Development and dynamics of the normal gut microbiota

Lisa Olsson

Department of Molecular and Clinical Medicine
Institute of Medicine
Sahlgrenska Academy, University of Gothenburg

UNIVERSITY OF GOTHENBURG

Gothenburg 2020

Cover illustration: Anna Hallén

Development and dynamics of the normal gut microbiota

lisa.olsson@wlab.gu.se

ISBN 978-91-7833-886-3 (PRINT)
ISBN 978-91-7833-887-0 (PDF)

Printed in Gothenburg, Sweden 2020
Printed by STEMA SPECIALTRYCK AB

To Elias

“It’s not the mountain we conquer, but ourselves.”

– Edmund Hillary

Development and dynamics of the normal gut microbiota

Lisa Olsson

Department of Molecular and Clinical Medicine, Institute of Medicine
Sahlgrenska Academy, University of Gothenburg
Gothenburg, Sweden

ABSTRACT

Altered gut microbiota configurations have been linked to human diseases. To identify mechanistic links between altered gut microbiota and disease states, definitions of the healthy gut microbiota need to be established. Therefore, in this thesis, we investigated how the gut microbiota develops in Swedish children up to 5 years of age and characterized dynamics of the adult gut microbiota in a normal Swedish population. Using a longitudinal design to study the gut microbiota in both the Swedish children and adults, we identified complex sets of bacteria acquired by the children during their development and compared them to the gut microbiota of the adult population. We identified features of the gut microbiota that were associated to richness at different stages of a child's gut microbiota development.

In the adult Swedish population, we analyzed how the composition and functional potential of the gut microbiota fluctuate over the course of a year in normal population aged 50-64 years. We characterized the total variability of the gut microbiota and determined to which extent gut microbiota variability between individuals is due to intra-individual variability over time. We observed large fluctuations in abundance of facultative anaerobes and in potential bacterial functions, identified from metagenomic analysis, linked to these bacteria. Interestingly, large fluctuations of the facultative anaerobes were indicative of highly variable individual gut microbiota composition.

In the third study in this thesis, we investigated the gut microbiota in relation to obesity and insulin resistance. Here we characterized the gut microbiota in morbidly obese individuals with the genetic Prader-Will syndrome and in obese people matched for fat mass composition. Less insulin resistance and healthier blood lipid in the individuals with Prader-Willi were associated with a less heterogeneous gut microbiota composition as well as higher diversity, which are important ecological features of a stable and resilient microbial community. Importantly, these potentially beneficial microbes were also observed to link to community richness in the children and adult Swedish populations. In summary, we identified gut microbes that associate to community stability and community richness in children as well as adults, and that may play a key role for metabolic health.

Keywords: dynamics, ecology, gut microbiome, gut microbiota development, microbiota, richness, Prader-Willi Syndrome, stability, variation

ISBN 978-91-7833-886-3 (PRINT)
ISBN 978-91-7833-887-0 (PDF)

SAMMANFATTNING PÅ SVENSKA

De mikroorganismer som växer och samverkar i en specifik miljö kallas för mikrobiota. Vi har mikroorganismer, till största delen bakterier, både på och i våra kroppar, på huden, i munnen samt i våra tarmar, överallt där vår kropp möter omvärlden. De flesta bakterierna finns i våra tarmar, även kallad tarmmikrobiotan och utgör 1–1,5 kg av vår kroppsvikt. I livmodern har fostret inga bakterier, men koloniseras under födseln och de närmsta dagarna. Sammansättningen av bakterier har utvecklats tillsammans med oss, till exempel av den kost vi äter, och bidrar till att vi håller oss friska och mår bra. Tarmmikrobiotan producerar vitaminer och utbildar immunsystemet samt bryter ner fibrer i vår kost som vår egen kropp inte kan bryta ner. Beroende på sammansättningen av olika bakterier, och vilka ämnen de producerar när de växer, påverkar de inte bara tarmen utan dessa ämnen kan även transporteras med blodet till andra delar av vår kropp.

Forskningen har kunnat koppla en förändrad sammansättning av tarmmikrobiota från patienter jämfört med friska. En förändrad tarmmikrobiota har setts i flera olika sjukdomar så som inflammatoriska tarmsjukdomar, kardiovaskulära sjukdomar och typ-2 diabetes. Däremot är det inte känt om detta beror på att den förändrade tarmmikrobiotan kan orsakar sjukdom eller om sjukdomen i sig förändrar tarmmikrobiotans sammansättning.

Efter födseln och den första koloniseringen har vi en väldigt enkel tarmflora som är anpassad för att bryta ner bröstmjölk, vilket är den föda vi främst får i oss under vårt första levnadsår. I takt med att vi börjar äta mer och mer fast föda börjar vår tarmmikrobiota utvecklas och utökas med flera olika typer av mikroorganismer som kan utföra mer komplexa uppgifter. Mitt arbete har visat att friska barn genomgår den här förändringen med olika hastighet. Fram till nu har man trott att barn har en vuxen tarmmikrobiota vid 3 års ålder men vi visar att barn som är 5 år fortfarande har en tarmmikrobiota som är enklare, med lägre artrikedom än en vuxens tarmmikrobiota. Femåriga barn har dessutom lägre halter än vuxna av vissa mikroorganismer som vi såg introduceras sent i tarmmikrobiotans utveckling.

Jämfört med barns tarmmikrobiota förändras vuxnas tarmmikrobiota betydligt mindre över tid. Även om tarmmikrobiotan har samma uppgifter i oss alla så kan dessa utföras av olika bakterier i dig och mig. När vi undersöker tarmmikrobiotan studerar vi oftast sammansättningen av bakterier i vår avföring som även i friska individer kan se väldigt olika ut. Detta är väldigt

viktigt att ta hänsyn till när man skall försöka identifiera vad som karaktäriserar en sjuk tarmmikrobiota.

Vår tarmmikrobiota är som vilket annat ekosystem, till exempel en blandskog eller ett korallrev. Dessa gör sitt bästa för att anpassa sig och återhämta sig efter potentiella förändringar, likt skogar om våren eller efter en skogsbrand. För att lyckas med detta har alla arter i ekosystemet olika roller för att tillsammans utföra ekosystemets viktiga funktioner. Om förändringen blir för stor, eller om små förändringar tar död på arter, så tappar ekosystemet sin förmåga att återställa sig, likt den korallblekning vi ser i runt om i världens hav. För att identifiera ett sjukt ekosystem behöver vi särskilja mellan de förändringar som, till exempel återkommande årstiderna utgör, från de förändringar som ger ekosystemet bestående men. På motsvarande sätt utsätts tarmmikrobiotan av olika förändringar i miljön. Beroende på vad vi äter, om vi tar antibiotika eller andra läkemedel, samt förändringar i kroppen när vi blir sjuka förändras förutsättningarna för tarmmikrobiotan. För att förstå hur en sjuk tarmmikrobiota reagerar mot förändringar behöver vi till en början veta hur en frisk tarmflora reagerar. I arbete inkluderade i den här avhandlingen har vi tittat på hur mycket tarmmikrobiotan förändras i friska individer, i åldrarna 50–64 år, genom att studera deras tarmflora vid 4 tillfällen under ett år. Vi såg att varje individ har en egen specifik sammansättning och förändringen mellan varje individs prov är betydligt mindre än mellan individers. Vi noterade även att olika bakterier varierar olika mycket över tid. Vissa bakterier har samma nivå i alla 4 prover medan andra varierar lika mycket inom en individ som nivån mellan individer. För att kunna använda människors tarmmikrobiota för att utvärdera sjukdom och hälsa måste markörer undvikas som har en stor variation inom friska individer.

I det tredje arbete i den här avhandlingen jämförde vi individer med ett genetiskt syndrom, kallat Prader-Willi syndrom vilket leder till fetma, och individer med fetma orsakad av livsstil. Dessa individer med Prader-Willi syndrom har, trots sin fetma, färre följdsjukdomar. Vi såg att deras tarmmikrobiota var mer homogen, vilket skulle kunna vara ett tecken på en mer stabil tarmmikrobiota. Dessa individer hade även en högre artrikedom, vilket är kopplat till en frisk tarmmikrobiota i flertal studier. Vi såg även att mikroorganismer kopplat till färre följdsjukdomar till fetman samt hög artrikedom var bland de mikroorganismer som introduceras sent i barns tarmmikrobiotautveckling.

LIST OF PAPERS

This thesis is based on the following studies, referred to in the text by their Roman numerals.

I. Josefine Roswall*, Lisa M Olsson*, Petia Kovatcheva-Datchary, Staffan Nilsson, Rozita Akrami, Valentina Tremaroli, Marie-Christine Simon, Manuela Krämer, Mathias Uhlén, Göran Bergstöm, Karsten Kristiansen, Jovanna Dahlgren, Fredrik Bäckhed. **Developmental trajectory of the healthy human gut microbiota during the first 5 years of life.**
Manuscript

II. Lisa M Olsson, Fredrik Boulund, Valentina Tremaroli, Staffan Nilsson, Anders Gummesson, Linn Fagerberg, Lars Engstrand, Mathias Uhlén, Göran Bergström, Fredrik Bäckhed. **Dynamics of the gut microbiota in a normal population: a prospective 1-year study.**
Manuscript

III. Lisa M Olsson, Christine Poitou, Valentina Tremaroli, Muriel Coupaye, Judith Aron-Wisnewsky, Fredrik Bäckhed, Karine Clément, Robert Caesar**. Gut microbiota of obese subjects with Prader-Willi syndrome is linked to metabolic health.**
Gut (2019), doi:10.1136/gutjnl-2019-319322

CONTENT

ABBREVIATIONS

ASV	Amplicon sequence variant
CAZy	Database of Carbohydrate-Active enzymes (CAZymes)
COG	Clusters of Orthologous Groups of proteins
CRT	Conditional rare taxa
ESV	Exact sequence variant
GO	Gene Ontology
KEGG	Kyoto Encyclopedia of Genes and Genomes
KO	KEGG ontology
metaCyc	Database with metabolic pathways from all domains in life
MPS	Massively parallel sequencing
OTU	Operational taxonomical unit
PWS	Prader-Willi syndrome
SCFA	Short-chain fatty acid
zOTU	Zero-radius Operational Taxonomical Unit

DEFINITIONS IN SHORT

Alpha diversity	Measurement of within sample diversity. Richness or evenness of microorganism within a sample
Beta diversity	Measurement of between sample diversity. How similar or different the microbial composition in one sample is compared to another
Community types	Clusters of samples with similarities in relative abundance of different genera
Conditionally rare taxa	Rare microbial taxa that occasionally become very abundant. Defined by Shade et al. in 2014
Core microbiota	Microorganisms in a microbiota shared by a large majority of individuals
Enterotype	Community types in adults as defined by Arumugam et al. in 2011
Functional potential	Encoded functions present in a metagenome
Gene richness	Measure of within sample diversity. Number of genes, with more than one count, within a sample. Defined by Le Chatelier et al. in 2013
Keystone species	Microorganisms performing a key function for the ecosystem in the microbiota
Metagenome	The collective genetic content from microorganism in a specific environment
Microbiome	Microorganisms, their genomes and specific conditions in an environment
Microbiota	Collection of all microorganisms present in a specific environment

1 INTRODUCTION

In this thesis I will discuss different types of dynamics in the gut microbiota, the microorganism that live in our gastrointestinal tract. Dynamics that will be addressed are how the gut microbiota is assembled during the gut microbiota development in childhood but also fluctuations in the gut microbiota in adulthood, when the composition has stabilized. Since the gut microbiota is an ecological system that constantly is exposed to environmental fluctuations, for example from what we eat and do, we need to understand how the gut microbiota vary in the context of non-disease in order to understand the gut microbiota in disease.

1.1 THE MICROBIOTA

We live in symbiosis with diverse communities of microbes. The number of microbes on our bodies correspond to at least the number of human cells (Sender et al., 2016). These microbes colonize almost all surfaces of our body, ours skin, teeth, airways and our gastrointestinal tract (Human Microbiome Project, 2012) but also the stomach (Nardone and Compare, 2015) and the vagina (Greenbaum et al., 2019). A microbiota is defined as the community living in a specific environment. Thus, each body site has its unique microbiota that can be affected by environmental factors and has the potential to interact with the host.

During normal pregnancy the fetus is considered sterile while still in the womb (de Goffau et al., 2019). After birth the newborn is immediately exposed to bacteria, originating from the mother and the environment. During the first weeks the microbiota expands and diversifies and at 6 weeks of age body site-specific microbiotas can be differentiated (Chu et al., 2017).

The adult gut harbors the largest and most complex community on the human body. The microbial density in the gastrointestinal tract increases from the stomach to the distal end and comprise a biomass of 1.5-2kg. The gut microbiota is dominated by anaerobic Bacteria and Archaea. Eukarya and viruses are also present. However, knowledge about their influence on composition and human health is limited and will not be addressed in this thesis. The vast majority of the around 1000 bacterial and archaeal species in the gut microbiota belong to 5 phyla (Firmicutes, Bacteroidetes, Actinobacteria, Proteobacterium and Verrucomicrobia) (Qin et al., 2010).

These microbes found on human bodies encode a diverse range of genes and this collective genetic potential is called the metagenome.

1.2 THE GUT MICROBIOTA FUNCTIONS

The gut microbiota plays critical roles for the development and physiology of the host (Sommer and Bäckhed, 2013). Through interaction and co-development with the host it affects maturation of the immune system (Belkaid and Hand, 2014). It also influences the innate immune system and provides protection against pathogenic organisms (Kamada et al., 2013). The presence of bacteria also affects local physiology in the gut such as proliferation of host cells and vascular remodeling (Reinhardt et al., 2012). Normal colonic epithelial differentiation requires metabolism through the nuclear receptor PPARγ, for which the microbial derived short chain fatty acid (SCFA) butyrate is a substrate (Byndloss et al., 2017).

An non-leaky intestinal barrier with efficient tight junctions between epithelial cells is important to avoid that bacteria or bacterial products translocate to the circulation (Ghosh et al., 2020). For example, increased levels of inflammatory microbial products, such as lipopolysaccharide (LPS), in the blood give rise to metabolic endotoxemia, which is linked to metabolic complications (Caesar et al., 2012; Cani et al., 2007). The colonic mucin layers are also part the of the physical barrier between the gut bacteria and epithelial cells. Different bacterial composition affects both the penetrability and growth rate of the mucin layers, which prevents microbes from reaching the epithelium (Schroeder et al., 2018).

The gut microbiota affects the host's metabolism through several mechanisms. For example, it contributes to energy harvest by producing Short-chain fatty acids (SCFA) from carbohydrates that cannot be digested by the host. Thus, the absence of a gut microbiota give rise to lower body fat and increased energy excretion in the feces. To compensate for the energy loss germ free mice have increased food intake (Bäckhed et al., 2004; Bäckhed et al., 2007).

Metabolites from the gut microbiota can also influence systemic metabolism by acting as signaling molecules. By translocating from the gut to the systemic circulation metabolites can affect distant organs directly. Alternatively they can stimulate hormone secretion and neural signaling (Schroeder and Bäckhed, 2016). For example, SCFAs can regulate host metabolism through the release of GLP-1 from intestinal L-cells by binding to G-protein-coupled receptors. Microbial regulated GLP-1 has been shown to affect gut transit, insulin release and energy intake (Greiner and Bäckhed, 2016). In addition, SCFAs can also

alter histone modifications, resulting in changes in transcription, in different tissues (Krautkramer et al., 2016).

Another group of microbially modified metabolites that can affect host physiology are bile acids (Wahlstrom et al., 2016). Primary bile acids are released by the host in the small intestine after a meal. Members of the gut microbiota can modify primary bile acids through de-conjugation in the small intestine. Those that have not been reabsorbed are further transformed in the colon resulting in a variety of secondary bile acids. These bile acids have different ability to activate or inhibit the nuclear receptor farnesoid X receptor (FXR) and the membrane bound G-coupled receptor (TGR5), which both regulate host metabolism.

Finally, amino acid-derived metabolites produced by the gut microbiota can affect host physiology. For example, the gut microbiota of type-2-diabetes patients have altered histidine metabolism compared to healthy subjects, resulting in the histidine derived metabolite imidazole propionate. Imidazole propionate has recently been shown to impair insulin signaling through reduction of insulin receptor substrate in the liver (Koh et al., 2018).

1.3 THE NORMAL GUT MICROBIOTA

The micro-organisms on and inside our body have been studied for centuries. Antonie van Leeuwenhoek was the first to describe ‘animalcules’ in the 1670’s, which he found in his own and other people’s mouth and feces (Dunn and Jones, 2004). Before the 1990’s studies of the human gut microbiota were dependent on culturing methods. Since the majority of microorganisms in our gut are challenging to culture the diversity of the communities had been underestimated (Eckburg et al., 2005). In the 1990’s studies using molecular methods were introduced. Sequencing of marker genes was initially performed using Sanger sequencing but through the introduction of massively paralleled sequencing (MPS), in the beginning of this millennium, it has been possible to study the human microbiota in much larger scale.

1.3.1 GUT MICROBIOTA ESTABLISHMENT

From birth the gut microbiota co-develops together with the rest of the physiology of the child and the maturity of the immune system (Gensollen et al., 2016) and other metabolic and physiological processes (Belkaid and Hand, 2014). The early development of the gut microbiota is a dynamic process which are strongly influenced by maternal microbiome transmission (Ferretti et al., 2018; Korpela et al., 2018) and external factors, such as mode of birth

(Dominguez-Bello et al., 2010; Dominguez-Bello et al., 2016; Reyman et al., 2019; Shao et al., 2019). After the first weeks, the development of the microbiota is linked to macronutrient intake and thus breastfeeding has significant effects on the gut microbiota compared to formula feeding (Baumann-Dudenhoeffer et al., 2018; Bokulich et al., 2016), by for example increase abundance of *Bifidobacterium*. The infant gut microbiota is characterized by a low community richness and heterogenous microbiota. Facultative anaerobes are the first colonizers followed by more oxygen sensitive bacteria such as *Bacteroides* and *Bifidobacterium* (Bäckhed et al., 2015; Eggesbo et al., 2011; Mackie et al., 1999).

It has been shown that the gut microbiota in children matures into an adult-like configuration after 2-3 years (Bergstrom et al., 2014; Koenig et al., 2011; Yatsunenko et al., 2012). Indeed, in infancy, species richness of the gut microbiota is low and its overall composition is highly heterogeneous, as estimated by dissimilarity indexes of beta diversity (e.g., Bray-Curtis and UniFrac). However, with the introduction of solid foods, and cessation of breastfeeding, community richness and complexity of the microbiota increase and with that an altered bacterial functional potential (Bäckhed et al., 2015; Yatsunenko et al., 2012). However, the knowledge about the assemble of an adult-like microbiota after the introduction of solid food and after the first 2-3 years of life is much more limited compare to the microbiota in infancy (Derrien et al., 2019).

Continuous sampling of healthy children from Bangladesh (Subramanian et al., 2014), Malawi (Blanton et al., 2016) and United States (Planer et al., 2016) until 24 to 36 months of age, have provided models of the gut microbiota maturity and identified important age-discriminatory taxa for normal gut microbiota development. Models based on the individual cohorts across geographical locations have several age-discriminatory taxa in common and the model based on the children from United States performed consistently across the three different cohorts (Planer et al., 2016), suggesting similar dynamics independent of geography. Among the age-discriminatory taxa, consistent between cohorts, they observed different *Bifidobacterium* taxa dominated in the young ages whereas *Faecalibacterium*, *Ruminococcus* and *Clostridium* increased with age. Using these models, they connected a less mature gut microbiota with undernourished growth phenotypes with a lower age-adjusted alpha diversity observed in severe malnourished children (Subramanian et al., 2014). By administering diets designed to promote age-discriminatory taxa, which is underrepresented in children with acute malnourished children, the gut microbiota development could be improved in malnourished children (Gehrig et al., 2019).

There are several factors affecting the composition of the gut microbiota, and these factors differ during the different stages of human life. Factors at birth usually have strong effect on the gut microbiota composition during the first year of life, such as mode of birth (Dominguez-Bello et al., 2010; Dominguez-Bello et al., 2016; Reyman et al., 2019; Shao et al., 2019) and maternal microbiome transmission (Ferretti et al., 2018; Korpela et al., 2018). Other factors in infancy such as feeding type (breast milk or formula feeding) also affect the developing gut microbiota during infancy (Bäckhed et al., 2015; Bokulich et al., 2016). Additional factors which affect the gut microbiota development are antibiotic use (Bokulich et al., 2016; Korpela et al., 2016). Korpela et al. observed a long-term effect of antibiotic on the gut microbiota use in 7-year-old children. However, in this study it was not clear if it was early exposure during important periods of development or multiple treatment in the first 4 years which were the most contributing factor. Together with the development of the immune system these factors affect the microbiota assembly, the order of species arrival and the timing of their arrival (also called priority effects) during the first years of life (Sprockett et al., 2018).

