolescentis* (M9), *Eubacterium rectale* (M29), *Faecalibacterium prausnitzii* (M31), *Prevotella copri* (M41), and *Roseburia inulinivorans* (M43) were compared to the *in silico* predictions. The simulation of these GEMs was demonstrated using the nutrient composition of the YCFA growth medium. As shown in Figure 11, the growth rates of the GEM predictions were consistent with that of the experimental results.

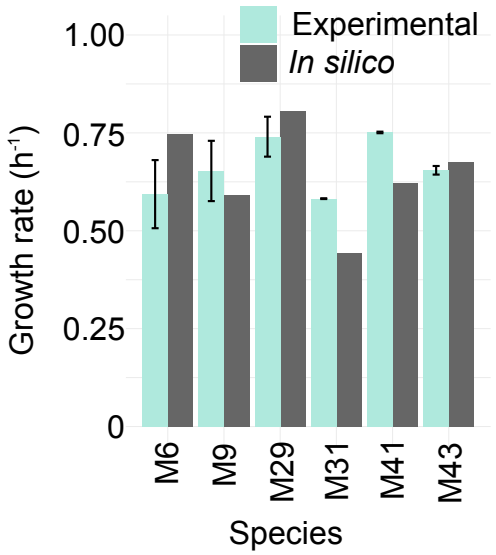

Figure 11. Experimental validation of the GEM predictions for the growth rate of six gut bacterial species cultured in YCFA medium under anaerobic growth conditions. The species are as follows: *Bacteroides thetaiotaomicron* (M6), *Bifidobacterium adolescentis* (M9), *Eubacterium rectale* (M29), *Faecalibacterium prausnitzii* (M31), *Prevotella copri* (M41), and *Roseburia inulinivorans* (M43).

III. Differential estimation of metabolite production potential in the representative community members of healthy and malnourished subjects

To evaluate the metabolic variations for SCFAs and amino acids in the healthy and malnourished gut microbial community, manually curated GEMs were analyzed. These GEMs were used to predict the production capacities of SCFAs (acetate, propionate and butyrate), and amino acids, (which includes L-glutamate, L-glycine, L-alanine, L-lysine, L-aspartate, L-arginine, L-glutamine, L-serine, L-methionine, L-tryptophan, L-phenylalanine, L-tyrosine, L-cysteine, L-leucine, L-histidine, L-proline, L-valine, L-threonine, L-isoleucine, and L-asparagine). In each species, the metabolite production potential, i.e., Estimated Maximal Production Potential (EMPP) was estimated by multiplying the abundance of the given species and the reaction flux for the metabolite production.

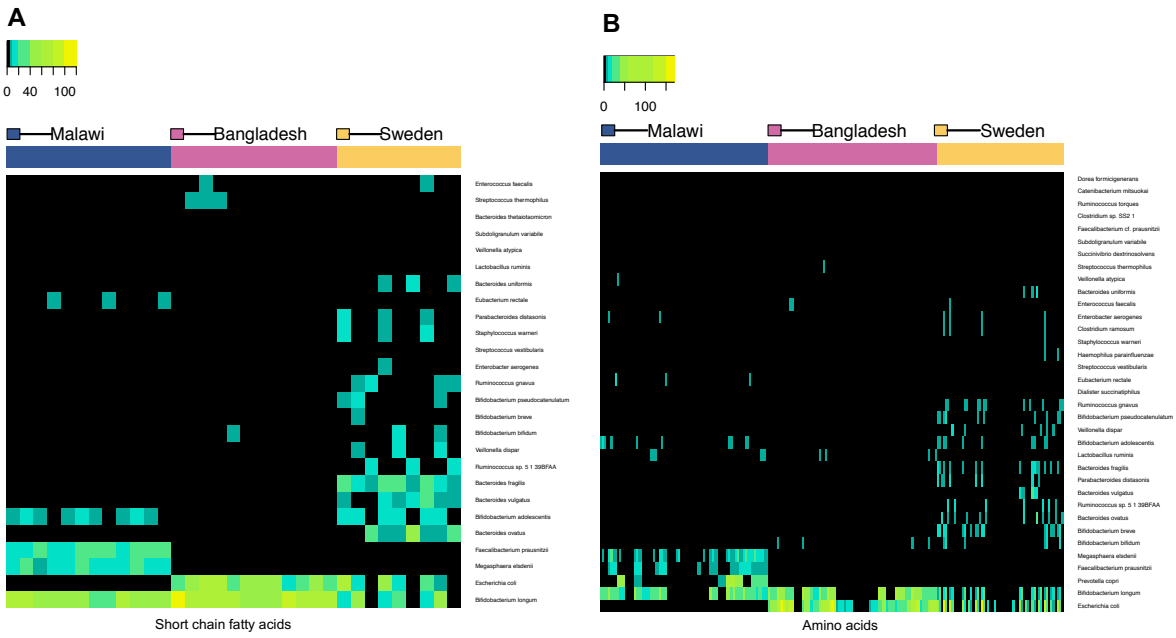

Figure 12. *In silico* predictions of estimated maximal production potential (EMPP) of **A.** short-chain fatty acids and **B.** amino acids using GEMs of gut bacteria (on the Y-axis) from Malawian, Bangladeshi, and Swedish children (on the X-axis).

On comparison of country-matched Malawian healthy and malnourished subjects, no significant difference was observed in their metabolic capacity. But, on comparison of country-matched Bangladeshi healthy and malnourished subjects, significant differences were observed between them for their EMPP of SCFAs and amino acids (Mann-Whitney U test, $p < 0.01$ for all comparisons). However, when the metabolic diversity was compared between Swedish, Malawian and Bangladeshi subjects as presented in Figure 12, the EMPP of the SCFAs and amino acids in the gut microbiota of Swedish children were observed to higher than that of the Malawian and Bangladeshi children. This observation highlights that country-dependent metabolic variations in the gut microbiota of different children groups are suggested to be driven by differences in environment and lifestyle (especially dietary choices). Furthermore, reduced metabolic diversity has been associated with an increased risk for a dysbiotic gut, which was consistent with the reduced microbial diversity in Malawian and Bangladeshi children versus healthy Swedish children (Kumar *et al.*, 2016).

PAPER III – Uniform randomized sampling of microbial communities.

OBJECTIVES

This study aimed:

- To design and implement a random sampling algorithm in microbial communities
- To study the feasible metabolic phenotypes in microbial communities using GEMs
- To characterize the metabolic capabilities of several microbial communities comprised of representatives of the human gut commensals with experimental validations

MOTIVATION

As elucidated before, GEMs have been used in modeling of single gut microbial species. They could be extended and applied for pairwise or higher community simulations. In a modeling investigation of a two-species microbial community, the growth rate of each member, substrate utilization and product production profile, and their contribution levels could be computationally quantified, provided the relevant experimental data is available. Following to model evaluation, they could be used as a predictive tool to model the community framework. Flux through a biomass reaction is often selected as an objective function to mimic the cellular physiological state, such that bacterial cells tend to maximize their growth rate due to evolutionary pressure. However, the possible objective function for a microbial community of a large size is uncertain due to lack of knowledge of a proper objective function. Alternatively, one could analyze the community-level fluxes in an unbiased way through uniform randomized sampling method, where the feasible metabolic phenotypes are sampled in a particular microbial community. However, randomly sampling a high-dimensional space is computationally intensive and time-consuming, therefore, an alternative coarse-grain approach is to consider only the community level (inputs, outputs and linking reactions) reactions and not the internal reactions occurring within a single species. Hence, to understand how the community level fluxes distribute and cluster, several microbial communities were formulated in the form of a linear programming (LP) problem and uniform sampling approach was implemented to investigate the feasible solution space of the LP problem.

ANALYSIS, RESULTS, AND DISCUSSION

I. Sampling of solution space for co-culture between *B. adolescentis* and *F. prausnitzii*

The growth dynamics and metabolite concentration changes were investigated over time for *B. adolescentis* and *F. prausnitzii* co-culture *in vitro*. *B. adolescentis* has been regarded as an acetate producer, and *F. prausnitzii* is a known acetate-converting butyrogenic colon bacterium (Rios-Covian *et al.*, 2015). As predicted and validated in Chapter I, *B. adolescentis* model produced acetate, and *F. prausnitzii* model used this metabolite (supplemented in the growth

medium) for growth. Therefore, acetate could act as a linking metabolite in the co-culture simulations.

On co-culturing *B. adolescentis* and *F. prausnitzii*, experimental measurements for the concentration of SCFAs were found in the following order: with acetate and lactate being the most highly produced metabolites, followed by butyrate and formate. When this co-culture was simulated through GEMs, the co-culture model predicted a higher capacity for acetate and lactate production followed by that of butyrate and formate (Figure 13A). As predicted, acetate was observed to be as a linking metabolite in the topology detection of the co-culture simulation. The samples related to this linking reaction have a lower median compared to the samples related to the acetate producing reaction (released to the media) (Figure 13A). This implied that the acetate produced from *B. adolescentis* is partly cross-fed to the *F. prausnitzii* in the co-culture model. As predicted, these observations were found to be consistent with experimental co-culture evaluation.

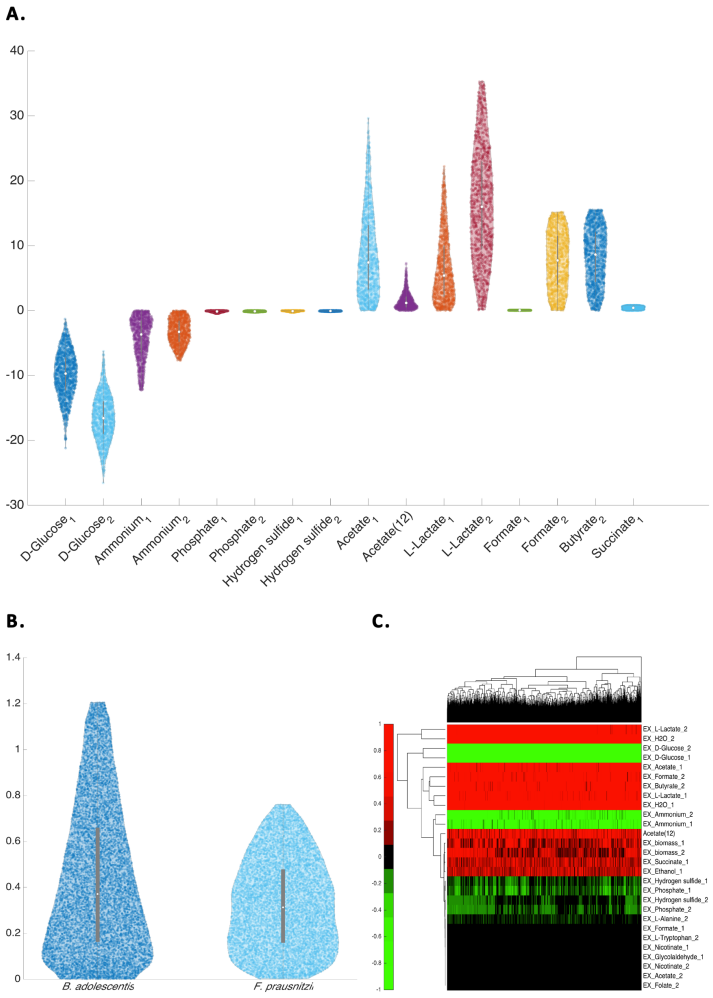

Figure 13. Sampling of solution space for co-cultures between *B. adolescentis* and *F. prausnitzii.* **A.** Distribution of the community-level metabolic samples. **B.** Distribution of the biomass reactions' samples. **C.** Hierarchical clustering of the community-level reactions.

Regarding the carbon substrates in the medium, each species had similar substrate uptake rates between the two models except for glucose. For glucose, the model predicted that the consumption rate of *F. prausnitzii* was higher than that of *B. adolescentis* (Figure 13A). However, it was noted that even though *F. prausnitzii* had a higher glucose consumption rate and also received acetate as a second carbon source from *B. adolescentis*, the former had a lower growth rate compared to the latter as measured in their *in vitro* co-culture. When the distribution of the biomass reaction flux samples was plotted, *B. adolescentis* model showed higher growth rate capacity compared to that of *F. prausnitzii* under the same set of constraints (Figure 13B). These growth predictions were consistent with the *in vitro* co-culture data measurements, where *B. adolescentis* had higher growth rate than that of *F. prausnitzii*. In the clustergram, the biomass reaction of *B. adolescentis* clustered tightly with the biomass reaction of *F. prausnitzii*, followed by the co-culture model's two outputs, namely ethanol and succinate, and the linking reaction (Figure 13C). The tightness of the biomass reaction samples indicates the level of influence on each other.

II. Sampling of solution space for co-cultures between *B. thetaiotamicron* and *F. prausnitzii*

Similar to the previous simulation, the growth dynamics and metabolite concentration (Figure 14A) were investigated for co-culture of *B. thetaiotamicron* and *F. prausnitzii in vitro*. *B. thetaiotamicron* is an acetate producer (Das *et al.*, 2018), and *F. prausnitzii* is an acetate converting butyrogenic strain. Experimental evaluation of co-culture with these species resulted in higher levels of butyrate production compared to that of *F. prausnitzii* mono-culture. The amount of detected acetate in the co-culture was lower compared to *B. thetaiotaomicron* mono-culture.

As predicted and validated in the mono-cultures, *B. thetaiotaomicron* model produced acetate, and *F. prausnitzii* model utilizes this metabolite for growth (Figure 14A). Furthermore, when the samples related to the production of these two metabolites from mono-culture versus co-culture models were plotted (Figure 14B), the co-culture model predicted higher butyrate production capacity than that of its mono-culture model. However, on analyzing the biomass reactions, we observed that the acetate producer could reach higher growth rates within the community constraints while the butyrate producer of the co-culture has a considerably lower median (Figure 14C). Regarding clustering of the community-level reactions, the biomass reactions clustered tightly with the linking acetate reaction, which suggests that, the clustering pattern is community dependent (Figure 14D). The clustering patterns for a simple community consisting of two members, one acetate producer and one butyrogenic strain seemed to be dependent on the type of strain present in the community.

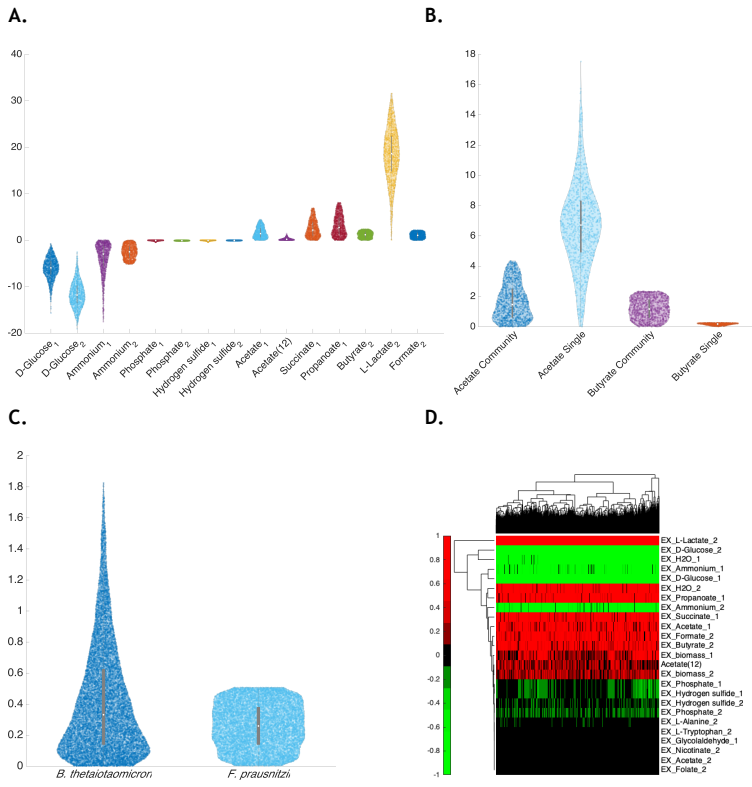

Figure 14. Sampling of solution space for co-cultures between *B. thetaiotamicron* and *F. prausnitzii*. **A.** Distribution of the community-level samples. **B.** Distribution of samples related to acetate production in *B. thetaiotaomicron* and butyrate production in *F. prausnitzii* in mono and co-culture simulations. **C.** Distribution of the biomass reactions' samples. **D.** Hierarchical clustering of the community-level reactions.

