JONNA JALANKA

Characterization of Intestinal Microbiota in Healthy Adults and the Effect of Perturbations

DEPARTMENT OF VETERINARY BIO SCIENCES
FACULTY OF VETERINARY MEDICINE
DOCTORAL PROGRAMME IN FOOD CHAIN AND HEALTH
UNIVERSITY OF HELSINKI
Faculty of Veterinary Medicine,
Department of Veterinary Biosciences,
University of Helsinki, Finland

CHARACTERIZATION OF INTESTINAL MICROBIOTA
IN HEALTHY ADULTS AND THE EFFECT OF
PERTURBATIONS

Jonna Jalanka

ACADEMIC DISSERTATION

To be presented, with the permission of the Faculty of Veterinary Medicine of
the University of Helsinki, for public examination in the Walter lecture room, EE-building,
Agnes Sjöbergin katu 2, Helsinki, on December 5th, 2014, at 12 o’clock.

Helsinki 2014
All disease begins in the gut

Hippocrates
Contents

ORIGINAL ARTICLES

ABBREVIATIONS

ABSTRACT

1. INTRODUCTION............................................................................................................................1

2. REVIEW OF THE LITERATURE..................................................................................................2
   2.1 An Overview of the Molecular Methods Used to Study the Intestinal Microbiota .... 2
       2.1.1 DNA Extraction ...................................................................................................................2
       2.1.2 Molecular Methods for Analysing the Intestinal Microbiota Community .......... 4
   2.2 Gastrointestinal Microbiota in Healthy Subjects...............................................................7
       2.2.1 Characteristics of Healthy Adult Intestinal Microbiota....................................................7
   2.3 Effects of Perturbation to Healthy Intestinal Microbiota...............................................11
       2.3.1 Effect of Bowel Cleansing to the Intestinal Microbiota ................................................. 13
       2.3.2 Differences in the Intestinal Microbiota in Diseases - Irritable Bowel Syndrome ...........14

3. AIMS OF THE STUDY ..............................................................................................................18

4. MATERIALS AND METHODS ................................................................................................19
   4.1 Study Subjects....................................................................................................................19
   4.2 Analysis Protocol................................................................................................................ 19

5. RESULTS AND DISCUSSION ..................................................................................................22
   5.1 Comparison of Faecal DNA Extraction Methods (Study I) ...........................................22
       5.1.1 Differences in the Microbiota Profiles Between the DNA Extraction Methods...............22
   5.2 Characterization of the Healthy Microbiota (Study II) ..................................................24
       5.2.1 Healthy Microbiota Composition and Temporal Variation ............................................24
       5.2.2 Definition of the Common Core Microbiota ..................................................................26
       5.2.3 Mild Intestinal Symptoms Associated with Microbial Composition .................28
   5.3 Perturbations to the Healthy Microbiota – the Effect of Bowel Cleansing (Study III) ..........................................................29
       5.3.1 The Effect of Bowel Cleansing on the Intestinal Microbiota Composition ........30
       5.3.2 Changes in the Intestinal Microbiota Composition Immediately After Lavage ....30
       5.3.3 Impact of Different Doses of the Laxative Treatment on the Microbiota ...........31
   5.4 Perturbations to the Healthy Microbiota – Microbial Changes in Post-Infectious Irritable Bowel Syndrome (Study IV) ..............................................................32
       5.4.1 The Index of Microbial Dysbiosis ..................................................................................32
       5.4.2 Associations Between Microbiota and Host Gene Expression ................................35

6. CONCLUSIONS AND FUTURE ASPECTS ............................................................................37

7. ACKNOWLEDGEMENTS.........................................................................................................40

8. REFERENCES ..........................................................................................................................42
Original Articles

This thesis is based on the following studies, which are referred to in the text by their Roman number:


The original articles are reprinted with the kind permission of the publishers.
Abbreviations

CoV    coefficient of variation
DCL    differential centrifugation and lysis method
DGGE   denaturing gradient gel electrophoresis
FISH   fluorescent in situ hybridization
FODMAP fermentable, oligo-, di-, mono-saccharides and polyols
FMT    Faecal microbial transplant
GI     gastrointestinal tract
GSA    gene set analysis
HAD    hospital anxiety and depression score
HC     healthy control
HITChip human intestinal track chip
HMP    human microbiome project
HRQoL  health related quality of life
IBD    inflammatory bowel disease
IBS    irritable bowel syndrome
IBS-A  irritable bowel syndrome with alternating symptoms
IBS-C  constipation predominant irritable bowel syndrome
IBS-D  diarrhoea predominant irritable bowel syndrome
IMD    index on microbial dysbiosis
MetaHIT metagenomics of the human intestinal tract
MPR    modified version of the Promega Genomic Wizard DNA Purification kit
NGS    next generation sequencing
PI-BD  post C. jejuni infection bowel dysfunction
PI-IBS post-infectious irritable bowel syndrome
PI-nonBD post C. jejuni infection with out bowel dysfunction
qPCR   quantitative polymerase chain reaction
QSK    QIAamp* DNA stool mini kit
RBB    repeated bead-beating
rDNA   ribosomal deoxyribonucleic acid
rRNA   ribosomal ribonucleic acid
SDS    sodium dodecyl sulphate
T-RFLP terminal restriction fragment length polymorphism
TLR    toll-like receptor
TNF    tumour necrosis factor
ZO-1   tight junction protein 1
Abstract

In the research described in this thesis, the intestinal microbiota of healthy subjects was characterized in detail and benchmarked against two perturbations, bowel cleansing and irritable bowel syndrome, both hypothesized to alter the microbial composition. First, four commonly used faecal DNA extraction protocols were compared to determine the optimal method to be utilized in all of the following projects. The extraction of DNA from faecal samples is a crucial step in molecular analyses as incomplete cell lysis of certain members of the diverse intestinal ecosystem can bias the resultant community composition. We detected performance differences between the mechanical and enzymatic methods as well as with commercial kits. The DNA yields varied by 35-fold, moreover the abundance and diversity of different bacterial groups were found to be significantly different between the extraction methods. The levels of certain Firmicutes and Archaea were on average 20-fold lower with the tested commercial kit as compared to the method with a mechanical lysis step that also efficiently disrupted the more difficult-to-lyse Gram positive and Archaeal cell walls. The best performing method involved mechanical lysis based on repeated bead-beating and thus this was utilized in the subsequent studies.

To identify alterations in the microbiota of patients, it is essential to characterize microbiota of healthy subjects. Hence, the intestinal microbiota of healthy adults and specifically its temporal stability were investigated. Moreover, the commonly shared, the core microbiota as well as the relation of intestinal symptoms and microbiota composition were addressed; first by performing a longitudinal study, which monitored a group of 15 healthy Finnish subjects for seven weeks and regularly assessed their intestinal bacteria. Additionally, the participants’ perception of health and the occurrence of intestinal symptoms were recorded with a questionnaire at each sampling point. The healthy microbiota was characterized as displaying high subject-specificity and temporal stability. Approximately 35% of the detected bacterial phylotypes were detected in all participants, emphasizing their contribution to the core microbiota. Moreover, significant correlations between the microbiota and mild intestinal symptoms were identified, including abdominal pain and bloating; these symptoms were correlated to the scarcity of bifidobacteria. This study provided the baseline data of the microbiota of healthy adults and thus can be used in further work aimed at benchmarking perturbations.

Adequate bowel cleansing such as used prior to colonoscopy or gastric surgery has been shown to be safe for patients. However, the long-term effects of the use of purgatives on the intestinal microbiota, and especially the potential differences arising from different dosing regimens have not been addressed previously. We found that in healthy adults the intestinal microbiota was generally resilient to purging with the tested osmotic laxative, and the majority of the bacterial levels that changed, returned to the subject’s baseline already two weeks after the lavage. However, the rate of recovery was dependent on the dosing of the purgative, since the consumption of a single dose of the purgative had a more severe and long-lasting effect on the microbiota composition than the split dose. Moreover, bacterial phylotypes belonging to Proteobacteria, a phylum that contains potential pathogens, were increased in the group consuming a single dose of the purgative. It was concluded that the use of two separate doses of the purgative resulted in fewer changes in the intestinal microbiota than a single dose and therefore should be preferred in the clinical practice.

There is growing evidence of the involvement of the intestinal microbiota in the pathophysiology of irritable bowel syndrome (IBS). However, the microbial component in post-infectious (PI) IBS patients has previously remained uncharacterised. The spectrum of symptoms
in PI-IBS is heterogeneous and in parallel to the characterization of intestinal microbiota of post-infectious patients and its comparison with other IBS patients, one aim of this work was to address the associations between the intestinal microbiota and a patient’s clinical characteristics. We identified a bacterial signature, consisting of 27 genus-like bacterial taxa that separated the healthy controls from PI-IBS and other IBS patients. The differences mainly consisted of increased levels of Bacteroidetes as well as uncultured members of the Clostridia that were decreased in the patients. The abundances of these bacteria correlated with clinical markers and expression levels of several host gene pathways, including amino acid synthesis, cell junction integrity and inflammatory response, all suggesting an impaired epithelial barrier function in IBS. Identification of these specific associations between the host and intestinal microbiota may provide novel insights into the origin and mechanistic background of intestinal symptoms in IBS as well as enables novel stratification of the IBS patient material with a different aetiology.

In conclusion, this thesis characterised the normal healthy intestinal microbiota and how perturbations such as IBS and bowel cleansing can disrupt the microbial composition. The results have clinical relevance by providing novel and a much needed tool for segregating the IBS patient material objectively as well as providing grounds for choosing the split-dosing regime of the purgative agent as an optimal bowel cleansing method. Furthermore, associations between the microbiota and health markers of the host were detected that will give grounds for future research on the aetiology of IBS. Finally, we determined and validated an optimal faecal DNA extraction method.
1. INTRODUCTION

The human gastrointestinal (GI) microbiota is a complex ecosystem harbouring trillions of organisms, mainly bacteria but also viruses and eukaryotic organisms. It is estimated that the intestinal microbiota contains $10^{14}$ bacterial cells, outnumbering the host cells by tenfold. Moreover, the genetic potential of the intestinal microbiome, the collective genome of the microbiota, is 100-fold greater than that of the human genome (Bäckhed et al., 2005; Huttenhower et al., 2012; Ley et al., 2006a; Qin et al., 2010). Unlike our own genome, the intestinal microbiome is not only vertically transmitted and fixed, but can be modified by early life events, diet and pharmaceutical treatments that affect the composition, stability and function of the gut ecosystem.

The co-evolution between the host and microbiota has resulted into a mutually beneficial relationship in which gut bacteria make essential contributions to human health e.g. by producing beneficial metabolites and vitamins and in return are allowed to occupy a nutrient-rich environment (Salonen and de Vos, 2014). The microbiota is also involved in the maturation of the immune system (Nylund et al., 2014) and in regulation of innate and adaptive immune responses (Maynard et al., 2012), for instance via initiation of specific T cell responses (Round et al., 2011). Although the mucosal barrier, which consists of the mucus layer and epithelial cells, prevents direct contact with the intestinal bacteria, there is constant cross-talk between the two systems allowing mutual benefits.

Additionally, there is growing evidence to indicate that the intestinal microbiota, and its alterations that are referred to as dysbiosis, may be an important factor in the pathogenesis of various diseases. Compositional changes in the microbiota have been associated with several conditions ranging from disorders limited to the GI tract, such as IBS, to systemic diseases, such as diabetes (de Vos and de Vos, 2012). In many cases, the implicated microbiological and host-microbe interaction processes are only beginning to be understood, making the research in this field pivotal in achieving a better understanding of human health.

It is important to characterize the microbial composition and function in a healthy host before understanding how the dysbiosis might affect pathogenesis of the related diseases. The aim of this doctoral thesis was to study what typifies a normal, healthy intestinal microbiota, and how the microbiota would be altered by a common medical procedure, bowel cleansing, as well as gastroenteritis-provoked IBS. In addition to the microbiological changes, the associations of the microbiota with subjective and objective host parameters, including occurrence of intestinal symptoms and clinical laboratory tests, were investigated.
2. **REVIEW OF THE LITERATURE**

2.1 **AN OVERVIEW OF THE MOLECULAR METHODS USED TO STUDY THE INTESTINAL MICROBIOTA**

Much of the current knowledge on the diversity of the intestinal microbiota has been dependent on the development of high-throughput technologies, most of which are based on the analysis of the small sub-unit ribosomal 16S rRNA gene, as a tool to classify the different microbial phylotypes. This gene is present in all bacteria and it contains both conserved and variable regions that allow taxonomic identification of bacterial species (Woese, 1987). To date, over 1000 different bacterial species residing in the human gastrointestinal tract have been characterised with cultivation (Rajilic-Stojanovic and de Vos, 2014). However, the majority of intestinal bacteria have been characterized only by their 16S rRNA gene sequence and therefore the function and relevance of these uncultured phylotypes remain unresolved. With the aid of high-throughput techniques, over 3000 gut-derived prokaryotes have been recognized and 9.8 million bacterial genes identified from the intestines of healthy adults, increasing our understanding of the bacterial diversity and its relation with health and disease (Li et al., 2014).

Achieving access to the genomic potential of the intestinal microbiota requires collecting of the bacteria and disrupting the cells. Faecal samples are the most widely used material, due to their easy and non-invasive access, while biopsies from the intestinal wall are occasionally sampled. The high microbial diversity in the faecal sample as well as the complexity and heterogeneity of the starting material makes the study of intestinal microbiota challenging. Although the major flaws in the pre-analytical steps typically originate from incomplete or biased DNA extraction, also other factors, such as storage conditions, can affect the microbial composition. The optimal preservation conditions for stool samples prior to DNA extraction is that they should be kept at room temperature for maximum of 24h prior to DNA extraction. Alternatively the samples could be placed immediately after collection at −20°C and stored at −80°C for longer periods of time (Cardona et al., 2012). Freezing of the faecal sample prior to the DNA extraction has been shown to affect the Firmicutes/Bacteroidetes ratio, in most cases significantly by lowering the Bacteroides proportion significantly (Bahl et al., 2012). However, extracting DNA from fresh samples is difficult to perform in large-scale studies, and most of the published work has been conducted from frozen samples.