1.3.2 GUT MICROBIOTA VARIATION

In recent years several studies have sought to characterize the human gut microbiota and its metagenome. In particular the bacterial components of the gut microbiota, their structure and function in healthy adult subjects.

From large studies, mainly from American, European and Chinese populations, we know that there is a large variation in composition between individuals (Falony et al., 2016; Human Microbiome Project, 2012; Qin et al., 2010). The effect size of different factors on compositional variation have been studied and medication and stool consistency have been identified as the most contributing factors (Falony et al., 2016). Genetic ancestry in a population with similar lifestyles was identified to have a minor contribution to the compositional variation (Rothschild et al., 2018). Other environmental factors, such as diet, medications and anthropometric measurements (such as BMI) were responsible for around 20% of the compositional variation between individuals. Although individual taxa, which contributes to the total composition to lesser extent, have been identified as heritable (Goodrich et al., 2016; Goodrich et al., 2014). Large regional differences in the gut microbiota composition have been found. These differences are present both between regions in the same country and between different ethnic backgrounds in the same city (Deschasaux et al., 2018; He et al., 2018b). Immigration from southeast Asia to the United States have also been found to have large effects on the gut microbiota (Vangay et al., 2018). Regional differences have been

found to explain more inter-individual variation than disease, highlighting the importance of local baselines (He et al., 2018b).

To understand fundamental properties of the gut microbiota and identify species that are essential for gut microbiota function, efforts in identification of a core microbiota have been done (Qin et al., 2010). From one study of a European population the core microbiota, which was defined as taxa shared between 95% of the individuals, consisted of 35 genera (Falony et al., 2016). These genera contributed, in median, to 90% of the total abundance in this population. When extending the population to also including other western populations this core microbiota decreased to 17 genera with a median core abundance of 72%. When further extending the population to include samples from Papa New Guinea, Peru and Tanzania the core microbiota was further reduced to 14 genera.

1.3.3 GUT MICROBIOTA DYNAMICS

In contrast to the gut microbiota composition in children the gut microbiota composition in adults are considered stable over time. This has been seen in a number of studies where the composition is on average more similar between samples from the same individual compare to samples from other individuals (Caporaso et al., 2011; Costello et al., 2009; Faith et al., 2013; Rajilic-Stojanovic et al., 2012; Schloissnig et al., 2013; Zoetendal et al., 1998). This is seen over the course of a year up to 10 years (Rajilic-Stojanovic et al., 2012).

Due to the stability of the gut microbiota composition it has been suggested that individuals can be distinguished by stable and unique fingerprints based on their microbiota profile. Franzosa et al. constructed codes from variation in clade-specific marker genes from which individuals could be identified when repeatedly sampled in more than 80% of the time. The codes were based on stable features that positively correlated with features abundance and prevalence. They found gene-level codes to be more stable compared to taxon-level codes (Franzosa et al., 2015). Schloissing et al. also conclude that individuals can be distinguished based on variation patterns of the genomic content in the metagenome but not on abundance on species level (Schloissnig et al., 2013).

Low abundant species that occasionally become abundant member of a community have been defined as conditionally rare taxa (CRT) (Shade et al., 2014). In other environments, these CRTs have been seen to affect over all community composition (Shade and Gilbert, 2015). In the human gut many of the CRT are facultative anaerobes (Gibbons et al., 2017). Two types of different dynamics have been proposed in the human gut microbiota based on

densely sampled time-series (Gibbons et al., 2017). The first dynamics was characterized by day-to-day variation that cannot be predicted from previous samples, these effects are most likely due to external factors such as diet. The second dynamics was observed for abundances that were predictable from previous samples, which was followed by large deviations in composition. This dynamic involved bloom of facultative anaerobes, followed by re-establishment of strict anaerobes. The pattern of blooms was different in the 4 individuals from which these time-series came from. In one time-series no blooms were seen while in another they were frequent, all 4 time-series were from healthy individuals.

1.3.4 FUNCTIONAL REDUNDANCY

Although large variation in species abundance between individuals in normal populations the variation in functional potential is in general small (Human Microbiome Project, 2012; Turnbaugh et al., 2009). This observation indicate that the microbiome includes specific functional processes, which are important for the host, but can be performed by different microbial constellations under different conditions. This is called functional redundancy, or functional response diversity, and is suggested to be important for ecological stability of a microbial community (Lozupone et al., 2012). Functional redundancy is acquired already during the development of the gut microbiota. Intra-individual compositional variation responsible for functions increases from infancy up to 3 years of age along with the overall community richness of the microbiota and the richness of taxa responsible for functions (Vatanen et al., 2019). This ‘minimal gut genome’ consists of functions that are present in all bacteria, such as functions of microbial reproduction and structural components, but also functions which are potentially specific for the gut. Among the functions in the ‘minimal gut genome’, which are potentially gut specific, more than 70% are not known. In the known fraction of gut-specific functions are the majority within potential for degradation of sugar or complex polysaccharides from the diet or mucosa lining Examples are degradation and uptake of pectin, sorbitol, mannose, fructose, cellulose and sucrose (Qin et al., 2010). Thus it has been suggested that characterization of a “healthy” gut microbiome should be focused on functions necessary to fulfil all functional niches in the ecosystem (Gibbons, 2019).

It has been suggested that the inter-individual functional stability on pathway level has been overestimated partly due to methodology biases but also due to that genes that are invariable within pathways masks variation in other genes. This is potentially important for detecting altered functionality in metagenomes (Manor and Borenstein, 2017). When studying variable genes within the gut microbiota, compared to functional pathways, it has been

observed that the majority of the variable genes can be assigned to the phylum Proteobacteria (Bradley and Pollard, 2017).

1.3.5 COMMUNITY RICHNESS

Increased alpha diversity, the community richness within a sample, is considered an important marker of a healthy gut ecosystem (Bäckhed et al., 2012; Lloyd-Price et al., 2016). The ecological concept alpha diversity can describe the richness of taxa, and the evenness of the composition of taxa present in a sample, using different indices (Lozupone et al., 2012). The richness of genes in the metagenome is also a measurement found linked to more healthier phenotypes (Le Chatelier et al., 2013). In ecology, diversity is a fundamental property and are in general used as an indicator of a community function, productivity and stability (Naeem et al., 1994). The 'insurance hypothesis' implies that biodiversity maintains the functionality of the ecosystem. High diversity is suggested to be linked to a more stable gut microbiota that is more resistant to change as well as more resilience to perturbation (Lozupone et al., 2012). The community richness of the gut microbiota is suggested to be influenced by several factors. Niche availability and variation in substrates for growth are factors that would increase richness. Whereas environmental factors which limit growth, such as temporal disturbances or chronically extreme conditions, would have negative influence on community richness (Reese and Dunn, 2018). Our industrialized society has been suggested to have a strong negative impact on our gut microbiota diversity (Sonnenburg and Sonnenburg, 2019). We know that use of antibiotics has short-term effects on the richness of the gut microbiota (Dethlefsen et al., 2008; Palleja et al., 2018). However, depending on when in life antibiotics are used could also have long-term effects (Blaser, 2016).

Our sanitary improvements and the use of antibiotics have saved lives. However, extensive limitation to microbial exposure can, along with the hygiene hypothesis, affect the function and regulation of our immune system (Sonnenburg and Sonnenburg, 2019). We also have altered dietary patterns in our industrialized society, with the major alteration being less diverse and more refined diet that is depleted in fiber (Sonnenburg and Sonnenburg, 2014). Increased gut microbiota richness is observed in communities with more traditional lifestyles and a more diverse diet, rich in fiber, compared to westernized societies (Clemente et al., 2015; Schnorr et al., 2014). These differences in community richness between lifestyles are also observed in children, over 3 year of age (De Filippo et al., 2010; Yatsunenko et al., 2012). Decreased in community richness is also seen in the gut microbiota of individuals originating from southeast Asia after immigration to the United States (Vangay et al., 2018). Many of these changes has occurred in parallel

with the increase in non-communicable inflammatory and metabolic disease (Sonnenburg and Sonnenburg, 2019).

In metabolic diseases a decreased gene richness has been associated with more adiposity, insulin resistance and dyslipidemia (Le Chatelier et al., 2013). In a mixed obese and non-obese population, the number of genes in the metagenomes were counted. The distribution of number of genes was bimodally distributed and, when the population was divided into high and low gene richness, they found 46 genera significant different between individuals with high and low gene richness. Individuals with low gene richness had higher relative abundance of *Bacteroides*, *Ruminococcus torques*, *Ruminococcus gnavus* and *Campylobacter*. Whereas, individuals with high gene richness had higher relative abundance of *Faecalibacterium*, *Bifidobacterium*, *Lactobacillus* and *Methanobrevibacter*. When searching for genes that contributed to this difference Le Chatelier et al. also observed genes from opportunistic pathogens such as *Clostridium bolteae*, *Clostridium symbiosum* and *Clostridium clostridioforme* in individuals with low gene richness. They observed negative correlations between gene richness and parameters of insulin resistance and dyslipidemia but no significant correlation with BMI and weight. Gene richness was also found to have an impact on the improvement of metabolic parameters over a dietary intervention (Cotillard et al., 2013). Individuals with low gene richness had not only worse parameters relating to adiposity, adipose tissue inflammation and systemic inflammation from the start but they also had lower likelihood of normalizing these parameters at the end of the intervention.

Gut microbiota richness has also been linked to stool consistency and colonic transit time. These factors, along with diet, contribute to the nutrient availability for the microbiota during transit. Slow transit time can shift microbial metabolism from saccharolytic to more proteolytic fermentation and niche differentiation with increase richness (Falony et al., 2018). Therefore, Falony et al. emphasize the importance of viewing the fecal sample as a snap shot at the end of a dynamic system and that a high richness instead could be an indicator of gut ecosystem age, without any large perturbation, and not necessarily of a stable community in the lumen.

1.3.6 MICROBIOTA VARIATION AND DIET

The gut microbiota metabolizes dietary components that reach the distal gut unabsorbed and undigested. The potential for degradation of proteins and fats is less understood but start to be explored and understood (Koh et al., 2018). The largest family of undigested carbohydrates are glycans, consisting of carbohydrates such as resistant starch, inulin, lignin, pectin, cellulose and

fructo-oligosaccharides from the diet. The collective genomic potential from the gut microbiome can encode tens of thousands of carbohydrate active enzymes (Cantarel et al., 2012). Due the large variation in potential for these enzymes, and the resulting effects on host, the gut microbiota diet interaction has thus been extensively studied in relation to health and disease (Makki et al., 2018; Oliphant and Allen-Vercoe, 2019; Salonen et al., 2014; Sonnenburg and Bäckhed, 2016). The intake and degradation potential of fibers has been suggested to be the main driver of the compositional differences between individuals, defined as the enterotypes by Arumugam et al. in 2011. Gut enterotypes have been described as a result of long-term eating habits (Wu et al., 2011). The Prevotella enterotype was associated to diets rich in fibers while the Bacteroides enterotype was associated to diets rich in animal products. However, in large population studies diet has not been identified as a large contributing factor to variation in the human gut microbiota composition (Falony et al., 2016; Rothschild et al., 2018). In a longitudinal study 34 healthy individuals was followed over 17 days with daily samplings together with records of daily dietary intake. They observed that change in diet was associated with changes in gut microbiota. However, these diet-microbiota interactions were individual specific. They could predict the gut microbiota composition in a sample based on previous sample's composition and dietary record but failed to use the same model across individuals (Johnson et al., 2019). However, in other short-term longitudinal studies that introduce extreme changes in diets, such as complete exclusion of primary carbohydrates, show consistent alterations in the gut microbiota among individuals (David et al., 2014b; Mardinoglu et al., 2018).

An important observation from several studies is that different individuals respond differently, to both dietary interventions and probiotics, with specific changes in microbiota composition as well as host physiology (Korem et al., 2017; Kovatcheva-Datchary et al., 2015; Krumbeck et al., 2018). Because of this, the relationship between gut microbiota variation and diet have been studied to develop a personalized nutrition approach (Kolodziejczyk et al., 2019). Using the knowledge about inter-individual variation, of both gut microbiota composition and glucose response to food components, individual specific diets can be designed, which can control a person's glucose response after meals (Mendes-Soares et al., 2019; Zeevi et al., 2015).

1.4 SHORT-CHAIN FATTY ACIDS

Through anaerobic fermentation of mainly carbohydrates, the gut microbiota generates short-chain fatty acids (SCFA), which are one major group of

metabolites from microbial metabolism. Dietary polysaccharides can be constructed in diverse and complex configurations. The capacity to degrade and utilize this diversity of substrates is an important function for the gut microbiota reflected by the large number of carbohydrate-active enzymes found in the human metagenome (Bhattacharya et al., 2015). The gut microbiota efficiently degrades substrates, humans thus relay on the gut microbiota for harvesting the energy from the remaining complex carbohydrates (Singh et al., 2017). The main SCFAs are acetate, propionate and butyrate. These SCFAs are rapidly absorbed by the large intestine and are estimated to provide humans with 6-10% of the total daily energy requirement (Mcneil, 1984). Acetate is the most abundant SCFA followed by propionate and butyrate. The proportion of acetate are though increasing from the gut lumen, via portal vein and circulation. Butyrate is the main energy source for the host epithelium (Donohoe et al., 2011) and most of the butyrate is consumed by the epithelium and 75% propionate is metabolized in the liver (Cummings et al., 1987). SCFAs can act locally, or be transported to the circulation, and function as signaling molecules through interaction with receptors or regulate gene expression levels (Koh et al., 2016).

The metabolism of complex carbohydrates to SCFAs, is performed through an interplay between different species with different functional capacity. The first step, primary degradation, is the rate limiting step in which polysaccharides are degraded into monosaccharide or oligosaccharides. Although the gut microbiome have been described as a system with high functional redundancy this function has been highlighted to be performed by a few keystone species (Ze et al., 2013). After primary degradation, resulting sugars can quickly be consumed by other members of the gut microbiota, for energy generation through glycolytic pathways. From these pathways pyruvate is produced and used in different fermentation processes, where Acetyl-CoA is a central molecule (Wolfe, 2015). Through these pathways can end-products, such as acetate and lactate, be used as substrates for other bacterial species and, through cross-feeding, produce end-products such as butyrate.

In the fermentation process the fermentation products need to be removed for the process to proceed. SCFAs and alcohols are rapidly absorbed by the host but gaseous fermentation products, such as carbon dioxide and hydrogen, are mainly utilized as substrates trough cross-feeding by other members of the microbiota (Miceli et al., 2016). Three strategies of utilizing hydrogen are known in the human gut (Smith et al., 2019). Reductive acetogenesis; where acetate is produced from carbon dioxide and hydrogen through the Wood-Ljungdahl pathway. The second strategy is methanogenesis; where carbon dioxide and hydrogen are converted to methane by Archaea such as

Methanobrevibacter. Lastly, dissimilatory sulfate reduction; where hydrogen and sulfate, either from the diet or host mucin, is converted to hydrogen sulfide by sulfate reducing bacteria such as *Desulfovibrio*.

1.5 MICROBIOTA, HEALTH AND DISEASE

The number of studies investigating how the gut microbiota composition relates to different diseases have rapidly increased during the past 10 years. The gut microbiota has been associated with a large number of diseases ranging from a variety of physiological processes; metabolically, inflammatory and neurological (Lynch and Pedersen, 2016). However, it is still not established any mechanistic links of structures and functions to a healthy gut microbiota (McBurney et al., 2019).

1.5.1 MICROBIOTA AND DISEASE

The altered composition of the gut microbiota associated to different diseases has by large parts of the field been stated as dysbioitic. This term comes from the concepts dysbiosis, which is the altered state that could be corrected by targeted interventions and return to the state of eubiosis (Brussow, 2020). The use of this concept has had some criticism (Brussow, 2020; Olesen and Alm, 2016). The main criticism is that identification of dysbiosis alone is not a useful result. Majority of studies identifying a dysbiotic microbiota are cross-sectional and thus it is not established if the dysbiotic community state is the cause, consequence or a combination of physiological alterations in disease. There are also few causal hypotheses in studies identifying dysbiosis, which would be necessary if information of the microbiota would add value in diagnosis or be used in interventions.

This dysbiotic gut microbial state has mainly been described in inflammatory conditions (Byndloss and Baumler, 2018). Meta-analysis has identified that the majority of signal between disease individuals and controls are not specific for one disease but rather common to different disease (Duvallet et al., 2017). This common disease-associated gut microbiota signal is characterized by decreased abundance of obligate anaerobes, such as butyrate producers, and increased abundance of facultative anaerobes. A dysbiotic microbiota is also associated to a microbiota with reduced diversity (Kriss et al., 2018). Similar patterns are also observed in metabolic diseases (He et al., 2018a; Le Chatelier et al., 2013) and insulin resistance (Khan et al., 2014).

Identifying a state of the gut microbiota specifically associated to a certain disease is also problematic when we do not have a definition of a healthy

microbiota. Due to the large inter-individual variation in gut microbiota composition it is not possible to define one healthy community configuration. Ideas so far about what characterizes a healthy microbiota is ecological stability, which is defined as resistance to change and resilience to recover from a perturbation (Bäckhed et al., 2012).

Since the gut microbiota is a dynamic system and responds to changes in the environment, we need to continue to build on the knowledge from cross-sectional studies. Through longitudinal studies we would gain more information about these dynamics and how stability and dynamics could be altered in disease. If we aim to be able to therapeutically alter the microbiota to sustain health or treat disease we need to understand the health-associated dynamics of gut microbiota in the short-term as well as in the long-term and the following functional microbiome variation (McBurney et al., 2019).

1.5.2 METABOLIC DISEASE AND PRADER-WILLI SYNDROME

Common metabolic diseases associated with obesity are acquired diseases such as type-2-diabetes, cardiovascular disease and liver steatosis. Risk factors for these metabolic diseases are often summarized into the metabolic syndrome, where obesity is one of the risk factors. The metabolic syndrome also include dyslipidemia which imply an altered ratio of low to high-density lipoproteins and/or increased triglycerides in the circulation. High blood pressure and changes in glucose metabolism are also risk factors to the metabolic syndrome (Mendrick et al., 2018).

Prader-Willi syndrome is the most common genetic syndrome linked to development of severe obesity. The syndrome is caused by lack of expression of the paternal allele on chromosome 15, due to different genetic alterations, where deletions in the paternal allele is the most common (Butler, 2011). Individuals with Prader-Willi syndrome develop hyperphagia and often have a rapid weight gain. They are shorter and develop altered body composition characterized by altered distribution of adipose tissue, including reduced visceral fat and increased ratio of excess body fat mass to lean body mass (Goldstone et al., 2001; Lacroix et al., 2015). Although often being severely obese, individuals with Prader-Willi syndrome have a phenotype with fewer metabolic complications from obesity (Talebizadeh and Butler, 2005).

Obese children with Prader-Willi syndrome have lower fasting insulin (Lindgren et al., 1999) and higher insulin sensitivity compared to BMI-matched children (Haqq et al., 2007; Haqq et al., 2011; Schuster et al., 1996). Adult individuals with Prader-Willi syndrome have been found to have lower insulin resistance and lower fasting insulin (Talebizadeh and Butler, 2005) and

reduced insulin release in intra-venous glucose tolerance test (Schuster et al., 1996). The complete picture of the mechanisms contributing to the improved glucose metabolism in individuals with Prader-Willi syndrome are not known but the differences in body distribution of adipose tissue to more subcutaneous and less visceral fat storage, with decreased adipose tissue inflammation as well as increased levels of the insulin sensitive hormone adiponectin could contribute (Lacroix et al., 2015). There is one previous study of the gut microbiota in individuals with Prader-Willi syndrome, this study was in children (Zhang et al., 2015). In this study they found no differences in the gut microbiota between children with Prader-Willi syndrome compare to children with simple obesity. There was neither any differences in the response to a dietary intervention.

2 AIMS

The aims of this thesis are to investigate the temporal dynamics in the gut microbiota in normal populations, without clinical evidence of disease, at different stages of life. Also, to identify characteristics of the normal gut microbiota that could contribute to the gut microbiota's resistance to disease.

The specific aims of this thesis are:

1. Identify how the gut microbiota is established and developed in young children and how they compare to adults (**Paper I**).

2. Determine the normal variation of the gut microbiota within an individual and identify different patterns of variability for both gut microbiota species and functional potential (**Paper II)**.

3. Characterize features of the gut microbiota that can be linked to metabolic health (**Paper III**).

3 METHODOLOGICAL CONSIDERATIONS

This section includes general discussion about the methodology used in the papers included in my thesis. A more detailed description about the methods can be found in the Method sections for each individual paper.