One common pattern in clustering of the community-level reactions is that in all the co-culture simulations, the biomass reactions and the linking acetate clustered together. This pattern could point to the fact that there is a commensalistic relationship between the species, as observed in this study. Another common observation in the co-culture simulations is the higher capacity of growth for the primary degrader strains. The reason behind the dominance of the acetate producers might be due to their vast enzymatic arsenal to degrade various complex di- and polysaccharides, as for example in case of *B. longum* at least 8% of the genome is dedicated to carbohydrate metabolism (Schell *et al.*, 2002) and this percentage can reach to 20% in case of *B. thetaiotaomicron* (Degnan and Macfarlane, 1995; Xu *et al.*, 2003). It was also observed that the acetate released to the media, was higher than that of the linking reaction in all two-member co-culture simulations. This might be due to the butyrogenic strains' lower capacity to absorb the produced acetate compared to the amount of its production. Altogether, this methodology allowed us to investigate the achievable metabolic phenotypes within the boundaries of a microbial community, visualize the distribution of samples associated to each community-level reaction and detect the reactions that cluster together.

Chapter III – Functional Analysis of Bile Acid Metabolism

PAPER IV – Metagenomic analysis of bile salt biotransformation in the human gut microbiome

OBJECTIVES

This study aimed:

- To identify bile salt biotransformation protein homologs through sequence and domain conservation in microbial species
- To estimate the distribution profile of the identified homologs at phyla level
- To calculate and compare the abundance of bile salt biotransformation genes between healthy and diseased subjects
- To calculate and compare the abundance of primary and secondary bile acids between healthy and diseased subjects

MOTIVATION

Based on the published studies related to bile acid metabolism, four features have been realized. First, IBD subjects differed in gut microbial composition and fecal metabolite profile from healthy subjects (Jansson *et al.*, 2009; Le Gall *et al.*, 2011; Bjerrum *et al.*, 2015; Jacobs *et al.*, 2016). Second, IBD-associated taxonomic changes included a phylum-level decrease in Firmicutes (especially of class Clostridia and family Lachnospiraceae), and phylum-level increase in Proteobacteria (especially of family Enterobacteriaceae) (Huttenhower *et al.*, 2014). Third, mouse models of inflammation revealed changes in bile acid levels and physiological response mediated through a dysbiotic gut microbiome (Marcobal *et al.*, 2013; Wahlström *et al.*, 2016, 2017). Fourth, partial set of microbial-mediated bile acid enzymatic genes were assessed to evaluate the role of gut microbial contribution in bile acid metabolism (Jones *et al.*, 2008; Gothe *et al.*, 2014; Labbé *et al.*, 2014). In this regard, this study aimed to provide a comprehensive view of gut microbial-mediated bile acid metabolism based on sequence and structural domain conservation. Once the entire set of relevant enzymatic homologs were identified, their gene and metabolite abundances were compared in the gut microbiota (shotgun metagenomes) of healthy and IBD subjects.

ANALYSIS, RESULTS, AND DISCUSSION

I. Identification of bile salt biotransformation protein homologs

Clostridium scindens, and few other Clostridial species have been well characterized for bile acid metabolic genes (Ridlon *et al.*, 2006). However, there is very limited experimental knowledge related to the growth sensitivity of various species to bile acids. Hence, this study employed protein sequences from these well-studied species to identify BSBP homologs in other organisms

(Ridlon *et al.*, 2006, 2016) using protein BLAST (Altschul *et al.*, 1990) from the UniProt database (UniProt Consortium, 2018).

Protein sequences with known BSB function and non-BSB function were used as test data to determine the threshold parameters for BLASTp search. The accuracy of BLASTp was estimated to be >92% at e-value < 1e-10, sequence coverage > 70% and sequence identity > 30%. To select potential candidate sequences, BSBP sequences from *Eggerthella lenta* were used as positive control, while BSBP sequences from the genus Helicobacter, Prevotella and Porphyromonas were used as negative control based on their growth sensitivity tests in the presence of bile acids in the culture medium (Shah *et al.*, 1995; Finegold, 1996; Han *et al.*, 1996; Itoh *et al.*, 1999; Yokota *et al.*, 2012; Ridlon *et al.*, 2013; Harris *et al.*, 2018). Next, these candidate sequences were analyzed for their functional domains using the default GA (gathering cut-offs) (Finn *et al.*, 2016). Sequences that had identical domain organization to that with the respective query proteins were retained and cross-verified for their functional annotation using eggNOG algorithm (Huerta-Cepas *et al.*, 2016) with DIAMOND (Buchfink *et al.*, 2015) as mapping mode and retaining other default parameters. Finally, sequences that had passed through each of these filtering steps were estimated with a total count 10,613 protein homologs. Species from six major phyla were identified to have bile biotransformation protein homologs (Table 6).

Table 6. Distribution of the total number of bile biotransformation protein homologs grouped by respective phylum. The values indicate the total number of strains in each category.

PHYLUM	BaiA	BaiB	BaiCD	BaiE	BaiF	BaiG	BaiH	BaiI	BaiJ	BaiK	BaiL	BSH	HSD
Actinobacteria	86	317	25	538	146	475	67	0	15	68	44	36	182
Bacteroidetes	158	1	2	1	0	1	2	0	3	0	23	4	264
Firmicutes	385	13	435	20	96	244	424	2	130	92	576	889	409
Fusobacteria	5	0	2	0	0	0	2	0	44	0	0	0	11
Proteobacteria	114	152	92	104	460	73	85	0	31	534	84	2	598
Verrucomicrobia	1	0	0	1	0	0	0	0	0	0	2	0	2

Based on the distribution pattern of species with Bai or BSH homologs, it could be inferred that not necessarily all the species might have each of the structural proteins as compared to the reference strain (Table 6). This observation was in line with the pattern from other Clostridial species, such as *C. hiradonis* that carries baiBCDEA2FGHJ while *C.sordellii* VPI 9048 carries only baiCDA2HE (Long *et al.*, 2017). In contrast, species with sequences of less than half of the total set of reference proteins might contribute to similar function and substrates resembling that of bile acids.

II. Differential abundance analysis of bile-salt biotransformation genes

The inherent challenging task of metagenomic mapping is to assemble the structural genes of the *bai* operon from whole genome metagenomes without considering a verified reference library of bile acid metabolic genes. This makes it difficult to estimate the absolute abundance of the *bai* operon in each species. Hence, for simplification, we assumed that the genes involved in the secondary bile acid metabolic pathway (*bai*A to *bai*I) are physically unclustered across the genome (Devlin and Fischbach, 2015) and thereby estimate the total bile acid biotransformation capacity of a microbial community. In the following sections, the differential gene abundance of the identified BSBP homologs was calculated by mapping shotgun metagenomes to UniProt database (for the list of identified or candidate homologs) and compared against country-matched healthy controls in two different IBD cohorts: MetaHIT Spanish cohort (Qin *et al.*, 2010), and HMP2 American cohort (The Integrative HMP Research Network Consortium, 2014).

a. Case-control groups

In the American cohort, the differential abundance analysis was performed in the following two steps. First, the genes abundances were compared between healthy and IBD subjects. Second, we sectioned the IBD group into two based on their subtypes i.e., Ulcerative Colitis (UC) and Crohn's Disease (CD) subjects. We observed no significant difference for the normalized gene abundance of the total BSBGs between the healthy and IBD populations (Figure 15A). However, a pairwise comparison between healthy controls and CD subjects of the IBD group, but not UC subjects, showed a significant difference (Mann Whitney–Wilcoxon test, $p < 0.05$) (Figure 15A). The mean of the normalized abundance of total BSBGs was lower in the CD subjects (3.68e-05) than that of the healthy controls (4.34e-05) (Mann–Whitney–Wilcoxon test, $p < 0.05$) Figure 15A), suggesting a reduced abundance of species with BSB potential in the CD subjects.

For a similar analysis performed in the Spanish cohort, a significant difference was observed for the normalized gene abundance of the total BSBGs between healthy and IBD subjects. The mean of the normalized abundance of total BSBGs was lower in IBD subjects (7.77e-05) than that of healthy controls (1.08e-04) (Mann–Whitney–Wilcoxon test, $p < 0.05$) Figure 15C), suggesting a reduced abundance of species with BSB potential in the IBD subjects. This was consistent with previous findings, where similar analysis with a different approach (where only three enzymes of bile acid metabolism were considered) was performed on the same set of samples (Labbé *et al.*, 2014). However, due to the low sample size of CD subjects, further division of the IBD subjects could not be carried out for pairwise statistical comparisons. Replication of results in an additional cohort with similar case-control group further validates our method workflow and the association between this metabolism and IBD subjects.

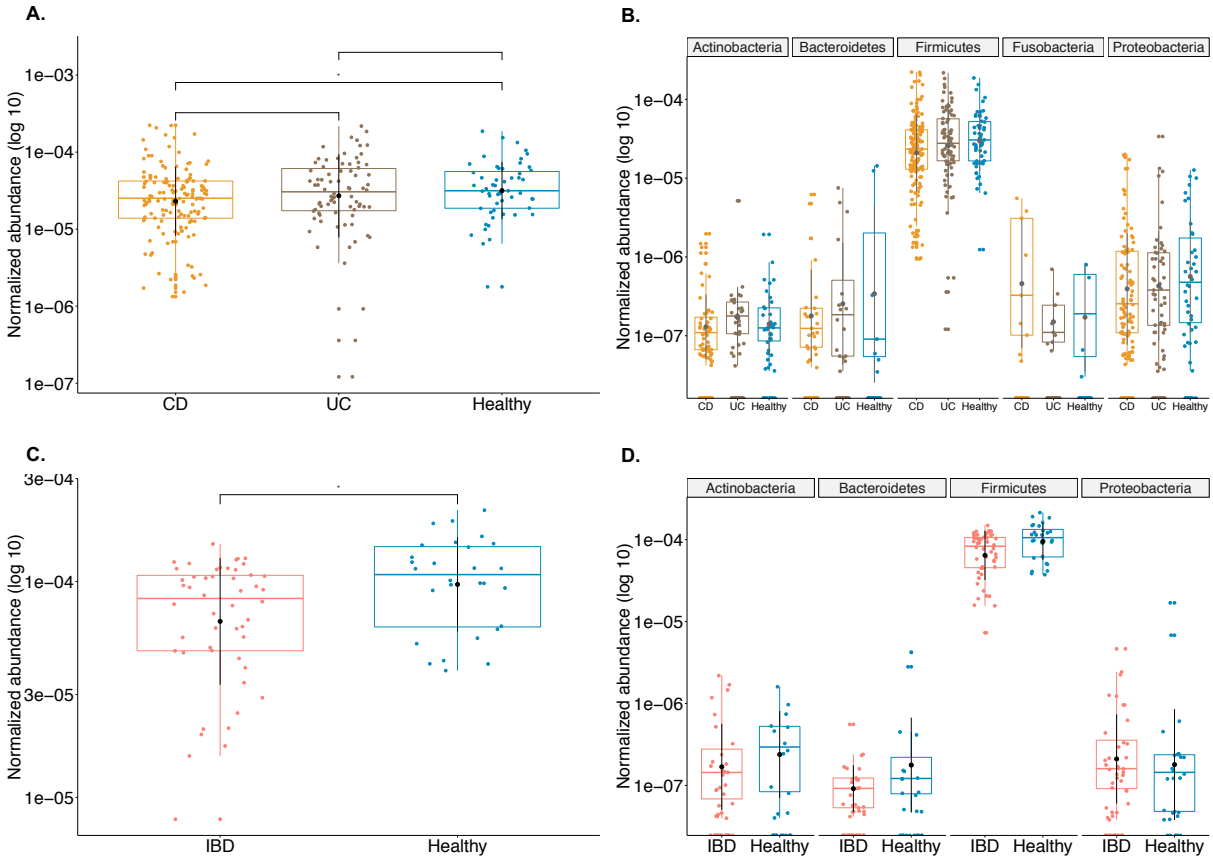

Figure 15. Normalized abundance of total BSBGs and their phylum (Y-axis) in healthy and IBD subjects (X-axis) sampled from **A. and B.** the USA, and **C. and D.** Spain. In the USA cohort, subjects with Crohn's disease and Ulcerative colitis are abbreviated as CD and UC respectively). The asterisks indicate ns: p > 0.05, *: p <= 0.05, **: p <= 0.01, ***: p <= 0.001, ****: p <= 0.0001 (Mann-Whitney Wilcoxon test).

Next, to assess the contribution of phyla accounting for BSBG abundances, the taxonomic information for each mapped gene was retrieved, and the relative abundances of each phyla-level BSBGs were computed accordingly for each group in both the cohort. Genes from Firmicutes phylum accounted for the most considerable abundance with BSB potential relative to other phyla (Figure 15B and Figure 15D). This observation was in line with the existing data that only a subset of species from Firmicutes phylum could metabolize BAs (Gérard, 2013; Ridlon *et al.*, 2013). However, the low abundance of these genes, originating from Firmicutes, in IBD subjects compared to healthy controls (Mann-Whitney-Wilcoxon test, $p < 0.01$), further confirms that the typical signature of gut microbiome in IBD patients (Rigottier-Gois, 2013; Huttenhower *et al.*, 2014; Imhann *et al.*, 2018; Vich Vila *et al.*, 2018).

b. Multi-cohort, multi-country control groups

Following to country-matched analysis of these genes between case and control groups, healthy subjects from different geographical regions were compared against each other. The normalized abundances of the total BSBGs among healthy individuals of different countries increased in the

following order: USA > China (Qin *et al.*, 2012) > Spain > Denmark for all possible pairwise combinations (Figure 16). While individuals from USA showed a significant difference from that of Denmark and Spain (Mann Whitney Wilcoxon, $p < 0.01$), no significant difference was found between the healthy individuals of Denmark and Spain (Figure 16).

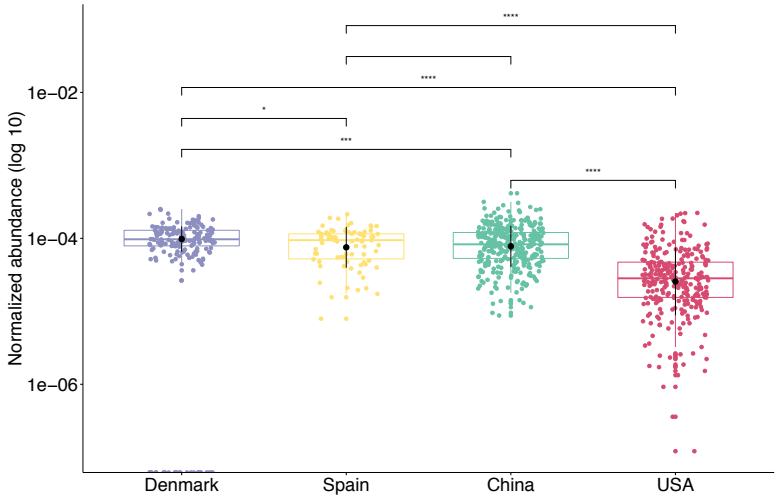

Figure 16. Normalized abundance of total BSBGs (Y-axis) in healthy individuals sampled from Denmark, Spain, China, and the USA (X-axis). The asterisks indicate ns: p > 0.05, *: p <= 0.05, **: p <= 0.01, ***: p <= 0.001, ****: p <= 0.0001 (Mann-Whitney Wilcoxon test).