2.1.1 **DNA EXTRACTION**

The majority of the molecular studies, which have attempted to characterize microbial communities, have been based on the analysis of DNA. Therefore, the quality of the work is largely dependent on the efficiency of the faecal DNA extraction protocol. Large species diversity poses major challenges to the DNA extraction methods, especially because cell wall structures and thus their propensity to lyse vary among the microbiota. The field is currently lacking any standard operating procedures for DNA extraction, and the use of several different methods complicates the comparison of individual studies. The protocols of all DNA extraction methods have two main objectives, disruption the bacterial cell wall followed by recovery of the DNA. The key points of both of the steps will be detailed below.

The two major bacterial phyla of the intestinal microbiota, Gram-positive Firmicutes and Gram-negative Bacteroidetes have very different cell wall structures, hence they differ in the optimal DNA extraction methods. The lysis of the Gram-positive bacteria requires a rigorous
extraction method, which poses the risk that this group will be underrepresented in the resultant community structure due to failure to lyse the bacterial cells. The Gram-negative bacteria lyse more easily due to their thinner peptidoglycan cell wall but if an excessively harsh extraction protocol is applied, this may result in degraded DNA and lower diversity and abundance of these bacteria. Such pronounced differences in the bacterial cell wall properties of the mixed community, and the challenges posed by the starting material (i.e. presence of inhibitors, variability of the matrix, heterogeneity of the sample) places limits on the ability to collect a representative sample of the faecal microbiota. Therefore, it is unlikely that one can develop a completely unbiased DNA extraction method. The majority of the utilized protocols rely on mechanical, enzymatic or chemical disruption of the cell wall, or a combination of these procedures. There have been several studies, including Study I in this thesis, which have attempted to compare the available faecal extraction methods. Their results clearly indicate that the combination of mechanical procedures, typically achieved by bead-beating, with either chemical or enzymatic lysis is the most effective way to acquire a holistic view of the microbiota (Claassen et al., 2013; Maukonen et al., 2012; Yuan et al., 2012). For example the numbers of bifidobacteria are shown to be reduced in DNA extracts lacking the mechanical lysis step (Claassen et al., 2013; Santiago et al., 2014b). Mechanical cell disruption can be achieved in several different ways, including sonication. Perhaps the most common protocol utilizes the combination of lysis buffer and small beads (typically 0.1 mm in diameter) that are added to the faecal material and shaken extensively in order to break the cell wall structures, although this can result in fragmentation of DNA. Therefore, when the methods used in research requires the extracted DNA to be intact or in as large fragments as possible (e.g. followed by %GC-profiling or large DNA libraries), vigorous mechanical lysis protocols are not suitable. In these cases, the preferred DNA extraction protocol is usually based on differential centrifugation and lysis (Apajalahti et al., 1998) or chemical and enzymatic disruption of the cells (Ahlroos and Tynkkynen, 2009). However, most of the current next generation sequencing (NGS) methods fragment the target DNA and therefore do not require the presence of large intact DNA molecules. When using chemical or enzymatic lysis, different organic solvents (e.g. methanol), detergents (e.g. sodium dodecyl sulfate, SDS) or enzymes (e.g. lysozyme) are used with the aim being to disrupt the cell walls of both Gram-negative and -positive bacteria. This approach results in less fragmented DNA, but due to the large variations in the properties of the outer cell surface structure, there is a greater risk of bias towards accessing only the easily breakable bacterial cells.

While the incomplete disruption of the cell wall of the community members can introduce a selective bias, the following steps of the DNA extraction protocols (removal of inhibitors, RNA and proteins as well as collection of DNA) are not thought to be biased towards any specific bacterial group. However, those procedures may affect the yield and quality of the collected DNA, which in turn can impact the success of the downstream applications, such as PCR. It has been shown that DNA extraction methods using phenol-chloroform purification and ethanol precipitation will harvest relatively more bacterial DNA than DNA extraction methods based on silica columns for DNA recovery (Study I, Morita et al., 2007).
Table 1: Overview of the analytical methods used to study the intestinal microbiota community.

<table>
<thead>
<tr>
<th>Technique</th>
<th>16S rDNA based</th>
<th>Sensitivity</th>
<th>Advantages</th>
<th>Limitations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Culturing</td>
<td>No</td>
<td>Moderate</td>
<td>Ability to gain functional information from the organism and to characterise novel species</td>
<td>Labour intensive, most of the intestinal microbiota cannot yet be cultured</td>
</tr>
<tr>
<td>FISH(^a)</td>
<td>Yes</td>
<td>Good</td>
<td>Targets bacterial taxa at both the species and group level. Enumeration of bacteria not dependent on 16S copy number</td>
<td>No novel taxa identified. Requires reference strains for validation. Laborious. Detection not at the community level</td>
</tr>
<tr>
<td>qPCR</td>
<td>Yes</td>
<td>Good</td>
<td>Detection of both specific or higher bacterial taxa</td>
<td>Labour intensive for larger validation, requires reference strains for the quantitative analysis. 16S copy number varies Detection not at the community level</td>
</tr>
<tr>
<td>DGGE(^b)</td>
<td>Yes</td>
<td>Poor</td>
<td>Provides a snapshot of the predominant bacterial species in an ecosystem, fast and inexpensive</td>
<td>Poor reproducibility and sensitivity</td>
</tr>
<tr>
<td>Phylogenetic microarrays</td>
<td>Yes</td>
<td>Good</td>
<td>Fast and deep detection of intestinal microbiota samples based on previously known sequences</td>
<td>No novel phylotypes detected</td>
</tr>
<tr>
<td>Next generation sequencing</td>
<td>Yes</td>
<td>Good</td>
<td>High-throughput, detects novel species</td>
<td>Requires extensive bioinformatic analysis, PCR chimera represents the most relevant bias</td>
</tr>
<tr>
<td>Metagenomics</td>
<td>Genomewide</td>
<td>Good</td>
<td>Ability to gain information on both the composition and genetic potential of the intestinal microbiota community. High-throughput, detects novel genomes</td>
<td>Requires extensive bioinformatic analysis. Annotation limited by the available reference genomes and hence limited phylogenetic information. Activity of genetic potential cannot be addressed</td>
</tr>
</tbody>
</table>

\(^a\)FISH = Fluorescent in situ hybridization  
\(^b\)DGGE = Denaturing gradient gel electrophoresis

2.1.2 MOLECULAR METHODS FOR ANALYSING THE INTESTINAL MICROBIOTA COMMUNITY

The methods used for the microbiota analysis depend on the research question and the project framework. Table 1 provides an overview of the most commonly used analytical techniques; high-throughput methods for the microbial community analysis, namely NGS and phylogenetic microarrays, will be discussed in more detail below.
PHYLOGENETIC MICROARRAYS

Microarrays were originally developed for the detection of differences in the expression of eukaryotic or prokaryotic genes. However, currently the technique is also implemented in the characterization of the microbial communities within the intestine (Brodie et al., 2006; Kang et al., 2010; Manges et al., 2010; Paliy et al., 2009; Palmer et al., 2007; Rajilic-Stojanovic et al., 2009; Tottey et al., 2013; Wang et al., 2004). The analysis pipeline includes a universal PCR targeting the conserved regions of 16S rRNA to achieve high primer coverage of all bacteria. As with all PCR based techniques, the choice of universal primers and the amplification bias will affect the detected phylotypes (Lee et al., 2012). It is now well established that there is no truly universal primers that would equally amplify all bacterial sequences. This limitation affects all techniques based on detection of PCR amplified 16S rRNA gene sequences, including NGS approaches (Frank et al., 2008; Huws et al., 2007). Although phylogenetic microarrays are semi-quantitative, and posses both a high analytical depth and high-throughput capability, the drawback of this technique is that no novel sequences can be detected. Additionally, the initial development of a phylogenetic microarray is technically demanding and requires extensive testing, optimization and validation. Nevertheless, the advantage of this method is its rapid capability to process samples and the bioinformatic requirements are considerably lower than encountered with the high-throughput NGS techniques. Moreover, the reproducibility of phylogenetic microarray protocol is high, allowing the identification of even small biological differences.

The phylogenetic microarray used in this thesis, the Human Intestinal Track Chip (HITChip), is based on oligonucleotide probes that have been designed from human gut-derived bacterial sequences gathered from the sequence databases (Rajilic-Stojanovic et al., 2009). The HITChip allows deep and accurate detection of up to 1140 species-level phylotypes on a highly reproducible manner (Study I, Rajilic-Stojanovic et al., 2009), allowing the detection of one of the larges collection of intestinal bacterial phylotypes. Only the recently developed HuGChip (Tottley et al., 2013) exceeds in the amount of detected phylotypes. Moreover, the method has been benchmarked and validated against pyro-sequencing (Claesson et al., 2009; van den Bogert et al., 2011), showing that the HITChip results can be replicated with other high-throughput methods. The data gathered from HITChip, so far over 6000 samples, is stored into a MySQL database, and represents a unique set of highly comparable data due to standardized analysis pipeline from the sample handling to analytical steps and data mining. This represents the largest existing data set of intestinal microbiota samples that have gone through the same preanalytical pipeline. The analysis of HITChip data is conducted with custom made scripts developed for this platform (Salojarvi and Lahti, 2014). To date, there has been 38 publications utilizing the HITChip microarray, including characterization of the microbiota in various disease states such as obesity (Verdam et al., 2013) and celiac disease (Cheng et al., 2013) as well as in different age groups such as preterm noenates (Moles et al., 2013), pre-school children (Ringel-Kulka et al., 2013), adults (Lahti et al., 2013) and elderly (Biagi et al., 2010). The large database collection provides an opportunity to unique meta-analysis of the intestinal microbiota, giving insight to its general features, such as the abundance distributions of different bacterial groups (Lahti et al., 2014) or the specifications of the core microbiota (Salonen et al., 2012).

NEXT GENERATION SEQUENCING

The NGS methods have dominated the microbiota research in recent years; the different platforms are detailed in Table 2 (NGS methods also reviewed in Liu et al., 2012; Quail et al.,
Review of the Literature

NGS based techniques can be either used to sequence the variable regions of choice of the 16S rRNA gene, or metagenomic fragments (shotgun sequencing). In compositional microbiota analysis, the sequencing is preceded with a PCR amplification of the 16S rRNA gene, which introduces the same PCR biases as discussed before. Only the PacBio sequencing platform can be utilized without any prior PCR amplification as it targets single molecules. The majority of the microbiota research utilising the NGS has been performed with the 454-platform. Large scale studies such as the US based Human Microbiome project (HMP), sampling 242 healthy adults at 15 or 18 body sites, have utilised this sequencing method (Methe et al., 2012). The other, European based large sequencing effort, metagenomics of the human intestinal tract (MetaHIT), utilised the Illumina Genome Analyser sequencing producing slightly shorter but many more sequencing reads than the 454, and provided the first large-scale metagenomic reference dataset of 3.3 M microbial genes (Qin et al., 2010). Today, NGS allows cost-effective and deep sampling of large cohorts, but the post-analytical burden, removal of chimeras, sequencing errors and sequence alignment, is higher than what is encountered for instance with phylogenetic microarrays. However, the price of sequencing has rapidly decreased due to the development of new techniques, meaning that the sequencing will become an even more favourable technique in the future. Several bioinformatic pipelines have been developed for analysing NGS reads of the 16S rRNA gene, including MOTHUR (Schloss et al., 2009) and QIIME (Caporaso et al., 2010), making the accumulating sequencing data more comparable when such pipelines are utilized.

Table 2: Overview of the next generation sequencing (NGS) methods.

<table>
<thead>
<tr>
<th>Platform</th>
<th>Read length</th>
<th>Reads per run</th>
<th>Advantages</th>
<th>Disadvantages</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sanger-sequencing</td>
<td>750 bp</td>
<td>-</td>
<td>Long reads, high accuracy</td>
<td>Slow, not high-throughput, requires biased clone libraries</td>
</tr>
<tr>
<td>454 (Roche)</td>
<td>700 bp</td>
<td>1 mil</td>
<td>Long reads allows good covering several variable regions of 16S rRNA in a single run</td>
<td>Chimeras. Runs are expensive. Homopolymer errors</td>
</tr>
<tr>
<td>MiSeq (Illumina)</td>
<td>200 bp</td>
<td>40 mil</td>
<td>High coverage and low cost</td>
<td>Short reads and chimeras</td>
</tr>
<tr>
<td>SOLiD (Life technologies Thermo)</td>
<td>75 + 35 bp</td>
<td>1.2 bil</td>
<td>Low cost per base</td>
<td>Short sequence fragments. Slower than other methods. Not good coverage of palindromic sequence</td>
</tr>
<tr>
<td>Ion-Torrent</td>
<td>400 bp</td>
<td>80 mil</td>
<td>Uses pH changes for detections, fast performance</td>
<td>Chimeras, Homopolymer errors</td>
</tr>
<tr>
<td>PacBio</td>
<td>1kb - 30kb</td>
<td>50 000</td>
<td>Single molecule sequence technology based on visualization of each of the fluorescence labelled nucleotide. No PCR required, High accuracy and long read lengths and fast</td>
<td>High sensitivity to contaminants and hence bias. Need for long DNA fragments. Moderate throughput. Equipment can be very expensive</td>
</tr>
</tbody>
</table>
2.2 GASTROINTESTINAL MICROBIOTA IN HEALTHY SUBJECTS

The 7-meter long GI tract varies greatly in terms of its microbial diversity and composition (Figure 1). The bacterial load increases and the oxygen levels decline towards the distal parts of the digestive tract, while the pH rises from acidic towards neutrality. The intestinal tract can be divided into the lumen, containing the indigested foods and the bulk of microbes, and the mucosal layers with its attached microbiota. Due to the large amount of bacteria inhabiting the human GI tract, the colonic epithelium is essential in the successful maintenance of the segregation between host and bacteria to avoid unnecessary immune responses and to guarantee homeostasis between the microbiota and the host (Belkaid and Hand, 2014). The gut microbiota is crucial for proper development of the immune system, for example the immune system must evolve adaptations that allow for the preservation of the beneficial relationship but protect the host from e.g. pathogens by compartmentalizing the bacteria (Hooper et al., 2012). One such method is the mucus layer separating the microbiota from epithelial cells. It has been shown that both mice (Johansson et al., 2008) and humans (Johansson et al., 2014) possess a mucus barrier consisting of two layers. These are the loosely adherent mucus layer closer to the lumen and the firmly adherent mucus layer protecting the epithelial cells. The inner mucus layer is impervious to microbes, preventing most of the contact between microbes and host. The inner mucus layer is constantly converted into the outer layer, which provides nutrients and attachment sites for several bacteria. The mucus layer acts as a physical barrier, together with tight junctions between the epithelial cells to prevent contact between intestinal microbiota and host cells.