3.1 STUDY THE GUT MICROBIOME USING MOLECULAR METHODS

There has been, and there still are, a debate about the vocabulary when describing an environment and its associated microbial community. The following terminology has been used within the studies of this thesis (Marchesi and Ravel, 2015). Microbiota is used when referring to the composition of different microorganisms which are present in an environment. In contrast, the microbiome is used when referring to the microorganism, their genomes and the environmental conditions they are in. The gut microbiome has been referred to as an organ (Baquero and Nombela, 2012), an ecological system that interacts with the host, and the collective genomes of the microbiome, the metagenome, is thought of as the genetic potential of this organ. A shift in the metagenome functional potential does not fully represent changes in what the metagenome actually transcribes, it reflects a shift in functions that a metagenome has the potential to transcribe.

The majority of the microorganisms in the gut are challenging to isolate and culture *in vitro*. Therefore, the development of molecular methods, and especially high-throughput sequencing and bioinformatics analyzes, has increased the number and size of microbiome studies dramatically in the past years and provided a deeper understanding of the composition of the gut microbiota. The majority of these studies in humans have characterized microbial communities through the analysis of DNA extracted from fecal samples. Today, three main culture-independent methods, using molecular biology, are commonly used for studying the microbiota. (1) The first method is the taxonomical profiling of microbial communities by sequencing the hyper-variable regions of marker genes, which for bacteria usually is the 16S ribosomal RNA (16 rRNA) gene (Hamady and Knight, 2009). (2) The second method is whole-genome metagenome sequencing of the extracted DNA, where the collective genomic DNA extracted from a fecal sample is sequenced. Whole-genome metagenome sequencing provides information not only on microbial taxonomic composition but also on the functional genetic potential of the metagenomes, and if sequencing is sufficiently deep, it may also allow recovery of individual genomes.

When studying the gut microbiome there are several layers of factors which affect a correct representation of the community of a sample. In this thesis I have only studied fecal samples and I will only discuss factors affecting microbiota composition in these samples (Figure 1).

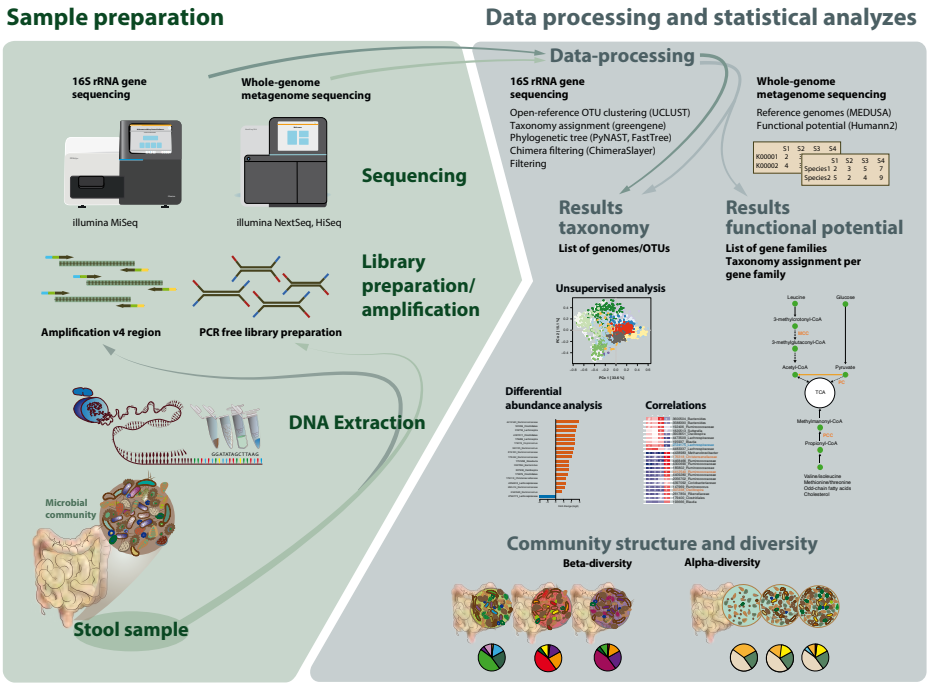

Figure 1. Overview of molecular methods used to study the gut microbiota.

3.1.1 FECAL SAMPLES

The majority of gut microbiota studies are based on the fecal microbiota composition, which practically is a representation of what is leaving the gastrointestinal system. However, the microbiota composition varies along the intestinal tract (Eckburg et al., 2005). There is also a compositional gradient with different niches in the lumen compared to close to the mucosal lining. For example, the abundance for some bacterial communities close to the mucosal lining, which grow on host-derived substrate such as mucins, are independent of dietary changes (Donaldson et al., 2016). Thus, when analyzing the microbiota composition in fecal samples we might not have a correct representation of the microbes with closest proximity to the host or in the small intestine, which need to be acknowledged when hypothesizes of the microbiota are generated.

We also study the genetic material in the fecal samples and do not know if the bacteria are viable or not. Almost all studies that have been performed so far are not quantitative, which means that the features in the data table are not independent but proportional to the unknown total amount. Suggested methods are available to generate quantitative abundance (Vandeputte et al., 2017).

3.1.2 DNA EXTRACTION

Extraction of microbial DNA from samples has been identified as the most important step in the sample processing and different extraction methods generate the largest bias in metagenomic analysis (Costea et al., 2017). It is thus important when comparing samples and cohorts that the samples have been processed with the same protocol. Due to the different cell wall properties in different bacteria, it is challenging to lyse all bacteria without damaging the DNA of more easily lysed bacteria, thus maintaining the representation of the sample. A combination of chemical and mechanical, in form of repeated bead-beating, lysis protocols have been found to generate the largest diversity in samples and was also important for recovery of bacteria from *Clostridium cluster IV* and Archaea (Salonen et al., 2010). Furthermore, bead-beating was required for detection of *Bifidobacterium* in samples obtained from infants (Walker et al., 2015). For all samples in this thesis genomic DNA has been extracted using double bead-beating protocols adapted from Salonen et al., 2010.

3.1.3 16S rRNA GENE PROFILING

16S rRNA gene profiling has been the dominating method in large studies until recent years, and is still widely used in clinical studies. However, 16S rRNA gene profiling has several limitations such as the restriction to the phylogenetic characterization of microbial communities based on small genomic regions and the presence of technical bias, as well as choice of region (Gohl et al., 2016). In **Paper I** and **Paper III** the microbiota was studied using 16S rRNA gene sequencing after amplification of the V4 region using primers previously described (Kozich et al., 2013). This region has been reported to have sufficient resolution for all phyla present in the human gut and, in contrast to V1-2 regions, can identify the genus *Bifidobacterium* from the phylum Actinobacteria, which is an important genus in the gut of infants and children (Sim et al., 2012). Using the V4 region alone, compared to using V3 and V4 combined, the shorter region makes it possible to fully cover the region twice using paired-end sequencing. By constructing a consensus sequence from the two reads, and thus reducing the influence of low sequence quality, there is a limit to the introduction of spurious OTUs (Bokulich et al., 2013; Kozich et al., 2013). In **Paper I** and **Paper III** generated sequences were merged into a

consensus sequence and clustered into OTUs (operational taxonomic units) at a 97% identify threshold using an open-reference method in UCLUST (Edgar, 2013) and the Greengenes database (DeSantis et al., 2006)(13_8 release). All sequences that failed to cluster against the Greengenes database were used to cluster OTUs de novo based on their pairwise similarities.

Traditionally, a threshold of 97% identity over the 16S rRNA marker gene sequence has been used to define a operational taxonomical unit (OTU) for taxa on species level (Schloss and Handelsman, 2005). In recent years, new methods have been developed that aim to increase the threshold of identity for species definition to 100% (Callahan et al., 2017; Edgar, 2018). Technical noise is filtered out from biological variation in the marker gene sequences, denoising, which results in identification of exact sequence variants (ESV). Similar methodology identifies amplicon sequence variants (ASV) using DADA2 (Callahan et al., 2016) and ZOTU using UNOISE (Edgar, 2016).

Both OTU and ESV methods have advantages and disadvantages and should perhaps be seen as different methods rather than one method being better than the other. In my work in **Paper I** and **Paper III** I have not compared the different methods. However, due to the lack of knowledge on how the differences in methods affect distributions of individual taxa among samples and how these new methodologies affect the possibility to detect differences between groups I could not assume that the biological interpretations from the two methods would be the same. I made the decision to use similar methodology used in previous literature, which I wanted to relate my results to, since the moment there at are no clear disadvantages of using operational taxonomical units (OTU).

Independent of method used for identification of taxa, the next critical step is the taxonomic classification. For this step, there are also different methods available but the taxonomic classification is only as good as the database that is used. Databases such as RDP, Greengenes and Silva are databases frequently used in metagenomics studies. Factors such as updates and database curation vary between the databases (Hugerth and Andersson, 2017).

3.1.4 WHOLE-GENOME METAGENOME SEQUENCING

In **Paper II**, the method for profiling of the gut microbiota was whole genome metagenomic sequencing, in which the total DNA extracted is sequenced. Compared to other sequencing applications using DNA, here quantification of sequences is also important, similar to RNA sequencing. Thus, it is important to prepare libraries that keep the representation of the DNA and do not skew the composition of the sample. Also, since many bacteria that have biological

importance often are present in low abundance, the amount of input DNA need to be sufficient to capture the full complexity of the sample. The sequencing depth, together with the amount of DNA, and PCR bias was sequencing associated factors which have been found to influence results in whole genome metagenomics studies (Jones et al., 2015). In **Paper II** the extracted DNA was prepared for sequencing using a PCR-free protocol to reduce the influence of GC-bias and duplication rate. This protocol also uses 1000 time more genomic DNA compared to other commonly used protocols in metagenomic studies, for example Illumina Nextera XT.

After quality filtering and removal of sequences originating from the human DNA or the host, several different approaches can be applied for analyzing taxonomy and functional potential. For taxonomy there are both reference based methods that use reference genomes but also reference free methods. In **Paper II** sequences were aligned to a non-redundant species catalogue using Bowtie2 (Langmead and Salzberg, 2012) as described by Karlsson et al. (Karlsson et al., 2014). In other strategies sequences are aligned to marker genes as in Metaphlan (Segata et al., 2012) or to generate mOTUs (Sunagawa et al., 2013). Using these methods, data analysis can be made on species, or in some cases, strain level. Compared to aligning to the whole genomes, strategies of aligning to marker genes are less sensitive for low abundant genomes and require more reads per sample.

In reference free approaches sequences are assembled into genomes, or parts of genomes (contigs), using several available metagenomic adapted assemblers (Vollmers et al., 2017). Assembly is followed by gene prediction and functional annotation. Using the fact that the abundance of different genomes varies between individuals, strategies have been developed for grouping or binning contigs into genomes based on, for example, correlation of gene abundance among individuals (Nielsen et al., 2014) or clustering contigs into genomes based on sequence composition and abundance (Alneberg et al., 2014).

Functional annotation can be assigned to contigs or raw sequences using a wide range of tools available and to gene annotation from databases such as KEGG, COG, PFAM or GO. More specialized databases for microbial functions are also available, such as the database of enzymes for carbohydrate degrading enzymes; CAZy (Lombard et al., 2014). Irrespective of the database used, the functional annotations of metagenomes are strongly dependent on the functions that are available in the reference databases. Although large sequencing efforts have increased the number of genes in databases, a large portion of the functions encoded in the metagenomes remains unknown and

more than 50% of microbial genes are not found in databases (Abubucker et al., 2012).

In **Paper II** functional potential and responsible taxonomy was assessed using the method Humann2 (Franzosa et al., 2018). In this method, sequences are first used to assign taxonomy using Metaphlan, where sequences are mapped to a database of marker genes of known species. For the identified species, functional potential is retrieved by mapping sequences to functionally annotated pangenomes. Unmapped sequences are aligned to protein databases, in **Paper II** to a version of Uniprot, with nonredundant protein sequences with more than 90% identity. Uniprot annotation was then collapsed to Uniprot gene families and associated to metaCyc enzymatic reactions followed by pathway reconstruction (Caspi et al., 2018). Since the aim of the analysis in **Paper II** was to describe an overview of the variability pattern of microbial functions in the metagenomes, the broadly described gene families in metaCyc was the method chosen instead of analyzing more detailed levels. This method was also chosen for the assignment of taxonomy to the functional potential. Although a low percentage of function could be assigned a taxonomy, and the taxonomy assigned has a bias towards the content in databases, this method provides an idea of which taxa are responsible for stable and variable functions.

3.2 HUMAN MICROBIOTA-ASSOCIATED MICE MODEL

In order to functionally test direct influence of the microbiota on host physiology, we used human microbiota-associated mouse models in **Paper III**. In this model, fecal material from human patients is transferred into mice models lacking bacteria (germ-free).

The mouse is a widely used experimental model for studies of metabolism and can be easily be rederived as germ-free and maintained in plastic isolators. However, mouse experimental models have limitations. Although there is similarity in physiology between mice and humans, there are also significant differences in for example the immune system, which results in differences between responses to mouse and human microbiota (Ivanov et al., 2009). Large proportion of the taxa in the human gut microbiota fail to transfer into mice (Zhang et al., 2017) and also to induce immunological responses (Ivanov et al., 2009). The microbes which do transfer do not assemble into communities which completely resemble the donor community (Staley et al., 2017). Many strains of bacteria are either found in mice or human, which makes them host specific. This limits the transfer between hosts (Ley et al., 2005). Compared to

molecular methods that detect the genomic material of both live and dead microbial cells, establishment of a microbial community into a mouse model is influenced by the viability of the microbes that are transplanted. Storage, anaerobic preparation and solutions for the preparation of slurries used in the transfer of samples, are important factors that can affect transfer success. The environmental factors, such as diet, lifestyle or physiological or genetic predisposal, which also could have caused the changes in gut microbiota in the donors, are not replicated in this type of experiment (Arrieta et al., 2016). Due to the large intra-individual variation of the human microbiota, it is important to use sufficient number of donors to be able to conclude causality and also to report the negative results. To be able to draw conclusions based on results from transfers, it is also important to perform the analysis of the microbiota on the donors and mice, to address which features in the donors' microbiota can be transferred to the mice (Walter et al., 2020).

Despite the limitations described above, transfer of human gut microbiota into germ-free mice is the best models to study the effects of donors' phenotypes on physiology. We do need to adapt the questions we aim to answer using this method, with a focus on individual bacterial effects on the host, rather than effects from ecological changes (Arrieta et al., 2016).

3.3 STATISTICAL CONSIDERATIONS

Due to large intra-individual variation, metagenomics data is in general sparse and individual bacteria have for biological reasons different distributions.

Since number of raw reads in a sample do not reflect a biological meaning, only how well that sample was quantified in the sequenced pool, data need to be normalized before statistical testing. Number of reads from an undefined amount of extracted DNA biased by PCR should not be considered quantitative but need to be related to all observed taxa and thus metagenomic data is always compositional. This implies that, although the theoretical absolute abundance is constant, the relative abundance of a taxa can be expected to change due to changes in other taxa.

Due to the non-normality of the majority of microbial taxa, hypothesis testing and correlation of individual taxa in papers in this thesis were made using non-parametric methods based on rank, such as Wilcoxon rank-test, Wilcoxon signed-rank test and Spearman's correlation.

Since community data is high dimensional compositional differences between samples (beta diversity) is in ecology traditionally described using different

dissimilarity measurements. Euclidean distance is not appropriate for compositional and sparse datasets and for this type of data other distances are used. In this thesis the Bray-Curtis dissimilarity was used and, when phylogenetic relationship between taxa was available, the Unifrac distance was used (Lozupone and Knight, 2005).

In order to simplify the complexity of the data, and unsupervised explore differences between samples, clustering methods based on the relative abundance of features rather than dissimilarity measures have been included in this thesis. This is a widely used method to understand compositional structures in the field. Almost ten years ago the adult gut microbiota configuration, due to its community stability, was considered as three discrete states. These were called enterotypes and identified using k-mer clustering (Arumugam et al., 2011). These enterotypes were differentiated by the sample relative abundance of the dominant genera *Ruminococcus*, *Bacteroides* and *Prevotella*. This approach has later been disputed and these three enterotypes of the normal gut microbiota are no longer considered discrete but instead considered influenced by gradients of genera abundances (Costea et al., 2018; Knights et al., 2014). With improved methodology using Dirichlet Multinomial Mixtures a fourth enterotype in the adult gut microbiota have been found, by using Laplace approximation of the model fit (Ding and Schloss, 2014; Holmes et al., 2012; Vieira-Silva et al., 2019). Using the same approach, I investigated the stability of the enterotypes in the adult population over the course of a year in **Paper II**. To simplify the complexity of the microbial communities, I have also used the same approach to cluster the samples from different ages, in **Paper I**, into community types to understand how children travel through several community configurations during the gut microbiota development. Since the microbiota configurations in the children are different from the enterotypes originally defined by Arumugam et al. in 2011 I have chosen to call them community types.

4 RESULTS AND DISCUSSION

4.1 GUT MICROBIOTA DYNAMICS IN CHILDREN

In **Paper I,** we investigated how the gut microbiota develops from birth, and during the first 5 years of age, in a Swedish birth cohort sampled longitudinally, with a stronger focus on the development from 12 months to 5 years, to expand our previous study (Bäckhed et al., 2015). We also investigated how a 5-year-old microbiota compares to the microbiota of the mothers and an adult normal population age 50-64 years (S3WP adults). We observed, as others before us, that as the child grow the microbiota complexity, measured as the alpha diversity, increases and the children become less heterogenous, measured as beta diversity. At 5 years the compositional differences to the S3WP adults, measured as the weighted unifrac distance, had decreased compare to the children at younger ages and the adults. However, there was still a significant difference in the composition between the children and adults. The children also had a lower alpha diversity, measured as Faith's phylogenetic diversity.

4.1.1 MICROBIAL GENERA TRAJECTORIES

To identify how microbial genera are incorporated into the developing gut microbiota, we clustered genera based on their change in abundance over time. To this end we used infants with complete sample series from 4 months to 5 years. We identified 4 different trajectories that the genera followed (Figure 2). The first trajectory included genera dominating the infant gut microbiota but decreased over time. The second trajectory, was characterized by genera with the highest abundance at 12 months. The third trajectory was dominated by genera such as *Prevotella, Akkermansia,* and several Clostridiales such as *Faecalibacterium*, *Ruminococcus*, *Blautia, Lachnospira, Roseburia* and *Coprococcus*. These genera were highly abundant in adults and the majority of them are considered part of the core microbiota in adults (Falony et al., 2016). These genera also had higher abundance at 12 months, compared to 4 months old infants, but the abundance continued to increase up to 3 years of age when the abundance stabilized. We also identify a fourth trajectory dominated by genera that had low prevalence and low abundance in infancy and at 12 months but had increased average abundance at 3 years, which continued to increase until 5 years of age. However, they still had reduced abundance at 5 years compared to an adult population.

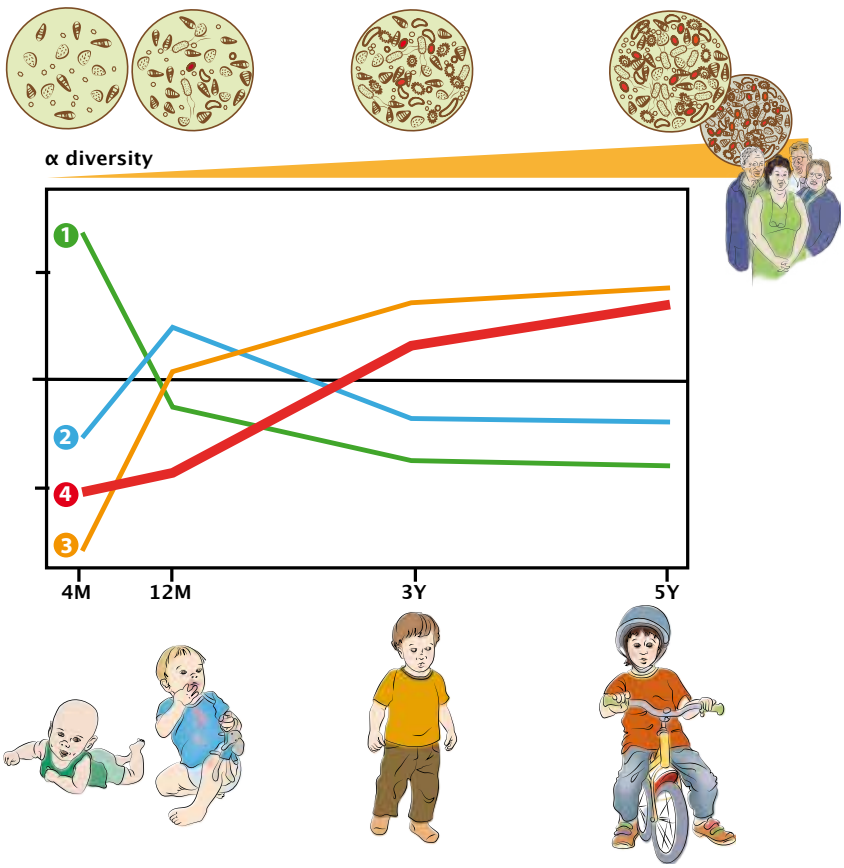

Figure 2. Trajectories which genera followed during gut microbiota development from 4 months to 5 years.