This observation suggests that dietary and lifestyle choices could be a potential factor in impacting the composition of the gut microbiota in general, thus leading to the differences in the BSBG abundance at functional level. It also highlights the consideration of demographically matched samples as a prerequisite for comparative functional studies.

III. Differential abundance analysis of primary and secondary metabolites

To verify our predictions if the IBD individuals had reduced levels of secondary BAs corresponding to their low abundance of total BSBGs, metabolomics data from the American cohort was analyzed. Primary and secondary BAs were sampled from the feces and measured with LC-MS (C18 negative ion mode analysis) (Franzosa *et al.*, 2018). BA quantification was expressed as relative proportion for each BA metabolite to the total level of BAs. The proportion of conjugated primary BAs (i.e., cholate and chenodeoxycholate) and secondary BAs (i.e., deoxycholate and lithocholate) were increased and decreased in IBD subjects compared to that of healthy controls respectively (Figure 17). The decrease in the level of secondary BAs compared to the level of primary BAs suggests a reduction in the microbial potency of BA metabolism consistent to the findings at gene level. In connection to this, low levels of secondary BAs were measured in IBD subjects who were in the active phase and clinical remission state (Duboc *et al.*, 2013). Among different forms of BAs, the proportion of conjugated BA metabolites were found at higher levels in IBD subjects than that of healthy controls, which further supports

our predictions, i.e., decreased bile acid potency in IBD subjects in comparison to healthy controls.

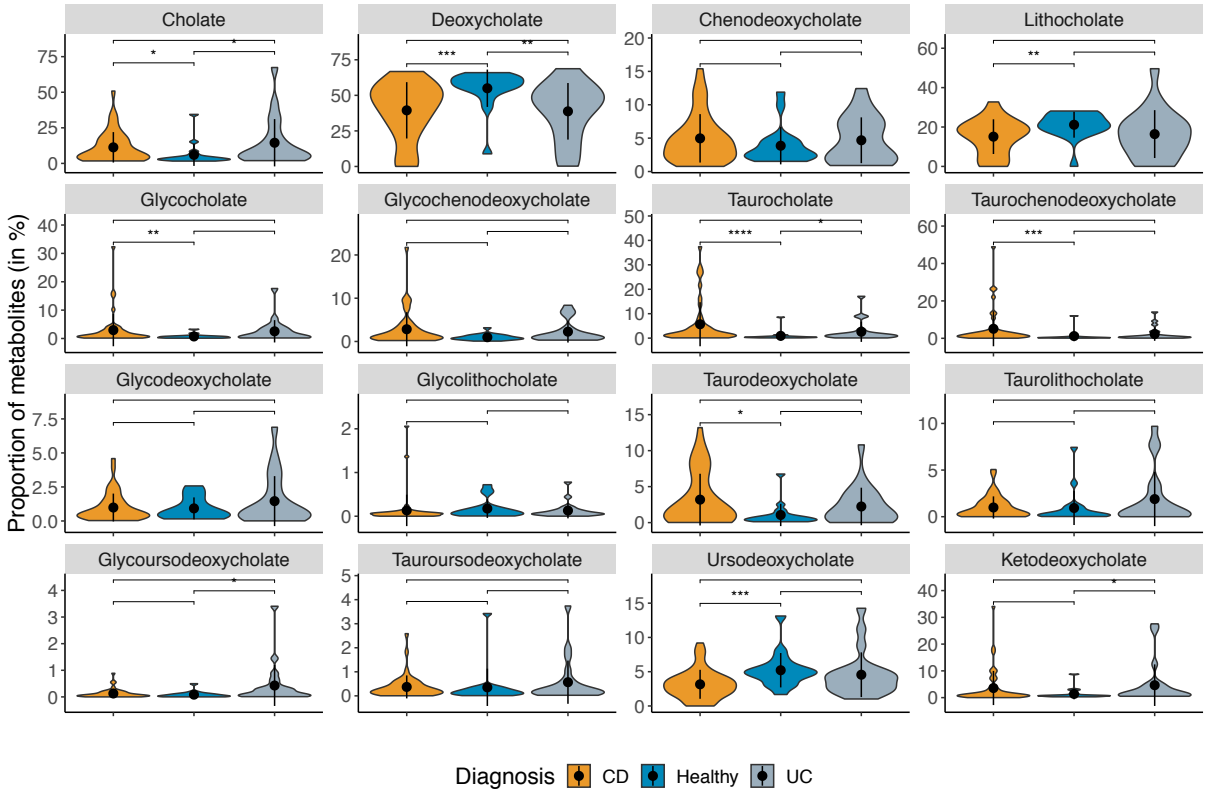

Figure 17. Proportion of conjugated and unconjugated bile acids (Y-axis) in healthy and IBD individuals (X-axis) from American cohort. The shape represents the kernel probability density of the data across the proportion of each bile acid metabolites (in percentage) along the Y-axis. Subjects with Crohn's disease and Ulcerative colitis are abbreviated as CD and UC respectively. The asterisks indicate ns: p > 0.05, *: p <= 0.05, **: p <= 0.01, ***: p <= 0.001, ****: p <= 0.0001 (Mann-Whitney Wilcoxon test).

Computation of BA metabolites and BSBGs in a coordinated set of IBD samples showed that a reduced abundance of Firmicutes has led to a reduced abundance of BSBGs, thereby reduced the conversion of primary to secondary BAs, irrespective of early-onset IBD or in clinical remission stage. Beside IBD subjects, cirrhotic patients were also found to have reduced levels of secondary BAs and increased levels of conjugated BAs in the feces. This alteration was mostly due to the breakdown of BSB bacterial population (due to low-level input of primary BAs in the gut) (Kakiyama *et al.*, 2013; Ridlon *et al.*, 2013; Bajaj *et al.*, 2014). Contrarily, colorectal cancer has been associated with high physiological levels of secondary BAs (Ajouz *et al.*, 2014). Thus, this study proposes that any deviation from the physiological range of secondary BAs could indicate a dysfunctional bile salt biotransformation.

Chapter IV – Functional Analysis of Vitamin Metabolism

PAPER V – Metagenomic analysis of microbe-mediated vitamin metabolism in the human gut microbiome

OBJECTIVES

This study aimed:

- To assess the gene-protein-reaction coverage of vitamin pathways in the GEMs of abundant bacterial species
- To estimate the relationship between the abundance of vitamin producers and vitamin consumers in a community
- To calculate the abundances of vitamin biosynthetic and transporter genes between healthy and diseased subjects
- To calculate the abundances of vitamin biosynthetic and vitamin-dependent transcripts between healthy and diseased subjects

MOTIVATION

Conventional and germ-free mice experiments have revealed the contribution of gut microbial-mediated vitamin metabolism in various studies (Sumi *et al.*, 1977; Hill, 1997; Rossi *et al.*, 2011; Degnan *et al.*, 2014). Based on genomic analyses, three features have been observed through systematic evaluations. First, the commensal bacteria differed in prevalence for vitamin biosynthetic and transporter genes. Second, for some of the species, the pathways were identified with incompleteness in their genome. Third, complementary patterns (presence and absence of a vitamin pathway) were predicted to exist between a pair of species (Magnúsdóttir *et al.*, 2015). This available information provided an opportunity to further expand the field with evaluations for the same from metagenomics and metatranscriptomics datasets. These evaluations include describing aspects of relationship between vitamin producers and their consumers, abundance of these functional genes in fecal metagenomes, and the relationship between vitamin production enzymes and vitamin-dependent enzymes in the metatranscriptomes.

ANALYSIS, RESULTS, AND DISCUSSION

I. Genome coverage of vitamin pathways in the abundant gut species

With the advent of prokaryotic microbial genomics, it is now possible to probe for specific features in the genomes of microbial species. Kyoto Encyclopedia of Genes and Genomes (KEGG) is a source of a network of enzyme-catalyzed chemical reactions, consisting of the metabolic pathways in organisms, manually curated from the literature (Kanehisa *et al.*, 2004). Microbial-

mediated vitamin metabolic KEGG Orthologies (KOs) were retrieved from KEGG database based on the numerous literature studies. Reconstruction of metabolic pathways for the most abundant species were inspected to estimate an overview of vitamin metabolic potential in a microbial community.

To assess the coverage of vitamin biosynthetic pathways, this study employed our previously published repository of human gut bacterial GEMs as described in Chapter II (Kumar *et al.*, 2018). These models have a gene-protein-reaction (GPR) association component that translates gene profiles to reactions through logical associations. All the identified microbial KOs were extracted and mapped to the nomenclature of the metabolic models to maintain consistency with the reaction identifiers in them (Kanehisa *et al.*, 2011; Arkin *et al.*, 2018). Once the related reactions were identified, a list of biochemical reactions for each vitamin category, namely, biotin, cobalamin, folate, niacin, pantothenate, pyridoxine, riboflavin, thiamine, and menaquinone, was compiled. Presence and absence of these reactions were evaluated for their coverage and pathway conservation across species and populations of each cohort. By pathway coverage, it is denoted as the likelihood of all genes required to operate the pathway are present.

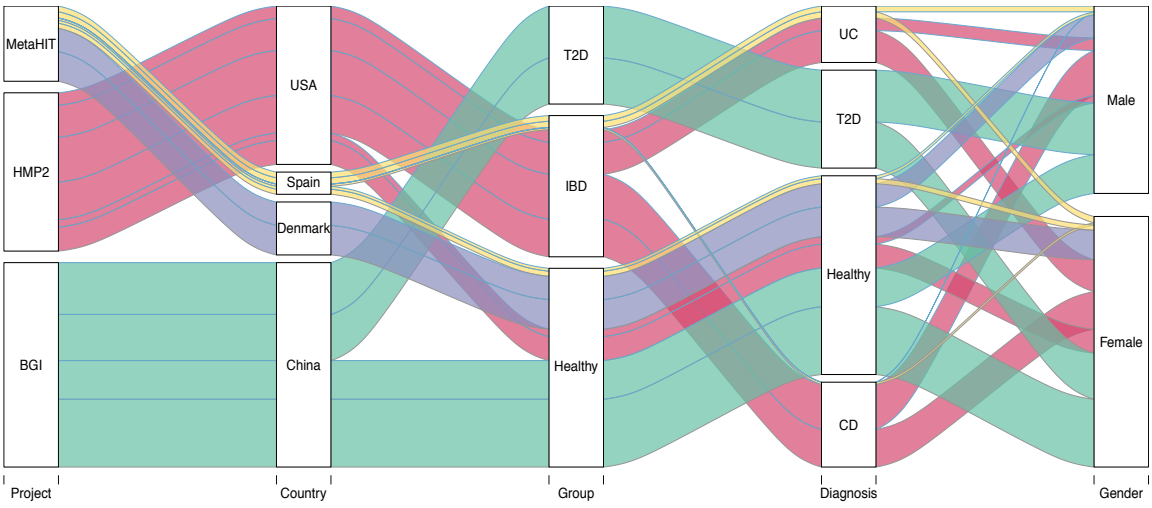

Figure 18. An overview of categorical metadata from the cohorts employed in this study.

Species abundance of the abundant gut bacterial species were identified in each of the cohort (Figure 18). Later on, unique and common species were categorized for the case and control groups, and pathway coverage for each of the identified species was plotted (Figure 19). Regardless of the health status, riboflavin and pantothenate seemed to be well represented and conserved among all the abundant species across the cohorts (Figure 19). This pathway conservation could be either due to the limited number of reactions involved in riboflavin pathway compared to other vitamin biosynthesis or evolutionary dependence on these vital reactions for bacterial cell growth. Contrarily, menaquinone has the least number of species with its

biosynthesis pathway present (Figure 19). Perhaps, human gut microbiome might have evolved to have a reduced dependence of menaquinone production from the gut microbiota as they fulfill their requirements through dietary means (Ichihashi et al., 1992; Davidson et al., 1998). In the case of folate and cobalamin, the majority of the reactions involved in its biosynthesis were present (Figure 19). Biotin and thiamine were the vitamins with a significant part of their biosynthetic reactions with their annotations missing (Figure 19). These missing reactions were related to fatty acid metabolism, and the upstream region of the biotin synthesis pathway provides a long-chain acyl as the starting point for the biosynthesis. In the case of thiamine, three phosphatase reactions involving interconversion of thiamine monophosphate, -diphosphate, -triphosphate and thiamine were absent (Figure 19). For the vitamins, the incompleteness of their biosynthesis pathway as observed in many species was is in line with previous studies; suggesting that vitamin biosynthesis could be carried out as complementary tasks between several bacteria, each harboring a part of the pathway (LeBlanc et al., 2013). These different combinations of functional differences among gut microbial members thus reflect different combinations of microbial lineages with a possible impact on synergistic interdependencies within themselves and ultimately on the host.

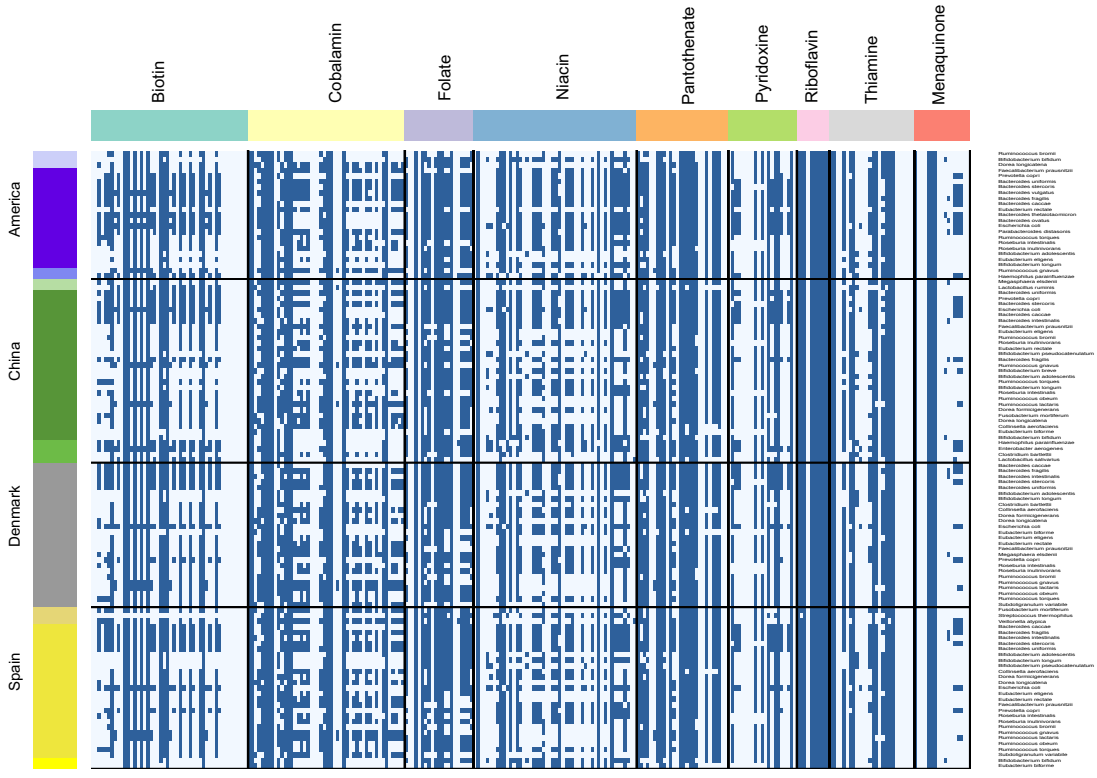

Figure 19. Presence (dark blue) and absence (light blue) of reactions associated with gene annotations of vitamin biosynthetic pathways in genome-scale metabolic models (GEMs) of abundant gut bacteria. The x-axis represents the reaction IDs associated with each vitamin pathway. Y-axis represents the list of abundant gut bacteria that are unique for health, core microbial species between groups from each country and unique for disease (as shown in a three gradient color scale under each country vertical axis).