The co-existence of the microbiota and host starts to develop at birth. The infant gut was previously thought to be sterile, however there is evidence that the infant receives his/hers initial inoculum from the mother before birth, since commensals have been found from both placenta (Satokari et al., 2009) and the meconium (Moles et al., 2013). Nevertheless, most of the microbiota development occurs in the postnatal phase with a major contribution from maternal microbes (Belkaid and Hand, 2014). Therefore it is thought that life events such as mode of delivery (Dominguez-Bello et al., 2010), feeding patterns and mothers’ specific microbiota composition have an effect on infants’ microbial signatures (Penders et al., 2006). The maturation of microbiota continues when solids are introduced to the infant’s diet (Palmer et al., 2007) and gradually develops all the way to adolescents (Agans et al., 2011; Ringel-Kulka et al., 2013).

The composition and the amount of bacteria in the intestine of a healthy adult vary along the digestive tract (Figure 1). The vast majority of the intestinal bacteria reside in the colon, reaching values between $10^{12}$ to $10^{13}$ bacteria per gram of faeces. The composition of the different locations of the intestinal tract varies due to different environmental factors, such as pH, transit time, mucus layer and peristaltic movements. Although the GI tract microbiota predominantly contains bacteria (94%), there are also viruses (4.8%) and archaea (0.5%) (Arumugam et al., 2011).

2.2.1 CHARACTERISTICS OF HEALTHY ADULT INTESTINAL MICROBIOTA

In order to study the intestinal microbiota of healthy subjects, a definition of what constitutes intestinal health and how to measure it must be determined. The World Health Organization specified in 1948, “Health is a state of complete physical, mental and social well-being and not merely the absence of disease or infirmity”. When related to GI health, this definition could be refined to include several other aspects, such as the absence of GI complaints, normal digestion and absorption of foods, normal immune status, normal intestinal microbiota and a good quality
of life (Bischoff, 2011). This multifaceted health status can be assessed with diagnostic measures covering both subjective perceptions of the GI health as well as objective parameters, both summarized in Table 3.

**Table 3: Assessment of the gastrointestinal health status.**

<table>
<thead>
<tr>
<th>Measure</th>
<th>Method</th>
<th>Description</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Subjective parameters</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>General health</td>
<td>Health related quality of life -questionnaires</td>
<td>Validation of the general health and well-being</td>
<td>Hays 1993, Barry 2007</td>
</tr>
<tr>
<td>Psychological status</td>
<td>Hospital anxiety and depression score (HAD)</td>
<td>Widely used assessing depression and anxiety measures in healthy controls and patients</td>
<td>Zigmond 1983</td>
</tr>
<tr>
<td>GI symptoms</td>
<td>IBS severity scoring</td>
<td>Assessing intestinal symptoms and their severity</td>
<td>Spiller 2010, Talley 1989</td>
</tr>
<tr>
<td>Stool consistency</td>
<td>Bristol stool scale</td>
<td>7-stage scale helps to diagnose subjects with constipation (type 1 and 2), normal stool (type 3 and 4) and diarrhoea (type 5 to 7)</td>
<td>Lewis 1997</td>
</tr>
<tr>
<td>Diet</td>
<td>Food frequency questionnaires</td>
<td>Evaluating the subjects’ habitual diet</td>
<td>Riboli 2002</td>
</tr>
<tr>
<td><strong>Objective parameters</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GI function</td>
<td>Transit rate</td>
<td>Hinton’s test based on ingestion of radioactive beads and the progress is monitored with X-rays</td>
<td>Hinton 1969</td>
</tr>
<tr>
<td>Digestion</td>
<td>Measuring nutrients such as fats or macronutrients from the stool or blood</td>
<td>Khouri 1989, Benzie 2014</td>
<td></td>
</tr>
<tr>
<td>Permeability</td>
<td>Evaluation of the integrity of intestinal barrier based on the movement of small indigestible particles from the intestine to urine</td>
<td>Teshima 2008</td>
<td></td>
</tr>
<tr>
<td>Motility</td>
<td>Measures of peristaltic movements, and gut wall tension</td>
<td>Whitehead 1997</td>
<td></td>
</tr>
<tr>
<td>pH</td>
<td>Stool pH reflects the microbial production rate of organic acids</td>
<td>Walker 2005</td>
<td></td>
</tr>
<tr>
<td>Immunity &amp; Inflammation</td>
<td>Cytokine measures</td>
<td>Measurement of the inflammatory cytokines (e.g. IL-6, IL-1 and TNF-a) or the anti-inflammatory cytokines (e.g. IL-10) or the expression of these genes as a means of evaluating inflammation</td>
<td>Swan 2013</td>
</tr>
<tr>
<td>Epithelial integrity</td>
<td>Visualisation of the epithelial layer e.g. with histological means</td>
<td>Martinez 2012, Johansson 2014</td>
<td></td>
</tr>
<tr>
<td>Calprotectin</td>
<td>Detection of neutrophilic protein, calprotectin, from faecal samples as an indication of intestinal inflammation</td>
<td>Roseth 1992</td>
<td></td>
</tr>
</tbody>
</table>
MICROBIAL COMPOSITION AND DIVERSITY

The bacterial composition, though demonstrating great variation from subject to subject, contains over 90% of Firmicutes and Bacteroidetes, additionally Actinobacteria, Proteobacteria, and Verrucomicrobia are among the most abundant phyla (Li et al., 2014; Qin et al., 2010). Healthy subjects are reported to harbour around 400-600 different species in their colon (Qin et al., 2010; Rajilic-Stojanovic et al., 2007; Salonen et al., 2012; Tap et al., 2009). The bacteria reside in the lumen or become attached to the mucosal layer of the gut. The definition of health and healthy intestine does not take into account the vast natural variation and subject-specificity of the intestinal microbiota. It has been shown that the relative abundance of both Firmicutes and Bacteroidetes can vary between 0 to 99% (Huttenhower et al., 2012) and the most abundant phylotypes can differ by as much as 5000-fold between subjects (Huse et al., 2008). To date the intestinal microbiota of over 1500 healthy subjects has been characterized (Huttenhower et al., 2012; Lahti et al., 2014; Qin et al., 2010) and no clear species collection specific for health has been identified. Due to the vast individual variation in composition, a detailed catalogue of a healthy microbial composition is not feasible. However, there are several components that can be used to distinguish the healthy intestinal microbiota; these will be discussed in the next sections.

The diversity of the intestinal microbiota is an ecological measure of the community structure consisting of species richness, the amount of different species, and species evenness, the relative abundance of those species. High diversity is thought to be beneficial in an ecosystem, allowing it to cope better with stressful conditions. High diversity has been associated to low temporal variation and was shown to be a subject-specific feature of the microbiota (Flores et al., 2014). Moreover, there are several reports indicating that decreased diversity can be associated with disease conditions such as IBS (Noor et al., 2010), inflammatory bowel disease (IBD) (Manichanh et al., 2006; Rajilic-Stojanovic et al., 2013) and obesity (Tap et al., 2009), also nutrition and alterations to the habitual diet are known to drastically affect the microbial diversity and composition (Salonen et al., 2014). The intestinal microbiota can rapidly adapt to the specific nutrient abundance, producing specific metabolites that generate different metabolic responses in the host (David et al., 2014). Together this indicates that high diversity could be beneficial to health and result to more stable microbiota.
SUBJECT SPECIFICITY

Each individual’s microbiota is most similar to their own samples (Costello et al., 2009; Turnbaugh et al., 2009; Zoetendal et al., 1998) followed by their close relatives and then by unrelated individuals (Tims et al., 2013). This high subject-specificity has been shown to be a combination of environmental factors shaping the microbiota immediately after birth as well due to genetic variation (Benson et al., 2010; Kashyap et al., 2013b; Wacklin et al., 2011). It appears that the functions performed by the microbiota and its composition do correlate with each other, indicating that there is redundancy in the functional properties of the intestinal microbiota that allow the observed individuality but retain the important functions (Huttenhower et al., 2012; Muegge et al., 2011). It has been shown that the minimal set of functions contains pathways necessary for bacterial homeostasis as well as gut specific pathways such as adhesions to the host proteins and harvesting energy in the gut environment (Qin et al., 2010).

Figure 1: Characteristics of the adult intestinal tract and microbiota (Lozupone et al., 2013; Walker and Lawley, 2013; Zoetendal et al., 2012)
TEMPORAL STABILITY AND RESILIENCE

From the ecological point of view, resilience represents a community’s ability to respond to perturbations by resisting change and the ability to recover to its original form (Holling, 1973). The capability for the microbiota of healthy subjects to maintain its composition in time has been associated to health in several studies. It has been shown that the microbiota retains its similarity over time, both the short and long term (Delgado et al., 2006; Krogius-Kurikka et al., 2009; Martinez et al., 2013; Maukonen et al., 2006; Scanlan et al., 2006; Tannock et al., 2000; Vanhoutte et al., 2004; Yatsunenko et al., 2012; Zoetendal et al., 1998). It is also known that there is little flux in the presence of the species but their abundances vary in time (Rajilic-Stojanovic et al., 2012) and that the temporal variation was associated to subjects’ microbial diversity (Flores et al., 2014). The majority of the current data supports the view that the intestinal microbiota of a healthy adult operates in a state of homeostasis and shows resilience to perturbations such as lifestyle changes, medication and diet.

CORE MICROBIOTA

Regardless of the high individuality, if the hallmarks of a healthy intestinal microbiota are temporal stability, high diversity and conserved key functions, it would be plausible to think that at least to a certain degree there would be similarities in the microbiota composition among healthy individuals. In a large European cohort, all healthy adults were found to harbour approximately 160 high-abundance species, several of these were shared between individuals, although in highly variable abundances (Qin et al., 2010). It would be tempting to assume that species detected in the majority of healthy subjects would have been conserved throughout evolution and would be beneficial for health. There are several studies showing that this kind of common microbial core can be detected in healthy individuals (Claesson et al., 2010; Martinez et al., 2013; Qin et al., 2010; Rajilic-Stojanovic et al., 2009; Salonen et al., 2012; Sekelja et al., 2011; Tap et al., 2009; Turnbaugh et al., 2009; Turnbaugh et al., 2007; Willing et al., 2010). The largest studies have estimated that the core microbiome contains approximately 30% of the detected phylotypes consisting mainly of Firmicutes (Ruminococcus, Eubacterium and Faecalibacterium spp.), Bacteroides spp. and Actinobacteria (Bifidobacterium spp.) (Qin et al., 2010; Salonen et al., 2012).

Another, more holistic approach for identifying conserved microbiota and similarities among healthy individuals was introduced by Arumugam and colleagues who revealed the existence of enterotypes (Arumugam et al., 2011). They hypothesised that the intestinal microbiota of healthy adults could be divided into three high-level ecosystem solutions based on the abundance of the so-called driver species. These included Bacteroides (Enterotype 1), Prevotella (Enterotype 2) or Ruminococcus (Enterotype 3). They have been shown to remain relatively stable over a period of at least 6 months (Roager et al., 2014) and there is evidence of fluctuations between enterotypes; these were revealed in a 10-year follow-up study (Rajilic-Stojanovic et al., 2012). Moreover the enterotypes appeared to be independent of gender and nationality, however long term dietary habits are known to be associated with enterotypes 1 and 2 (Wu et al., 2011).

2.3 EFFECTS OF PERTURBATION TO HEALTHY INTESTINAL MICROBIOTA

There is overwhelming evidence that changes to the intestinal microbiota can be associated with a variety of disease states (detailed in Table 4). In addition, perturbations such as consumption of
antibiotics are known to alter the intestinal microbiota of healthy individuals (Dethlefsen et al., 2008; Jernberg et al., 2007; Vrieze et al., 2013). The phylotypes affected by an antibiotic treatment tend to be subject-specific, but there is evidence that some phylotypes, such as members from the Bacteroidetes phylum, do not recover even several months after the treatment, and a long-term reduction in the diversity has also been observed (Dethlefsen and Relman, 2011; Jakobsson et al., 2010). It has been hypothesized that the increased use of antibiotics might have altered the microbial composition of Western people, and associations between obesity and use of antibiotics are emerging (Thuny et al., 2010).

This thesis concentrated on the evaluation of two different perturbations; the effect of bowel cleansing of the healthy intestine and the characterisation of the intestinal microbiota of IBS patients. The current literature of both cases will be discussed next.

Table 4: Disease states with associations to altered intestinal microbiota.