Among the late 'bloomers' of the fourth trajectory were hydrogen consumers, such as the Archaea *Methanobrevibacter* and the Proteobacteria *Desulfovibrio*, the Firmicutes family Christensenellaceae and also genera within the Coriobacteriaceae family of the Actinobacteria phylum, such as *Collinsella* and *Adlercreutzia*. As well as unspecified genera within the orders ML615J-28 and RF39 in the Tenericutes phylum.

Methanobrevibacter is the most common Archaea in the human adult (Goodrich et al., 2016) and child gut (van de Pol et al., 2017). Consistent with our findings lower abundance is observed in school children compared to

adults (Zhong et al., 2019) and with lower prevalence in children compared to adults (Stewart et al., 2006). Archaea, such *Methanobrevibacter*, as has been detected in fecal samples (Dridi et al., 2009; Stewart et al., 2006) and from gastric juice in infants and are thought to be acquired from the mother (Grine et al., 2017). However, using 16S rRNA profiling in **Paper I**, *Methanobrevibacter* have been detected at low prevalence and low abundance during infancy. The same age discriminatory observation has been observed in rats, where *Methanobrevibacter* was first observed after weaning suggesting that fermentation is a prerequisite for colonization with this taxa (Maczulak et al., 1989). Increased redox potential, lower bacterial load and depletion of *Methanobrevibacter* has been observed in fecal samples from children with severe acute malnutrition (Million et al., 2016). These children have previously been identified with an immature gut microbiota, characterized with lower abundance of obligate anaerobes dominating the adult microbiota (Subramanian et al., 2014). Since methanogenesis is a process which requires a strongly reduced environment (Hirano et al., 2013), and requires hydrogen for growth, this data indicates that establishment of obligate anaerobes and a fully reduced gut environment is required for hydrogen consumption through methanogenesis.

The species *Desulfovibrio piger* is also a hydrogen consuming bacterium, although through dissimilatory sulfate reduction. This bacterium requires a source of sulfate for sulfate reduction, and although this can be available through diet, most comes from sulfated glycans in the host mucosa (Tailford et al., 2015). This can be accessible through degradation of host mucin by species encoding sulfatases such as *Bacteroides thetaiotaomicron* (Rey et al., 2013).

In the TwinUK population Tenericutes as phylum, unclassified genera within RF39 and Christensenellaceae together with *Methanobrevibacter* were found heritable and positively associated to alpha diversity. In repeated sampling the levels of these taxa were considered stable within an individual (Goodrich et al., 2016). Christensenellaceae and *Methanobrevibacter* have been found to co-occur and to cross-feed *in vitro* (Goodrich et al., 2014; Ruaud et al., 2020).

4.1.2 INDIVIDUAL DEVELOPMENT PACE

By clustering the samples from the children, based on the abundance of individual genera, we observed that the children pass through different community types through the development of their microbiota. The samples from newborn and at 4 months classified into community types that were mostly seen in samples from these ages. The majority of samples from the older ages, 3 years and 5 years, classified into community types dominated with

samples from 3 and 5-year-old children. The large majority of the adult samples formed an adult community type with few samples from children.

The age specific community types we observed at 12 months indicated that the children go through a transitional phase. The 12 months specific community types were also observed in samples from 3- and 5-year old children, the frequency of samples in these community types though decreased over time. This transition phase has previously been described as the period between 15 and 30 months in where all five phyla change abundance and alpha diversity continues to increase (Stewart et al., 2018). The phase after 30 months has been defined as the stable phase where phyla abundance or alpha diversity do not change. In contrast to the study of Stewart et al., all children in **Paper I** had the same chronological age but at 12 months. We identified children in community types seen both in mostly 4-month-old, 12-month-old children and children 3 years and older, indicating that at this age children were at different stages within the transitional phase. This suggest that children in a normal population have an individual pace in their transition and maturation of their gut microbiota development. We identified an association between the alpha diversity in the children at 12 months and the alpha diversity in their 5-year microbiota, indicating a connection with the microbiota established at 12 months and the continuing development of the gut microbiota up to 5 years. This need to be accounted for when investigating factors which can affect gut microbiota maturation.

4.1.3 GUT MICROBIOTA IN TRANSITIONAL PHASE

The prevalence of facultative anaerobes before the introduction of solid food indicates a higher redox potential in the infant gut compared to adults (Stark and Lee, 1982). The high abundance of *Eubacterium*, *Veillonella* and *Megasphaera* at 12 months, several were among the genera following the second trajectory with peak abundance at 12 months, indicates high availability of lactate. These genera include species which have the ability to convert lactate to acetate. *Megasphaera* and *Eubacterium hallii* have also been shown to produce butyrate from lactate, whereas other highly abundant butyrate producers in the adult gut microbiota, such as *Roseburia intestinalis*, *Eubacterium rectale* and *Faecalibacterium prausnitzii* have no, or limited, lactate utilization (Duncan et al., 2004).

Bifidobacterium is one of the primary degraders in the human gut, degrading resistant starch to acetate and lactate (Macfarlane and Englyst, 1986). Two routes of cross-feeding have been suggested between degradation of starch and fructo-oligosaccharides, by *Bifidobacterium adolescentis*, to butyrate forming bacteria in the gut (Belenguer et al., 2006). The first was through cross-feeding

of lactate utilized by *E. hallii* for butyrate production and the second was through degradation of partial breakdown of fructo-oligosaccharides, which species such as *Roseburia sp.* can use for butyrate production.

In a study of children's microbiota at 9 months in Denmark the abundance of *Megasphaera*, *Veillonella* and *Haemophilus* was associated with exclusively breastfeeding the first 6 months (Laursen et al., 2016). In this study they did not see associations between how early or late solid food was introduced in complement to breastmilk. Instead the age when exclusively breastfeeding ceased had significant effects on the microbiota at 9 months of age. This reflect observations form the metagenome study of a subset of the children in the cohort in **Paper I** from their first year (Bäckhed et al., 2015) where the duration of breastfeeding affects their gut microbiota composition at 12 months. In **Paper I** we observed that the subset of children still breastfeeding at 12 months had a significant different gut microbiota with higher abundance of *Bifidobacterium*. We did however not observe any significant differences in alpha diversity in their microbiota and they are not overrepresented in any of the community types seen at 12 months. The differences observed in community types at 12 months, with higher alpha diversity in the community types characterized by higher abundance of unclassified *Ruminococcaceae*, *Faecalibacterium* and *Roseburia* and lower abundance of facultative anaerobes, have previously been linked to progression toward family food and introduction of more fiber and protein rich food (Laursen et al., 2016).

To summarize, the development of the children's gut microbiota from a low diverse gut microbiota dominating of *Bifidobacterium* and facultative anaerobes, in our study exemplified by the samples from the children at 4 months, and into an adult like gut microbiota, exemplified by the microbiota composition in mainly 3-years and 5-years samples, passes an altered state exemplified by the samples from the children at 12 months. How factors such as breast feeding and different dietary components shape the community to prepare for acquiring abundant adult core microbes is still not known. This acquisition also seems necessary for gaining low abundant species, with specific functions in the gut, which have been found linked to a high community richness and healthy phenotypes in adults.

4.2 GUT MICROBIOTA DYNAMICS IN ADULTS

In **Paper II** we investigate dynamics in the gut microbiota by longitudinal sampling during a time period of a year in a normal population (S3WP), age 50-64 years, using whole genome metagenomics. In addition to investigation of compositional variability and different types of variability pattern in the gut

microbiota we also identified factors that are associated with high and low variability.

4.2.1 COMPOSITIONAL VARIABILITY

The core microbiota can be defined in different ways, that depend on the taxonomical levels but also the percentage of individuals who present the taxa, as well as the methodology used. In **Paper II** we elaborated our results with respect to the results of previous large cross-sectional population studies as well as results of longitudinally densely sampled time-series, which generally contained small number of subjects. Overall, we observe that the core taxa identified in the majority of samples is only a fraction of all taxa present in the combined samples (Caporaso et al., 2011; David et al., 2014a). In our study of individuals sampled four times over a year we observed that the core microbiota is increased if the criteria of temporally presence was used. Adding this information implies that temporal dynamics for many taxa is part of the normal gut microbiota. This is also observed in studies with consistent seasonal fluctuations, due to seasonal dietary patterns in individuals within the Hadza hunter-gatherers in Tanzania (Smits et al., 2017). In our Swedish population there was no such strict seasonal dietary pattern and we thus did not find any seasonal variation.

In **Paper II** we could demonstrate, using whole genome metagenomics, that the gut microbiota compositional variability is an individual feature in an adult population. Some individuals, within a normal population without known diseases, had similar microbiota composition between repeated samplings, whereas other individuals were almost as different between their repeated samplings as they were to other individuals. This has previously been observed using 16S rRNA gene profiling (Flores et al., 2014). In their study Flores at al. demonstrated that increased alpha diversity was associated with increased compositional stability. Although the gut microbiota is considered stable over time in adult individuals, the intra-individual compositional variability in our study accounted for 23% of the total variation between samples. We could not link any metabolic and inflammatory markers or measurable dietary patterns to intra-individual variability. Individuals with a more stable composition had higher abundance of *Faecalibacterium prausnitzii* and several *Bifidobacterium* species as well as a trend for higher gene richness.

F. prausnitzii in one of the most abundant species in the normal gut microbiota and was the most abundant species in our cohort in **Paper II**. It produces butyrate via butyryl coenzyme A (CoA):acetate-CoA transferase primarily from acetate (Duncan et al., 2002). It is an obligate anaerobe but it has the possibility to grow under low oxygen tension, in the presence of reduced

compounds such as riboflavin, cysteine and glutathione (Khan et al., 2012). Through this property *F. prausnitzii* can possibly exploit a niche for growth in proximity of the colonic wall, where oxygen can diffuse from the circulation. The oxygen consumption in *F. prausnitzii* is accompanied with decreased butyrate production, so it appears to come with a cost.

In the validation population from the extended Human microbiome program (Lloyd-Price et al., 2017) we did not observe any species linked to compositional variability and potentially representing a stability factor in the fecal communities. However, this population is more heterogeneous in the range of age and in the time intervals of samplings. In another longitudinally sampled American population, aged of 20-40 years, the authors identified that different *Alistipes* species and *Bacteroides uniformis* were positively correlated to compositional stability (Johnson et al., 2019). We observed differences between the Swedish population and the American population with higher abundances of species within the phylum Bacteroidetes in the American population, which also has been observed when compared to other European populations (Falony et al., 2016). Potentially, the same stabilizing functions could be represented by different species in different populations.

4.2.2 VARIABILITY PATTERNS IN THE GUT MICROBIOTA

In **Paper II** we also investigated different variability pattern of microbial species in the gut. We observed, as others have seen before us, that the ranges of abundances of individual bacterial species vary between individuals (Human Microbiome Project, 2012; Turnbaugh et al., 2009). Some species had high abundance with a small range, while other species had lower abundance and small ranges. However, there were also species that could be detected at high abundances in some individuals but low abundances in others, with a bimodal distribution. We could validate these results in the extended Human microbiome project. Next, we investigated if these large spans in abundance were due to actual large differences between individuals or due to fluctuations within the individuals over time (Figure 3).

Among the species with most stable abundance within individuals, we found species with the potential to produce butyrate but also many *Clostridium* species and *Ruminococcus gnavus*. In general, high abundant species were less variable, which has previously been observed in other longitudinal studies (Mehta et al., 2018). However, we observed also low abundant species with very consistent abundance within individuals over time. Among these we observed *Clostridium scindens*, which can dehydroxylate primary bile acids into secondary bile acids in the large intestine (Kitahara et al., 2000).

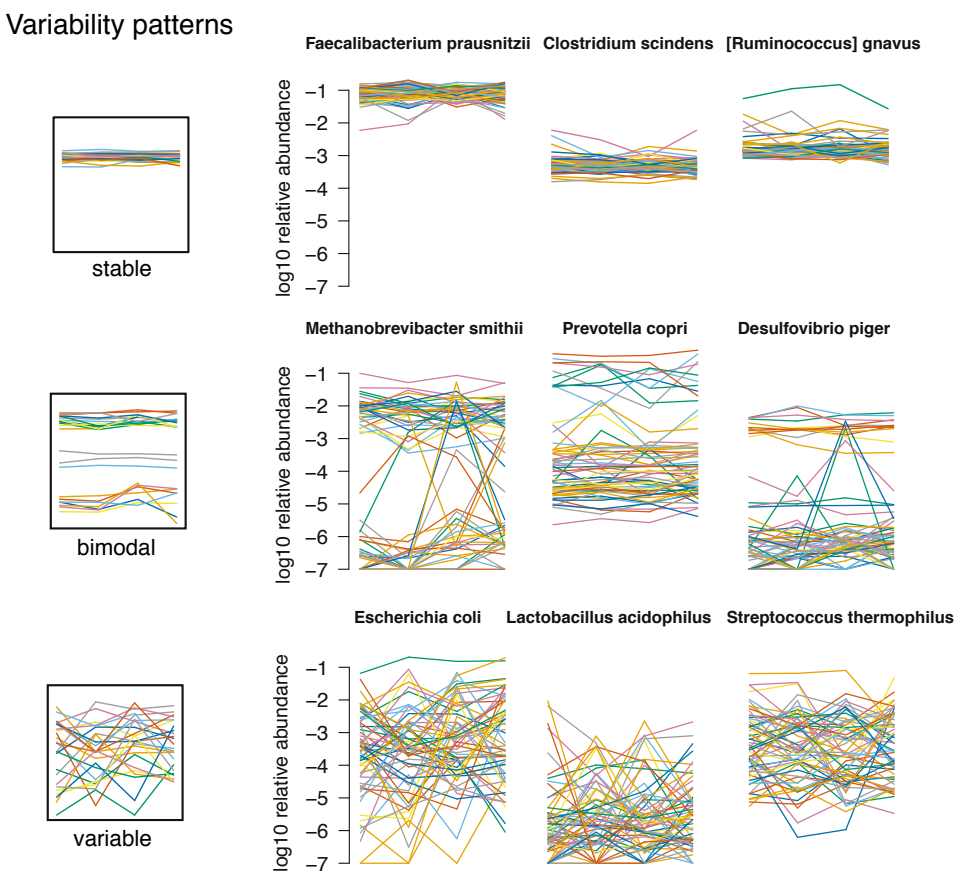

Figure 3. Examples of species with different variability patterns. Randomly subsampled from data from Paper II. Each line is one person's abundances at 4 visits.

Species such as *Prevotella copri* was identified as a stable species with most of the variation between individuals. This species was prevalent among the species containing microbial 'codes' from which individual's microbiomes could be recognized (Franzosa et al., 2015). *Methanobrevibacter smithii* was also identified as the most variable species where most of the variability was due to inter-individual variation. Species that are stable within an individual but with large differences between individuals suggest that they may be interesting biomarkers in longitudinal or intervention studies. However, their large inter-individual variation requires large observed effect sizes to be powered in cross-sectional settings.

In **Paper II** we could show that species with large variation and with the majority of variation due to fluctuation within an individual were mostly facultative anaerobes, such as species from the Proteobacteria and particularly Enterobacteriaceae family, but also lactic acid producing bacteria such as *Lactobacillus*. The abundances of Enterobacteriaceae could be measured within the whole range of detection in the Swedish normal population. This was also validated in the extended human microbiome project population. Intra-individual compositional variation was weakly associated with higher average abundance of several facultative anaerobes in this normal population. Nevertheless, we observed negative correlations between changes in *F. prausnitzii* and Enterobacteriaceae such as *Escherichia coli*, *Escherichia albertii* and *Citrobacter youngae*. We thus concluded that these blooms of facultative anaerobes such as Enterobacteriaceae are part of the normal dynamics of the gut microbiota and occurs when *F. prausnitzii* levels are reduced. This has also been seen in densely sampled time series of healthy individuals (Gibbons et al., 2017).

It is important to keep a low oxygen tension, or a low redox potential, to maintain the abundance of strict anaerobes and butyrate producing bacteria, thus maintaining normal gut microbiota functions. Host-bacteria interactions are suggested to be responsible for maintaining low redox potential and sustained butyrate production in the colon (Byndloss et al., 2018). A low redox potential is necessary for growth of obligate anaerobes and the production of fermentation products such as butyrate, which is the major nutrient source for the colonocytes. Oxygen from the circulation that is passively diffusing through the intestinal epithelium increases the redox potential in the gut lumen (Albenberg et al., 2014; Espey, 2013). Butyrate activates epithelial PPRAγ which shifts energy metabolism in the epithelium towards oxygen consuming beta-oxidation (Byndloss et al., 2017). This process reduces the influx of oxygen from the circulation to the intestinal lumen, thus maintaining the low oxygen environment that is required for sustained butyrate production. Increased oxygen, or higher redox-potential in the gut lumen is associated with a shift from obligate anaerobes to facultative anaerobes (Rivera-Chavez et al., 2017). This loss of hypoxia in the gut is suggested to contribute to the signals of increase abundance of facultative anaerobes, such as Proteobacteria, and has negative effects on host epithelial function (Litvak et al., 2017).

Since blooms in facultative anaerobes is part of a normal dynamics of the gut microbiota, changes in these dynamics in disease can only be determined through repeated samplings of individuals with and without disease. Repeated sampling is required to determine if blooms of facultative anaerobes are potentially more frequent in a disease gut microbiota.

We did not find any significant correlations between dietary intake, or change in dietary intake on macronutrient level, and compositional variability in the participants in **Paper II.** This could be due to individuality in response to diet. In a study of daily sampled individuals along with investigating dietary records the authors could predict the gut microbiota composition based on the dietary intake with one day lag (Johnson et al., 2019). However, this prediction was individual-specific and could not be applied to other individuals. This indicates that the effects of diet composition are highly dependent on the individual gut microbiota composition. In this study it was also observed that individual food items had better prediction compared to if they were summarized into micro- and macronutrients. These results suggest that components in the diet, which influences the gut microbiota, are not included in these tools or that categories, such as fiber, are too general to have resolution for the gut microbiota variability. The lack of associations found between diets and gut microbiota dynamics in **Paper II** most likely were due to lack of consistent response to diet and limitations in tools used for dietary records.

4.2.3 VARIATION IN MICROBIOME FUNCTION

In **Paper II** we observed that the total variation for functional gene families in the metagenomes was much smaller compared to variation of microbiota species, which is not surprising since functional redundancy is an important feature of the normal gut microbiota. However, we observed that the majority of variation in gene families was due to intra-individual fluctuation. This is consistent with longitudinal multi-omics studies with healthy individuals (Zhou et al., 2019).

Among the functions with low variation were functions within the 'minimal gut genome' (Qin et al., 2010). Pathways involved in butyrate production from carbohydrate metabolism (CENTFERM-PWY: pyruvate fermentation to butanoate and PWY-5676: acetyl-CoA fermentation to butanoate II) were not among the least variable but had low fluctuation within individuals over time. Whereas, pathways of butyrate production from amino acids (PWY-5022: 4-aminobutanoate degradation V and P163-PWY: L-lysine fermentation to acetate and butanoate) had a large component of within individual fluctuation. Butyrate production from amino acid is just a limited fraction of the total potential for butyrate production in the human gut but has been found increased in disease populations, such as type-2-diabetes (Vital et al., 2017).

Functions with the largest total variation also had a large component of intra-individual variation and were dominated by functions for catabolism of sugars, fermentative processes (e.g., production of lactate, acetate, propionate and butyrate), glyoxylate cycle, TCA and modified TCA cycles. Biosynthesis of

components for electron transfer chains (e.g., phylloquinol, menaquinones and demethylmenaquinones) as well as potential functions for the synthesis of amino acids (e.g., arginine, tyrosine and tryptophan) and for production of co-factors, such as vitamin K, biotin and folate were also among the functions with large intra-individual variation. We found that these functions were linked to less redundancy in the metagenomes and with larger intra-individual variation in the taxa with these functions. The main contributors to these pathways were often Gammaproteobacteria, such as *E. coli* and *Citrobacter freundii*. This suggests that intra-individual fluctuation in functional potential could reflect more intra-individual variation in facultative anaerobes such as Enterobacteriaceae.

