Diversity of the Human Gastrointestinal Microbiota

Novel Perspectives from High Throughput Analyses

Mirjana Rajilić-Stojanović

Promotor

Prof. Dr W. M. de Vos

Hoogleraar Microbiologie

Wageningen Universiteit

Samenstelling Promotiecommissie

Prof. Dr T. Abee

Wageningen Universiteit

Dr J. Doré

INRA, France

Prof. Dr G. R. Gibson

University of Reading, UK

Prof. Dr S. Salminen

University of Turku, Finland

Dr K. Venema

TNO Quality for life, Zeist

Dit onderzoek is uitgevoerd binnen de onderzoekschool VLAG.

Diversity of the Human Gastrointestinal Microbiota

Novel Perspectives from High Throughput Analyses

Mirjana Rajilić-Stojanović

Proefschrift

Ter verkrijging van de graad van doctor
op gezag van de rector magnificus
van Wageningen Universiteit,
Prof. dr. M. J. Kropff,
in het openbaar te verdedigen
op maandag 11 juni 2007
des namiddags te vier uur in de Aula

Mirjana Rajilić-Stojanović, Diversity of the Human Gastrointestinal Microbiota – Novel Perspectives from High Throughput Analyses (2007)

PhD Thesis, Wageningen University and Research Centre, Wageningen, The Netherlands – with Summary in Dutch – 214 p.

ISBN: 978-90-8504-663-9

Abstract

The human gastrointestinal tract is densely populated by hundreds of microbial (primarily bacterial, but also archaeal and fungal) species that are collectively named the microbiota. This microbiota performs various functions and contributes significantly to the digestion and the health of the host. It has previously been noted that the diversity of the gastrointestinal microbiota is disturbed in relation to several intestinal and not intestine related diseases. However, accurate and detailed defining of such disturbances is hampered by the fact that the diversity of this ecosystem is still not fully described, primarily because of its extreme complexity, and high inter-individual variability. The work presented in this thesis summarises the hitherto gathered knowledge on the diversity of the human gastrointestinal microbiota based on an extensive literature search that aimed to gather all reports of microbial life within the gastrointestinal tract starting from the first isolation in 1885. This resulted in the construction of the ecosystem-specific database of small subunit ribosomal RNA sequences of over 1,000 human gastrointestinal inhabitants based on which a phylogenetic microarray was designed. This newly developed HITChip (Human Intestinal Tract Chip) enabled comprehensive and high throughput analyses, and provided some novel perspectives on the diversity of this ecosystem, which include the following: (a) the first molecular follow up of the long-term microbiota dynamics showed that genetic, environmental and stochastic factors are shaping the microbial community within the gastrointestinal tract, and revealed that some members of Actinobacteria and Bacteroidetes are highly influenced by the genotype of the host, while determening the host-specific microbiota signature; (b) the analysis of the microbial diversity of intestinal contents of patients suffering from irritable bowel syndrome (IBS) and ulcerative colitis (UC) showed that both diseases impacted the microbiota, as its composition was distinctive from that of healthy adults, although the disturbance of the microbiota in relation to UC appeared to be more severe when compared to IBS, and was characterised, in addition to the different relative abundance of numerous members of the microbiota, by dramatically reduced richness of the microbial community; (c) finally, studying the development of the microbial community in a model of the human colon (TIM-2) provided novel guidelines for improving the handling procedures that will ensure more accurate representation of the microbial diversity of healthy adults. In conclusion, the results indicate that a high throughput and comprehensive studying of the human gastrointestinal microbiota, as facilitated by the application of phylogenetic microarrays is a promising approach for further describing and understanding of this exceptionally relevant microbial ecosystem.

Keywords: gastrointestinal microbiota, diversity, phylogenetic microarray, SSU rRNA

Contents

Outline

Evidence for the presence of specific microbes in the human gastrointestinal tract was provided at the end of the XIX century when Teodor Escherich isolated *Escherichia coli* from children's diarrhoeal faeces. Although this bacterial presence was considered as an infection, it soon became clear that a complex microbial community permanently resides in the gastrointestinal tract, and forms an essential element of the human body. A plethora of functions that contribute to the host's nutrition and health has been assigned to these microbes that collectively are termed microbiota. It has even been suggested that - because of its functions, dimensions and appearance - the microbiota can be considered as a virtual organ within the gastrointestinal tract. Furthermore, it seems that the disturbance of the delicate balance between numerous members of this ecosystem is associated with a number of intestinal and non-intestinal disorders. The appreciation of the importance of the gastrointestinal microbiota has generated significant interest in a further understanding of its dynamics, structure-function relationship and the possibility to modulate those. Although it is prerequisite for understanding and exploiting this ecosystem, the description of its enormous diversity is still not completed.

The work presented in this thesis aims to systematise the generated knowledge on the gastrointestinal microbiota diversity, and provide an improved insight into its relation to health and disease using a newly developed, high-throughput technology. Specific questions that were addressed by each of the chapters are listed below.

Chapter 1 summarises the present perception of functions and dimensions of the gastrointestinal microbiota, and provides an overview of the currently available techniques for studying the diversity of this ecosystem.

The following two chapters deal with the diversity of the human gastrointestinal microbiota. The results of cultivation-based studies, which are an important source of information because of their extensive use for the analysis of a large amount of samples are summarised in **Chapter 2**. This chapter presents a thorough analysis of gastrointestinal isolates, embedded in a curated phylogenetic framework, given through a historical perspective with the aim to clarify their importance. An alternative to time-consuming, laborious and expensive cultivation-based analysis of the microbiota's diversity are molecular approaches, such as those based on small subunit ribosomal RNA (SSU rRNA) gene sequences. The application of molecular techniques has revolutionised our view of many microbial ecosystems, and their impact on the knowledge of human gastrointestinal ecology is portrayed in **Chapter 3**. This chapter also provides an up-to-date, systematic overview of the gastrointestinal microbiota's diversity, built by integrating results that were generated by the application of culturing and molecular approaches, and that accumulated during more than a century long studying of this ecosystem.

Chapter 4 describes a generic oligonucleotide probe design strategy based on two hypervariable regions of the SSU rRNA gene sequence. This approach is high throughput, can be readily applied to any complex ecosystem, and allows easy addition of probes that target its newly reported inhabitants, which is essential for not yet fully described ecosystems, such as the one that resides in the human intestine. The application of the described strategy on the human gastrointestinal microbiota resulted in the development of the HITChip (Human Intestinal Tract Chip), a phylogenetic microarray that was applied for the analysis of the human intestinal microbiota's diversity described in the further chapters of this thesis.

In vitro fermentation of intestinal contents is a widely used approach to study the metabolic activity of the human gastrointestinal microbiota. To clarify whether such systems can adequately mimic the gastrointestinal microbiota in the absence of the human host, the microbial diversity was monitored during inoculum preparation and fermentation in an *in vitro* model of the human colon. The results of this study are presented in **Chapter 5**.

In **Chapter 6**, an in-depth analysis of the microbiota dynamics in healthy subjects is presented. It has previously been established that the intestinal microbiota is stable and individual-specific, although the long-term microbiota development has not been previously addressed. In addition to a short time interval, we have assessed the potential changes of the microbiota after a period of four years and included an analysis of the microbiota of monozygotic twin pairs, as it provides information about the impact of the host genotype. The ultimate goal of

this study was to reveal the phylogeny of the carriers of long-term stability and individual-specificity of the gastrointestinal microbiota.

Chapter 7 and **Chapter 8** present the results of the application of the analysis of the microbiota's diversity in relation to two intestinal disorders with unknown aetiology. Irritable Bowel Syndrome (IBS) is one of the most common functional intestinal disorders the incidence of which is increasing in the developed world. The multi-factorial nature of IBS pathogenesis and pathophysiology makes its definition and treatment particularly challenging. Among numerous factors that are potentially involved in the development of the IBS the gastrointestinal microbiota seems to be a relevant one. **Chapter 7** describes assessed differences in the microbiota composition of IBS patients and corresponding healthy controls.

Ulcerative colitis (UC) is an acute and chronic inflammatory disease of the large bowel characterised by an aggressive immune response to a subset of commensal intestinal microbes. Although both environmental and genetic factors are relevant for the development of UC, the intestinal microbes are essential for the occurrence of this disease. To define possible specificities of the microbiota composition in UC patients, the total microbiota of UC patients and corresponding controls was analysed. Additional effort was put into defining the diversity and identification of intestinal sulphate reducing bacteria, and results of this study are presented in **Chapter 8**.

Finally, in the general discussion (**Chapter 9**), the results of the work described in this thesis were summarised and discussed. Bearing in mind the most recent achievements concerning the human gastrointestinal ecology, the author's view on the future developments in this dynamic field of research are given.

CHAPTER 1

Mirjana Rajilić-Stojanović and Willem M. de Vos

Modified version of this chapter is accepted for publication in Agro Food Industry Hi-Tech

General Introduction

The human gastrointestinal tract is populated by reportedly the densest microbial ecosystem, which is essential for the digestion of foods and affects health. The basic understanding of the human microbiota is rapidly expanding by the implementation of molecular approaches that obviate the need for microbial cultivation. This report aims to provide an overview of the most relevant molecular techniques to describe the community structure of the microbiota in our gastrointestinal tract within a historic perspective and with specific attention for its association with health and disease.

Introduction

Ever since we are born, our body becomes colonised by myriads of microbes. This also holds for the human gastrointestinal tract that represents the most densely colonised interface. While this notion of microbial presence has developed over more than a century, it is only during the last decade that we have become to appreciate the diversity of the gastrointestinal microbiota. In retrospect, this reflects the revival of the interest in microbes, that although invisible to the naked eye, represent the most abundant and diverse life form on our planet (518).

One of the first reviews that reported the complexity and importance of the gastrointestinal microbiota for the host's nutrition, health and wellbeing dates from 1977 (418). Many other reviews have followed and recently, it has even been suggested that because of its functions, dimensions and appearance, the microbiota can be considered as a virtual organ within the gastrointestinal tract (353). A plethora of functions have been assigned to the human gastrointestinal microbiota that relate to its interactions with the host (124, 348) and the mechanisms underlying these are just beginning to be understood. A major obstacle for a better comprehension of those interactions is the incomplete description of the microbiota's enormous diversity, notably as each of us is inhabited by an individual-specific microbial community composed of hundreds of distinct organisms (129, 275, 332). Therefore, there is a need for the improvement of already established, and the development of novel, creative approaches that enhance our insight into the microbiota. This communication provides summary of the present perception of functions and dimensions of the gastrointestinal microbiota, and provide an overview of the currently available techniques for studying the diversity of this ecosystem.

Functions of the human gastrointestinal microbiota

The human gastrointestinal tract is responsible for food intake and digestion, and represents a system of organs of which colon is traditionally seen as the sewage of the human body, responsible for recovery of water, and removal of waste material. Although those functions have been well established, this part of the human body, is inhabited by up to 10^{12} cells per gram of intestinal content (418), which makes it the densest microbial system reported (23), and expands tremendously the functions performed by this part of the human body. Microbial ecosystem present in the colon but also along the entire gastrointestinal tract, is so active, that it has been proposed that it equals that of the liver (124). Some of those numerous functions are summarised below.

The most basal function of the gastrointestinal microbiota includes its contribution to the digestion of compounds resistant to the human digestive enzymes (99, 288). Substrates available for microbial conversion include food components, such as resistant carbohydrates,

proteins, lipids, nucleic acids and other polymers. It has been estimated that approximately 100 g per day of these substances are digested by the microbiota (99). Additionally, the microbiota is processing large quantities, up to one kg per day, of substances secreted or simply shed from the human host, such as mucins, epithelial cells or enzymes (Table 1). Conversion of the food and host-derived substrates by a complex microbial community results in more than 100 volatile compounds (168). Only a fraction of those complex conversions was studied in more detail, and better insight has been provided for the metabolism of carbon, nitrogen and sulphur compounds as summarised below.

The principal products of the gastrointestinal microbiota are short-chain fatty acids (SCFAs), and the gasses hydrogen, carbon dioxide and, in case of one third of the human population, methane (58, 98). SCFAs provide the human body with energy, and are supposedly impacting health (68, 99). Although acetate is produced in the largest quantity, butyrate has been proposed as the SCFA of the greatest importance since it provides about 60% of energy for the colonic epithelium (99, 290). However, no *in vivo* data have been reported to confirm this, while the steady state concentration of butyrate is below the detection level in humans (Venema, K., personal communication). Furthermore, some novel findings indicated that butyrate can have a necrotic effect on the colonic cell lines (307). Based on this conflicting data, the beneficial role of butyrate is presently being reconsidered (422).

Table 1 Substances of external and internal origin present in the human gastrointestinal tract and potentially available for microbial conversion.

Substance	Amount	Reference
Carbohydrates (total intake)	~ 410 g/day	(94)
Proteins (total intake)	~ 120 g/day	(94)
Fat (total intake)	~140 g/day	(94)
Bile	~1 g dry matter/day	(36)
Intestinal epithelial cells	~150 g dry weight/day	(96, 97, 161)
Mucus	~200 g glycoprotein/day	(7, 21, 145, 277, 360)
Pancreatic juice	~12 g protein/day	(404, 479)
Immunoglobulins	~4 g/day	(226, 382)

The active microbial community within the gastrointestinal tract contributes significantly to the host's amino acids homeostasis by the *de novo* synthesis of essential amino acids (221, 317). Experiments with the ^{15}N urea supplement in the diet, showed that the gastrointestinal microbiota incorporates this trace atom into all amino acids, although with

different levels of enrichment (173). Remarkably, the levels of the amino acids enrichment found in the plasma albumin were substantially reduced after antibiotics treatment. Due to the fact that microbes can synthesise amino acids from different precursors including dietary nitrogen, urea synthesised by liver, and endogenous and exogenous amino acids, the quantitative contribution of microbial amino acids synthesis remains unknown. Nevertheless, it was shown that the contribution of microbial lysine derived from dietary nitrogen varies between 5 and 44% of the total plasma lysine in healthy and ileostomy patients respectively (318).

The metabolism of sulphur by the human gastrointestinal microbiota was studied in detail due to the fact that substantial quantities of highly toxic hydrogen sulphide, which might be involved in colon cancer and ulcerative colitis aetiology, are produced in the human gastrointestinal tract (239, 375). Moreover, hydrogen sulphide has recently been identified as genotoxic agent (20) and a neuromodulator affecting the gut signalling (423). Production of hydrogen-sulphide is traditionally associated with sulphate-reducing bacteria, a minor group of the gastrointestinal microbiota that utilise sulphates derived from food or mucin ethers (171). However, since the increased protein intake has direct correlation with the increase of hydrogen-sulphide levels in faeces (297), and other sulphur compounds (168), proteins and amino-acids seem to be another relevant sulphur source in the human intestine. The analysis of the oral microbiota showed that a number of bacterial species, including *Eubacterium*, *Bacteroides* spp. and *Fusobacterium* spp., can release sulphide from L-cysteine (366). Some of these bacteria are abundant members of the human gastrointestinal microbiota, while the production of hydrogen-sulphide was associated to a number of gastrointestinal isolates which were characterised in pure culture (23, 207, 457), which suggests that metabolism of sulphur should be studied also from the perspective of protein utilisation.

In addition to the contribution to the digestion, the gastrointestinal microbiota takes part in the metabolism of several complex biomolecules. For example, it is involved in the metabolism of the sex steroid hormones. Steroid hormones are taken up by the liver, conjugated with glucuronic or sulphuric acid, and excreted via bile into the gastrointestinal tract. In the intestine, the microbiota is responsible for deconjugation of these esters and therefore it promotes their enterohepatic circulation (2). It is believed that suppression of the intestinal microbiota by antibiotics may induce significant reduction of sex hormones concentrations in the serum, which can explain the failure of oral contraceptive pills in combination with antibiotics (438). The gastrointestinal microbiota is also capable of metabolising bile (36) and cholesterol (124), although the microbiota-induced metabolic changes of those compounds lead to their excretion.

Other functions of the gastrointestinal microbiota include the production of several essential compounds, such as the vitamins K or B_{12} (4, 387). The microbiota also provides the first line of defence against invasion of exogenous pathogens or indigenous opportunists

(484). Additionally, the interaction between the microbiota and the host is an important factor in the development of the innate and adaptive immune responses (78).

A great number of other interactions between the host and the microbiota is to be expected, as it was shown that a single intestinal bacterium modulates the expression of important genes in a germ-free mouse model and improves its nutritional and defensive functions (210). Inclusion of different microbial species into the scope of host-microbe interaction studies showed profound communication between the host and intestinal microbes while the type of interaction was dependent on the phylogeny of the microbe (108, 450).

Although the gastrointestinal microbiota contributes to the host's health and wellbeing, there is increasing evidence that, when disturbed, it can play a role in either the aetiology or maintenance of several intestinal or other disorders. In this context, correlations have been suggested between the composition and activity of the microbiota and a variety of clinical statuses, of which some are listed in Table 2.

Table 2 Relation between the human gastrointestinal microbiota activity or diversity and several diseases or disorders.

Clinical status	Specific microbiota activity/ diversity	Reference
Acute Diarrhoea	Direct association with *Clostridium difficile, Campylobacter* spp., members of Enterobacteriaceae	(17)
Irritable Bowel Syndrome	Alterations of hydrogen-consuming bacteria; instability of microbiota, deviated abundance of some Firmicutes	(247, 308)
Crohn's Disease	Reduced *Clostridium leptum* group; increased Bacteroidetes and Proteobacteria and decreased Firmicutes	(177, 443)
Ulcerative Colitis	Reduced bacterial diversity; reduced abudance of *Clostridium coccoides* group, increased incidence of sulphate reducing bacteria	(170, 356, 443)
Colon cancer	Increased bacterial colonisation, and abundance of some *Clostridium* spp. and *Lactobacillus* spp.; deviated metabolic activity includes high sulphide production	(239)
Allergy	Decreased *Bifidobacterium* spp. and *Enterococcus* spp. and increased levels of *Clostridium* spp.	(248, 351)
Obesity	Low abundance of Bacteroidetes accompanied by an increase in Firmicutes	(275)
Autism	Higher incidence of *Clostridium histolyticum* group.	(361)

Due to the great inter-individual variation of the microbiota composition, and wide variety of the techniques that can be used for assessing the potential deviations of the microbiota's activity or diversity, the reports in the literature are diverse, and Table 2 summarises only a subset of the numerous findings. Ultimate question of this research is whether there is a causal relation between the presence of specific microbiota and the clinical symptoms and, if so, what mechanisms underlie the observed effects. Considering the impact of these issues, this area is being researched actively and partially explains the revival in the interest of the gastrointestinal microbiota.

Dimensions of the human gastrointestinal microbiota

After the respiratory tract, the human gastrointestinal tract is the second largest interface between the human body and the external environment. Its surface of approximately 300 square meters in adults is densely populated by the microbiota along the entire gastrointestinal tract, but it is in the colon where the majority of the microbiota resides (40). The density of the microbial community in the colon is approximately 10^{12} cells per gram of intestinal content (418), which makes it the most densely populated ecosystem reported (518). The collective microbial weight is estimated to be about 1.5 kg (531) and it is comprising about 30% of the total volume of the intestinal contents (333).

Bacteria absolutely dominate the gastrointestinal ecosystem, although they are accompanied by a viral community and archaeal and eukaryotic species (63, 150, 340). Thus, it is not a surprise that the first intestinal isolate was a bacterium – *Bacterium coli commune* — nowadays known as *Escherichia coli* (this and other highly relevant findings for the description of the microbial diversity within the human gastrointestinal tract are presented on Fig. 1). The detection of the microbial presence in the gastrointestinal tract triggered series of studies of the microbiota composition, and already in the beginning of the 20th century, the representative species of some dominant gastrointestinal groups, such as *Bacteroides* spp. or *Clostridium* spp., were isolated (362, 492). However, as the vast majority of the human gastrointestinal inhabitants are strict anaerobes, only the improved cultivation techniques and media applied during the 70s of the previous century, enabled better insight into the gastrointestinal microbiota's diversity (150, 333). Among numerous novel isolates that have been obtained with the improved cultivation techniques (207, 335), an archaeal species of the human intestine was detected (349).

Based on the findings of cultivation-based studies it was estimated that the human gastrointestinal tract harbours as many as 400-500 species (333). This perception did not change until a decade ago when molecular techniques were applied for the analysis of the gastrointestinal microbiota.

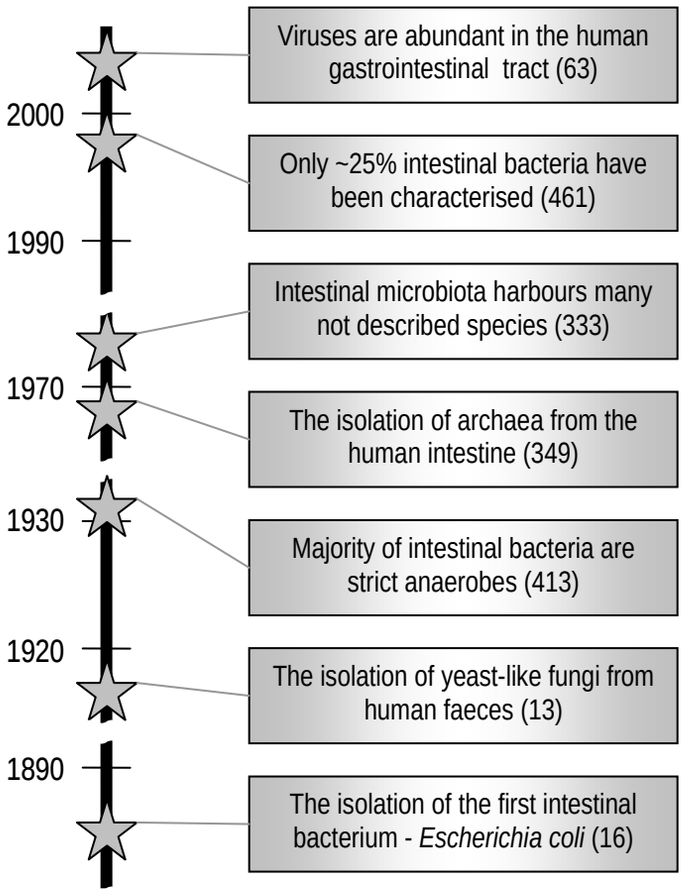


Figure 1 The time line that shows the most important discoveries concerning the composition of the human gastrointestinal microbiota, including the first isolation of archaeal, bacterial and eukaryotic representatives of this ecosystem.

Molecular, small subunit ribosomal RNA (SSU rRNA)-based approaches showed that only a small fraction of the intestinal microbes corresponds to the fully described microorganisms (461). During a decade of application, until 2005, molecular techniques revealed the presence of approximately 800 bacterial phylotypes (23). Since this is a very dynamic research area and each new study reports a proportion of novel species level types (phylotypes) (129), the diversity of this ecosystem certainly exceeds the value provided only two years ago. The complexity of the intestinal ecosystem rises further since the bacteria are complemented by an abundant viral community (63).

The numerous intestinal species belong to a limited number of taxonomic groups that seem to be highly adjusted to the conditions of their ecological niche (23). Therefore, many intestinal microbes are specific for the gastrointestinal tract, and when found in other ecosystems indicate faecal contaminations (261, 392, 400). Interestingly, with the improvements of the bacterial systematics, some gastrointestinal genera, *e.g. Bacteroides* and *Bifidobacterium*, have been distinguished from species that belong to other ecosystems but were previously assigned to the intestinal taxonomic groups (223, 431).

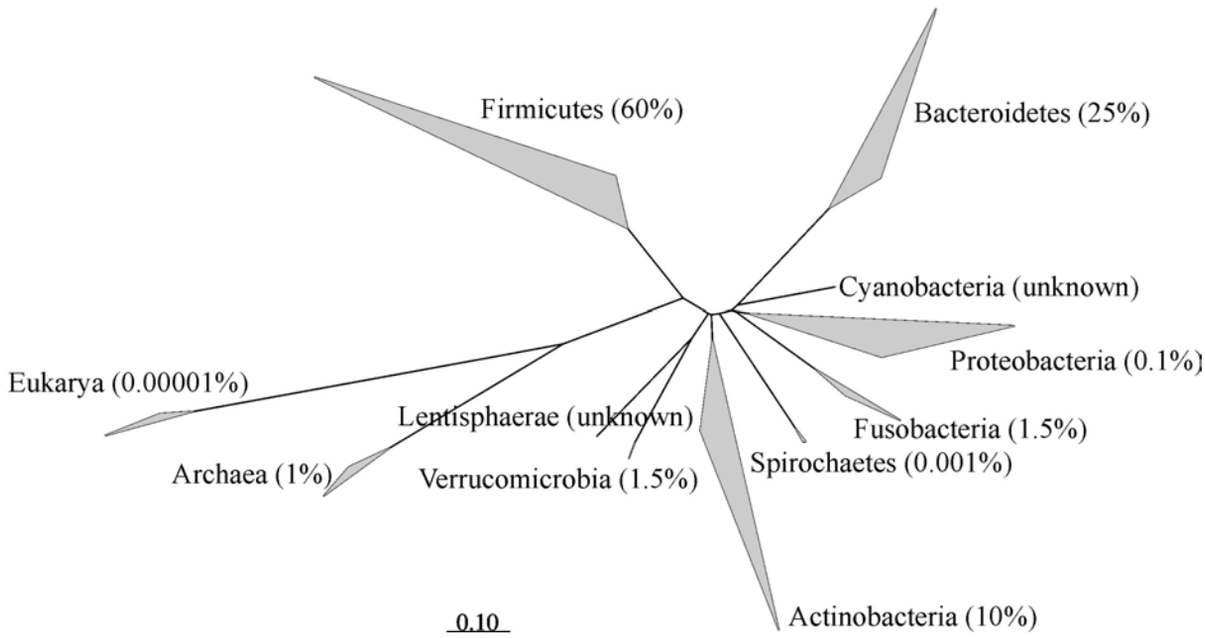


Figure 2 SSU rRNA-based phylogenetic tree representing the diversity of the human gastrointestinal microbiota. In parentheses, an estimated relative abundance for a healthy adult is indicated. The reference bar indicates 10% sequence divergence.

In total, members of nine bacterial phyla have been up to now found in the human gastrointestinal tract, while Archaea and Eukarya are represented by a single phylum (Fig. 2). Due to the large inter-individual variations in the microbiota composition, the quantitative contribution of different microbiota fractions can differ considerably between individuals (268). Despite those variations, it is clear that the most abundant intestinal inhabitants belong to only three phyla – Firmicutes, Bacteroidetes and Actinobacteria, while other significant phyla include Fusobacteria and Verrucomicrobia. The composition of the gastrointestinal microbiota was extrapolated on the basis of findings from several studies and illustrates the community in an average healthy adult (Fig. 2).

Techniques for studying the human gastrointestinal microbiota's diversity

The analysis of the human gastrointestinal microbiota, as of any other microbial ecosystem, has started with its cultivation. However, the microbial world is extremely complex, and as already observed by Pasteur, "similar forms can mask profound differences" (426). Therefore, the identification of the individual species in a complex ecosystem, such as the gastrointestinal microbiota, has been an extremely challenging task (334). Consequently, only few detailed cultivation-based studies of the gastrointestinal microbiota's diversity have been reported, while the majority of reports suffers from an incomplete identification of the

detected microbes that were systematised into groups based on their appearance or physiology (*e.g.*(543). In addition to the technical challenges which are related to the complexity of this ecosystem, the human gastrointestinal tract is a highly specific ecological niche, with a variety of interactions to different host tissues, gradients of oxygen, pH and solutes, and a continuous flux of food and host-produced components. These can not be mimicked in the laboratory settings, even with highly sophisticated equipment. Hence, cultivation-based analysis of the gastrointestinal microbiota is laborious, time-consuming, technically challenging. All this led to the omission of the vast majority of intestinal microbes in cultivation-based studies.

Alternative to the cultivation is the analysis of various biomarkers using so-called molecular techniques. Any biological component could serve as biomarker including different metabolites, proteins, RNA, DNA or even cells (337). The prime biomarker used for the phylogenetic analysis of complex ecosystems is the SSU rRNA gene. The following properties of this gene justify its wide use (526): (a) SSU rRNA has high degree of functional constancy, (b) it occurs in all organisms, (c) the change of its sequence is an indicator of phylogenetic relationship, (d) it can be sequenced directly, and (e) the SSU rRNA gene sequence databases provide the most complete information about the diversity of many ecosystems.

Several SSU rRNA-based techniques have enabled relative rapid profiling and quantification of the microbial community in the gastrointestinal tract (Table 3). Production of clone libraries and SSU rRNA gene sequencing is the principle SSU rRNA-based technique, as it provides considerable diagnostic information about the diversity of an ecosystem. Although this approach is limited by library biases, incomplete sequence coverage, sequence analysis of SSU rRNA gene libraries represents the bases for the development of other SSU rRNA-based techniques. Those include enumerative fluorescent in situ hybridisation (FISH), which can be coupled with microscopy or flow cytometry, quantitative real time PCR (qRT-PCR); and fingerprinting denaturant gradient gel electrophoresis (DGGE) and terminal-restriction fragment length polymorphism (T-RFLP). The principles and applications of those techniques have been previously summarised and reviewed (491). The application of these molecular techniques, during only a single decade, has revolutionised our view of the gastrointestinal microbiota (163). Some of the relevant findings consider the microbiota is more complex than observed with the use of cultivation approaches (461), that members of Firmicutes phylum dominate in this ecosystem (159, 537), and that the host's genotype strongly influences the microbiota composition (536).

Each of the previously mentioned techniques has contributed to better description of the gastrointestinal microbiota, and due to their specificities any technique can be a desirable choice for answering specific scientific questions (Table 3). Some of the most relevant advantages and limitations of the techniques presently available for studying human gastrointestinal microbiota listed in Table 3 may serve as an indicator for future studies. In

addition to the cultivation and well-established molecular techniques, Table 3 also lists the recently developed phylogenetic microarray systems for the gastrointestinal microbiota (359).

Table 3 Overview of the techniques available for studying the human gastrointestinal microbiota's diversity with indicated qualities and limitations. Table is adapted from (112, 538)

Technique	Qualities	Limitations
Cultivation	Identification and quantification; pure culture isolates available for functional analysis	Slow; laborious; selective for the subgroups that can grow on the developed media.
Sequence Analysis of SSU cDNA Clone Libraries	Identification and relative quantification; semi-automatic; the most detailed picture of the diversity.	Costly and affected by library bias, coverage and sensitivity limitations and misassembly
FISH of SSU rRNA	Absolute and relative quantification expressed in cell numbers; in combination with flow cytometry semi-automatic and allows cell sorting	Sequence knowledge is a prerequisite; available probe sets do not cover total gastrointestinal diversity; laborious and low throughput; differential permeability leads to biases.
qRT-PCR of SSU cDNA/DNA	Relative quantification; semi-automatic; highly sensitive	Sequence knowledge is prerequisite; available primers sets do not cover total gastrointestinal diversity; dependency on rRNA copy number
DGGE & T-RFLP of SSU rRNA gene amplicons	Rapid profiles of total microbial community or groups; semi-quantitative; rapid comparative analysis	No phylogenetic information unless when joined to clone libraries which makes analysis laborious and costly.
Phylogenetic microarrays	High throughput, rapid profiling, direct phylogenetic identification and relative quantification.	Sequence knowledge is prerequisite, limited sensitivity

The present generation phylogenetic microarrays is based on small fragments of synthetic single-stranded DNA on solid surfaces that are complementary to specific SSU RNA sequences. They enable rapid profiling of the gastrointestinal microbiota with phylogenetic positioning and semi-quantitative information for the detected organisms. The detection limit of such system can be as low as 0.1% of the total community, and this feature together with the fact that phylogenetic microarrays enable comprehensive and high throughput analyses, indicate a great potential of this technique for providing a detailed

insight into the composition of the gastrointestinal microbiota by rapid analysis of large amount of samples.

Particularly relevant advantage of the phylogenetic microrray analysis considers the high resolution of profiling of complex ecosystems, which is performed on the level of phylotype. Other, presently used SSU rRNA based techniques, principally target microbial groups at high phylogentic level (268, 306), and despite the fact that there are qPCR or FISH assays for detection of individual bacterial species (306), because of the complexity of the gastrointestinal microbiota, such analysis cannot be performed for comprehensive, high-throughput profiling of this ecosystem. This becomes particularly relevant since some, usually targeted gastrointestinal groups, such as *Clostridium leptum* or *Clostridium coccoides*, share as little as 85% of SSU rRNA gene sequence similarity. Therefore, it can be expected that members of these groups, once cultivated and described, will form a number of functionally different bacterial genera and families. The analysis on lower phylogenetic level will provide enhanced insights into the diversity of the gastrointestinal microbiota, as shown in the recently reported studies that have employed large scale analysis of the SSU rRNA gene clone libraries (52, 129, 275). However, this approach is very laborious and costly, which hampers its wide application for the analysis of the microbial diversity in relation to the different factors, such as health status, diet, age; and geographic or genetic origin.

The phylogenetic microarray-based analyses have a great potential for identifying important microbial groups and species in relation to the functionality of the gastrointestinal microbiota. However, as any other technique, for the complete description of the microbial diversity within the human gastrointestinal tract, the phylgenetic microarray analyses should be complemented with traditional cultivation-based and other molecular techniques. Furthermore, one should keep in mind that the criteria for differentiation between species in the prokaryotic world are rather broad. Therefore, different strains of the very same microbial species can exhibit substantially different functionality. An example is *E. coli* that contains number of pathogenic, diarrhoea causing strains (400), while another strain is used as probiotic in treatment of prematurely born neonates (280). This indicates that community structure analysis alone would not provide sufficient information about the gastrointestinal microbiota if not coupled with functional analysis and stresses the need for development of novel creative approaches for studying the exceptionally complex and important ecosystem in the gastrointestinal tract.

CHAPTER 2

Mirjana Rajilić-Stojanović and Willem M. de Vos

Manuscript in preparation

Characterised Microbes of the Human Gastrointestinal Tract

The majority of the human gastrointestinal microbes are highly adjusted to the living conditions of this ecosystem, and therefore difficult to obtain in pure culture in the laboratory settings. This, and the fact that the gastrointestinal microbiota is extremely complex and subject-specific ecosystem make its description particularly challenging. Hence, even after a century of studying, the microbial diversity within the human gastrointestinal tract remains only partially described. In the view of the rapidly growing interest in the human microbiota and its function an overview of its diversity is a prerequisite. Here we provide a systematic overview of the fully described gastrointestinal microbes, which were obtained since the first isolation of *Escherichia coli*, with the intention to identify the dominant and abundant gastrointestinal groups. The presented thorough analysis of over 400 characterised gastrointestinal microbes through a historical perspective is coupled to their phylogenetic analysis, which is performed on the basis of the small subunit ribosomal RNA (SSU rRNA) sequence. Based on the statistical predictions a considerable number of novel gastrointestinal isolates can be expected to arise from the application of traditional cultivation techniques alone. However, to enable the complete description of this ecosystem novel cultivation approaches should continue to be developed.

Introduction

Human beings, similarly to other higher organisms, live in symbiosis with co-evolved microbiota (23). The vast majority of the human microbiota resides in the gastrointestinal tract, where, besides contributing to the digestion, performs several other functions that are essential for the human host. The mechanisms of the microbiota's interaction with its host and the microbiota's impact on health and disease are beginning to be understood. However, because of the specificities of the gastrointestinal tract and the difficulties to mimic the conditions of this ecological niche in laboratory settings, the microbiota remains moderately described even a century after its discovery.

The gastrointestinal microbiota is composed of members of all three domains of life – Bacteria, Archaea and Eukarya – of which Bacteria by far dominate. During the years of studying of this ecosystem, the perception of the diversity of the gastrointestinal microbiota has gradually changed. Until the 1970s, bifidobacteria and bacteroides were considered to be the dominant groups in human intestinal contents. Aerobes, referred as coliforms, streptococci and lactobacilli, were found as minor groups, while clostridia, staphylococci and aerobic spore-formers were reported as rare and not always detectable (185). Currently the following cultivated groups are considered as predominant in the human gastrointestinal tract: *Bacteroides, Clostridium, Eubacterium, Veillonella, Ruminococcus, Bifidobacterium, Fusobacterium, Lactobacillus, Peptostreptococcus* and *Peptococcus* (333, 531). However, recent studies, especially those based on the SSU rRNA gene sequence analysis; provide strong evidence that even this representation of the gastrointestinal composition is neither truthful nor complete.

In view of the rapidly growing interest in the human microbiota and its function (274, 353) an overview of its diversity is a prerequisite. Here we provide a thorough analysis of phylogenetic groups of microbes that were isolated from human gastrointestinal samples through a historical perspective with the aim to clarify their importance. In addition, information about other groups, including clinical isolates, is provided. Finally, the known species of all groups are listed, positioned into a curated phylogenetic framework based on SSU rRNA sequence information, and connected to the reference of their isolation.

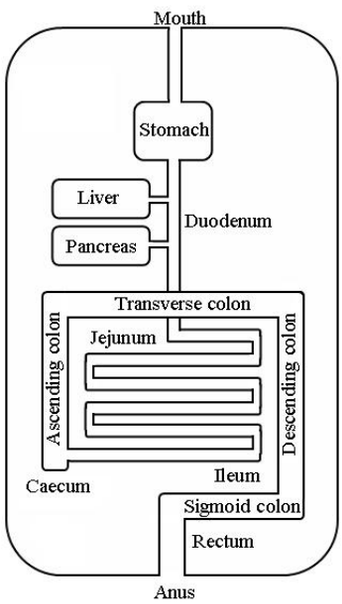
General remarks on the gastrointestinal microbiota

It is generally accepted that gastrointestinal microbiota is established at birth, when the virtually sterile neonate is colonised by microbes, usually derived from the mother's microbiota or the environment. The composition and the development of the microbiota depends on a variety of factors including the mode of delivery, the feeding and the therapeutic treatments that baby receives (143, 146, 181). In the first months of life the gastrointestinal

microbiota is dramatically changing, its complexity is increasing, and usually at the age of around two it reaches adult-like composition (255).

The composition of the gastrointestinal microbiota changes from the stomach to the colon and both density and diversity are increasing along the same axes (Table 1). The stomach and the proximal small intestine have relatively low numbers of microbes due to the low pH and the rapid flow of intestinal contents in this region. Because of the slower passage, colonisation intensity increases in the ileum and the microbial community in the colon reaches a climax with 10^{11} cells per gram of intestinal content. This makes the intestine of a single individual so densely populated that its microbiota outnumbers the total human population of the world by a factor of 1,000 (334).

Table 1 Approximate dimensions of the human gastrointestinal tract, passage time of intestinal contents and corresponding microbiota density (160, 183, 228, 333).



Gastrointestinal segment	Length, cm	Passage time, h	Density of microbiota, cells/ml(g)
Stomach	12	2-6	10^0-10^4
Duodenum	25		10^4-10^5
Jejunum	160	3-5	10^5-10^7
Ileum	215		10^7-10^8
Caecum	6		
Ascending colon	15		
Transverse colon	50	10-20	10^{10}-10^{11}
Descending colon	25		
Sigmoid colon	40		
Rectum	18	1	10^{10}-10^{11}

The gastrointestinal microbiota is apparently stable over time (537), although it can be altered by different factors, such as aging, disease, antibiotic treatment, diet or stress (174, 325). Furthermore, it has been shown that genetic background is influencing the microbiota composition (536). Since each individual has a unique genetic background, inhabits a specific environment, and usually follows an individual diet, the intestinal microbiota is consequently host-specific. In fact, it is unique for each subject and therefore can be considered as kind of fingerprint, which can even be used for forensic analysis.

The diversity of the gastrointestinal microbiota

Historical perspective

The characterisation of the gastrointestinal microbiota composition started in 1885, when Teodor Escherich isolated *Bacterium coli commune* from children's faeces, a bacterium which was later renamed to *Escherichia coli* (16). Since then, numerous studies aiming to describe this ecosystem have been performed. In retrospect, three periods in which there was an increased scientific interest in the gastrointestinal microbiota can be recognised.

The initial period in which the human gastrointestinal microbiota was extensively studied followed the isolation of *E. coli* and led to the description of the representatives of the major intestinal bacterial groups (116, 155, 338, 362, 477, 478, 492). However, the vast majority of the intestinal microbes are strict anaerobes, as it was for the first time shown in 1931 (413). Thus, the early cultivation procedures enabled isolation of only 10-25 percent of the gastrointestinal microbes (149).

With the improvement of the anaerobic cultivation techniques by Hungate (220) the gastrointestinal microbiota was again recognised as an important and yet unexplored ecosystem. The novel techniques allowed cultivation of up to 88% of the total microscopic counts in faeces (333). However, many of the obtained isolates were not characterised below the genus level (45, 150, 333). This is mainly due to the extreme complexity of the ecosystem, as when processing a single sample an enormous amount of different isolates is obtained, for which it was physically difficult to make a full description (334). Hence, even nowadays, novel intestinal isolates are obtained by the application of traditional cultivation techniques (27, 328, 444).