<table>
<thead>
<tr>
<th>Disease</th>
<th>Primary findings</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Celiac disease</td>
<td>Increased amounts of Firmicutes and Proteobacteria such as <em>Sutterella wadsworthensis</em> and lower proportions of Actinobacteria detected in children with risk for genetic predisposition for the disease</td>
<td>Cheng 2013, Olivares 2014</td>
</tr>
<tr>
<td><em>Clostridium difficile</em> infection</td>
<td>Decreased diversity, faecal material transplant from healthy individual cures patients</td>
<td>Fuentes 2014, van Nood 2013</td>
</tr>
<tr>
<td>Colorectal cancer</td>
<td>Reduced amounts of SCFA-producing Firmicutes, low-grade inflammation induced by the microbiota, increased <em>Fusobacteria</em></td>
<td>Kostic 2012, Plottel 2011, Zhu 2013</td>
</tr>
<tr>
<td>Crohn’s disease</td>
<td>Reduced diversity, temporal instability, decreased amount of Firmicutes such as <em>Faecalibacterium prausnitzii</em> and increased amounts of Proteobacteria</td>
<td>Loh 2012, Manichanh 2006, Sokol 2008</td>
</tr>
<tr>
<td>Necrotizing enterocolitis</td>
<td>Increased immune reaction in preterm infants induced by members of <em>Enterobacteriaceae</em> family</td>
<td>Wang 2009</td>
</tr>
<tr>
<td>Ulcerative colitis</td>
<td>Reduced diversity, decreased levels of Firmicutes and <em>Akermannia muciniphila</em></td>
<td>Lepage 2011, Loh 2012, Rajilic-Stojanovic 2013</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Disease</th>
<th>Findings</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arthritis</td>
<td>Increased levels of Proteobacteria and Bacteroidetes</td>
<td>Taneja, 2014</td>
</tr>
<tr>
<td>Autism</td>
<td>Increased levels of <em>Lachnospiraeae</em> spp. Ingestion of <em>Bacteroides fragilis</em> reduced symptoms in mice. Increased levels of <em>Sutterella</em> spp. and <em>Ruminococcus torques</em> in humans</td>
<td>Hsiao 2013, Wang 2013,</td>
</tr>
<tr>
<td>Depression</td>
<td><em>Oscillibacter</em> and <em>Alistipes</em> increased in patients, probiotics seem to alleviate symptoms</td>
<td>Naseribafroui 2014, Dinan 2013</td>
</tr>
<tr>
<td>Metabolic syndrome and type 2 diabetes</td>
<td>Lower diversity, lower level of SCFA-producing Firmicutes. Insulin sensitivity improved after faecal microbiota transplant</td>
<td>Karlsson 2013, Vrieze 2013</td>
</tr>
<tr>
<td>Obesity</td>
<td>Inconclusive results regarding the Firmicutes/Bacteroidetes ratio. Increased capacity for energy harvest and low-grade inflammation.</td>
<td>Ley 2006b, Moreno-Indias 2014, Turnbaugh 2009, Qin 2010</td>
</tr>
<tr>
<td>Type 1 Diabetes</td>
<td>Lower diversity, larger Bacteroidetes/Firmicutes ratio</td>
<td>Kriegel 2011, Murri 2013</td>
</tr>
</tbody>
</table>
2.3.1 EFFECT OF BOWEL CLEANSING TO THE INTESTINAL MICROBIOTA

Colonoscopy is an endoscopic procedure used for examining the large bowel and distal parts of the small bowel. It is intended for visual diagnosis of several gastroenterological diseases such as IBD, polyps and colon cancer. An essential preceding step for a successful colonoscopy is to remove all faecal particles and to reduce the bacterial load in the GI tract. This is achieved by administration of purgatives such as polyethylene glycol (PEG) supplemented with electrolytes. Consumption of osmotically active bowel cleansing agents increases the amount of water in the intestinal tract, washing out the luminal matter and reducing the colonic faecal material, including the intestinal bacteria, and introduces oxygen into the normally anaerobic colonic ecosystem (Strocchi et al., 1990). Moreover, the rapid increase in bowel movements will further flush out those bacteria incapable of adhering to the gut mucosa, distorting the microbiota composition compared to the normal state. All these rapid changes may affect the microbial ecosystem and impede its restoration.

Although consumption of purgatives has been shown to be safe for the patient, there is little consensus about how the intestinal microbiota is affected by the procedure, especially regarding the potential long-term consequences. It has been reported that the intestinal microbial composition is altered momentarily after the consumption of purgatives (Gorkiewicz et al., 2013; Harrell et al., 2012; Mai et al., 2006; O’Brien et al., 2013), but the observed changes have been shown to be relatively subject-specific, with little agreement on which bacterial taxa are affected. The main reason for the inconclusive evidence is likely to be the small number of study subjects or the lack of analytical depth in the previous studies. Surprisingly, there are several reports indicating that lavage does not alter the microbial diversity even when the total bacterial load was halved (Gorkiewicz et al., 2013; Harrell et al., 2012; O’Brien et al., 2013). Moreover, there are indications that the bowel cleansing might cause temporary changes in the mucus layer (Johansson et al., 2014) and reports of the microbial changes in the mucosal tissue after bowel preparation have shown a trend towards increased amounts of pathobionts from the Proteobacteria phylum (Gorkiewicz et al., 2013; Harrell et al., 2012).

2.3.2 DIFFERENCES IN THE INTESTINAL MICROBIOTA IN DISEASES - IRRITABLE BOWEL SYNDROME

Irritable bowel syndrome is one of the most common gastrointestinal disorders in the Western world, affecting approximately 10-15% of the population (Spiller et al., 2007). This patient group is very heterogeneous and it is characterized by abdominal pain and discomfort. IBS has not been associated with the development of serious diseases or higher mortality rate, with the exception of increased suicidal behaviour in patients suffering the chronic symptoms of abdominal pain (Spiegel et al., 2007). The patient groups can experience a variation of abnormal bowel habits and these are used to subdivide the patients into smaller groups of IBS with diarrhoea (IBS-D), IBS with constipation (IBS-C) and IBS with alternating bowel habits (IBS-A) (Longstreth et al., 2006). The above-mentioned symptoms can be common even in healthy subjects, but the IBS patients experience these chronically, reducing their quality of life. The diagnosis is based on symptom evaluation and elimination of more serious, organic diseases. The patient evaluation can be based on several different diagnostic criteria including the Manning (Manning et al., 1978) and the most frequently used Rome criteria with three versions, Rome I (Drossman et al., 1990) Rome II (Thompson et al., 1999) and Rome III (Longstreth et al., 2006). All distinguish the IBS patients in approximately 75% of the cases, however only 50% of patients with IBS meet all 4
symptom-based diagnostic criteria, emphasizing the need for developing more accurate ways of diagnosing IBS (Ford et al., 2013).

There is no single unifying cause reported for IBS (Camilleri, 2012). Several indicative factors have been identified, including genetic predisposition (Saito et al., 2010), prior episodes of anxiety and depression and female sex, which all have been shown to increase the risk of developing IBS (Koloski et al., 2012). The condition is heterogeneous in terms of its clinical presentation so it is tempting to hypothesize that there are several reasons for the underlying aetiology and pathophysiology. Moreover, there is strong evidence indicating that intestinal microbial composition could be one of the contributing factors in the aetiology of the disease. This is supported by the fact that the strongest predictor of onset of IBS is a prior gastroenteritic episode which will increase by 7-fold the risk for developing IBS (Rodriguez and Ruigomez, 1999).

MICROBIAL CHANGES IN IBS AND THEIR RELEVANCE TO THE SYMPTOM DEVELOPMENT

The microbial component in the aetiology of IBS has been studied extensively in the past decade and the development of molecular techniques in the microbiota analysis have provided evidence for the involvement of microbiota in the pathogenesis of IBS. Several studies have reported differences in diversity, both increase (Saulnier et al., 2011) and decrease (Carroll et al., 2011), and temporal stability of the intestinal microbiota (Durban et al., 2013; Kajander et al., 2007; Kajander et al., 2008; Mättö et al., 2005). Most of the studies (summarized in Table 5) show changes in the intestinal microbiota composition between the IBS patients and the healthy controls. However, the reported changes are inconclusive or even contradictory and show little overlap from study to study. This is likely due to both lack of standard operation procedures in the data collection and analysis of the intestinal microbiota as well as in the patient recruitment. The observed differences may derive from heterogeneous patient material due to difficulties in the diagnosis of IBS patients, the high individuality of the faecal microbiota, and therefore inadequate sample sizes. Regardless of the lack of coherent changes in the compositional analysis, the evidence for the role of microbiota in the pathogenesis of IBS is supported by the following arguments. Firstly, the IBS symptoms can be improved with several different agents targeting the intestinal microbiota, including some probiotics and prebiotics (Whelan, 2011) as well as antibiotics such as rifaximin (Menees et al., 2012). Secondly, though there is a lack of clear microbial signature, the microbial composition of the IBS patients is shown to be different from that of the healthy controls (detailed in Table 5). Thirdly, about 10% of the IBS cases are reported to have began after an episode of gastroenteritis, causing post-infectious IBS (PI-IBS), indicating a cause and effect relationship (Chaudhary and Truelove, 1962; Spiller and Garsed, 2009). Several different pathogens causing the gastroenteritis prior the onset of PI-IBS have been identified, including Campylobacter jejuni (Spiller et al., 2000), Escherichia coli (Okhuysen et al., 2004) and Salmonella (Mearin et al., 2005). In addition, there are studies showing a potential genetic predisposition for development of the PI-IBS, showing alterations in genes encoding for proteins involved in epithelial cell barrier function and the innate immune response to enteric bacteria (Swan et al., 2013; Villani et al., 2010).

The presence of low-grade inflammation or immune activation is well documented in IBS patients but the cause is not known (Chadwick et al., 2002; Spiller, 2004). Furthermore, it has been speculated that the altered microbial composition could be a driver for these changes. IBS patients with alternating bowel symptoms (IBS-A) have been demonstrated to have increased

Review of the Literature
expression of Toll-Like Receptors (TLRs) (Belmonte et al., 2012; Brint et al., 2011). TLRs are structures recognising e.g. the outer cell wall structures of the intestinal microbiota and other structurally conserved regions of bacteria, highlighting the potential of the microbiota in the development of low-grade inflammation. In other studies, mast cells and lymphocytes have been associated with the symptom development, although the reasons for their increased numbers are not fully understood (Spiller et al., 2000). There have been speculations about the role of intestinal microbiota, since rats with a dysbiosis due to antibiotic treatment display increased levels of mast cells (Nutten et al., 2007). Moreover, the altered intestinal microbiota has been associated with host immune markers, including the increased numbers of mast cells in the patients (Study IV). Furthermore, the increased expression of TLR2 and TLR4 have been associated with IBS subtypes providing further support for the hypothesis that there is altered intestinal immune activation and a role for intestinal microbiota in the aetiology of IBS (Belmonte et al., 2012).

Closely related to the development of low-grade inflammation is the maintenance of the mucosal barrier separating the microbial content from the intestinal epithelium. The integrity of the epithelial cell barrier has been linked to the development of IBS where the tight junction proteins, such as ZO-1, were shown to have decreased expression and this phenomena has also been demonstrated in intestinal tissue cross sections (Martinez et al., 2012). Although the microbial composition of these subjects was not determined, if there was leaky epithelium there would be an outflow of antigens causing stimulation of the mucosal immune system and low-grade inflammation, irrespective of the actual microbial composition.

The altered abundance of methanogenic archaea has been associated with both IBS-D and IBS-C subtypes. Organisms in this group recycles intestinal hydrogen into methane and the levels of this organic material have been shown to be increased in the constipation predominant patients (Menees et al., 2012; Chatterjee et al., 2007; Furnari et al., 2012; Pimentel et al., 2003) and reduced in the diarrhoea-predominant patients (Study IV; Rajilic-Stojanovic et al., 2011). Methane is known to slow down the intestinal transit and the levels of methane excretion have been positively associated with constipation in IBS patients. This suggests that the methanogens might play a role in the symptom development in the IBS-C patients. However, since the organisms grow slowly and are sensitive to increased transit time (Pimentel et al., 2006) their role in constipation might be secondary rather than causative.

One of the most consistently IBS associated bacterial phylotypes is an uncultured relative of *Ruminococcus torques* (Kassinen et al., 2007). This species has been shown to be more abundant in several IBS cohorts in comparison to controls (Kassinen et al., 2007; Lyra et al., 2010; Rajilic-Stojanovic et al., 2011; Saulnier et al., 2011). This, still uncultured, phylotype has been associated with an increased sensation of pain (Malinen et al., 2010) and its numbers were reduced when IBS patients received probiotic supplements (Lyra et al., 2010). The closest relative of this phylotype (*Lachnospiraceae bacterium A4*) is known to possess the pro-inflammatory flagellin proteins, which has been associated with IBS (Schoepfer et al., 2008). This bacteria has the capability to degrade mucus (Hoskins et al., 1985) and its abundance is believed to be increased in Crohn’s disease (Prindiville et al., 2004) and Ulcerative colitis (Png et al., 2010).
### Table 5: Summary of studies characterising the intestinal microbiota in IBS patients.

<table>
<thead>
<tr>
<th>Microbial result (IBS vs. HC)</th>
<th>Study size (IBS/HC)</th>
<th>Diagnostic criteria</th>
<th>Sample source</th>
<th>Method</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>IBS patients clustered to both IBS-like microbiota and HC-like microbiota</td>
<td><strong>Increased:</strong> Firmicutes (<em>Dorea</em>, <em>Blautia</em>), Actinobacteria (<em>Bifidobacteria</em>)</td>
<td><strong>Decreased:</strong> Bacteroidetes</td>
<td>37/20</td>
<td>Rome II (D=15, C=10, A=12)</td>
<td>Pyrosequencing (16S rRNA V4) – 30000 reads/subject</td>
</tr>
<tr>
<td><strong>Increased:</strong> Enterobacteriaceae, Fusobacterium, Pseudomonadaceae, Lactobacillaceae</td>
<td><strong>Decreased:</strong> stability, richness, <em>Fusobacterium</em></td>
<td>23/23</td>
<td>Rome III (D=23)</td>
<td>F</td>
<td>qPCR, Pyrosequencing (16S rRNA V1-3 and V6) – 7400 reads/subject</td>
</tr>
<tr>
<td>Smaller core microbiota in IBS children,</td>
<td><strong>Increased:</strong> β- and δ-Proteobacteria, <em>Megasphaera</em>, <em>Parasporobacterium</em>, <em>B. thetaiotaomicron</em>, <em>B. ovatus</em>,</td>
<td><strong>Decreased:</strong> <em>Dehalobacter, Fusobacter, Oxobacter, Bifidobacterium</em></td>
<td>22/22</td>
<td>Rome II (D=22)</td>
<td>F</td>
</tr>
<tr>
<td>General effects in IBS</td>
<td><strong>Increased:</strong> <em>Papillibacter</em>, relatives of <em>C. orbiscindens</em>, <em>Blautia</em>, Peptococcus, <em>C. difficile</em>, <em>C. neulate</em>, <em>C. symbiontium</em>, <em>C. eutactus</em>, <em>Lachnospira</em>, <em>R. gnarus</em> and <em>Dorea</em></td>
<td><strong>Decreased:</strong> <em>bifidobacteria, Bacteroidetes, Uncultured Clostridiales</em></td>
<td>62/46</td>
<td>Rome II (D=22, C=18, A=19)</td>
<td>F</td>
</tr>
<tr>
<td></td>
<td><strong>Increased:</strong> Diversity, δ-Proteobacteria, <em>Dorea, Veillonella, Bacteroides, Anaerovorax, Ruminococcus</em> like phylotype</td>
<td><strong>Decreased:</strong> <em>Eubacterium, Bacteroides</em></td>
<td>22/22</td>
<td>Rome III (D=1, C=13, U=8)</td>
<td>F</td>
</tr>
<tr>
<td></td>
<td><strong>Increased:</strong> <em>Staphylococcus aureus</em></td>
<td><strong>Decreased:</strong> Diversity (F)</td>
<td>96/23</td>
<td>Rome I and II (D or A=81, C=15)</td>
<td>F</td>
</tr>
<tr>
<td></td>
<td><strong>Increased:</strong> <em>Pseudomonas aeruginosa</em></td>
<td><strong>Decreased:</strong> Diversity (F)</td>
<td>37/20</td>
<td>Rome II (D=13, C=11, A=13)</td>
<td>F, M</td>
</tr>
<tr>
<td></td>
<td><strong>Decreased:</strong> Diversity (F)</td>
<td>16/21</td>
<td>Rome III (D=16)</td>
<td>F, M</td>
<td>T-RLFP</td>
</tr>
<tr>
<td></td>
<td><strong>Increased:</strong> Diversity, total bacterial load, <em>Bacteroidetes, Lactobacillus</em></td>
<td><strong>Decreased:</strong> <em>Clostridium cocooides</em> et rel., <em>Bifidobacterium</em></td>
<td>11/8</td>
<td>Rome II (U=8)</td>
<td>F</td>
</tr>
<tr>
<td></td>
<td><strong>Increased:</strong> <em>Lactobacillus, Veillonella</em></td>
<td><strong>Decreased:</strong> Diversity</td>
<td>26/26</td>
<td>Rome II (C=11, D=8, A=7)</td>
<td>F</td>
</tr>
<tr>
<td></td>
<td><strong>Increased:</strong> <em>Bacteroides</em></td>
<td><strong>Decreased:</strong> Diversity</td>
<td>11/22</td>
<td>Rome II (U=11)</td>
<td>F</td>
</tr>
</tbody>
</table>
### Table 5 cont.