Since a stable functional potential in the gut microbiome characterizes a healthy microbiome (Bäckhed et al., 2012; Gibbons, 2019), altered variability in functional potential could be seen in diseases. More variation in functional potential of metagenomes was observed in populations with type-2-diabetes and prediabetes compare to a normal population (Bradley and Pollard, 2017). Due to intra-individual variation repeated measurements have been recommended for accurate prediction of taxonomic and functional potential abundance (Poyet et al., 2019). Poyet et al. concludes that the variance of estimation was greatly reduced when including 5 to 9 longitudinal samples.

4.3 FEATURES OF A HEALTHY GUT MICROBIOTA

In **PaperIII** we investigated the gut microbiota of obese individuals with the genetic syndrome Prader-Willi syndrome (PWS) and compared it to the gut microbiota of fat mass matched individuals with simple obesity (OC), as well as a non-obese group (PWS-parents). Since less metabolic complications have been observed in individuals with Prader-Willi syndrome compared to individuals with simple obesity, we hypothesized that PWS may have features in the gut microbiota linked to metabolic health.

4.3.1 HEALTHY ECOLOGICAL SYSTEM

In ecology theory, a stable ecological system is characterized by a community in equilibrium that is resistant to effects from perturbations and high resilience to recover after a perturbation (Sommer et al., 2017). Complex ecosystems, such as the microbial community in the gut, are thought to exist in limited number of stable states (Beisner et al., 2003). When a perturbation is too large, or the resilience limited, the community may adopt a new compositional state (Holling, 1973). This is a proposed explanation for the gut microbiota

compositional differences seen in disease (Sommer et al., 2017). However, this explanation is also debated. Particularly, current studies are not sufficient to determine whether the shift in disease occur from a healthy stable state to a new altered stable state that characterizes the disease, or whether disease communities may originate from a variety of less stable states, stochastically determined (Zaneveld et al., 2017). For the latter scenario, the 'Anna Karenina principle' is one alternative explanation to the disease microbiota configuration. From Tolstoy's "all happy families are alike but each unhappy family is unhappy in its own way", the Anna Karenina principle applied to microbial communities states that gut microbiota composition of individuals with a disease is more heterogenous compared to the composition of controls (Diamond, 1997). This mean that the gut communities of healthy people are more similar, measured by lower beta-diversity metrics between samples, than among individuals within a disease group. The Anna Karenina effect is the result of microbial communities' responses to stress or perturbations and the communities' possibilities to regulate community composition. In ordination representation of beta diversity measurements, these effects result in that disease individuals form more heterogenous groups overlapping with healthy controls rather than forming new groups separated from healthy controls (Zaneveld et al., 2017).

Consistent with that individuals with PWS are more metabolically healthy compared to subjects with OC, we found compositional differences in the gut microbiota between the PWS group and OC but not between the PWS and PWS-parents. The heterogeneity of the samples in the OC group was also larger compared to the other two groups, measured as larger beta diversity between the samples within the OC group compared to samples within the other two groups. These results indicate that the homogeneous composition observed in the individuals with PWS was a feature of the more metabolically healthy obese state. The composition in the gut communities in the OC group was different from the PWS and non-obese group but this difference was due to the larger dispersion of OC samples rather than formation of a new group representing the composition of OC samples, which suggested a more instable composition according to the Anna Karenina principle.

Differences in composition, due to heterogeneity among disease individuals rather than complete shift in composition, has also been observed in individuals with inflammatory bowel disease (Pascal et al., 2017; Ryan et al., 2020). In the work by Halfvarson el al. they define a 'healthy plane' in the ordination representation and follow patients with different types of inflammatory bowel disease over time (Halfvarson et al., 2017). They observed that healthy individuals move within this 'healthy plane' but that patients with

inflammatory bowel disease occasionally move outside of this plane. The compositional variability outside of the healthy plane was associated to increase in Enterobacteriaceae and decrease in *Prevotella copri* and *Faecalibacterium prausnitzii* suggesting that fluctuations of abundance in these species affect the gut microbiota composition and that the variation in abundance of these species is outside of normal variation.

Similar conclusions have also been made when clustering samples using Dirichlet multinomial mixtures (Holmes et al., 2012). In this study they found that samples from obese individuals distributed differently in clusters than samples from lean controls. However, the clusters observed in lean and obese were the same. The obese individuals also formed a less homogenous group in the ordination representation compared with the lean controls. These observations could either imply that the gut microbiota is not affected by the obesity status in all obese individuals or that the gut microbiota in obese individuals is less stable. The same is also observed in individuals with inflammatory bowel disease (Vieira-Silva et al., 2019). They also identified four clusters of samples, here called enterotypes, overlapping with the three enterotypes originally defined by Arumugam et al. in 2011. Distribution in the four enterotypes was different for samples from patients with inflammatory bowel disease and a normal population. The highest frequency of samples within the new enterotype B2 was observed from samples originating from patients with Crohn´s disease. However, the enterotypes were overlapping and there was no new compositional configuration without samples from controls. This B2 'enterotype' has also, apart from inflammatory bowel disease, been associated to depression status (Valles-Colomer et al., 2019). In **Paper II** we investigated the stability of the community types in adults, often called enterotypes. We could identify three enterotypes in the Swedish normal population, missing the enterotype named B2. This is most likely due to lack of enough samples which would cluster into this configuration, due to the healthy status of the individuals in this population. In our population 45% of the individuals changed enterotype over the course of a year. This indicates that all three enterotypes could be associated to healthy gut microbiota and that enterotypes should not be used as a predictive tool. Instead it can be used as a way to understand how different gut microbiota configurations are distributed among different groups.

In a study investigating the effects of a dietary intervention in children with Prader-Willi syndrome and simple obesity the authors did not find any differences in the gut microbiota between patients with and without the syndrome (Zhang et al., 2015). The patients in Paper **III** were adults and the discrepancy in results between our results and the results in the children could

be due to age. The effects on the microbiota due to increased adiposity could either develop over time and might not be observed in younger ages. The range of ages in the Zhang et al. study is also quite large and the heterogeneity of ages could also contribute to the lack of result.

4.3.2 COMMUNITY RICHNESS

Since community richness is a consistent marker in the gut microbiota found linked to health in cross-sectional studies we also investigated taxa linked to alpha diversity, measured as Faith's phylogenetic diversity (Faith, 1992), as well as to metabolic parameters in the obese individuals in **Paper III**. We found correlations between taxa and Faith's phylogenetic diversity, fasting insulin, plasma triglycerides as well as to HOMA-IR, a measurement of insulin resistance in fasting state. We found positive correlations between increased alpha diversity, as well healthier metabolic values, and the taxa annotated as the Archaea *Methanobrevibacter*, *Oscillospira*, *Coprococcus*, taxa from the Christensenellaceae family and several taxa from the Ruminococcaceae family. Negative correlations where sparser but among them where the correlation to taxa annotated as *Ruminococcus gnavus*, which was positively correlated with increase fasting insulin and increase plasma triglycerides. Lower community richness is a consistent marker found in metabolic diseases such as the metabolic syndrome (Le Chatelier et al., 2013), glucose deregulation (Allin et al., 2018; Forslund et al., 2015) and obesity (Aron-Wisnewsky and Clement, 2014; Cotillard et al., 2013) but also in inflammatory disease such as inflammatory bowel disease (Duvallet et al., 2017).

Using fecal microbiota transplantations into germ-free mice we could verify that the richness phenotype that discriminated communities of obese individuals with Prader-Willi and simple obesity was transferred in the mice and influenced the physiology, especially on the insulin sensitivity.

Since alpha diversity is acquired over time in the developing gut microbiota we also correlated alpha diversity to taxa in the children at the different ages, as well as in the adults, in **Paper I**. At 12 months taxa with high identity to *Faecalibacterium* and *Eubacterium rectale* associated to increase alpha diversity. These are taxa, which in studies of children up to 2 years of age, contributed to an increased maturity of the gut microbiota and associated to age-adjusted richness (Blanton et al., 2016; Subramanian et al., 2014). In the children at the older ages, at 3 years and 5 years and in the adults, we found different taxa correlating to alpha diversity. Increased alpha diversity correlated to taxa with unspecific species classification from the Firmicutes phylum and the families Ruminococcaceae, Lachnospiraceae and Christensenellaceae as well as *Methanobrevibacter*. In **Paper I**

Methanobrevibacter and Christensenellaceae were among the late blooming genera, following the fourth trajectory of genera, which increasing abundance between 3 year and 5 year and with significant higher abundance in the adult community types compared to communities dominated by 3-year and 5-year-old children. Within all ages we found taxa with high identity to *Ruminococcus gnavus* negatively correlating to alpha diversity. These taxa decreased in abundance from 12 months to 5 year in **Paper I**.

Taxa within Christensenellaceae are identified as among the most heritable bacteria. The abundance of taxa within Christensenellaceae family are linked genetic traits but not vertical transmitted or inherited from parent to child. In the same study, and replicated in other populations, Christensenellaceae was also associated to lower BMI (Goodrich et al., 2014; Waters and Ley, 2019). *Methanobrevibacter* has also been identified as heritable (Goodrich et al., 2014) and in twin studies found with higher concordance in monozygotic twins (Hansen et al., 2011). In a following study in 1313 twins within the TwinUK study the authors associated higher abundance of these heritable taxa and higher alpha diversity to less visceral fat and other measurements of adiposity (Beaumont et al., 2016). Others have also seen negative association between Christensenellaceae and trunk and android fat (Hibberd et al., 2019). Christensenellaceae have also been linked to a more healthier lipid profile, lower plasma triglycerides and reduced in individuals with metabolic syndrome (Fu et al., 2015; He et al., 2018a; Hibberd et al., 2019) and to healthy glucose metabolism (Lim et al., 2017; Lippert et al., 2017). Oscillospira is an uncultured bacterial genus, belonging to the Ruminococcaceae family, which is frequently detected in the human gut microbiota using 16S rRNA gene profiling and has been linked to leanness and health (Konikoff and Gophna, 2016).

In a meta-analysis of individuals with inflammatory bowel disease including a total of 3000 individuals, in 28 studies of patients with Crohn's disease, ulcerative colitis and pseudomembranous colitis compared with controls, the most consistent signal among studies was a decrease in alpha diversity, depletion of Christensenellaceae and unclassified *Ruminococcus* and an increase in facultative anaerobes (Mancabelli et al., 2017). In a large study of individuals with inflammatory bowel disease they also observe that patients with Crohn's disease have a less stable microbiota with a reduced diversity and identify a signature of lower abundance of *Faecalibacterium*, an unknown Peptostreptococcaceae, *Anaerostipes*, *Methanobrevibacter*, an unknown Christensenellaceae and *Collinsella*, but an increased abundance of *Fusobacterium* and *Escherichia* (Pascal et al., 2017).

5 SUMMARY AND CONCLUSIONS

In **Paper I** we followed a large cohort of children longitudinally from birth to the age of 5 years. Children at 5 years of age, although displaying a more adult-like microbiota than at younger ages, still had a less diverse microbiota and altered composition compared to an adult Swedish population, suggesting that the microbiota is still developing at the age of 5.

The microbiota at 12 months represented the transitional phase of the gut microbiota from infant to an adult lite microbiota. We observed that the speed of transition through this phase was different in different children.

In **Paper I** we observed that bacterial genera followed different trajectories during the succession of the gut microbiota. We identified a group of late bloomers which increased in prevalence and abundance after the establishment of adult core microbiota species. Several of these we found associated with increased richness in children and in adults, and still at 5 years significantly lower in abundance compared to adults. We also found these associated do reduced insulin resistance, healthier lipid profiles and microbiota richness in obese individuals in **Paper III.**

In **Paper II** we investigated the dynamics of the gut microbiota in a normal adult population and identified species following different variability patterns. We found that large fluctuations in facultative anaerobes such as the Enterobacteriaceae family, both on functional and taxonomical level, were part of the normal gut microbiota dynamics. This expansion may be buffering increased oxygen influx in the gut to maintain the ecosystem. This observation also indicates that aerobe respiration and a dysbiosis-associated microbiota based on the expansion of Proteobacteria, (Litvak et al., 2017), might be misleading, as large fluctuations in Proteobacteria normally occur also in the absence of diseases.

However, the environmental conditions do not only result in taxonomic alterations they can also cause changes in the metabolism of species in the microbiota (Yoo and Byndloss, 2020). Kitamoto et al. suggests that in an inflammatory environment the metabolic pathways for *Escherichia coli LF82* is reprogramed to adapt to these conditions which results in a shift in metabolism to a catabolize L-serine for maximizing growth. However, L-serine catabolism had very limiting effect on the fitness in the healthy gut (Kitamoto et al., 2020). This show that it is not only important to understand

which bacteria are there but also how different environmental states affect their metabolism.

Facultative anaerobes in the microbial community of infants and adults might have different roles. In a model of late-onset sepsis in mice *Lactobacillus sp.* and *Escherichia coli* were required for development of the intestinal microbiota to prevent expansion of species causing late-onset sepsis (Singer et al., 2019). The authors also identified that decreased oxygen levels in the gut is a major driver of beneficial colonization dynamics.

However, similarities in gut microbiota observed in healthy children and adults with disease could reflect similarities in gut environment. A disease associated microbiota in cross-sectional settings share similarities to a low-diverse infant microbiota (younger than 2 years) and literature suggest a re-succession after perturbations which are similar to the succession in children (Kriss et al., 2018). A secondary succession, with bloom of facultative anaerobes, could help restore abundance of obligate anaerobes and diversity. Facultative anaerobes are also identified as the first responder in recovery of an antibiotic treatment in adults (David et al., 2015; Dethlefsen et al., 2008; Jakobsson et al., 2010; Palleja et al., 2018). Shortly after antibiotic treatment, an increase in redox potential was observed in the gut in mice followed by a bloom of Enterobacteriaceae before returning to normal redox potential (Reese et al., 2018). Therefore, these results suggest that the dynamics of Enterobacteriaceae is important for maintaining redox homeostasis in the adult gut.

We also concluded that repeated sampling is required to account for temporal variation, when searching for functional relationships between the gut microbiota and disease. To correctly capture all dynamics, both stable and variable features, analysis of 5-9 samples spaced out 3-5 days apart is necessary in order to be able to correctly estimate the abundance of any given microbial taxa in a fecal sample (Poyet et al., 2019).

In **Paper III** we found a more heterogenous gut microbiota composition in obese controls compared to obese individuals with Prader-Willi syndrome, with less insulin resistance and a healthier lipid profile, which, according to the Anna Karenina principle, indicate a more variable gut microbiota composition. Among the differently abundant taxa, between obese controls and individuals with Prader-Willi syndrome, in **Paper III**, we observed the majority decreased in obese controls. Similarly, we observed a sparse number of negative correlations to richness. This indicates that loss of microbial taxa constituted a consistent signal in the obese controls. However, other non-consistent variation must also be present which can explain the larger

compositional variation we observed in the obese controls. During the blooms of conditional rare taxa (CRT) in healthy individuals Gibbons et al. also observed a temporal decrease in alpha diversity (Gibbons et al., 2017). This raises the question if the average lower alpha diversity observed in cross-sectional studies of health and disease is due to an altered dynamic, with more prevalent blooms of CRT? Alpha diversity explained a large part of the compositional variation in **Paper III** but we did not observe any negative correlations between alpha diversity and facultative anaerobes. However, this does not necessarily mean that there were no associations between alpha diversity and blooms of CRT. An alternative explanation could be that the taxa blooming is not consistent in the obese controls but different in different individuals. The Anna Karenina principle states that heterogenous disease microbiota is due to stochastic alterations and these differences are thus not detected in correlation analysis.

However, *R. gnavus* was one of the few taxa observed as negatively correlated to alpha diversity in **Paper III**. We also saw a clear relationship with *R. gnavus* and alpha diversity in **Paper I**, both between different ages and within the ages in the children and within the adults. This species is also seen frequently associated increased in inflammatory bowel disease (Png et al., 2010) and seen bloom in abundance corresponding to disease periods (Hall et al., 2017). Hall et al. also characterized strains of *R. gnavus* in patients and found them more resistant to oxidative stress. Functional potential for handling oxidative stress was also seen associated to low gene richness (Le Chatelier et al., 2013). Compared to the high intra-individual variation of facultative anaerobes observed in **Paper II** we identified *R. gnavus* as a stable feature in the gut microbiota, which could explain why this is a more robust marker of a low diverse community in a potentially less reduced environment.

6 FUTURE PERSPECTIVES

In this thesis I have studied the dynamics of the normal gut microbiota in the view of what we know today is different in a disease gut microbiota. In the future we need to understand what differentiates disease dynamics from healthy and which functional changes characterize a disease-associated gut ecology. After this we can start to detangle if these changes are the cause or consequence of disease.

To validate if there is more heterogeneity in a disease gut microbiota, compare to a non-disease due to less stable gut microbiota composition, repeated sampling in disease or pre-disease cohorts need to be performed to detect changes in dynamics. This is seen in the context of inflammatory bowel disease (Halfvarson et al., 2017) but could potentially be true in other disease, such as type-2-diabetes and other metabolic diseases.

We observe, in line with a large amount of literature, a correlation between high gut microbiota richness and absence of disease. We also find Bacteria and Archaea associated with a high richness microbiota but we do not know which conditions are needed for these to grow in the community. We also do not know if they are just associated to a microbiota in a certain state or if or how they also contribute to this healthy state and to positive effects on the host physiology.

For more reliable results in microbiota studies, towards the functional understanding of the microbiota in relation in disease and potential clinical and therapeutically implementations, we need to continue study the gut microbiota as an ecological system with a focus on stable markers.

The microbial community changes along with the environment it lives in. By studying similarities in the gut microbiota during succession, where there is a decrease in redox-potential over time, and the gut microbiota in disease we could learn about environmental conditions that are key to gut microbiota stability and resilience. With in-vitro studies in gut simulators, we might be able to investigate how gut communities are affected by small changes in redox potential and identify key microbial homeostatic mechanisms. Important questions to answer will be whether the gut microbiota itself has the capacity to return to a reduced state, and which factors affect this process.

7 ETHICAL CONSIDERATIONS

The Swedish research council guide to good research practice (2017) covers all aspect of how research should be guided and performed according to different ethical codes, guidelines and legal regulations. A central part in research ethics is how research participants are handled and one of the most central ethical codes are that no research should cause harm. A second central aim is that research should be useful and transparent. In 2016 the European union decided that the member states should implement open science for publicly funded research, which include open access of both publications and research data, following the FAIR-principles (2018). According to the FAIR-principles the research data should be Findable, Accessible, Interoperable and Reusable. This to ensure optimal research integrity, open access of research publication and reuse of research data generated.

In this thesis research data, in form of metagenomic data, has been generated from several human studies, all with approved ethical applications including informed consent from participants or legal guardian (ethical applications can be found in each individual paper). This data, will be submitted to public databases for nucleotide sequencing information (European Nucleotide Archive, ENA). This is important for the continued development of knowledge about the gut microbiota and disease and also provide another scientists opportunity to validate our analysis. This is especially important for normal populations to be used as a reference for how the gut microbiota are distributed in disease population. The datasets in this thesis, include cohorts with repeated samples, which are rare among published datasets of the gut microbiota. However, due to the data protection regulation (EU GDPR) 2016/679 only limiting amount of information together with the data will be available open access. This will limit the reusability of the data since the value of the data, in terms of for example the possibility to draw conclusion, adjust for co-varying factors and find associations, increase with the amount of additional data.

According to General Data Protection Regulation (EU GDPR) 2016/679 for research on sensitive personal data require voluntarily informed consent, approved ethical application and that the research have fulfil general interest. Among the definitions of sensitive personal data are information about health, genetic data and biometric data which can be used to uniquely identify a person. This type of data is very valuable additional data to metagenomic data. The question is how valuable the open accessible data is without this data which due to the GDPR cannot be open access? To fulfill the larger goal of

open research accessibility to sensitive personal data but due to regulations of personal integrity (GDPR) is limited.

Metagenomic data, even whole genome sequencing from all DNA extracted from samples as in **Paper II**, is today not classified as sensitive personal data. While other types of human genetic material have this classification. This discrepancy is for me a bit unclear.

In article 4(13) GDPR 'genetic data' is defined as "personal data relating to the inherited or acquired genetic characteristics of a natural person which give unique information about the physiology or the health of that natural person and which result, in particular, from an analysis of a biological sample from the natural person in question". 'genetic data' contain unique information about the individual which can differentiate them from other individuals.

In fecal samples the amount of DNA from the human genome is quite low, around 1%, but in individual samples this can vary up to 20%. From data deposited to public databases researches have been able to access enough host genomic material to cover 5-20x of the host genome (Blekhman et al., 2015). In this study the metagenome data is from other body sites, containing larger amounts of host DNA in the samples, and not fecal sample. This however show that metagenomic samples may fulfill the requirement as 'genetic data'. For this reason, the data deposited for **Paper II** have host data been filtered out.