Finally, the third revival of the scientific interest in the gastrointestinal microbiota came with the introduction of molecular 16S rRNA-based techniques. Application of molecular techniques revealed that the microbiota was even more complex than previously anticipated (461). The proportion of the reported uncultivated bacterial phylotypes varies between studies, and it ranges from 80% for the mucosal samples of healthy individuals (129) to 54% for the elderly Japanese subjects (193).

Analysis of the diversity of the human gastrointestinal microbiota

Although the human gastrointestinal microbiota has been studied for over a century, the constant reports of novel intestinal inhabitants indicate that a complete description has not yet been achieved. Based on an extensive literature search, we have identified 442 bacterial, 3 archaeal, and 17 eukaryotic isolates of human gastrointestinal origin. Therefore, the number of known, fully characterised gastrointestinal inhabitants is within the range of the most frequently cited diversity prediction of 400-500 species (333). The reason for overseeing the fact that presently known gastrointestinal diversity has already reached its estimate is that the

description of the gastrointestinal microbiota is going on for very long period of time and the results of the different studies have not been previously integrated in a systematic way.

Integration of results obtained in different studies is confronted with many difficulties even with currently available systematics framework for cultivated organisms (453, 510). One of the major inconveniences derives from the dramatic taxonomic changes that have occurred since the first isolation of *E. coli*. Some species have been reported under many different names, *e.g.* 26 synonyms of *Bacteroides fragilis* can be found in the NCBI taxonomy browser. Other species appeared to be heterogeneous, such as the *Lactobacillus acidophilus* group, which actually contains 6 distinct species (265). As this was revealed 80 years after the initial isolation of this organism, it is impossible to reconstruct, which species were actually reported in the preceding studies. Finally, some of the isolates cannot be traced in the presently available databases. In total 33 fully described human gastrointestinal isolates cannot be found in the present databases, of which the majority was reported in three studies from the first half of the 20th century (33, 133, 134). The most probable reason for losing those organisms is that they have not been included in the Approved List of Bacterial Names, as they were isolated prior to the establishment of this list. In this way species such as *Bacillus fusus* or *Bacillus tritus* (33) have disappeared from the current databases, while other isolates, obtained in the very same study, were revived in more recent publications (344). However, the reason for the loss of certain isolates, such as *Lactobacillus rogosae* (147), remains unclear.

Taxonomic analysis of the human gastrointestinal microbes showed that the most of the gastrointestinal isolates could be classified within the abundant phylum of Firmicutes (Table 2). Bacteroidetes and Actinobacteria are also abundant, but considerably less diverse. In addition, many gastrointestinal isolates belong to the scarce phylum of Proteobacteria. It should be noted that the taxonomic position of the 33 organisms that are not systematised in the current databases could not be reliably determined.

Phylogenetic analysis of the gastrointestinal isolates based on the SSU rRNA gene sequence could not be assessed for 46 bacterial and 7 eukaryotic species as they lack the SSU rRNA gene sequence. Phylogenetic trees, based on the SSU rRNA gene sequence, presented in this review were constructed using Parsimony procedures as implemented in ARB software package (286). The ARB database release 2002 was used, and the sequences that were not present in the database were added using appropriate filters that fit sequence phylogeny and length. The SSU rRNA gene-based phylogenetic analysis indicated probable misclassification of some isolates (see below). Although number of human gastrointestinal isolates could not be integrated in the phylogenetic analysis, all isolates were included in the overview of the human gastrointestinal diversity. Microbes that are absent from the current databases were presented in gray-filled cells when listed in the corresponding Tables (see below).

Table 2 The distribution of the human gastrointestinal isolates among different phylogenetic groups (phylum, class, order, and family) expressed as cumulative number of fully chracterised isolates. Average detected abundance per gram of intestinal content of memebers of each phylum is given in bold.

Phylum		Class		Order		Family	
Actinobacteria (**~10^{10}**)	32	Actinobacteria	32	Actinomycetales	14	Actinomycetaceae	2
						Corynebacteriaceae	4
						Micrococcaceae	2
						Propionibacteriaceae	6
				Bifidobacteriales	12	Bifidobacteriaceae	12
				Coriobacteriales	6	Coriobacteriaceae	6
Bacteroidetes (**~10^{11}**)	35	Bacteroidetes	35	Bacteroidales	35	Bacteroidaceae	20
						Porphyromonadaceae	3
						Prevotellaceae	8
						Rikenellaceae	4
Firmicutes (**~10^{11}**)	214	Bacilli	81	Bacillales	16	Bacillaceae	9
						Paenibacillaceae	3
						Staphylococcaceae	4
				Lactobacillales	65	Aerococcaceae	1
						Lactobacillaceae	33
						Leuconostocaceae	3
						Streptococcaceae	28
		Clostridia	131	Clostrdiales	131	Acidaminococcaceae	9
						Clostridiaceae	68
						Eubacteriaceae	19
						Lachnospiraceae	26
						Peptococcaceae	1
						Peptostreptococcaceae	8
		Mollicutes	2	*Incertae sedis* 8	2	Erysipelotrichaceae	2
Fusobacteria (**~10^{10}**)	10	Fusobacteria	10	Fusobacteriales	10	Fusobacteriaceae	9
						Incertae sedis 11	1
Lentisphaerae (**~10^{8}**)	1	Lentisphaerae	1	Victivallales	1	Victivallaceae	1
Proteobacteria (**~10^{8}**)	104	α-Proteobacteria	1	Rhizobiales	1	Hyphomicrobiaceae	1
		β-Proteobacteria	5	Burkholderiales	5	Alcaligenaceae	4
						Oxalobacteriaceae	1
		γ-Proteobacteria	73	Aeromondales	4	Aeromonadaceae	2
						Succinivibrionaceae	2
				Enterobacteriales	52	Enterobacteriaceae	52
				Pasteurellales	2	Pasteurellaceae	2
				Pseudomonadales	8	Moraxellaceae	4
						Pseudomonadaceae	4
				Vibrionales	4	Vibrionaceae	4
		δ-Proteobacteria	7	Desulfovibrionales	7	Desulfovibrionaceae	7
		ε-Proteobacteria	18	Campylobacterales	18	Campylobacteraceae	12
						Helicobacteraceae	6
Spirochaetes (**~10^{6}**)	2	Spirocaetes	2	Spirochaetales	2	Serpulinaceae	2
Verrucomicrobia (**~10^{9}**)	1	Verrucomicrobia	1	Verrucomicrobiales	1	Verrucomicrobiaceae	1
Euryarchaeota (**~10^{9}**)	3	Methanobacteria	3	Methanobacteriales	3	Methanobacteriaceae	3
Ascomycota (~10^{4})	14	Eurotiomycetes	7	Eurotiales	7	Trichocomaceae	7
		Saccharomycetes	7	Saccharomycetales	7	Saccharomycetaceae	7

Bifidobacteria

Bifidobacteria represent a dominant fraction of the human intestinal microbiota, which is the most pronounced in the intestine of infants (43). Bifidobacteria can reach densities of up to 10^{10} cells per gram of intestinal content (45, 150, 153, 333). The majority of bifidobacteria were recovered exclusively from human or animal intestinal samples, while for three species, which were isolated from sewage and an anaerobic digester (50), an intestinal origin can be suspected. Therefore, this genus appears to be specific for the gastrointestinal tract.

Phylogenetically bifidobacteria are a homogenous group, with 11 fully characterised species linked to the human gastrointestinal tract (see Table 3). A bifidobacterial species was for the first time recovered from infant faeces in 1900 by Tissier, as a part of his Ph.D. thesis work (477). It was named *Bacillus bifidus-communis.* Already in 1924 this bacterium was renamed to *Bifidobacterium bifidum*, but *Bifidobacterium* was not recognised as an independent genus until 1974 (50).

Table 3 *Bifidobacterium* species isolated from the human gastrointestinal tract.

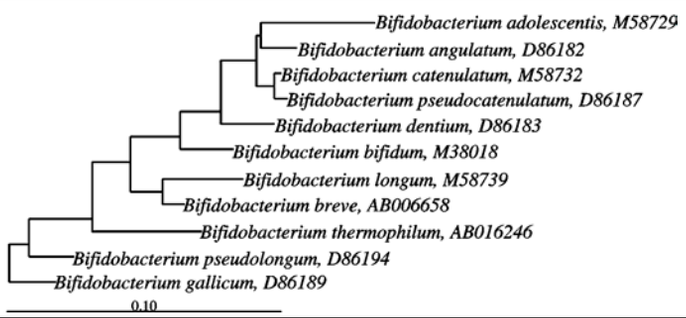
Species	Reference	Phylogenetic position
B. adolescentis	(390)	
B. angulatum	(419)	
B. catenulatum	(419)	
B. pseudocatenulatum	(420)	
B. dentium	(419)	
B. bifidum	(477)	
B. longum	(390)	
B. breve	(390)	
B. thermophilum	(150)	
B. pseudolongum	(333)	
B. gallicum	(264)	
B. eriksonii	(150)	Unknown, no 16S rRNA gene sequence available
B. cornutm	(123)	

Table lists species name and the reference that provides the link with the human gastrointestinal tract as ecological niche.

Bacteroides

Bacteroides represent a diverse and abundant group of the human gastrointestinal bacteria that can reach densities of up to 10^{11} cells per gram of intestinal material (45, 134, 333). Forty-eight *Bacteroides* and *Bacteroides*-like species of the human gastrointestinal tract have been, until now, fully described (Table 4). The first *Bacteroides* species – *Bacteroides fragilis* was isolated in 1898 as a human pathogen linked to appendicitis among other clinical cases (492).

Table 4 *Bacteroide*s and *Bacteroides*- like species isolated from the human gastrointestinal tract.

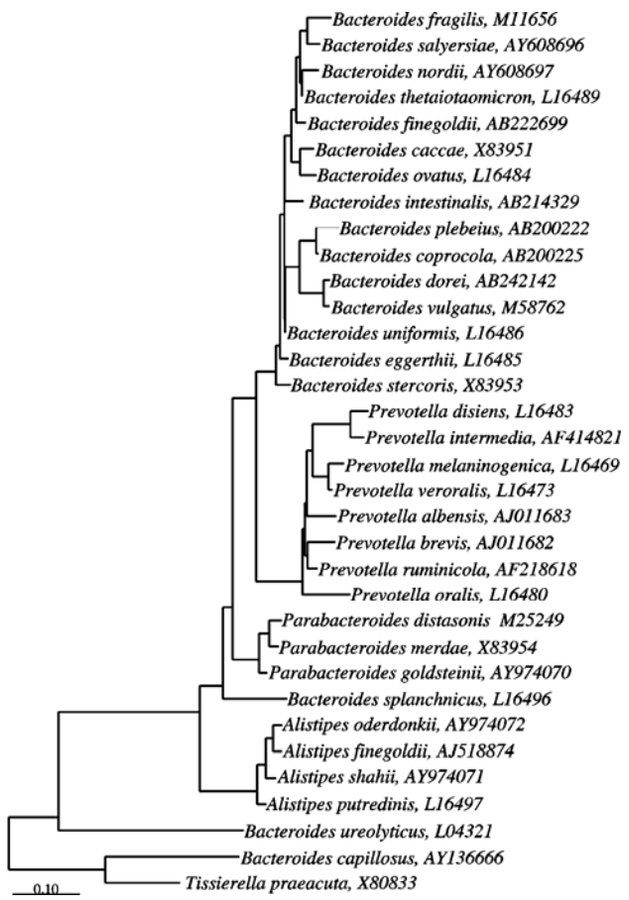
Species	Reference	Genus	Phylogenetic position
B. fragilis	(492)	*Bacteroides*	
B. salyersiae	(448)		
B. nordii	(448)		
B. thetaiotaomicron	(116)		
B. finegoldii	(25)		
B. caccae	(225)		
B. ovatus	(134)		
B. intestinalis	(26)		
B. plebius	(249)		
B. coprocola	(249)		
B. dorei	(27)		
B. vulgatus	(134)		
B. uniformis	(134)		
B. eggerthii	(207)		
B. stercoris	(225)		
P. disiens	(387)	*Prevotella*	
P. intermedia	(150, 431)		
P. melaninogenica	(325, 431)		
P. veroralis	(45)		
P. albensis	(22, 150)		
P. brevis	(22, 206)		
P. ruminicola	(22, 206)		
P. oralis	(150, 431)		
P. distasonis	(134, 406)	*Parabacteroides*	
P. merdae	(225, 406)		
P. goldsteinii	(406, 445)		
B. splanchnicus	(514)	*Bacteroides*	
A. oderdonkii	(444)	*Alistipes*	
A. finegoldii	(389)		
A. shahii	(444)		
A. putredinis	(389)		
B. ureolyticus	(42)	*Bacteroides*	
B. capillosus	(333)	*Bacteroides*	
T. praeacuta	(91, 477)	*Tissierella*	
B. bullosus	(116)	*Bacteroides*	Unknown, no 16S rRNA gene sequence available
B. pectinophilus	(222)		
B. galacturonicus	(222)		
B. gulosus	(134)		
B. variabilis	(134)		
B. tumidus	(134)		
B. vescus	(134)		
B. convexus	(134)		
B. exiguous	(134)		
B. insolitus	(134)		
B. siccus	(134)		
B. coagulans	(134)		
P. bivia	(42)	*Prevotella*	
A. furcosa	(333, 430)	*Anaerorhabdus*	

Table lists species name, the reference that provides link with the human gastrointestinal tract as ecological niche and the genus in which each species is classified.

The most comprehensive culture-based study aiming at the description of intestinal bacteroides was performed in 1933, when 18 *Bacteroides* spp. were isolated from 65 stool samples (134). Only five of the reported species are nowadays recognised as distinct species as four have been renamed and for nine there is no information in the publicly available databases. Recently, eight novel intestinal *Bacteroides* spp. were described (25, 26, 249, 389,

445, 448). Four of them derive from a single, yet detailed study of the microbiota composition of three healthy individuals (191). The reason that recent studies describe species that have not been previously reported might be that subjects analysed harbour different *Bacteroides* spp., but also that different criteria were applied for the definition of species, as discussed above.

Some of the species that were once described as *Bacteroides* spp. were reclassified into distant phylogenetic groups, but the majority is distributed over four families of Bacteroidetes phylum namely Bacteroidaceae, Prevotellaceae, Rikenellaceae, and Porphyromonadaceae (91, 389, 431, 445). Hence, it is more appropriate to refer to those organisms as members of phylum Bacteroidetes, which is also often termed as the Bacteroides-Cytophaga-Flavobacterium group. Since some of the *Bacteroides* spp. form deep branches in the 16S rRNA gene based phylogenetic tree (see Table 4) further reclassification within this genus can be expected.

Lactobacilli

In 1900, a *Lactobacillus* species was for the first time isolated from the human intestine and it was named *Bacillus acidophilus* (338). The description of this species is vague, based on the currently accepted standards, and as the strain of Moro was lost, it is not clear if the first intestinal *Lactobacillus* spp. was indeed *L. acidophilus* or one of the other 5 species derived from the so-called *L. acidophilus* group (324).

Lactobacilli are group of intestinal inhabitants that has received particular scientific attention (469), mainly because of the health claims proposed by Metchnikoff in the beginning of the previous century (316). Specific media, developed already in the 1950s (399), enabled isolation and description of many gastrointestinal lactobacilli. Lactobacilli are a minor fraction of the microbiota in faeces where they can reach counts of up to 10^8 cells per gram (437), while in the small intestine of humans they represent one of the predominant groups (391). The most recent report of four novel gastric *Lactobacillus* spp. (401) indicates that the 32 known gastrointestinal members of *Lactobacillus* genus (Table 5) are probably not covering the group's full diversity. There has been no major reclassification of the lactobacilli, although, based on the 16S rRNA gene sequence, the genus appears to be rather heterogeneous. Only two species previously described as belonging to the *Lactobacillus* genus are nowadays assigned to other genera, namely *Weissella* and *Atopobium* (53, 92). It should be noted that *Atobium* spp. are distantly releated to *Lactobacillus* spp. as they belong to the Actinobacteria phylum. Furthermore, *L. catenaformis* and *L. vitulinus* form a separate phylogenetic cluster within Firmicutes phylum - cluster XVII (87), which is distant from the rest of the genus, as it does not even cluster within the Bacilli class.

Table 5 *Lactobacillus* and *Lactobacillus*-like species isolated from the human gastrointestinal tract.

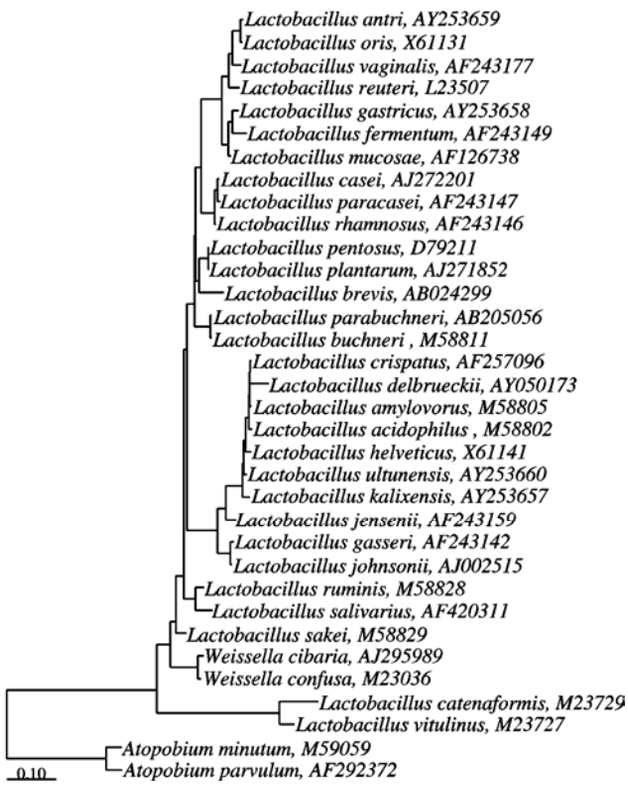
Species	Reference	Genus	Phylogenetic position
L. antri	(401)	*Lactobacillus*	
L. oris	(391)		
L. vaginalis	(100)		
L. reuteri	(42)		
L. gastricus	(401)		
L. fermentum	(333)		
L. mucosae	(104)		
L. casei	(333)		
L. paracasei	(100)		
L. rhamnosus	(100)		
L. pentosus	(100)		
L. plantarum	(150)		
L. brevis	(42)		
L. parabuchneri	(100)		
L. buchneri	(326)		
L. crispatus	(153)		
L. delbrueckii	(150)		
L. amylovorus	(323)		
L. acidophilus	(338)		
L. helveticus	(153)		
L. ultunensis	(401)		
L. kalixensis	(401)		
L. jensenii	(72)		
L. gasseri	(42)		
L. johnsonii	(391)		
L. ruminis	(272)		
L. salivarius	(333)		
L. sakei	(100)		
W. cibaria	(53)	*Weissella*	
W. confusa	(53)		
L. catenaformis	(133)	*Lactobacillus*	
L. vitulinus	(324)		
A. minutum	(92, 333)	*Atopobium*	
A. parvulum	(92, 150)		
L. leichmanii	(333)	*Lactobacillus*	Unknown,
L. rogosae	(207)		no 16S rRNA gene sequence available

Table lists species name of bacteria that were either initially reported as *Lactobacilli* or reclassified into this genus, with an exception of *Atopobium parvulum*, which was initially recognised as a *Streptococcus* (92). In addition, the reference that links each species with the human gastrointestinal tract as ecological niche, and the genus in which each species is classified is provided.

Streptococci and enterococci

Streptococci and enterococci have only recently been classified into two distinct genera, although the presence of the subgroup within the genus *Streptococcus* was noticed in the 1930s (433). They are one of the dominant bacterial fractions in the upper part of the small intestine (391, 438). Twenty-eight species of these two genera are known to be gastrointestinal inhabitants (Table 6). Such an extensive representation of this group can be explained by the fact that the species are oxygen tolerant and easily cultivable. The oldest isolate of the group, *Enterococcus faecalis*, was for the first time plated in 1899 from a case of endocarditis (287), and only seven years later it was recovered from intestinal samples (15).

Table 6 *Streptococcus* and *Enterococcus* species isolated from the human gastrointestinal tract.

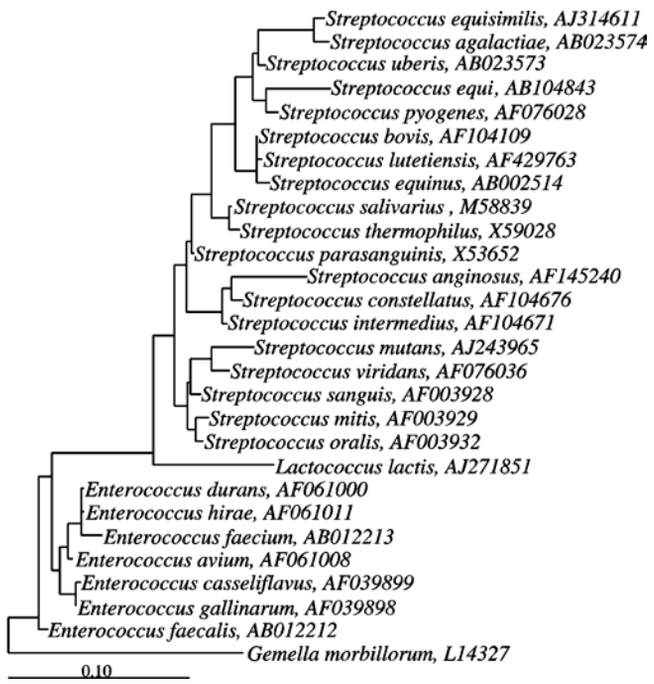
Species	Reference	Genus	Phylogenetic position
S. equisimilis	(150)	*Streptococcus*	
S. agalactiae	(150)		
S. uberis	(150)		
S. equi	(150)		
S. pyogenes	(150)		
S. bovis	(150)		
S. infantarius	(424)		
S. lutetiensis	(379, 424)		
S. equinius	(150)		
S. salivarius	(333)		
S. thermophilus	(150)		
S. parasanguinis	(517)		
S. anginosus	(516)		
S. constellatus	(207)		
S. intermedius	(207)		
S. mutans	(153)		
S. viridans	(123)		
S. sanguis	(150)		
S. mitis	(150)		
S. oralis	(24)		
L. lactis	(150)	*Lactococcus*	
E. durans	(10)	*Enterococcus*	
E. hirae	(394)		
E. faecium	(150)		
E. avium	(278)		
E. casseliflavus	(166)		
E. gallinarum	(89)		
E. faecalis	(15)		
G. morbillorum	(207, 246)	*Gemella*	
S. infrequens	(10)	*Streptococcus*	Unknown, no 16S rRNA gene sequence available

Table lists species name, the reference that provides the link with the human gastrointestinal tract as ecological niche, and genus in which each species is classified.

Veillonellae

The first record of a *Veillonella* spp. dates from 1898 when Veillon and Zuber isolated *Staphylococcus parvulus* from infected appendix tissue (492). *S. parvulus* was later reclassified as *Veillonella parvula*. *Veillonella* is a relatively small genus which members are typically isolated from the oral cavity, but are also frequently isolated from the proximal small bowel (397, 438). The genus consists of seven species, of which three, namely *V. atypica*, *V. dispar*, and *V. parvula*, were isolated from the lower gastrointestinal tract. *V. parvula* is regularly isolated, even in high densities of up to 10^{11} cells per gram of faeces (153). The fact that there are other related species, which are frequently isolated from the human intestine, and which used to be classified within the family Veillonellaceae (398) probably contributed to the recognition of the genus as dominant. Those species, also including veillonellae, have nowadays been reclassified into the family Acidaminococcaceae and are listed in Table 7 (165).

Table 7 The human gastrointestinal isolates that belong to the Acidaminococcaceae family.

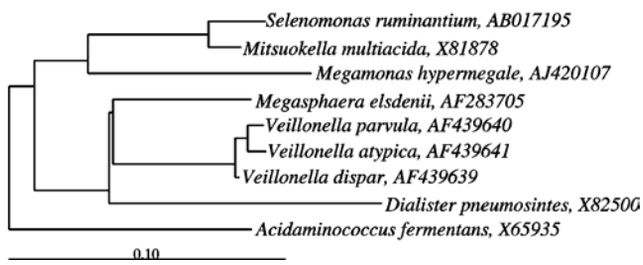
Species	Reference	Genus	Phylogenetic position
S. ruminantium	(45)	*Selenomonas*	
M. multiacida	(327, 429)	*Mitsuokella*	
M. hypermegale	(333, 428)	*Megamonas*	
M. elsdenii	(150)	*Megasphaera*	
V. parvula	(492)	*Veillonella*	
V. dispar	(100)		
V. atypica	(100)		
D. pneumosintes	(150)	*Dialister*	
A. fermentans	(333)	*Acidaminococcus*	

Table lists species name, the reference that provides link with the human gastrointestinal tract as ecological niche, and the genus in which each species is classified.

Clostridia

Clostridium spp. are very diverse group of the human gastrointestinal bacteria which can be found in densities up to 10^{10} cells per gram of intestinal content (43, 150, 333). The vague definition of the genus, which groups all Gram-positive, rod-shaped, spore-forming anaerobic bacteria, resulted in the assembly of a range of highly heterogeneous species. Based on 16S rRNA gene sequence analysis, *Clostridium* spp. are distributed over ten distinct phylogenetic clusters (87, 454). None of the clusters is composed exclusively of *Clostridium* spp. although *Clostridium* cluster I is considered as the core of the genus as it groups the majority of the described clostridia (Table 8) (454).

Table 8 *Clostridium* species isolated from the human gastrointestinal tract that based on the 16S rRNA gene sequence belong to *Clostridium* cluster I.

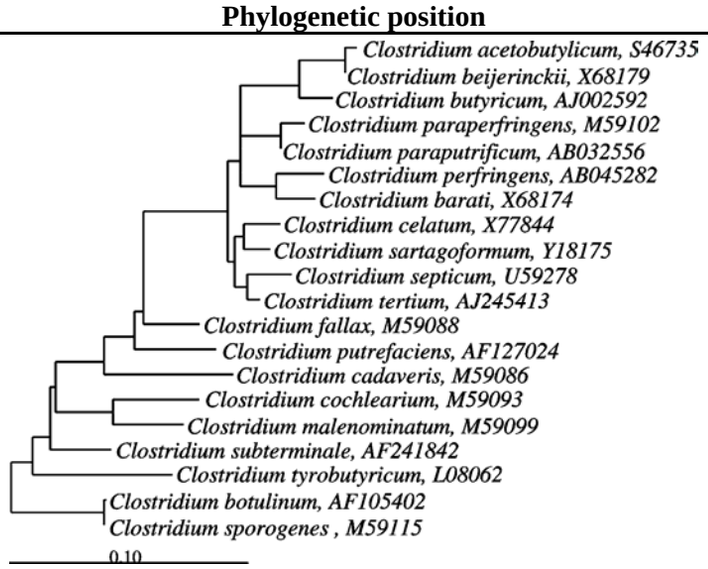
Species	Reference	Phylogenetic position
C. acetobutylicum	(473)	
C. beijerinckii	(153)	
C. butyricum	(150)	
C. paraperfringens	(153)	
C. paraputrificum	(150)	
C. barati	(150)	
C. perfringens	(362)	
C. celatum	(190)	
C. sartagoforme	(150)	
C. septicum	(150)	
C. tertium	(42)	
C. fallax	(150)	
C. putrefaciens	(150)	
C. cadaveris	(325)	
C. cochlearium	(150)	
C. malenominatum	(150)	
C. subterminale	(473)	
C. tyrobutyricum	(473)	
C. botulinum	(150)	
C. sporogenes	(325)	

Table lists species name and the reference that provides link with the human gastrointestinal tract as ecological niche.

Even the *Clostridium* cluster I is not consisting of only *Clostridium* species as it also contains several *Eubacterium* spp. and *Sarcina ventriculi*, which is the type species of *Sarcina* genus (520). The most heterogeneous *Clostridium* cluster is the cluster XIVa, as its members are distributed over seven gastrointestinal genera: *Clostridium*, *Coprococcus*, *Ruminococcus*, *Eubacterium*, *Dorea*, *Roseburia* and *Anaerostipes* (see below). Although so far only one gastrointestinal *Clostridium* spp. was reclassified into genus of *Paenibacillus* (87), it is reasonable to expect major reclassification of the genus *Clostridium*.

Table 9 *Clostridium* species isolated from the human gastrointestinal that based on the 16S rRNA gene sequence belong to other *Clostridium* clusters but *Clostridium* cluster I (Table 8).

Species	Reference	Cluster/Genus	Phylogenetic position
C. hylemonae	(250)		
C. scindens	(250)		
C. glycyrrhizinlyticum	(408)		
C. nexile	(207)		
C. oroticum	(150)		
C. coccoides	(45)		
C. bolteae	(447)	Cluster XIVa	
C. clostridiiformes	(242)		
C. aspargiforme	(328)		
C. symbiosum	(242)		
C. hathewayi	(458)		
C. sphenoides	(150)		
C. indolis	(150)		
C. aminovalericum	(150)		
C. lactatifermentans	(446)		
C. difficile	(150)		
C. hiranonis	(251)		
C. irregularis	(150)		
C. bartlettii	(449)		
C. glycolicum	(150)		
C. biferementans	(150)	Cluster XI	
C. ghoni	(150)		
C. sordellii	(150)		
C. sticklandii	(45)		
C. felsineum	(123)		
C. leptum	(335)		
C. sporosphaeroides	(333)	Cluster IV	
C. methylpentosum	(203)		
C. orbiscindens	(525)		
C. limosum	(133)	Cluster II	
C. cocleatum	(241)		
C. ramosum	(333)	Cluster XVIII	
C. spiroforme	(241)		
C. innocuum	(150)	Cluster XVI	
P. durus	(45, 87)	*Paenibacillus*	
C. rectum	(43)	Cluster XIX	
C. cellobioparum	(42)	*Clostridium*	Unknown, no 16S rRNA gene sequence available
C. inulinum	(150)		

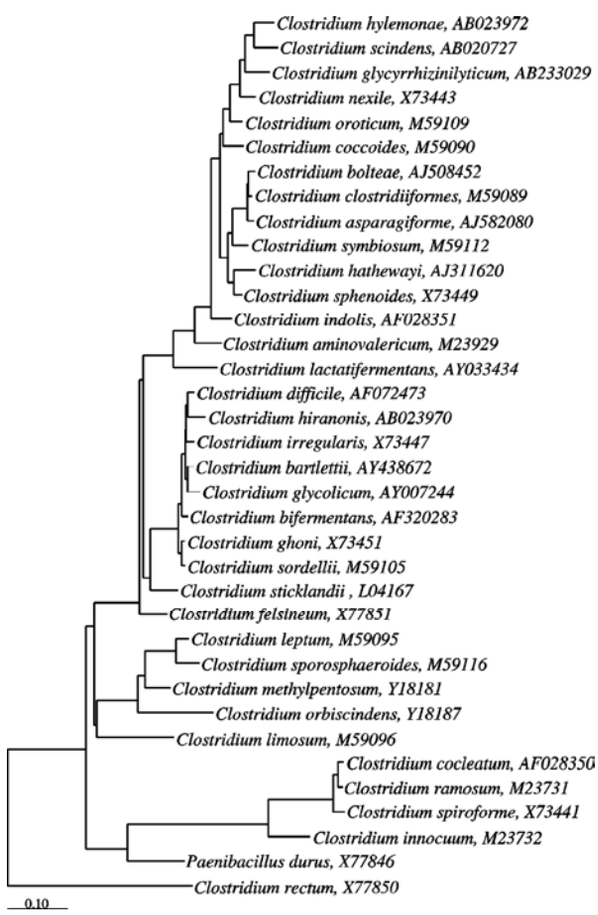


Table lists species name, the reference that provides link with the human gastrointestinal tract as ecological niche, and the *Clostridium* cluster in which, based on the 16S rRNA gene sequence, species belongs (87, 454).

The first *Clostridium* isolate of the human gastrointestinal origin was *C. perfringens*, which was recovered in 1905 (362). The same species, previously known as *Bacillus*

aerogenes capsulatus and *Clostridium welchii*, was earlier isolated from a case of endocarditis in 1892 by Welch *et al.* (513). Both isolation sites fit with the nowadays known properties of *C. perfringens*, which is a commensal gastrointestinal bacterium that can cause bacteraemia out of its normal ecological niche (368). Up to now, 58 *Clostridium* spp. have been isolated from the human intestinal samples, and this makes it the most diverse group of bacteria in the human gastrointestinal tract (Table 9).

Eubacteria

Eubacteria are another abundant group of bacteria from the human intestine with densities of up to 10^{10} cells per gram of intestinal content (333). Genus *Eubacterium*, similarly to *Clostridium*, is genus that is very vaguely described. Defined as anaerobic, rod-shaped, Gram-positive bacteria that do not form endospores, eubacteria represent a consortium of distantly related species that are, as recently revealed, distributed over two phyla. While most of previously identified *Eubacterium* spp. cluster within the Firmicutes phylum, several species have been reclassified into distant genera, such as *Collinsella*, *Eggerthella* or *Slackia*, which belong to the Actinobacterium phylum. Until now, eight novel genera have derived from previously recognised *Eubacterium* spp., of which four are reported to be gastrointestinal inhabitants – *Dorea*, *Collinsella*, *Eggerthella* and *Pseudoramibacter* (233, 234, 472, 521). Further reclassification of the genus can be expected since the residual eubacteria are distributed over six distinct 16S rRNA gene based clusters (Table 10). The first isolation of an *Eubacterium* spp. from the human gastrointestinal sample occurred in 1908 when Tissier plated *Bacillus ventriousus*, which was later renamed into *E. ventriosum* (478).

Ruminococci

Ruminococci are abundant fraction of the human gastrointestinal microbiota that can reach concentrations of up to 10^{10} cells per gram of intestinal content (153). Being strictly anaerobic cellulolytic cocci, ruminococci were isolated from human intestinal samples only after the improvement of the anaerobic techniques and media for studying rumen anaerobes (219). The first human intestinal *Ruminococcus* sp. reported is *Ruminococcus bromii* isolated in 1972 (331). None of the 12 intestinal ruminococci has been reclassified, even though, based on the 16S rRNA gene sequence, they form two distinct groups. The more numerous group of ruminococci belongs to the cluster XIVa, while the remaining species are groping within the cluster IV (Table 11) (87).

Table 10 *Eubacterium* and *Eubacterium*-like species isolated from the human gastrointestinal tract.

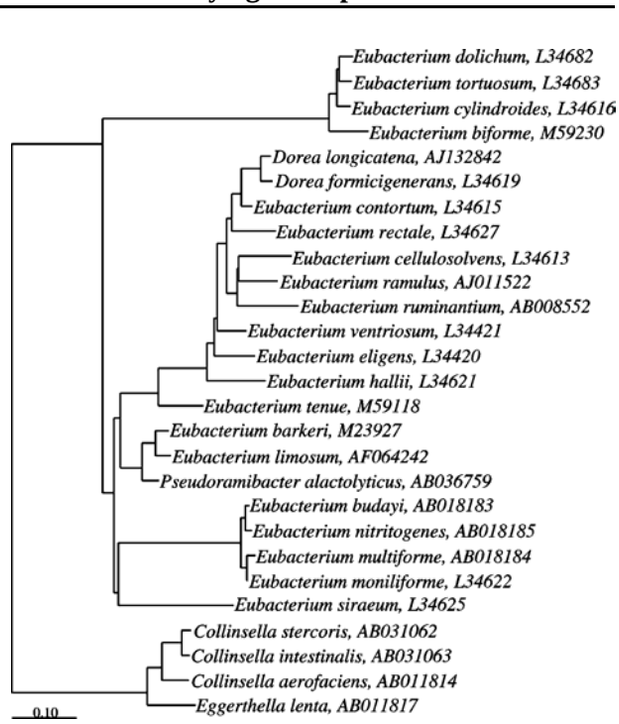
Species	Reference	Genus/Cluster	Phylogenetic position
E. dolichum	(335)	Cluster XVI	
E. tortuosum	(133)		
E. cylindroides	(77)		
E. biforme	(207)		
D. longicatena	(472)	*Dorea*	
D. formicigenerans	(207, 472)		
E. contortum	(150)	Cluster XIVa	
E. rectale	(333)		
E. cellulosolvens	(153)		
E. ramulus	(335)		
E. ruminantium	(333)		
E. ventriosum	(133)		
E. eligens	(207)		
E. hallii	(207)		
E. tenue	(115)	Cluster XI	
E. barkeri	(150)	Cluster XV	
E. limosum	(333)		
P. alactolyticus	(45, 521)	*Pseudoramibacter*	
E. budayi	(153)	Cluster I	
E. nitritogenes	(150)		
E. multiforme	(115)		
E. moniliforme	(115)		
E. siraeum	(335)	Cluster IV	
C. stercoris	(231)	*Collinsella*	
C. intestinalis	(231)		
C. aerofaciens	(233, 333)		
E. lenta	(133, 234)	*Eggerthella*	
E. hadrum	(335)	*Eubacterium*	
B. angulosus	(116)	*Bacteroides**	Unknown, no 16S rRNA gene sequence available
B. acuminatus	(116)		
B. cornutus	(133)		
B. pseudoramosus	(133)		
B. dimorfus	(116)		

Table lists species name, the reference that provides the link with the human gastrointestinal tract as ecological niche, and genus or phylogenetic cluster to which each species belongs based on the 16S rRNA gene sequence as proposed by Collins *et al.* (87)

* Bacteria were reported as Gram-positive *Bacteroides* sp., which does not comply with the current systematics. As several of those *Bacteroides* sp. reported in the same study were reclassified into the *Eubacterium* genus, all reported species are listed as eubacteria, which group the Gram-positive, non-sporulating anaerobes (381).

Table 11 *Ruminococus* species isolated from the human gastrointestinal tract.

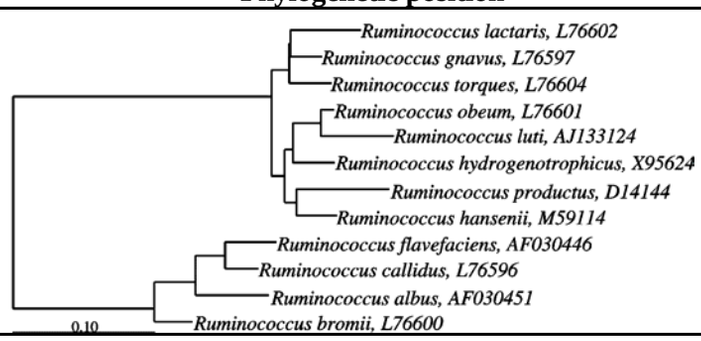
Species	Reference	Phylogenetic position
R. lactaris	(335)	
R. gnavus	(335)	
R. torques	(207)	
R. obeum	(335)	
R. luti	(436)	
R. hydrogenotrophicus	(46)	
R. productus	(141, 333)	
R. hansenii	(141, 207)	
R. flavefaciens	(150)	
R. callidus	(207)	
R. albus	(333)	
R. bromii	(331)	

Table lists species name and the reference that provides the link with the human gastrointestinal tract as ecological niche.

Peptococci and Peptostreptococci

Peptostreptococcus and *Peptococcus* spp. can reach high concentrations in the intestinal samples, even at the level of 10^{10} cells per gram of intestinal content; however, only a proportion of analysed subjects were carriers of these groups of organisms (44, 45, 153). Only eight species belonging to those two genera are reported in relation to the gastrointestinal tract of humans (Table 12), of which five species were reclassified into the following genera: *Anaerococcus*, *Finegoldia* and *Peptoniphilus* (140, 341). Additionally, *P. micros*, is still recognised as a *Peptostreptococcus* due to an unfortunate choice of an already taken genus name – *Micromonas*, to which organisms was meant to be reclassified (341). The most frequently isolated intestinal species of this group is now known as *Peptoniphilus asaccharolyticus*, and this was also the first cultured and described organism from this group (116).

Table 12 *Peptococcus*, *Peptococcus*-like, *Peptostreptococcus* and *Peptostreptococcus*-like species isolated from the human gastrointestinal tract.

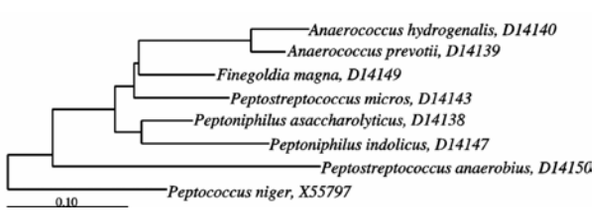
Species	Reference	Genus	Phylogenetic position
A. hydrogenalis *A. prevotii*	(141, 142) (140, 150)	*Anaerococcus*	
F. magna	(150, 341)	*Finegoldia*	
P. micros	(150)	*"Micromonas"*	
P. asaccharolyticus *P. indolicus*	(116, 140) (43, 140)	*Peptoniphilus*	
P.anaerobius	(42)	*Peptostreptococcus*	
P. niger	(485)	*Peptococcus*	
P. intermedius	(150)	*Peptostreptococcus*	Unknown, no 16S rRNA gene sequence available

Table lists species name, the reference that provides link with the human gastrointestinal tract as ecological niche, and the genus in which each species is classified.