<table>
<thead>
<tr>
<th>Microbial result (IBS vs. HC)</th>
<th>Study size (IBS/HC)</th>
<th>Diagnostic criteria</th>
<th>Sample source</th>
<th>Method</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Increased:</strong> aerobic bacteria, <em>Lactobacillus</em> (F)</td>
<td>10/10</td>
<td>F, M</td>
<td>culturing, qPCR</td>
<td>Carroll 2010</td>
<td></td>
</tr>
<tr>
<td><strong>Increased:</strong> Proteobacteria, <em>Lachnospiraceae</em></td>
<td>10/22 Rome II (D=10)</td>
<td>F</td>
<td>%G+C, sanger sequencing, qPCR</td>
<td>Krogius-Kurikka 2009</td>
<td></td>
</tr>
<tr>
<td><strong>Decreased:</strong> Actinobacteria, <em>Bacteroidetes</em></td>
<td>24/23 Rome II (D=10, C=8, M=6)</td>
<td>F</td>
<td>%G+C, sanger sequencing, qPCR</td>
<td>Kassinen 2007</td>
<td></td>
</tr>
<tr>
<td><strong>Increased:</strong> <em>R. torques</em> 94%</td>
<td>24/23 Rome II (D=10, C=8, M=6)</td>
<td>F</td>
<td>%G+C, sanger sequencing, qPCR</td>
<td>Kassinen 2007</td>
<td></td>
</tr>
<tr>
<td><strong>Decreased:</strong> <em>C. aerofaciens</em>, <em>C. eutactus</em> 97%, <em>C. cocleatum</em> 88%, <em>B. catenulatum</em></td>
<td>24/23 Rome II (D=10, C=8, M=6)</td>
<td>F</td>
<td>%G+C, sanger sequencing, qPCR</td>
<td>Kassinen 2007</td>
<td></td>
</tr>
</tbody>
</table>

D = diarrhoea-predominant IBS  
C = constipation-predominant IBS  
A = alternating-type IBS  
U = unclassified  
F = Faecal sample  
M = mucosal sample
3. AIMS OF THE STUDY

The general aim of the thesis was to characterise the intestinal microbiota of healthy subjects and to benchmark it to conditions with expected microbiota aberrations due to PI-IBS or bowel cleansing. The microbial profiles were correlated to several clinical measurements to identify interrelations between the microbiota and host health. The specific aims of the studies were:

I) Comparison of the efficacy of widely used faecal DNA extraction methods and the determination of the biases that they introduce to the microbial community in order to select the optimal method to be used in subsequent studies.

II) Characterization of the normal healthy adult microbiota and determination of the extent of temporal changes as well as determining the composition and stability of the microbiota shared by these individuals (the core microbiota).

III) Determination of the effects of bowel cleansing procedure on the healthy intestinal microbiota and the effects of different doses of the purgative agent.

IV) Comparison of the microbiota composition between healthy individuals and post-infectious irritable bowel syndrome patients with an attempt to facilitate the future use of microbiota for stratification of the patient material.
4. MATERIALS AND METHODS

4.1 STUDY SUBJECTS

All subjects taking part in the four studies provided written informed consent. The samples used in the studies I and II were part of a clinical trial (Kekkonen et al., 2008) and the ethics committee of the Hospital District of Helsinki and Uusimaa approved the study protocols. The Nottingham Research Ethics Committee approved studies III and IV, the latter was a subset of a larger clinical trial (Swan et al., 2013). A summary of the study subjects is provided in Table 6.

Unpublished data on intestinal symptoms of healthy subjects were collected from 46 Finnish adults taking part in a science outreach “Tiedekulma”, organized in 2012.

Table 6: The study subjects.

<table>
<thead>
<tr>
<th>Study</th>
<th>Number of subjects</th>
<th>Clinical status</th>
<th>Sex (f/m)</th>
<th>Mean age (range)</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>5 + 19</td>
<td>Healthy</td>
<td>2/3 + 14/5</td>
<td>42 (34-54)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>+ 42 (25-56)</td>
<td></td>
</tr>
<tr>
<td>II</td>
<td>15</td>
<td>Healthy *</td>
<td>10/5</td>
<td>42 (28-56)</td>
</tr>
<tr>
<td>III</td>
<td>23</td>
<td>Healthy</td>
<td>12/11</td>
<td>26 (19 - 50)</td>
</tr>
<tr>
<td>IV</td>
<td>57</td>
<td>IBS-D: 12,</td>
<td>36/21</td>
<td>42 (20 - 68)</td>
</tr>
<tr>
<td></td>
<td>PI-IBS: 11,</td>
<td>PI-BD*: 11,</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>PI-nonBD*: 12,</td>
<td>Healthy: 11</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thesis</td>
<td>54</td>
<td>Healthy</td>
<td>34/20</td>
<td>34 (25 - 62)</td>
</tr>
</tbody>
</table>

* During the trial, divided into subgroups: 5 compromised and 9 healthy, see study II Table 1

a PI-BD = post C. jejuni infection bowel dysfunction

b PI-nonBD = post C. jejuni infection with out bowel dysfunction

4.2 ANALYSIS PROTOCOL

All of the studies followed a similar analysis pipeline, starting from faecal DNA extraction with the DNA extraction method validated in the Study I, combining the mechanical lysis by repeated bead beating to chemical lysis of cells, and purification of DNA with the QIAamp DNA Mini Kit. The resulting DNA was amplified with universal primers targeting the hypervariable V1 and V6 regions of 16S rRNA gene (Rajilic-Stojanovic et al., 2009). The PCR product was transcribed into RNA, labelled with Cy3 or Cy5 followed by fragmentation and hybridization on the microarray. After scanning the arrays, the raw data was normalized not only to reduce outliers and technical biases within the samples but also to normalize the signals between the arrays. The phylogenetic microarray analysis was complemented with qPCR analysis (detailed in Table 7).

The bioinformatic analysis of the samples utilised the analytical methods summarized in Table 7, concentrating on finding differences in the microbial community between treatments or patient groups, or finding similarities between the microbiota of healthy subjects. The microarray and qPCR data were also correlated with different host parameters in order to evaluate their co-variance (studies II, III, IV).
### Table 7: Methods used in the thesis.

<table>
<thead>
<tr>
<th>Method</th>
<th>Study</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Sample preparation</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Faecal DNA extraction</td>
<td>I, II, III, IV</td>
<td>Study I</td>
</tr>
<tr>
<td>DNA quantification</td>
<td>I, II, III, IV</td>
<td></td>
</tr>
<tr>
<td><strong>HITChip sample-wise normalization methods</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Quantile between sample normalization</td>
<td>I, IV</td>
<td>Bolstad 2003</td>
</tr>
<tr>
<td>Assumption of normal-exponential (background-signal) distribution and quantile normalization</td>
<td>II</td>
<td>Bolstad 2003</td>
</tr>
<tr>
<td>Robust Probabilistic Averaging</td>
<td>III</td>
<td>Lahti 2011</td>
</tr>
<tr>
<td><strong>Quantitative PCR from microbiota</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Methanogenic archaea (<em>Methanobrevibacter</em> spp.)</td>
<td>I, II, III, IV</td>
<td>Ufnar 2006</td>
</tr>
<tr>
<td>Total bacteria</td>
<td>I, II, III</td>
<td>Nadkarni 2002</td>
</tr>
<tr>
<td><em>Clostridium cocoides-Eubacterium rectale</em>-group</td>
<td>I</td>
<td>Rinttila 2004</td>
</tr>
<tr>
<td><em>Bacteroides-Prevotella-Porphyromonas</em></td>
<td>I, IV</td>
<td>Rinttila 2004</td>
</tr>
<tr>
<td><em>Bifidobacterium</em> genus</td>
<td>I, II</td>
<td>Rinttila 2004</td>
</tr>
<tr>
<td><em>Faecalibacterium praunznitzii</em></td>
<td>II, III</td>
<td>Rinttila 2004</td>
</tr>
<tr>
<td><em>Ruminococcus torques</em>-like phylotype (94% similarity)</td>
<td>IV</td>
<td>Kassinen 2007</td>
</tr>
<tr>
<td><em>Collinsella aerofaciens</em>-like taxa</td>
<td>IV</td>
<td>Kassinen 2007</td>
</tr>
<tr>
<td><strong>Statistical analysis</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Linear mixed models</td>
<td>I, II, III, IV</td>
<td></td>
</tr>
</tbody>
</table>
### Materials and Methods

Table 7 cont.

<table>
<thead>
<tr>
<th>Method</th>
<th>Study</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Statistical analysis</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(Tukeys Honest significant differences,</td>
<td></td>
<td></td>
</tr>
<tr>
<td>q-value estimation or Benjamini-Hochberg</td>
<td></td>
<td></td>
</tr>
<tr>
<td>false discovery rate)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Correlations between data sets</td>
<td>II, III, IV</td>
<td></td>
</tr>
<tr>
<td>(Spearman or Pearson)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hierarchical clustering</td>
<td>I, II, III</td>
<td></td>
</tr>
<tr>
<td>Principal component analysis</td>
<td>I</td>
<td>Venables and Ripley, 2002</td>
</tr>
<tr>
<td>Coefficient of variation</td>
<td>II</td>
<td></td>
</tr>
<tr>
<td>Species enrichment with Fisher’s exact test</td>
<td>II</td>
<td>Rivals 2007</td>
</tr>
<tr>
<td>Redundancy analysis</td>
<td>III</td>
<td></td>
</tr>
<tr>
<td>Gene set analysis</td>
<td>IV</td>
<td>Efron 2007, Subramanian 2005</td>
</tr>
<tr>
<td>Bootstrap aggregated redundancy analysis</td>
<td>IV</td>
<td>Study IV</td>
</tr>
<tr>
<td><strong>Other analysed data</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Clinical data and laboratory test of patients</td>
<td>III, IV</td>
<td></td>
</tr>
<tr>
<td>Quantitative PCR from the host</td>
<td>I, III, IV</td>
<td>Swan 2013</td>
</tr>
<tr>
<td>Questionnaires regarding GI symptoms</td>
<td>II, III, IV,</td>
<td>Hays 1993, Melzack, 1975, Spiller 2010,</td>
</tr>
<tr>
<td></td>
<td>science outreach (unpublished)</td>
<td>Study II</td>
</tr>
<tr>
<td>Questionnaires regarding mental status</td>
<td>III, IV</td>
<td>Zigmond and Snaith, 1983</td>
</tr>
</tbody>
</table>
5. RESULTS AND DISCUSSION

5.1 COMPARISON OF FAECAL DNA EXTRACTION METHODS (STUDY I)

When working with DNA-based molecular methods using complex study material, such as faeces, it is essential to use an unbiased DNA extraction method. Although the recovery of bacterial DNA from a faecal sample is a crucial step in obtaining a representative view of the intestinal microbiota, there is currently no consensus about what represents a standard operating procedure. The aim of this work was to examine the effects of different faecal DNA extraction protocols on the composition and diversity of the recovered microbial DNA. The goal was to validate a standard DNA extraction protocol, which could be used in the subsequent studies.

The evaluated DNA extraction methods represented different principles for cell lysis and DNA purification. The first technique enriched the bacteria from faecal solids by differential centrifugation, followed by enzymatic, chemical and physical cell lysis, hereafter referred as to the Differential Centrifugation and Lysis method (DCL, Apajalahti et al., 1998). The second method was a modified version of the Promega Genomic Wizard DNA Purification kit (MPR), relying primarily on enzymatic cell lysis (Ahlroos and Tynkkynen, 2009). The third method, subsequently referred to as repeated bead-beating (RBB), was initially develop for extracting DNA from bovine ruminal and faecal samples, which we optimised for human samples. This method employs physical disruption of bacteria, followed by purification of DNA first by precipitation and then through a column (Yu and Morrison, 2004). The fourth method was based on the commercial QIAamp® DNA Stool Mini Kit (QSK) preceded with brief mechanical cell lysis (Zoetendal et al., 2006).

5.1.1 DIFFERENCES IN THE MICROBIOTA PROFILES BETWEEN THE DNA EXTRACTION METHODS

The efficiency of the DNA extraction protocols can be compared by assessing the quantity and quality of the extracts. We quantified the amounts of total bacteria and methanogenic archaea with qPCR and found up 35-fold differences in the same sample, depending on the method. In general, the MPR yielded the highest amount of bacterial DNA and the QSK the lowest. In addition, there were statistically significant differences in the amount of archaeal DNA. Again the QSK detected the lowest amounts of 16S rRNA copies, whereas the DLC and RBB methods yielded over 20-fold more archaeal DNA than the QSK method (p<.05). Moreover, we documented higher archaeal prevalence when applying mechanical lysis, indicating that the efficiency of DNA extraction method affects the detection limit for hard-to-lyse organisms.