To investigate the vitamin potential in a population-specific microbiome, similar analysis as presented in Figure 19 was performed on the American and Chinese cohort. Figure 20 shows higher coverage for riboflavin, pantothenate, and pyridoxine related pathways relative to other vitamin pathways. Species of Bacteroides genus, such as *B. uniformis, B. stercoris, B. vulgatus, B. fragilis,* and *B. caccae* were enriched with a majority of B-vitamin related reactions. The abundant species profile differed in the Chinese cohort (Figure 20B), from that of the American cohort (Figure 20A), thereby reflecting a difference in their pathway coverage, owing to a population-specific (which is mostly due to country variation) gut microbiota. However, in the Chinese cohort, species from Firmicutes were high in number compared to American cohort, suggesting that based on the Bacteroidetes/Firmicutes ratio, the abundance and function of vitamin metabolic genes could be exhibited accordingly.

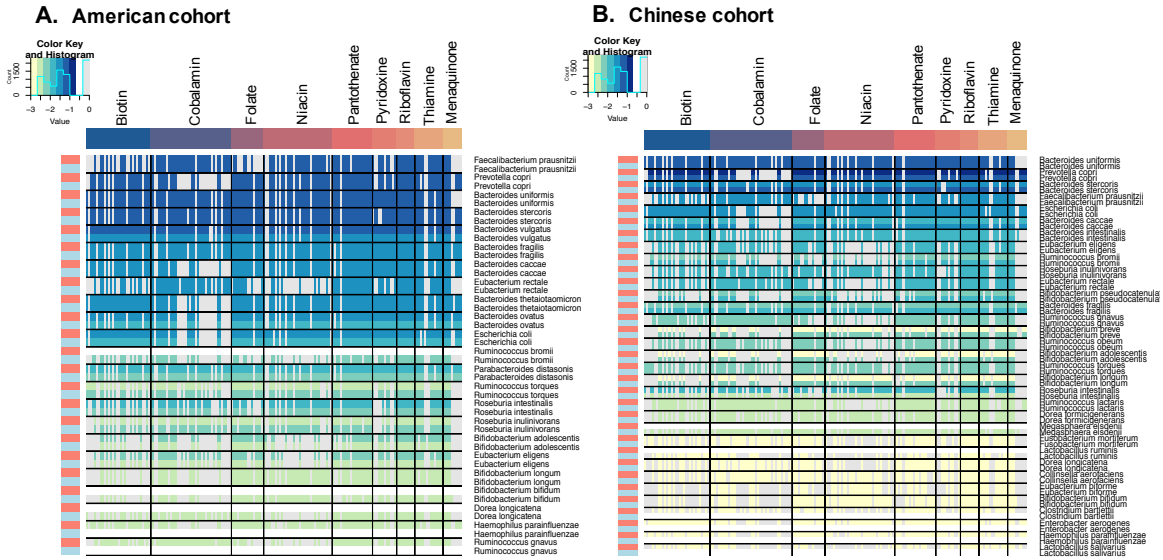

Figure 20. Relative abundance of vitamin biosynthetic reactions based on the prevalence of associated reactions, from GEMs of species from (a) American cohort and (b) Chinese cohort. The x-axis represents the reaction IDs associated to each vitamin pathway and on the y-axis the list of abundant bacteria is shown. Blue and orange (on the left) represent healthy- and diseased-associated reaction abundance in each microbial species (on the right) respectively. White color in every row represents absence of species in the cohort and the color shade (from dark in top towards light shade in the bottom) represents high to low abundance values.

II. The relationship between vitamin producers and vitamin consumers

It is well-known that vitamins serve as precursors in the biosynthesis of other coenzymes or other biochemical reactions driving a diverse set of metabolic reactions in a cell (Monteverde *et al.*, 2017). Provided this information, absence for a vitamin metabolic pathway renders an organism auxotrophic for that specific vitamin (as observed in Figure 19 and Figure 20). For the identified vitamin auxotrophs, vitamin uptake from producer organism through transporters becomes a vital source of vitamin intake. By mapping metagenomic sequences to the KEGG protein sequences,

species with or without biosynthetic and/or transporter-related proteins were identified. Based on the prevalence of genes related to vitamin biosynthesis or transporter mechanisms or both, species were assigned as either vitamin prototrophs (producers) or vitamin consumers (consumers) or producer-consumer (dual) respectively (Figure 21). Then, between each of the identified species, a correlational analysis estimated the interaction between vitamin prototrophs and consumers. From all possible pairwise combination of 171 species (675 strains), 28.9 % were significant and positively correlated, while 8 % were significant and negatively correlated (Mann–Whitney test, $p < 0.01$).

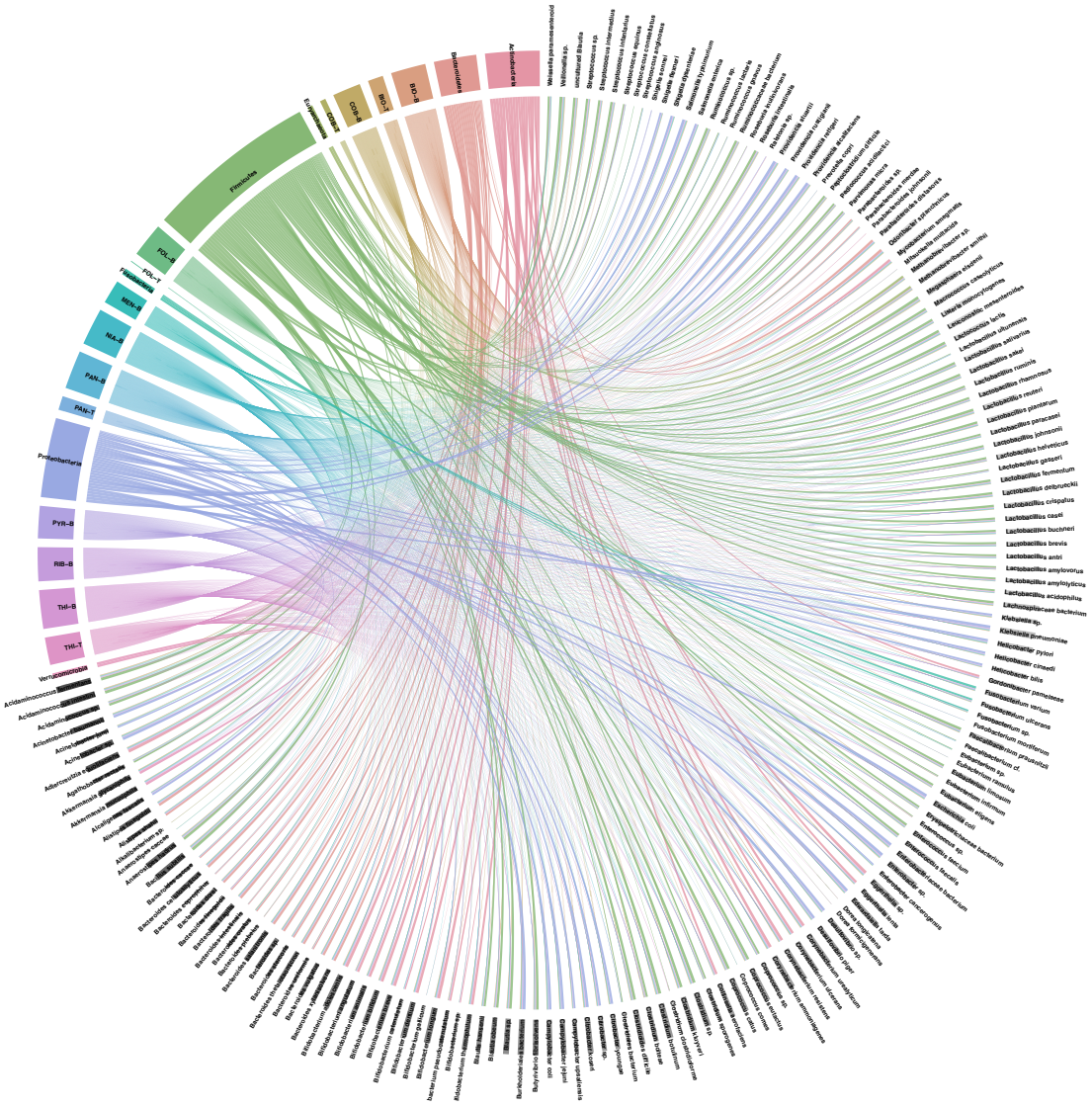

Figure 21. Prevalence of vitamin pathways in human gut microbial species. Each phylum and vitamin type are colored, and each microbial species is shown in grey.

Akin to the cycle of food webs, the equation between vitamin producers and auxotrophs cannot be analogous as that in predator-prey relationships for a "public good" (vitamin). This is because a combination of multiple factors could drive the coexistence of two species. This study observed

that pairs of species with both producer-consumer (dual) features were predominant and with a tendency to co-occur with each other. This observation was in line with the idea that species equipped with a diverse set of functional traits tend to establish resilient communities (McGill *et al.*, 2006; de Bello *et al.*, 2010; Lefcheck *et al.*, 2015). However, this study also cautions that the positive and negative associations as predicted from the correlation-analysis might not hold in every circumstance. For instance, even though *Bacteroides thetaiotamicron* could produce biotin and *Bifidobacterium adolescentis* encodes an uptake transporter protein for biotin, we found that *B. thetaiotamicron* and *B. adolescentis* were negatively correlated (Figure 6). In another example of positive correlation, *Faecalibacterium prausnitzii* and *Roseburia inulinivorans* were predicted to produce cobalamin (Magnúsdóttir *et al.*, 2015; Das *et al.*, 2018). These observations suggest that there could be a combination of variables involved in driving the species co-occurrence in a community. The parameters that drive the cohabitation of species in a specific environment seems to support as a significant working hypothesis in the determination of species interdependencies, as bacterial species evolve in response to the environmental perturbations (Lawrence *et al.*, 2012).

III. Differential gene abundance analysis of vitamin biosynthetic and transporter genes

To estimate if and to which level, the gut microbiome offers functional features that were not evolved by its host, the vitamin pathway-related sequences from the metagenomic reads were analyzed based on the KEGG annotations. However, as it is assumed that distantly related orthologs species seldom exhibits high sequence identity, the mapping accuracy of alignment was considered reliable.

a. Case-control groups

An American cohort with IBD subjects and a Chinese cohort with T2D subjects were analyzed for their differential gene abundances (Figure 22). Microbial genes annotated to KOs were grouped into the respective type of B and K_2 vitamins. On comparing the gene abundances between the country-matched groups, the gut microbiota of Chinese T2D subjects differed for most of the vitamin types associated with biosynthetic enzymes and transporters in reference to their country-matched healthy controls (Mann–Whitney test, $p < 0.01$). On the other hand, American IBD subjects were not found to be significantly different in vitamin metabolic gene abundances in comparison to their country-matched healthy controls (Figure 22).

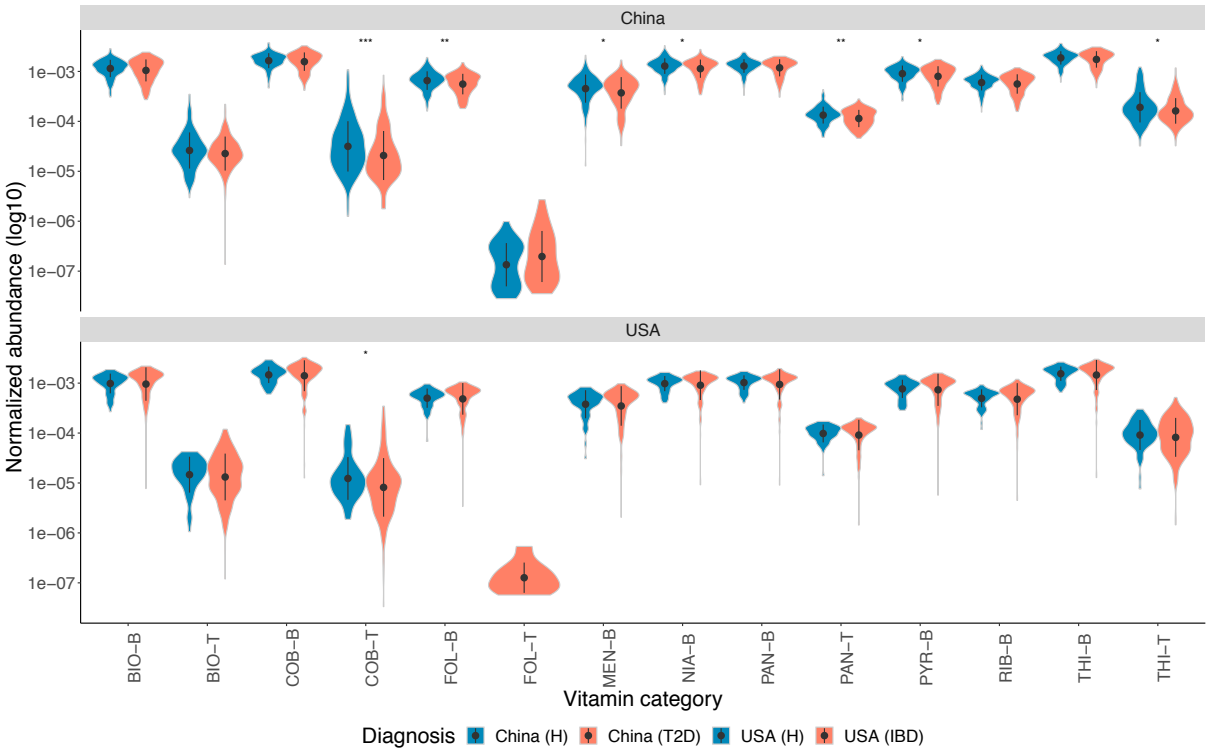

Figure 22. Differential normalized abundance of vitamin metabolic genes between case and control group subjects from the USA and China. The shape represents the kernel probability density of the data across different vitamin types are abbreviated as biotin (BIO), cobalamin (COB), folate (FOL), menaquinone (MEN), niacin (NIA), pantothenate (PAN), pyridoxine (PYR), riboflavin (RIB), thiamine (THI). Suffixes that end with -B and -T are related to biosynthetic and transporter related genes respectively. The asterisks indicate ns: p > 0.05, *: p <= 0.05, **: p <= 0.01, ***: p <= 0.001, ****: p <= 0.0001 (Mann-Whitney Wilcoxon test).

b. Multi-cohort, multi-country control groups

To evaluate if dietary and lifestyle choices had an impact on the vitamin metabolic potential of the gut microbiota, gene abundances of each vitamin categories were compared among healthy individuals from four different country populations (Figure 23). Microbial genes annotated to KOs were grouped into the respective type of B and K_2 vitamins. The abundances for each vitamin 's biosynthetic and transporter related genes fell within a similar range of border values. However, gene abundances of individuals from China and the USA differed significantly from that of Denmark and Spain (Mann–Whitney test, $p < 0.01$) (Figure 23). Although, it seems the genes involved in the biosynthetic pathways were higher than their corresponding transport related pathways, it is to be noted that functional annotation of these transporters remains challenging as they tend to diverge evolutionarily belonging to a diverse class of protein families (Eitinger *et al.*; Rodionov *et al.*, 2009; Jaehme and Slotboom, 2015).