Fusobacteria

The genus *Fusobacterium* groups pointed, non-sporulating, Gram-negative, anaerobic bacilli (256). The first record of a *Fusobacterium* spp. originates from 1886 when *Bacillus fusiforme* (*F. necrophorum*) was reported as a pathogen in relation to appendicitis (155). Despite the fact that the first isolate was obtained from appendix tissue, fusobacteria were considered to be exclusively oral inhabitants until 1966 when it was shown that fusobacteria can be cultivated from faecal material (486). Eleven *Fusobacterium* spp., have been recognised as the human gastrointestinal inhabitants (Table 13). Although, fusobacteria can reach densities of up to 10^{10} cells per gram of faeces (42), it is likely that the misclassification of the abundant *Faecalibacterium prausnitzii* into the genus *Fusobacterium* (127, 333) considerably contributed to their recognition as one of the most dominant intestinal fractions.

Table 13 *Fusobacterium* and *Fusobacterium*-like species isolated from the human gastrointestinal tract.

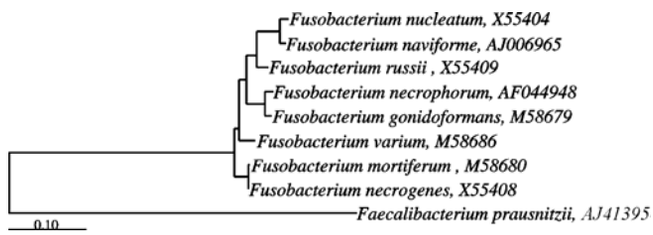
Species	Reference	Genus	Phylogenetic position
F. nucleatum	(206)	*Fusobacterium*	
F. naviforme	(206)		
F. russi	(333)		
F. necrophorum	(150)		
F. gonidoformans	(150)		
F. varium	(42)		
F. mortiferum	(333)		
F. necrogenes	(150)		
F. prausnitzii	(127, 333)	*Faecalibacterium*	
F. bullosum	(42)	*Fusobacterium*	Unknown, no 16S rRNA gene sequence available
F. girans	(123)		
F. glutinosum	(42)		
F. symbiosum	(333)		

Table lists species name, the reference that provides link with the human gastrointestinal tract as ecological niche, and the genus in which each species is classified.

Enterobacteriaceae

Quantitatively, Enterobacteriaceae represent only a small fraction of the total gastrointestinal microbiota. Their occurrence increases with age, but even in elderly subjects they cover only about 1% of the total intestinal bacteria (214). Moreover, the vast majority of the Enterobacteriaceae species are infrequently isolated from intestinal samples (70, 340). Still, this is one of the most comprehensively described gastrointestinal families, which is probably because of their relatively easy cultivation in laboratory conditions but also suspected clinical significance as pathogenic organisms (144, 179, 198, 199, 201). For instance, *Morganella-Proteus-Providencia* group is considered as potential agent in diarrhoea development, but the isolation from faeces of over a thousand healthy persons and patients suffering from enteric diseases showed that all species of this group could be found both in healthy and diseased subjects, albeit with a different frequency (340). Such species can be considered as members of the normal gastrointestinal microbiota and are listed in Table 14. Other members of the Enterobacteriaceae family, such as *Salmonella* spp. or *Shigella* spp. have been isolated exclusively from clinical samples and were identified as pathogenic organisms, and therefore are reported in the following section of this review.

Clinical isolates with potential pathogenic properties

Many of the human gastrointestinal inhabitants were initially reported as clinical isolates for which pathogenic properties were suspected, while today most of them are recognised as members of the normal gastrointestinal microbiota. For instance, the first intestinal isolates of the genera *Bacteroides* and *Fusobacterium* isolates that were initially recognised as pathogens (492).

Table 14 Human gastrointestinal isolates that belong to the Enterobacteriaceae family.

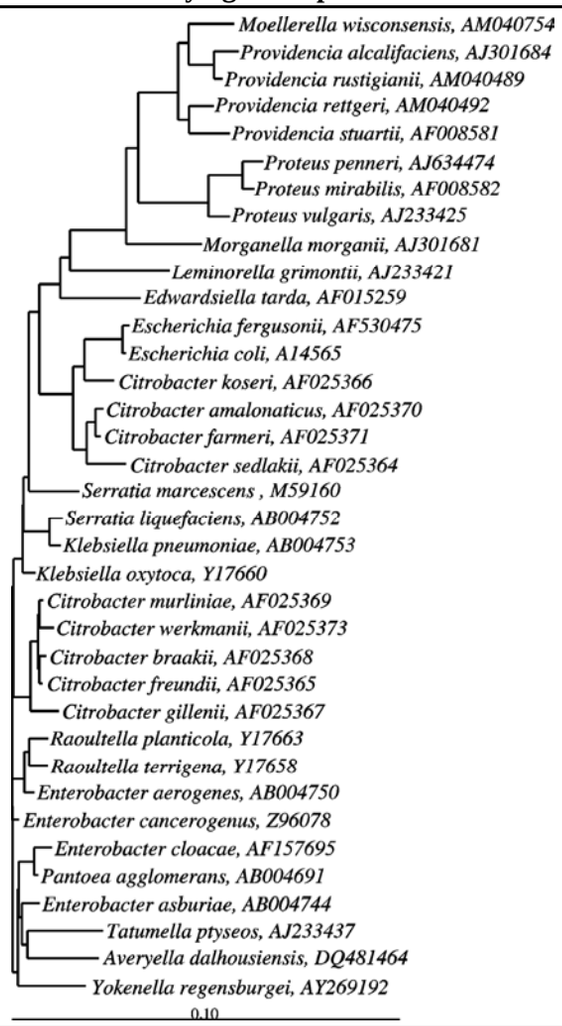
Species	Reference	Genus	Phylogenetic position
M. wisconsensis	(199)	*Moellerella*	
P. alcalifaciens	(340)		
P. rustigianii	(340)	*Providencia*	
P. rettgeri	(340)		
P. stuartii	(138)		
P. penneri	(340)		
P. mirabilis	(340)	*Proteus*	
P. vulgaris	(340)		
M. morganii	(340)	*Morganella*	
L. grimontii	(201)	*Leminorella*	
E. tarda	(139)	*Edwardsiella*	
E. coli	(333)	*Escherichia*	
E. fergusonii	(144)		
C. koseri	(67)		
C. amalonaticus	(64)	*Citrobacter*	
C. farmeri	(64)		
C. sedlakii	(64)		
S. marcescens	(180)	*Serratia*	
S. liquefaciens	(333)		
K. pneumoniae	(333)	*Klebsiella*	
K. oxytoca	(480)		
C. murliniae	(67)		
C. werkmanii	(64)		
C. braakii	(64)	*Citrobacter*	
C. freundii	(150)		
C. gillenii	(67)		
R. planticola	(121, 377)	*Raoultella*	
R. terrigena	(121, 376)		
E. aerogenes	(153)		
E. cancerogenus	(144)	*Enterobacter*	
E. cloacae	(150)		
P. agglomerans	(473)	*Pantoea*	
E. asburiae	(66)	*Enterobacter*	
T. ptyseos	(208)	*Tatumella*	
A. dalhousiensis	(224)	*Averyella*	
Y. regensburgei	(481)	*Yokenella*	
E. hafnia	(333)	*Escherichia*	Phylogenetic position unknown No 16S rRNA gene sequence available
L. richardii	(201)	*Leminorella*	
C. intermedius	(259)	*Citrobacter*	
C. youngae	(64)		

Table lists species name, the reference that provides link with the human gastrointestinal tract as ecological niche, and genus in which each species is classified.

Probable reasons for isolation of non-pathogenic species from an infected tissue include the fact that infected site represents their natural ecological niche, or that they are opportunistic pathogens, such as the very first intestinal isolate *E. coli*, with an example of pathogenic *E. coli* 0157H7 and non-pathogenic *E. coli* K12. Considering the fact that clinical microbiology was established earlier than microbial ecology it is likely that many species, which are at present time only clinical isolates with potential pathogenic properties, will be recovered from healthy intestinal samples. In any case, the distinction between clinical isolates with potential pathogenic properties and other gastrointestinal isolates given in this review should be taken with some reserve, as only genera for which the majority of species

were found to be clinically significant pathogens, are reported under this section (Table 15). However, some genera, which are primarily associated with intestinal diseases, contain species that have already been isolated from healthy intestinal samples. An example of such genus is *Campylobacter*, which in principally considered to group pathogenic microorganisms of which nine species were isolated from the human intestinal samples. However *C. hominis* is a species isolated from intestinal sample of healthy subject (266), while *C. concisus* is a clinical isolate that has been recently recovered from a healthy individual (137). Moreover, *Helicobacter* spp. that are principally recognised as pathogens inducing gastric and duodenal ulcers, can be detected in the faeces of healthy individuals when sensitive techniques are applied (79, 295).

Table 15 Clinical isolates with potential pathogenic properties from the human gastrointestinal tract.

Species	Reference	Phylum	Phylogenetic position
S. dysenteriae	(508)		
S. flexneri	(76)		
T. guamensis	(314)		
S. enterica	(5, 476)		
C. davisae	(179)		
Y. pseudotuberculosis	(81)		
Y. bercovieri	(509)		
Y. rohdei	(6)		
Y. kerstensenii	(118)		
Y. frederiksenii	(481)	γ-Proteobacteria	
Y. enterocolitica	(118)		
H. alvei	(150)		
V. furnisii	(65)		
V. fluvialis	(269)		
V. parahaemolyticus	(330)		
G. hollisae	(202)		
A. enteropelogenes	(73)		
A. veronii	(200)		
A. succiniciproducens	(300)		
A. thomasii	(300)		
H. pylori	(312)		
H. pullorum	(456)		
H. canadensis	(158)		
H. winghamensis	(315)		
H. cinaedi	(137)		
F. rappini	(18)		
C. hominis	(266)		
C. gracilis	(515)		
C. rectus	(515)	ε-Proteobacteria	
C. concisus	(488)		
C. fetus	(497)		
C. lari	(439)		
C. jejuni	(358)		
C. upsaliensis	(176)		
C. coli	(365)		
A. cryaerophilus	(137)		
A. butzleri	(137)		
B. pilosicoli	(319)	Spirochaetes	
B. aalborgi	(319)		
Yersinia aleksiciae	(452)	γ-Proteobacteria	Unknown, no 16S rRNA gene sequence available

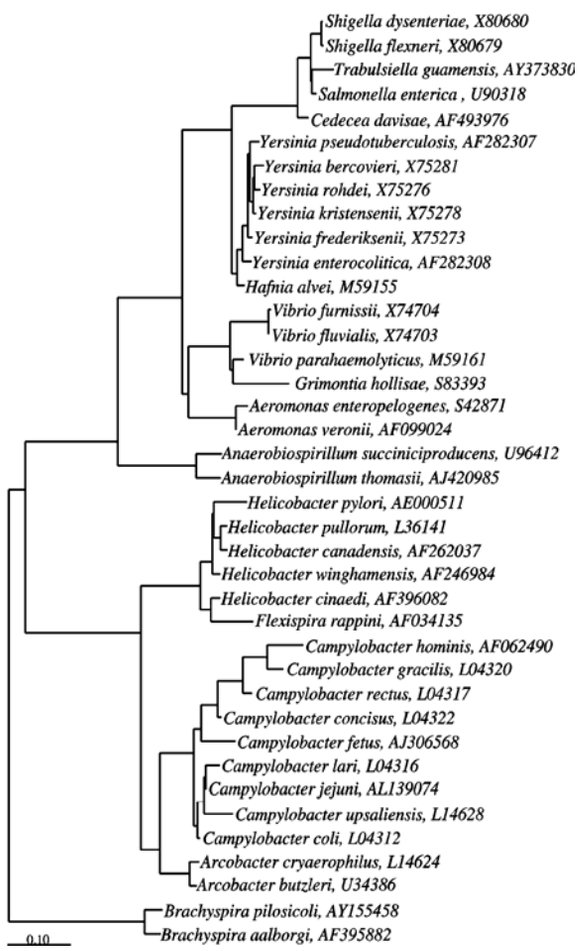


Table lists human gastrointestinal isolates of clinical origin, the reference that provides link with the human gastrointestinal tract as ecological niche, and phylum in which each species is classified.

Other human gastrointestinal isolates

A number of human gastrointestinal inhabitants does not belong to any of the major groups listed above. Those isolates are phylogenetically very heterogeneous, but also differ in their distribution and abundance in the human intestine. The majority, however, belongs to the most diverse gastrointestinal phylum Firmicutes (Table 16).

Table 16 Human gastrointestinal isolates of the Firmicutes phylum that are not members of groups recognised as dominant in the human gastrointestinal tract.

Species	Reference	Cluster	Phylogenetic position
C. catenaformis	(230)	cluster XVIII	
C. mitsuokai	(229)	cluster XVII	
S. moorei	(232)	cluster XVI	
H. filiformis	(522)		
A. viridans	(153)	Bacilli	
P. acidilactici	(32)		
P. pentosaceus	(100)		
L. mesenteroides	(100)		
B. cereus	(33)		
B. flexus	(33)		
B. megaterium	(473)		
B. badius	(33)		
B. circulans	(473)		
B. sphaericus	(473)		
B. subtilis	(33)		
B. pumilus	(33)		
S. saccharolyticus	(43)		
S. epidermidis	(333)		
S. aureus	(150)		
B. haldurans	(61)		
P. lautus	(33, 197)		
B. brevis	(434, 473)		
A. aneurinolyticus	(434, 435)		
B. fibrisolvens	(69)	cluster XIVa	
R. intestinalis	(126)		
B. formatexigenes	(528)		
B. crossotus	(335)		
C. eutactus	(207)		
A. caccae	(425)		
L. pectinoschiza	(523)		
C. catus	(207)		
A. stercorihominis	(151)	cluster XV	
S. variabile	(209)	cluster IV	
A. colihominis	(267)		
S. ventriculi	(153)	cluster I	
Staphylococcus albus	(35)	Unknown, no 16S rRNA gene sequence available	
Bacillus fusus	(33)		
Bacillus tritus	(33)		
Bacillus albolactus	(33)		
Bacillus pseudotetanus	(33)		
Bacillus petasites	(33)		
Bacillus vulgatus	(33)		
Coprococcus comes	(207)		

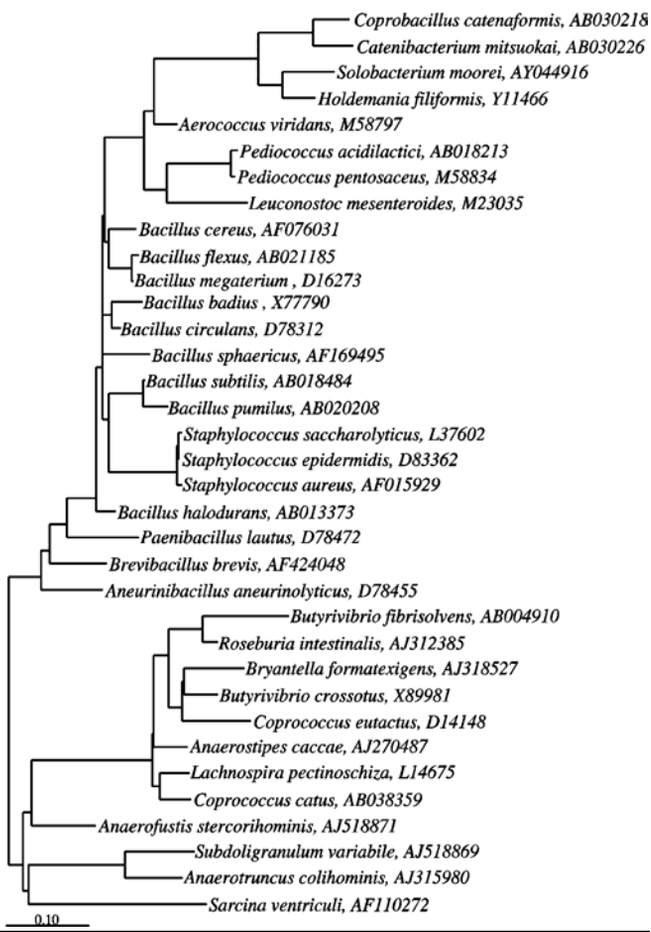


Table lists species name, the reference that provides link with the human gastrointestinal tract as ecological niche, and phylogenetic cluster to which each species belongs based on the 16S rRNA gene sequence as proposed by Collins et al. (87)

A great proportion of those Firmicutes species are representatives of novel genera such as *Roseburia, Anaerofustis, Subdoligranulum*, *Coprobacillus* or *Solobacterium*. Other species belong to well established, frequently isolated but simple genera such as *Coprococcus* that is composed of only three species, which are all known gastrointestinal inhabitants. Furthermore, several bacteria primarily found in other ecosystems, such as *Bacillus* spp., typical soil organisms, or *Pediococcus* spp, used in fermented food production, were occasionally isolated from the human gastrointestinal samples.

Other than the Firmicutes gastrointestinal isolates which are not recognised as dominant members of the gastrointestinal microbiota are distributed between five phyla – Actinobacteria, Fusobacteria, Lentisphaerae, Proteobacteria and Verrucomicrobia (Table 17).

Table 17 The human gastrointestinal isolates, other than Firmicutes, that are not members of groups recognised as dominant in the human gastrointestinal tract.

Species	Reference	Phylum	Phylogenetic position
C. somerae	(152)	Fusobacteria	
L. buccalis	(123)		
H. haemolyticus	(279)		
H. parainfluenzae	(279)		
P. fluorescens	(260)		
P. putida	(260)		
P. montelii	(135)		
P. aeruginosa	(150)	γ-Proteobacteria	
P. stutzeri	(123)		
A. calcoaceticus	(384)		
A. johnsonii	(76)		
A. haemolyticus	(371)		
M. catarrhalis	(123)		
S. wadsworthia	(515)		
A. faecalis	(369)		
A. dentrificans	(150)	β-Proteobacteria	
K. gyiorum	(85)		
O. formigenes	(9)		
D. desulfuricans	(170)		
D. fairfieldensis	(281)	δ-Proteobacteria	
D. piger	(335)		
B. wadsworthia	(30)		
C. pseudodiphtheriticum	(123)		
C. ammoniagenes	(95)		
C. ulcerans	(123)		
C. xerosis	(123)		
P. acnes	(333)		
P. avidum	(133)		
P. acidipropionici	(387)	Actinobacteria	
P. jensenii	(473)		
P. granulosum	(150)		
P. propionicum	(387)		
M. luteus	(150)		
R. dentocariosa	(473)		
A. naeslundii	(153)		
A. pyogenes	(473)		
V. vadensis	(540)	Lentisphaerae	
A. muciniphila	(110)	Verrucomicrobia	
Gemmiger formicilis	(206)	α-Proteobacteria	Unknown, no 16S rRNA gene sequence available

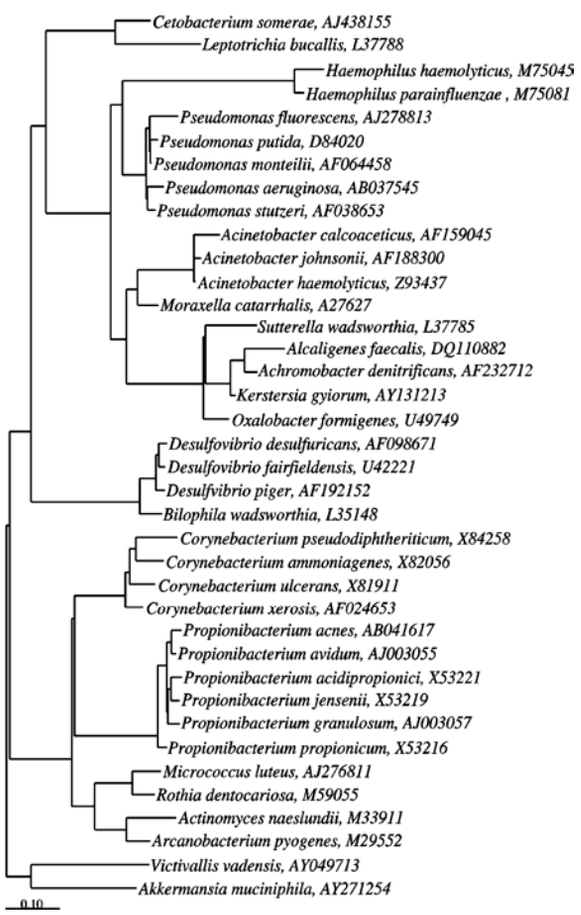


Table lists species name, reference that provides link with the human gastrointestinal tract as ecologic niche, and the phylum in which the species is classified.

The majority of the scarce gastrointestinal groups belong to the Proteobacteria phylum, and most of those species are principally found in low numbers, and consequently rarely detected. However, even low abundant bacteria can be widely distributed, *e.g. Oxalobacter formigenes* is normally present in healthy individuals although its richness reaches at maximum 10^7 cells per gram of faeces (8). Some of the low abundant bacteria were specific target of different studies and thus were recovered from the gastrointestinal samples. An example of those includes *Desulfovibrio* spp. that are interesting because of their involvement in the competition for hydrogen available in the gut and their potential role in the aetiology of ulcerative colitis (170, 171, 281). Furthermore, the only cultivated member of phylum Verrucomicrobia – *Akkermansia muciniphila* – is an abundant and widely distributed in human gastrointestinal inhabitant (108). However, as it is the single representative of its phylum, which has been recently isolated, this group is not yet widely acknowledged as dominant. Another single representative of its phylum – *Victivallis vadensis* representative of Lentisphaerae – is also reported under this section. In addition, here we report the only two human gastrointestinal isolates that belong to phylum Fusobacteria but do not group within the genus *Fusobacterium*, and several Actinobacteria primarily belonging to two genera – *Propionibacterium* and *Corynebacterium* that are occasionally isolated from gastrointestinal samples.

As mentioned earlier, bacteria are not an exclusive component of the human gastrointestinal microbiota, although they are the most dominant and well-studied fraction. Microorganisms inhabiting the colon of about one third of the human population are responsible for methane production in these individuals (58). Methanogenesis is carried out by archaeal species whose density reaches up to 10^{10} cells per gram of faeces (321). Although a single methanogenic species, isolated from four out of five individuals, was identified as *Methanobrevibacter ruminantium* in 1968 (349), today, *Methanobrevibacter smithii* is considered as the most abundant methanogen of human intestine (321). *Methanosphaera stadtmaniae* is atypical methanogenic archaea that reduces methanol and can be found in human faeces in low concentrations (320).

The yeasts belonging to *Candida* genus are other known inhabitants of the human gastrointestinal tract. *C. albicans* and *C. rugosa* are often found to be part of normal microbiota (13, 150, 153), while other candidae, namely *C. glabrata*, *C. krusii*, *C. parapsilosis*, *C. tropicalis*, *C. colliculosa*, *C. kefyr* and *C. lustianiae* were isolated from faeces of patients with digestive disorders or allergic problems (48). In addition, another yeast – *Cladosporium sphaerospermum* – was detected as transient gastrointestinal inhabitant (473).

Finally, filamentous fungi can be present in the human gastrointestinal tract and although they are not widely distributed, their presence was noticed in a number of studies (45, 150, 153). Several *Aspergillus* and *Penicillium* spp. were identified during a nine-year long monitoring of an immunodeficient child's microbiota (473). The identified isolates

include *P. citrinum*, *P. notatum*, *P. steckii*, and *P. decumbens*; *A. sydowi*, *A. versicolor* and *A. niger*.

Species richness estimation

In order to visualise the representation of the diversity of the gastrointestinal microbiota, a rank abundance curve was constructed by plotting the frequency of isolation of each described intestinal species (Figure 1). Based on the shape of this curve, which can visualise how well an ecosystem has been sampled (218), the cultivable fraction of the gastrointestinal microbiota seems to be well represented.

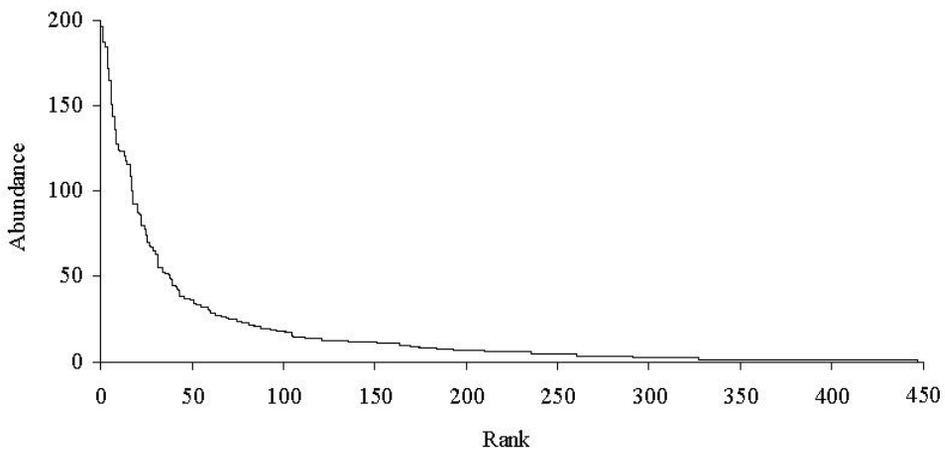


Figure 1 Rank-abundance curve for the human gastrointestinal tract cultivated and characterised inhabitants. The abundance of the species is defined as number of individuals that harboured each of the identified gastrointestinal inhabitants using the identification of isolate as provided by an original reference.

Taking into account the number of obtained isolates and the frequency of their isolation it is possible to estimate the expected total diversity of the ecosystem (218). In this particular case, the analysis of the literature data showed that, out of 442 cultivated species, 119 are recovered from only one individual and 37 from two individuals. The expected diversity of the human gastrointestinal microbiota calculated on the basis of this data (expressed as Chao1 estimate) is ~630 species. If the isolates that are reported only in the clinical studies are excluded - as it is still not proven that they are part of the normal microbiota - the expected diversity is ~580 species.

It should be noted that in the vast majority of descriptive studies of the human gastrointestinal microbiota only already known species were identified, while novel isolates were characterised at the genus level. Due to the different coding systems that are used in different studies, the number of those isolates and the frequency of their isolation could not be exactly determined. Taking into account large inter-individual variations in the

gastrointestinal microbiota composition, it is likely that the unidentified isolates would contribute considerably to the variability of the data set. This possibility is supported by the analysis of data from a single study of the microbiota composition of 40 individuals (150). Here, out of 266 obtained isolates 155 were not fully identified, and from the latter, only 22 were detected in more than single individual. Therefore, the good representation (Fig. 1) might correspond to the representation of characterised fraction of the cultivable microbiota alone. Moreover, under currently used cultivation conditions, only a part of the gastrointestinal microbes can be cultivated. Therefore, once the cultivation procedures have improved and allow recovery of the species that cannot be cultured at this stage, this expected diversity will evidently increase.

Concluding remarks

Cultivation-based studies are an important source of information on the human gastrointestinal diversity because of their extensive use, notably for the analysis of a large amount of samples. Moreover, cultivation precedes taxonomic description and allows analysing representative of a species as a pure culture. The present literature indicates that the total microbiota composition of more than 300 individuals has been assessed by cultivation studies. Hundreds of subjects took part in the analysis of subgroups of gastrointestinal microbiota and, finally, there were attempts to isolate species of specific interest from thousands of intestinal samples (81, 340).

Although cultivation techniques have been applied in high throughput analysis of the human gastrointestinal microbiota, these approaches are time consuming, laborious and costly and therefore not the desirable choice especially for large scale studies. Alternatively, molecular techniques that enable quantitative and qualitative analysis of the human gastrointestinal microbiota at high throughput level have been developed and successfully applied (491). However, the cultivation approach should not be abandoned when analysing the human gastrointestinal microbiota until the full description of this ecosystem has been achieved. Although a considerable number of novel isolates is expected to be discovered even with the application of traditional cultivation techniques, the traditional techniques are only partially covering the total gastrointestinal diversity. Recently, a member of a novel, widely distributed and abundant intestinal genus *Akkermansia* was isolated with the use of a specific medium, based on intestinal mucin (110, 129, 502). This example should serve as a guide in the development of creative approaches for the assessment of the complete diversity of the microbial ecosystem within the human gastrointestinal tract.

CHAPTER 3

Mirjana Rajilić-Stojanović, Hauke Smidt and Willem M. de Vos

Accepted for publication in Environmental Microbiology

Diversity of the Human Gastrointestinal Microbiota Revisited

Since the early days of microbiology, more than a century ago, representatives of over 400 different microbial species have been isolated and fully characterised from human gastrointestinal samples. However, during the past decade molecular ecological studies based on ribosomal RNA (rRNA) sequences have revealed that cultivation has been able only to access a small fraction of the microbial diversity within the gastrointestinal tract. The increasing number of deposited rRNA sequences calls for the setting up a curated database that allows handling of the excessive degree of redundancy that threatens the usability of public databases. The integration of data from cultivation-based studies and molecular inventories of small subunit (SSU) rRNA diversity, presented here for the first time, provides a systematic framework of the microbial diversity in the human gastrointestinal tract of more than 1,000 different species-level phylogenetic types (phylotypes). Such knowledge is essential for the design of high-throughput approaches such as phylogenetic DNA microarrays for the comprehensive analysis of gastrointestinal tract microbiota at multiple levels of taxonomic resolution. Development of such approaches is likely to be pivotal to generating novel insights in microbiota functionality in health and disease.

Introduction

The human gastrointestinal microbiota represents a complex ecosystem that consists of bacteria, archaea, yeasts and filamentous fungi (150, 321). While Archaea and Eukarya are represented by members of a single phylum each, bacterial community within the human gastrointestinal tract is exceptionally diverse. Members of nine bacterial phyla were found to inhabit the human gastrointestinal tract of which Firmicutes, Bacteroidetes and Actinobacteria are dominant (23). Members of Proteobacteria are also common and diverse, but they are usually secondary to the above. Recent reports indicated that viruses represent another important constituent, as more then 1,200 viral genotypes were identified in human faeces with a density of up to 10^9 virions per g of dry material (63, 533).

Many studies have attempted to describe the normal gastrointestinal microbiota in terms of microbial species that inhabit healthy humans. However, individual differences in the microbiota composition make the definition of this concept very challenging. Since only a limited number of individuals has been subjected to the analysis of intestinal microbial diversity, the description of the normal composition remains incomplete. Both molecular, SSU (16S and 18S) rRNA-based, and traditional cultivation studies show that the mere selection of some different subjects is sufficient for the discovery of novel intestinal inhabitants (129, 209, 249, 502).

Even though it is clear that the present view of the microbial diversity within the human gastrointestinal tract is incomplete, it is difficult to estimate to what degree. Most work has been performed in an uncoordinated way over more than a century, by many different groups, and by using a wide variety of approaches. As a coherent meta-analysis of the findings of those numerous studies is still lacking, there is no general agreement even concerning the number of different bacterial species that can inhabit the healthy human intestine. Still, the vast majority of authors refer to an expected diversity of about 400-500 species, but in some of the recent reports the number of 1,000 gastrointestinal species appears (348, 350, 370). Recently, Bäckhed and co-authors determined that 800 distinct 16S rRNA gene sequences present in the GenBank were derived from the human gastrointestinal samples (23). However, a comprehensive picture of the diversity of the microbiota in the gastrointestinal tract is still lacking. This is because the results from cultivation-based studies have never been curated, analysed and integrated with molecular data. Hence, we provide a systematic overview of the gastrointestinal microbiota diversity, with specific attention on the community structure, based on a large body of the critical and phylogenetic analysis of the literature and public databases entries.

The present standpoint of the diversity of the human gastrointestinal microbiota

The first identified bacterial species recovered from a human gastrointestinal sample was *Escherichia coli* isolated in 1885 from children's diarrhoeal faeces (16). The description of the human gastrointestinal inhabitants continued throughout the previous century and resulted in the full characterisation of over 400 cultivated species (**Chapter 2**). However, it was the application of molecular techniques that has revolutionised the view of this ecosystem (163). Application of molecular techniques revealed that the microbiota was significantly more complex than previously anticipated, as only a fraction of the bacteria living in the human intestine can currently be cultured (461). The proportion of reported cultivable bacteria varies between studies from 20% for the mucosal samples of three healthy individuals (129) to 46% for the three elderly Japanese subjects (193). The other major outcome of the molecular ecology revolution was that the composition of the microbiota is subject-specific, and dominated by yet uncharacterised phylotypes (29, 129, 191-193, 204, 301). This observation is in contrast to the often cited (*e.g.* (255, 504)), but likely incorrect, statement that the human gastrointestinal microbiota is diverse but dominated by a very limited number of bacterial species (122).

The first comprehensive study aiming at the characterisation of the microbial community of the human intestine based on the SSU rRNA gene sequence analysis, was published in 1999 (461). However, prior to that date a preliminary insight into the diversity of the intestinal ecosystem based on SSU rRNA gene sequences was reported in a study that provided a rapid method for profiling the intestinal microbiota (537). Since then and until the beginning of 2006, the total gastrointestinal microbiota of 34 subjects and specific groups of 28 subjects have been studied based on the SSU rRNA sequencing approach (Table 1).

Although relatively new, molecular studies have been very productive, and until the beginning of 2006 about 15,000 SSU rRNA gene sequences, which were obtained from human gastrointestinal samples, could be retrieved from public databases. However, most of those sequences originate from a single study, in which each cloned 16S rRNA gene insert was sequenced and deposited, resulting in a public catalogue of 13,335 sequences (129). This study of Eckburg and colleagues represents the first application of high throughput sequencing applied on the samples of the human gastrointestinal origin, which have enabled better insight into the diversity and abundance of 395 detected phylotypes along the gastrointestinal tract of three individuals. Several subsequent studies confirmed that high throughput sequencing is a powerful approach of analysing a complex microbial ecosystem, such as the human gastrointestinal microbiota (52, 275). An undesired consequence of such massive sequencing approach is the deposition of large number of identical sequences that has caused extensive redundancy in public databases. This jeopardizes the usefulness of public

databases, as it limits, for instance, the recognition of the closest cultivated relative during the BLAST search (11). Such extensive redundancy can be omitted by the deposition of only one representative sequence per identified phylotype. Several strategies can be used for this purpose and they include analysing the uniqueness of the SSU rRNA gene sequence, for instance by restriction analysis of the amplicons, (194, 461), phylogenetic analysis of partial SSU rRNA (191, 192, 502), or by simple selection of distinct sequences based on *in silico* identification of unique phylotypes. For the purpose of the analysis of the human gastrointestinal diversity presented in this paper all redundant sequences have been excluded.

Table 1 Number of subjects whose gastrointestinal microbiota composition was described by SSU rRNA gene sequencing. Inventory period from 1998 until the beginning of 2006.

Subject	Fraction of microbiota	Number of subjects	Reference
Healthy adult	Total	11	(59, 60, 129, 191, 192, 461, 502, 537)
Healthy infant		4	(146, 501)
Healthy elderly		13	(193, 204, 505)
Crohn's disease patient		4	(301)
Ulcerative colits patient		2	(R. A. Hutson, unpublished)
Healthy adult	*Lactobacilli*	12	(194)
Healthy infant		1	
Healthy adult	*Bifidobacteria*	5	(417)
Healthy adult	Bacteria enriched on mucin-based medium	6	(109)
Healthy adult	Butyrate producing bacteria	3	(29, 128)
Healthy infant		1	

Phylogenetic analysis of the diversity of the gastrointestinal microbiota

To assess the diversity of the human gastrointestinal microbiota, we integrated the results obtained in the reported molecular studies of the human gastrointestinal diversity by a careful phylogenetic analysis of the SSU rRNA gene sequences. This step was indispensable because each clone of the SSU rRNA gene is given a distinct name, even when it contains only another copy of the very same gene. Although most studies include phylogenetic analysis of the obtained sequences, the results are not preserved in a publicly available format. Consequently, integration of the results from different studies is impossible without performing extensive phylogenetic analysis. Repetition of this effort leads to an unreasonable loss of energy and time. Therefore, there is an urgent need for establishing a framework for

the curation of fast-growing databases of SSU rRNA gene sequences as it is attempted here for those obtained from human gastrointestinal samples.

In addition to the high level of redundancy within public databases, phylogenetic analysis of the SSU rRNA gene sequences obtained in different studies is hampered by the fact that sequences might differ in size and the region of the SSU rRNA gene that was sequenced. A perfectly appropriate tool for the analysis of such data set is not available, but the ARB software seems to be the most suitable (286). For the purpose of this review, the ARB database release from 2002 was updated with SSU rRNA gene sequences obtained from human gastrointestinal tract samples that were not present in this release but could be recovered in other SSU rRNA gene databases, such as RDP (86), or GeneBank (www.ncbi.nlm.nih.gov). For the selected sequences a distance matrix was calculated and a threshold of 2% sequence divergence was used to define distinct phylotypes. Calculation of similarity (distance) matrices as implemented in ARB has a major advantage compared to the routinely performed BLAST search (11), as an alignment with an ambiguous nucleotide is not considered as a mismatch. As a result, sequences of low quality can have high similarity scores. This is particularly relevant, as for some cultivated organisms SSU rRNA gene sequences are of low quality, and contain a relevant number of ambiguous sequences. However, the quality of the ARB similarity matrix is highly dependent on the quality of the alignment, which is edited by the operator and thus more error-prone than the automatically generated alignment in BLAST.

A similar approach, based on the ARB sequence similarity matrix, was previously used by Bäckhed and co-workers (2005). However, in that study only the number of different phylotypes was determined. In this review, we have identified the distinct phylotypes, and for each monophyletic group with ≥ 98% sequence identity, a representative type sequence was assigned. Since sequences of gastrointestinal isolates were combined with those retrieved in cultivation-independent studies, the representative phylotype sequence was chosen as high quality and the most complete SSU rRNA sequence of the monophyletic group, if possible from the cultured representative. Based on the sequence information and using Parsimony procedures, as implemented in ARB, a phylogenetic tree was constructed. The sequences which were not present in the ARB database release 2002 were added to the tree with the use of appropriate filters fitting the sequence phylogeny and length.

The analysis of the human gastrointestinal SSU rRNA gene sequences showed that, until the beginning of 2006, 898 distinct phylotypes were reported in cultivation-independent studies alone. The vast majority of the recovered phylotypes (622; representing 69%) appeared to be subject-specific, which is due to the high level of inter-individual variation of the microbiota composition, but also influenced by the fact that the microbiota of only a limited number of individuals was studied. The majority of the phylotypes represents organisms that have not yet been cultured. Among 898 phylotypes, only 158, representing

18% of the total number of phylotypes, correspond to fully characterised isolates. Integration of data from different studies revealed that the contribution of the cultivated organisms in the total diversity assessed by sequencing of the SSU rRNA gene is lower than within individual studies. This is mainly due to the fact that each attempt to describe the gastrointestinal microbiota by SSU rRNA gene sequencing renders a proportion of novel phylotypes that may be as high as 62% (129). This has been leading to an ever increasing number of novel phylotypes that are only represented by sequences, rather than by cultured isolates. Moreover, a large proportion of the SSU rRNA sequences corresponding to cultivated species is repeatedly detected (59%; 93 out of 158) when compared to uncultivated repeatedly detected uncultured phylotypes (25%; 183 out of 740).

Cultured versus uncultured

Our current knowledge of the human gastrointestinal microbiota diversity originates from cultivation-based and molecular studies. Thus, generating the most complete view of the known gastrointestinal diversity is possible only through the integration of the results from both types of studies. As SSU rRNA gene sequences are in most cases the only available information for the phylotypes reported in molecular analyses, the integration is possible only on the basis of these sequences.

However, for 46 bacterial and 8 eukaryotic human gastrointestinal isolates the SSU rRNA gene has not yet been sequenced (**Chapter 2**). Consequently, these organisms could not be included in the SSU rRNA based phylogenetic analysis. They were considered as distinct species, which were not recovered in the cultivation-independent studies, although it has to be noted that it can not be excluded that isolates with lacking SSU rRNA sequence match with molecular phylotypes.

The integration of datasets from cultivation-dependent and -independent studies showed that 1,184 unique bacterial, 3 archaeal, and 17 eukaryotic phylotypes are reported as inhabitants of the human gastrointestinal tract. The compressed version of this phylogenetic tree is depicted in Figure 1 and summarised distribution of the phylotypes is given in Table 2.