ACKNOWLEDGEMENT

I would like to send my biggest gratitude's to

Fredrik, my main supervisor. I am enormously grateful that you included me in this environment and that you believed in me. I came to you more than 5 years ago asking for challenges. And I got that for sure! Even if I have struggled I have always felt your support. Being in this environment with these impressive, supportive and loving people have been fantastic.

Valentina, my co-supervisor, for all the time, support and patience when there was no time. I hope I have not caused you too many gray hairs. Also, for being a good friend and always caring.

Staffan, also co-supervisor, for taking the time and taking things from the beginning and never letting me forget that the devil is in the details.

Robert, for being my in-official co-supervisor, for discussions, explanations and support. And off course, for always supplying us with fantastic coffee.

Ankan, **Lojsbojs** och **Lollis**, for all the fun, laughs and parties. But also, for all the coffees, support and encouragements.

Everyone in Bäckhed group, past and present members. **Lill-Annika**, **Gunnel**, **Filipe**, **Tulika**, **Matte B** and **Petia**. **Björn** for the best writing advises and drink recipes. **Ara**, your determination is so inspiring. My bioinformatics colleagues **Tony**, **Harald**, **Meenakhsi** and **Chinmay**, I am so grateful to have you to learn new things with. **Antonio**, for being a doctor you are quite alright <3. **Anna**, for all the fantastic artwork. **Manuela** and **Robert J** for being so patient with me in the lab and providing me with pens :P. **Tanweer**, for always teaching me new things. **Thomas**, without you and me the lab would be too quiet. **Kim**, **Kassem**, **Christina**, **Marc**, **Estelle**, **Alba**, **Matthias** and **Annika L** you are not only great colleagues you are also awesome friends!

All the wonderful people at wlab. My always fabulous officemate **Lilliana**. **Rosie**, for support when times are tough and always good advice. My fellow phd-students **Ismena**, **Ying**, **Marta**, **Tony**, **Angela** and **Matias,** we did it! **Martin**, for taking the time and not always running away. **Magnus**, having you around is like having sunshine inside.

Kaisa and **Fredrik Boulund**, making science with friends is the best way of making science, especially with you. I so hope that we can continue in the future.

Malin, **Anna**, **Hannah**, **Emma**, **Mia** and **Fia**. A girl cannot have better friends that the six of you. Thanks for always listening and caring.

Mats, thanks for fantastic collaboration.

My family, **Mamma** och **Pappa**, **Magnus** och **Joel**. Thank you for the stability of always having you close and always having what I need.

Elias my joy and pride, my absolute biggest achievement. Always remember that when life is tough you know you are alive, regardless if it is when you are conquering a mountain or writing a thesis.

REFERENCES

(2017). God forskningssed [Elektronisk resurs] (Stockholm: Vetenskapsrådet).

(2018). Kriterier för FAIR forskningsdata [Elektronisk resurs] (Stockholm: Vetenskapsrådet).

Abubucker, S., Segata, N., Goll, J., Schubert, A.M., Izard, J., Cantarel, B.L., Rodriguez-Mueller, B., Zucker, J., Thiagarajan, M., Henrissat, B.*, et al.* (2012). Metabolic reconstruction for metagenomic data and its application to the human microbiome. PLoS computational biology *8*, e1002358.

Albenberg, L., Esipova, T.V., Judge, C.P., Bittinger, K., Chen, J., Laughlin, A., Grunberg, S., Baldassano, R.N., Lewis, J.D., Li, H.*, et al.* (2014). Correlation between intraluminal oxygen gradient and radial partitioning of intestinal microbiota. Gastroenterology *147*, 1055-1063 e1058.

Allin, K.H., Tremaroli, V., Caesar, R., Jensen, B.A.H., Damgaard, M.T.F., Bahl, M.I., Licht, T.R., Hansen, T.H., Nielsen, T., Dantoft, T.M.*, et al.* (2018). Aberrant intestinal microbiota in individuals with prediabetes. Diabetologia *61*, 810-820.

Alneberg, J., Bjarnason, B.S., de Bruijn, I., Schirmer, M., Quick, J., Ijaz, U.Z., Lahti, L., Loman, N.J., Andersson, A.F., and Quince, C. (2014). Binning metagenomic contigs by coverage and composition. Nature methods *11*, 1144-1146.

Aron-Wisnewsky, J., and Clement, K. (2014). The effects of gastrointestinal surgery on gut microbiota: potential contribution to improved insulin sensitivity. Curr Atheroscler Rep *16*, 454.

Arrieta, M.C., Walter, J., and Finlay, B.B. (2016). Human Microbiota-Associated Mice: A Model with Challenges. Cell Host Microbe *19*, 575-578.

Arumugam, M., Raes, J., Pelletier, E., Le Paslier, D., Yamada, T., Mende, D.R., Fernandes, G.R., Tap, J., Bruls, T., Batto, J.M.*, et al.* (2011). Enterotypes of the human gut microbiome. Nature *473*, 174-180.

Bäckhed, F., Ding, H., Wang, T., Hooper, L.V., Koh, G.Y., Nagy, A., Semenkovich, C.F., and Gordon, J.I. (2004). The gut microbiota as an environmental factor that regulates fat storage. P Natl Acad Sci USA *101*, 15718-15723.

Bäckhed, F., Fraser, C.M., Ringel, Y., Sanders, M.E., Sartor, R.B., Sherman, P.M., Versalovic, J., Young, V., and Finlay, B.B. (2012). Defining a healthy human gut microbiome: current concepts, future directions, and clinical applications. Cell Host Microbe *12*, 611-622.

Bäckhed, F., Manchester, J.K., Semenkovich, C.F., and Gordon, J.I. (2007). Mechanisms underlying the resistance to diet-induced obesity in germ-free mice. P Natl Acad Sci USA *104*, 979-984.

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

Baquero, F., and Nombela, C. (2012). The microbiome as a human organ. Clin Microbiol Infect *18 Suppl 4*, 2-4.

Baumann-Dudenhoeffer, A.M., D'Souza, A.W., Tarr, P.I., Warner, B.B., and Dantas, G. (2018). Infant diet and maternal gestational weight gain predict early metabolic maturation of gut microbiomes. Nat Med *24*, 1822-1829.

Beaumont, M., Goodrich, J.K., Jackson, M.A., Yet, I., Davenport, E.R., Vieira-Silva, S., Debelius, J., Pallister, T., Mangino, M., Raes, J., *et al.* (2016). Heritable components of the human fecal microbiome are associated with visceral fat. Genome Biol *17*, 189.

Beisner, B.E., Haydon, D.T., and Cuddington, K. (2003). Alternative stable states in ecology. Front Ecol Environ *1*, 376-382.

Belenguer, A., Duncan, S.H., Calder, A.G., Holtrop, G., Louis, P., Lobley, G.E., and Flint, H.J. (2006). Two routes of metabolic cross-feeding between Bifidobacterium adolescentis and butyrate-producing anaerobes from the human gut. Appl Environ Microb *72*, 3593-3599.

Belkaid, Y., and Hand, T.W. (2014). Role of the microbiota in immunity and inflammation. Cell *157*, 121-141.

Bergstrom, A., Skov, T.H., Bahl, M.I., Roager, H.M., Christensen, L.B., Ejlerskov, K.T., Molgaard, C., Michaelsen, K.F., and Licht, T.R. (2014). Establishment of intestinal microbiota during early life: a longitudinal, explorative study of a large cohort of Danish infants. Appl Environ Microbiol *80*, 2889-2900.

Bhattacharya, T., Ghosh, T.S., and Mande, S.S. (2015). Global Profiling of Carbohydrate Active Enzymes in Human Gut Microbiome. PloS one *10*.

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

Blaser, M.J. (2016). Antibiotic use and its consequences for the normal microbiome. Science *352*, 544-545.

Blekhman, R., Goodrich, J.K., Huang, K., Sun, Q., Bukowski, R., Bell, J.T., Spector, T.D., Keinan, A., Ley, R.E., Gevers, D., *et al.* (2015). Host genetic variation impacts microbiome composition across human body sites. Genome Biology *16*.

Bokulich, N.A., Chung, J., Battaglia, T., Henderson, N., Jay, M., Li, H., A, D.L., Wu, F., Perez-Perez, G.I., Chen, Y., *et al.* (2016). Antibiotics, birth mode, and diet shape microbiome maturation during early life. Sci Transl Med *8*, 343ra382.

Bokulich, N.A., Subramanian, S., Faith, J.J., Gevers, D., Gordon, J.I., Knight, R., Mills, D.A., and Caporaso, J.G. (2013). Quality-filtering vastly improves diversity estimates from Illumina amplicon sequencing. Nature methods *10*, 57-59.

Bradley, P.H., and Pollard, K.S. (2017). Proteobacteria explain significant functional variability in the human gut microbiome. Microbiome *5*, 36.

Brussow, H. (2020). Problems with the concept of gut microbiota dysbiosis. Microb Biotechnol *13*, 423-434.

Butler, M.G. (2011). Prader-Willi Syndrome: Obesity due to Genomic Imprinting. Curr Genomics *12*, 204-215.

Byndloss, M.X., and Baumler, A.J. (2018). The germ-organ theory of non-communicable diseases. Nature reviews Microbiology *16*, 103-110.

Byndloss, M.X., Olsan, E.E., Rivera-Chavez, F., Tiffany, C.R., Cevallos, S.A., Lokken, K.L., Torres, T.P., Byndloss, A.J., Faber, F., Gao, Y.*, et al.* (2017). Microbiota-activated PPAR-gamma signaling inhibits dysbiotic Enterobacteriaceae expansion. Science *357*, 570-575.

Byndloss, M.X., Pernitzsch, S.R., and Baumler, A.J. (2018). Healthy hosts rule within: ecological forces shaping the gut microbiota. Mucosal Immunol *11*, 1299-1305.

Caesar, R., Reigstad, C.S., Bäckhed, H.K., Reinhardt, C., Ketonen, M., Lunden, G.O., Cani, P.D., and Bäckhed, F. (2012). Gut-derived lipopolysaccharide augments adipose macrophage accumulation but is not essential for impaired glucose or insulin tolerance in mice. Gut *61*, 1701-1707.

Callahan, B.J., McMurdie, P.J., and Holmes, S.P. (2017). Exact sequence variants should replace operational taxonomic units in marker-gene data analysis. ISME J *11*, 2639-2643.

Callahan, B.J., McMurdie, P.J., Rosen, M.J., Han, A.W., Johnson, A.J., and Holmes, S.P. (2016). DADA2: High-resolution sample inference from Illumina amplicon data. Nature methods *13*, 581-583.

Cani, P.D., Amar, J., Iglesias, M.A., Poggi, M., Knauf, C., Bastelica, D., Neyrinck, A.M., Fava, F., Tuohy, K.M., Chabo, C.*, et al.* (2007). Metabolic endotoxemia initiates obesity and insulin resistance. Diabetes *56*, 1761-1772.

Cantarel, B.L., Lombard, V., and Henrissat, B. (2012). Complex Carbohydrate Utilization by the Healthy Human Microbiome. PloS one *7*.

Caporaso, J.G., Lauber, C.L., Costello, E.K., Berg-Lyons, D., Gonzalez, A., Stombaugh, J., Knights, D., Gajer, P., Ravel, J., Fierer, N.*, et al.* (2011). Moving pictures of the human microbiome. Genome Biol *12*, R50.

Caspi, R., Billington, R., Fulcher, C.A., Keseler, I.M., Kothari, A., Krummenacker, M., Latendresse, M., Midford, P.E., Ong, Q., Ong, W.K.*, et al.* (2018). The MetaCyc database of metabolic pathways and enzymes. Nucleic acids research *46*, D633-D639.

Chu, D.M., Ma, J., Prince, A.L., Antony, K.M., Seferovic, M.D., and Aagaard, K.M. (2017). Maturation of the infant microbiome community

structure and function across multiple body sites and in relation to mode of delivery. Nat Med *23*, 314-326.

Clemente, J.C., Pehrsson, E.C., Blaser, M.J., Sandhu, K., Gao, Z., Wang, B., Magris, M., Hidalgo, G., Contreras, M., Noya-Alarcon, O.*, et al.* (2015). The microbiome of uncontacted Amerindians. Sci Adv *1*.

Costea, P.I., Hildebrand, F., Arumugam, M., Bäckhed, F., Blaser, M.J., Bushman, F.D., de Vos, W.M., Ehrlich, S.D., Fraser, C.M., Hattori, M.*, et al.* (2018). Enterotypes in the landscape of gut microbial community composition. Nat Microbiol *3*, 8-16.

Costea, P.I., Zeller, G., Sunagawa, S., Pelletier, E., Alberti, A., Levenez, F., Tramontano, M., Driessen, M., Hercog, R., Jung, F.E.*, et al.* (2017). Towards standards for human fecal sample processing in metagenomic studies. Nat Biotechnol *35*, 1069-1076.

Costello, E.K., Lauber, C.L., Hamady, M., Fierer, N., Gordon, J.I., and Knight, R. (2009). Bacterial community variation in human body habitats across space and time. Science *326*, 1694-1697.

Cotillard, A., Kennedy, S.P., Kong, L.C., Prifti, E., Pons, N., Le Chatelier, E., Almeida, M., Quinquis, B., Levenez, F., Galleron, N.*, et al.* (2013). Dietary intervention impact on gut microbial gene richness. Nature *500*, 585-588.

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

David, L.A., Materna, A.C., Friedman, J., Campos-Baptista, M.I., Blackburn, M.C., Perrotta, A., Erdman, S.E., and Alm, E.J. (2014a). Host lifestyle affects human microbiota on daily timescales. Genome Biol *15*, R89.

David, L.A., Maurice, C.F., Carmody, R.N., Gootenberg, D.B., Button, J.E., Wolfe, B.E., Ling, A.V., Devlin, A.S., Varma, Y., Fischbach, M.A.*, et al.* (2014b). Diet rapidly and reproducibly alters the human gut microbiome. Nature *505*, 559-563.

David, L.A., Weil, A., Ryan, E.T., Calderwood, S.B., Harris, J.B., Chowdhury, F., Begum, Y., Qadri, F., LaRocque, R.C., and Turnbaugh, P.J. (2015). Gut microbial succession follows acute secretory diarrhea in humans. mBio *6*, e00381-00315.

De Filippo, C., Cavalieri, D., Di Paola, M., Ramazzotti, M., Poullet, J.B., Massart, S., Collini, S., Pieraccini, G., and Lionetti, P. (2010). Impact of diet in shaping gut microbiota revealed by a comparative study in children from Europe and rural Africa. Proc Natl Acad Sci U S A *107*, 14691-14696.

de Goffau, M.C., Lager, S., Sovio, U., Gaccioli, F., Cook, E., Peacock, S.J., Parkhill, J., Charnock-Jones, D.S., and Smith, G.C.S. (2019). Human placenta has no microbiome but can contain potential pathogens. Nature *572*, 329-334.

Derrien, M., Alvarez, A.-S., and de Vos, W.M. (2019). The Gut Microbiota in the First Decade of Life. Trends in microbiology.

DeSantis, T.Z., Hugenholtz, P., Larsen, N., Rojas, M., Brodie, E.L., Keller, K., Huber, T., Dalevi, D., Hu, P., and Andersen, G.L. (2006). Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. Appl Environ Microbiol *72*, 5069-5072.

Deschasaux, M., Bouter, K.E., Prodan, A., Levin, E., Groen, A.K., Herrema, H., Tremaroli, V., Bakker, G.J., Attaye, I., Pinto-Sietsma, S.J., *et al.* (2018). Depicting the composition of gut microbiota in a population with varied ethnic origins but shared geography. Nat Med *24*, 1526-1531.

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.

Diamond, J. (1997). 'Guns, germs, and steel'. New York Rev Books *44*, 69-69.

Ding, T., and Schloss, P.D. (2014). Dynamics and associations of microbial community types across the human body. Nature *509*, 357-360.

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*, 11971-11975.

Dominguez-Bello, M.G., De Jesus-Laboy, K.M., Shen, N., Cox, L.M., Amir, A., Gonzalez, A., Bokulich, N.A., Song, S.J., Hoashi, M., Rivera-Vinas, J.I., *et al.* (2016). Partial restoration of the microbiota of cesarean-born infants via vaginal microbial transfer. Nat Med *22*, 250-253.

Donaldson, G.P., Lee, S.M., and Mazmanian, S.K. (2016). Gut biogeography of the bacterial microbiota. Nature reviews Microbiology *14*, 20-32.

Donohoe, D.R., Garge, N., Zhang, X., Sun, W., O'Connell, T.M., Bunger, M.K., and Bultman, S.J. (2011). The microbiome and butyrate regulate energy metabolism and autophagy in the mammalian colon. Cell metabolism *13*, 517-526.

Dridi, B., Henry, M., El Khechine, A., Raoult, D., and Drancourt, M. (2009). High Prevalence of Methanobrevibacter smithii and Methanosphaera stadtmanae Detected in the Human Gut Using an Improved DNA Detection Protocol. PloS one *4*.

Duncan, S.H., Hold, G.L., Harmsen, H.J., Stewart, C.S., and Flint, H.J. (2002). Growth requirements and fermentation products of Fusobacterium prausnitzii, and a proposal to reclassify it as Faecalibacterium prausnitzii gen. nov., comb. nov. Int J Syst Evol Microbiol *52*, 2141-2146.

Duncan, S.H., Louis, P., and Flint, H.J. (2004). Lactate-utilizing bacteria, isolated from human feces, that produce butyrate as a major fermentation product. Appl Environ Microbiol *70*, 5810-5817.

Dunn, G.A., and Jones, G.E. (2004). Cell motility under the microscope: Vorsprung durch Technik. Nat Rev Mol Cell Biol *5*, 667-672.

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. Nature communications *8*, 1784.

Eckburg, P.B., Bik, E.M., Bernstein, C.N., Purdom, E., Dethlefsen, L., Sargent, M., Gill, S.R., Nelson, K.E., and Relman, D.A. (2005). Diversity of the human intestinal microbial flora. Science *308*, 1635-1638.

Edgar, R.C. (2013). UPARSE: highly accurate OTU sequences from microbial amplicon reads. Nature methods *10*, 996-998.

Edgar, R.C. (2016). UNOISE2: improved error-correction for Illumina 16 S and ITS amplicon sequencing. bioRxiv.

Edgar, R.C. (2018). Updating the 97% identity threshold for 16S ribosomal RNA OTUs. Bioinformatics *34*, 2371-2375.

Eggesbo, M., Moen, B., Peddada, S., Baird, D., Rugtveit, J., Midtvedt, T., Bushel, P.R., Sekelja, M., and Rudi, K. (2011). Development of gut microbiota in infants not exposed to medical interventions. APMIS *119*, 17-35.

Espey, M.G. (2013). Role of oxygen gradients in shaping redox relationships between the human intestine and its microbiota. Free Radic Biol Med *55*, 130-140.

Faith, D. (1992). Conservation evaluation and phylogenetic diversity. Biological Conservation *61*, 1-10.

Faith, J.J., Guruge, J.L., Charbonneau, M., Subramanian, S., Seedorf, H., Goodman, A.L., Clemente, J.C., Knight, R., Heath, A.C., Leibel, R.L.*, et al.* (2013). The long-term stability of the human gut microbiota. Science *341*, 1237439.

Falony, G., Joossens, M., Vieira-Silva, S., Wang, J., Darzi, Y., Faust, K., Kurilshikov, A., Bonder, M.J., Valles-Colomer, M., Vandeputte, D.*, et al.* (2016). Population-level analysis of gut microbiome variation. Science *352*, 560-564.

Falony, G., Vieira-Silva, S., and Raes, J. (2018). Richness and ecosystem development across faecal snapshots of the gut microbiota. Nat Microbiol *3*, 526-528.

Ferretti, P., Pasolli, E., Tett, A., Asnicar, F., Gorfer, V., Fedi, S., Armanini, F., Truong, D.T., Manara, S., Zolfo, M.*, et al.* (2018). Mother-to-Infant Microbial Transmission from Different Body Sites Shapes the Developing Infant Gut Microbiome. Cell Host Microbe *24*, 133-145 e135.

Flores, G.E., Caporaso, J.G., Henley, J.B., Rideout, J.R., Domogala, D., Chase, J., Leff, J.W., Vazquez-Baeza, Y., Gonzalez, A., Knight, R.*, et al.* (2014). Temporal variability is a personalized feature of the human microbiome. Genome Biol *15*, 531.