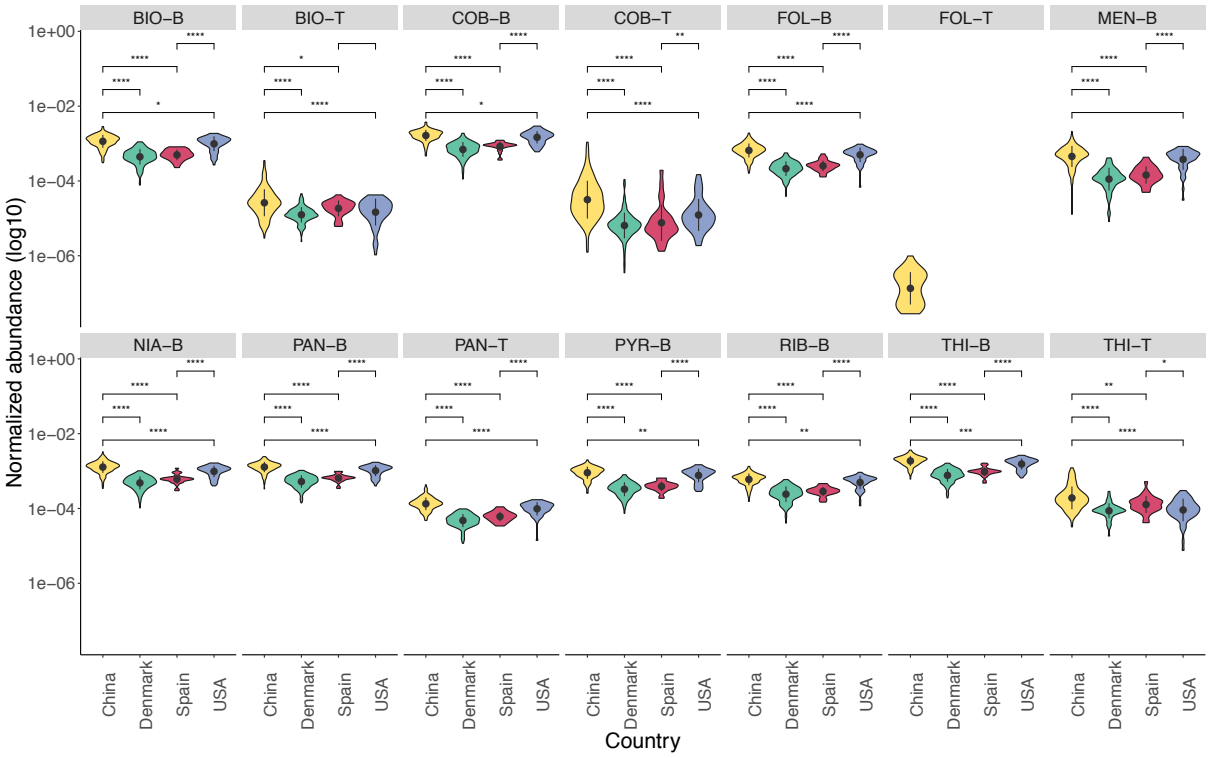

Figure 23. Differential normalized abundance of vitamin metabolic genes in healthy individuals sampled from four different countries (i.e., USA, China, Denmark, Spain). The shape represents the kernel probability density of the data across different vitamin types are abbreviated as biotin (BIO), cobalamin (COB), folate (FOL), menaquinone (MEN), niacin (NIA), pantothenate (PAN), pyridoxine (PYR), riboflavin (RIB), thiamine (THI). Suffixes that end with -B and -T are related to biosynthetic and transporter related genes respectively. The asterisks indicate ns: p > 0.05, *: p <= 0.05, **: p <= 0.01, ***: p <= 0.001, ****: p <= 0.0001 (Mann-Whitney Wilcoxon test).

IV. The relationship between vitamin biosynthetic enzymes and vitamin-dependent enzymes

Quantification of vitamin production by bacterial species of the human microbiome using HPLC or other traditional methods is challenging. However, metatranscriptomics could provide evidence on transcriptional investment or *in situ* vitamin production to understand the expression level of these pathways. Therefore, to investigate the relationship between vitamin synthesis and vitamin-dependent metabolism, metatranscriptomes were analyzed as computed by HUMAnN2 (Abubucker *et al.*, 2012). Microbial genes annotated to EC numbers were grouped into the respective type of B and K_2 vitamins.

On the synthesis side, only biotin, cobalamin, and thiamine were found to be significantly different between the metatranscriptomes of healthy and IBD subjects (Mann–Whitney test, $p < 0.01$) (Figure 24). These observations were congruent to that of a previous study where several of these vitamins were found to be lower in plasma samples of IBD subjects (Fernandez-Banares *et al.*, 1989; Kuroki *et al.*, 1993; Huang *et al.*, 2017). On further inspection of the utilization of these three vitamins, the abundance of biotin-dependent transcripts had an opposite pattern to their

synthesis transcript level, i.e., high and low abundance levels at metagenomics and metatranscriptomics scale in the IBD subjects when compared to healthy controls respectively (Figure 24A). Likewise, biotin synthesizers were equivalently high in IBD. This observation suggested that it was not only the difference in species abundance that circled two groups apart; it was also the functional features of the abundant species that contributed in determining a healthy gut microbiome. However, the abundance of thiamine-dependent transcripts had a similar pattern to its level of synthesis, i.e., low in IBD subjects compared to healthy controls (Mann–Whitney test, $p < 0.01$); while cobalamin showed no difference in the abundance of vitamin-dependent metabolisms between healthy and IBD subjects (Figure 24B).

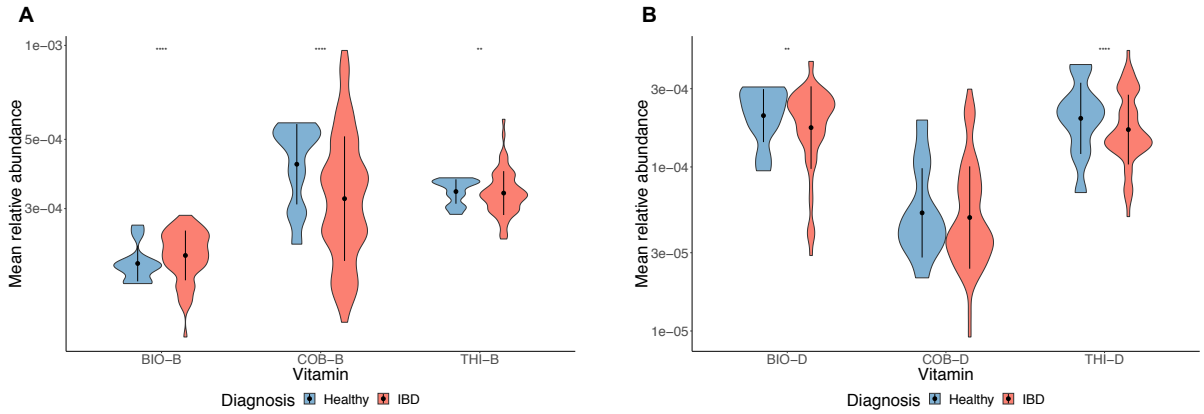

Figure 24. Mean relative abundance of (a) vitamin-biosynthetic gene transcripts (-B) and (b) vitamin-dependent gene transcripts (-D) in the healthy and IBD subjects from American cohort. The shape represents the kernel probability density of the data across different vitamin types are abbreviated as: biotin (BIO), cobalamin (COB), and thiamine (THI). The asterisks on the top indicate ns: $p > 0.05$, *: $p <= 0.05$, **: $p <= 0.01$, ***: $p <= 0.001$, ****: $p <= 0.0001$ (Mann-Whitney Wilcoxon test).

Although there was no significant difference in the metabolic potential of the gut microbiota in American healthy and IBD subjects at metagenomics level, probing into the vitamin-related expression level from metatranscriptomics provided the following insights. The ability to discern the correlation between the vitamin gene expression and its associated reaction metabolisms in a microbial community does not seem straightforward. The variations observed in vitamin synthesis and their utilization pattern suggests the presence of non-linear or complex mode of resource allocation and reuse of the enzymes. However, these findings do highlight the contribution of gut microbiota to human health.

In summary, assuming fecal metagenomics and metatranscriptomics as a representative proxy to gain insights into microbe-mediated vitamin metabolism in a broader context, samples from four different countries were studied. Ecological relationship between vitamin prototrophs and consumers highlighted possible metabolic correlations between species of various phenotypes. Species with dual phenotypes predominated amongst other possible combination of species co-occurrences. Quantitative analysis of the vitamin biosynthetic and transporter related genes

between different groups revealed alteration in vitamin metabolic potential of T2D subjects (Chinese cohort). However, only a part of B-vitamins was found to differ in their abundance between country-matched healthy controls and IBD subjects (American cohort). GEMs deployed to visualize the community-level metabolic potential of abundant gut species in diverse cohorts reflected the contribution of the population-specific microbial profile to the variation in the abundance of these genes. Quantifying the abundance of vitamin biosynthetic enzymes and vitamin-dependent enzymes in an IBD cohort in comparison to healthy controls emphasized the microbial contribution to human host.

CONCLUSIONS

The higher number of unknown than known factors drives the complexity of microbial communities. A microbial community consists of several species with diverse kinds of interaction, which makes experimental validations of community-wide interactions difficult if possible. Hence, to simplify the process of understanding a community-wide network, we constructed a simplified co-culture based on the co-occurrence network predictions presented in Paper I. *In vitro* cultivations are limited as the physiological capabilities or ecological niche of the species under investigation are challenging or might not be closely related to the wild system. However, as long as the growth medium supports their growth, experimental interpretations could be applied to these networks. Chapter I resulted in showing how metabolite-mediated pH change and its associated chemical ecology could determine the growth and metabolism of a species in a simplified community.

However, from a holistic perspective, mathematical metabolic reconstructions (GEMs) represent the topology between different metabolic reactions. When these reactions and pathways form into a network, they could be integrated with metagenome-wide omics data. GEMs provide a scaffold upon which metagenomics data can be mapped, augmenting the data interpretations with the connectivity information encoded within the network architecture. By applying this concept in Paper II, we reconstructed GEMs for the abundant gut bacterial species from the healthy and malnourished cohorts and verified their predictions *in vitro*. By mapping the abundance of species to the predicted reaction flux values for each reaction that comprised the pathway, metabolic potential of SCFAs and AAs was estimated in abundant species. In doing so in Paper II, we observed that the malnourished subjects (from Bangladesh and Malawi) were metabolically less diverse than the healthy subjects (from Sweden), suggesting how country-dependent diet variations could reflect the microbial diversity and thereby the functional diversity between infants with matured and immatured gut microbiome.

Paper III applied the concept of the random sampling of genome-scale metabolic networks to model defined gut microbial communities. This method allows consideration of the global level of community reactions, excluding the internal reaction fluxes present within the members of a community. Through co-culture modeling, we investigated the possible metabolic phenotypes within the sample space. Moreover, it identified that acetate producing reaction clustered together with biomass reaction samples, which suggests the association of a linking metabolite with the co-culture growth. Although the size of microbial communities and the *in silico* media in the present study is relatively small and simple, future research with a larger microbial community in more complex media could be employed to understand interactions at an advanced level. This approach can be used to engineer the growth and metabolism of a community through modeling to simulate community behaviors.

Next to SCFAs, investigation of second-abundant class of bile acid metabolism in Paper IV was primarily driven (i) to profile the distribution of BSBPs based on sequence homology and structure conservation, and (ii) to study the abundance of BSBGs in the gut microbiota of healthy and IBD subjects. Country-matched metagenomic analysis of shotgun metagenomes revealed a low

abundance of BSBGs in the gut microbiota of IBD subjects, from two different cohorts. Parallel to that, low abundance of secondary bile acids was found in the fecal metabolome of IBD subjects. These observations validate the identified biological insights as gathered from the integrative analysis of multi-layer omics data. Future research could address two open questions that arise from this study, and those are, (i) which combination of BSBP (Bai/BSH) contributes to bile acid resistance or sensitivity? (ii) Experimental methodologies to elucidate the requirement for minimum bai gene set could clear the distribution of these genes at least across few gut species. Nonetheless, environmental selection of microbial species in the gut could potentially tip a compensated state of immune balance in favor of chronic disease in IBD hosts.

In Paper V, exploration of vitamin metabolism in the gut microbiome was undertaken for two reasons: (i) connectivity of vitamin pathways with the central carbon metabolic reactions makes it an integral functioning system for exploration in the gut microbiome, and (ii) network analysis of gut microbial species between vitamin producers and consumers, could reveal interdependencies between species of different phenotypes. This study found that the wide prevalence of vitamin metabolic genes across the well-known abundant phyla could confer its role as part of the core functional system. Co-occurrence of bacterial species with both biosynthesis and transporter-related genes predominated amongst other possible pairs of the combination. Besides that, metagenomic and metatranscriptomic analysis of vitamin-related genes in IBD cohort showed statistically significant differences in their expression level but not in their potency than that of the healthy subjects. The nonlinear relationship of expression between vitamin synthesis and their utilization level probably suggested a complex mode of their usage. Mapping the metagenomics data onto GEMs of abundant species highlighted the difference in the population-specific gut microbiome, perhaps mediated through lifestyle and dietary choices. However, the presence of partial genome coverage of these pathways stresses if this observation was based on lack of annotation or absence of such reactions constituting the pathway.

Altogether, this thesis shows how the net co-occurrence of bacterial species in the succession of a microbial community and their functional outcome is determined based on the country-specific dietary substrates and metabolism-mediated environmental changes subjected under a specific condition. Starting from infants with immature gut microbiota to adults with perturbed gut microbiota, low abundance of their metabolic genes contributes to lower functional potential of gut microbial role in host health compared to the country-matched healthy controls; revealing the inter-relation between the microbial diversity and their functional diversity in various scenarios.

FUTURE PERSPECTIVES

Microbiome genome-wide association studies have resulted in several variants that influence the composition of the gut microbiome (Duvallet *et al.*, 2017). This number will continue to increase in future studies. However, which of these factors has a crucial role in disease-causing or health-promoting effect on the gut microbiome remains a challenging and open question to address. The primary challenge is to discriminate microbiome features that are causal for a specific condition from those that are a consequence of the disease itself — further followed by distinguishing features from confounding factors that show statistical correlation. The secondary challenge is to provide an alternative approach to a generic solution that considers high inter-individual variability of the gut microbiome.

Manipulation of complex microbial communities has become a primary goal in the field of microbiome research with an increasing appreciation of microbial ecosystems that impact human health. From an engineering perspective, the gut microbiome is highly “plastic.” Their plasticity offers an alternative solution to pharmacological interventions through treatment with the defined commensal consortium (Tanoue *et al.*, 2019). Rational design of a synthetic consortium offering clinical benefits are formulated based on properties such as gut colonization, repairing of microbiome imbalances, and stimulation of targeted immune responses to a specific disease. Nevertheless, the challenge is the selection of “correct” bacterial species in designing an effective consortium, as a single species has been found to have a positive or negative association in different cohorts and conditions. A promising approach would be to identify, test, and verify communities with potent pharmacological effects through systematic investigation of human clinical intervention studies.

Diet has been a primary driver in influencing the stability of the microbial composition and function (Kovatcheva-Datchary *et al.*, 2019). So, based on the microbiome and clinical characteristics of an individual, the personalized dietary regimen has the potential of modifying the microbiome. However, this kind of intervention requires regular re-evaluation and restructuring of dietary design in an individual. These analyses could be done through machine-learning prediction algorithms, where the associative trend between nutritional compounds and specific bacterial taxa could be identified over time (Zeevi *et al.*, 2015). However, the feasibility and accuracy of this approach in a large cohort over longer duration warrants further investigation. Before shifting to P4 (Personalized, Predictive, Preventative, and Participatory) medicine from conventional therapeutic practices, it is a challenging yet promising task to gather sufficient and necessary data to be personalized for a specific treatment (Gutin *et al.*, 2019; Zeevi *et al.*, 2019).

Advancements and parallelized nature of high-throughput omics profiling methods provide a holistic view to a system in contrast to a reductionist description. The data-rich nature of the gut

microbiome field stems from the high efficiency of omics data generation and complex interpretation of these datasets. It is therefore critical to interpret and extract accurate biological information through the integration of various data layers. To gain mechanistic insights into the underlying biological principles, omics datasets are often integrated with systems modeling and network analysis. With this perspective, this thesis has aimed to realize an understanding of the growth and metabolism of gut bacterial species. Integrative analysis of metagenomic species abundances with co-occurrence network analysis in Chapter I details the interaction between bacterial species in a simplified *in vitro* mixture. Exploitation of top-down and bottom-up approach in Chapter II leverages the advantages of metagenomics datasets and GEMs of a microbial community. Chapter III captures the biological information of a specialized functional system (bile acid metabolism) through assessment of multi-omics datasets. Along similar lines, Chapter IV analyzes a core functional system (vitamin metabolism) across diverse conditions through an integrative approach. The work described here contributes to a better understanding of the emergent properties of the gut microbiome from its parts.