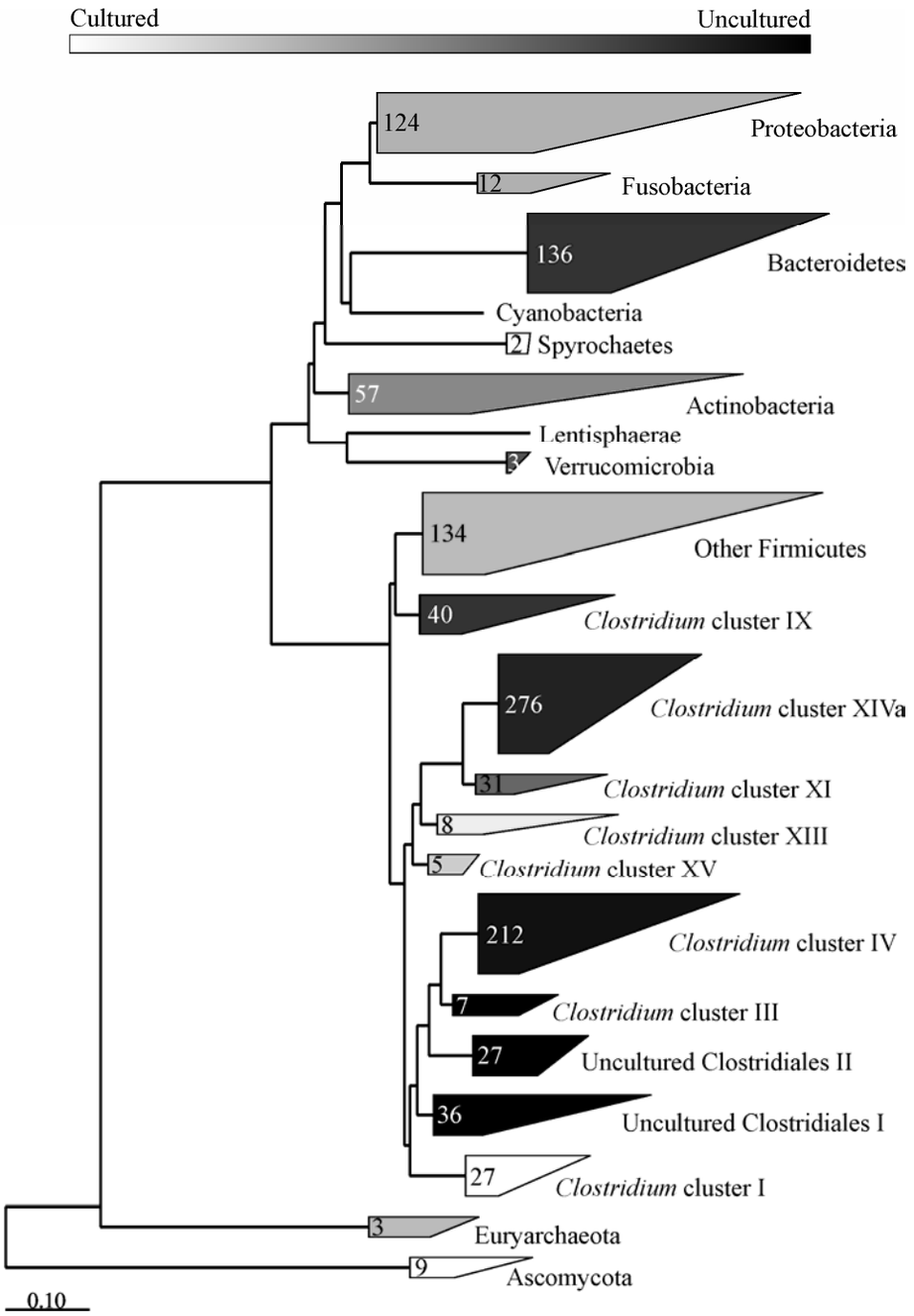


Figure 1 SSU rRNA-based phylogenetic tree of the distinct phylotypes that have been found in the human gastrointestinal tract. The relative proportion of phylotypes that correspond to cultured representatives is indicated by different darkness of the filling. Black fills indicate phylotypes detected in cultivation-independent studies, while white indicates species detected in cultivation-based studies. The reference bar indicates 10% sequence divergence. Numbers of distinct phylotypes are given for each phylogenetic group.

Table 2 SSU rRNA gene sequence based phylogenetic distribution of the human gastrointestinal prokaryotic phylotypes given for taxonomic levels of phylum, class, order and family or cluster as proposed by Collins and colleagues (1994). Inventory period from 1998 until the beginning of 2006.

Phylum		Class		Order		Family/Cluster	
Actinobacteria	57	Actinobacteria	57	Actinomycetales	22	Actinomycetaceae	6
						Corynebacteriaceae	6
						Micrococcaceae	4
						Propionibacteriaceae	6
				Bifidobacteriales	20	Bifidobacterium	20
				Coriobacteriales	15	Coriobacteriaceae	15
Bacteroidetes	136	Bacteroidetes	136	Bacteroidales	136	Rikenellaceae	15
						Bacteroidaceae	45
						Prevotellaceae	36
						Porphyromonadaceae	21
						Unclassified	19
Cyanobacteria	1	Cyanobacteria	1	Chroococcales	1	Unclassified	1
Firmicutes	818	Asteroleplasma	1	Anaeroplasmatales	1	Anaeroplasmataceae	1
		Bacilli	92	Bacillales	18	Bacillaceae	14
						Staphylococcaceae	4
				Lactobacillales	74	Aerococcaceae	1
						Carnobacteriaceae	1
						Lactobacillaceae	36
						Leuconostocaceae	2
						Lactococcaceae	2
						Streptococcaceae	32
		Clostridia	669	Clostridiales	669	Cl. cluster I	27
						Cl. cluster III	7
						Cl. cluster IV	212
						Cl. cluster IX	40
						Cl. cluster XI	31
						Cl. cluster XIII	8
						Cl. cluster XIVa	276
						Cl. cluster XV	5
						Unclassified	63
		Mollicutes	42	Unclassified	42	Cl. cluster XVI	14
						Cl. cluster XVII	5
						Cl. cluster XVIII	9
		Unclassified	14	Unclassified	14	Unclassified	14
Fusobacteria	12	Fusobacteria	12	Fusobacteriales	12	Fusobacteriaceae	11
						Incertae sedis 11	1
Proteobacteria	124	Alphaproteobacteria	6	Rhizobiales	1	Unclassified	1
				Sphingomonadales	1	Unclassified	1
				Unclassified	4	Unclassified	4
		Betaproteobacteria	19	Burkholderiales	18	Alcaligenaceae	8
						Oxalobacteriaceae	3
						Burkholderiaceae	1
						Incertae sedis 5	1
						Unclassified	5
				Neisseriales	1	Neisseriaceae	1
		Gammaproteobacteria	74	Aeromonadales	4	Aeromonadaceae	2
						Succinivibrionaceae	2
				Enterobacterales	51	Enterobacteraceae	51
				Pasteurellales	2	Pasteurellaceae	2
				Pseudomonadales	11	Moraxellaceae	6
						Pseudomonadaceae	5
				Vibrionales	4	Vibrionaceae	4
				Xanthomonadales	2	Xanthomonadaceae	2
		Deltaproteobacteria	7	Desulfovibrionales	7	Desulfovibrionaceae	7
		Epsilonproteobacteria	18	Campylobacterales	18	Campylobacteraceae	12
						Helicobacteraceae	6
Verrucomicrobia	3	Verrucomicrobia	3	Verrucomicrobiales	3	Verrucomicrobiaceae	3
Euryarchaeota	3	Methanobacteria	3	Methanobacteriales	3	Methanobacteriaceae	3

Only 138 species, 12% of the total species richness, were recovered by application of both molecular and cultivation-based approaches (Fig. 2). Remarkably, another 20 phylotypes detected by the cultivation-independent studies matched cultivated and fully described bacterial species, which were not previously isolated from the human gastrointestinal tract (Table 3). Examples of such species include *Allisonella histaminiformans* and *Phascolarctobacterium faecium*, which are intestinal isolates from other mammals (105, 164). There are, however, also isolates such as *Aquabacterium commune*, which are typical for non-intestinal ecosystems, in this particular case drinking water (238). Thirteen phylotypes corresponded to organisms recovered from the human oral cavity, and were mainly recovered in studies of microbiota in the upper gastrointestinal tract, rather than faecal samples.

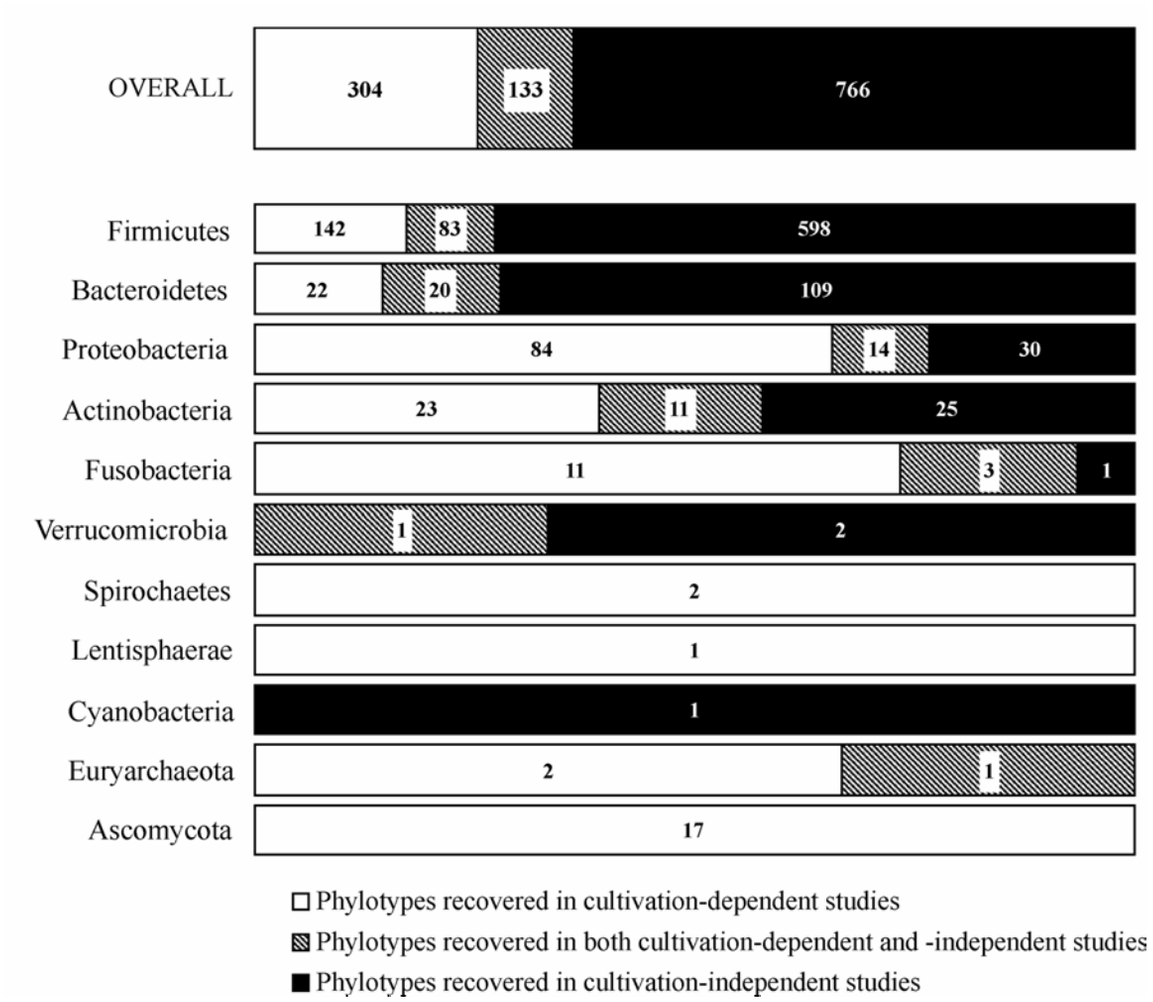


Figure 2 Distribution of gastrointestinal phylotypes over type of the study, given overall and for eight bacterial, one archaeal and one eukaryal phylum. Gastrointestinal isolates that have been fully characterised but lack the SSU rRNA gene sequence were taken into account for construction of this figure.

Cultivation-independent studies of gastrointestinal microbiota have revealed the presence of not-yet described phylotypes, of which some are, based on the SSU rRNA gene sequence, only distantly related to any cultured microbe with known SSU rRNA gene sequence. Those phylotypes form novel, monophyletic clusters that are well separated from cultivated species. In the phylogenetic tree of the human gastrointestinal microbiota those clusters are marked as Uncultured Clostridiales I and II (Fig. 1). Both clusters together comprise 63 distinct phylotypes, which will be, once they are cultured, organised in a number of families and genera, as indicated by their SSU rRNA gene similarity. Another cluster of 14 phylotypes within the Firmicutes phylum could not be classified even at the level of order (Table 2). Finally, 19 uncultured phylotypes within the phylum Bacteroidetes form another distinct cluster of exclusively uncultured organisms (Table 2).

Table 3 Bacterial species typically isolated from other ecosystems but the human gastrointestinal tract that were detected in the human gastrointestinal samples, based on SSU rRNA gene sequence analysis.

Species	**Known ecological niche, (Reference)**	**Reference of study in which SSU rRNA sequence was detected in the human intestine**
Acinetobacter lwoffi	Oropharynx, (388)	(502)
Actinomyces graevenitzii	Saliva, bronchia, (386)	(129)
Actinomyces odontolyticus	Tooth root canal, (367)	(129)
Allisonella histaminiformans	Cecum of horse, (164)	(129, 192)
Anaerococcus vaginalis	Vaginal discharges, (140)	(129, 301)
Aquabacterium commune	Drinking water, (238)	(505)
Bifidobacterium ruminantium	Bovine rumen, (49)	(59)
Burkholderia cepacia	Saliva, (405)	(505)
Corynebacterium durum	Subgingival plaque, (31)	(129)
Corynebacterium sundsvallense	Sinus, (88)	(129)
Dialister invisus	Oral cavity, (120)	(59, 129, 301, 502)
Fusobacterium peridonticum	Gingival crevice, (175)	(502)
Granulicatella adiacens	Throat flora, (90)	(K. Saunier, unpublished)
Leuconostoc argentinum	Raw milk, (113)	(194)
Mogibacterium vescum	Peridontal pocket, (345)	(502)
Neisseria mucosa	Subgingival plaque, (186)	(502)
Phascolarctobacterium faecium	Koala faeces, (105)	(191, 193)
Prevotella shahii	Saliva, (407)	(502)
Rothia mucilaginosa	Oral cavity, (402)	(502)
Stenotrophomonas maltophilia	Oral cavity, (464)	(505)

Cultivation-based and molecular studies provide a somewhat different description of the diversity of the human gastrointestinal tract microbiota, as indicated by the degree of the overlapping findings (Fig. 2). Remarkably, both approaches show unequivocally that Firmicutes are by far the most diverse group, although the community structure of this group is highly underestimated in the reports of cultivation-dependent studies. Similarly, the Bacteroidetes community in human gastrointestinal samples is more diverse than assessed with cultivation-based approaches. However, about two thirds of Bacteroidetes isolates, for which the SSU rRNA gene sequence is available, have also been detected by molecular approaches. This indicates that some of the widely distributed and abundant members of this phylum have already been isolated. Representatives of the Proteobacteria, and particularly subdivisions ε and γ, have so far been detected relatively poorly by molecular approaches. However, as sequencing of the SSU rRNA gene amplicons from an intestinal sample generally describes only the dominant fractions of the microbiota (541), and it is known that Proteobacteria have low abundance in the human gastrointestinal tract (214), such low recovery of this phylum had to be expected.

Molecular studies showed that members of the phylum Cyanobacteria (one phylotype) can be found in the human gastrointestinal samples (129), even though isolates from this group have not been reported. Furthermore, six phylotypes of the subphylum of α-Proteobacteria were detected by SSU rRNA gene sequencing, while there is only one known α-proteobacterial gastrointestinal isolate – *Gemmiger formicilis* (45, 206, 293, 336). Unfortunately, no 16S rRNA gene sequence is available for this strain. Phylogenetic analysis revealed that five of the six α-proteobacterial phylotypes did not cluster within the order Rhizobiales to which *Gemmiger* belongs. Strikingly, all seven cyano- and α-proteobacterial phylotypes were detected in studies where mucosa samples from the upper gastrointestinal tract were analysed (129, 502). It was previously shown that the microbiota associated with rectal mucosa may differ from that present in faeces (542), and recently, two studies indicated that mucosa at upper gastrointestinal parts contains specific phylotypes that are distantly related to any known gastrointestinal isolate (129, 502). These data reinforce the notion that faecal samples do not comprehensively represent the microbiota that can be found in the lumen or attached to mucosal surfaces along the gastrointestinal tract, and that sampling along the gastrointestinal tract, both for cultivation-based analysis and molecular profiling, would enable better description of the microbial diversity of this ecosystem.

As previously mentioned, the diverse community of bacteria in the human gastrointestinal tract is accompanied by representatives of archaea and eukarya. According to our current knowledge, the archaeal community of the human intestine is very simple and consists of only three isolates, of which *Methanosphaera stadtmanae* is rare (320), and *Methanobrevibacter ruminantium* was reported in a single study dating from 1968 (349). The

presence of only one archaeon - the abundant *Methanobrevibacter smithii*- was confirmed by molecular studies (129, 321).

The diversity of eukarya in the human intestine was assessed exclusively by cultivation-dependent approaches. Seventeen *Candida*, *Aspergillus*, and *Penicillium* species were isolated from human intestinal samples (13, 48, 150, 153, 473). However, except for *C. albicans* and *C. rugosa*, eukarya are neither widely distributed nor abundant in the human intestine, according to the results of the cultivation-based approaches.

Species richness estimation

The recent studies of the human gastrointestinal microbiota diversity by sequencing have provided an exploding amount of SSU rRNA gene sequence information. However, due to the fact that only a limited number of individuals has been subjected to the analysis of the human gastrointestinal microbiota and since each subject harbours a unique microbial community, a good representation of the ecosystem has not yet been achieved. Consequently the rank-abundance curve, which can visualise how well an ecosystem has been sampled (218), has a very different shape when compared to the one obtained with data from cultivation-based studies (Fig. 3). The curve is characterised by very strong tailing, as the majority of the phylotypes were found in intestinal samples of only one individual.

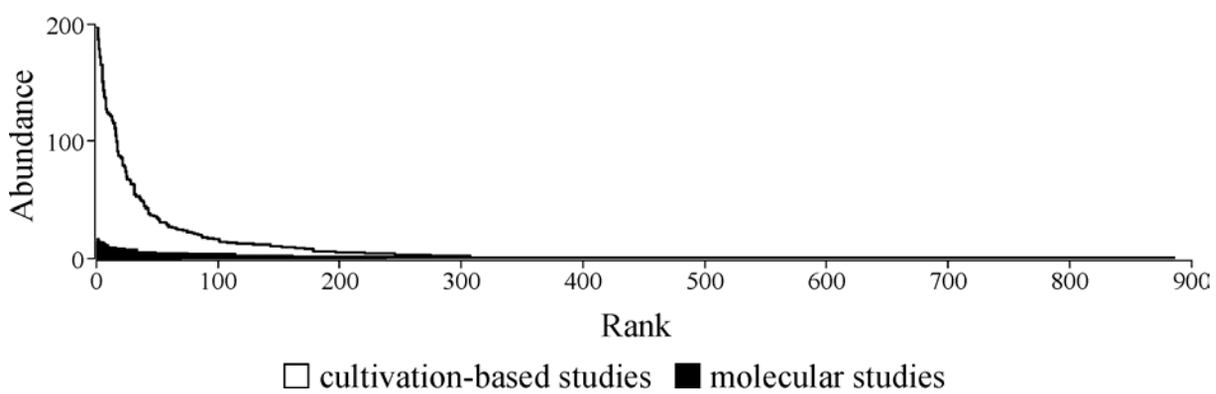


Figure 3 Rank-abundance curve for the phylotypes of the human gastrointestinal tract found in cultivation-based and molecular studies. The abundance of the phylotype is defined as number of individuals that have been reported to harbour the particular phylotype.

Nevertheless, it is remarkable that molecular studies performed to date revealed almost 900 distinct phylotypes based on the microbiota of only 50 individuals, while cultivation studies, which were employed for characterisation of the microbiota of hundreds of subjects have recovered less than half of this species richness. Based on Good's coverage estimates, our current knowledge of the normal human gastrointestinal microbiota diversity

overall is about 30% (333), which contrasts with the good coverage of the microbiota diversity that was achieved in individual studies (129, 461).

Since a good representation of the intestinal diversity has not yet been achieved, it is not possible to come to reliable number of the total species richness on the basis of the present data set (218). Nevertheless, in for the purpose of this review, the Chao1 index was calculated to provide its estimate. Because the archaeal and eukaryal community in the human intestine are rather simple, only the frequency of bacterial phylotypes was taken into account for the calculation. Three values are needed for estimating the species richness based on the Chao1 index: the total number of phylotypes, and the number of the rare phylotypes – these which were detected in only one or two individuals. Based on data from molecular studies that yielded a total of 898 bacterial phylotypes, of which 622 were found in only one individual while 121 in two, the expected diversity of the human gastrointestinal microbiota amounts to about 2,500 phylotypes. This estimate exceeds the estimate based on cultivation-based techniques by one order of magnitude but the true diversity of the human gastrointestinal tract might be even greater, since the Chao1 estimator tends to underestimate the true richness at low sample sizes (218).

Furthermore, the Chao1 estimate, calculated on the basis of frequency of bacterial phylotypes reported in both types of approaches, where out of 1,184 phylotypes, 686 were found in one and 128 in two individuals, showed that expected diversity of the gastrointestinal tract exceeds 3,000 phylotypes. However, this estimate, similar to the one calculated on the basis of molecular studies alone, probably underestimates the true diversity. Moreover, there are several limitations of the applied approach that originate from the dataset that describes the diversity of the human gastrointestinal microbiota, which are discussed below.

Limitations of the approach

The gastrointestinal diversity was analysed on the bases of the SSU rRNA sequence. Unique phylotypes were defined as groups of sequences that have sequence similarity of 98% or less with any other SSU rRNA sequence derived from the human gastrointestinal tract. There is no general agreement about the most appropriate cut-off similarity for a phylotype, and different cut-off values have been applied even in the most recent studies (129, 172). This has a major effect on the obtained measure of diversity. As previously reported, within the same clone library, the number of distinct phylotypes ranges from 148 to 643 when the cut-off value is changed from 97% to 99% (496). The sequence similarity of different SSU rRNA gene copies in the same genome is rarely below 99% (1), although it can vary substantially (506). This variation increases between different strains of the same species but rarely reaches values of 97% (403). Hence, in this review a cut-off value of 98% was chosen, which has been widely applied by others (192, 461) and represents a compromise between three commonly used

values. However, it is well known that many species, especially among Enterobacteriaceae, have more than 98% SSU rRNA gene sequence similarity with other related species. In fact, from 382 distinct gastrointestinal species for which the SSU rRNA gene is available, only 305 would be recognised as unique phylotypes based on their SSU rRNA gene sequence, when using the 98% similarity criterion. Thus, it is likely that some of the identified distinct phylotypes are, in fact, groups of different phenotypes. Until cultured representatives of these phylotypes are obtained, and while SSU rRNA gene sequence remains the only available information, absolute accuracy will not be achieved.

Another factor affecting the accuracy of our analyses originates from the fact that many of the available SSU rRNA gene sequences are partial, and unless they correspond to the same region within the gene, they do not allow proper direct comparison. Only if two sequences match a common, fully sequenced “type phylotype”, ARB software allow grouping of partial sequences into a single distinct phylotype. Otherwise the sequences are recognised as distinct phylotypes even if they correspond to a different region of the very same SSU rRNA gene. Hence, an overestimation of the true diversity might occur. It would be convenient if there were a general agreement that, if partial sequences are provided, they are always obtained from the same variable region of the SSU rRNA gene. As Ludwig and Klenk indicated, the sequence of the V1 region appears to be the most variable and consequently the most informative (285). Therefore, selection of this region seems the most appropriate. However, it has to be acknowledged that sequences generated for analysis by fingerprinting approaches, such as denaturing gradient gel electrophoresis (DGGE) or amplicons obtained by group-specific PCR amplification are often confined to different specific regions of the SSU rRNA gene, due to restricted options for primer design.

Each of the techniques used for the assessment of the diversity of the human gastrointestinal microbiota has its limitations and induces biases in the obtained portray of the microbial diversity. The limitations of the cultivation-based approaches to adequately resemble the diversity of the gastrointestinal microbiota are nowadays well acknowledged (468, 491). Moreover, the SSU rRNA gene sequencing-based analyses of the microbial diversity are also giving biased picture, and among many factors that influence the obtained information, effectiveness of cell lyses, and PCR biases are highly relevant. Although it is generally considered that SSU rRNA gene sequence assessed diversity of an ecosystem is representing the most dominant part of the microbial community, this is is necessarily accurate as the sequence of the used PCR primers also induces biases. The reason for this effect is that, even when the total bacterial microbiota is targeted, different PCR primers are used, and primer sets can have preferences for different phylogenetic groups (129, 461). This is because none of the universal primer sets has the claimed specificity or universality (215). This effect, together with the problems related to the cell lyses, have led to the underestimation of the members of Actinobacteria phylum within the human gastrointestinal

microbiota, which appear to be quantitatively relevant members of this ecosystem, based on the results of enumerative techniques such as florescent *in situ* hybridisation (268). Finally, cloning and sequencing procedures are in most cases performed only once. Remarkably, the analysis of microbiota of a single individual by the use of different numbers of PCR cycles showed some differences in the recovered clone libraries (60, 461). This might be an exclusive effect of the number of PCR cycles, as suggested by authors, but it might also indicate limitations of the cloning and sequencing technique itself. If so, the microbiota of each individual might be even more complex than revealed in a single cloning and sequencing experiment.

Concluding remarks

The human gastrointestinal microbiota is an important element of the human body, which is receiving increasing attention. Based only on the data which was generated until the beginning of the 2006, the human gastrointestinal tract harbours almost 1,200 distinct microorganisms. The complete coverage of the diversity has not yet been achieved, but the total number of already reported distinct phylotypes exceeds the most ambitious estimates previously given for this ecosystem. However, so far the majority of the phylotypes have been detected only in a single individual, indicating that many more gastrointestinal inhabitants are to be discovered, and the total diversity will probably be measured in thousands of species. Therefore, the human gastrointestinal microbiota, often referred to as forgotten organ of the human body (353), is still insufficiently described even at the composition level. Since reliable insight into the gastrointestinal diversity is essential for providing a reference framework to study its dynamics in time and space, analyse its functions and characterise host-microbe interactions, more attention should be given to the simple description of this ecosystem. To this end, the availability of curated databases of rRNA sequences, allowing for adequate organisation of redundant data without loosing useful information from all previously deposited and future sequence entries such as source of isolation, is an absolute requirement. The inventory of currently available data described here is the first important step towards this goal.

Acknowledgements

The authors would like to thank Dr Colin Ingham for reviewing this manuscript and useful suggestions concerning phrasing and sentence structure.

CHAPTER 4

Mirjana Rajilić-Stojanović, Hans H. G. J. Heilig, Douwe Molenaar,

Hauke Smidt and Willem M. de Vos

Modified version of this chapter was submitted for publication

A Generic Design Strategy for Phylogenetic Microarrays Based on SSU rRNA Gene Sequence Diversity - the Human Gastrointestinal Microbiota as a Case

A generic oligonucleotide probe design strategy was developed, based on two hypervariable regions of the small subunit ribosomal RNA (SSU rRNA) gene, as an efficient approach for selection of probes with similar predicted hybridisation behaviour suitable for use in microarray technology. The systematic probe design approach can be readily applied to any complex ecosystem and allows easy addition of probes that target newly discovered members of the ecosystem. Applied to the human gastrointestinal microbiota, the presented probe design strategy was used to develop the HITChip (Human Intestinal Tract Chip), containing more than 5,000 oligonucleotides targeting over 1,000 intestinal microbial phylotypes. The HITChip illustrates the potential of the used approach of phylogenetic microarrays design, as by assigning of each probe to its phylogenetic specificity it enables rapid profiling of intestinal samples and provides information on the microbial diversity. With the possibility for a relative quantification even for species-level phylotypes, the HITChip will bridge the gaps in our knowledge in the quantitative and qualitative description of the human gastrointestinal microbiota diversity.

Introduction

Ribosomal RNA (rRNA) gene sequence-based analysis of the microbial diversity was for the first time applied more than two decades ago (455). Since then, a diverse range of rRNA–based techniques, mostly targeting its small subunit (SSU), that include PCR sequence analysis and *in situ* hybridisation have become standard procedures for describing complex microbial ecosystems (12). Diagnostic approaches based on the SSU rRNA gene or other molecular biomarkers proved to be more accurate, faster and less laborious when compared to traditional cultivation techniques. Subsequent molecular surveys have revealed an unexpected complexity in the microbial diversity. As an example, more than 1,000 microbial species-level phylogenetic types (phylotypes) have been shown to inhabit the human gastrointestinal tract, one of the most densely populated microbial ecosystems (**Chapter 3**). In order to effectively study such complex ecosystems and adequately address the effect of biotic and abiotic environmental factors on microbial communities, there is a need for establishing novel, molecular techniques with improved resolution and high throughput, such as phylogenetic microarrays.

Microarrays have initially been developed for gene expression profiling of individual organisms but this technology has been expanded for strain typing, and the analysis of diversity and functionality of complex ecosystems (167, 182, 499). The subset of microarrays that can provide a phylogenetic information are phylogenetic microarrays which are principally based on the SSU rRNA gene (55). Although other genes can be used as bases for the phylogenetic identification, in many microbial ecosystems a large proportion of inhabitants has only been detected by SSU rRNA gene sequencing, and therefore, techniques targeting this gene or the encoded rRNA product provide the most complete information about its diversity. A recent application of a comprehensive SSU rRNA gene-based phylogenetic microarray showed that this technology provides superior diagnostic power for the analysis of the microbial community structure when compared to the clone library (548).

Phylogenetic microarrays, similar to any SSU rRNA gene–based technique, primarily depend on the design of probes at different levels of specificity. However, as the microbial world is for a large proportion unknown, an accurate design of universal or specific signature sequences within the SSU rRNA gene is not yet feasible. In this context, following the redesign of universal primer sets to optimise universal amplification of bacterial sequences (507), it appeared that the most these matched maximally 74% of the presently available SSU rRNA gene sequences (215). Furthermore, specific probe design for comprehensive coverage of the occurring microbial diversity is a laborious process and therefore feasible only for ecosystems of limited complexity. One alternative to the conventional probe design is the “tiling” of the whole gene sequence (80), which in a case of thousands of SSU rRNA gene sequences is not feasible with the current array technology. Another option is the design of

platforms with probes that target only a specific hypervariable region of the SSU rRNA gene (524). However, these do not have the capacity to target all of the more than 300,000 SSU rRNA gene sequences present in today's databases (W. Ludwig, personal communication). Hence, there is a need to design ecosystem-specific phylogenetic microarrays.

In this paper we present a generic oligonucleotide probe design strategy based on two hypervariable regions of the SSU rRNA gene. This approach is high throughput, can produce probes with similar predicted hybridisation behaviour, and can be readily applied to any complex ecosystem. Moreover, it allows easy addition of probes that target newly reported inhabitants of the ecosystem. The potential of this probe design strategy for phylogenetic microarrays was tested on the human gastrointestinal microbiota in a microarray, further on referred to as HITChip (Human Intestinal Tract Chip).

Materials and Methods

In silico analysis of SSU rRNA gene sequences

The SSU rRNA gene sequences of the human gastrointestinal microbiota were extracted from the ARB database release 2002 (286), amended with sequences of newly reported isolates and sequences amplified from gastrointestinal samples. Furthermore, the database was upgraded by assigning each of the human gastrointestinal sequences to their phylogenetic position. However the use of nomenclatural taxonomy is hampered by the fact that many SSU rRNA sequences obtained from the human gastrointestinal samples represent uncultured organisms which are distantly related to any cultured representative, and therefore their taxonomic position cannot be accurately determined. In addition, phylogenetic position of many of the cultivated inhabitants should be reconsidered (**Chapter 2**). Therefore, the sequences were organised into alternative SSU rRNA-based phylogenetic groups similar to those previously proposed for *Clostridium* genus (87).Three levels of phylogeny were defined: 1) Level one corresponds to the phylum level, or in case of Firmicutes to the *Clostridium* cluster 2) Level two includs the group of sequences with 90% or more sequence similarity, and 3) Level three represents unique phylotypes that were defined as species for cultivated microorganisms, or representatives of each monophyletic group with $\geq$ 98% sequence identity for clones corresponding to uncultured microorganisms. Nine variable regions of the SSU rRNA gene were exported for 1,140 sequences representative of prokaryotic gastrointestinal phylotypes. The sequences were exported using the termini filter as implemented in ARB (286), and appropriate bordering nucleotide positions (Table 1). These were determined based on the alignment positions of primers available for the amplifications of each of the variable regions. Unique SSU rRNA sequences were identified using Pivot Tables built from the exported sequence data in Microsoft Excel.

Table 1 Position and diversity of variable regions of the SSU rRNA gene extracted based on the SSU rRNA gene sequence of 1,140 human gastrointestinal phylotypes.

Variable region	Starting and ending position (*E. coli*)	Reference	Number of sequences	Number of distinct sequences	Percentage of distinct sequences
V1	65-107	(47)	849	743	87.51
V2	174-235	(262)	875	759	86.74
V3	442-492	(216, 262)	1107	687	62.06
V4	705-763	(216)	1002	556	55.49
V5	822-879	(216, 262)	988	682	69.03
V6	985-1047	(196)	895	754	84.25
V7	1115-1175	(Roest, unpublished)	874	635	72.65
V8	1253-1313	(352)	865	699	80.81
V9	1420-1481	(262)	446	365	81.84

Probe design

Antisense oligonucleotide probes were designed based on the reverse complement of two variable regions that were selected based on their low redundancy. Reverse complements of the SSU rRNA gene sequences were produced in ARB prior to export. Variable regions were split into three overlapping probes of 24 nucleotides (nt) with the use of available Microsoft Excel functions. The melting temperature (Tm) of probe-perfect match target duplexes was predicted using an earlier reported equation (411), assuming a sodium ion concentration of 0.5 M. The probe length was adjusted by extending or restricting the overlap between probes until the corresponding Tm fitted a narrow range of 70 +/-2^{o}C. The size of the probes was kept between 18nt and 30nt to enable similar duplex kinetics. Following this probe design procedure, six probes were designed per each complete SSU rRNA gene sequence.

Preparation of artificial mixtures

Artifical mixes were prepared by combination of human gastrointestinal phylotypes obtained in two SSU rRNA clone libraries previously prepared from human gastrointestinal samples (461, 537). SSU rRNA gene inserts were amplified by the use of appropriate vector-targeting primers using GO Taq PCR amplification kit (Promega, Leiden, The Netherlands). PCR was performed on 1 µl of cell-lysates (containing approximately 10 ng of plasmid DNA) from

clones using a set of Sp6 and T7 primers for clones obtained by Zoetendal et al. (11) and the pUAg primer set for clones obtained by Suau et al. (12).

Faecal DNA extraction

Total DNA was extracted from faecal material using a modified protocol of the QiaAmp DNA Mini Stool Kit (Qiagen, Hilden, Germany) (539). DNA yield was quantified using NanoDrop ND-1000 spectrophotometer (NanoDrop® Technologies, Wilmington, DE). DNA concentration was adjusted to 10 ng/μl and was used as a template for PCR amplification.

RNA preparation

The SSU rRNA gene was re-amplified from obtained clone inserts or amplified on obtained faecal DNA using the *T7prom*-Bact-27-for (5'-TGA ATT GTA ATA C GA CTC ACT ATA GGG GTT TGA TCC TGG CTC AG–3') and Uni-1492-rev (5'-CGG CTA CCT TGT TAC GAC-3'), which ensured the introduction of a T7 promoter sequence at the 5' terminus of the rRNA gene amplicon. PCR reactions were carried out in a final volume of 50 μl, and 10 ng of DNA samples were used as template. Samples were initially denatured at 94˚C for 2 min followed by 35 cycles of 94˚C (30 sec), 52˚C (40 sec), 72˚C (90 sec) and a final extension at 72˚C for 7 min. The PCR products were purified by using the DNA Clean and Concentrator kit (Zymo Research, Orange, USA), according to the manufacturer's instruction. Final DNA concentration was determined by using a NanoDrop spectrophotometer as described above..

In vitro transcription of T7-promoter carrying SSU rRNA gene was performed according to the manufacturer's protocol using the Riboprobe System (Promega, La Jolla, USA), 500 ng of the T7-16S rRNA gene amplicon, and, besides rATP, rGTP, rCTP, a 1:1 mix of rUTP and aminoallyl-rUTP (Ambion Inc., Austin, Tx, USA). The transcription reaction was performed at room temperature for 2 h., the template DNA was digested applying the Qiagen RNAse-free DNAse kit (Qiagen, Hilden, Germany), and RNA was purified using the RNeasy Mini-Elute Kit (Qiagen, Hilden, Germany). RNA yield was measured as described above.

Amino-allyl-modified nucleotides were coupled with CyDye using the Post-Labeling Reactive Dye (Amersham Biosciences, Little Chalfont, UK), previously dissolved in 84 μl dimethyl sulfoxide. Labelling reaction was performed in a 25 mM sodium bicarbonate buffer (pH 8.7) by adding 20 μl of dissolved CyDye to 2 μg of purified RNA in a final volume of 40 μl. Samples were incubated for 90 min in the dark at room temperature. The reaction was stopped by adding 15 μl of 4M hydroxyl-amine and incubating for 15 min in the dark. RNAse-free water was added to 100 μl and labeled RNA was purified and quantified as described above.

Microarray production and hybridisation

Microarrays were custom synthesised by Agilent Technologies (Agilent Technologies, Palo Alto, CA). The oligonucleotide probes were extended at the 3' end (at the array support side) by 10-nt long T spacers and were printed on the array using in situ surface-attached oligonucleotide probe synthesis (54). Arrays used in this study were of the 2×11K format, with two arrays per glass slide. Each array was hybridised with two samples, labelled with Cy3 and Cy5, respectively. Combined Cy3- and Cy5-labelled target mixtures were fragmented by adding 1 μl of Ambion 10× fragmentation reagent (Ambion Inc., Austin, Tx, USA), and incubation for 20 minutes at 70˚C, according to the manufacturer's instruction. Fragmentation was stopped by adding 1 μl of Ambion stop solution. Hybridization mix was prepared by adding to the RNA mixture 31.5 μl of 20× SSC (412), 6.3 μl of 10% SDS (412), 25 μl of Agilent Control Target mix and RNAse-free water to a total volume of 210 μl. Hybridization was carried out at 62.5^{o}C in a rotation oven (Agilent) for 16 h. Slides were washed at room temperature in 2× SSC, 0.3% SDS for 10 min and at 38˚C in 0.1× SSC, 0.3% SDS for 10 min. SDS was completely removed by washing the slides in 0.06× SSPE (412) for 5 min, followed by a quick dry with compressed air.

Support database construction

The data were stored in a custom-designed relational database, which runs under the MySQL database management system (http://www.mysql.com/). The database design was divided into sections covering sample annotation, microarray raw and normalized data, microarray design, probe characteristics, probe-target pairs, and the phylogenetic position of the SSU rRNA target molecules. Using the sequences of the SSU rRNA molecules stored in the database that, in addition to unique phylotypes, contained sequences with more than 98% sequence similarity to one of the unique phylotypes (total 2,681 sequences), putative probe-target pairs were selected using a standalone version of the BLAST program from NCBI (11) with the following parameter settings: gapopen=3, gapextend=2, matchreward=1, mismatchpenalty=-1, wordsize=7, and expect=1. A theoretical melting temperature (Tm) of each probe-target pair was calculated using the nearest neighbour algorithm (415) assuming sodium ion concentration of 0.5 M and a nucleotide concentration of 10^{-12} M. In this probe-target pair hybrid calculation the whole probe, including the poly-T spacer was taken into account, leaving open the possibility that some nucleotides from the spacer took part in the hybrid pair. The nearest neighbour algorithm was also used to calculate the theoretical melting temperature (pmTm) for a perfectly matching target for each probe, excluding the poly-T spacer region. Raw microarray data files from the feature extraction software as well as tab-delimited files containing information about the sample annotations were imported into the database by converting them to SQL scripts that were subsequently executed.

Microarray data extraction and analysis

Data were extracted from microarray images using the Agilent Feature Extraction software, version 7.5 (www.agilent.com). Normalisation of the microarray data was performed in three steps: spatial normalisation, outlier detection and sample-wise quantile normalisation.

Spatial normalisation was carried out for each channel separately to correct spatial defects in the arrays. Spatial defects were especially clear in a number of arrays in the lower signal ranges. Therefore, the 70% lower quantile of foreground signals were fitted to a two-dimensional polynomial surface using an implementation written by B.D. Ripley in R statistical software (http://www.r-project.org/) of the loess algorithm (84). A spatially normalised signal for each spot was calculated by subtracting the local loess predicted signal from the raw foreground signal and adding the global minimum of the fitted polynomial surface. The resulting signals were stored as the spatially normalised signal in the database.

Outlier measurements in the spatially normalised signals were detected and flagged automatically. Since each sample was labelled with both dyes and hybridised to two different arrays printed on different slides, and since each probe was spotted at least twice on the array, at least 4 replicate measurements per probe and sample were available. Outliers in these replicate measurements were detected using an implementation in the "outlier" package in R of a chi-square test proposed by Dixon (117). The test was applied to the logarithmically transformed signals, and the average over the whole probe set of the variances per probe was substituted in the variance parameter for this test. The significance level of outlier detection used was set at 0.001.

The reproducibility of the hybridisation was assessed by calculating the Pearson's linear correlation of spatially normalised signals excluding outlier spots (466). A correlation coefficient of minimally 0.98 was considered to indicate a satisfactory reproducibility. Further normalisation was performed using only arrays that contained data of satisfactory correlation.