To further clarify the potential sources of the differences in the bacterial community structure attributing to the DNA extraction protocols, we analysed the samples with a phylogenetic microarray and the results were confirmed by qPCR analysis targeting specific bacterial groups. The individuality of subject’s microbiota contributed most to the variation between the samples, as all methods and their technical replicates clustered by individual. When comparing the microbial profiles within a sample it was evident that the microbial profile extracted with the QSK method differed the most from those extracted with the other methods (Figure 2). The compositional analysis of the different extracts showed that all methods recovered all the predominant bacterial taxa, although the QSK-extracted DNA was enriched with Bacteroidetes in contrast to the other methods that retrieved relatively more Firmicutes. Additionally, the QSK method extracted less Actinobacteria as compared to the other DNA extraction protocols. A
recent study supported our results and indicated that the detected microbial composition was affected by DNA extraction, especially with the use of mechanical lysis. Similar to the present study, in the samples that were not subjected to extensive mechanical lysis, various Gram-positive bacteria, especially bifidobacteria and Blautia were severely underrepresented (Santiago et al., 2014a). These results confirm and extend the earlier observations from Zoetendal and colleagues that increasing the mechanical lysis time will result in the appearance of new Gram-positive bacterial species (Zoetendal et al., 1998).

Figure 2: Comparison of the microbial profiles obtained from two subjects with four DNA extraction methods. A) Hierarchical clustering visualizing the similarity of the phylogenetic fingerprints. a) and b) depict technical replicates. The abundance of B) Actinobacteria and C) Bacteroidetes phylum, as measured with microarray signal intensities.

The observed differences in the microbial composition are likely to be explained by the inefficient mechanical cell breaking capabilities in those methods lacking the mechanical lysis step. Gram-negative bacteria lyse more easily due to their thinner cell wall and hence these species become overrepresented in the mixed community if the more recalcitrant Gram-positive bacteria are not lysed by the applied extraction procedure. This highlights the importance of employing mechanical cell lysis in order to obtain a representative sample of the intestinal microbiota. However, another factor to consider is the conditions at cell lysis since when DNA degradation is not prevented rapidly lysing cells liberate DNA, which will be rapidly degraded.

Due to the above-mentioned biases introduced by the different DNA extraction protocols, studies with dissimilar DNA extraction methods are not well comparable. In a recent meta-analysis microbiota samples tended to cluster together study-wise rather than according to the ethnicity or health status suggesting that there are technical biases in the results of different studies (Lozupone et al., 2013). Although the DNA extraction procedure is only one of several potential sources of biases, it has been shown here and by others that it can influence the results profoundly (Claassen et al., 2013; Kennedy et al., 2014). However, further large-scale studies comparing the most commonly used extraction methods in sufficiently sized cohorts need to be conducted in order to achieve standard operating procedures for faecal DNA extraction methodology. This would allow better comparison of the data across different studies and yield better quality results.

In conclusion, we have compared four conceptually different faecal DNA extraction methods of which two, RBB and MPR, performed highly reproducibly and achieved a good representation of the intestinal microbiota, as reflected in high DNA yield, high diversity and a
Results and Discussion

ratio of hard-to-lyse Gram-negative bacteria and archaea, and microbial composition comparable to that described in the literature (Qin et al., 2010, Turnbaugh et al., 2009). The major difference between these two extraction methods was the higher microbial diversity of RBB extracts and superiority in extracting the methanogens, meaning that the RBB was the optimal DNA extraction method for the following studies.

Study I was one of the pioneering studies in validation of different faecal DNA extraction protocols based on untargeted, community-wide microbiota analysis. Since the publication of our study, efforts with larger cohorts and further methods have been made, taking steps towards the standardisation of DNA extraction protocols. The European Union-funded project International Human Microbiome Standards (IHMS) is aiming to for this. In this international collaboration, aliquots of pooled faecal samples were extracted with different DNA extraction protocols (including the RBB methods described in Study I) and the extracted DNA was subsequently analysed by metagenomic sequencing analysis. The RBB method was found to meet the all quality criteria specified in the first IHMS progress report (IHMS, 2012) and was ranked among the top three of the DNA extraction protocols tested within the project (Joel Dore and Anne Salonen, personal communication).

5.2 CHARACTERIZATION OF THE HEALTHY MICROBIOTA (STUDY II)

Despite the overwhelming accumulation of data of the intestinal microbiota, a definition of the healthy microbiota is still lacking. The aim of Study II was to expand current knowledge by characterising the normal temporal variation of the healthy adult microbiota, its composition, diversity and stability as well as the common core microbiota, shared among most or all studied individuals. In addition, our objective was to identify potential associations between common and mild intestinal complaints and the microbial composition. To achieve this, we studied the intestinal microbiota of 15 Finnish adults, by collecting six faecal samples over a 7-week time period. The subjects evaluated their health status with a Health Related Quality of Life (HRQoL)-questionnaire completed at the time of each faecal sampling. The participants were divided into groups of healthy (n=9) and compromised (n=6) subjects based on the results of the HRQoL-questionnaire. Subjects without significant deviations from Finnish reference values (Aalto, 1999) and any self-reported life style changes that might effect the microbial composition were considered healthy. The compromised group consisted of subjects who reported significant complaints of GI symptoms (n=2), antibiotic use (n=1), or who had travelled outside Finland (n=3).

5.2.1 HEALTHY MICROBIOTA COMPOSITION AND TEMPORAL VARIATION

The compositional analysis of the participants’ microbiota detected in all healthy subjects studied in this thesis (193 samples from 54 subjects) revealed on average 410 species-level bacterial phylotypes per sample. The most abundant phyla were Firmicutes (80%) followed by Bacteroidetes (10%) and Actinobacteria (1.5%). Members of the Proteobacteria and Verrucomicrobiota represented less than 1% of the total community. However, there was great differences between the subjects, as can be seen from Figure 3A; for example the relative abundance of Bacteroidetes varied from less than 1% to over 75% between the healthy individuals. The phylogenetic microbiota composition detected in these subjects was similar to that of healthy adults reported in previous publications (Qin et al., 2010) as well as studies appearing afterwards (Huttenhower
et al., 2012; Schnorr et al., 2014; Yatsunenko et al., 2012). When examining all of the healthy subjects of this thesis, the diversity showed large variation among participants. Moreover, when compared to the IBS patients there was a significant decline ($p<.05$) in the Shannon diversity index as compared to all of the healthy controls evaluated here (Figure 3B).

Results and Discussion

![Figure 3: Microbial composition of healthy subjects. A) Composition of the healthy subjects studied in the four sub-studies of the thesis (N=193 from 54 subjects) ordered by the relative abundance of Bacteroidetes. B) Diversity of all healthy subjects (N=54) studied in the thesis compared with the IBS patients (N= 24) from the Study IV.](image)

Study II was set up to follow the temporal variation of the intestinal microbiota, which showed only a slight and non-significant decline within the seven-week study period. This indicated that the healthy subjects’ microbiota was intrinsically stable, a view that has been supported also in later studies (Stein et al., 2013). A novel finding was that the taxa remaining most stable within an individual varied from person to person. In other words, the same bacterial group (e.g. bacteria related to Faecalibacterium prausnitzii) could be one of the most variable in one individual and among of the most stable in another, highlighting the fact that there is subject-specificity in the microbial composition as well as stability of microbes in healthy subjects. However, some common trends in the temporal dynamics of the intestinal bacteria could be observed. The temporal variation of individual taxa (measured as Coefficient of Variation, CoV) was on average 6.3% and the same measured over the whole microbiota profile was 4%, indicating that the healthy subjects microbiota had high temporal stability. Nonetheless, phylotypes related to the Bacteroides vulgatus and uncultured Clostridiales were some the least stable taxa (CoV 10.0% and 13.3% respectively) in most of the subjects whereas phylotypes related to Clostridium symbiosum and Ruminococcus obeum remained stable over the 7-week study period in most of the participants (average CoV 4.4% and 3.7%, respectively). The observed high temporal stability indicates that there is host-specific selection of the intestinal microbial composition and is evidence for an individual core microbiota. One of the characteristics of a
Results and Discussion

healthy intestinal microbiota has been suggested to be the ability to be resilient to change and to maintain its composition and more importantly, the function of the microbiota (Lozupone et al., 2012). There has been several studies supporting our findings of temporally stable microbiota with both short- (Scanlan et al., 2006; Vanhoutte et al., 2004) and long-term follow-up (Rajilic-Stojanovic et al., 2012). The ability to understand the temporal stability maintained by the microbiota would represent an important step in being able to predict diseases that resulted in alterations in the microbial community and for the long term in the development of therapies to retain the balance.

5.2.2 DEFINITION OF THE COMMON CORE MICROBIOTA

One way of defining the healthy microbiota is to concentrate on the commonly shared bacteria, conserved through the co-evolution of the host and intestinal inhabitants. It would be plausible to hypothesize that the composition of the common core microbiota would have been selected to include bacterial taxa with health-associated properties. Therefore, once characterized, they could possess diagnostic and therapeutic potential.

Previously the common core microbiota has been determined with somewhat arbitrary pre-defined criteria for the prevalence of the detected phylotypes, i.e. a phylotype was required to be detected in a certain percentage of study subjects, ranging from 50% (Tap et al., 2009) to 100% (Qin et al., 2010) of the subjects. In the absence of established prevalence thresholds, it is difficult to compare studies examining the common core. In addition, in these studies the number of study subjects varied and their health status had not been determined. Moreover, the results may well have been affected by methodological issues including coverage of the selected primers, and the choice of hypervariable regions of the 16S rRNA gene (Turnbaugh et al., 2009) as well as the analysis depth (Hamady and Knight, 2009) and the phylogenetic assignment (Claesson et al., 2009).

The approach followed in Study II was to consider the common microbial core as an outcome of continuous prevalence and detection thresholds, i.e. to refrain from using pre-set thresholds since it was unclear which of these thresholds would be biologically most meaningful. The result revealed the deterministic role of such threshold in defining the core size. If we required for a taxon to be present in all participants at high abundance, then the size of the common core microbiota was practically zero. Hence, we included all healthy subjects of Study II (n=9) in the analysis and utilized the full detection range of the phylogenetic microarray, including also phylotypes with a low abundance (minimum relative abundance 0.05%), while covering a large dynamic range. By using this detection threshold and requirement for 100% prevalence we concluded that the core microbiota of 9 healthy subjects consisted of 288 species-like taxa representing about 40% of the detected phylotypes. The common core microbiota mainly included members from the Firmicutes phylum, specifically bacteria related to R. obeum, F. prausnitzii, and O. guillermondii but there were also phylotypes from the other major phyla, Actinobacteria and Bacteroidetes. However, the abundance of these latter phylotypes varied greatly from subject to subject. When combining the data from all healthy subjects of this thesis and computing the core microbiome we could conclude that the core contained 208 phylotypes and was dominated again with Firmicutes from the Clostridium clusters IV and XIVa (Figure 4).
The estimation of core microbiota as a continuum of prevalence and abundance, introduced in Study II, has been followed-up in a large-scale study involving 115 healthy individuals (Salonen et al., 2012). They verified our initial findings of the core size and its composition and could not detect any core taxa with a relative abundance over 2.5%, indicating that the abundant taxa vary between individuals. However, when setting the detection limit to 0.05% of the relative abundance, the common core of these 115 healthy individuals included 30% of all detected phylotypes.

In addition to the common core composition, we assessed the temporal stability of the shared phylotypes, since this had not been addressed in previous studies. The majority (70%) of the core phylotypes fluctuated by less than 10% during the 7-week study period. The high temporal stability of the individual core could suggest a strong selection by the healthy host for a particular microbial composition, but the temporal fluctuation of certain taxa, such as Bacteroidetes (Eckburg et al., 2005), partly explains the lack of phylotypes from these genera in the core analysis. Previous findings either supported (Sekelja et al., 2011; Tap et al., 2009; Turnbaugh et al., 2009) or contradicted (Claesson et al., 2010; Qin et al., 2010; Willing et al., 2010) our findings of low numbers of Bacteroidetes in the core microbiota. There are various reasons to explain this finding, including several technical issues, such as choice of DNA extraction method and its effect on the relative share of Bacteroidetes (Study I) as well as the depth of the sequencing analyses. Other factors such as subjects’ age, nationality and diet may affect the obtained results, but there are presently too few data to address these in a coherent way.

The core microbiota identified in Study II contained more phylotypes than formerly reported. The previous estimations of the size of the common microbial core size have varied from a few percentages (Turnbaugh et al., 2009) to over 30% (Huse et al., 2008; Qin et al., 2010). The within-subject temporal variation of certain taxa could be one of the reasons why the commonly core microbiota have not been detected in some of the previous studies. However,
with a analytical method able to detect low abundant species, accommodating the large individual variation, the core microbiota of healthy subjects could be addressed. In our study, the use of phylogenetic microarray facilitated the detection of the core microbiota as the microarray has a low detection threshold and only quantifies the fixed set of previously known phylotypes, minimizing the overall variation in microbiota profiles. Taxa with a relative abundance as low as 0.05% are not readily detectable with the NGS approaches, hence they may well have remained undetected in the previous studies (Claesson et al., 2010; Qin et al., 2010).