ACKNOWLEDGMENTS

The journey to my doctoral thesis and publications is definitely a team effort derived from a bunch of supportive and encouraging colleagues.

Concerning my arrival here at SysBio in Sweden, thank you, **Jens**, for your green signal to my Ph.D. application. I appreciate the belief you had in me while recruiting and let me explore my own research journey. The freedom and resources that you have provided have definitely shaped and matured my research and administrative skills over the years. I have always admired the stories and experiences that you had shared on various occasions. I don't think I have ever met so much of an optimistic person other than you!

Thank you, **Fredrik** for co-supervising me over the first two years of my Ph.D. in Sahlgrenska, GU. Your critical thinking and story-telling skills with specific experimental pieces of evidence in every research project have absolutely molded my questioning skills before I execute an experiment or analysis. I am sincerely grateful for the experience gained in in vitro cultivation of gut microbial species from your lab facility.

Petia, right from providing lab-training to manuscript submission, you have been there literally next to me in each step. Even after you moved to China, you were always there for me, with just an email away. You have motivated me continuously during my hardships to go beyond the glass ceiling. I have always enjoyed and loved your proofreads. I want to express my gratitude for your positive mentorship that helped me achieve here today, it has been and will always be a pleasure working with you and knowing you as a person. I would love to see our collaboration growing in future too! XOXO!

Manish, and **Boyang** – Somehow or the other, we have always shared our workspaces nearby, perhaps leading to fun conservations and all the fantastic collaborations we've worked on. You two have assisted me with my project in the starting years. I will cherish the times that we have had through MetaCardis project meetings. We three have always been keen on trying out different cuisine and drinks making it a tradition after every monthly subgroup meeting. Thank you, guys, for the excellent company!

Parizad – You are my co-cartoon :D We started and shared our Ph.D. journey together, starting from attending my first conference together, followed by vacations, check-listing restaurants, parties, potluck, photo sessions, dramas, cultural exchanges, silly explorations, and much more with a long list of activities. Wrapping all of them into one sentence, I would like to say, "Cheers to our friendship in and out of SysBio!"

Simonas – Thanks for being so patient catalyzing me (when I was a lab-student) to basic programming and all other computational stuff in my first dry-lab project. I appreciate your dedication to our collaboration. And yes, "I sveikata" to the times being a fantastic buddy in the initial days and thereupon!

Jon – I remember the times where I have been like a question-bank to you, kudos to you for sharing your thoughts and dealing patiently. I am immensely grateful for having you as "G2-M checkpoint" in my Ph.D. years. Thank you for your valuable time to provide thorough inputs!

I thank all the 18 unknown reviewers (so far) whose critical comments and feedback has definitely benefitted the manuscript a lot in terms of scientific clarity. Also, thanks to those reviewers who appreciated and highlighted the strengths of the work.

The lab members in Gothenburg University and Sahlgrenska **Karin**, **Manuela**, **Sven**, **Tanweer**, **Hannah**, **George**, **Brendan**, **Koviljka** – I thank each one of you for your hospitality and lab support during my stay in Wallenberg lab and Mucin Biology Group.

Thank you, **Marie**, **Julia**, **Jacob**, and **Lisa** for the wet lab HPLC and GC-MS support at CMSI.

Computational support team – **Mihai**, **Pierre**, **Shaq**, and **Dimitra**: thank you guys for responding to all of my door-knocks ever since I became a dry lab student. Haha! Besides that, I will remember our pleasant and enjoyable conversations (ranging from weather to what not) over lunch and fika in the first-floor kitchen.

Martina – Just one word, hats-off to you! Thank you for being a fantastic colleague and a darling friend. I hope we continue to stay connected through social media.

Anna Bolling, **Anne-Lise**, **Erica**, **Josefine**, **Charlotte** – I thank you all the ladies for a seamless administration conducive to a Ph.D. student in SysBio.

Elias, **Gopal**, **Aida**, **Tatiana**, **Filip**, **Benjamin**, **Yasaman**, **Leif**, **Ivan**, **Facundo**, **Andy**, **Santosh**, **Jan**, **Alexandra, Takayoshi, Kate** – Warm and extended thanks to each one of you for all the on and off chat sessions that we ‘ve had about random things (academic and non-academic discussions) and for a piece of encouragement and support when required.

To all other members of the SysBio family, it has been a fantastic experience for me in being part of this dynamic group. I will remember our mantra “We Respect, We Share, and We Science.”

With a personal touch, I am grateful and blessed to have you, **Pom**. If it weren’t for you, I wouldn’t have done my Ph.D. for sure :). Despite your jam-packed schedule, thanks for always being a call or a text away, every day, no matter what!

REFERENCES

Abubucker, S., Segata, N., Goll, J., Schubert, A.M., Izard, J., Cantarel, B.L., et al. (2012) Metabolic Reconstruction for Metagenomic Data and Its Application to the Human Microbiome. *PLoS Comput. Biol.* **8**: e1002358.

De Aguiar Vallim, T.Q., Tarling, E.J., and Edwards, P.A. (2013) Pleiotropic roles of bile acids in metabolism. *Cell Metab.* **17**: 657–669.

Ajouz, H., Mukherji, D., and Shamseddine, A. (2014) Secondary bile acids: an underrecognized cause of colon cancer. *World J. Surg. Oncol.* **12**: 164.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., and Lipman, D.J. (1990) Basic local alignment search tool. *J. Mol. Biol.* **215**: 403–410.

Arkin, A.P., Cottingham, R.W., Henry, C.S., Harris, N.L., Stevens, R.L., Maslov, S., et al. (2018) KBase: the United States department of energy systems biology knowledgebase. *Nat. Biotechnol.* **36**:.

Arumugam, M., Raes, J., Pelletier, E., Le Paslier, D., Yamada, T., Mende, D.R., et al. (2011) Enterotypes of the human gut microbiome. *Nature* **473**: 174–180.

Attene-Ramos, M.S., Wagner, E.D., Plewa, M.J., and Gaskins, H.R. (2006) Evidence that hydrogen sulfide is a genotoxic agent. *Mol. Cancer Res.* **4**: 9–14.

Bäckhed, F., Roswall, J., Peng, Y., Feng, Q., Jia, H., Kovatcheva-Datchary, P., et al. (2015) Dynamics and Stabilization of the Human Gut Microbiome during the First Year of Life. *Cell Host Microbe* **17**: 690–703.

Bajaj, J.S., Heuman, D.M., Hylemon, P.B., Sanyal, A.J., White, M.B., Monteith, P., et al. (2014) Altered profile of human gut microbiome is associated with cirrhosis and its complications. *J. Hepatol.* **60**: 940–7.

Bartoloni, A., Pallecchi, L., Rodríguez, H., Fernandez, C., Mantella, A., Bartalesi, F., et al. (2009) Antibiotic resistance in a very remote Amazonas community. *Int. J. Antimicrob. Agents* **33**: 125–129.

Battaglioli, E.J. and Kashyap, P.C. (2018) Diet Effects on Gut Microbiome Composition, Function, and Host Physiology. *Physiol. Gastrointest. Tract* 755–766.

de Bello, F., Lavorel, S., Díaz, S., Harrington, R., Cornelissen, J.H.C., Bardgett, R.D., et al. (2010) Towards an assessment of multiple ecosystem processes and services via functional traits. *Biodivers. Conserv.* **19**: 2873–2893.

Berry, D. and Widder, S. (2014) Deciphering microbial interactions and detecting keystone species with co-occurrence networks. *Front. Microbiol.* **5**: 219.

Bindels, L.B., Delzenne, N.M., Cani, P.D., and Walter, J. (2015) Towards a more comprehensive concept for prebiotics. *Nat. Rev. Gastroenterol. Hepatol.* **12**: 303–310.

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

Blanton, L. V, Charbonneau, M.R., Salih, T., Barratt, M.J., Venkatesh, S., Ilkaveya, O., et al. (2016) Gut bacteria that prevent growth impairments transmitted by microbiota from malnourished children. *Science (80-.).* **351**: aad3311.

Blekhman, R., Goodrich, J.K., Huang, K., Sun, Q., Bukowski, R., Bell, J.T., et al. (2015) Host genetic variation impacts microbiome composition across human body sites. *Genome Biol.* **16**: 191.

Bonder, M.J., Kurilshikov, A., Tigchelaar, E.F., Mujagic, Z., Imhann, F., Vila, A.V., et al. (2016) The effect of host genetics on the gut microbiome. *Nat. Genet.* **48**: 1407–1412.

Buchfink, B., Xie, C., and Huson, D.H. (2015) Fast and sensitive protein alignment using DIAMOND. *Nat. Methods* **12**: 59–60.

Cantarel, B.L., Lombard, V., and Henrissat, B. (2012) Complex Carbohydrate Utilization by the Healthy Human Microbiome. *PLoS One* **7**: e28742.

Clemente, J.C., Pehrsson, E.C., Blaser, M.J., Sandhu, K., Gao, Z., Wang, B., et al. (2015) The microbiome of uncontacted Amerindians. *Sci. Adv.* **1**:.

Cummings, J.H. and Macfarlane, G.T. (1991) The control and consequences of bacterial fermentation in the human colon. *J. Appl. Bacteriol.* **70**: 443–459.

Cummings, J.H., Pomare, E.W., Branch, H.W.J., Naylor, C.P.E., and MacFarlane, G.T. (1987) Short chain fatty acids in human large intestine, portal, hepatic and venous blood. *Gut* **28**: 1221–1227.

Darzi, Y., Eeckhaudt, L., Zhernakova, A., Vieira-Silva, S., Vandeputte, D., Falony, G., et al. (2016) Population-level analysis of gut microbiome variation. *Science (80-.).* **352**: 560–564.

Das, P., Ji, B., Kovatcheva-Datchary, P., Bäckhed, F., and Nielsen, J. (2018) In vitro co-cultures of human gut bacterial species as predicted from co-occurrence network analysis. *PLoS One* **13**: e0195161.

Davidson, R.T., Foley, A.L., Engelke, J.A., and Suttie, J.W. (1998) Conversion of Dietary Phylloquinone to Tissue Menaquinone-4 in Rats is Not Dependent on Gut Bacteria. *J. Nutr.* **128**: 220–223.

Degnan, B.A. and Macfarlane, G.T. (1995) Carbohydrate utilization patterns and substrate preferences in Bacteroides thetaiotaomicron. *Anaerobe* **1**: 25–33.

Degnan, P.H., Taga, M.E., and Goodman, A.L. (2014) Vitamin B12 as a modulator of gut microbial ecology. *Cell Metab.* **20**: 769–778.

Dethlefsen, L., Huse, S., Sogin, M.L., and Relman, D.A. (2008) The Pervasive Effects of an Antibiotic on the Human Gut Microbiota, as Revealed by Deep 16S rRNA Sequencing. *PLoS Biol.* **6**: e280.

Dethlefsen, L. and Relman, D.A. (2011) Incomplete recovery and individualized responses of the human distal gut microbiota to repeated antibiotic perturbation. *Proc. Natl. Acad. Sci.* **108**: 4554–4561.

Devlin, A.S. and Fischbach, M.A. (2015) A biosynthetic pathway for a prominent class of microbiota-derived bile acids. *Nat. Chem. Biol.* **11**: 685–690.

Dominguez-Bello, M.G., Costello, E.K., Contreras, M., Magris, M., Hidalgo, G., Fierer, N., and Knight, R. (2010) Delivery mode shapes the acquisition and structure of the initial microbiota across multiple body habitats in newborns. *Proc. Natl. Acad. Sci. U. S. A.* **107**: 11971–5.

Duboc, H., Rajca, S., Rainteau, D., Benarous, D., Maubert, M.-A., Quervain, E., et al. (2013) Connecting dysbiosis, bile-acid dysmetabolism and gut inflammation in inflammatory bowel diseases. *Gut* **62**: 531–9.

Duvallet, C., Gibbons, S.M., Gurry, T., Irizarry, R.A., and Alm, E.J. (2017) Meta-analysis of gut microbiome studies identifies disease-specific and shared responses. *Nat. Commun.* **8**: 1784.

Eckburg, P.B., Bik, E.M., Bernstein, C.N., Purdom, E., Dethlefsen, L., Sargent, M., et al. (2005) Diversity of the human intestinal microbial flora. *Science* **308**: 1635–8.

Eitinger, T., Rodionov, D.A., Grote, M., and Schneider, E. Canonical and ECF-type ATP-binding cassette importers in prokaryotes: diversity in modular organization and cellular functions.

Ekmekciu, I., von Klitzing, E., Fiebiger, U., Neumann, C., Bacher, P., Scheffold, A., et al. (2017) The probiotic compound VSL# 3 modulates mucosal, peripheral, and systemic immunity following murine broad-spectrum antibiotic treatment. *Front. Cell. Infect. Microbiol.* **7**: 167.

Etxeberria, U., Arias, N., Boqué, N., Macarulla, M.T., Portillo, M.P., Martínez, J.A., and Milagro, F.I. (2015) Reshaping faecal gut microbiota composition by the intake of trans-resveratrol and quercetin in high-fat sucrose diet-fed rats. *J. Nutr. Biochem.* **26**: 651–660.

Evenepoel, P., Claus, D., Geypens, B., Maes, B., Hiele, M., Rutgeerts, P., and Ghoos, Y. (1998) Evidence for impaired assimilation and increased colonic fermentation of protein, related to gastric acid suppression therapy. *Alimentary pharmacology & therapeutics*, *12*(10), 1011-1019.

Fernandez-Banares, F., Abad-Lacruz, A., Xiol, X., Gine, J.J., Dolz, C., Cabre, E., et al. (1989) Vitamin status in patients with inflammatory bowel disease. *Am. J. Gastroenterol.* **84**: 744–8.

Finegold, S.M. (1996) Anaerobic Gram-Negative Bacilli University of Texas Medical Branch at Galveston.

Finn, R.D., Coggill, P., Eberhardt, R.Y., Eddy, S.R., Mistry, J., Mitchell, A.L., et al. (2016) The Pfam protein families database: towards a more sustainable future. *Nucleic Acids Res.* **44**: D279–D285.

Flowers, S.A., Evans, S.J., Ward, K.M., McInnis, M.G., and Ellingrod, V.L. (2017) Interaction Between Atypical Antipsychotics and the Gut Microbiome in a Bipolar Disease Cohort. *Pharmacother. J. Hum. Pharmacol. Drug Ther.* **37**: 261–267.

Forslund, K., Hildebrand, F., Nielsen, T., Falony, G., Le Chatelier, E., Sunagawa, S., et al. (2015) Disentangling type 2 diabetes and metformin treatment signatures in the human gut microbiota. *Nature* **528**: 262–266.

Forslund, K., Sunagawa, S., Coelho, L.P., and Bork, P. (2014) Metagenomic insights into the human gut resistome and the forces that shape it. *BioEssays* **36**: 316–329.

Frank, D.N., St Amand, A.L., Feldman, R.A., Boedeker, E.C., Harpaz, N., and Pace, N.R. (2007) Molecular-phylogenetic characterization of microbial community imbalances in human inflammatory bowel diseases. *Proc. Natl. Acad. Sci. U. S. A.* **104**: 13780–5.

Franzosa, E.A., Sirota-Madi, A., Avila-Pacheco, J., Fornelos, N., Haiser, H.J., Reinker, S., et al. (2018) Gut microbiome structure and metabolic activity in inflammatory bowel disease. *Nat. Microbiol.* , *4*(2), 293.