A quantile normalisation was applied to the collective spatially normalised measurements of each sample (57). The rationale assumption behind this normalisation is that the distribution of signals should be the same for each measurement of a sample. The sample-wise quantile normalised signals were also stored in the database and used for further analyses.

Calculation of probe profiles

For each probe, the specific target SSU rRNA sequences in the database were identified by selecting those targets for which the theoretical melting temperature was equal to or higher than the perfect-match melting temperature of the probe (Tm $\geq$ pmTm). The phylogenetic position of these targets was scored based on the ARB database assigned clustering of human gastrointestinal sequences described above. The lowest phylogenetic level to which all theoretically hybridising targets belonged was recorded as the specific phylogenetic level of

the probe. It should be realised that this specific phylogenetic level for a probe is not an absolute entity, but is influenced by the size and diversity of the SSU rRNA sequence target set in the database. The larger and more diverse this target set, the better the specific phylogenetic level represents the true phylogenetic specificity of the probe.

Hierarchical clustering of probe profiles was carried out by calculating a distance matrix between the samples based on the squared difference between each pair of profiles (Euclidian distance). The distance matrix was used in the hclust implementation in R of a hierarchical clustering algorithm. The agglomeration method used in this algorithm was Ward's minimum variance method.

Probes were considered responsive if signals obtained after hybridisation with perfect match target sequences were at least two-fold higher than the hybridisation signal obtained with Agilent negative control probes.

Results and Discussion

Identification of hypervariable SSU rRNA regions

In this paper we present the *in silico* assessment of the diversity of variable regions of the SSU rRNA gene based on a curated database of a specific ecosystem; describe a probe design procedure based on two hypervariable regions with minimal redundancy; and test the potential of such probe design strategy for the design of a flexible microarray platform. As a model, the human intestinal microbiota was selected as it has considerable complexity, shows significant coverage of SSU rRNA sequences, and is subject to intensive studies (129, 275).

The sequence of all nine SSU rRNA variable regions was extracted from an ARB database that contained 1,140 SSU rRNA gene sequences representing the known prokaryotic diversity found in the human gastrointestinal tract. Since the majority of these sequences does not cover the full length SSU rRNA gene, an inconsistent number of sequences per variable region, varying from 446 to 1,107, could be extracted from the database (Table 1). The level of variability differed between variable regions as the percentage of distinct sequences was in a range from 55.5% to 87.5%. Variable regions V1, V2, V6, V8 and V9 had more than 80% of distinct sequences and were considered hypervariable. Further sequence analysis showed that there were only 12 pairs of phylotypes that had identical sequence of both the V1 and the V6 variable regions. Therefore, the joined sequence information of those two variable regions reduces dramatically the proportion of redundant sequences - from 12.5% and 15.5% for V1 and V6, respectively, to 3.4%. Furthermore, the 12 pairs of phylotypes that had identical sequence of both variable regions are highly similar along the whole SSU rRNA gene sequence, with average similarity of 99.35% (Table 2). This indicated that merged sequence information of the V1 and the V6 variable regions contains sufficient sequence information to

discriminate between SSU rRNA gene sequences with less than 98% sequence similarity. In addition, 63 phylotypes (5% of the total dataset) shared identical V1 orand V6 regions with other phylotypes, which were different per variable region.

These results expand previous findings on the hypervariability of the V1 and the V6 SSU rRNA gene regions with a limited set of SSU rRNA sequences (47, 196). In addition, they indicate that also in an ecosystem that is composed of only nine prokaryotic phyla (23, 129), these two variable SSU rRNA regions provide sufficient and discriminating information.

Table 2 Pairs of human gastrointestinal species and yet uncultured phylotypes that can not be distinguished based on their variable region V1 and V6 SSU rRNA gene sequence.

Species	**GenBank accession number**	**Percentage of overall SSU rRNA gene similarity**
Bifidobacterium sp. PL1	AF306789	98.38
uncultured *Bifidobacterium* sp.15D	AF275886	
Bifidobacterium breve	AB006658	97.69
Bifidobacterium infantis	D86184	
Bacillus flexus	AB021185	98.79
Bacillus megaterium	D16273	
Enterococcus casseliflavus	AF039899	99.86
Enterococcus gallinarum	AF039898	
Eubacterium budayi	AB018183	99.32
Clostridium barati	X68174	
Clostridium paraperfringens	M59102	98.94
Clostridium paraputrificum	AB032556	
Uncultured bacterium HUCA2	AJ408958	98.96
Uncultured bacterium HUCA26	AJ408977	
Citrobacter braakii	AF025368	99.74
Citrobacter freundii	AB006658	
Raoultella planticola	Y17663	99.09
Enterobacter aerogenes	AB004750	
Klebsiella pneumoniae	AB004753	99.37
Serratia liquefaciens	AB004752	
Vibrio fluvialis	X74703	99.51
Vibrio furnissii	X74704	
Yersinia rohdei	X75276	98.78
Yersinia kristensenii	X75278	

Probe design

The feasibility of the use of the V1 and the V6 variable regions for phylotype-specific identification was initially tested by dot-blot hybridisations on a macro-array. The complete SSU rRNA gene of 98 human gastrointestinal phylotypes that were obtained in two studies (461, 537), was blotted and hybridised to PCR amplified variable regions of a selection of these phylotypes. The removal of the primer sites in enzymatic reactions (196) enabled hybridisation with the core of the variable region, and indicated specific hybridisation (results not shown). However, probe preparation through PCR and primer site restriction proved to be very laborious, and not feasible to be applied at a large scale beyond pilot-macroarrays. In addition, the optimisation of hybridisation conditions for the V1 region was challenging due to the fact that besides variation of the sequence, the length of the V1 region varies considerably. Hence, a microarray approach was developed with an uniform hybridisation behaviour, as detailed below.

The specific hybridisation of the variable regions indicated the possibility of use of pre-selected hypervariable segment of the SSU rRNA gene sequence for microarray probe design that would enable systematic probe production. Presently available algorithms for microarray probe design, for example (136, 276), produce probes which specificity is determined on the bases of set of target and non-targeted sequences. Even though this is considered to be a necessary step for the production of specific probes, the resulting probe sequence will largely depend on the size of database. Similarly to the universal primer design, as discussed above (215), it is likely that each optimally designed oligonucleotide probe sequence should be reconsidered with the increased database size. Probes, designed using the strategy presented in this paper, do not require such reconsideration since no specificity claim was set during their design, although phylogenetic specificity of those probes will be more accurately if the SSU rRNA sequence database is more informative.

To allow for uniform hybridisation behaviour, the sequence of the V1 and V6 hypervariable regions was separated into three 24-nt probes. As the length of the variable regions was in the range from 24 to 82 base pairs, the size of overlap between the probes varied. The predicted melting temperature of 4,809 probes followed a unimodal symmetrical distribution over a wide range of temperatures. After manipulation of the probe size by extending or shortening by maximally 6 nucleotides, 97% of the probes fitted the desired range of 70 +/-2^{o}C.

With the application of the probe design strategy described above, only a proportion of probes can have distinct sequences. In the presented study, 3699 out of 4,809 oligonucleotide probes based on the SSU rRNA gene sequence of 1,140 gastrointestinal phylotypes were distinct probes (76.94%). Recently, another microarray platform for the rapid profiling of complex microbial ecosystems applied on the gastrointestinal microbiota was reported (359). The probe design strategy reported there was based on 359 gastrointestinal phylotypes found

in intestinal samples of three individuals and consisted of tiling the whole SSU rRNA gene, and selecting the five most specific 40 nt-long probes for each targeted phylotype. Remarkably, the percentage of unique sequences in this microarray system (68.50%) was similar to the one obtained described here based on sequence information of 3 times more phylotypes Although the two systems are not directly comparable as the length of the probes and the phylotype cut-off value were different, this result indicates that probe design based on hypervariable segments of the SSU rRNA sequence will produce a similar proportion of distinct sequences when compared to one where the whole gene is taken into account.

Reproducibility assessment

The influence of the following three factors on the reproducibility of the HITChip readouts was tested: 1) DNA extraction; 2) PCR amplification and *in vitro* transcription; and 3) microarray hybridisation. The maximal obtained reproducibility values were 0.967 for different DNA extractions, 0.996 for different PCR amplifications and *in vitro* transcription and 0.999 for different hybridisations (Fig. 1). In general, the reproducibility of data obtained with different PCR products was only slightly poorer than those obtained with the same PCR product, and can be explained by failure to effectively normalise the spatial defects of individual arrays rather than PCR introduced bias. Therefore, the standard procedure for sample hybridisation included a single PCR, although an additional PCR amplification was performed on samples that did not reach satisfactory reproducibility with two initial hybridisations. The experimental variation in DNA extraction only slightly affected the hybridisation results and the reproducibility of experiments performed with different DNA extractions was in the range of the technical reproducibility of expression microarrays (263, 511). The median reproducibility of 230 samples analysed so far using the HITChip platform was found to amount 0.995 (data not shown).

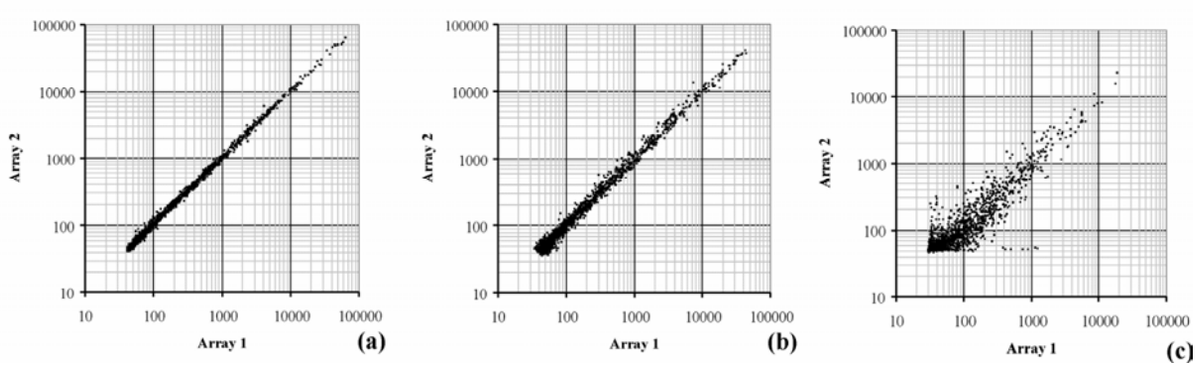


Figure 1 Plots showing the correlation of HITChip hybridisation signals obtained with (a) different hybridisation (Pearson's correlation 0.999), (b) different PCR amplification and in vitro RNA transcription (Pearson's correlation 0.996) and (c) different DNA extraction (Pearson's correlation 0.967).

Identification of phylotypes

The possibility of phylotype-level identification was tested by analysis of artificial mixes containing in vitro transcribed SSU rRNA of up to 40 distinct phylotypes. Theoretically, a given phylotype should be detected by all six probes targeting the two variable regions V1 and V6. Hybridisation results showed that probes hybridising with their respective perfect match targets yielded variable signal intensity with up to 100-fold difference. In addition, a few probes did not yield a significant hybridisation signal under the conditions used in this study.

The strong variation of the hybridisation signal intensity from probe-target duplexes with similar predicted hybridisation behaviour observed in our study is in fact a widely reported phenomenon (56, 284, 380). Several factors are expected to influence probe-target duplex stability. One of the relevant factors is the presence of a stable secondary structure at the position in the SSU rRNA molecule that is targeted by a probe (527). This strong influence of the probe position was also shown experimentally for microarrays (359), and is a well-known phenomenon in FISH analysis (37). However, the probe design strategy described here produces oligonucleotides that target only two short segments of the SSU rRNA, i.e. the V1 and V6 hypervariable regions, and therefore a strong influence of this factor was not expected. The formation of hairpin structures has been reported as another relevant parameter that influences the behaviour of oligonucleotide probes (347). Therefore, the mfold web tool (544) was used to predict potential hairpin structures for probes used in this study. However, no evidence was found for a correlation between the presence of predicted hairpin structures and the signal intensity.

The criterion introduced above, according to which all six probes designed for phylotype should give positive signal for the detection of individual phylotype had to be adapted because some probes did not yield a significant signal. More specifically, we required that at least two out of three probes per variable region should show a signal above the background to call the presence of a phylotype. This modification still allowed for accurate identification of simple mixtures composed of ten species (data not shown), which indicated that phylotypes of relatively simple ecosystems, such as gastrointestinal microbiota of babies or ileostomy patients can be identified unambiguously. However, in complex ecosystems, where a large degree of cross-hybridisations with similar sequences from related organisms (167) can be expected, it is likely that the accuracy of the detection will be affected.

The variation of the probe signal intensity that hampers accurate phylotype identification can certainly be improved by inclusion of other important parameters in the prediction of the probe-target duplex behaviour. Although for some Tm prediction algorithms, such as nearest neighbour algorithm (415), are reported to relatively reliably predict probe behaviour (195, 519) other algorithms that for modelling experimental data based on the Gibson’s free energy of perfect match duplex formation continue to be developed (195).

According to some authors the oligonucleotide probe behaviour seems to be influenced by yet unidentified factors (380), and efforts are being made to define new parameters that determine hybridisation kinetics (519). Therefore, the increasing knowledge of hybridisation kinetics coupled with increasing knowledge of gastrointestinal diversity might improve our ability to design efficient algorithms for the deconvolution of signal patterns, allowing for more accurate identification at phylotype-level resolution in more complex communities.

Phylogenetic fingerprinting

The human gastrointestinal microbiota, as many other microbial ecosystems, is a still insufficiently described ecosystem, and each new study of the diversity of this ecosystem fetches a significant proportion of novel phylotypes (129). Additionally, it has been found that about 70% of the human gastrointestinal microbiota is subject-specific (275). Therefore, as the probes can only be designed based on the detected diversity of a limited number of individuals, several phylotypes occurring in the gastrointestinal tract of other individuals will potentially not be identified even with the most optimal methods for deconvolution of hybridisation profiles.

Even though the identification at the species level is an ultimate goal of diversity microarrays, such goals should be set only when more is known about the diversity of the targeted ecosystem. Presently, when samples that contain unknown sequences are being analysed on a microarray, probes designed to be specific to known sequences can cross-hybridise to unknown sequences from related organisms (167). Such cross-hybridisations can provide meaningful diagnostic information about the composition of complex ecosystem that is being analysed. Hence, the HITChip could also be used as profiling or fingerprinting tool, where assignment of each probe to its phylogenetic specificity upgrades the obtained profiles with the phylogenetic information. An example of profiles obtained for eight subjects at three time points over a period of two months indicated that the HITChip profiles provide specific signatures of each individual (Fig. 2). This finding is consistent with results of other SSU rRNA-based techniques such as DGGE or sequencing of SSU rRNA gene clone libraries (275, 537). In contrast to other fingerprinting techniques the HITChip profiles contain phylogenetic information, and therefore, enable an immediate identification of discriminating bands in a profile. Although the accuracy of the HITChip technology is not yet on the level of the clone libraries, obtaining of the phylogenetic fingerprints can provide relevant information noticeably faster, less laborious and at lower costs, which gives an advantage to the presented microarray technology especially for large scale trials.

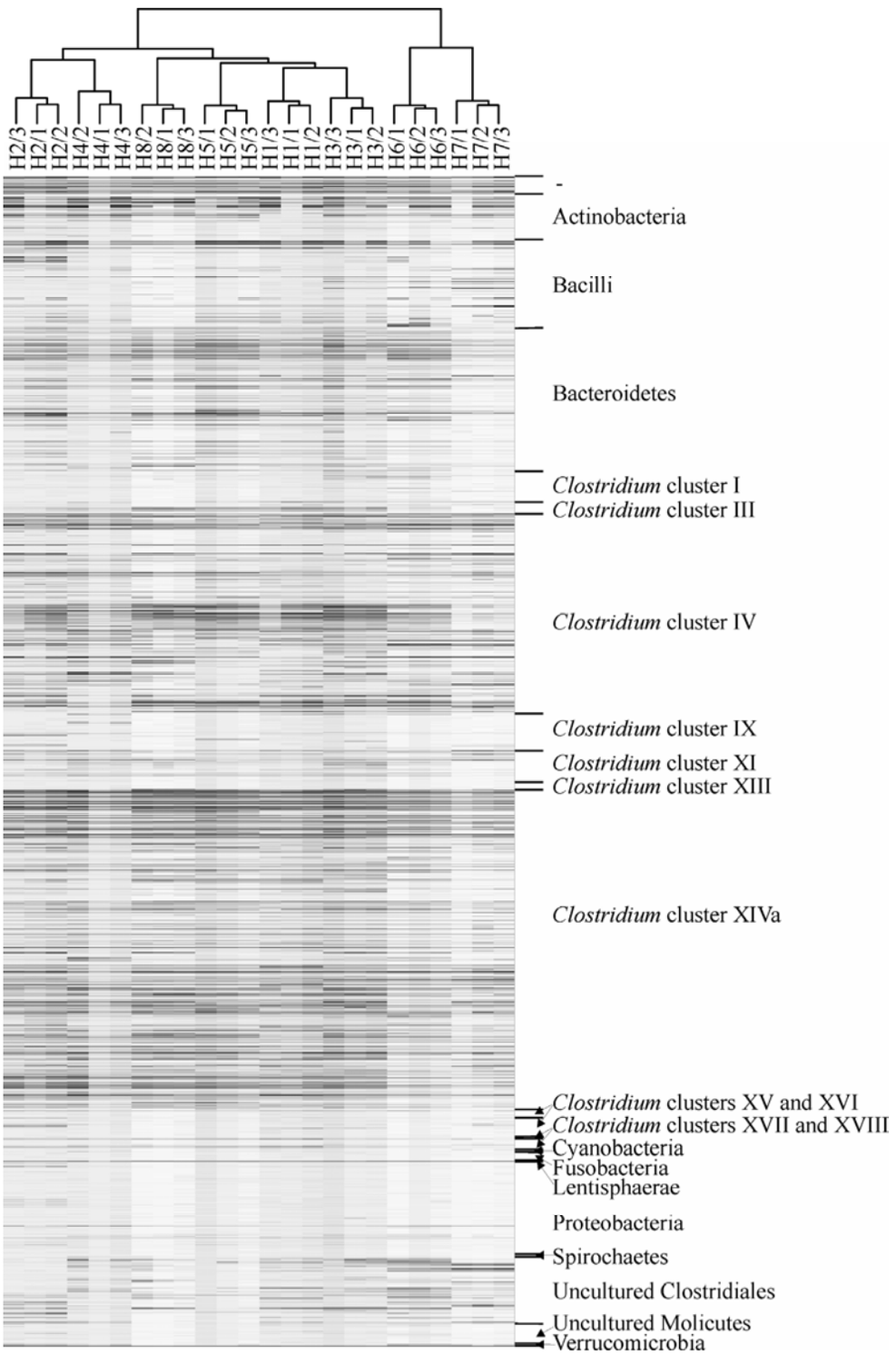


Figure 2 Clustering of the HITChip phylogenetic fingerprints of the human gastrointestinal microbiota of faecal samples of eight subjects of three samples collected during two months period with one month time interval between collection dates. Samples are encoded by H1-H8, while 1, 2 or 3 indicate sample collection sequence. The highest phylogenetic level of specificity of probes (Level 1) is depicted on the right panel of the figure.

Quantification potential

Due to the significant variation of the hybridisation signal, direct quantification of individual phylotypes in a mixture was found to be impossible. However, the relative change of hybridisation signal of phylotype-specific probes in response to the quantitative change of the respective phylotype in a complex mixture appeared to be directly proportional (Fig. 3). The experiments were performed with ten phylotypes varying in quantity from 0.1 to 3%, and 30 phylotypes as a background mixture of constant composition. The change of hybridisation signals could be used as an indicator of relative quantitative change of a phylotype with the detection limit of 0.1% of the total SSU rRNA pool hybridised to the array.

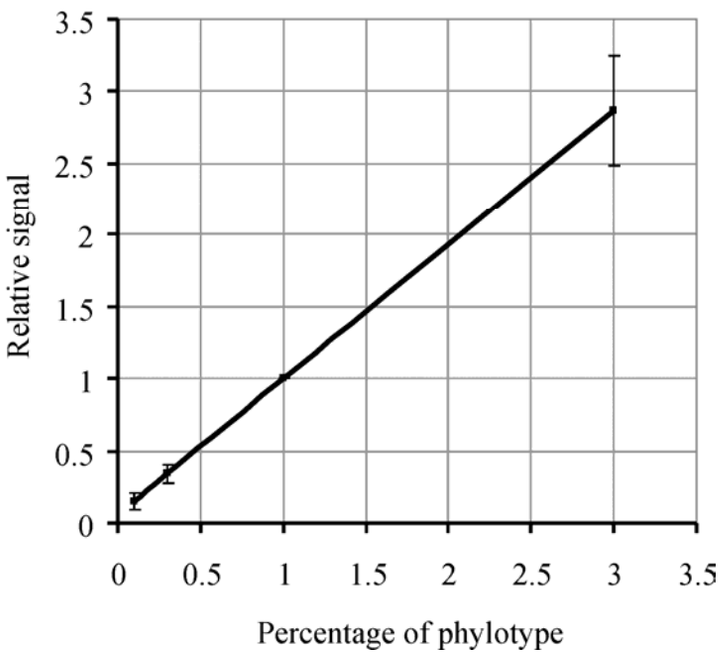


Figure 3 The relative change of the hybridisation signal of the phylotype-specific HITChip probes with the change of the relative abundance in the artificial mix, assessed for ten gastrointestinal phylotypes. phylotypes in the target mix. Average data with the respective standard error is presented.

The observed variation of the probe signal intensity was not dependent on the phylogeny of the targeted phylotype. Hence, it was assumed that the variation of the hybridisation signal will level out within a group of probes, and consequently allow the quantitative comparison between groups. However, due to the fact that the number of SSU rRNA operons differs between species such group, quantification cannot be directly converted to cell numbers present in a biological sample. In fact, the number of genomic SSU rRNA gene copies can vary significantly between 1 and 15 (253). Once genomes of more members of the human gastrointestinal microbiota are sequenced, including the current majority of yet uncultured microorganisms, it will be possible to correct for such differences and hence directly use signal intensity for the quantification of a given phylogenetic group of interest. To validate the possibility of the use of HITChip results for group level quantification, the average of the HITChip hybridisation signals per phylogenetic group were compared with

quantification data obtained by another often used, but low throughput technique to quantify bacterial groups, fluorescent in situ hybridisation (FISH) (268). The relative quantitative data obtained for ten bacterial groups of five individuals were compared and showed that the proportion of several human gastrointestinal bacterial groups assessed with the HITChip correlated to various extent with the results of FISH (Fig. 4).

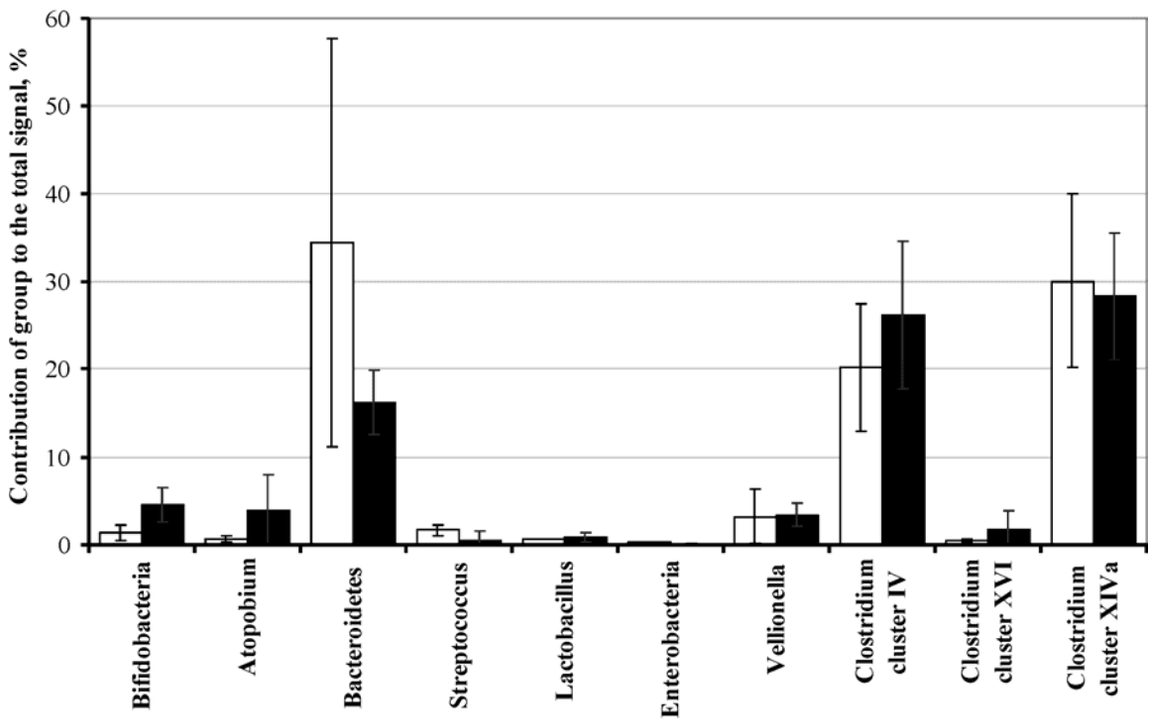


Figure 4 Correlation between average of the group specific HITChip hybridisation signals (empty columns) and FISH quantification (filled columns) based on the analysis of faecal microbiota of five subjects.

The results for members of two dominant human gastrointestinal groups, *Clostridium* cluster XIVa and *Clostridium* cluster IV showed an almost complete correlation. Similarly, the HITChip data of other quantitatively smaller groups, including *Veillonella* and *Lactobacillus*, correlated also almost completely with the FISH results. Nevertheless, the proportion of Bacteroidetes and Streptococci as assessed by the HITChip was significantly higher than assessed with FISH, while the opposite trend was observed for Actinobacteria that include *Bifidobacterium* and *Atopobium* group. As the HITChip analysis was performed on DNA and the FISH analysis on cells that were quantified by flow-cytometry, the number of SSU rRNA gene copies per cell is a crucial factor for DNA-based quantification of complex ecosystems.

Based on very limited number of gastrointestinal organisms for which the number of SSU rRNA genes has been determined or estimated (157, 254), most gastrointestinal organisms have an average of three copies of this gene. However, Bacteroidetes and Streptococci have an average of six SSU rRNA gene copies, which could explain their

overestimation by DNA-based quantification. This difference may become more pronounced as the number of genome copies per cell is related to the growth rate (546) – from various studies (*e.g.* (39)) it is know that different bacterial groups are not equally active along the gastrointestinal tract, and this fact will certainly effect the number copies of the SSU rRNA gene per cell, leading to the overrepresentation of the active bacterial groups.

Interestingly, similar relation between relative quantification signals for the analysed bacterial groups was obtained when, in another study, faecal samples were quantified using SSU RNA dot-blot hybridisation and FISH (547). Dot-blot hybridisation showed relatively higher proportion of *Bacteroides* and relatively lower proportion of *Atopobium* group, and this was explained by the possible different activity of these bacterial groups in faecal samples. However, the suspected higher activity *Bacteroides* spp. does not seem to be supported by more recent findings that showed that members of *Bacteroides* were more frequently identified within the subset of dead faecal bacteria (39). Furthermore, similar to rRNA dot-blot hybridisation, FISH technique is dependent on the activity of the cell, as this is directly linked to the number of ribosomes available for hybridisation, which implies that results of both techniques would be affected in similar manner, although at different extent.

The similar trend of the results observed using both DNA (as shown in our study) and RNA based methods (as shown before (547)) indicates that sample processing might have introduced biases. Cell fixation and nucleic acid extraction are performed under conditions that cannot be optimal for all members of a complex ecosystem such as the human gastrointestinal microbiota. Generally, the lower proportion of the Gram-negative Bacteroidetes detected by FISH (471) when compared to HITChip analysis, rRNA dot-blot hybridisation or cultivation (119, 333) suggests that a fraction of this, already injured or dead bacteria (39), might have been lost during sample preparation for FISH analysis. An alternative explanation is that the Gram-positive *Bifidobacterium* and *Atopobium* group (*Atopobium*, *Collinsella* and *Eggerthella*) are underrepresented by nucleic acids based techniques due to the possible failure of the nucleic acids extraction procedures of those robust organisms because of their strong cell wall. Finally, it has been shown that the application on FISH on an intestinal isolate has various effectiveness dependent on the growth phase of the microbe (103), and that in some cases the optimal hybridisation was obstructed. It cannot be ruled out that this or similar undefined parameters are introducing robustness in quantification of other, not yet characterised human gastrointestinal microbes. In fact, there are many factors that influence the quantification results obtained by different techniques, and presently no golden standard technique is available.

Despite all technical biases of the HITChip and FISH which can introduce some inconsistence of the results obtained using those two techniques; it was shown that the absolute signals obtained by both techniques correlate very well. As an example, bifidobacteria in 60 faecal samples were quantified using both the HITChip and FISH and the

results had positive linear correlation with Pearson's index of 0.72 (Fig. 5). Even better correlation was obtained when comparing the relative increase or decrease of this group in time per individual (Pearson's index of 0.93).

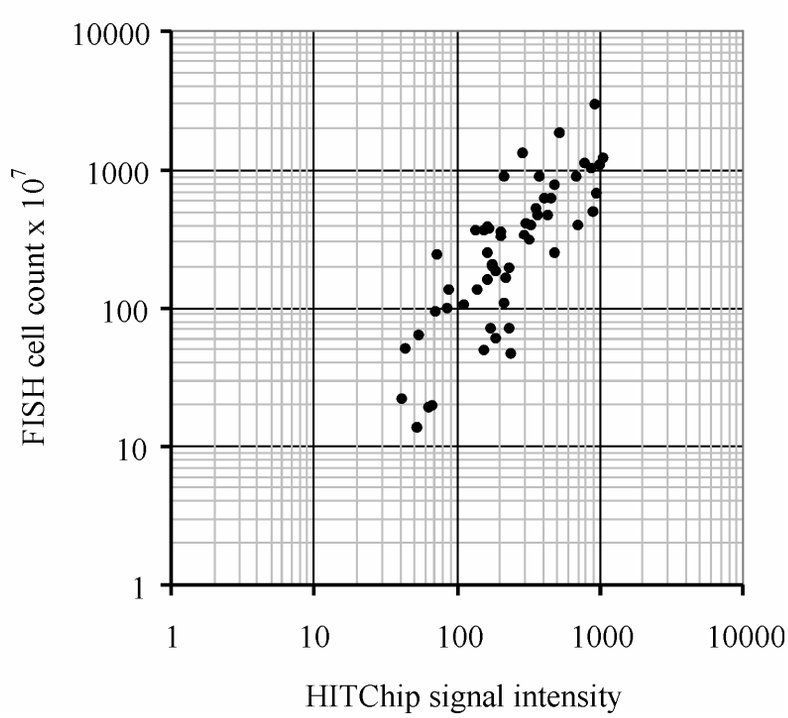


Figure 5 Correlation of results obtained by the *Bifidobacterium* specific FISH quantification and sum of the *Bifidobacterium* specific HITChip hybridisation signals for 56 faecal samples. Faecal samples obtained from 20 individuals in three time points were analysed, of which four samples that did not reach detection limit with the FISH are were not plotted on the graph.

These results indicated that the HITChip hybridisation signals can be used for the assessment of quantitative information on the human gastrointestinal microbiota, and that relative change of microbial groups or phylotypes can be the most reliably predicted, as confirmed in experiments with artificial mixtures and by comparing with the results of other SSU rRNA based technique.

In conclusion, we present here a simple, flexible and easy to use system for the systematic design of oligonucleotide probes that can be easily applied on any complex ecosystem of interest. The systematic probe design approach allows easy addition of probes that target newly discovered members of an ecosystem, which is very valuable for yet insufficiently described ecosystems such as the human gastrointestinal microbiota. Although the reliable prediction of the probe behaviour is hampered by the fact that not all relevant parameters that influence the probe-duplex hybridisation behaviour have yet been identified, microarrays based on the hypervariable regions of the SSU rRNA do enable rapid profiling of human gastrointestinal samples. With the possibility for relative quantification, microarray systems have a great potential to provide better insight into the human gastrointestinal microbiota, to identify the effect of time or dietary changes, and to compare groups of individuals that differ in the genotype or health status, age and geographic origin.

Acknowledgements

This study was carried out with financial support from the Commission of the European Communities, specifically the RTD programs Quality of Life and Management of Living Resources, QLK1-2000-00108, Microbe Diagnostics, coordinated by Michaël Blaut (Dife, Germany). The study does not necessarily reflect the Commission's views and in no way anticipates its future policy in this area. The authors would further like to thank Christophe Lay and Anu Surakka, for providing FISH data. Finally, the authors express their gratitude to Elena Biagi for the excellent technical assistance and Erwin G. Zoetendal for his interest, useful suggestions, and critical reading of this paper.

CHAPTER 5

Mirjana Rajilić-Stojanović, Annet Maathuis, Hans H. G. J. Heilig,

Koen Venema, Willem M. de Vos and Hauke Smidt

Manuscript submitted for publication

Evaluating the Microbial Diversity in an *In Vitro* Model of the Human Large Intestine by Phylogenetic Microarray Analysis

A high density phylogenetic microarray targeting small subunit ribosomal RNA (SSU rRNA) sequences of over 1,000 microbial phylotypes described for the human gastrointestinal tract, the HITChip, was used to assess the impact of faecal inoculum preparation and operation conditions of an in vitro model of the human large intestine (TIM-2) on the microbial diversity. This revealed that propagation of mixed faecal donations for the production of standardised inocula has only a limited effect on the microbiota composition, with slight changes observed mainly within the Firmicutes phylum. Adversely, significant shifts in several major groups of intestinal microbiota were observed after inoculation of the in vitro model. Hierarchical cluster analysis could show that microbiota profiles of samples taken throughout the inoculum preparation grouped with microbiota profiles assesed for faecal samples of healthy adults. In contrast, TIM-2 microbiota was distinct. While members of the Bacteroidetes phylum and some groups within the Bacilli order were increased in TIM-2 microbiota, a strong reduction in abundance of other microbial groups, including *Bifidobacterium* spp., *Streptococcus* spp., and *Clostridium* cluster IV and XIVa was observed. The changes detected with the HITChip could be confirmed, although at low resolution, by denaturing gradient gel electrophoresis of small subunit rRNA amplicons.

Introduction

The composition and functionality of the human intestinal microbiota is intimately linked to the nutrition and health of its host (23, 541). Hence, considerable interest has focused on the intestinal microbiota and more than 1,000 different microbial phylotypes have been reported to inhabit the human intestinal tract (**Chapter 3**). To better understand composition and functionality of this complex microbial ecosystem, *in vitro* model systems have been developed to bypass obvious restrictions with respect to accessibility of the intestine for frequent sampling (131). Currently available models of the large intestine range from simple batch fermentation systems to more or less sophisticated, well-controlled single- and multi-vessel continuous bioreactor systems (for a recent overview and discussion of the different models, see (131, 291)). In contrast to most of the models, the TIM-2 model of the large intestine is characterised by physiological water content and constant removal of metabolites via a dialysis system, allowing for recording of metabolite production kinetics and prevention of product inhibition of fermentation processes (131, 322). This model has been used to investigate several microbiota-mediated processes occurring in the colon, such as the fermentation of a variety of potentially prebiotic carbohydrates (487, 494, 495). Recently, stable isotope probing of microbial RNA has been proven instrumental for the identification of microbial populations responsible for the degradation of carbohydrates in the TIM-2 model, using glucose as a model substrate (130).

Intestinal models have, in general, been validated based on their representativeness of the metabolite production by comparison to the *in vivo* situation obtained from sudden death victims (289, 292, 322). Nevertheless, only limited information is available with respect to the diversity and stability of the microbiota present in such model systems. The reproducibility and compositional stability of *in vitro* model systems have been evaluated using various approaches, such as group-specific cultivation on selective media (294, 322, 493, 494), and cultivation-independent small subunit ribosomal RNA (SSU rRNA)-based approaches, including fluorescent *in situ* hybridisation (FISH) (82, 294, 383), dot blot hybridisation (213), quantitative real-time PCR (483), and microbiota fingerprinting by denaturing gradient gel electrophoresis (DGGE) of PCR-amplified SSU rRNA gene fragments (378, 483, 493, 494). Nevertheless, such approaches are in general not suitable to simultaneously provide data at sufficient phylogenetic and spatiotemporal resolution, as they either allow for analysis of only a few samples at high phylogenetic resolution by clone library analysis, or profiling of multiple samples for only a restricted number of organisms (qPCR, FISH) or without any direct phylogenetic information (DGGE). Such information, however, can now be generated by phylogenetic microarrays, which are in general based on SSU rRNA-targeted diagnostic oligonucleotide probes that allow for the detection of microorganisms at different levels of taxonomic resolution (182, 283, 442, 499, 534). Phylogenetic microarrays have now been

developed for a wide range of complex microbial ecosystems, including the human intestine and the oral cavity (359, 442). The recently developed Human Intestinal Tract Chip (HITChip) offers the possibility to follow the microbiota composition in intestinal model systems at unprecedented phylogenetic resolution to elucidate shifts in community composition as a result of operating procedures of intestinal models (**Chapter 4**). Its application here with the microbiota in the TIM-2 model reveals specific community shifts that may be avoided when specific precautions are taken.

Materials and methods

The TIM-2 model of the human colon: inoculum preparation, operating conditions and sampling

The TNO *in vitro* model of the human large intestine, TIM-2, represents the conditions found in the proximal colon. Standard operation procedures for inoculum preparation and model operation have previously been established and validated (322, 487). The inoculum used in this study was derived from faecal donations of 7 healthy volunteers. For the purpose of this study, two slightly different procedures of inoculum preparation were used. For procedure IN1, 80 g of fresh mixed faecal material was mixed in a 2l-fermentor with 670 ml of a complex medium representing terminal ileal chyme (modified from (169) as described by (487)) and incubated at pH 5.8 and 37°C. Additional medium was added after an initial 4h batch incubation during 40h to reach a final volume of 2000 ml. For procedure IN2, 80 g of faecal material was mixed with 670 ml of dialysis liquid (487). Medium as described above, but concentrated 1.5-fold and amended with 1.7% of pig gastric mucin (Type II, Sigma-Aldrich), was added during 44h of incubation, yielding the same final volume as described for IN1. Aliquots of 1000 ml were removed from the reactor, mixed with 10% w/v glycerol, divided in aliquots inside an anaerobic cabinet, frozen in liquid N_2 and stored at -80^{o}C. For both procedures, samples were taken at 0, 4, 20, 24, 28 and 44 h of incubation, and stored at -80°C until further analysis.

Duplicate modules of the TIM-2 model were inoculated with material obtained by procedure IN1 after thawing for 1 h at 37°C, and operated at pH 6.4 as previously described (487). Microbiota was allowed to adapt to model conditions during an initial incubation for 16 h, after which the actual run was initiated. Luminal and dialysate samples (each ~ 2 ml) were taken at 0, 24, 48, and 72 h, and stored at -80°C until further analysis. The 0 h time point corresponds to initiation of the run after initial adaptation.

Concentrations of the short chain fatty acids acetate, propionate, and n-butyrate, and branched chain fatty acids iso-butyrate and iso-valerate were determined by gas chromatography as described previously (487).

For evaluation of the microbial diversity during *in vitro* preparation, the faecal samples of healthy human volunteers which were analyised in a separate study, were used (**Chapter 7**).

DNA extraction, PCR amplification and denaturing gradient gel electrophoresis (DGGE) analysis

Samples were thawed at room temperature, and DNA was extracted from 200 µl using the FastDNA SPIN Kit for Soil (Qbiogene, Carlsbad, CA, USA) and a FastPrep disruptor (FP120, Savant Instruments, Farmingdale, NY, USA). DNA yield was quantified using a NanoDrop ND-1000 spectrophotometer (NanoDrop® Technologies, Wilmington, DE, USA).

DGGE of PCR-amplified 16S rRNA gene fragments was used to initially compare overall bacterial diversity in samples taken throughout inoculum preparation and model operation. To this end, primers GC-968-f (5'- CGC CCG GGG CGC GCC CCG GGC GGG GCG GGG GCA CGG GGG GAA CGC GAA CCT TAC-3') and 1401-r (5'- GCG TGT GTA CAA GAC CC-3') were used to amplify the V6-V8 region of the 16S rRNA gene as described previously (352). Amplicons were then separated by DGGE (342), using a Dcode™ system (Bio-Rad, Veenendaal, The Netherlands). Electrophoresis was done as described previously (194). Briefly, samples were loaded on 8% polyacrylamide gels with a gradient of 30-60% of denaturant (100% defined as 40% formamide and 7M urea), pre-run for 5min at 200V, and subsequently electrophoresed at 85V for 16h at 60°C. Gels were developed by silver staining according to Sanguinetti et al. (414), scanned at 400 dpi, and further analysed by Bionumerics 4.5 (Applied Maths BVBA, Sint-Martens-Latem, Belgium). Pearson product-moment correlation based on densitometric curves was used to assess the similarity of DGGE-patterns obtained for different samples (189, 536).