5.2.3 MILD INTESTINAL SYMPTOMS ASSOCIATED WITH MICROBIAL COMPOSITION

It is very common for healthy individuals to occasionally experience intestinal discomfort. A survey mapping the prevalence on intestinal complaints in healthy Finnish adults (n=46), revealed that the most frequent intestinal symptom was flatulence, experienced at least weekly by 66% of the healthy subjects (Jalanka, unpublished data). Moreover, 29% of the participants experienced intestinal pain at least monthly (Figure 5). Therefore, not surprisingly the participants of Study II, though in good general health, experienced occasional, mild intestinal complaints throughout the study period (Study II, Table 1). The most frequent complaints were mild to moderate feelings of abdominal bloating or pain. We aimed to find correlations between the intestinal complaints and the intestinal microbiota composition. We identified a significant positive association between bloating and the abundance of uncultured bacteria related to *Anaerotruncus colihominis* and *Ruminococcus callidus* (*p*<.05). Both have been associated previously with rumen microbiota and cellulose fermentation (Li et al., 2012). This process is known to produce both H₂ and CO₂ and hence could explain the observed association with bloating. Moreover, previously

![Figure 5: Commonly experienced intestinal symptoms recorded from 46 healthy subjects participating in a science outreach.](image-url)
the bacterial overproduction of H$_2$ has been claimed to contribute to intestinal symptoms in healthy subjects (Kleessen et al., 2007). Other phylotypes with significant associations to bloating included uncultured relatives of *Ruminococcus flavefaciens* (*p* < .05). The abundance of these *R. flavefaciens*-like bacteria was elevated in individuals with increased perception of symptoms. *R. flavefaciens* is an anaerobic bacterium, known to obtain nutrients by breaking down cellulose resulting in the production of organic acids, acetate as well as CO$_2$ and H$_2$ (Julliand et al., 1999).

The other frequently recorded symptom among the subjects was abdominal pain, which correlated with several phylotypes from the *Bifidobacterium* genus, but also a phylotype from the genus *Tannerella*. The latter was also associated in Study IV with the IBS-type microbiota and increased mucosal expression of pro-inflammatory tumor necrosis factor (TNF) pathway (Study IV, Figure 3). In Study II, the subjects experiencing abdominal pain had 6-fold less bifidobacteria than subjects with no pain (*p* < .05). These results were verified with qPCR and provide further correlation-based evidence for the impact of bifidobacteria on intestinal health.

The inverse correlation between bifidobacterial abundance and abdominal pain was a novel finding in healthy adults but has been previously but less pronouncedly reported in IBS patients (Malinen et al., 2010; Tana et al., 2010) and in IBS patients consuming probiotics with bifidobacterial species (Guglielmetti et al., 2011; Whorwell et al., 2006). Several studies have reported decreased levels of bifidobacteria in IBS patients (reviewed by (Salonen et al., 2010) and a probiotic *B. infantis* is documented to reduce visceral pain both in mice (McKernan et al., 2010) and humans (Brenner et al., 2009). On the other hand, there is also a study in IBS patients where the decrease of bifidobacteria coincided with decreased the experienced GI symptoms during a FODMAP (Fermentable, Oligo-, Di-, Mono-saccharides and Polyols) restriction diet (Staudacher et al., 2012). However, the reduction of bifidobacteria in that study was likely due to the reduced intake of bifidogenic fibres during FODMAP restriction rather than being causally associated to the reduced symptoms of flatulence, bloating and pain, as bifidobacteria do not produce gas during growth.

### 5.3 PERTURBATIONS TO THE HEALTHY MICROBIOTA – THE EFFECT OF BOWEL CLEANSING (STUDY III)

The stability of the intestinal microbiota in a healthy subject is high and resilient to change since it has been shown to recover back to its original composition relatively rapidly after perturbations (Study II). Bowel cleansing introduces an osmotic flow of fluids into the intestine reducing the amount of luminal bacteria, introduces oxygen (Strocchi et al., 1990) into an anaerobic environment, and reduces the amount of available nutrients. All of these changes are likely to affect the microbial ecosystem.

One of the main objectives of the bowel cleansing procedure is to clear the intestinal tract from faecal material as well as from its microbial content prior to colonoscopy or gastric surgery. Although the effectiveness and safety of this procedure has been extensively characterized (Connor et al., 2012), very little is known about its effects on the intestinal microbiota. Study III aimed to address this aspect of lavage and, moreover, to determine whether different doses of a purgative would cause long-lasting effects in the recovery rate of the microbiota. A total of 23 healthy subjects were randomly assigned into groups receiving two litres of the laxative in two different dosage protocols (Garsed et al., 2012). One group consumed a split dose of the study substance where the first litre was consumed in the evening and the second taken in the
following morning. The other group consumed the 2 litres of the PEG solution as single dose in the morning of the study. Four consecutive stool samples were collected; at baseline, immediately after the lavage and two follow-up samples 14 and 28 days after the bowel cleansing.

5.3.1 THE EFFECT OF BOWEL CLEANSING ON THE INTESTINAL MICROBIOTA COMPOSITION

It has been shown that the best result in bowel cleansing prior gastric operation is achieved by using a split dose over a single, larger dose of the purgative (Cohen, 2010). However, the effectiveness of the two different doses on the microbiota had not been addressed previously, though this procedure is widely used in clinical practice. In this study, both doses of the purgative significantly reduced the total bacterial amount quantified at the point of lavage ($p<.05$). However, the split dosing introduced a higher reduction to the total microbial load than observed with the single dose procedure. Since one of the main aims of the bowel cleansing is to reduce the microbial load (Strocchi et al., 1990), the split dosing procedure was a superior method in that respect.

To investigate the individual differences in the microbial perturbation and recovery, we performed a hierarchical clustering of the total microbial profiles. The composition of the intestinal microbiota is very subject-specific, which results in the unsupervised grouping of consecutive samples from the same individual in a hierarchical clustering analysis (Study II) (Rajilic-Stojanovic et al., 2012). In this cohort, in 5 out of 23 (22%) subjects, the sample taken immediately after lavage did not cluster with the earlier sample from the same individual, indicating that the composition was drastically altered. However, already 14 days after the lavage the microbial composition of these subjects had returned its original state and resembled the baseline sample with 94% similarity. This indicated that the majority of the microbiota returned to its original composition after the laxative treatment, although the lavage introduced a transient dramatic change.

5.3.2 CHANGES IN THE INTESTINAL MICROBIOTA COMPOSITION IMMEDIATELY AFTER LAVAGE

Consumption of laxatives is known to increase the water content in the intestine resulting into removal of the luminal content and faecal material (Strocchi et al., 1991). Accordingly, we observed a large, 31-fold reduction of the total microbial load caused by the laxative treatment as well as changes in the faecal microbiota composition. Samples collected immediately after the lavage were significantly different when compared to the baseline ($p<.05$), according to the mean relative abundances of genus-like phylogenetic groups. The levels of bacteria affiliated to the Bacilli and Clostridium cluster IV were significantly decreased at the point of lavage ($p<.05$). The latter included abundant butyrate-producing bacteria related to F. prausnitzii as well as less known and so far uncultured representatives of Oscillospira guillermondii. Interestingly, we also noted that the proportion of some members of the Clostridium cluster XIVa increased (bacteria related to R. obeum and Dorea formicigenerans) in addition to the increase of several Proteobacteria (all $p<.05$).

The GI tract contains high levels of endogenous and exogenous proteases (Antalis et al., 2007). These enzymes are involved in the breakdown of proteins as well as in the regulation of several other cellular processes such as cell-cycle progression, cell death and DNA replication (Turk, 2006). Elevated levels of faecal serine proteases (FSP) are detected in some
IBS patients, and are considered to increase the intestinal permeability and potentially lead to visceral hypersensitivity (Gecse et al., 2008). In this study, the laxative treatment resulted into momentary increased levels of FSP, previously shown to be of pancreatic origin (Tooth et al., 2013). To determine whether the elevated FSP could be associated to compositional changes of the intestinal microbiota, we correlated the bacterial taxa with the sudden change in the FSP levels. The strongest correlations were observed between bacteria related to *F. prausnitzii* \( (r=-0.40) \), whose levels decreased during the lavage, in addition to the positive association detected between bacteria related to *D. formicigenerans* and FSP \( (r=0.45) \). Our findings are in line with a recent report where the increased FSP amounts were associated with lowered amounts of *F. prausnitzii* (Carroll et al., 2013). It has been shown in experimental rat studies that the majority of the secreted FSP is degraded by the intestinal bacteria (Genell and Gustafsson, 1977). Hence, the correlation of FSP amounts with reduced counts of important bacterial groups as *F. prausnitzii* suggests that having normal levels of such bacteria ensures an adequate degradation of endogenous pancreatic protease thus avoiding sensitisation of the rectum.

### 5.3.3 IMPACT OF DIFFERENT DOSES OF THE LAXATIVE TREATMENT ON THE MICROBIOTA

One aim of this study was to evaluate whether dosing of the purging agent would affect recovery of the microbiota after the laxative treatment. When analysing the differences in the abundance between the baseline and follow-up samples, we identified taxa from 5 different genera to be statistically significantly different 14 or 28 days after the ingestion of PEG solution \( (p<0.05) \). There was a 19% overall increase in all of the involved bacteria. Specifically, there was increase in Clostridial species related to *D. formicigenerans* as well as an increase in the facultative aerobic bacteria belonging to Fusobacteria and Proteobacteria, indicating that these potentially pathogenic bacteria had increased in abundance after the bowel cleansing. These organisms, although all part of the normal microbiota, harbour pathogenic potential and their increased abundance have been associated with adverse health outcomes (Chow et al., 2011). The increased abundance of phylotypes related to *D. formicigenerans* has been described in IBS (Rajilic-Stojanovic et al., 2011), Ulcerative colitis (Nomura et al., 2005) as well as in non-alcoholic fatty liver disease (Raman et al., 2013). Similarly, Fusobacteria and Proteobacteria are Gram-negative bacteria often implicated in inflamed intestinal tissue (Swidsinski et al., 2011). Although the increase of these potential pathogens might not have any health implications in healthy individuals, in conditions involving decreased microbiota diversity the increase in these potential pathogens could be more substantial and their metabolic products may evoke to intestinal symptoms or health impairment.

It has been claimed that the split dose of PEG solution is superior to the single dose in the bowel cleansing with an osmotic laxative treatment (Cohen, 2010). The present microbiota results support these findings. On the day of lavage, there was a significantly higher reduction in the total bacterial counts with the split dose treatment as compared to that achieved with the single dose. Moreover, there was a higher microbiota recovery with the split dose treatment than with the single dose, where the bacterial phyla such as Proteobacteria remained elevated and did not return to their original levels after 28d of consumption of the purgative (Figure 6). In conclusion, the split dose of the laxative has a less disturbing effect on intestinal microbiota than a single dose.
5.4 PERTURBATIONS TO THE HEALTHY MICROBIOTA - MICROBIAL CHANGES IN POST-INFECTIONOUS IRRITABLE BOWEL SYNDROME (STUDY IV)

The aim of this study was to characterize differences in the microbial composition of PI-IBS patients with varying degrees of symptoms, and to address the associations between the faecal microbiota and the clinical features of IBS. As the development of IBS is multifactorial, we hypothesized that the observed symptoms and development of PI-IBS may arise from the interplay between the faecal microbiota, the host immune response, and psychological factors. In order to find these associations we conducted a detailed characterization of the faecal microbiota of patients in different degrees of recovery from gastroenteritis and compared these to healthy controls and patients with IBS-D.

5.4.1 THE INDEX OF MICROBIAL DYSBIOSIS

We hypothesized that instead of limiting the comparative microbiota analysis to individual taxa, it would be relevant to search for a microbial signature that potentially differs between the patients and healthy controls. By using a bioinformatic approach that allowed this, we identified a combination of 27 genus-like groups that separated the study groups (Figure 7). The microbial profile of the PI-IBS and PI-BD groups resembled those of the IBS-D patients. The major difference in the patients' microbiota was the 12-fold increase of Bacteroidetes phylum and the 35-fold decrease of Uncultured Clostridia from the Firmicutes phylum. Furthermore, when all participants were indexed according to the abundance of the discriminating bacterial taxa (Figure 7), a specific ordering of the study subjects was obtained, reflecting their health status, a so called index of dysbiosis (IMD).

Figure 6: The microbial recovery of the Proteobacteria phyla calculated as the correlation of the microbial profile against the baseline sample, in the two dosing groups. Statistical significance is indicated with an asterisk.
Figure 7: Index of microbial dysbiosis (IMD) was created using Bootstrap aggregated Redundancy analysis (RDA) of the microbiota composition showing the bacterial groups significantly contributing to separation between the healthy controls and the patient groups. The position of each study subject in the plot was transposed on the first axis of the RDA, which explains the majority of the microbial separation, creating an index of microbial dysbiosis (IMD).

When the Firmicutes/Bacteroidetes ratio was measured, there was a significant decrease in the PI-IBS group when compared to all of the healthy controls examined in the thesis work ($p<.05$, unpublished data, Figure 8). The increase of Bacteroidetes was confirmed with qPCR of the Bacteroides-Prevotella-Porphyromonas-group. Interestingly, similar microbiota changes have been associated with susceptibility of enteric pathogens (Haag et al., 2012). It was thought that the infection predisposition of an individual could increase due to a certain microbial composition, since the complex interactions between the microbiota and host immune system might expose certain individuals to gastroenteric episodes.
The role of Bacteroides ssp. in the aetiology or pathophysiology of IBS is unclear. Previous studies have reported opposing results regarding the relevance of Bacteroidetes in IBS patients. Some have reported increased amounts of Bacteroides ssp. in patients with IBS (Kassinen et al., 2007; Parkes et al., 2012; Ponnusamy et al., 2011) and also in related functional bowel disorders, (Manichanh et al., 2013) whereas others have detected a higher abundance in healthy controls (Jeffery et al., 2012; Krogus-Kurikka et al., 2009; Rajilic-Stojanovic et al., 2011). Explanations to the discrepancy are many, but such extensive variation in the microbiota findings may suggest that rather than a single pathobiont, it is the altered microbial community as such that is the important factor in the pathophysiology of IBS. Changes in the host might trigger the observed changes as for example accelerated transit is known to modify the microbiota (Kashyap et al., 2013a) and Study III). Currently, it cannot be clearly stated whether the microbiota change is the cause of IBS or whether it is secondary to IBS pathophysiology.

Previous studies have shown that the microbial composition of IBS patients could be used to further stratify IBS patients and thus complement the clinical subgroupings (Jeffery et al., 2012). Our analysis on the patients in the post-infectious subcategories of this cohort verified these previous findings. We tested the clinical significance of the IMD by correlating it with the diagnostic variables of IBS patients and found positive associations with the IBS-type microbiota and several intestinal complaints. However, there were no significant correlations with the psychological symptoms often associated with IBS. This was also replicated with another cohort where IBS patients with a microbial profile similar to healthy controls had significantly more psychological problems than the IBS patients with an altered microbiota (Jeffery et al., 2012). Therefore, it could be speculated that the IBS patients with an altered microbiota and those with psychological symptoms have different disease aetiology. Together, these results indicate that the discriminant microbial profile could be used as an objective measure of disturbed bowel functions and a potential novel way of stratifying this heterogeneous patient material.
The putative associations between the discriminant microbiota and host responses were determined by mapping the gene expression data derived from rectal biopsies to known biological functions using Gene Set Analysis (GSA), as well as other clinical measurements reflecting inflammatory responses. The IMD correlated positively with several host expression pathways associated with inflammation as well as with increased cytokine production, whereas the patient disease status or separate microbial abundances alone did not show any significant associations.