Le Gall, G., Noor, S.O., Ridgway, K., Scovell, L., Jamieson, C., Johnson, I.T., et al. (2011) Metabolomics of Fecal Extracts Detects Altered Metabolic Activity of Gut Microbiota in Ulcerative Colitis and Irritable Bowel Syndrome. *J. Proteome Res.* **10**: 4208–4218.

Gardner, R.C., Howarth, A.J., Hahn, P., Brown-Luedi, M., Shepherd, R.J., and Messing, J. (1981) The complete nucleotide sequence of an infectious clone of cauliflower mosaic virus by M13mp7 shotgun sequencing. *Nucleic Acids Res.* **9**: 2871–2888.

Gérard, P. (2013) Metabolism of cholesterol and bile acids by the gut microbiota. *Pathog. (Basel, Switzerland)* **3**: 14–24.

Gibson, P.R. (2017) The evidence base for efficacy of the low FODMAP diet in irritable bowel syndrome: is it ready for prime time as a first-line therapy? *J. Gastroenterol. Hepatol.* **32**: 32–35.

Gill, S.R., Pop, M., Deboy, R.T., Eckburg, P.B., Turnbaugh, P.J., Samuel, B.S., et al. (2006) Metagenomic analysis of the human distal gut microbiome. *Science* **312**: 1355–9.

Goodrich, J.K., Davenport, E.R., Beaumont, M., Jackson, M.A., Knight, R., Ober, C., et al. (2016) Genetic Determinants of the Gut Microbiome in UK Twins. *Cell Host Microbe* **19**: 731–743.

Goodrich, J.K., Waters, J.L., Poole, A.C., Sutter, J.L., Koren, O., Blekhman, R., et al. (2014) Human Genetics Shape the Gut Microbiome. *Cell* **159**: 789–799.

Gothe, F., Beigel, F., Rust, C., Hajji, M., Koletzko, S., and Freudenberg, F. (2014) Bile acid malabsorption assessed by 7 alpha-hydroxy-4-cholesten-3-one in pediatric inflammatory bowel disease: Correlation to clinical and laboratory findings. *J. Crohn's Colitis* **8**: 1072–1078.

Gutin, L., Piceno, Y., Fadrosh, D., Lynch, K., Zydek, M., Kassam, Z., et al. (2019) Fecal microbiota transplant for Crohn disease: A study evaluating safety, efficacy, and microbiome profile. *United Eur. Gastroenterol. J.* 205064061984598.

Han, S.W., Evans, D.G., el-Zaatari, F.A., Go, M.F., and Graham, D.Y. (1996) The interaction of pH, bile, and Helicobacter pylori may explain duodenal ulcer. *Am. J. Gastroenterol.* **91**: 1135–7.

Harris, S.C., Devendran, S., Méndez- García, C., Mythen, S.M., Wright, C.L., Fields, C.J., et al. (2018) Bile acid oxidation by Eggerthella lenta strains C592 and DSM 2243 T. *Gut Microbes* 1–17.

Hill, M.J. (1997) Intestinal flora and endogenous vitamin synthesis. *Eur. J. Cancer Prev.* **6 Suppl 1**: S43-5.

Hofmann, A.F. (2011) Enterohepatic Circulation of Bile Acids. *Compr. Physiol.* 567–596.

Holmes, E., Li, J. V., Marchesi, J.R., and Nicholson, J.K. (2012) Gut microbiota composition and activity in relation to host metabolic phenotype and disease risk. *Cell Metab.* **16**: 559–564.

Hu, Y., Yang, X., Qin, J., Lu, N., Cheng, G., Wu, N., et al. (2013) Metagenome-wide analysis of antibiotic resistance genes in a large cohort of human gut microbiota. *Nat. Commun.* **4**: 2151.

Huang, S., Ma, J., Zhu, M., and Ran, Z. (2017) Status of serum vitamin B12 and folate in patients with inflammatory bowel disease in China. *Intest. Res.* **15**: 103–108.

Huerta-Cepas, J., Szklarczyk, D., Forslund, K., Cook, H., Heller, D., Walter, M.C., et al. (2016) eggNOG 4.5: a hierarchical orthology framework with improved functional annotations for eukaryotic, prokaryotic and viral sequences. *Nucleic Acids Res.* **44**: D286–D293.

Huttenhower, C., Kostic, A.D., and Xavier, R.J. (2014) Inflammatory Bowel Disease as a Model for Translating the Microbiome. *Immunity* **40**: 843–854.

Hylemon, P.B., Zhou, H., Pandak, W.M., Ren, S., Gil, G., and Dent, P. (2009) Bile acids as regulatory molecules. *J. Lipid Res.* **50**: 1509–20.

Ichihashi, T., Takagishi, Y., Uchida, K., and Yamada, H. (1992) Colonic Absorption of Menaquinone-4 and Menaquinone-9 in Rats. *J. Nutr.* **122**: 506–512.

Imhann, F., Bonder, M.J., Vila, A.V., Fu, J., Mujagic, Z., Vork, L., et al. (2016) Proton pump inhibitors affect the gut microbiome. *Gut* **65**: 740–748.

Imhann, F., Vila, A.V., Bonder, M.J., Fu, J., Gevers, D., Visschedijk, M.C., et al. (2018) Interplay of host genetics and gut microbiota underlying the onset and clinical presentation of inflammatory bowel disease. *Gut* **67**: 108–119.

Itoh, M., Wada, K., Tan, S., Kitano, Y., Kai, J., and Makino, I. (1999) Antibacterial action of bile acids against Helicobacter pylori and changes in its ultrastructural morphology: effect of unconjugated dihydroxy bile acid. *J. Gastroenterol.* **34**: 571–6.

Jackson, M.A., Goodrich, J.K., Maxan, M.-E., Freedberg, D.E., Abrams, J.A., Poole, A.C., et al. (2016) Proton pump inhibitors alter the composition of the gut microbiota. *Gut* **65**: 749–56.

Jacobs, J.P., Goudarzi, M., Singh, N., Tong, M., McHardy, I.H., Ruegger, P., et al. (2016) A Disease-Associated Microbial and Metabolomics State in Relatives of Pediatric Inflammatory Bowel Disease Patients. *Cell. Mol. Gastroenterol. Hepatol.* **2**: 750–766.

Jaehme, M. and Slotboom, D.J. (2015) Diversity of membrane transport proteins for vitamins in bacteria and archaea. *Biochim. Biophys. Acta - Gen. Subj.* **1850**: 565–576.

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

Jernberg, C., Löfmark, S., Edlund, C., and Jansson, J.K. (2007) Long-term ecological impacts of antibiotic administration on the human intestinal microbiota. *ISME J.* **1**: 56–66.

Jones, B. V, Begley, M., Hill, C., Gahan, C.G.M., and Marchesi, J.R. (2008) Functional and comparative metagenomic analysis of bile salt hydrolase activity in the human gut microbiome. *Proc. Natl. Acad. Sci. U. S. A.* **105**: 13580–5.

Kakiyama, G., Pandak, W.M., Gillevet, P.M., Hylemon, P.B., Heuman, D.M., Daita, K., et al. (2013) Modulation of the fecal bile acid profile by gut microbiota in cirrhosis. *J. Hepatol.* **58**: 949–55.

Kanehisa, M., Goto, S., Kawashima, S., Okuno, Y., and Hattori, M. (2004) The KEGG resource for deciphering the genome. *Nucleic Acids Res.* **32**: D277–D280.

Kanehisa, M., Goto, S., Sato, Y., Furumichi, M., and Tanabe, M. (2011) KEGG for integration and interpretation of large-scale molecular data sets. *Nucleic Acids Res.* **40**: D109–D114.

Karlsson, F.H., Nookaew, I., Nielsen, J., Ribeca, P., and Berger, S. (2014) Metagenomic Data Utilization and Analysis (MEDUSA) and Construction of a Global Gut Microbial Gene Catalogue. *PLoS Comput. Biol.* **10**: e1003706.

Karlsson, F.H., Tremaroli, V., Nookaew, I., Bergström, G., Behre, C.J., Fagerberg, B., et al. (2013) Gut metagenome in European women with normal, impaired and diabetic glucose control. *Nature* **498**: 99–103.

Koenig, J.E., Spor, A., Scalfone, N., Fricker, A.D., Stombaugh, J., Knight, R., et al. (2011) Succession of microbial consortia in the developing infant gut microbiome. *Proc. Natl. Acad. Sci. U. S. A.* **108 Suppl**: 4578–85.

Koh, A., De Vadder, F., Kovatcheva-Datchary, P., and Bäckhed, F. (2016) From Dietary Fiber to Host Physiology: Short-Chain Fatty Acids as Key Bacterial Metabolites. *Cell* **165**: 1332–1345.

Kovatcheva-Datchary, P., Shoaie, S., Lee, S., Wahlström, A., Nookaew, I., Hallen, A., et al. (2019) Simplified Intestinal Microbiota to Study Microbe-Diet-Host Interactions in a Mouse Model. *Cell Rep.* **26**: 3772-3783.e6.

Kovatcheva-Datchary, P., Tremaroli, V., and Bäckhed, F. (2013) The gut microbiota. In, *The prokaryotes*. Springer, pp. 3–24.

Kristensen, N.B., Bryrup, T., Allin, K.H., Nielsen, T., Hansen, T.H., and Pedersen, O. (2016) Alterations in fecal microbiota composition by probiotic supplementation in healthy adults: a systematic review of randomized controlled trials. *Genome Med.* **8**: 52.

Kumar, M., Babaei, P., Ji, B., and Nielsen, J. (2016) Human gut microbiota and healthy aging: Recent developments and future prospective. *Nutr. Heal. Aging* **4**: 3–16.

Kumar, M., Ji, B., Babaei, P., Das, P., Lappa, D., Ramakrishnan, G., et al. (2018) Gut microbiota dysbiosis is associated with malnutrition and reduced plasma amino acid levels: Lessons from genome-scale metabolic modeling. *Metab. Eng.* **49**:.

Kurilshikov, A., Wijmenga, C., Fu, J., and Zhernakova, A. (2017) Host Genetics and Gut Microbiome: Challenges and Perspectives. *Trends Immunol.* **38**: 633–647.

Kuroki, F., Tominaga, M., Matsumoto, T., Hirakawa, K., Sugiyama, S., and Fujishima, M. (1993) Multiple Vitamin Status in Crohn's Disease Correlation with Disease Activity.

Kurtz, Z.D., Müller, C.L., Miraldi, E.R., Littman, D.R., Blaser, M.J., and Bonneau, R.A. (2015) Sparse and Compositionally Robust Inference of Microbial Ecological Networks. *PLoS Comput. Biol.* **11**: 1–25.

Labbé, A., Ganopolsky, J.G., Martoni, C.J., Prakash, S., and Jones, M.L. (2014) Bacterial Bile Metabolising Gene Abundance in Crohn's, Ulcerative Colitis and Type 2 Diabetes Metagenomes. *PLoS One* **9**: e115175.

Lamas, B., Richard, M.L., Leducq, V., Pham, H.-P., Michel, M.-L., Da Costa, G., et al. (2016) CARD9 impacts colitis by altering gut microbiota metabolism of tryptophan into aryl hydrocarbon receptor ligands. *Nat. Med.* **22**: 598–605.

Lawrence, D., Fiegna, F., Behrends, V., Bundy, J.G., Phillimore, A.B., Bell, T., and Barraclough, T.G. (2012) Species Interactions Alter Evolutionary Responses to a Novel Environment. *PLoS Biol.* **10**: e1001330.

Laxminarayan, R., Duse, A., Wattal, C., Zaidi, A.K.M., Wertheim, H.F.L., Sumpradit, N., et al. (2013) Antibiotic resistance—the need for global

solutions. *Lancet Infect. Dis.* **13**: 1057–1098.

LeBlanc, J.G., Milani, C., de Giori, G.S., Sesma, F., van Sinderen, D., and Ventura, M. (2013) Bacteria as vitamin suppliers to their host: A gut microbiota perspective. *Curr. Opin. Biotechnol.* **24**: 160–168.

Lefcheck, J.S., Bastazini, V.A.G., and Griffin, J.N. (2015) Choosing and using multiple traits in functional diversity research. *Environ. Conserv.* **42**: 104–107.

Lefebvre, P., Cariou, B., Lien, F., Kuipers, F., and Staels, B. (2009) Role of bile acids and bile acid receptors in metabolic regulation. *Physiol. Rev.* **89**: 147–91.

Long, S.L., Gahan, C.G.M., and Joyce, S.A. (2017) Interactions between gut bacteria and bile in health and disease. *Mol. Aspects Med.* **56**: 54–65.

Macfarlane, G.T., Gibson, G.R., and Cummings, J.H. (1992) Comparison of fermentation reactions in different regions of the human colon. *J. Appl. Bacteriol.* **72**: 57–64.

Magee, E.A., Richardson, C.J., Hughes, R., and Cummings, J.H. (2000) Contribution of dietary protein to sulfide production in the large intestine: an in vitro and a controlled feeding study in humans. *Am. J. Clin. Nutr.* **72**: 1488–1494.

Magnúsdóttir, S., Ravcheev, D., de Crécy-Lagard, V., and Thiele, I. (2015) Systematic genome assessment of B-vitamin biosynthesis suggests co-operation among gut microbes. *Front. Genet.* **6**: 148.

Maier, L., Pruteanu, M., Kuhn, M., Zeller, G., Telzerow, A., Anderson, E.E., et al. (2018) Extensive impact of non-antibiotic drugs on human gut bacteria. *Nature* **555**: 623–628.

Marcobal, A., Kashyap, P.C., Nelson, T.A., Aronov, P.A., Donia, M.S., Spormann, A., et al. (2013) A metabolomic view of how the human gut microbiota impacts the host metabolome using humanized and gnotobiotic mice. *ISME J.* **7**: 1933–1943.

Martinot, E., Sèdes, L., Baptissart, M., Lobaccaro, J.-M., Caira, F., Beaudoin, C., and Volle, D.H. (2017) Bile acids and their receptors. *Mol. Aspects Med.* **56**: 2–9.

McGill, B.J., Enquist, B.J., Weiher, E., and Westoby, M. (2006) Rebuilding community ecology from functional traits. *Trends Ecol. Evol.* **21**: 178–185.

McIntosh, K., Reed, D.E., Schneider, T., Dang, F., Keshteli, A.H., Palma, G. De, et al. (2017) FODMAPs alter symptoms and the metabolome of patients with IBS: a randomised controlled trial. *Gut* **66**: 1241–1251.

Mihrshahi, S., Ding, D., Gale, J., Allman-Farinelli, M., Banks, E., and Bauman, A.E. (2017) Vegetarian diet and all-cause mortality: Evidence from a large population-based Australian cohort - the 45 and Up Study. *Prev. Med. (Baltim).* **97**: 1–7.

Molinaro, A., Wahlström, A., and Marschall, H.-U. (2018) Role of Bile Acids in Metabolic Control. *Trends Endocrinol. Metab.* **29**: 31–41.

Montel, M.-C., Buchin, S., Mallet, A., Delbes-Paus, C., Vuitton, D.A., Desmasures, N., and Berthier, F. (2014) Traditional cheeses: Rich and diverse microbiota with associated benefits. *Int. J. Food Microbiol.* **177**: 136–154.

Monteverde, D.R., Gómez-Consarnau, L., Suffridge, C., and Sañudo-Wilhelmy, S.A. (2017) Life's utilization of B vitamins on early Earth. *Geobiology* **15**: 3–18.

Moreno-Indias, I., Sánchez-Alcoholado, L., Pérez-Martínez, P., Andrés-Lacueva, C., Cardona, F., Tinahones, F., and Queipo-Ortuño, M.I. (2016) Red wine polyphenols modulate fecal microbiota and reduce markers of the metabolic syndrome in obese patients. **7**: 1775.

Nielsen, J. (2017) Systems Biology of Metabolism. *Annual review of biochemistry*, *86*, 245-275.