The HITChip analysis

PCR

The SSU rRNA gene amplified from faecal DNA using the *T7prom*-Bact-27-for (5'-TGA ATT GTA ATA C GA CTC ACT ATA GGG GTT TGA TCC TGG CTC AG–3') and Uni-1492-rev (5'-CGG CTA CCT TGT TAC GAC-3'), which ensured the introduction of a T7 promoter sequence at the 5' terminus of the rRNA gene amplicon. PCR reactions were carried out in a final volume of 50 µl, and 10 ng of DNA samples were used as template. Samples were initially denatured at 94˚C for 2 min followed by 35 cycles of 94˚C (30 sec), 52˚C (40 sec), 72˚C (90 sec) and a final extension at 72˚C for 7 min. The PCR products were purified by using the DNA Clean and Concentrator kit (Zymo Research, Orange, USA), according to the manufacturer's instruction. The final DNA concentration was determined by using a NanoDrop spectrophotometer as described above.

RNA production and labelling

In vitro transcription of the T7-promoter carrying SSU rRNA gene was performed according to the manufacturer's protocol using the Riboprobe System (Promega, La Jolla, USA), 500 ng of the T7-16S rRNA gene amplicon, and, besides rATP, rGTP, rCTP, a 1:1 mix of rUTP and aminoallyl-rUTP (Ambion Inc., Austin, Tx, USA). The transcription reaction was performed at room temperature for 2 h., the template DNA was digested applying the Qiagen RNAse-free DNAse kit (Qiagen, Hilden, Germany), and RNA was purified using the RNeasy Mini-Elute Kit (Qiagen, Hilden, Germany). RNA yield was measured as described above.

Amino-allyl-modified nucleotides were coupled with CyDye using the Post-Labeling Reactive Dye (Amersham Biosciences, Little Chalfont, UK), which was previously dissolved in 84 µl DMSO. Labelling reaction was performed in a 25 mM sodium bicarbonate buffer (pH 8.7) by adding 20 µl of dissolved CyDye to 2 µg of purified RNA in a final volume of 40 µl. Samples were incubated in the dark at room temperature for 90 min. The reaction was stopped by adding 15 µl of 4M hydroxyl-amine and incubating in the dark for 15 min. RNAse-free water was added to 100 µl and labelled RNA was purified and quantified as described above.

Microarray production, hybridisation and data extraction

The HITChip microarrays were produced by Agilent Technologies (Agilent Technologies, Palo Alto, CA). The oligonucleotide probes were extended at the 3' end (at the array support side) by 10-nt long T spacers and were printed on the array using *in situ* surface-attached oligonucleotide probe synthesis (54). The arrays used in this study were of the 2×11K format, with two arrays per glass slide. Each array was hybridised with two samples, labelled with Cy3 and Cy5, respectively. Combined Cy3- and Cy5-labelled target mixtures were fragmented by adding 1 µl of Ambion 10× fragmentation reagent (Ambion Inc., Austin, Tx, USA), and incubation at 70˚C for 20 min, according to the manufacturer's instruction. Fragmentation was stopped by adding 1 µl of Ambion stop solution. Hybridization mix was prepared by adding to the RNA mixture 31.5 µl of 20× SSC (412), 6.3 µl of 10% SDS (412), 25 µl of Agilent Control Target mix and RNAse-free water to a total volume of 210 µl. Hybridization was carried out at 62.5^{o}C in a rotation oven (Agilent) for 16 hours. Slides were washed at room temperature in 2× SSC, 0.3% SDS for 10 min and at 38˚C in 0.1× SSC, 0.3% SDS for 10 min. SDS was completely removed by washing the slides in 0.06× SSPE (412) for 5 min, followed by a quick dry with compressed nitrogen.

Data were extracted from microarray images using the Agilent Feature Extraction software, version 7.5 (www.agilent.com). Data normalisation was performed using set of R based scripts (http://www.r-project.org/), while the further microarray analysis was performed in a custom designed relational database which runs under the MySQL database management system (http://www.mysql.com/) using a series of custom made R scripts, as previously described (**Chapter 4**). Similarity of the total microbiota composition based on the

HITChip profiles was assessed by calculating Pearson's product moment correlation (Pearson's correlation) that reflects the degree of linear relationship between analysed data sets. Similarity of the DGGE profiles was assessed as previously described (162). Ward's minimum variance method was used for the generation of hierarchical clustering of probe profiles by calculating a distance matrix between the samples based on the squared difference between each pair of profiles (squared Euclidian distance – E^2) (74).

Statistical analysis

To assess correlations of microbial groups detected by HITChip analysis with sample characteristics, principal component analysis (PCA) was used as implemented in Canoco for Windows 4.5 (271). Average signal intensities for 131 bacterial groups defined on the bases of the SSU rRNA gene sequence similarity were used as species data, PCA was performed focusing on inter-species correlation and diagrams were plotted by using the CanoDraw for Windows utility. For the evaluation of statistical significance of observed correlations, Student's t-test was used. Calculated p-values were two tailed, and correlations were considered significant at $p < 0.05$. The Monte Carlo Permutation Procedure (212) as implemented in the Canoco package was used to assess statistical significance of the variation in large datasets in relation to sample origin, i.e. human faeces, inoculum preparation and TIM-2 model operation.

Results and Discussion

The composition and stability of intestinal microbiota maintained in the TIM-2 in vitro model of the human colon was evaluated by bacterial SSU rRNA gene-targeted DGGE and phylogenetic microarray analysis using the HITChip. To this end a number of steps during the preparation and preservation of standardised inocula, as well as the actual operation of the model system were investigated. In addition to the monitoring of the inoculum production process, we also compared two methods, which differed i) in the presence or absence of mucin, and ii) with respect to the conditions of medium addition to the cultivation vessel. The intestinal mucus layer provides an ecological niche, and mucus can be seen as a rich carbon and energy source for intestinal microbiota (109). Remarkably, mucin beads added to a two-stage intestinal model system were not only shown to be readily colonised by bacterial biofilms, but mucin was also largely degraded by luminal microbiota, reinforcing the notion that mucin can serve as substrate for a variety of intestinal microorganisms (294).

Production of standardised faecal inocula

To provide a standardised inoculum for a large number of experiments, allowing for high reproducibility and comparability between runs, a protocol has previously been developed, in

which faecal donations from ten healthy volunteers are mixed, propagated and preserved by freezing in the presence of glycerol as a cryo-protectant (487, 493). The extent by which this procedure affected the microbial composition was readdressed using the HITChip, although it had previously been assessed by DGGE and selective plating, because of the limitations of both techniques. DGGE analysis does not directly provide any information with respect to the phylogenetic affiliation of populations represented by bands in diversity profiles. Moreover, plating is laborious and seriously impacted by our inability to cultivate the majority of intestinal microorganisms with currently available media (538).

Differences in feeding regime between the two procedures, i.e. initial batch addition of 30% of the medium in the case of IN1 and linear addition of a more concentrated medium for IN2, were reflected by different profiles of SCFA production during the 44 h of inoculum propagation. Acetate, propionate and n-butyrate were more readily produced in the initial phase of IN1, in line with the initial batch addition, while final concentrations reached in IN1 and IN2 were comparable (80 mM acetate, 30 mM propionate, 25 mM n-butyrate). Similar kinetics was observed for the production of iso-butyrate and iso-valerate, although these branched-chain fatty acids were produced at 10-fold lower concentrations. DGGE profiling, as well as HITChip analysis based on the signal intensity of approximately 3,700 distinct oligonucleotide probes, revealed that none of procedures IN1 and IN2 dramatically affected community composition (Fig. 1). The microbiota present after 44 h was 85.0% (IN1) and 79.2% (IN2) similar to the initial mixture of the faecal donations based on Pearson product moment correlations of DGGE profiles.

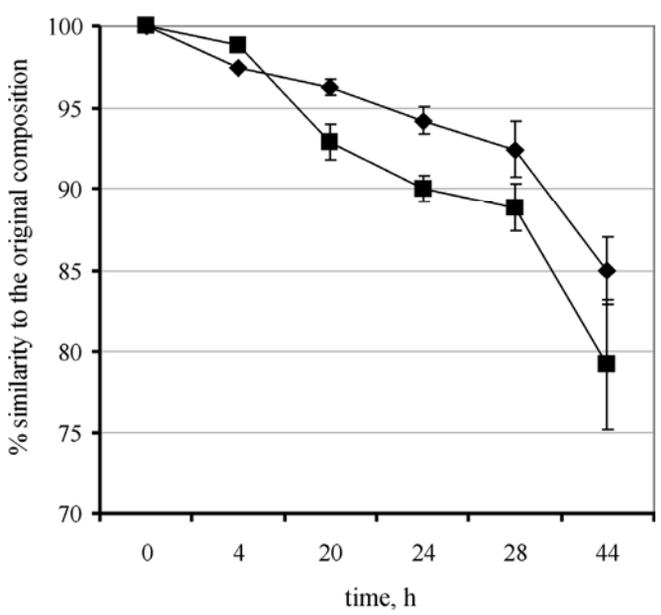


Figure 1 Similarity indices based on Pearson product-moment correlation coefficients for DGGE fingerprints generated from the initial inoculum compared to those taken at 0, 4, 20, 24, 28, and 44 h during the propagation of faecal inocula for *in vitro* intestinal models by two different procedures, IN1 (◆) and IN2 (■). Average data from duplicate DGGE gels are shown with their respective standard error.

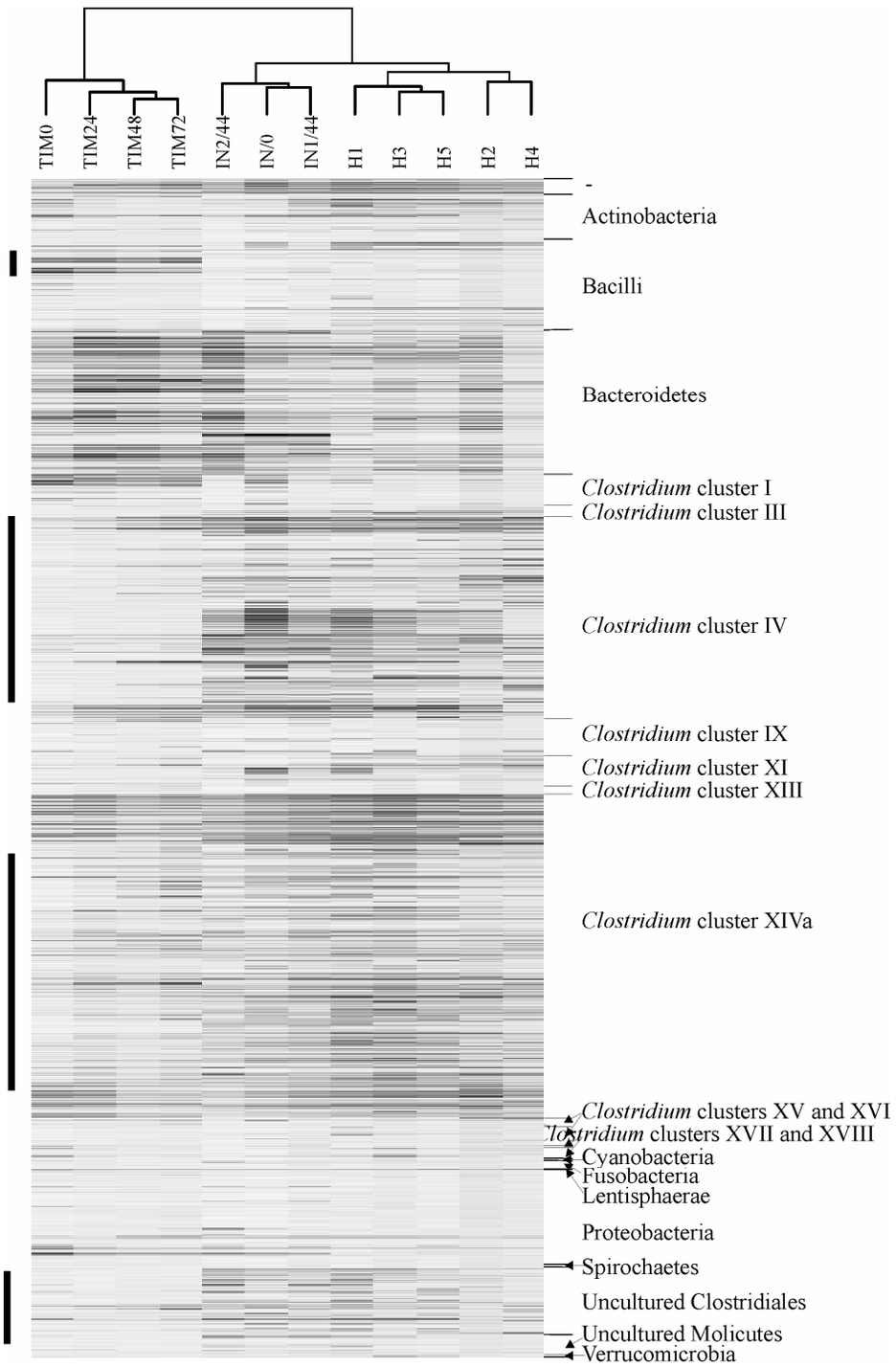


Figure 2 Hierarchical cluster analysis of phylogenetic HITChip fingerprints generated from faecal mix sample (IN/0), and after sample taken after 44h of inoculum preparation using both procedures (IN1/44, and IN2/44), as well as samples taken at 0, 24, 48 and 72 h of TIM-2 operation after an initial 16 h stabilisation (TIM/0-TIM/72), and faecal samples of five randomly selected healthy adult volunteers (H1-H5). Highest phylogenetic ranks of probe specificity are given at the right side of the figure. Vertical bars at the left indicate microbial groups most affected by *in vitro* model operation.

The similarity between the initial microbiota composition and the outcome of the inoclum preparation using both procedures, assessed by the HITChip analysis were comparable to the results obtained by the analysis of the DGGE gels. The calculated similarity for IN1 was 82.5 (85.2%, based on DGGE), whereas for IN2 it was 74.9% (79.2% based on DGGE). The same trend of the results, and slightly lower values assessed by the HITChip showed the good agreement between two profiling techniques and indicated better sensitivity of the HITChip profiles. Moreover, during propagation of both inocula the bacterial community composition seemed to be similarly affected as the change in the relative intensity of identical DGGE bands was observed for IN1 and IN2 (data not shown). Analysis of the phylogenetic fingerprints obtained using the HITChip, provided a more detailed insight into subtle differences between faecal mix, and inocula IN1 and IN2 (Fig. 2).

It appeared that both procedures affected individual populations within the Firmicutes phylum in a similar manner, while procedure IN2 had an additional pronounced impact on members of the Bacteroidetes (Fig. 2). The ability to extract these data directly from the microbiota profiles stresses the advantage of phylogenetic microarray-based fingerprints in comparison to those obtained using other, well established fingerprinting techniques, such as DGGE, and enforces previously established evidence for superiority of phylogenetic microarrays (346).

DGGE analysis of samples obtained before and after freezing of the inocula indicated that inocula produced by procedure IN1 were less prone to freeze-thaw induced changes in detectable community composition. Similarity values amounted to 96.5 and 67.1% for IN1 and IN2, respectively. Consequently, inoculum prepared by procedure IN1 was chosen for subsequent experiments with the TIM-2 model as described below.

Impact of i*n vitro* model operation on microbiota composition

IN1-derived material was used to inoculate duplicate modules of the TIM-2 model of the proximal colon, and samples taken at regular 24 h intervals were analysed for production of metabolites, as well as composition and stability of the microbiota. Acetate, propionate and n-butyrate were the main metabolites observed, at approximate molar ratios of 3:1:1, respectively, which is similar to values found during inoculum propagation (Fig. 3A). Similarity indices, calculated on the bases of DGGE and the HITChip profiles, between the microbiota of the starting faecal mix and number of samples taken during the course of *in vitro* model operation at pH 6.4 showed similar trends (Fig. 3B).

A notable shift of the microbiota composition during the initial 16 h of stabilisation of the inoculum was observed, as similarity as compared to the inoculum dropped to values around 40%, which were maintained throughout the subsequent model operation for 72 h. Such dramatic changes can have their origin in any of the steps preceding the actual *in vitro* model experiment, i.e. freezing and thawing of the standardised inoculum, traces of oxygen

entering the system during the actual inoculation of the model as well as the adaptation period of 16 h. In a previous study, where stable isotope probing of bacterial RNA was used to identify major players in colonic carbohydrate fermentation using glucose as a model substrate for methods development, T-RFLP fingerprinting and SSU rRNA clone library analysis revealed that the microbiota that was enriched in the TIM-2 model, resembled more the small rather than the large intestinal situation (130).

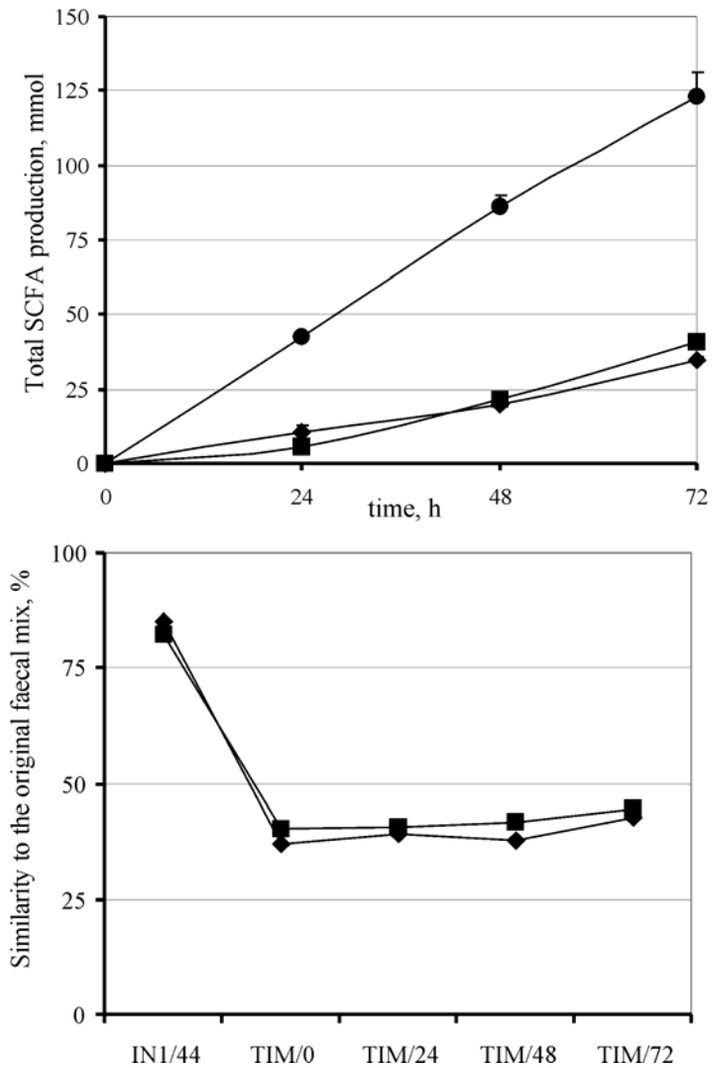


Figure 3 A Cumulative production of the short chain fatty acids acetate (●), propionate (■) and n-butyrate (◆) during TIM-2 operation at pH 6.4 and 37°C. Average data from duplicate TIM-2 runs are shown with their respective standard error. B Similarity of the TIM-2 model microbiota compared to the initial faecal mixture used for inoculum preparation, based on DGGE (◆) and HITChip (■) analyses. IN 44, inoculum prepared according to procedure IN1, after 44 h of cultivation; TIM 0, 24, 48, 72, samples taken from the TIM-2 model in 24 h intervals, starting after an initial adaptation for 16h.

Hierarchical clustering of HITChip profiles obtained for samples taken from TIM-2 with those taken throughout the preparation of inocula, as well as randomly selected profiles generated from faecal samples of healthy adults, revealed that the microbiota selected by TIM-2 operation conditions considerably deviated from normal human colonic microbiota (Fig. 2). Interestingly, and in contrast to TIM-2 microbiota, samples taken during inoculum preparation still clustered with profiles of human faecal microbiota. This finding reinforced our previous observation that cluster analysis allows for analysis beyond inter-individual variation, discerning consensus profiles of healthy adult faecal microbiota (**Chapter 4**). To

further identify those microbial populations responsible for the observed deviation of TIM-2 from the *in vivo* situation, Student's T-tests were performed. This revealed that a total of 22 out of the 131 bacterial groups were significantly different between TIM-2 samples and human faecal profiles (Table 1). Most pronounced changes could be attributed to an increase in the abundance of several groups within the Bacteroidetes (*Prevotella tannerae, Bacteroides intestinalis*) and the Bacilli (*Enterococcus*, *Granulicatella*). In addition, numerous other groups within these two phyla were strongly increased in TIM-2 samples, albeit not significant (data not shown). This was also confirmed by principal component analysis, showing that a number of groups within the Bacteroidetes were positively correlated with TIM-2 operation (Fig. 4).

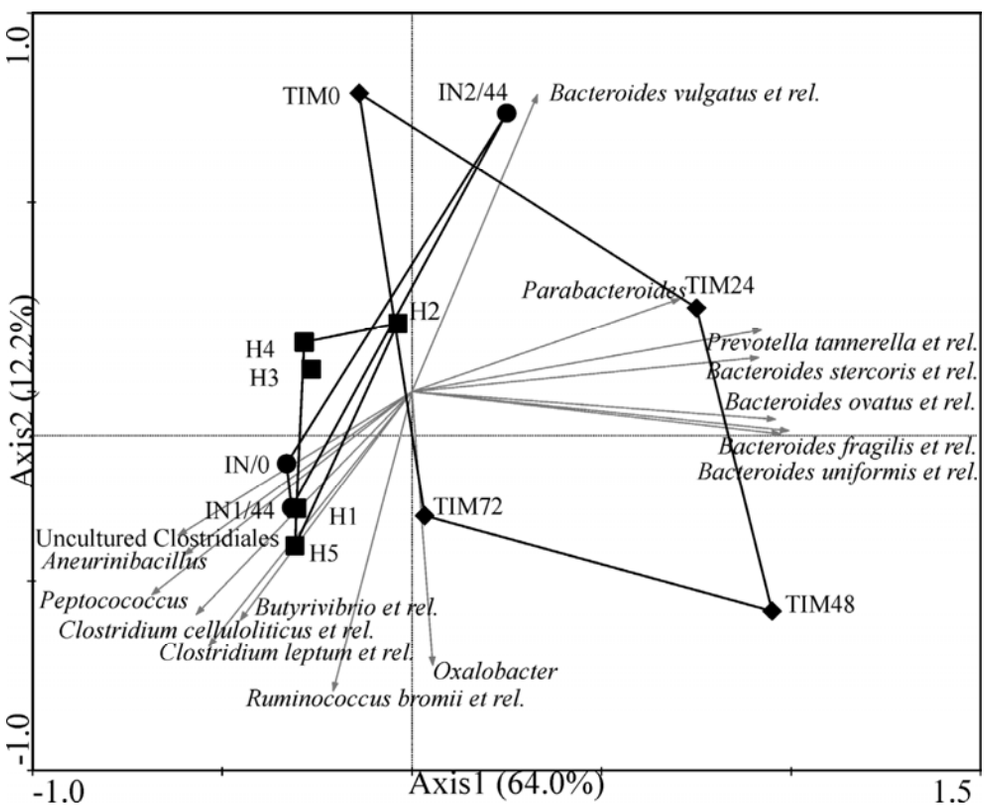


Figure 4 Principal component analysis of the microbiota composition as measured by the average hybridisation signals for 131 phylogenetic groups in samples taken during inoculum preparation (●; IN/0, IN1/44, IN2/44), samples taken during TIM-2 model operation (◆; TIM0-TIM72), and faecal samples of 5 randomly selected healthy adult volunteers (■; H1-H4). Microbial groups that contributed at least 40% to the first two principle components used as explanatory axes in the plot are shown as vectors. For the details about coding see Fig. 2. MCPP revealed that the environmental variables Inoculum, Human and TIM were correlated to the variation in microbiota composition with a p-value of 0.062.

Groups that significantly decreased during model operation included streptococci, bifidobacteria, several populations within the *Clostridium* clusters IV (*Clostridium leptum*

group) and XIVa (*Clostridium coccoides* group), as well as members of *Peptococcus*, *Clostridium stercorarium* and *Anaerovorax* (Table 1, Figs. 2, 4). Some of these groups that were negatively affected in the present study represent fastidious anaerobes, which in the past largely eluded cultivation. This holds specifically for members of the Firmicutes, such as the *Clostridium* cluster IV and XIVa groups (258, 461, 467). Interestingly, the strong reduction in *Streptococcus* spp. was previously observed for a three-stage intestinal model, while this was not the case for *Bifidobacterium* spp. and *Clostridium* clusters IV and XIV (82), indicating similarities, but also differences of the effect of the standard operating procedures for both model systems on the microbiota composition.

Table 1 Microbial groups that were significantly affected by TIM-2 operation as compared to faecal microbiota of healthy volunteers, judged based on average HITChip signal intensities for 131 phylogenetic groups based on their SSU rRNA sequence similarity. Average signal intensities and p values are provided.

Higher taxonomic group		Group	P-value	Average signal intensity	
				TIM	faeces
Bacteroidetes		*Prevotella tannerae*-like	0.030	237	37
		Bacteroides intestinalis-like	0.050	309	74
Actinobacteria		*Bifidobacterium*	0.041	31	180
Firmicutes	Bacilli	*Enterococcus*	<0.001	304	24
		Granulicatella	0.002	121	16
		Streptococcus bovis-like	0.002	28	189
		Streptococcus mitis-like	0.002	30	95
		Streptococcus intermedius-like	0.006	32	74
		Aneurinibacillus	0.027	21	74
	Clostridium cluster IV	*Ruminococcus callidus*-like	0.001	15	57
		Oscillospira guillermondii-like	0.007	56	177
		Clostridium orbiscindens-like	0.009	31	119
	Clostridium cluster XI	*Anaerovorax odorimutans*-like	0.005	28	57
	Clostridium cluster XIVa	*Anaerostipes caccae*-like	0.001	17	228
		Eubacterium hallii-like	0.001	63	368
		Ruminococcus gnavus-like	0.010	45	158
		Coprococcus eutactus-like	0.021	37	147
		Clostridum symbiosum-like	0.032	63	97
		Dorea	0.034	54	174
		Lachnospira pectinoschiza-like	0.036	19	69
		Ruminococcus obeum-like	0.042	147	456
		Outgrouping *C.*cluster XIVa	0.049	18	68

In conclusion, two SSU rRNA gene based techniques were employed for identifying shifts in the faecal microbiota composition during the preparation of standardised inoculum and the operation of TIM-2 *in vitro* model of the human large intestine. The results obtained using DGGE and the HITChip were consistent, while more detail was provided with the HITChip, phylogenetic microarray analyses. The phylogenetic insight into the microbiota composition showed that, despite small changes during the preparation procedure, the microbiota of the standardised faecal inoculum resembles that of the fresh faecal samples. However, the initial adaptation of the faecal inoculum in the TIM-2 *in vitro* model has induced considerable shifts in the microbiota composition, which have affected the relative abundance of several microbial groups, and resulted in the noticeably different microbial community in the *in vitro* model when compared when compared to that of healthy adults ($p=0.062$). The results reported in this study provide new leads for adjustments in standard operating procedures of TIM-2 *in vitro* model to better simulate the diversity that is normally present *in vivo*, in healthy adults. Such studies are currently underway, focusing on the effect of individual steps during inoculum preparation and model operation, including initial propagation of a mixed faecal sample, freezing and thawing, inoculation of the model, and stabilisation of the inoculum under the *in vitro* model conditions.

CHAPTER 5

Mirjana Rajilić-Stojanović, Hans G. H. J. Heilig, Erwin G. Zoetendal and Willem M. de Vos

Manuscript in preparation

Dynamics of the Adult Gastrointestinal Microbiota

The microbiota in the human gastrointestinal tract represents a complex community that was shown to be host-specific and stable. Since the microbiota's dynamics was hitherto analysed only in short term studies, we have readdressed the question of the microbiota's stability during longer period, up to nine years, using a newly developed phylogenetic microarray, the method that combines the power of fingerprinting, phylogenetic and quantitative population analysis. Our results confirmed that each individual harbours an unique consortium of microbes, despite noted changes in time. Within a time span changing from six months to four years the similarity between the microbiota profiles decreased from 79.5% to 63.2%. Monozygotic twins shared 55.6% of the microbiota, which made their microbiota notably more similar than that of unrelated individuals (41.7%, p=0.067), but close to the level of temporal variation in the 4 years time span (p=0.17). Subpopulations of the microbiota were differently affected by the temporal variation, and notably members of Actinobacteria and Bacteroidetes exhibited significantly higher similarity when compared to the variation of the total microbiota. The similarity of these groups had the same pattern of change as that of the total microbiota in respect to temporal variation and genetic relatedness. Based on these results, and insights provided by nine-year long follow up of the microbiota's dynamics of a single individual, we hypothesise that the human gastrointestinal microbiota fluctuates around a core community of permanent colonisers, the composition of which is affected by the host genotype.

Introduction

The human gastrointestinal microbiota is a complex ecosystem that consists of microbes co-evolved with its host and highly adapted to this specific ecological niche (23). Only few microbial groups can survive in strictly anaerobic conditions of the gastrointestinal tract that is fuelled by a combination of food-borne nutrients and host-derived substrates. Therefore, on a high taxonomic level the microbiota is relatively simple as members of only four phyla, namely Actinobacteria, Bacteroidetes, Firmicutes and Verrucomicrobia, are regularly detected as abundant colonisers in gastrointestinal samples. Nevertheless, the intestinal microbiota shows a high diversity and more than thousand species- level types (phylotypes), which also belong to the other less dominant phyla, have been reported to inhabit this incompletely described ecosystem (**Chapter 3**). The global analysis of the microbiota revealed that it is not specific only for the gastrointestinal tract, but also for each individual (129, 275, 537).

The initial microbial innoculum derives from the neonates' mother, environment and diet, and it develops into a complex ecosystem throughout dramatic changes primarily in the first two years of life (146, 325). In contrast, the microbiota of adults exhibits a high level of stability (537). Nevertheless, the fact that the microbiota of the healthy middle-aged adults differs significantly from that of healthy elderly (339) indirectly suggests that the changes of the microbiota occur throughout the entire life span of humans. Since the established differences between age groups are of quantitative nature, it is likely that the individual-specific signature of the microbiota composition is preserved during the entire life, while the relative ratio of microbial groups is changing. This is supported by the observation that abundance of the dominant bacterial groups shows considerable fluctuations within a period of less than a year (306), while during the same period fingerprinting techniques showed a stable profile (489, 537). However, due to the large inter-individual variations the direct coupling of these experimental data is impossible. Therefore, the recently risen question of the gastrointestinal microbiota stability (112) was readdressed using a newly developed microarray, the HITChip, the method that combines the power of fingerprinting, phylogenetic and quantitative population analysis (**Chapter 4**). In addition to a short time interval of half a year, we have assessed changes of the microbiota after period of four years. More insight into the behaviour of the microbiota in time was assessed by analysing microbiota of an individual during a period of nine years. Finally, the obtained results were compared to similarity of the microbial composition of intestinal contents of five monozygotic twin pairs, since these can provide a measure of genotype influence on the gastrointestinal microbiota (536). The obtained results allowed us to hypothesise that the human gastrointestinal microbiota fluctuates around a core community of permanent colonisers, which composition is affected by the host genotype.

Materials and Methods

Subjects

Faecal samples were collected from five unrelated individuals (three male and two female), aged between 28 and 38, and all living in the Netherlands. If not processed immediately faecal samples were stored at -80^{o}C. In addition to newly collected samples we have used faecal samples collected as part of previously published work of Zoetendal and colleagues (536, 537) of healthy individuals that were followed in time while their major life style changes were retrospectively recorded.

Faecal DNA extraction

Total DNA was extracted from faecal material using a modified protocol of the QiaAmp DNA Mini Stool Kit (Qiagen, Hilden, Germany) (539). DNA yield was quantified using NanoDrop ND-1000 spectrophotometer (NanoDrop® Technologies, Wilmington, DE). DNA concentration was adjusted to 10ng/µl and was used as a template for PCR amplification.

Denaturant gradient gel electrophoresis (DGGE)

SSU rRNA gene-fragments suitable for the separation on the DGGE were generated by PCR according to published protocols (352). DGGE was performed using a 30-60% gradient gel under previously described conditions (194). DGGE fingerprints were normalised and clustered using the Bionumerics software (http://www.applied-maths.com).

The HITChip analysis

PCR

The SSU rRNA gene amplified from faecal DNA using the *T7prom*-Bact-27-for (5'-TGA ATT GTA ATA C GA CTC ACT ATA GGG GTT TGA TCC TGG CTC AG–3') and Uni-1492-rev (5'-CGG CTA CCT TGT TAC GAC-3'), which ensured the introduction of a T7 promoter sequence at the 5' terminus of the rRNA gene amplicon. The SSU rRNA gene of species that belong to *Bifidobacterium* genus was amplified using the primers Im26 equipped with a T7-promotor (5'-TGA ATT GTA ATA CGA CTC ACT ATA GGG GAT TCT GGC TCA GGA TGA ACG-3') and Im3 (5'-CGG GTG CTI CCC ACT TTC ATG-3') using the described reaction-conditions (243). PCR reactions were carried out in a final volume of 50 µl, and 10 ng of DNA samples were used as template. Samples were initially denatured at 94˚C for 2 min followed by 35 cycles of 94˚C (30 sec), 52˚C (40 sec) when universal bacterial primers were used or 57˚C (40 sec) when *Bifidobacterium* specific primers were used, 72˚C (90 sec) and a final extension at 72˚C for 7 min. The PCR products were purified by using the DNA Clean and Concentrator kit (Zymo Research, Orange, USA), according to the

manufacturer's instruction. The final DNA concentration was determined by using a NanoDrop spectrophotometer as described above.

RNA production and labelling

In vitro transcription of the T7-promoter carrying SSU rRNA gene was performed according to the manufacturer's protocol using the Riboprobe System (Promega, La Jolla, USA), 500 ng of the T7-16S rRNA gene amplicon, and, besides rATP, rGTP, rCTP, a 1:1 mix of rUTP and aminoallyl-rUTP (Ambion Inc., Austin, Tx, USA). The transcription reaction was performed at room temperature for 2 h., the template DNA was digested applying the Qiagen RNAse-free DNAse kit (Qiagen, Hilden, Germany), and RNA was purified using the RNeasy Mini-Elute Kit (Qiagen, Hilden, Germany). RNA yield was measured as described above.

Amino-allyl-modified nucleotides were coupled with CyDye using the Post-Labeling Reactive Dye (Amersham Biosciences, Little Chalfont, UK), which was previously dissolved in 84 µl DMSO. Labelling reaction was performed in a 25 mM sodium bicarbonate buffer (pH 8.7) by adding 20 µl of dissolved CyDye to 2 µg of purified RNA in a final volume of 40 µl. Samples were incubated in the dark at room temperature for 90 min. The reaction was stopped by adding 15 µl of 4M hydroxyl-amine and incubating in the dark for 15 min. RNAse-free water was added to 100 µl and labelled RNA was purified and quantified as described above.

Microarray production, hybridisation and data extraction

The HITChip microarrays were produced by Agilent Technologies (Agilent Technologies, Palo Alto, CA). The oligonucleotide probes were extended at the 3' end (at the array support side) by 10-nt long T spacers and were printed on the array using *in situ* surface-attached oligonucleotide probe synthesis (54). The arrays used in this study were of the 2×11K format, with two arrays per glass slide. Each array was hybridised with two samples, labelled with Cy3 and Cy5, respectively. Combined Cy3- and Cy5-labelled target mixtures were fragmented by adding 1 µl of Ambion 10× fragmentation reagent (Ambion Inc., Austin, Tx, USA), and incubation at 70˚C for 20 min, according to the manufacturer's instruction. Fragmentation was stopped by adding 1 µl of Ambion stop solution. Hybridization mix was prepared by adding to the RNA mixture 31.5 µl of 20× SSC (412), 6.3 µl of 10% SDS (412), 25 µl of Agilent Control Target mix and RNAse-free water to a total volume of 210 µl. Hybridization was carried out at 62.5^{o}C in a rotation oven (Agilent) for 16 hours. Slides were washed at room temperature in 2× SSC, 0.3% SDS for 10 min and at 38˚C in 0.1× SSC, 0.3% SDS for 10 min. SDS was completely removed by washing the slides in 0.06× SSPE (412) for 5 min, followed by a quick dry with compressed nitrogen.

Data were extracted from microarray images using the Agilent Feature Extraction software, version 7.5 (www.agilent.com). Data normalisation was performed using set of R based scripts (http://www.r-project.org/), while the further microarray analysis was performed

in a custom designed relational database which runs under the MySQL database management system (http://www.mysql.com/) using a series of custom made R scripts, as previously described (**Chapter 4**).

Similarity calculation

Similarity of the total microbiota composition based on the HITChip profiles was assessed by calculating Pearson's product moment correlation (Pearson's correlation) that reflects the degree of linear relationship between analysed data sets. Similarity of the DGGE profiles was assessed as previously described (162). P-values were calculated using Student's t-test. All presented values are one tailed, and if lower than 0.05 were considered significant and were marked with an asterix in the figures, while p values lower than 0.01 were considered as highly significant and were marked with a double asterix.

In addition to the assessment of the total microbiota similarity, Pearson's correlation was calculated for sets of probes that correspond to different phylogenetic groups. Strong positive association (ρ=0.999) was obtained by calculating of Pearson's correlation for total HITChip profiles obtained on *in vitro* transcribed *Bifidobacterium* specific PCR with Pearson's correlation calculated for *Bifidobacterium* specific probes for five pairs of samples (Fig. S1). This correlation could not be assessed for other relevant gastrointestinal groups since the group specific primers do not allow amplification of the SSU rRNA segment which is used for the HITChip analysis and is between 64 to 1,047 *E. coli* position (156, 305, 432). Therefore, the similarity of phylogenetic groups was calculated exclusively by Pearson's correlation of the hybridisation signals of the HITChip probes which are specific for the analysed phylogenetic group.

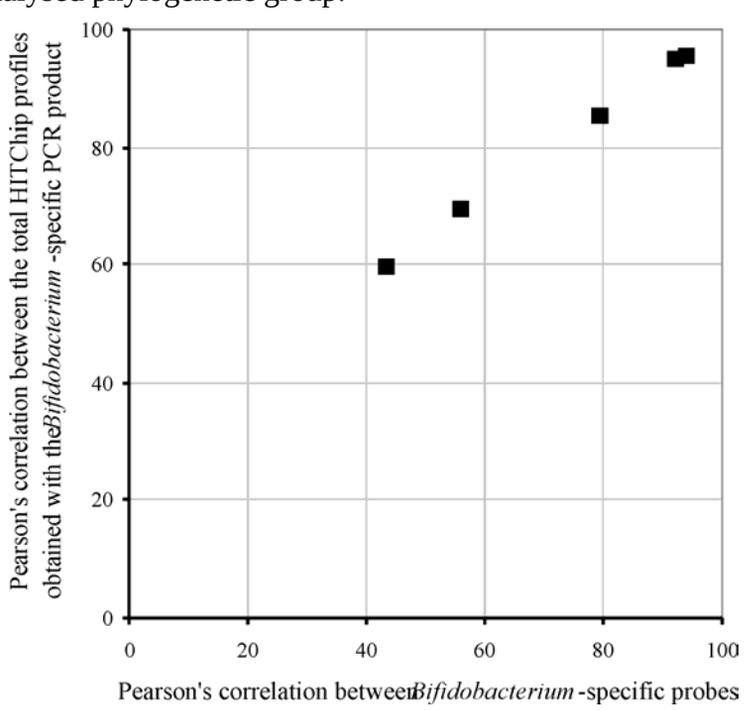


Figure S1 The plot of the correlation between the similarity of the *Bifidobacterium* communities in the faecal microbiota of five pairs of monozygotic twins calculated as Pearson's correlation between the total hybridisation profiles obtained with *in vitro* transcribed *Bifidobacterium* specific PCR product (y axis) and between hybridisation signals of *Bifidobacterium* specific probes obtained with *in vitro* transcribed universal bacterial PCR product (x axis).

Results and Discussion

Assessment of the host specific signature

The faecal microbiota composition of five unrelated and healthy individuals was profiled using two SSU rRNA-based techniques, the HITChip phylogenetic microarray and DGGE of SSU rRNA amplicons. Clustering of phylogenetic fingerprints obtained by the HITChip analysis resulted in grouping of the faecal microbiota profiles of the same individual collected at different times, and was presented for the samples collected after four years interval (Fig. 1). This result was in line with the previously established high similarity of the microbiota profiles (489, 537), although the time interval between sampling dates was never as long as presented in this study. This individual-based clustering of the microbiota profiles was observed despite an active life of the selected individuals, which included travelling oversees, antibiotic therapies, and in two cases pregnancy, and showed that the gastrointestinal microbiota represents a unique fingerprint of each subject, which can be supported by the findings of other studies (275, 311)

In contrast to the results assessed by the HITChip, the individual-based clustering was not detected when analysing the DGGE profiles of the faecal samples that were taken with the time span of four years (data not shown). However, it should be noted that DGGE, and other similar profiling techniques, does not have sufficient resolution for complex ecosystems such as the human gastrointestinal microbiota, as it was shown that DGGE bands sometimes represent mixtures of phylotypes (71, 482).