5.4.2 ASSOCIATIONS BETWEEN MICROBIOTA AND HOST GENE EXPRESSION

To further analyse the host–microbiota crosstalk we concentrated on the 27 taxa of IMD and separately studied their associations between the host expressional changes. There were several independent measures linking the abundance of the IBS-type microbiota to changes in the host gene expression. We identified associations between the discriminant microbiota and the physical barrier integrity of the host. The most prominent negative correlation was observed between seven Bacteroidetes ssp. and the expression of glycine, serine, threonine metabolism pathway. It is known that majority of the dietary threonine is utilized for synthesizing the secretory mucin and that its dietary restriction results in the impaired gut barrier (Faure et al., 2005). In general, these amino acids are known to be important in the maintenance of gut integrity and the efficacy of the mucin layer (Hansson, 2012). Fucosyltransferase-2 is involved in the formation of the ABO blood groups and also acts as an adhesion receptor for several microbes (Moulds et al., 1996). It has been claimed that it is the mucus structure and the extent of the glycosylation, which shapes the microbiota composition (Wacklin et al., 2014). Fut2 deficient mice (Fut2−/−) have been shown to display differences in microbial diversity and the community composition as they harbour increased amounts of several Bacteroides species and increased levels of Clostridiales (Kashyap et al., 2013b). Moreover, conditional knock-out mice lacking the enzyme responsible for o-glycosylation in mice, core 1 β1,3-galactosyltransferase (C1galt1) presented subtle changes in their intestinal microbiota composition, where Bacteroidetes was increased and the Firmicutes were reduced (Sommer et al., 2014).
Additionally to the alterations in the expression of glycine, serine, threonine metabolism pathway, there were several other pathways associated with the IMD bacterial taxa and cell junctions, or cell junction integrity as well as the pathways indicating increased inflammatory response (see Figure 3 from Study IV). Based on these findings, our hypothesis is that the physical barrier of subjects with IBS-type microbiota has been compromised resulting in a low-grade inflammation, a typical symptom of IBS patients. Moreover, we could show that the IMD correlated with the IBS symptoms and clinical features whereas in the patients with a more psychological basis for the disease, the effect of microbial dysbiosis was smaller (Figure 9).
6. CONCLUSIONS AND FUTURE ASPECTS

The work presented in this thesis represents new knowledge in the field of intestinal microbiota research from several different aspects. We addressed a source of technical biases in the microbiota research and validated an efficient faecal bacterial DNA extraction method. Moreover the characteristics of the healthy intestine were benchmarked and compared with two common perturbations. Our study on IBS was among the first to perform an integrated analysis of the different aspects of the disease etiology allowing us to obtain a holistic view of the patients and to identify several novel associations between the intestinal microbiota and the host’s clinical status.

The Study I was focused on finding an optimal DNA extraction procedure to be used throughout the thesis. The optimal extraction protocol, repeated bead-beating (RBB), achieved the highest diversity and high DNA extraction efficiency, especially its ability to extract DNA from methanogens as well as Clostridium cluster IV phylotypes was superior. Our study was pioneering in that it was one of the first to exploit community-wide microbiota analyses comparing different DNA extraction protocols. Our results emphasized the importance of pre-analytical steps in the microbiome research. Recently an on-going, large comparative trial evaluating several DNA extraction procedures ranked the RBB method one of the top three faecal DNA extraction protocols (personal communication with Anne Salonen and Joel Dore, IHMS, 2012). This project validated our results, and responds to the high need for standard operating procedures in human microbiome research. Standardized methodology on all of the analysis steps, from sample collection and storage to DNA extraction and analysis pipelines is imperative for the efficient use of rapidly cumulating data as it enables meta-analyses and can provide a more holistic and conclusive view of the intestinal microbiota properties.

The characteristics of the healthy intestinal microbiota were addressed in Study II. We showed that the microbiota of a healthy adult possessed high individuality remains stable during consecutive weeks and was resilient to change. By studying the bacteria shared by most of the studied individuals, we determined a common microbial core. Additionally we observed associations between microbiota composition and common intestinal symptoms giving a novel insight in the potential functions of specific taxa such as bifidobacteria. Altogether, these findings provide new approaches to define intestinal health and to characterize the microbial communities inhabiting the human gut.

It has been shown that the function and composition of the intestinal microbiota of healthy subjects correlate with each other. The microbial composition displays high individual variation whereas the functions of the microbiota have been shown to be fairly constant (Huttenhower et al., 2012). This indicates that the key functions for health performed by the microbiota can be carried out by various different phylotypes and that the function is more relevant to health rather than “who” performs it. Therefore, the future research should focus on determining the differences in the microbial functions in health and disease and aberrations possibly contributing to disorders. Since the microbiome is also transcriptionally regulated, it would be important to conduct more studies on the metatranscriptomics and metaproteomics to obtain a more comprehensive view on the capacity of the microbiota. Currently very few studies utilizing this technique (Booijink et al., 2010; Gosalbes et al., 2011; Perez-Cobas et al., 2013) due to several challenges, such as the instability of RNA and a small amount (2%) of messenger RNA (Morgan and Huttenhower, 2014).
Conclusions and Future Aspects

The microbiota of healthy subjects were further benchmarked against perturbations, induced osmotic diarrhoea and a disease state IBS. Bowel cleansing is a standard procedure prior to colonoscopy or gastric surgery. Often the aim is to reduce the amount of intestinal microbiota to avoid bacterial complications. The results of Study III showed that the split dosing of the purgative agent reduced the total bacterial load significantly more when compared to the single dose. Moreover, there were significant differences in the recovery rates of certain bacterial groups between the two different doses. Bacterial genera including potential pathogens were increased in healthy subjects when the single dose approach was applied. No such similar phenomenon was detected with the double dosing regime. Our results provide novel insight into the effects of purgative agents on the patient and may become incorporated into the clinical practice via the recommendation to split the purgative treatment into two dose administrations. However, our present results were conducted in healthy individuals. Patients requiring colonoscopy or gastric surgery often have an underlying disease, such as inflammatory bowel disease or colon cancer, both conditions known to be associated with unstable microbiota (Table 4). The effects of bowel cleansing on the host will need to be tested in these patient groups separately, in order to achieve a better understanding of the effects of purging on their health.

The relevance of microbiota in the pathophysiology of IBS was addressed in Study IV. We could show that a group of 27 bacterial taxa separated the post-infected and the IBS patients from healthy control. Moreover, this index of microbial dysbiosis could be associated to gastrointestinal symptom severity, and clinical markers, allowing stratification of patients on the basis of the IMD. Interestingly the severity of anxiety or depression did not correlate with the microbial changes. The applicability of the IMD for stratification of the patient material should be tested in another, larger cohort containing both PI-IBS patients as well as other IBS-subtypes. The aim of this trial would be to validate the diagnostic potential of IMD species and development of a fast method for the detection of their abundance. Currently, the diagnosis and stratification of IBS patients relies greatly on subjective and symptom based diagnosis. Objective markers are mainly used to exclude other, more severe diseases. The intestinal microbiota could be used to further stratify patients, i.e. differentiating those with a psychological rather than organic cause. Thus, it would be highly relevant for both patients as well as for further research.

Study IV also focused on finding host-microbiota interactions in IBS patients. It identified several cellular pathways, such as protein metabolisim and cell junction expression pathways to be decreased when the levels of IMD species were increased. This work could be used as a starting point for future functional studies where both host and microbiota functional responses would be addressed. The effect of the aberrations in the microbiota could be tested in animal models such as gnobiotic rodents. Such studies have already revealed that IBS symptoms could be transferred with the microbiota suggesting a causal link between the microbiota dysbiosis and symptom development (Crouzet et al., 2013). However, in these studies the dysbiosis has been caused by different set of taxa therefore the relevance of the IMD species should be verified this in similar animal experiments, to confirm the effect of the observed dysbiosis. Moreover the significance of health related strains could be potentially tested in human trials. One example would be the supplementation of a novel, potential probiotics \textit{F. prausnitzii}, which have been shown to be significantly decreased or even absent in IBD patients in remission (Fujimoto et al., 2013; Varela et al., 2013). Therefore the re-establishment of these phyiotypes could be beneficial for these patients. However, due to the complexity of the microbiota, the effects of a single strain may only be moderate. Therefore, restoring the healthy microbiota composition via faecal microbial transplants (FMT) could be more efficient and beneficial to the patient. There is
very promising evidence of the positive effect of FMT in patients with antibiotic induced *Clostridium difficile* infection (van Nood et al., 2013) and has been used to treat also other GI disorders, for example IBD (Brandt and Aroniadis, 2012) as well as non-GI disorders such as multiple sclerosis (Borody et al., 2011).

Several health and disease related taxa identified in this thesis represent uncharacterized, so far uncultured phylotypes. One of the most consistent markers of IBS, detected in several independent cohorts including study IV, is the uncultured bacterium related to *Ruminococcus torques*. It would be highly relevant for future research and patient care to characterise this as well as several other uncultured phylotypes important for health. For example, Study IV demonstrated that the healthy subjects had 35-fold more uncultured *Clostridiales* than the post-infectious or the IBS patient groups. Therefore, it would be relevant to culture and characterise this group of bacteria.

The temporal stability of the healthy microbiota was covered in Study II, however no large-scale studies have been conducted in IBS patients with similar extensive follow-up to study the temporal fluctuations of the patient microbiota. It is very common for the IBS patient to experience day-to-day fluctuations in their symptoms, and the basis for this feature is unknown (Cremonini and Talley, 2005). Temporal instability of the patients’ microbiota could alter the physiology of the host and thus provide insight in the symptom fluctuation, characteristic of IBS patients. Moreover, repeated sampling of IBS patients during and between the symptom episodes would make it possible to investigate whether the discrepancy of the IBS-related microbial changes in the current literature could be partly attributed to the large temporal variation in symptom occurrence.
7. ACKNOWLEDGEMENTS

This study was carried out in the Department of Veterinary Biosciences, Faculty of Veterinary Medicine, University of Helsinki. The work was financially supported by the Finnish Funding Agency for Technology and Innovation, the Academy of Finland, Doctoral Program in Food Chain and Health and University of Helsinki. The Dean, Prof. Antti Sukura and the Head of the Department, Prof Airi Palva, are thanked for providing excellent facilities. I would also like to thank my supervisor Airi for her advice and support throughout this thesis work.

The individual to whom I am most grateful is my supervisor Prof. Willem de Vos. It has been a privilege to work in his cutting-edge research group and to be part of so many interesting research projects. Throughout the thesis, you have always given me responsibility and trusted that I could handle things, even when I have had no trust in myself. Thank you for this! This is also why I now know how to eat an elephant - one bite at a time. I also would like to express my deep gratitude to my third supervisor Anne Salonen for all the practical advice she has provided. I could always ask you for help, even with the most stupid questions. I am also very grateful for all of the red markings in my manuscripts; through your help, I have learned so much about scientific writing. I admire Anne's attitude towards science, and I can only hope that some of that will stay with me in the future. Lastly would like to thank Jarkko Salojärvi for his long-suffering patience with me. Thank you for teaching me so much about statistics and programming.

I would like to extend my gratitude to the external reviewers of this thesis, Jaana Mättö and Harri. I appreciate all of your hard work and excellent comments. I would also like to acknowledge my follow-up group members, Arthur Ouwehand and Janne Nikkilä for sharing their time and taking an interest in my thesis projects.

I wish to express my sincere gratitude to all collaborators and co-authors of the thesis projects. In particular, I would like to thank Prof. Robin Spiller for sharing his expertise and genuine enthusiasm for the science. All of the thesis projects have benefitted from Outi Vepsäläinen's excellent practical experience. Not only do I admire your organization skills, but I am also very happy that we have become such good friends during these years.

I have had the privilege to work with a great group of people. I would like to thank the past and present members of Mikro for their companionship and creating enjoyable atmosphere around the lunch table. Especially I am grateful for the peer support and friendship provided by the other PhD students, Pia Rasinkangas, Tanja Lähteinen, Ravi Kant, Katri Korpela and Carolin Kolmeder. In addition the excellent researchers Reetta Satokari, Veera Kainulainen, Justus Reunanen, Jarmo Ritari and Leo Lahti are thanked for the stimulating scientific discussions and their help and support through out the project. Lastly I would like to thank Francois Douillard with whom I have shared an office for several years. Thank you for all of the good music and chats that we have had, not only about science but also about life in general.

I would like to thank all of my friends for their love and support, but especially Kati Ahlqvist and Sofia Ahola. I am truly grateful for your friendship and the peer support you both have provided. Kati, you are the strongest and most trustworthy person I know and Sofia you are the bravest,
who will tackle life’s obstacles with a smile on her face. We all have had the best and the worst of times whilst completing our theses and you have always been there to share the load or with a champagne glass to celebrate the victories. I thank you for all of the Apulanta concerts and cups of tea we have shared. This thesis could have not been completed without these distractions. I would also like to thank Mikko Björklund for giving me the strength to finish the book and for all the long conversations about science, religion, the meaning of life and everything in-between. You are one of the most inspiring individuals that I have ever met and I cannot wait for all of the conversations that are ahead.

Lastly, I would like to thank my mother for always believing in me. She has said that I could accomplish anything I set my mind into and it would only take perseverance to achieve it. Thanks for always being there for me; you have been my rock throughout this project. As a child I dreamed about becoming a vet, like my father. He defended his thesis in May 1991, and sadly passed away shortly after. I feel that the circle has in some way closed, I am graduating from the veterinary faculty, but in a field that is my own and more suited to me. I would like to think I have made my father proud. But most of all I would like to thank my beautiful children, Selina and Henri, for teaching me what is important in life - the pair of you. My best and most significant work is being your mother.

Helsinki, November 2014
8. REFERENCES


References


References


References


References


References


References


References


References


References


References