Forslund, K., Hildebrand, F., Nielsen, T., Falony, G., Le Chatelier, E., Sunagawa, S., Prifti, E., Vieira-Silva, S., Gudmundsdottir, V., Krogh Pedersen, H.*, et al.* (2015). Disentangling type 2 diabetes and

metformin treatment signatures in the human gut microbiota. Nature *528*, 262-266.
Franzosa, E.A., Huang, K., Meadow, J.F., Gevers, D., Lemon, K.P., Bohannan, B.J., and Huttenhower, C. (2015). Identifying personal microbiomes using metagenomic codes. Proc Natl Acad Sci U S A *112*, E2930-2938.
Franzosa, E.A., McIver, L.J., Rahnavard, G., Thompson, L.R., Schirmer, M., Weingart, G., Lipson, K.S., Knight, R., Caporaso, J.G., Segata, N.*, et al.* (2018). Species-level functional profiling of metagenomes and metatranscriptomes. Nature methods *15*, 962-968.
Fu, J., Bonder, M.J., Cenit, M.C., Tigchelaar, E.F., Maatman, A., Dekens, J.A., Brandsma, E., Marczynska, J., Imhann, F., Weersma, R.K.*, et al.* (2015). The Gut Microbiome Contributes to a Substantial Proportion of the Variation in Blood Lipids. Circ Res *117*, 817-824.
Gensollen, T., Iyer, S.S., Kasper, D.L., and Blumberg, R.S. (2016). How colonization by microbiota in early life shapes the immune system. Science *352*, 539-544.
Ghosh, S.S., Wang, J., Yannie, P.J., and Ghosh, S. (2020). Intestinal Barrier Dysfunction, LPS Translocation, and Disease Development. J Endocr Soc *4*, bvz039.
Gibbons, S.M. (2019). Defining Microbiome Health through a Host Lens. mSystems *4*.
Gibbons, S.M., Kearney, S.M., Smillie, C.S., and Alm, E.J. (2017). Two dynamic regimes in the human gut microbiome. PLoS computational biology *13*, e1005364.
Gohl, D.M., Vangay, P., Garbe, J., MacLean, A., Hauge, A., Becker, A., Gould, T.J., Clayton, J.B., Johnson, T.J., Hunter, R.*, et al.* (2016). Systematic improvement of amplicon marker gene methods for increased accuracy in microbiome studies. Nat Biotechnol *34*, 942-949.
Goldstone, A.P., Thomas, E.L., Brynes, A.E., Bell, J.D., Frost, G., Saeed, N., Hajnal, J.V., Howard, J.K., Holland, A., and Bloom, S.R. (2001). Visceral adipose tissue and metabolic complications of obesity are reduced in Prader-Willi syndrome female adults: evidence for novel influences on body fat distribution. J Clin Endocrinol Metab *86*, 4330-4338.
Goodrich, J.K., Davenport, E.R., Beaumont, M., Jackson, M.A., Knight, R., Ober, C., Spector, T.D., Bell, J.T., Clark, A.G., and Ley, R.E. (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., Beaumont, M., Van Treuren, W., Knight, R., Bell, J.T.*, et al.* (2014). Human genetics shape the gut microbiome. Cell *159*, 789-799.
Greenbaum, S., Greenbaum, G., Moran-Gilad, J., and Weintraub, A.Y. (2019). Ecological dynamics of the vaginal microbiome in relation to health and disease. Am J Obstet Gynecol *220*, 324-335.

Greiner, T.U., and Bäckhed, F. (2016). Microbial regulation of GLP-1 and L-cell biology. Mol Metab *5*, 753-758.

Grine, G., Boualam, M.A., and Drancourt, M. (2017). Methanobrevibacter smithii, a methanogen consistently colonising the newborn stomach. Eur J Clin Microbiol Infect Dis *36*, 2449-2455.

Halfvarson, J., Brislawn, C.J., Lamendella, R., Vazquez-Baeza, Y., Walters, W.A., Bramer, L.M., D'Amato, M., Bonfiglio, F., McDonald, D., Gonzalez, A.*, et al.* (2017). Dynamics of the human gut microbiome in inflammatory bowel disease. Nat Microbiol *2*, 17004.

Hall, A.B., Yassour, M., Sauk, J., Garner, A., Jiang, X.F., Arthur, T., Lagoudas, G.K., Vatanen, T., Fornelos, N., Wilson, R.*, et al.* (2017). A novel Ruminococcus gnavus clade enriched in inflammatory bowel disease patients. Genome Medicine *9*.

Hamady, M., and Knight, R. (2009). Microbial community profiling for human microbiome projects: Tools, techniques, and challenges. Genome research *19*, 1141-1152.

Hansen, E.E., Lozupone, C.A., Rey, F.E., Wu, M., Guruge, J.L., Narra, A., Goodfellow, J., Zaneveld, J.R., McDonald, D.T., Goodrich, J.A.*, et al.* (2011). Pan-genome of the dominant human gut-associated archaeon, Methanobrevibacter smithii, studied in twins. Proc Natl Acad Sci U S A *108 Suppl 1*, 4599-4606.

Haqq, A.M., Muehlbauer, M., Svetkey, L.P., Newgard, C.B., Purnell, J.Q., Grambow, S.C., and Freemark, M.S. (2007). Altered distribution of adiponectin isoforms in children with Prader-Willi syndrome (PWS): association with insulin sensitivity and circulating satiety peptide hormones. Clin Endocrinol (Oxf) *67*, 944-951.

Haqq, A.M., Muehlbauer, M.J., Newgard, C.B., Grambow, S., and Freemark, M. (2011). The metabolic phenotype of Prader-Willi syndrome (PWS) in childhood: heightened insulin sensitivity relative to body mass index. J Clin Endocrinol Metab *96*, E225-232.

He, Y., Wu, W., Wu, S., Zheng, H.M., Li, P., Sheng, H.F., Chen, M.X., Chen, Z.H., Ji, G.Y., Zheng, Z.D.*, et al.* (2018a). Linking gut microbiota, metabolic syndrome and economic status based on a population-level analysis. Microbiome *6*, 172.

He, Y., Wu, W., Zheng, H.M., Li, P., McDonald, D., Sheng, H.F., Chen, M.X., Chen, Z.H., Ji, G.Y., Zheng, Z.D.*, et al.* (2018b). Regional variation limits applications of healthy gut microbiome reference ranges and disease models. Nat Med *24*, 1532-1535.

Hibberd, A.A., Yde, C.C., Ziegler, M.L., Honore, A.H., Saarinen, M.T., Lahtinen, S., Stahl, B., Jensen, H.M., and Stenman, L.K. (2019). Probiotic or synbiotic alters the gut microbiota and metabolism in a randomised controlled trial of weight management in overweight adults. Benef Microbes *10*, 121-135.

Hirano, S., Matsumoto, N., Morita, M., Sasaki, K., and Ohmura, N. (2013). Electrochemical control of redox potential affects methanogenesis of

the hydrogenotrophic methanogen Methanothermobacter thermautotrophicus. Lett Appl Microbiol *56*, 315-321.
Holling, C.S. (1973). Resilience and Stability of Ecological Systems. Annu Rev Ecol Syst *4*, 1-23.
Holmes, I., Harris, K., and Quince, C. (2012). Dirichlet multinomial mixtures: generative models for microbial metagenomics. PloS one *7*, e30126.
Hugerth, L.W., and Andersson, A.F. (2017). Analysing Microbial Community Composition through Amplicon Sequencing: From Sampling to Hypothesis Testing. Front Microbiol *8*, 1561.
Human Microbiome Project, C. (2012). Structure, function and diversity of the healthy human microbiome. Nature *486*, 207-214.
Ivanov, II, Atarashi, K., Manel, N., Brodie, E.L., Shima, T., Karaoz, U., Wei, D., Goldfarb, K.C., Santee, C.A., Lynch, S.V*., et al.* (2009). Induction of intestinal Th17 cells by segmented filamentous bacteria. Cell *139*, 485-498.
Jakobsson, H.E., Jernberg, C., Andersson, A.F., Sjolund-Karlsson, M., Jansson, J.K., and Engstrand, L. (2010). Short-term antibiotic treatment has differing long-term impacts on the human throat and gut microbiome. PloS one *5*, e9836.
Johnson, A.J., Vangay, P., Al-Ghalith, G.A., Hillmann, B.M., Ward, T.L., Shields-Cutler, R.R., Kim, A.D., Shmagel, A.K., Syed, A.N., Walter, J*., et al.* (2019). Daily Sampling Reveals Personalized Diet-Microbiome Associations in Humans. Cell Host & Microbe *25*, 789-+.
Jones, M.B., Highlander, S.K., Anderson, E.L., Li, W.Z., Dayrit, M., Klitgord, N., Fabani, M.M., Seguritan, V., Green, J., Pride, D.T*., et al.* (2015). Library preparation methodology can influence genomic and functional predictions in human microbiome research. P Natl Acad Sci USA *112*, 14024-14029.
Kamada, N., Chen, G.Y., Inohara, N., and Nunez, G. (2013). Control of pathogens and pathobionts by the gut microbiota. Nat Immunol *14*, 685-690.
Karlsson, F.H., Nookaew, I., and Nielsen, J. (2014). Metagenomic data utilization and analysis (MEDUSA) and construction of a global gut microbial gene catalogue. PLoS computational biology *10*, e1003706.
Khan, M.T., Duncan, S.H., Stams, A.J., van Dijl, J.M., Flint, H.J., and Harmsen, H.J. (2012). The gut anaerobe Faecalibacterium prausnitzii uses an extracellular electron shuttle to grow at oxic-anoxic interphases. ISME J *6*, 1578-1585.
Khan, M.T., Nieuwdorp, M., and Bäckhed, F. (2014). Microbial modulation of insulin sensitivity. Cell metabolism *20*, 753-760.
Kitahara, M., Takamine, F., Imamura, T., and Benno, Y. (2000). Assignment of Eubacterium sp. VPI 12708 and related strains with high bile acid 7alpha-dehydroxylating activity to Clostridium scindens and proposal of Clostridium hylemonae sp. nov., isolated from human faeces. Int J Syst Evol Microbiol *50 Pt 3*, 971-978.

Kitamoto, S., Alteri, C.J., Rodrigues, M., Nagao-Kitamoto, H., Sugihara, K., Himpsl, S.D., Bazzi, M., Miyoshi, M., Nishioka, T., Hayashi, A.*, et al.* (2020). Dietary L-serine confers a competitive fitness advantage to Enterobacteriaceae in the inflamed gut. Nat Microbiol *5*, 116-125.

Knights, D., Ward, T.L., McKinlay, C.E., Miller, H., Gonzalez, A., McDonald, D., and Knight, R. (2014). Rethinking "enterotypes". Cell Host Microbe *16*, 433-437.

Koenig, J.E., Spor, A., Scalfone, N., Fricker, A.D., Stombaugh, J., Knight, R., Angenent, L.T., and Ley, R.E. (2011). Succession of microbial consortia in the developing infant gut microbiome. Proc Natl Acad Sci U S A *108 Suppl 1*, 4578-4585.

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.

Koh, A., Molinaro, A., Stahlman, M., Khan, M.T., Schmidt, C., Manneras-Holm, L., Wu, H., Carreras, A., Jeong, H., Olofsson, L.E.*, et al.* (2018). Microbially Produced Imidazole Propionate Impairs Insulin Signaling through mTORC1. Cell *175*, 947-961 e917.

Kolodziejczyk, A.A., Zheng, D., and Elinav, E. (2019). Diet-microbiota interactions and personalized nutrition. Nature reviews Microbiology *17*, 742-753.

Konikoff, T., and Gophna, U. (2016). Oscillospira: a Central, Enigmatic Component of the Human Gut Microbiota. Trends Microbiol *24*, 523-524.

Korem, T., Zeevi, D., Zmora, N., Weissbrod, O., Bar, N., Lotan-Pompan, M., Avnit-Sagi, T., Kosower, N., Malka, G., Rein, M.*, et al.* (2017). Bread Affects Clinical Parameters and Induces Gut Microbiome-Associated Personal Glycemic Responses. Cell metabolism *25*, 1243-+.

Korpela, K., Costea, P., Coelho, L.P., Kandels-Lewis, S., Willemsen, G., Boomsma, D.I., Segata, N., and Bork, P. (2018). Selective maternal seeding and environment shape the human gut microbiome. Genome research *28*, 561-568.

Korpela, K., Salonen, A., Virta, L.J., Kekkonen, R.A., Forslund, K., Bork, P., and de Vos, W.M. (2016). Intestinal microbiome is related to lifetime antibiotic use in Finnish pre-school children. Nature communications *7*, 10410.

Kovatcheva-Datchary, P., Nilsson, A., Akrami, R., Lee, Ying S., De Vadder, F., Arora, T., Hallen, A., Martens, E., Björck, I., and Bäckhed, F. (2015). Dietary Fiber-Induced Improvement in Glucose Metabolism Is Associated with Increased Abundance of Prevotella. Cell metabolism *22*, 971-982.

Kozich, J.J., Westcott, S.L., Baxter, N.T., Highlander, S.K., and Schloss, P.D. (2013). Development of a dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence data on the MiSeq Illumina sequencing platform. Appl Environ Microbiol *79*, 5112-5120.

Krautkramer, K.A., Kreznar, J.H., Romano, K.A., Vivas, E.I., Barrett-Wilt, G.A., Rabaglia, M.E., Keller, M.P., Attie, A.D., Rey, F.E., and Denu, J.M. (2016). Diet-Microbiota Interactions Mediate Global Epigenetic Programming in Multiple Host Tissues. Mol Cell *64*, 982-992.

Kriss, M., Hazleton, K.Z., Nusbacher, N.M., Martin, C.G., and Lozupone, C.A. (2018). Low diversity gut microbiota dysbiosis: drivers, functional implications and recovery. Curr Opin Microbiol *44*, 34-40.

Krumbeck, J.A., Rasmussen, H.E., Hutkins, R.W., Clarke, J., Shawron, K., Keshavarzian, A., and Walter, J. (2018). Probiotic Bifidobacterium strains and galactooligosaccharides improve intestinal barrier function in obese adults but show no synergism when used together as synbiotics. Microbiome *6*, 121.

Lacroix, D., Moutel, S., Coupaye, M., Huvenne, H., Faucher, P., Pelloux, V., Rouault, C., Bastard, J.P., Cagnard, N., Dubern, B.*, et al.* (2015). Metabolic and adipose tissue signatures in adults with Prader-Willi syndrome: a model of extreme adiposity. J Clin Endocrinol Metab *100*, 850-859.

Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. Nature methods *9*, 357-359.

Laursen, M.F., Andersen, L.B.B., Michaelsen, K.F., Molgaard, C., Trolle, E., Bahl, M.I., and Licht, T.R. (2016). Infant Gut Microbiota Development Is Driven by Transition to Family Foods Independent of Maternal Obesity. Msphere *1*.

Le Chatelier, E., Nielsen, T., Qin, J., Prifti, E., Hildebrand, F., Falony, G., Almeida, M., Arumugam, M., Batto, J.M., Kennedy, S.*, et al.* (2013). Richness of human gut microbiome correlates with metabolic markers. Nature *500*, 541-546.

Ley, R.E., Bäckhed, F., Turnbaugh, P., Lozupone, C.A., Knight, R.D., and Gordon, J.I. (2005). Obesity alters gut microbial ecology. Proc Natl Acad Sci U S A *102*, 11070-11075.

Lim, M.Y., You, H.J., Yoon, H.S., Kwon, B., Lee, J.Y., Lee, S., Song, Y.M., Lee, K., Sung, J., and Ko, G. (2017). The effect of heritability and host genetics on the gut microbiota and metabolic syndrome. Gut *66*, 1031-1038.

Lindgren, A.C., Hagenas, L., and Ritzen, E.M. (1999). Growth hormone treatment of children with Prader-Willi syndrome: effects on glucose and insulin homeostasis. Swedish National Growth Hormone Advisory Group. Horm Res *51*, 157-161.

Lippert, K., Kedenko, L., Antonielli, L., Kedenko, I., Gemeier, C., Leitner, M., Kautzky-Willer, A., Paulweber, B., and Hackl, E. (2017). Gut microbiota dysbiosis associated with glucose metabolism disorders and the metabolic syndrome in older adults. Benef Microbes *8*, 545-556.

Litvak, Y., Byndloss, M.X., Tsolis, R.M., and Baumler, A.J. (2017). Dysbiotic Proteobacteria expansion: a microbial signature of epithelial dysfunction. Curr Opin Microbiol *39*, 1-6.

Lloyd-Price, J., Abu-Ali, G., and Huttenhower, C. (2016). The healthy human microbiome. Genome Med *8*, 51.

Lloyd-Price, J., Mahurkar, A., Rahnavard, G., Crabtree, J., Orvis, J., Hall, A.B., Brady, A., Creasy, H.H., McCracken, C., Giglio, M.G.*, et al.* (2017). Strains, functions and dynamics in the expanded Human Microbiome Project. Nature *550*, 61-66.

Lombard, V., Golaconda Ramulu, H., Drula, E., Coutinho, P.M., and Henrissat, B. (2014). The carbohydrate-active enzymes database (CAZy) in 2013. Nucleic acids research *42*, D490-495.

Lozupone, C., and Knight, R. (2005). UniFrac: a new phylogenetic method for comparing microbial communities. Appl Environ Microbiol *71*, 8228-8235.

Lozupone, C.A., Stombaugh, J.I., Gordon, J.I., Jansson, J.K., and Knight, R. (2012). Diversity, stability and resilience of the human gut microbiota. Nature *489*, 220-230.

Lynch, S.V., and Pedersen, O. (2016). The Human Intestinal Microbiome in Health and Disease. N Engl J Med *375*, 2369-2379.

Macfarlane, G.T., and Englyst, H.N. (1986). Starch utilization by the human large intestinal microflora. J Appl Bacteriol *60*, 195-201.

Mackie, R.I., Sghir, A., and Gaskins, H.R. (1999). Developmental microbial ecology of the neonatal gastrointestinal tract. Am J Clin Nutr *69*, 1035S-1045S.

Maczulak, A.E., Wolin, M.J., and Miller, T.L. (1989). Increase in colonic methanogens and total anaerobes in aging rats. Appl Environ Microbiol *55*, 2468-2473.

Makki, K., Deehan, E.C., Walter, J., and Bäckhed, F. (2018). The Impact of Dietary Fiber on Gut Microbiota in Host Health and Disease. Cell Host Microbe *23*, 705-715.

Mancabelli, L., Milani, C., Lugli, G.A., Turroni, F., Cocconi, D., van Sinderen, D., and Ventura, M. (2017). Identification of universal gut microbial biomarkers of common human intestinal diseases by meta-analysis. FEMS Microbiol Ecol *93*.

Manor, O., and Borenstein, E. (2017). Revised computational metagenomic processing uncovers hidden and biologically meaningful functional variation in the human microbiome. Microbiome *5*.

Marchesi, J.R., and Ravel, J. (2015). The vocabulary of microbiome research: a proposal. Microbiome *3*, 31.

Mardinoglu, A., Wu, H., Bjornson, E., Zhang, C., Hakkarainen, A., Rasanen, S.M., Lee, S., Mancina, R.M., Bergentall, M., Pietilainen, K.H.*, et al.* (2018). An Integrated Understanding of the Rapid Metabolic Benefits of a Carbohydrate-Restricted Diet on Hepatic Steatosis in Humans. Cell metabolism *27*, 559-571 e555.

McBurney, M.I., Davis, C., Fraser, C.M., Schneeman, B.O., Huttenhower, C., Verbeke, K., Walter, J., and Latulippe, M.E. (2019). Establishing What Constitutes a Healthy Human Gut Microbiome: State of the Science,

Regulatory Considerations, and Future Directions. J Nutr *149*, 1882-1895.
Mcneil, N.I. (1984). The Contribution of the Large-Intestine to Energy Supplies in Man. American Journal of Clinical Nutrition *39*, 338-342.
Mehta, R.S., Abu-Ali, G.S., Drew, D.A., Lloyd-Price, J., Subramanian, A., Lochhead, P., Joshi, A.D., Ivey, K.L., Khalili, H., Brown, G.T.*, et al.* (2018). Stability of the human faecal microbiome in a cohort of adult men. Nat Microbiol *3*, 347-355.
Mendes-Soares, H., Raveh-Sadka, T., Azulay, S., Edens, K., Ben-Shlomo, Y., Cohen, Y., Ofek, T., Bachrach, D., Stevens, J., Colibaseanu, D.*, et al.* (2019). Assessment of a Personalized Approach to Predicting Postprandial Glycemic Responses to Food Among Individuals Without Diabetes. JAMA Netw Open *2*, e188102.
Mendrick, D.L., Diehl, A.M., Topor, L.S., Dietert, R.R., Will, Y., La Merrill, M.A., Bouret, S., Varma, V., Hastings, K.L., Schug, T.T.*, et al.* (2018). Metabolic Syndrome and Associated Diseases: From the Bench to the Clinic. Toxicol Sci *162*, 36-42.
Miceli, J.F., Torres, C.I., and Krajmalnik-Brown, R. (2016). Shifting the balance of fermentation products between hydrogen and volatile fatty acids: microbial community structure and function. Fems Microbiology Ecology *92*.
Million, M., Tidjani Alou, M., Khelaifia, S., Bachar, D., Lagier, J.C., Dione, N., Brah, S., Hugon, P., Lombard, V., Armougom, F.*, et al.* (2016). Increased Gut Redox and Depletion of Anaerobic and Methanogenic Prokaryotes in Severe Acute Malnutrition. Sci Rep *6*, 26051.
Naeem, S., Thompson, L.J., Lawler, S.P., Lawton, J.H., and Woodfin, R.M. (1994). Declining Biodiversity Can Alter the Performance of Ecosystems. Nature *368*, 734-737.
Nardone, G., and Compare, D. (2015). The human gastric microbiota: Is it time to rethink the pathogenesis of stomach diseases? United Eur Gastroent *3*, 255-260.
Nielsen, H.B., Almeida, M., Juncker, A.S., Rasmussen, S., Li, J., Sunagawa, S., Plichta, D.R., Gautier, L., Pedersen, A.G., Le Chatelier, E.*, et al.* (2014). Identification and assembly of genomes and genetic elements in complex metagenomic samples without using reference genomes. Nat Biotechnol *32*, 822-828.
Olesen, S.W., and Alm, E.J. (2016). Dysbiosis is not an answer. Nat Microbiol *1*, 16228.
Oliphant, K., and Allen-Vercoe, E. (2019). Macronutrient metabolism by the human gut microbiome: major fermentation by-products and their impact on host health. Microbiome *7*.
Palleja, A., Mikkelsen, K.H., Forslund, S.K., Kashani, A., Allin, K.H., Nielsen, T., Hansen, T.H., Liang, S., Feng, Q., Zhang, C.*, et al.* (2018). Recovery of gut microbiota of healthy adults following antibiotic exposure. Nat Microbiol *3*, 1255-1265.