O'Brien, E.J., Monk, J.M., and Palsson, B.O. (2015) Using genome-scale models to predict biological capabilities. *Cell* **161**: 971–987.

O'Keefe, S.J.D., Li, J. V., Lahti, L., Ou, J., Carbonero, F., Mohammed, K., et al. (2015) Fat, fibre and cancer risk in African Americans and rural Africans. *Nat. Commun.* **6**: 6342.

Ogilvie, L.A. and Jones, B. V (2012) Dysbiosis modulates capacity for bile acid modification in the gut microbiomes of patients with inflammatory bowel disease: a mechanism and marker of disease? *Gut* **61**: 1642–3.

Orlich, M.J., Singh, P.N., Sabaté, J., Jaceldo-Siegl, K., Fan, J., Knutsen, S., et al. (2013) Vegetarian Dietary Patterns and Mortality in Adventist Health Study 2. *JAMA Intern. Med.* **173**: 1230.

Ozdal, T., Sela, D.A., Xiao, J., Boyacioglu, D., Chen, F., and Capanoglu, E. (2016) The reciprocal interactions between polyphenols and gut microbiota and effects on bioaccessibility. *Nutrients* **8**: 1–36.

Paño-Pardo, J.R., Schüffelmann-Gutiérrez, C., Escosa-García, L., Laplaza-González, M., Moreno-Ramos, F., Gómez-Gil, R., et al. (2016) Opportunities to improve antimicrobial use in paediatric intensive care units: a nationwide survey in Spain. *Clin. Microbiol. Infect.* **22**: 171–177.

Pehrsson, E.C., Tsukayama, P., Patel, S., Mejía-Bautista, M., Sosa-Soto, G., Navarrete, K.M., et al. (2016) Interconnected microbiomes and resistomes in low-income human habitats. *Nature* **533**: 212–216.

Pitcher, M.C.L. and Cummings, J.H. (1996) Hydrogen sulphide: A bacterial toxin in ulcerative colitis?. *Gut*, *39*(1), 1.

Pols, T.W.H., Puchner, T., Korkmaz, H.I., Vos, M., Soeters, M.R., and de Vries, C.J.M. (2017) Lithocholic acid controls adaptive immune responses by inhibition of Th1 activation through the Vitamin D receptor. *PLoS One* **12**: e0176715.

Qin, J., Li, R., Raes, J., Arumugam, M., Burgdorf, K.S., Manichanh, C., et al. (2010) A human gut microbial gene catalogue established by metagenomic sequencing. *Nature* **464**: 59–65.

Qin, J., Li, Y., Cai, Z., Li, S., Zhu, J., Zhang, F., et al. (2012) A metagenome-wide association study of gut microbiota in type 2 diabetes. *Nature* **490**: 55–60.

Ridlon, J.M., Alves, J.M., Hylemon, P.B., and Bajaj, J.S. (2013) Cirrhosis, bile acids and gut microbiota: unraveling a complex relationship. *Gut Microbes* **4**: 382–7.

Ridlon, J.M., Harris, S.C., Bhowmik, S., Kang, D.-J., and Hylemon, P.B. (2016) Consequences of bile salt biotransformations by intestinal bacteria. *Gut Microbes* **7**: 22–39.

Ridlon, J.M., Kang, D.-J., and Hylemon, P.B. (2006) Bile salt biotransformations by human intestinal bacteria. *J. Lipid Res.* **47**: 241–59.

Ridlon, J.M., Kang, D.J., Hylemon, P.B., and Bajaj, J.S. (2014) Bile acids and the gut microbiome. *Curr. Opin. Gastroenterol.* **30**: 332–8.

Rigottier-Gois, L. (2013) Dysbiosis in inflammatory bowel diseases: the oxygen hypothesis. *ISME J.* **7**: 1256–1261.

Rios-Covian, D., Gueimonde, M., Duncan, S.H., Flint, H.J., and de Los Reyes-Gavilan, C.G. (2015) Enhanced butyrate formation by cross-feeding between Faecalibacterium prausnitzii and Bifidobacterium adolescentis. *FEMS Microbiol. Lett.* **362**(21).

Rivière, A., Selak, M., Lantin, D., Leroy, F., and De Vuyst, L. (2016) Bifidobacteria and Butyrate-Producing Colon Bacteria: Importance and Strategies for Their Stimulation in the Human Gut. *Front. Microbiol.* **7**: 979.

Rodionov, D.A., Hebbeln, P., Eudes, A., ter Beek, J., Rodionova, I.A., Erkens, G.B., et al. (2009) A novel class of modular transporters for vitamins in prokaryotes. *J. Bacteriol.* **191**: 42–51.

Rossi, M., Amaretti, A., and Raimondi, S. (2011) Folate production by probiotic bacteria. *Nutrients* **3**: 118–34.

Rothhammer, V., Mascanfroni, I.D., Bunse, L., Takenaka, M.C., Kenison, J.E., Mayo, L., et al. (2016) Type I interferons and microbial metabolites of tryptophan modulate astrocyte activity and central nervous system inflammation via the aryl hydrocarbon receptor. *Nat. Med.* **22**: 586–597.

Rothschild, D., Weissbrod, O., Barkan, E., Kurilshikov, A., Korem, T., Zeevi, D., et al. (2018) Environment dominates over host genetics in shaping human gut microbiota. *Nature* **555**: 210–215.

Said, H.M. and Mohammed, Z.M. (2006) Intestinal absorption of water-soluble vitamins: An update. *Curr. Opin. Gastroenterol.* **22**: 140–146.

Salyers, A.A., Gupta, A., and Wang, Y. (2004) Human intestinal bacteria as reservoirs for antibiotic resistance genes. *Trends Microbiol.* **12**: 412–416.

Sasaki, D., Sasaki, K., Ikuta, N., Yasuda, T., Fukuda, I., Kondo, A., and Osawa, R. (2018) Low amounts of dietary fibre increase in vitro production of short-chain fatty acids without changing human colonic microbiota structure. *Sci. Rep.* **8**: 435.

Savage, D.C. (2003) Microbial Ecology of the Gastrointestinal Tract. *Annu. Rev. Microbiol.* **31**: 107–133.

Schell, M.A., Karmirantzou, M., Snel, B., Vilanova, D., Berger, B., Pessi, G., et al. (2002) The genome sequence of Bifidobacterium longum reflects its adaptation to the human gastrointestinal tract. *Proc. Natl. Acad. Sci.* **99**: 14422–14427.

Sen, P., Orešič, M., Sen, P., and Orešič, M. (2019) Metabolic Modeling of Human Gut Microbiota on a Genome Scale: An Overview. *Metabolites* **9**: 22.

Shah, H.N., Collins, M.D., Olsen, I., Paster, B.J., and Dewhirsp, F.E. (1995) Reclassification of Bacteroides Zevii (Holdeman, Cato, and Moore) in the Genus Polphyromonas, as Porphyromonas Zevii comb. nov. *Int. J. Syst. Bacteriol.* **45**: 586–588.

Shapiro, H., Suez, J., and Elinav, E. (2017) Personalized microbiome-based approaches to metabolic syndrome management and prevention. *J. Diabetes* **9**: 226–236.

Smith, M.I., Yatsunenko, T., Manary, M.J., Trehan, I., Mkakosya, R., Cheng, J., et al. (2013) Gut microbiomes of Malawian twin pairs discordant for kwashiorkor. *Science* **339**: 548–54.

Sommer, M.O.A., Dantas, G., and Church, G.M. (2009) Functional characterization of the antibiotic resistance reservoir in the human microflora. *Science* **325**: 1128–1131.

Spearman, C. (1904) The Proof and Measurement of Association between Two Things. *Am. J. Psychol.* **15**: 72.

Subramanian, S., Huq, S., Yatsunenko, T., Haque, R., Mahfuz, M., Alam, M.A., et al. (2014) Persistent gut microbiota immaturity in malnourished Bangladeshi children. *Nature* **510**: 417–421.

Suez, J., Zmora, N., Zilberman-Schapira, G., Mor, U., Dori-Bachash, M., Bashiardes, S., et al. (2018) Post-Antibiotic Gut Mucosal Microbiome Reconstitution Is Impaired by Probiotics and Improved by Autologous FMT. *Cell* **174**: 1406-1423.e16.

Sumi, Y., Miyakawa, M., Kanzaki, M., and Kotake, Y. (1977) Vitamin B-6 Deficiency in Germfree Rats. *J. Nutr.* **107**: 1707–1714.

Tanoue, T., Morita, S., Plichta, D.R., Skelly, A.N., Suda, W., Sugiura, Y., et al. (2019) A defined commensal consortium elicits CD8 T cells and anti-cancer immunity. *Nature* **565**: 600–605.

Taur, Y., Coyte, K., Schluter, J., Robilotti, E., Figueroa, C., Gjonbalaj, M., et al. (2018) Reconstitution of the gut microbiota of antibiotic-treated patients by autologous fecal microbiota transplant.

The Integrative HMP Research Network Consortium (2014) The Integrative Human Microbiome Project: Dynamic Analysis of Microbiome-Host Omics Profiles during Periods of Human Health and Disease. *Cell Host Microbe* **16**: 276–289.

Thompson, S. V, Hannon, B.A., An, R., and Holscher, H.D. (2017) Effects of isolated soluble fiber supplementation on body weight, glycemia, and insulinemia in adults with overweight and obesity: a systematic review and meta-analysis of randomized controlled trials. *Am. J. Clin. Nutr.* **106**: 1514–1528.

Toden, S., Bird, A.R., Topping, D.L., and Conlon, M.A. (2005) Resistant starch attenuates colonic DNA damage induced by higher dietary protein in rats. *Nutr. Cancer* **51**: 45–51.

Turnbaugh, P.J., Hamady, M., Yatsunenko, T., Cantarel, B.L., Duncan, A., Ley, R.E., et al. (2009) A core gut microbiome in obese and lean twins. *Nature* **457**: 480–484.

Turnbaugh, P.J., Ley, R.E., Hamady, M., Fraser-Liggett, C.M., Knight, R., and Gordon, J.I. (2007) The human microbiome project. *Nature* **449**: 804–10.

Turpin, W., Espin-Garcia, O., Xu, W., Silverberg, M.S., Kevans, D., Smith, M.I., et al. (2016) Association of host genome with intestinal microbial composition in a large healthy cohort. *Nat. Genet.* **48**: 1413–1417.

UniProt Consortium (2018) UniProt: a worldwide hub of protein knowledge. *Nucleic Acids Res.* **47**: D506–D515.

De Vadder, F., Kovatcheva-Datchary, P., Goncalves, D., Vinera, J., Zitoun, C., Duchampt, A., et al. (2014) Microbiota-generated metabolites promote metabolic benefits via gut-brain neural circuits. *Cell* **156**: 84–96.

Vanholder, R., De Smet, R., and Lesaffer, G. (1999) p-Cresol: a toxin revealing many neglected but relevant aspects of uraemic toxicity. *Nephrol. Dial. Transplant.* **14**: 2813–2815.

Veronese, N., Solmi, M., Caruso, M.G., Giannelli, G., Osella, A.R., Evangelou, E., et al. (2018) Dietary fiber and health outcomes: an umbrella review of systematic reviews and meta-analyses. *Am. J. Clin. Nutr.* **107**: 436–444.

Vich Vila, A., Imhann, F., Collij, V., Jankipersadsing, S.A., Gurry, T., Mujagic, Z., et al. (2018) Gut microbiota composition and functional changes in inflammatory bowel disease and irritable bowel syndrome. *Sci. Transl. Med.* **10**: eaap8914.

Wahlström, A., Kovatcheva-Datchary, P., Ståhlman, M., Bäckhed, F., and Marschall, H.U. (2017) Crosstalk between Bile Acids and Gut Microbiota and Its Impact on Farnesoid X Receptor Signalling. *Dig. Dis.* **35**: 246–250.

Wahlström, A., Sayin, S.I., Marschall, H.-U., and Bäckhed, F. (2016) Intestinal Crosstalk between Bile Acids and Microbiota and Its Impact on Host Metabolism. *Cell Metab.* **24**: 41–50.

Wall, R., Ross, R.P., Ryan, C.A., Hussey, S., Murphy, B., Fitzgerald, G.F., and Stanton, C. (2009) Role of gut microbiota in early infant development. *Clin. Med. Pediatr.* **3**: 45–54.

Walter, J., Maldonado-Gómez, M.X., and Martínez, I. (2018) To engraft or not to engraft: an ecological framework for gut microbiome modulation with live microbes. *Curr. Opin. Biotechnol.* **49**: 129–139.

Wang, J., Thingholm, L.B., Skiecevičienė, J., Rausch, P., Kummen, M., Hov, J.R., et al. (2016) Genome-wide association analysis identifies variation in vitamin D receptor and other host factors influencing the gut microbiota. *Nat. Genet.* **48**: 1396–1406.

Wolever, T.M., Jenkins, D.J., Jenkins, A.L., and Josse, R.G. (1991) The glycemic index: methodology and clinical implications. *Am. J. Clin. Nutr.* **54**: 846–854.

Wright, G.D. (2007) The antibiotic resistome: the nexus of chemical and genetic diversity. *Nat. Rev. Microbiol.* **5**: 175–186.

Wu, G.D., Compher, C., Chen, E.Z., Smith, S.A., Shah, R.D., Bittinger, K., et al. (2016) Comparative metabolomics in vegans and omnivores reveal constraints on diet-dependent gut microbiota metabolite production. *Gut* **65**: 63–72.

Xie, H., Guo, R., Zhong, H., Feng, Q., Lan, Z., Qin, B., et al. (2016) Shotgun Metagenomics of 250 Adult Twins Reveals Genetic and Environmental Impacts on the Gut Microbiome. *Cell Syst.* **3**: 572-584.e3.

Xu, J., Bjursell, M.K., Himrod, J., Deng, S., Carmichael, L.K., Chiang, H.C., et al. (2003) A Genomic View of the Human-Bacteroides thetaiotaomicron Symbiosis. *Science (80-.).* **299**:.

Yatsunenko, T., Rey, F.E., Manary, M.J., Trehan, I., Dominguez-Bello, M.G., Contreras, M., et al. (2012) Human gut microbiome viewed across age and geography. *Nature* **486**: 222–227.

Yokota, A., Fukiya, S., Islam, K.B.M.S., Ooka, T., Ogura, Y., Hayashi, T., et al. (2012) Is bile acid a determinant of the gut microbiota on a high-fat diet? *Gut Microbes* **3**: 455–459.

Zeevi, D., Korem, T., Godneva, A., Bar, N., Kurilshikov, A., Lotan-Pompan, M., et al. (2019) Structural variation in the gut microbiome associates with host health. *Nature* **568**: 43–48.

Zeevi, D., Korem, T., Zmora, N., Israeli, D., Rothschild, D., Weinberger, A., et al. (2015) Personalized Nutrition by Prediction of Glycemic Responses. *Cell* **163**: 1079–1095.

Zhao, L., Zhang, F., Ding, X., Wu, G., Lam, Y.Y., Wang, Xuejiao, et al. (2018) Gut bacteria selectively promoted by dietary fibers alleviate type 2 diabetes. *Science (80-.).* **359**: 1151–1156.

Zheng, H., Yde, C.C., Clausen, M.R., Kristensen, M., Lorenzen, J., Astrup, A., and Bertram, H.C. (2015) Metabolomics Investigation To Shed Light on Cheese as a Possible Piece in the French Paradox Puzzle. *J. Agric. Food Chem.* **63**: 2830–2839.

Zhou, H. and Hylemon, P.B. (2014) Bile acids are nutrient signaling hormones. *Steroids* **86**: 62–8.

Zmora, N., Zilberman-Schapira, G., Suez, J., Mor, U., Dori-Bachash, M., Bashiardes, S., et al. (2018) Personalized Gut Mucosal Colonization Resistance to Empiric Probiotics Is Associated with Unique Host and Microbiome Features. *Cell* **174**: 1388-1405.e21