In conclusion, the HITChip phylogenetic microarray provides an unprecedented and high resolution profiling of the human gastrointestinal microbiota that is directly linked phylogenetic information (Fig. 1). Another, conceptually different, phylogenetic microarray for studying the same ecosystem has been recently described (359). This microarray was shown to be a powerful tool for identification and quantification of 359 gastrointestinal phylotypes. Although it cannot be ruled out that the microarray developed by Palmer and colleagues, has similar potentials, the HITChip is the first microarray to which high resolution fingerprinting has been appointed.

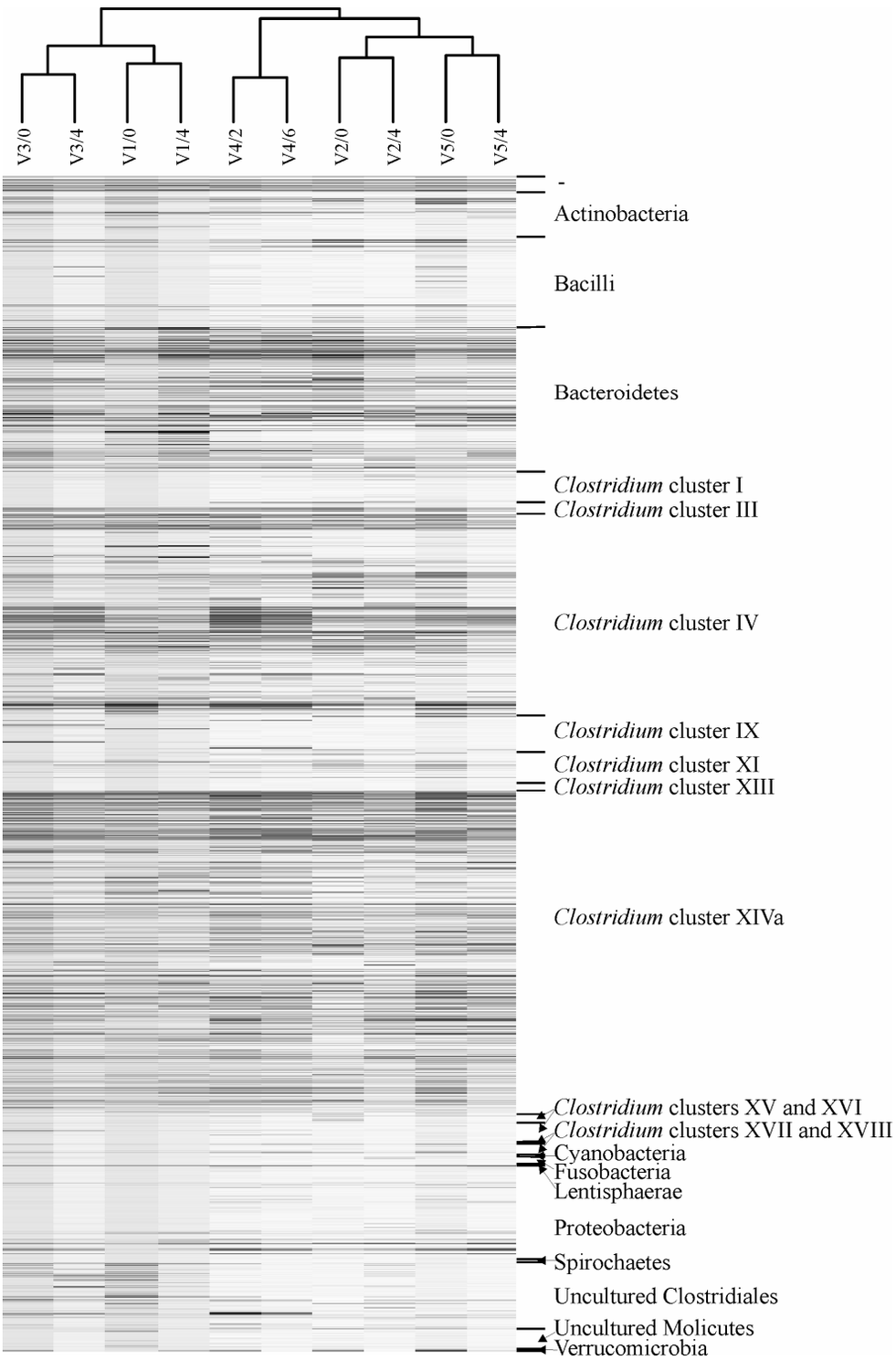


Figure 1 Phylogenetic fingerprints of the human gastrointestinal microbiota of five subjects in four years distant time points. The healthy volunteers are encoded by V1-V5, while the numbers subsequent to slash sign indicate the time span in years after the first sampling. The highest phylogenetic level of specificity of probes is depicted on the right panel of the figure.

Phylogenetic insight into the microbiota's dynamics

Based on the finding that faecal microbiota profiles, obtained with the HITChip, allowed identification of individual-specific pattern, an additional question was raised: which members of the complex microbial community in the human intestine are relevant for the preserving these patterns. Hence, the changes of the microbiota composition were quantified by calculating the similarity between the microbiota profiles of five subjects after six months and four years as well as between the microbiota of five pairs of monozygotic twins and unrelated individuals (Fig. 2). The calculated similarity index showed a negative correlation with the length of time between sampling and has decreased from 79.5% after six months to 63.2% after a period of four years. Even so, both values were found to be significantly higher when compared to the similarity of unrelated individuals, which was in average 41.7% ($p<0.001$ and $p=0.007$ for six months and four years, respectively). The average similarity between the microbiota of monozygotic twin pairs was 55.6%, and although this value was lower than that obtained for temporal variation in both analysed time intervals, it was notably higher than that of unrelated individuals, and this difference has almost reached significant level ($p=0.067$). These findings are in line with the results of the microbiota analysis by the other fingerprinting techniques (459, 536). After a time span of four years, the similarity of the microbiota within the same individual has decreased and although it did not reached the level of similarity between pairs of monozygotic twins (63.2% versus 55.6%), a significant difference between two data sets could not be found ($p=0.170$).

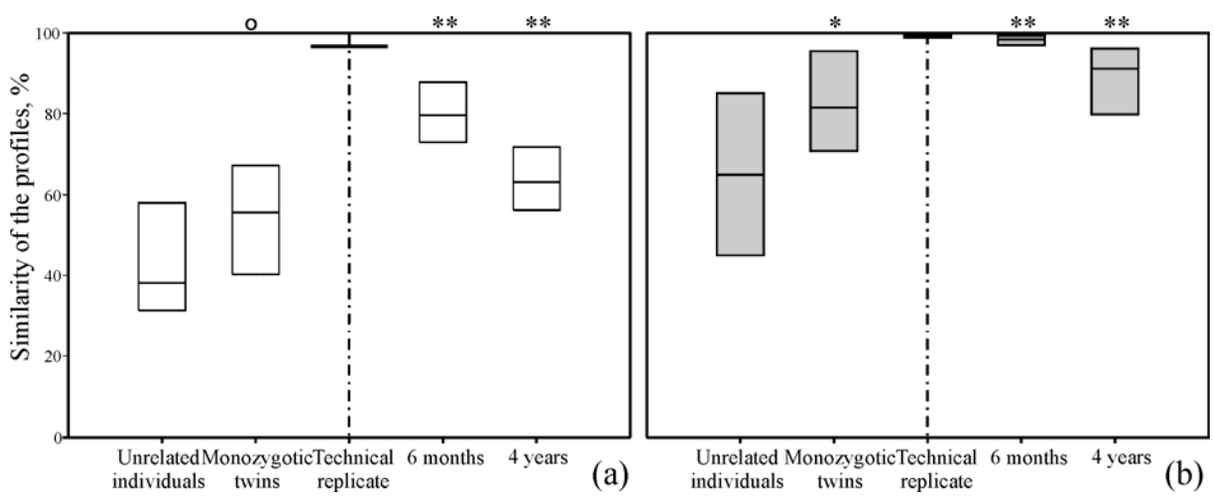


Figure 2 Similarity of the total microbiota (a) and *Bifidobacterium* profiles (b) calculated for temporal variation for two time intervals, and for monozygotic twin pairs and unrelated individuals presented as box plots. The box extends from 25th percentile to 75th percentile, with a line at the median. The technical replicate, defined as similarity of profiles obtained with two subsequent DNA extractions from one faecal sample, provides a measure of technically introduced biases of the obtained values. Double asterix indicates highly significant difference in similarity when compared to unrelated individuals ($p<0.01$), asterix indicates significant difference ($P<0.05$), while an empty circle indicates difference on the edge of significance ($p=0.067$).

The monozygotic twins, whose faecal microbiota was analysed in this study are adults (average age 37 years) that have been living in separated households for at least 10 years. Although we cannot exclude founder effects, the established high similarity of the twin microbiota profiles, which was confirmed by clustering of microbiota profiles (data not shown), suggests that the intestinal microbiota is strongly influenced by host's genotype. The changes of the microbiota composition can be attributed to the influence of two factors: host's genotype, and environmental stimuli. In addition, the variation of the gastrointestinal microbiota, at least partially, can be considered as stochastic event (112).

To provide phylogenetic insight into the subsets of the microbiota which are potentially influenced by either of those factors we have assessed the similarity of group-specific profiles. The assessed similarity values followed similar trend as the total microbiota in respect to the time span and genetic relatedness, as shown for *Bifidobacterium* subpopulation on Fig. 2. Furthermore, several members of two phyla, Actinobacteria and Bacteroidetes, have found to be significantly more similar when compared to the total microbiota (Table 1).

Despite the fact that this is the first report in which it is shown that specific subset of the human gastrointestinal microbiota is shown to be preserved from long-term changes, it has already been established that intestinal homeostasis is regulated by receptor mediated recognition of enteric bacteria (385). It can be speculated that the recognition of commensal bacteria includes specific identification of microbial populations which are the essential for performing some of numerous beneficial functions of the human gastrointestinal microbiota.

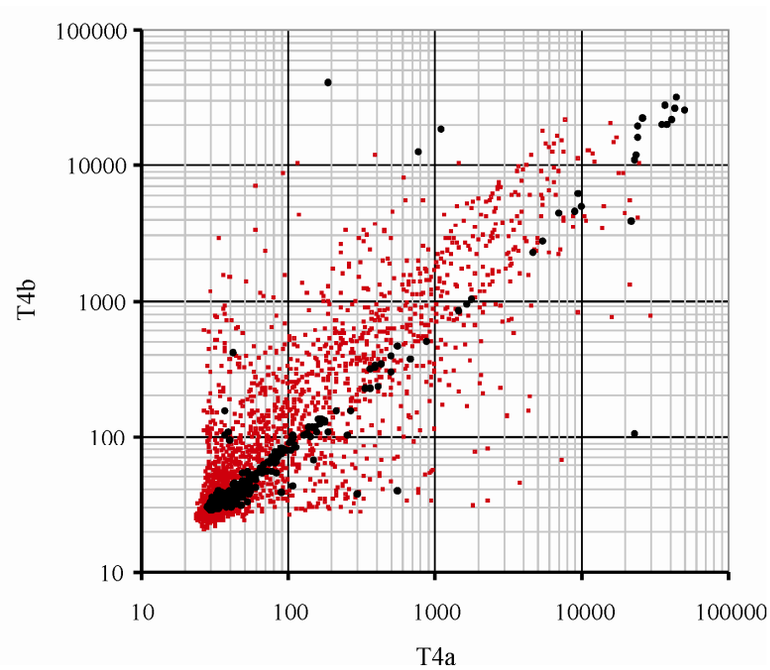


Figure 3 Plot showing the correlation between the HITChip hybridisation signals obtained for a pair of monozygotic twins with *in vitro* transcribed universal bacterial PCR (▪) and *Bifidobacterium* specific PCR products (●).

Table 1 Similarity between the faecal microbiota profiles of five individuals with time span between sampling of four years, calculated as Pearson's correlation for the total microbiota and different phylogenetic groups. P values are provided for the difference in the similarity of the total microbiota and specific phylogenetic group.

Higher taxonomic group	Lower taxonomic group	Average similarity	P value
Total microbiota		63.2%	
Actinobacteria	Overall	86.2%	**0.003**
	Bifidobacterium	88.6%	**0.001**
	Collinsella	78.2%	**0.048**
	Eggerthella	66.3%	0.434
Bacteroidetes	Overall	75.2%	**0.014**
	Alistipes	65.5%	0.400
	Bacteroides fragilis-like	76.7%	0.075
	Bacteroides intestinalis-like	90.2%	**<0.001**
	Bacteroides ovatus-like	79.7%	0.104
	Bacteroides plebius-like	81.5%	0.052
	Bacteroides stercoris-like	71.4%	0.267
	Bacteroides uniformis-like	81.6%	**0.037**
	Bacteroides vulgatus-like	97.3%	**<0.001**
	Parabacteroides	69.4%	0.183
	Prevotealla intermedia-like	88.0%	**0.001**
	Prevotella melaninogenica-like	78.2%	0.170
	Prevotella oralis-like	95.6%	**<0.001**
	Bacteroides splanchnicus-like	81.0%	**0.012**
Clostridium cluster IV	Overall	71.4%	0.198
Clostridium cluster XIVa	Overall	66.0%	0.386
Uncultured Clostridiales	Overall	63.6%	0.487

The groups that have exhibited higher similarity when compared to the total microbiota, both in the long time span and between monozygotic twin pairs, are the most profoundly influenced by the host's genotype, probably because of their beneficial functions. The similarity values between the group-specific profiles highlighted significant contribution of *Bifidobacterium* and *Collinsella* spp. and some Bacteroidetes in the fraction of genotype influenced microbiota. Remarkably, recent genomic characterisation of cultured representatives of these groups showed the presence of a prominent glycobiome involved in the degradation of host and diet specific sugars and sugar polymers (421, 530). The observed, strikingly higher level of similarity between the *Bifidobacterium* profiles when compared to the total microbiota similarity (p=0.001), which is supported by previously established short-term stability of this group (417, 489), was confirmed using *Bifidobacterium*-specific assay in

combination to the HITChip analysis for the assessment of similarity of the faecal samples of monozygotic twin pairs (Fig. 3).

There are several reports that indirectly suggest bifidobacterial adaptation for the conditions of the gastrointestinal tract, that include the fact that intestinal samples are almost an exclusive environmental source for isolation of these bacteria (50, 223), and that their detection in other ecosystems indicates faecal contamination (392). Recent genomic characterisation of *B.longum* confirmed its adaptation to the conditions of the gastrointestinal tract (421), while their prominent stability suggests an active crosstalk between this bacterial group and the human host. The beneficial role of bifidobacteria for the human host has been suggested from other ecological studies, which include the fact that *Bifidobacterium* spp. are on of the first colonisers of the human gastrointestinal tract (364). In addition, latest data indicate that enzymes that are characteristic for bifidobacteria are produced already after several weeks in the intestinal tract of neonates (E. Klaassens, unpublished data). Moreover, their low abundance is correlated with vitamin K deficiency and rectal bleeding in infants (44) or allergy incidence (237). Finally, the beneficial role of *Bifidobacterium* spp. is confirmed through their application as probiotics in treatment of irritable bowel syndrome (354), allergy (529) or diarrhoea (512).

In addition to bifidobacteria, several members of the phylum Bacteroidetes were found to have remarkable stability, which is in agreement with the previously published results of short term studies (489). Similarly to bifidobacteria, bacteroides are another typical intestinal inhabitants (261, 409), which are also one of the firstly developed groups of the neonates' microbiota (364). It is known that bacteroides are responsible for the fermentation of complex carbohydrates (409), although, unlike bifidobacteria, they represent a numerous and diverse group, the members of which have recently been reclassified into several other families (389, 406, 431). Our results showed that although Bacteroidetes as whole appeared to be more stable than the total microbiota, not all groups were affected in the similar manner. Whereas some *Bacteroides* spp. and *Prevotella* spp. were hardly affected by time-induced changes - similarity of *Bacteroides vulgatus*-like organisms between time span of four years was 97.3% - *Alistipes* and *Parabacteroides* spp. change was more pronounced as obtained similarity values were 64.5% and 69.9%, respectively (Table 1).

Ecological studies of Bacteroidetes are often performed using low level of phylogenetic resolution, and this might be a relevant reason for the fact the association of this group of bacteria to the human health status is not as clear as for bifidobacteria. Nevertheless, it has been shown that decreased levels of Bacteroidetes are associated with obesity (275). Similarly, decreased levels of *Bacteroides spp.*, more specifically *B. vulgatus*-, *B. ovatus*- and *B. uniformis*-like organisms, were associated with irritable bowel syndrome (**Chapter 7**). Interestingly, two of these *Bacteroides* subgroups have exhibited a remarkable level of long term stability in healthy adults. Furthermore, animal model experiments have shown that *B.*

vulgatus has a protective effect in *Escherichia coli*-induced colitis (500), while low abundance of this bacterium has been associated to Cronh's disease and ulcerative colitis mucosa (93). Little is known about mechanisms of interaction between this microbe and its host, as another model organism of this genus - *B. thetaiotaomicron* – has been studied in great detail. It has been shown that this microbe improves host's nutritional and defensive functions and can regulate production of host derived substrates required for its growth (210, 211). *B. thetaiotaomicron* is an abundant member of the human gastrointestinal microbiota, which, however, is considerably affected by temporal changes as previously found (206), and confirmed by our results (*B. fragilis*-like organisms, which include *B. thetaiotaomicron*), have shown considerable temporal variation. Therefore, it can be speculated that other interesting findings would rise from the analysis of interaction between the one of the selectively preserved members of the human gastrointestinal microbiota and its host.

Nine-year follow up of the microbiota changes

To provide better insight into character of the gastrointestinal microbiota dynamics, we have analysed five faecal samples of an individual collected during period of nine years. As similarity of the total profiles expresses both quantitative and qualitative differences between analysed samples, the distribution of positively hybridising probes was assessed, to provide a measure of qualitative changes alone. The results showed that approximately two thirds of responding probes were common for all analysed samples, while a number of probes was detected in more than one sample, which resulted in a complex network of the temporal variation of the microbiota's composition (Fig. 4).

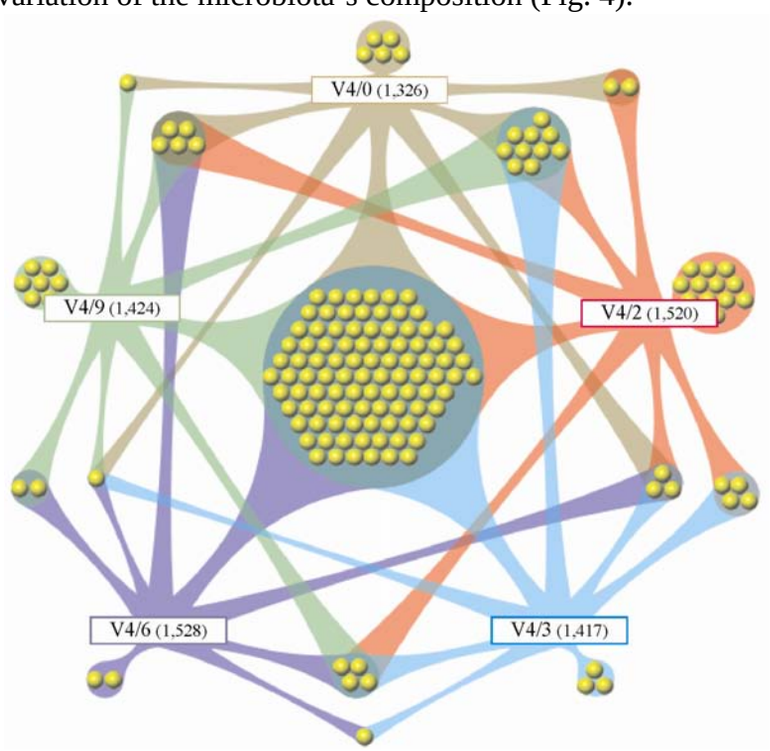


Figure 4 Venn diagram showing the distribution of the HITChip probes that had significant hybidisation signal when analysing five faecal samples of one subject. V4 is the code of the subject, numbers subsequent to slash sign indicate the time span in years after the first sampling, and numbers in parantheses indicate the total number of responding probes. Each yellow sphere represents ten responding probes. The image was rendered by AutoFocus software (Aduna B.V., the Netherlands).

Probes that had significant hybridisation signal in all analysed samples correspond to about 200 phylotypes, which represent the core microbiota, the permanent colonisers of this particular person, while the other detected phylotypes, correspond to occasional and incidental members of the microbiota. In average, only about 5% of the detected phylotypes appeared to be sample-specific. The detection of these organisms seems to be a stochastic event, and appearance of such, incidental colonisers, probably includes elements of chance.

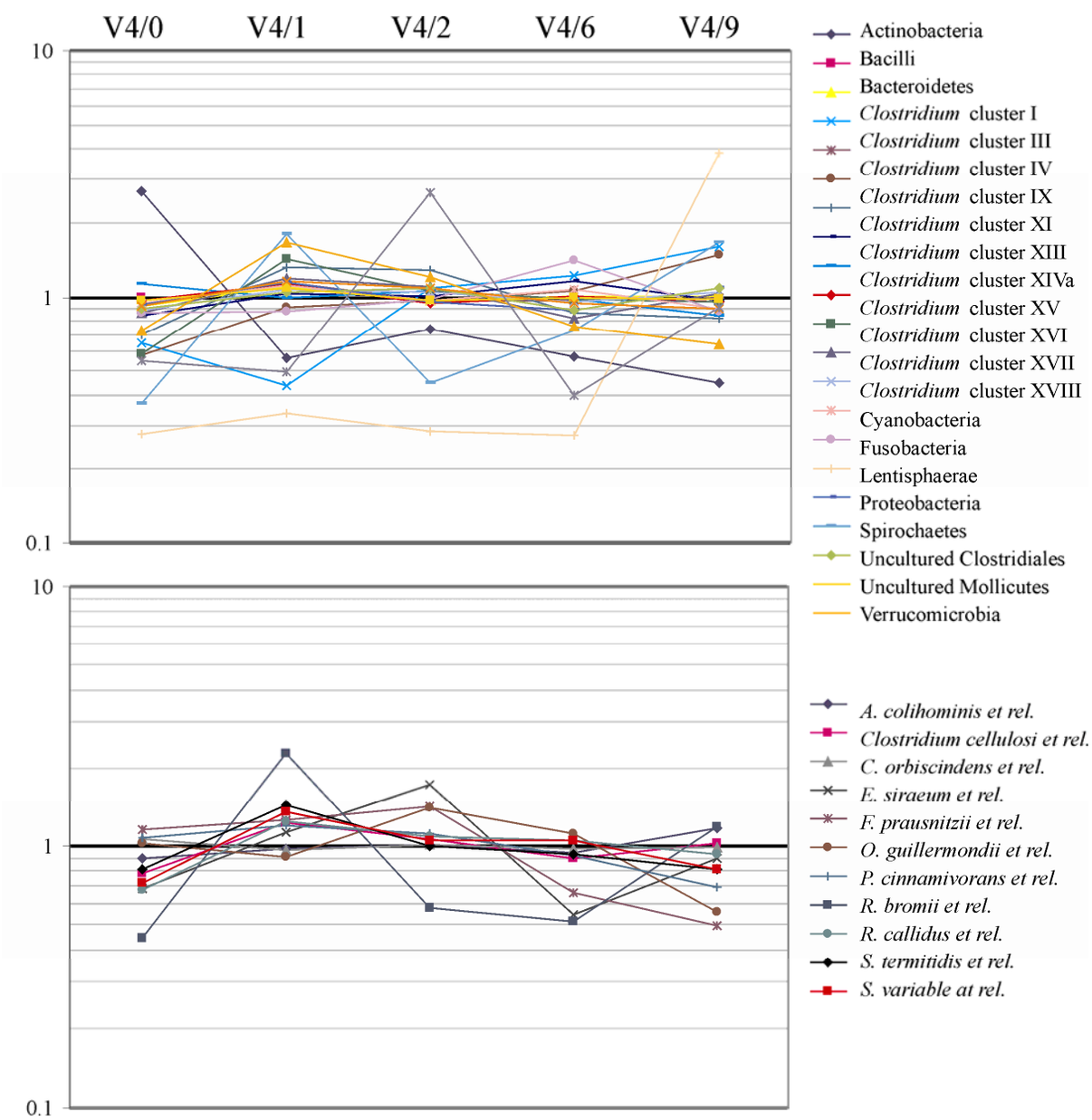


Figure 5 The relative quantitative changes of gastrointestinal groups during nine years follow up of the microbiota of an individual shown for two levels of phylogenetic specificity. The relative quantitative change is expressed as a ratio of average hybridisation signal obtained with the group specific probes for particular time point and arithmetic mean of average hybridisation signals for all analysed time points.

The proportion of sample specific phylotypes seemed larger if only two samples were compared as it had average value of 14%. Therefore, the assessment of the similarity between only two samples underestimates the component of individual-specific microbiota as it oversees the occasional colonisers of the particular individual. Using a cultivation approach, another nine-year study of the human microbiota composition has been performed and also indicated that microbiota contains permanent, occasional and incidental members (473). A negative correlation with the time span length characterised both the number of common hybridising probes and similarity between the total microbiota profiles (correlation coefficients -0.645 and -0.449, respectively). This result confirmed that environmental factors have an impact on shaping the microbiota's composition.

To address quantitative changes of the microbiota, each analysed sample was explained by average hybridisation signal obtained for probes specific for different phylogentic groups. The analysis of such datasets indicates relative abundance of groups and can be directly comparable between samples (**Chapter 4**). The changes in relative abundance were assessed for two phylogenetic levels. Higher taxonomic level was defined as phylum, and in case of Firmicutes, which are the most diverse and abundant group within the gastrointestinal microbiota, *Clostridium* cluster-level (87), while lower taxonomic level (genus-like) was defined based on SSU rRNA gene sequence similarity and was named after one of the cultured representatives of phylotypes with 90% or higher sequence similarity (**Chapter 4**). These results showed that microbiota composition was fluctuating, and that fluctuations were more pronounced on higher phylogenetic level (Fig. 5). On higher phylogenetic level, the abundance of each group was changing for 40% of its average value, while for the lower phylogenetic level this change was at level of 22%. These results are in agreement with previously assessed quantification changes in the microbiota, although they have been defined on the bases of changes during period less than a year, and for smaller number of phylogenetic groups (306).

Conclusions

This study is the first report of SSU rRNA-based analysis of the long term microbiota's dynamics. Our results show that the temporal variations of the microbiota are more pronounced with the extended time span, showing that environmental stimuli are an important factor in shaping the microbial community within the gastrointestinal tract. Despite environmentally introduced changes, the microbiota of each subject has unique profile which makes it distinct from the microbiota of unrelated individuals. Similar observations applied on the microbiota of monozygotic twin pairs, which suggested strong influence of the host genotype on the microbiota's composition. Furthermore, not all members of the microbiota were equally affected by the temporal variation, and the groups that exhibited higher level of

stability were fount to be the same groups that made the microbiota of monozygotic twins more similar. These results allowed us to hypothesise those subsets of the microbiota, primarily *Bifidobacterium* spp. and some members of the phylum Bacteroidetes, are having the most active crosstalk with the human host. Nine years long follow up of the microbiota of an individual showed that temporal variation of the microbiota can be explained by complex network of the changes in composition, and that dominant group of permanent colonisers is accompanied by occasional and incidental members of the gastrointestinal microbiota. Although, this study is based on a limited samples size, as the microbiota of only 15 individuals was analysed, the insight into the microbiota's dynamics assessed in over 100 hybridisation experiments allowed us to hypothesise that the human gastrointestinal microbiota fluctuates around a core community of permanent colonisers, which composition is affected by the host genotype.

Acknowledgements

This study was carried out with financial support from the Commission of the European Communities, specifically the RTD programs Quality of Life and Management of Living Resources, QLK1-2000-108, Microbe Diagnostics, coordinated by Michaël Blaut (Dife, Germany). The study does not necessarily reflect the Commission's views and in no way anticipates its future policy in this area.

CHAPTER 7

Mirjana Rajilić-Stojanović, Elena Biagi, Kajsa Kajander,

Hauke Smidt, Riitta Korpela, Erwin G. Zoetendal and Willem M. de Vos

Manuscript in preparation

Phylogenetic Microarray Analysis Reveals Specific Faecal Microbiota Composition in Patients Suffering from Irritable Bowel Syndrome

The intestinal microbiota has been suggested as a relevant aetiological factor of irritable bowel syndrome (IBS), although direct links have not been established primarily because of exceptional complexity of the microbial ecosystem within the human intestine. To address this, we have employed a comprehensive phylogenetic microarray, the HITChip, to study the faecal microbiota of primary care IBS patients, and age and sex-matched healthy controls. The total microbiota of IBS patients appeared to be distinctive from this of healthy subjects based on clustering of their respective phylogenetic profiles. Furthermore, the composition of the faecal microbiota of healthy volunteers was more homogenous than that of IBS patients, which could be explained by multi-factorial nature of IBS as patients recruited for this study that could be organised into three subgroups: alternating, diarrhoea or constipation-predominant. The microbiota of constipation-predominant IBS patients was the most similar to that of healthy volunteers, whilst the microbiota of diarrhoea-predominant IBS patients was the most distinct (p values 0.458 and 0.066, respectively). In general, faeces of IBS patients were characterised by significantly lower levels of *Bacteroides* species (p=0.018), and increased levels of members of the Bacilli order (p=0.002). Specifically, the microbiota of alternating and diarrhoea-predominant IBS patients differed in abundance of several subgroups of the *Clostridium* cluster XIVa, while levels of another major bacterial group, *Clostridium* cluster IV, were significantly higher in constipation-predominant when compared to diarrhoea-predominant IBS patients. In conclusion, HITChip analysis of the faecal microbiota of IBS patients indicated that it is disturbed and specific for each type of IBS.

Introduction

Irritable Bowel Syndrome (IBS) is one of the most common functional intestinal disorders and its definition includes symptoms of abdominal pain, bloating or discomfort, and diarrhea or constipation lasting over a period of time (474). IBS is most frequently diagnosed for women of the developed world (184). Its strong association with the developed countries suggests influence of environmental factors, while a recent population-based twin study indicates a strong genetic influence on the development of the IBS (41). Even though several currently available therapies, which include the consumption of dietary fiber, laxatives but also recently introduced probiotics, provide symptomatic relief, none have shown to alter the natural history of the disorder (3, 14, 235, 354). This is primarily due to the multi-factorial nature of IBS pathogenesis and pathophysiology coupled to the fact that the aetiology of this disorder is still unknown.

Various reports have implied the gastrointestinal microbiota as an aetiological factor of IBS. Indirect support for the involvement of bacteria in IBS derived from the observation that antibiotic therapy could alleviate IBS symptoms in patients with bacterial overgrowth syndrome (372), while - in contrast - the use of antibiotics for non-gastrointestinal complaints increased the likelihood for the development of functional abdominal symptoms (310). Moreover, distinct microbial fermentation patterns, associated with an increased hydrogen production, were observed in the bowel content of IBS patients (247). *In vivo* measurements confirmed the disturbed bowel metabolism in IBS patients, which was associated with an increased gas volume (257). More recent studies addressed the faecal microbial composition using profiling, sequencing, and qPCR analyses of bacterial small subunit ribosomal RNA (SSU rRNA) genes. These revealed a temporal instability of the microbiota and its misbalance expressed via an increase of total anaerobes (308), a decrease of *Bifidobacterium* spp. (28, 299) or lactobacilli and coliforms (28), or an increase of species belonging to the *Clostridium* cluster XIVa (308, 309). Moreover, it could be ruled out that well-established pathogens such as *Campylobacter* and *Helicobacter* spp. are involved in IBS (299). Altogether, these observations suggest a deviation in the composition and activity of the gastrointestinal microbiota in relation to IBS. The direct relations or causal effects have not been established, mainly because the used approaches only target specific microbial groups of the extremely complex ecosystem that resides in the human gastrointestinal tract (**Chapter 3**). Hence, a global and efficient analysis of the total microbial composition is an alternative approach to unravel specific characteristics of the gastrointestinal microbiota composition in IBS patients.

In this study, we have addressed the link between the global gastrointestinal microbiota composition and IBS, using the Human Intestinal Tract Chip (HITChip), a comprehensive, high throughput, phylogenetic microarray enabling the simultaneous qualitative and semi-quantitative analysis of over 1,000 gastrointestinal phylotypes (**Chapter**

4). A careful comparison between the microbiota in IBS and healthy subjects revealed specific differences of both composition and relative abundance of several microbial groups within the two abundant gastrointestinal phyla – Bacteroides and Firmicutes.

Materials and methods

Study subjects

Twenty primary care IBS patients that fulfilled the Rome II criteria (474) were recruited for this study. IBS patients were a sub set of a larger clinical trial that investigated the effects of probiotics in IBS (236). Inclusion criteria for the patients were: an IBS diagnosis consistent with the Rome II criteria; a clinical investigation with colonoscopy or barium enema of the colon performed during the preceding five years; an age between 28 and 65 years (average age 53); normal blood count (erythrocytes, haemoglobin, haematocrite, MCV, MCH, MCHC, thrombocytes, leucocytes) and within reference values for serum creatine, ALAT and AFOS. Subjects were excluded if they had organic intestinal diseases; previous major or complicated gastrointestinal surgery; severe endometriosis; complicated abdominal adhesions; malign tumours; were pregnant or lactating; had received antimicrobial medication during the previous one month or had dementia or were otherwise incapable of adequate co-operation. The patients were divided over three IBS subgroups: diarrhoea-predominant (seven patients), constipation-predominant (eight patients) and alternating IBS (five patients). None of the patients had a known history of gastroenteritis during the preceding three years. All patients were Finnish women. Additionally, 20 healthy, age-matched Finnish women, aged between 28 and 58 years (average 50 years), took part in this study. All the participants gave their written informed consent. The Ethics Committees of the Pirkanmaa Hospital District (PSHP, Finland) and of the Joint Authority for the Hospital District of Helsinki and Uusimaa (HUS, Finland) approved the study.

Faecal samples collection and DNA extraction

The faecal samples were collected in plastic containers and were immediately frozen to -20˚C by the subjects. Samples were transported to a -45˚C freezer, where they were stored until further processing. Approximately 2 g of faecal sample was diluted (1:10) in phosphate buffered saline (PBS, pH 7.4) and homogenized for 2 min in a stomacher bag containing the inner filter pouch (Seward, Norfolk, UK). Of this dilution, 1 ml was further diluted in PBS (1:10). From 1ml of this dilution (1:100), cells were collected (2 min, 14 000 g) and DNA was isolated with a Wizard Genomic DNA purification kit (Promega) according to the manufacturer's instruction with minor modifications (during the 1 h incubation at 37˚C,

lysozyme (8 mg/ml) and mutanolysin (165 U/ml) were used as lytic enzymes. DNA yield was quantified using a NanoDrop ND-1000 spectrophotometer (NanoDrop® Technologies, Wilmington, DE).

PCR

Bacterial SSU rRNA genes were amplified using the primers *T7prom*-Bact-27-for (5'-TGA ATT GTA ATA C GA CTC ACT ATA GGG GTT TGA TCC TGG CTC AG–3') and Uni-1492-rev (5'-CGG CTA CCT TGT TAC GAC-3') (262). PCR reactions were carried out in a final volume of 50 µl, and 10 ng of DNA samples was used as template. Samples were initially denatured at 94˚C for 2 min followed by 35 cycles of 94˚C (30 sec), 52˚C (40 sec), 72˚C (90 sec) and a final extension at 72˚C for 7 min. The PCR products were purified by using the DNA Clean and Concentrator kit (Zymo Research, Orange, USA), according to the manufacturer's instruction. The final DNA concentration was determined as described above.

RNA production and labelling

In vitro transcription of the T7-promoter carrying SSU rRNA gene was performed according to the manufacturer's protocol using the Riboprobe System (Promega, La Jolla, USA), 500 ng of the T7-16S rRNA gene amplicon, and, besides rATP, rGTP, rCTP, a 1:1 mix of rUTP and aminoallyl-rUTP (Ambion Inc., Austin, Tx, USA). The transcription reaction was performed at room temperature for 2 h., the template DNA was digested applying the Qiagen RNAse-free DNAse kit (Qiagen, Hilden, Germany), and RNA was purified using the RNeasy Mini-Elute Kit (Qiagen, Hilden, Germany). RNA yield was measured as described above.

Amino-allyl-modified nucleotides were coupled with CyDye using the Post-Labeling Reactive Dye (Amersham Biosciences, Little Chalfont, UK), which was previously dissolved in 84 µl DMSO. Labelling reaction was performed in a 25 mM sodium bicarbonate buffer (pH 8.7) by adding 20 µl of dissolved CyDye to 2 µg of purified RNA in a final volume of 40 µl. Samples were incubated in the dark at room temperature for 90 min. The reaction was stopped by adding 15 µl of 4M hydroxyl-amine and incubating in the dark for 15 min. RNAse-free water was added to 100 µl and labelled RNA was purified and quantified as described above.

Microarray production, hybridisation and data extraction

The HITChip microarrays were produced by Agilent Technologies (Agilent Technologies, Palo Alto, CA., USA). The oligonucleotide probes were extended at the 3'-end (at the array support side) with 10nt long T-spacer, and were printed on the array using *in situ* surface-attached oligonucleotide probe synthesis (54). Arrays used in this study were of the 2×11K format, with two arrays per glass slide. Each array was hybridised with two samples, labelled with Cy3 and Cy5. Combined Cy3- and Cy5-labelled target mixtures were fragmented by adding 1 µl of Ambion 10× fragmentation reagent (Ambion Inc., Austin, Tx, USA), and

incubation for 20 min at 70˚C, according to the manufacturer's instruction. Fragmentation was stopped by adding 1 µl of Ambion stop solution. Hybridization mix was prepared by adding to the RNA mixture 31.5 µl of 20× SSC (412), 6.3 µl of 10% SDS (412), 20 µl of Agilent Control Target mix and RNAse-free water to a total volume of 210 µl. The microarray, hybridisation-mixture and cover-slide were joined and placed in the hybridisation chamber according to the manufacturers'instructions. The hybridisation was carried out at 62.5^{o}C in a rotation oven (Agilent, Palo Alto, Ca., USA) speed 4, for 16 h. Slides were washed at room temperature in a 50 ml Greiner tube containing 50 ml 2× SSC, 0.3% SDS for 10 min and at 38˚C in 0.1× SSC, 0.3% SDS for 10 min. SDS was completely removed by washing the slides in 0.06× SSPE (412) for 5 min, followed by a quick dry with compressed nitrogen. All wash-steps were performed with the aid of a Stuart Rotator SB3 (Barloworld Scientific, Stone, UK), set at speed 40rpm.

Data were extracted from microarray images using the Agilent Feature Extraction software, version 7.5 (www.agilent.com). Data was normalised using a set of R based scripts (http://www.r-project.org/), microarray were analysed in a custom designed relational database which runs under MySQL database management system (http://www.mysql.com/) using a series of custom R scripts as previously described (**Chapter 4**).

Hierarchical clustering of probe profiles was carried out by calculating a distance matrix between the samples based on the squared difference between each pair of profiles (squared Euclidian distance – E^2) using Ward's minimum variance method. E^2 between profiles obtained on faecal samples of IBS subjects, of healthy subjects and between profiles of IBS patients and healthy subjects was calculated to assess the overall difference. For graphical representation the similarity between profiles was expressed as reciprocal squared Euclidean distance – 1/ E^2 (74).

Statistical analysis

Principal component analysis (PCA) was performed by using the software package Canoco for Windows 4.5 (271). Data matrices consisted of average signal intensities for 131 bacterial groups defined on the bases of the SSU rRNA gene sequence similarity, or single probe hybridisation intensity per sample. Data matrices were organised in Excel and converted into a Canoco-compatible format by using the WCanoImp utility. PCA was performed focusing on inter-species correlation and diagrams were plotted by using the CanoDraw program.

The Monte Carlo Permutation Procedure, further in the text referred as MCPP (212), was applied to assess statistical significance of the variation in large datasets in relation to the health status. Alternatively, for small datasets Student's t-test was performed, and calculated p-values were two tailed. P values lower than 0.05 were considered significant and were marked with an asterix in figures, while p values lower than 0.01 were considered as highly significant and were marked with a double asterix.