Pascal, V., Pozuelo, M., Borruel, N., Casellas, F., Campos, D., Santiago, A., Martinez, X., Varela, E., Sarrabayrouse, G., Machiels, K., *et al.* (2017). A microbial signature for Crohn's disease. Gut *66*, 813-822.

Planer, J.D., Peng, Y., Kau, A.L., Blanton, L.V., Ndao, I.M., Tarr, P.I., Warner, B.B., and Gordon, J.I. (2016). Development of the gut microbiota and mucosal IgA responses in twins and gnotobiotic mice. Nature *534*, 263-266.

Png, C.W., Linden, S.K., Gilshenan, K.S., Zoetendal, E.G., McSweeney, C.S., Sly, L.I., McGuckin, M.A., and Florin, T.H.J. (2010). Mucolytic Bacteria With Increased Prevalence in IBD Mucosa Augment In Vitro Utilization of Mucin by Other Bacteria. Am J Gastroenterol *105*, 2420-2428.

Poyet, M., Groussin, M., Gibbons, S.M., Avila-Pacheco, J., Jiang, X., Kearney, S.M., Perrotta, A.R., Berdy, B., Zhao, S., Lieberman, T.D., *et al.* (2019). A library of human gut bacterial isolates paired with longitudinal multiomics data enables mechanistic microbiome research. Nat Med *25*, 1442-1452.

Qin, J., Li, R., Raes, J., Arumugam, M., Burgdorf, K.S., Manichanh, C., Nielsen, T., Pons, N., Levenez, F., Yamada, T., *et al.* (2010). A human gut microbial gene catalogue established by metagenomic sequencing. Nature *464*, 59-65.

Rajilic-Stojanovic, M., Heilig, H.G., Tims, S., Zoetendal, E.G., and de Vos, W.M. (2012). Long-term monitoring of the human intestinal microbiota composition. Environ Microbiol.

Reese, A.T., Cho, E.H., Klitzman, B., Nichols, S.P., Wisniewski, N.A., Villa, M.M., Durand, H.K., Jiang, S., Midani, F.S., Nimmagadda, S.N., *et al.* (2018). Antibiotic-induced changes in the microbiota disrupt redox dynamics in the gut. Elife *7*.

Reese, A.T., and Dunn, R.R. (2018). Drivers of Microbiome Biodiversity: A Review of General Rules, Feces, and Ignorance. mBio *9*.

Reinhardt, C., Bergentall, M., Greiner, T.U., Schaffner, F., Ostergren-Lunden, G., Petersen, L.C., Ruf, W., and Bäckhed, F. (2012). Tissue factor and PAR1 promote microbiota-induced intestinal vascular remodelling. Nature *483*, 627-631.

Rey, F.E., Gonzalez, M.D., Cheng, J.Y., Wu, M., Ahern, P.P., and Gordon, J.I. (2013). Metabolic niche of a prominent sulfate-reducing human gut bacterium. P Natl Acad Sci USA *110*, 13582-13587.

Reyman, M., van Houten, M.A., van Baarle, D., Bosch, A., Man, W.H., Chu, M., Arp, K., Watson, R.L., Sanders, E.A.M., Fuentes, S., *et al.* (2019). Impact of delivery mode-associated gut microbiota dynamics on health in the first year of life. Nature communications *10*, 4997.

Rivera-Chavez, F., Lopez, C.A., and Baumler, A.J. (2017). Oxygen as a driver of gut dysbiosis. Free Radical Bio Med *105*, 93-101.

Rothschild, D., Weissbrod, O., Barkan, E., Kurilshikov, A., Korem, T., Zeevi, D., Costea, P.I., Godneva, A., Kalka, I.N., Bar, N., *et al.* (2018).

Environment dominates over host genetics in shaping human gut microbiota. Nature *555*, 210-215.

Ruaud, A., Esquivel-Elizondo, S., de la Cuesta-Zuluaga, J., Waters, J.L., Angenent, L.T., Youngblut, N.D., and Ley, R.E. (2020). Syntrophy via Interspecies H2 Transfer between Christensenella and Methanobrevibacter Underlies Their Global Cooccurrence in the Human Gut. mBio *11*.

Ryan, F.J., Ahern, A.M., Fitzgerald, R.S., Laserna-Mendieta, E.J., Power, E.M., Clooney, A.G., O'Donoghue, K.W., McMurdie, P.J., Iwai, S., Crits-Christoph, A.*, et al.* (2020). Colonic microbiota is associated with inflammation and host epigenomic alterations in inflammatory bowel disease. Nature communications *11*, 1512.

Salonen, A., Lahti, L., Salojarvi, J., Holtrop, G., Korpela, K., Duncan, S.H., Date, P., Farquharson, F., Johnstone, A.M., Lobley, G.E.*, et al.* (2014). Impact of diet and individual variation on intestinal microbiota composition and fermentation products in obese men. ISME J *8*, 2218-2230.

Salonen, A., Nikkila, J., Jalanka-Tuovinen, J., Immonen, O., Rajilic-Stojanovic, M., Kekkonen, R.A., Palva, A., and de Vos, W.M. (2010). Comparative analysis of fecal DNA extraction methods with phylogenetic microarray: effective recovery of bacterial and archaeal DNA using mechanical cell lysis. Journal of microbiological methods *81*, 127-134.

Schloissnig, S., Arumugam, M., Sunagawa, S., Mitreva, M., Tap, J., Zhu, A., Waller, A., Mende, D.R., Kultima, J.R., Martin, J.*, et al.* (2013). Genomic variation landscape of the human gut microbiome. Nature *493*, 45-50.

Schloss, P.D., and Handelsman, J. (2005). Introducing DOTUR, a computer program for defining operational taxonomic units and estimating species richness. Appl Environ Microbiol *71*, 1501-1506.

Schnorr, S.L., Candela, M., Rampelli, S., Centanni, M., Consolandi, C., Basaglia, G., Turroni, S., Biagi, E., Peano, C., Severgnini, M.*, et al.* (2014). Gut microbiome of the Hadza hunter-gatherers. Nature communications *5*, 3654.

Schroeder, B.O., and Bäckhed, F. (2016). Signals from the gut microbiota to distant organs in physiology and disease. Nat Med *22*, 1079-1089.

Schroeder, B.O., Birchenough, G.M.H., Stahlman, M., Arike, L., Johansson, M.E.V., Hansson, G.C., and Bäckhed, F. (2018). Bifidobacteria or Fiber Protects against Diet-Induced Microbiota-Mediated Colonic Mucus Deterioration. Cell Host Microbe *23*, 27-40 e27.

Schuster, D.P., Osei, K., and Zipf, W.B. (1996). Characterization of alterations in glucose and insulin metabolism in Prader-Willi subjects. Metabolism *45*, 1514-1520.

Segata, N., Waldron, L., Ballarini, A., Narasimhan, V., Jousson, O., and Huttenhower, C. (2012). Metagenomic microbial community profiling using unique clade-specific marker genes. Nature methods *9*, 811-814.

Sender, R., Fuchs, S., and Milo, R. (2016). Are We Really Vastly Outnumbered? Revisiting the Ratio of Bacterial to Host Cells in Humans. Cell *164*, 337-340.

Shade, A., and Gilbert, J.A. (2015). Temporal patterns of rarity provide a more complete view of microbial diversity. Trends Microbiol *23*, 335-340.

Shade, A., Jones, S.E., Caporaso, J.G., Handelsman, J., Knight, R., Fierer, N., and Gilbert, J.A. (2014). Conditionally rare taxa disproportionately contribute to temporal changes in microbial diversity. MBio *5*, e01371-01314.

Shao, Y., Forster, S.C., Tsaliki, E., Vervier, K., Strang, A., Simpson, N., Kumar, N., Stares, M.D., Rodger, A., Brocklehurst, P.*, et al.* (2019). Stunted microbiota and opportunistic pathogen colonization in caesarean-section birth. Nature *574*, 117-121.

Sim, K., Cox, M.J., Wopereis, H., Martin, R., Knol, J., Li, M.S., Cookson, W.O., Moffatt, M.F., and Kroll, J.S. (2012). Improved detection of bifidobacteria with optimised 16S rRNA-gene based pyrosequencing. PloS one *7*, e32543.

Singer, J.R., Blosser, E.G., Zindl, C.L., Silberger, D.J., Conlan, S., Laufer, V.A., DiToro, D., Deming, C., Kumar, R., Morrow, C.D.*, et al.* (2019). Preventing dysbiosis of the neonatal mouse intestinal microbiome protects against late-onset sepsis. Nat Med *25*, 1772-1782.

Singh, R.K., Chang, H.W., Yan, D., Lee, K.M., Ucmak, D., Wong, K., Abrouk, M., Farahnik, B., Nakamura, M., Zhu, T.H.*, et al.* (2017). Influence of diet on the gut microbiome and implications for human health. J Transl Med *15*.

Smith, N.W., Shorten, P.R., Altermann, E.H., Roy, N.C., and McNabb, W.C. (2019). Hydrogen cross-feeders of the human gastrointestinal tract. Gut Microbes *10*, 270-288.

Smits, S.A., Leach, J., Sonnenburg, E.D., Gonzalez, C.G., Lichtman, J.S., Reid, G., Knight, R., Manjurano, A., Changalucha, J., Elias, J.E.*, et al.* (2017). Seasonal cycling in the gut microbiome of the Hadza hunter-gatherers of Tanzania. Science *357*, 802-806.

Sommer, F., Anderson, J.M., Bharti, R., Raes, J., and Rosenstiel, P. (2017). The resilience of the intestinal microbiota influences health and disease. Nature reviews Microbiology *15*, 630-638.

Sommer, F., and Bäckhed, F. (2013). The gut microbiota--masters of host development and physiology. Nature reviews Microbiology *11*, 227-238.

Sonnenburg, E.D., and Sonnenburg, J.L. (2014). Starving our Microbial Self: The Deleterious Consequences of a Diet Deficient in Microbiota-Accessible Carbohydrates. Cell metabolism *20*, 779-786.

Sonnenburg, J.L., and Bäckhed, F. (2016). Diet-microbiota interactions as moderators of human metabolism. Nature *535*, 56-64.
Sonnenburg, J.L., and Sonnenburg, E.D. (2019). Vulnerability of the industrialized microbiota. Science *366*.
Sprockett, D., Fukami, T., and Relman, D.A. (2018). Role of priority effects in the early-life assembly of the gut microbiota. Nat Rev Gastroenterol Hepatol *15*, 197-205.
Staley, C., Kaiser, T., Beura, L.K., Hamilton, M.J., Weingarden, A.R., Bobr, A., Kang, J., Masopust, D., Sadowsky, M.J., and Khoruts, A. (2017). Stable engraftment of human microbiota into mice with a single oral gavage following antibiotic conditioning. Microbiome *5*.
Stark, P.L., and Lee, A. (1982). The microbial ecology of the large bowel of breast-fed and formula-fed infants during the first year of life. J Med Microbiol *15*, 189-203.
Stewart, C.J., Ajami, N.J., O'Brien, J.L., Hutchinson, D.S., Smith, D.P., Wong, M.C., Ross, M.C., Lloyd, R.E., Doddapaneni, H., Metcalf, G.A.*, et al.* (2018). Temporal development of the gut microbiome in early childhood from the TEDDY study. Nature *562*, 583-588.
Stewart, J.A., Chadwick, V.S., and Murray, A. (2006). Carriage, quantification, and predominance of methanogens and sulfate-reducing bacteria in faecal samples. Lett Appl Microbiol *43*, 58-63.
Subramanian, S., Huq, S., Yatsunenko, T., Haque, R., Mahfuz, M., Alam, M.A., Benezra, A., DeStefano, J., Meier, M.F., Muegge, B.D.*, et al.* (2014). Persistent gut microbiota immaturity in malnourished Bangladeshi children. Nature *510*, 417-421.
Sunagawa, S., Mende, D.R., Zeller, G., Izquierdo-Carrasco, F., Berger, S.A., Kultima, J.R., Coelho, L.P., Arumugam, M., Tap, J., Nielsen, H.B.*, et al.* (2013). Metagenomic species profiling using universal phylogenetic marker genes. Nature methods *10*, 1196-1199.
Tailford, L.E., Crost, E.H., Kavanaugh, D., and Juge, N. (2015). Mucin glycan foraging in the human gut microbiome. Front Genet *6*.
Talebizadeh, Z., and Butler, M.G. (2005). Insulin resistance and obesity-related factors in Prader-Willi syndrome: comparison with obese subjects. Clin Genet *67*, 230-239.
Turnbaugh, P.J., Hamady, M., Yatsunenko, T., Cantarel, B.L., Duncan, A., Ley, R.E., Sogin, M.L., Jones, W.J., Roe, B.A., Affourtit, J.P.*, et al.* (2009). A core gut microbiome in obese and lean twins. Nature *457*, 480-484.
Valles-Colomer, M., Falony, G., Darzi, Y., Tigchelaar, E.F., Wang, J., Tito, R.Y., Schiweck, C., Kurilshikov, A., Joossens, M., Wijmenga, C.*, et al.* (2019). The neuroactive potential of the human gut microbiota in quality of life and depression. Nat Microbiol.
van de Pol, J.A., van Best, N., Mbakwa, C.A., Thijs, C., Savelkoul, P.H., Arts, I.C., Hornef, M.W., Mommers, M., and Penders, J. (2017). Gut

Colonization by Methanogenic Archaea Is Associated with Organic Dairy Consumption in Children. Front Microbiol *8*, 355.

Vandeputte, D., Kathagen, G., D'Hoe, K., Vieira-Silva, S., Valles-Colomer, M., Sabino, J., Wang, J., Tito, R.Y., De Commer, L., Darzi, Y.*, et al.* (2017). Quantitative microbiome profiling links gut community variation to microbial load. Nature *551*, 507-511.

Vangay, P., Johnson, A.J., Ward, T.L., Al-Ghalith, G.A., Shields-Cutler, R.R., Hillmann, B.M., Lucas, S.K., Beura, L.K., Thompson, E.A., Till, L.M.*, et al.* (2018). US Immigration Westernizes the Human Gut Microbiome. Cell *175*, 962-+.

Vatanen, T., Plichta, D.R., Somani, J., Münch, P.C., Arthur, T.D., Hall, A.B., Rudolf, S., Oakeley, E.J., Ke, X., Young, R.A.*, et al.* (2019). Genomic variation and strain-specific functional adaptation in the human gut microbiome during early life. Nature Microbiology *4*, 470-479.

Vieira-Silva, S., Sabino, J., Valles-Colomer, M., Falony, G., Kathagen, G., Caenepeel, C., Cleynen, I., van der Merwe, S., Vermeire, S., and Raes, J. (2019). Quantitative microbiome profiling disentangles inflammation- and bile duct obstruction-associated microbiota alterations across PSC/IBD diagnoses. Nat Microbiol *4*, 1826-1831.

Vital, M., Karch, A., and Pieper, D.H. (2017). Colonic Butyrate-Producing Communities in Humans: an Overview Using Omics Data. mSystems *2*.

Vollmers, J., Wiegand, S., and Kaster, A.K. (2017). Comparing and Evaluating Metagenome Assembly Tools from a Microbiologist's Perspective - Not Only Size Matters! PloS one *12*, e0169662.

Wahlstrom, 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 metabolism *24*, 41-50.

Walker, A.W., Martin, J.C., Scott, P., Parkhill, J., Flint, H.J., and Scott, K.P. (2015). 16S rRNA gene-based profiling of the human infant gut microbiota is strongly influenced by sample processing and PCR primer choice. Microbiome *3*, 26.

Walter, J., Armet, A.M., Finlay, B.B., and Shanahan, F. (2020). Establishing or Exaggerating Causality for the Gut Microbiome: Lessons from Human Microbiota-Associated Rodents. Cell *180*, 221-232.

Waters, J.L., and Ley, R.E. (2019). The human gut bacteria Christensenellaceae are widespread, heritable, and associated with health. BMC Biol *17*, 83.

Wolfe, A.J. (2015). Glycolysis for Microbiome Generation. Microbiol Spectr *3*.

Wu, G.D., Chen, J., Hoffmann, C., Bittinger, K., Chen, Y.Y., Keilbaugh, S.A., Bewtra, M., Knights, D., Walters, W.A., Knight, R.*, et al.* (2011). Linking long-term dietary patterns with gut microbial enterotypes. Science *334*, 105-108.

Yatsunenko, T., Rey, F.E., Manary, M.J., Trehan, I., Dominguez-Bello, M.G., Contreras, M., Magris, M., Hidalgo, G., Baldassano, R.N., Anokhin, A.P., *et al.* (2012). Human gut microbiome viewed across age and geography. Nature *486*, 222-227.

Yoo, W., and Byndloss, M.X. (2020). How to thrive in the inflamed gut. Nat Microbiol *5*, 10-11.

Zaneveld, J.R., McMinds, R., and Vega Thurber, R. (2017). Stress and stability: applying the Anna Karenina principle to animal microbiomes. Nat Microbiol *2*, 17121.

Ze, X., Le Mougen, F., Duncan, S.H., Louis, P., and Flint, H.J. (2013). Some are more equal than others: the role of "keystone" species in the degradation of recalcitrant substrates. Gut Microbes *4*, 236-240.

Zeevi, D., Korem, T., Zmora, N., Israeli, D., Rothschild, D., Weinberger, A., Ben-Yacov, O., Lador, D., Avnit-Sagi, T., Lotan-Pompan, M., *et al.* (2015). Personalized Nutrition by Prediction of Glycemic Responses. Cell *163*, 1079-1094.

Zhang, C., Yin, A., Li, H., Wang, R., Wu, G., Shen, J., Zhang, M., Wang, L., Hou, Y., Ouyang, H., *et al.* (2015). Dietary Modulation of Gut Microbiota Contributes to Alleviation of Both Genetic and Simple Obesity in Children. EBioMedicine *2*, 966-982.

Zhang, L., Bahl, M.I., Roager, H.M., Fonvig, C.E., Hellgren, L.I., Frandsen, H.L., Pedersen, O., Holm, J.C., Hansen, T., and Licht, T.R. (2017). Environmental spread of microbes impacts the development of metabolic phenotypes in mice transplanted with microbial communities from humans. ISME J *11*, 676-690.

Zhong, H., Penders, J., Shi, Z., Ren, H., Cai, K., Fang, C., Ding, Q., Thijs, C., Blaak, E.E., Stehouwer, C.D.A., *et al.* (2019). Impact of early events and lifestyle on the gut microbiota and metabolic phenotypes in young school-age children. Microbiome *7*, 2.

Zhou, W., Sailani, M.R., Contrepois, K., Zhou, Y., Ahadi, S., Leopold, S.R., Zhang, M.J., Rao, V., Avina, M., Mishra, T., *et al.* (2019). Longitudinal multi-omics of host-microbe dynamics in prediabetes. Nature *569*, 663-671.

Zoetendal, E.G., Akkermans, A.D., and De Vos, W.M. (1998). Temperature gradient gel electrophoresis analysis of 16S rRNA from human fecal samples reveals stable and host-specific communities of active bacteria. Appl Environ Microbiol *64*, 3854-3859.