Results

Phylogenetic profiling of the total faecal microbiota

The microbiota profiles obtained from DNA extracted from faecal samples of 20 IBS patients with alternating (five patients), diarrhoea (seven patients) or constipation-predominant (eight patients) symptoms and 20 corresponding healthy controls were acquired based on the signal intensity of each of 3,699 distinct HITChip oligonucleotide probes. Further analysis provided direct insight into the phylogenetic identity of microbial populations differing between microbiota of IBS and healthy subjects. The profiles visualise the presence or absence of all targeted microbial species and were grouped using a hierarchical cluster analysis (Fig. 1). This unbiased grouping resulted in two distinct clusters, one of which consisted of profiles from 14 IBS subjects and 4 healthy volunteers, while the other cluster contained primarily profiles of healthy volunteers as 16 profiles of healthy volunteers were grouped with 6 profiles of IBS subjects (Fig. 1). When the composition of each cluster was defined using two dummy variables, 1 and 0 that were assigned to IBS patients and healthy subjects, respectively, the clustering appeared to be highly dependent on the health status ($p=0.001$). Clustering of the total microbiota profiles could not be correlated with any IBS subgroup.

The observed health status-dependent clustering of the microbiota profiles implied significant differences between the microbiota of IBS patients and healthy subjects. To assess the measure of overall variance between IBS patients, healthy subjects and members of these two groups, the squared Euclidean distance was calculated and presented as a reciprocal value on Fig. 2. A higher value of the squared Euclidean distance among IBS patients indicated stronger variation of the faecal microbiota composition. The total microbiota profiles within group of IBS patients were found to be notably more different when compared to group of healthy subjects, with the difference between the groups almost reaching statistical significance ($p=0.059$). These results indicated that the microbiota of healthy subjects was more homogenous, despite the known, strongly pronounced inter-individual variation of the microbiota of healthy adults (129, 275, 537).

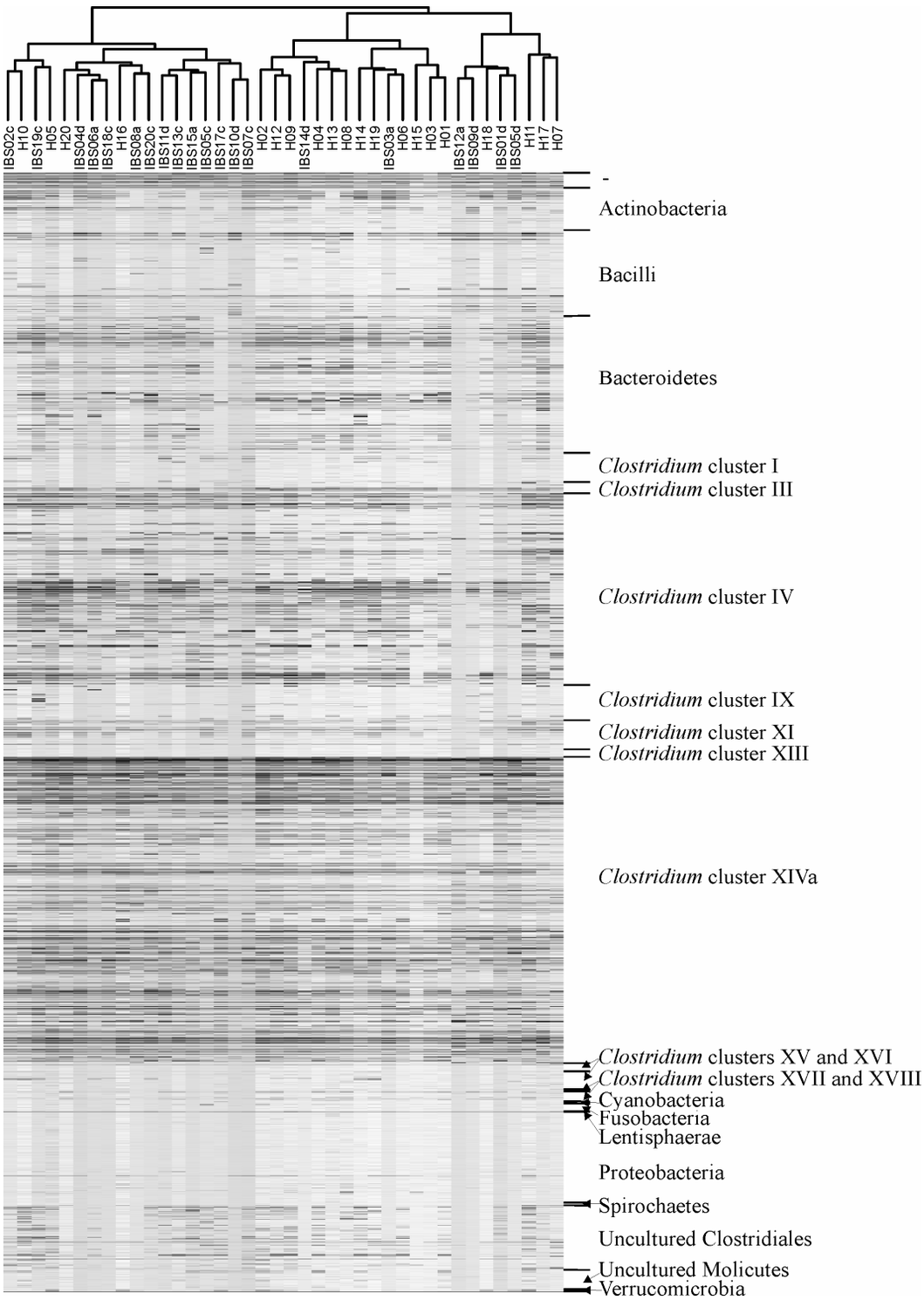


Figure 1 Phylogenetic fingerprints of the faecal microbiota of 20 IBS patients and 20 healthy volunteers. IBS patients profiles are encoded by IBS1-IBS20, and a, c or d for subgroups alternating, constipation and diarrhoea IBS, respectively. Samples of healthy control subjects are denominated by H1-H20. The highest phylogenetic level of specificity of probes is depicted on the right panel of the figure.

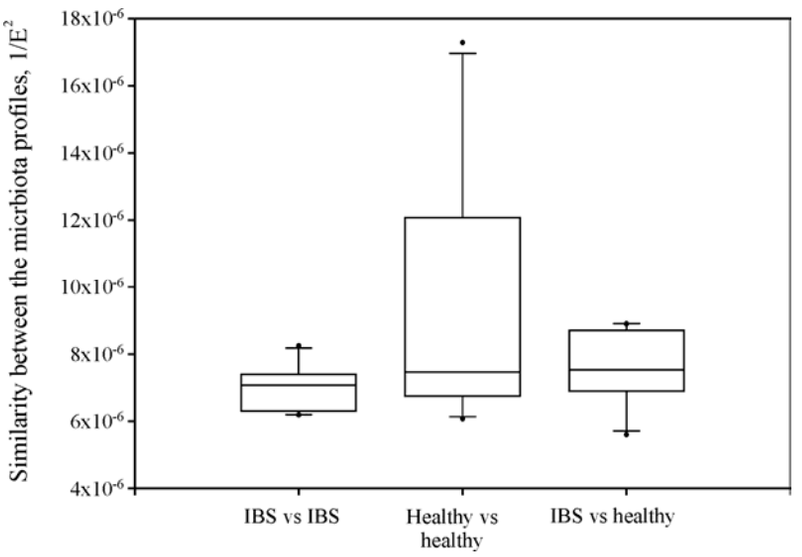


Figure 2 Similarity between the total microbiota profiles expressed as reciprocal Euclidean distance, between IBS patients, between healthy subjects, and between healthy subjects and IBS patients. The box extends from 25th percentile to 75th percentile, with a line at the median. The whiskers are showing the highest and the lowest value after exclusion of outliers, which are presented by dots.

PCA analysis of the total faecal microbiota

To detail the differences between the obtained microbiota profiles, the average hybridisation signal intensities were determined for those sets of the HITChip probes that target 131 SSU rRNA-based, phylogenetic groups (**Chapter 4**). These values have shown to be semi-quantitative and directly comparable between subjects. The analysis of such a reduced dataset could provide insight into the differences in the microbiota composition of IBS patients and healthy subjects at a level of phylogenetic resolution where inter-individual variation is strongly reduced. PCA showed that, despite the apparent overlap of the microbiota of both groups, the microbiota of IBS and healthy subjects groups around clearly separated centroides that indicate the respective health status (Fig. 3). Nevertheless, the overall microbiota of the IBS subjects was not found to be significantly different from the microbiota of healthy subjects as judged by MCPP ($p=0.152$). This may be attributed to the fact that the IBS groups were rather heterogeneous and consisting of different subgroups. Hence, the faecal microbiota of each IBS subgroup was compared to that of healthy controls. MCPP-based analysis of significance showed that the microbiota of diarrhoea type IBS displayed the highest difference when compared to the microbiota of healthy subjects, and that this difference was close to being significant with a p value of 0.066 (Table 1). On the contrary, the faecal microbiota of constipated-predominant type IBS was the least different from the microbiota of the healthy subjects.

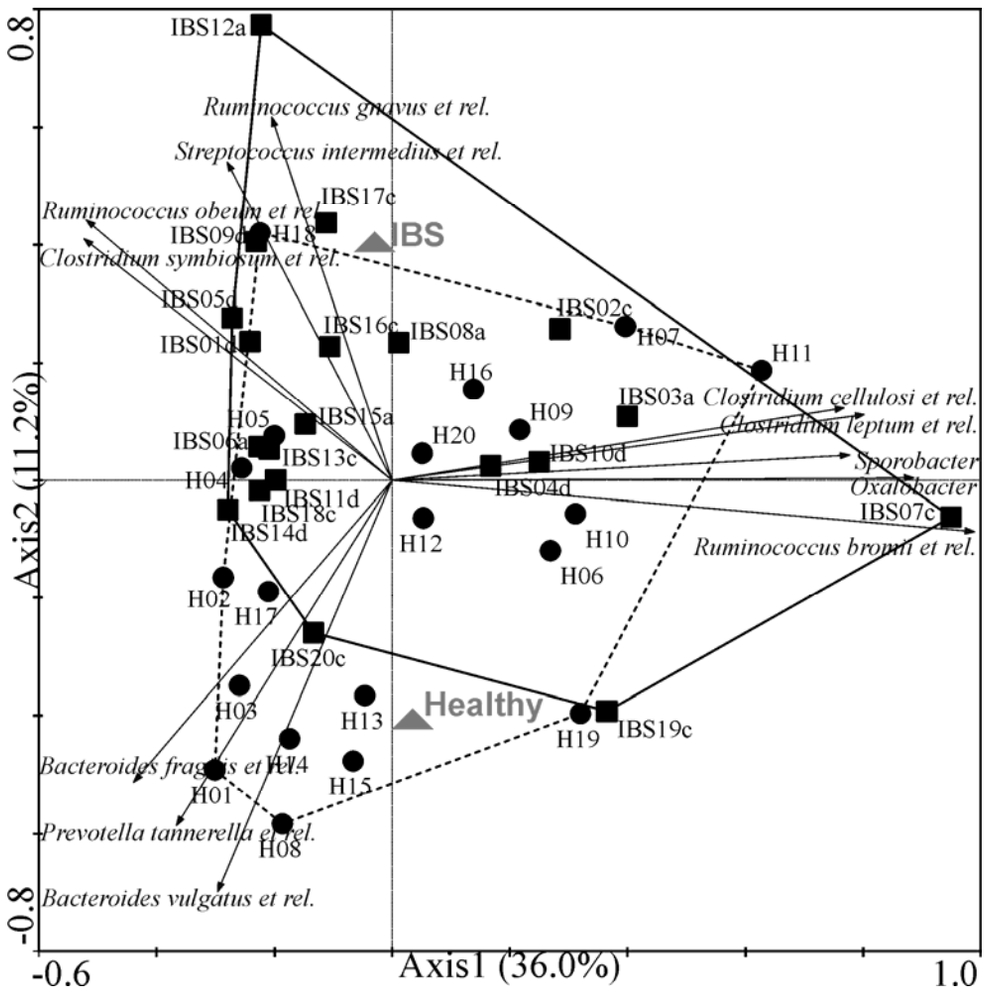


Figure 3 PCA triplot of the microbiota composition expressed by the average hybridisation signals for 131 phylogenetic groups for 20 IBS patients (■) and 20 healthy subjects (●). Species that contributed at least 40% to the explanatory axis used in the plot are presented as vectors, and health status (IBS or healthy) are centroides of the plot. Percentage values at the axes indicate contribution of the principle components to the explanation of total variance in the dataset. MCPP revealed that health status correlated to the variation in microbiota composition with a p-value of 0.152. See Figure 1 for samples codes.

P value	Healthy	IBSa	IBSc	IBSd
Healthy	1.000	0.338	0.458	0.066
IBSa		1.000	0.502	0.618
IBSc			1.000	0.176
IBSd				1.000

Table 1 Significance of difference between the overall microbiota composition of healthy subjects and patients suffering from all three subgroups of IBS, namely alternating, constipation and diarrhoea coded by IBSa, IBSc and IBSd, respectively. Entries show p values calculated by MCPP.

Assessment of differences between the faecal microbiota of healthy subjects and IBS patients

To identify the microbial populations responsible for the observed differences in the faecal microbiota composition of IBS patients, a Student's T test was performed on the average hybridisation signals of each of 131 SSU rRNA gene sequence-based phylogenetic groups. Initially, the microbiota of all IBS patients was compared to the microbiota of healthy subjects to assess common syndrome-specific denominators. Separate analysis of the microbiota of each of the IBS subgroups provided a more detailed portrait of the microbiota of this multi-factorial functional disorder.

IBS patients appeared to have a significantly lower level of *Bacteroides* spp. (p=0.018), while other genera of the Bacteroidetes phylum did not differ between the two analysed groups (Fig. 4). More detailed analysis showed that significant differences within the *Bacteroides* genus could be attributed to reduced numbers in IBS subjects of *B. vulgatus*-, *B. ovatus*- and *B. uniformis*-like organisms (Table 2). In addition, *Prevotella oralis*- like organisms showed an opposite trend for alternating type IBS and were significantly increased when compared to healthy subjects.

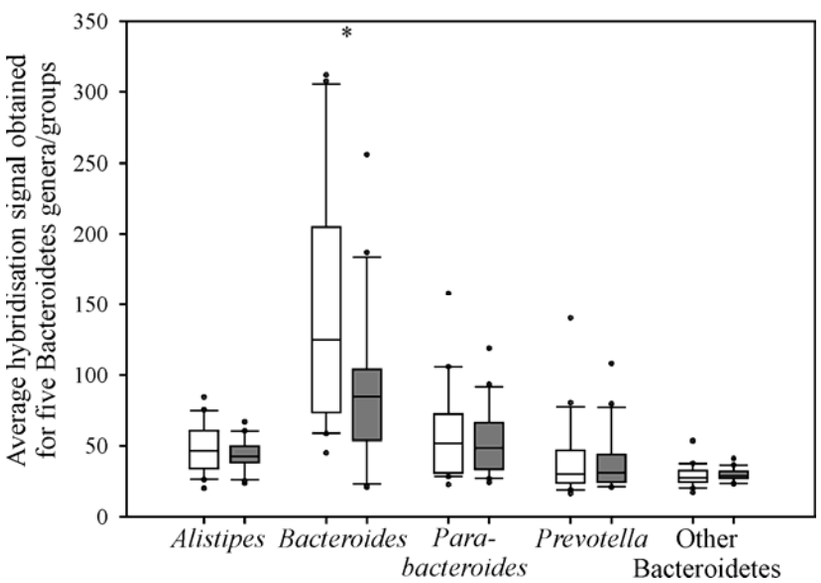


Figure 4 Average hybridisation signal obtained from members of the Bacteroidetes phylum, expressed in box and whisker plots (see Fig.2). Empty and filled boxes represent values obtained for 20 healthy subjects, and 20 IBS patients, respectively. Asterix indicates significant difference in average hybridisation signal intensity between healthy subjects and IBS patients (P<0.05).

Remarkably, the microbiota in the IBS subjects was found to be characterised by high levels of some groups within the Bacilli order (Fig. 5). A significant difference was observed between the microbiota of IBS and healthy subjects for members of the family Bacillaceae with p value of 0.001. Detailed analysis showed that in addition to *Bacillus* spp. that were significantly increased in all three types of IBS, *Aneurinibacillus* spp. displayed significantly higher levels in diarrhoea type IBS (Table 2). Even though no significant differences of Streptococcaceae were observed between the overall IBS microbiota and the healthy

microbiota, the levels of *Streptococcus intermedius*- and *Strepotococcus mitis*-like organisms were significantly increased in diarrhoea type IBS (Table 2).

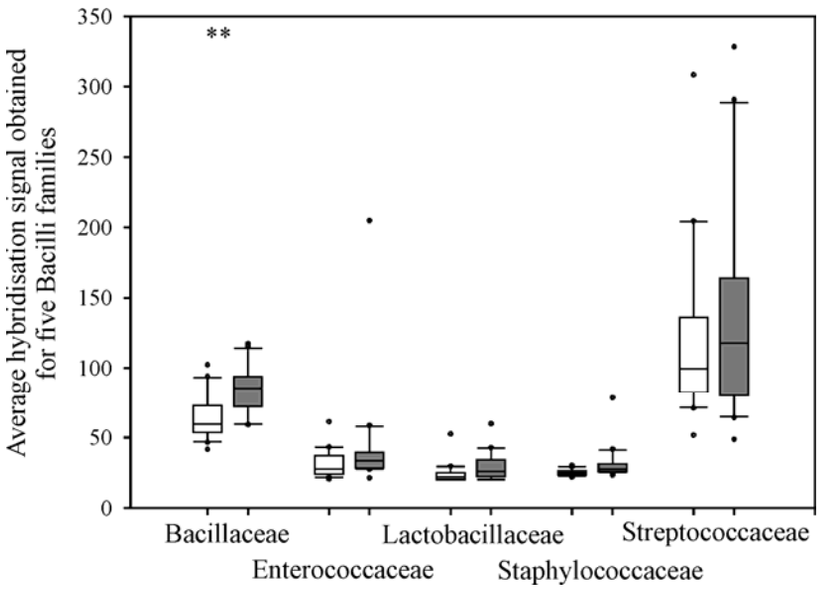


Figure 5 Average hybridisation signal obtained from members of the Bacilli order, expressed in box and whisker plots (see Fig.2). Empty and filled boxes represent values obtained for 20 healthy subjects, and 20 IBS patients, respectively. Double asterix indicates highly significant difference in average hybridisation signal intensity between healthy subjects and IBS patients (P<0.01).

Different levels of *Clostridium* cluster XIVa (also known as *Clostridium coccoides* group) appeared to be relevant for the overall observed differences in the microbiota composition of IBS patients and healthy volunteers. Unlike the difference observed for other taxa described above, the variation within the *Clostridium* cluster XIVa members appeared to be dependent on IBS subgroup. In addition to several groups, that were different between diarrhoea-predominant and alternating-predominant IBS patients and healthy subjects (Table 2), levels of *Anaerostipes caccae*-like organisms were decreased in the constipation type IBS patients and although the decrease was not significant when compared to healthy volunteers ($p=0.065$), it reached significance when compared to alternating type IBS ($p=0.027$).

Further differences between the microbiota of IBS patients and healthy subjects included significantly lower relative abundance of *Clostridium stercorarium*-like organisms (*Clostridium* cluster III) in constipation type IBS ($p=0.034$). Finally, constipation type IBS patients had significantly higher levels of *Clostridium* cluster IV members (also known as *Clostridium leptum* group) when compared to the diarrhoea type IBS patients ($p=0.017$). This difference was primarily caused by significantly higher hybridisation signals of several subgroups of *Clostridium* cluster IV, which were detected in faecal samples of patients that suffered from constipation-predominant type IBS (Table 2).

Table 2 SSU rRNA-based groups for which the relative abundance was found to be significantly different between healthy subjects and IBS patients of three subgroups of IBS, namely alternating, constipation and diarrhoea coded by IBSa, IBSc and IBSd, respectively. In addition to the group name, its physiological position is defined by the higher taxonomic group, the trend of the change observed, and p value is provided. Bold entries indicate a significance level below P=0.05.

Higher taxonomic group	Group	Change	P value		
			IBSa	IBSc	IBSd
Bacteroidetes	*Bacteroides ovatus*-like	IBS<healthy		**0.029**	
	Bacteroides uniformis-like	IBS<healthy		**0.036**	
	Bacteroides vulgatus-like	IBS<healthy		**0.021**	
	Prevotella oralis-like	IBS>healthy	**0.050**	0.633	0.651
Bacilli	*Aneurinibacillus* spp.	IBS>healthy	0.115	0.131	**0.004**
	Bacillus spp.	IBS>healthy		**0.002**	
	Streptococcus mitis-like	IBS>healthy	0.360	0.373	**0.017**
	Streptococcus intermedius-like	IBS>healthy	0.084	0.661	**0.005**
Clostridium cluster III	*Clostridium stercorarium*-like	IBS<healthy	0.521	**0.034**	0.651
Clostridium cluster IV	*Anaerotruncus colihominis*-like	IBSc>IBSd	-	**0.036**	
	Clostridium cellulosi-like	IBSc>IBSd	-	**0.039**	
	Oscillospira guillermondii-like	IBSc>IBSd	-	**0.011**	
	Sporobacter termitidis-like	IBSc>IBSd	-	**0.041**	
	Subdoligranulum variable-like	IBSc>IBSd		**0.041**	
Clostridium cluster XIVa	*Anaerostipes caccae*-like	IBSa>IBSc	**0.027**		-
	Clostridium colinum-like	IBS>healthy	**0.001**	0.114	**0.050**
	Clostridium symbiosum-like	IBS<healthy	**0.023**	0.226	0.090
	Roseburia intestinalis-like	IBS>healthy	0.762	0.238	**0.016**
	Ruminococcus gnavus-like	IBS>healthy	**0.006**	0.794	0.263

Discussion

In this study we have employed a comprehensive and high-throughput microarray tool for the assessment of potential syndrome-specific differences between the faecal microbiota of primary care IBS patients, and age and sex-matched healthy controls. Phylogenetic profiles of the total microbiota of IBS patients appeared to be distinctive from those of healthy subjects, which suggested significant differences in relative abundance of particular microbiota in both analysed groups. Furthermore, the composition of faecal microbiota of healthy volunteers was found to be more homogenous than that of IBS patients. This is probably due to the multi-

factorial nature of IBS and implies specificities of the microbiota of each IBS subgroup, which is in accordance with previous reports (308). Analysis of the total microbiota, which was explained by the average hybridisation signal of 131 phylogenetic groups, showed that the microbiota of constipation-predominant IBS patients was the most similar to the microbiota of healthy volunteers, while the microbiota of diarrhoea type IBS patients was the most distinct. Among IBS patients, constipation type and diarrhoea-predominant subgroups represented two extremes. Quantitative PCR (qPCR) analysis of the microbiota of IBS patients are principally in accordance with those findings (299). Further analysis revealed significant differences between the overall faecal microbiota of healthy subjects and all IBS patients, while a number of specific differences could be assigned to each of the IBS subgroups.

Some of our findings could not have been detected by analysis of the faecal microbiota on higher taxonomic level. This reinforces the advantage of in-depth and high throughput analysis using phylogenetic microarrays. Moreover, it also allows for discussing the relevance of our findings when compared to the results of other studies that dealt with the association of the microbiota composition to IBS. However, it should be kept in mind that any direct comparison is hampered by many factors of which a relevant is that the microbiota is individual-specific (275, 537). In each study, different subjects are recruited for the analysis of the microbiota composition, which has a certain influence on the results. Furthermore, each technique has a number of specificities, which in case of SSU rRNA-based approaches; include different target molecules (DNA versus RNA), different target sequences, and different taxonomic levels of specificity. Finally, the specific obstacle in the analysis of the IBS microbiota, rises from the fact that that the microbiota of IBS patients is significantly less stable than that of healthy volunteers (308), which makes it difficult to obtain a consistent pattern of microbiota changes in IBS patients by analysing only a single sample. The effect of temporal variation of the microbiota in IBS patients is illustrated by the observation that the relative abundance of several clostridial groups changes in time (309). Although, the extent of the changes did not reach a significant level when compared to healthy volunteers, it should be kept in mind that those variations might affect the microbial landscape presented in each study. Despite all those obstacles, the differences, and especially the significant ones, are certainly relevant for revealing the relation between IBS and the intestinal microbiota composition.

A new finding described in this study is that the microbiota of IBS patients is characterised by significantly lower levels of *Bacteroides* species. Although *Bacteroide*s spp. comprise a well established and frequently studied group of the gastrointestinal microbiota, in most of the studies term “bacteroides” is used as a synonym for the Bacteroidetes phylum that encompasses several distinct bacterial families (**Chapter 2**). This resolution of analysis, however, is not sensitive enough for the detection of profound differences such as those

observed in relation to IBS. An attempt to quantify only members of *Bacteroides* genus using Fluorescent *In Situ* Hybridisation (FISH) in mucosa-associated microbiota (in the particular paper referred as *B. fragilis* group) showed decreased levels of *Bacteroides* spp. in IBS patients when compared to healthy controls, despite equal levels of members of the Bacteroidetes phylum (463). Since the microbial composition found in faecal matter may differ significantly from mucosa-associated microbiota (129, 542), it is remarkable that the same trend of *Bacteroides* spp. relative abundance in relation to IBS is observed. A species-specific quantification of *Bacteroides fragilis* using qPCR assay showed that this species was reduced in faeces of patients belonging to all three IBS subgroups and especially alternating type IBS. However, this difference was not found to be significant ($p=0.269$), although the observed trend was in contrast to the uniform levels of Bacteroidetes ($p=0.934$) (299). Despite all previously discussed obstacles for the direct comparison of findings of different studies, the quantification of *Bacteroides* spp. in IBS always showed a trend of decrease in comparison to healthy subjects, which supports our finding. Interestingly, decreased relative abundance of Bacteroidetes have recently been found to be related to obesity (275). Obesity, similarly to the IBS, has higher incidence in the developed world (62) and the commonly observed negative correlation in levels of some members of the Bacteroidetes phylum indirectly suggests an effect of the modern life style on microbes of this group or some of its subgroups. Bacteroides are among the earliest and most typical colonisers of the human gastrointestinal tract (261, 364) and although their function has not yet been studied in great detail, some beneficial functions of *Bacteroides* spp. have already been reported (210, 387). Our findings call for further functional analysis of this group in relation to IBS.

Remarkably, the use of the phylogentic microarray allowed detecting increased levels of several members of the Bacilli in the microbiota of IBS patients. Although this group has not been previously studied in relation to IBS up to now, increased levels of aerobes have been associated with IBS (34, 308). No phylogenetic information in relation to the increased aerobes was provided, but since all Bacilli groups that were found to be significantly increased in this study (*Bacillus*, *Aneurinibacillus*, *and Streptococcus*) are aerobes, it could be speculated that the previously observed increase in aerobe counts was caused by a significant increase of genera detected in this study. Furthermore, *Streptococcus* spp. are among the dominant groups in the upper gastrointestinal tract (391). Bacterial overgrowth in the small bowel has been associated with IBS (373). Our finding of high levels of groups typical for the upper gastrointestinal tract in the faecal samples implies the overall disturbance of the gastrointestinal microbiota in IBS patients. Another explanation could be that a changed redox state in the gastrointestinal tract of, especially diarrhoea-predominant, IBS patients stimulates growth of facultative anaerobes.

We have revealed differences in the relative abundance of several subgroups of the *Clostridium* cluster XIVa between healthy subjects and IBS patients of each type, although

this group, as whole, was not significantly disturbed in any of the IBS subgroups. While a significant decrease of the entire *Clostridium* cluster XIVa was observed in all subgroups of IBS patients (299), another set of subjects analysed using a similar qPCR approach showed that only the constipation type IBS patients had significantly lower levels of this group (309). Moreover, FISH analysis of mucosa-associated bacteria showed a completely opposite trend as relative abundance of members of *Clostridium* cluster XIVa was higher in IBS patients (463). In our study, only one group of this cluster, namely *Anaerostipes caccae*-like organisms, was close to being significantly reduced in the constipation type IBS ($p=0.065$) when compared to healthy controls.

Increased relative abundance of several subgroups of the *Clostridium* cluster XIVa could be assigned to diarrhoea and alternating type IBS. Most of these groups, although referred as *e.g. Ruminococcus gnavus*-like organisms, are primarily represented by uncultivated species, and therefore very little is known about their functionality. More is known about physiology of recently described genus *Roseburia* (125, 126), which was found to be significantly increased in diarrhoea type IBS patients. *Roseburia* spp. comprise an abundant group within the gastrointestinal microbiota, and belong to one of the two phylogenetic groups that are responsible for the butyrate production in the human intestine (205). Butyrate has been claimed to be one of the most relevant microbiota's metabolites, as it is used as energy source by colonic cells and supposedly has an impact on health (68, 99). However, presently the beneficial role of butyrate is being reconsidered (422), and *in vivo* studies have shown that exposal of colonic cells to butyrate induces a series of oxidative stress responses in colonic cells (H. Hamer, personal communication). Therefore, it appears that the adjective beneficial is appropriately applied only for the delicate balance between constituents of the human gastrointestinal microbiota and their metabolic activities, which seems to be disturbed in IBS patients.

Levels of another predominant bacterial group, *Clostridium* cluster IV, were found to be significantly higher in constipation type IBS when compared to diarrhoea type IBS. It has been reported that relative abundance of members of this group were, in average, 40% higher in faeces of constipation type IBS patients than in diarrhoea type IBS patients (309). However, as this was a time-course study, in another set of samples, collected from the same subjects after period of 6 months, the levels of *Clostridium* cluster IV in the two IBS subgroups equalised.

Similarly to the *Clostridium* cluster XIVa group, members of *Clostridium* cluster IV have been recognised as dominant in human intestinal samples only during the last decade (461). This group is primarily represented by uncultivated, thus functionally yet uncharacterised species. One of the groups detected to be increased in constipation type IBS, *Oscillospira guillermondii*-like organisms, was also found to be increased in elderly subjects suffering from functional constipation (unpublished data). This group of organisms, which

appears to be relevant for the constipation-predominant disorders, does not have a single cultivated organism as *Oscillospira guillermondii* is not-yet cultivated inhabitant of the ruminants and other herbivorous animals, which detection was enabled by microscopic counting of the distinct, very large cells of this organism (532). The abundance of this bacterium was shown to be dependent on diet, as it decreases with indoor feeding of animals (329). However, SSU rRNA-gene based, species-specific assays appeared to be unspecific for the morphologically large *Oscillospira* cells and suggested either detection of other related species, or that this organism has other morphological forms (296). This SSU rRNA gene assay failed to confirm the effect of dietary changes on this microbe. Similarly, the SSU rRNA gene-based HITChip detection of different levels of the *Oscillospira guillermondii*-like organisms does not necessarily have to be driven by similar forces as those that induce change in relative abundance of *Oscillospira guillermondii*. Nevertheless, the fact that IBS is seems to be influenced by environmental factors (184) might imply an important dietary effect on the microbiota.

To conclude, the first comprehensive analysis of the faecal microbiota composition in relation to IBS has shown that the microbiota of IBS patients is disturbed and syndrome-specific. Furthermore, alternating, diarrhoea, and constipation type IBS patients appeared to be distinctive by the relative abundance of bacteria that inhabit the gastrointestinal tract of subjects suffering from various types of this disorder. The majority of the human intestinal microbes, in general, and those, which are potentially relevant for IBS, have not yet been cultured. Hence, it is difficult to speculate what are the driving forces for the observed differences or if these are cause or consequence of IBS. Therefore, other studies that address the response of specific microbial groups, which include *Bacteroides* spp., members of the Bacilli order, and members of *Clostridium* clusters IV and XIVa in relation to IBS should be conducted. Finally, the results presented in this study should be complemented by the analysis of the microbiota of more patients from different geographic locations, preferably from different gastrointestinal sites, and in different time points. This would provide more accurate picture of the disturbances in IBS gastrointestinal microbiota, which would enable for rational selection and evaluation of a treatment of this globally increasing disorder.

CHAPTER 8

Mirjana Rajilić-Stojanović, Hans H. G. J. Heilig,

Francisco Guarner, Fergus Shanahan and Willem M. de Vos

Manuscript in preparation

Diversity of the Human Gastrointestinal Microbiota in Relation to Ulcerative Colitis

Background: Gastrointestinal microbiota is considered to be an important aetiological factor of ulcerative colitis (UC), although the precise description of the disturbances of the microbiota in relation to UC is hampered by the complexity of this ecosystem and large inter-individual variations.

Aims: To compare and contrast the microbiota of UC patients in relapse and remission with that of healthy subjects using a global and comprehensive approach that is directed at the detection and relative quantification of gastrointestinal microbes.

Methods: The microbiota of faecal samples of 12 UC patients, collected at relapse and during remission, was compared to the faecal microbiota of 12 healthy subjects using a novel phylogenetic microarray that allows for simultaneous qualitative and semi-quantitative analysis of over 1,000 gastrointestinal phylotypes. Other small subunit ribosomal RNA based techniques were used to define the diversity of intestinal sulphate reducing bacteria (SRB) in relation to UC.

Results: The composition of the faecal microbiota of UC patients was found to be distinct from that of healthy subjects ($p=0.002$), and was characterised by reduced diversity of members of the Firmicutes phylum, both in relapse and remission (p values 0.012 and <0.001, respectively). In faecal samples collected during relapse *Peptostreptococcus* spp. (in five samples) and *Clostridium difficile* (six samples) could be identified. *Desulfovibrio piger* and *Bilophila wadsworthia* and other three uncultivated Desulfovibrionales species were identified in UC and healthy subjects, and no association between SRB and health status could be found.

Conclusions: Comprehensive analysis showed numerous deviations of the UC microbiota composition, which have indicated a multi-factorial character of UC.

Introduction

Ulcerative colitis (UC) is an acute and chronic inflammatory disease of the large bowel, which is characterised by an aggressive immune response to a fraction of commensal intestinal microbes (416). The aetiology of UC is unknown and complex, as twin studies showed that both environmental and genetic factors are relevant for its development (187). Moreover,studies with animal models indicated that intestinal microbes are an essential factor contributing to the disease (188). Hence, numerous studies have addressed the disturbances of the intestinal microbiota composition of UC patients when compared to that of healthy subjects. However, the reported results are not directly comparable and often not consistent, as the microbial diversity can be assessed by the analysis of mucosa-associated or faecal microbes, in samples from patients in relapse or remission, and by the application of different techniques, including cultivation-based or molecular approaches. Various recent studies addressed the microbiota deviation in relation to UC using molecular techniques and reported the reduction of the relative abundance of members of *Clostridium cluster* XIVa, both in mucosal and faecal samples of UC patients (443, 463), decreased *Bifidobacterium* and increased *Peptostreptococcus* spp. numbers (293), but also high incidence of *Bacteroides fragilis* (463, 503) and *Clostridium difficile* during relapse (343).

The perturbations in the intestinal microbiota composition are reflected in the deviation of the microbiota's metabolic activity (303). One of these deviations concerns the significant overproduction of hydrogen-sulphide (H_2S) (273, 374). H_2S is highly toxic for humans and it can induce a number of colitis-like symptoms in the intestinal epithelium. It has been shown that the exposure of the rat colon to sulphide increases mucosal apoptosis, goblet cell depletion and superficial ulceration (19, 107). Furthermore, interaction studies with intestinal cell lines have shown that H_2S-mediated inhibition of the respiratory activity of mitochondria triggers a series of transcription factors that up-regulate genes involved in the proliferation and apoptosis of cells, which might result in hyper-proliferation, a phenotype common to UC (107). Finally, it has been shown that H_2S is a genotoxic agent that causes damage of genomic DNA, even at concentrations that level those measured in the human colon (20), and that it acts as a neuromodulator that affects gut signalling (423).

The production of H_2S by the intestinal contents is traditionally associated with microbes that reduce sulphate into sulphide, the so-called sulphate reducing bacteria (SRB), which includes various groups of which *Desulfovibrio* and *Desulfobacter* are the typical inhabitants of the human gastrointestinal tract (170, 375). However, it is likely that other pathways of H_2S production are relevant for its accumulation in UC patients. It has been shown that sulphate supplementation of drinking water does not induce microbial changes in the mouse intestine, and only slightly affects the production of H_2S (106). On the contrary, an

increased protein intake is directly correlated with the H_2S production in human faeces (297). The analysis of oral H_2S producers showed that a number of bacterial species formed significant amounts of H_2S from L-cysteine. The highest activity was found for *Peptostreptococcus* spp., *Eubacterium* spp., *Selenomonas* spp., *Centipeda* spp., *Bacteroides* spp. and *Fusobacterium* spp. (366). Since some of these groups are abundant members of the microbiota in the lower gastrointestinal tract, it is possible that the production of H_2S in UC patients is carried by some of these or other, not yet identified microbes.

To further advance the insight in the gastrointestinal microbiota in relation to UC, a global and comprehensive approach that is directed at the detection and quantification of as many as possible intestinal microbes should be used. For that purpose, we have employed a newly developed phylogenetic microarray that targets over 1,000 gastrointestinal phylotypes – the HITChip (**Chater 4**) – for the analysis of the faecal microbiota composition in UC patients during remission and relapse. An additional effort was put into defining the diversity of intestinal SRB, as it was previously suggested that species-level specificities of this group might also be relevant for UC (535). The results of this analysis indicate the multi-factorial nature of UC, which is displayed by the disturbance of both diversity and relative abundance of a number of microbial groups.

Materials and Methods

Faecal samples

Twelve patients with UC were recruited at two sites of the European project PROGID (QLK1-2000-00563) including University College Cork, Ireland; and Hospital General Vall d'Hebron, Barcelona, Spain. All patients were in clinical and endoscopic remission (score <4) with regular episodes of relapse despite steroid treatment. Patients were screened for presence of *Clostridium difficile* glycosylating exotoxins A and B in the stool and they were excluded from the study if they were toxin positive. They were all younger than 75 years, and were on 5-aminosalicylic acid treatment. Faecal samples were collected during relapse of disease and from the same patients during remission, not less than four weeks following completion of steroid-induction treatment. Faecal samples from twelve healthy subjects originating from the same geographic localisations were used for comparative analysis. Subsequent to the delivery, faecal samples were frozen at -20^{o}C, transported on dry ice and stored at -20^{o}C until further processing.

Faecal DNA extraction

Total DNA was extracted from faecal material using a modified protocol of the QiaAmp DNA Mini Stool Kit (Qiagen, Hilden, Germany) (539). DNA yield was quantified using

NanoDrop ND-1000 spectrophotometer (NanoDrop® Technologies, Wilmington, DE). DNA concentration was adjusted to 10 ng/μl and was used as a template in PCR amplification.

The HITChip analysis

PCR

The SSU rRNA gene amplified from faecal DNA using the *T7prom*-Bact-27-for (5'-TGA ATT GTA ATA C GA CTC ACT ATA GGG GTT TGA TCC TGG CTC AG–3') and Uni-1492-rev (5'-CGG CTA CCT TGT TAC GAC-3'), which ensured the introduction of a T7 promoter sequence at the 5' terminus of the rRNA gene amplicon. PCR reactions were carried out in a final volume of 50 μl, and 10 ng of DNA samples were used as template. Samples were initially denatured at 94˚C for 2 min followed by 35 cycles of 94˚C (30 sec), 52˚C (40 sec), 72˚C (90 sec) and a final extension at 72˚C for 7 min. The PCR products were purified by using the DNA Clean and Concentrator kit (Zymo Research, Orange, USA), according to the manufacturer's instruction. Final DNA concentration was determined by as described above.

RNA production and labelling

In vitro transcription of T7-promoter SSU rRNA gene amplicon was performed according to the manufacturer's protocol using the Riboprobe System (Promega, La Jolla, USA), 500 ng of the T7-16S rRNA gene amplicon, and, besides rATP, rGTP, rCTP, a 1:1 mix of rUTP and aminoallyl-rUTP (Ambion Inc., Austin, Tx, USA). The transcription reaction was performed at room temperature for 2 h., the template DNA was digested applying the Qiagen RNAse-free DNAse kit (Qiagen, Hilden, Germany), and RNA was purified using the RNeasy Mini-Elute Kit (Qiagen, Hilden, Germany). RNA yield was measured as described above.

Amino-allyl-modified nucleotides were coupled with CyDye using the Post-Labeling Reactive Dye (Amersham Biosciences, Little Chalfont, UK), which was previously dissolved in 84 μl dimethyl sulfoxide. Labelling reaction was performed in a 25 mM sodium bicarbonate buffer (pH 8.7) by adding 20 μl of dissolved CyDye to 2 μg of purified RNA in a final volume of 40 μl. Samples were incubated for 90 min in the dark at room temperature. The reaction was stopped by adding 15 μl of 4M hydroxyl-amine and incubating for 15 min in the dark. RNAse-free water was added to 100 μl and labeled RNA was purified and quantified as described above.

Microarray production, hybridisation and data extraction

The HITChip microarrays were produced by Agilent Technologies (Agilent Technologies, Palo Alto, CA., USA). The oligonucleotide probes were extended at the 3'-end (at the array support side) with 10nt long T-spacer, and were printed on the array using *in situ* surface-attached oligonucleotide probe synthesis (54). Arrays used in this study were of the 2×11K

format, with two arrays per glass slide. Each array was hybridised with two samples, labelled with Cy3 and Cy5. Combined Cy3- and Cy5-labelled target mixtures were fragmented by adding 1 µl of Ambion 10× fragmentation reagent (Ambion Inc., Austin, Tx, USA), and incubation for 20 min at 70˚C, according to the manufacturer's instruction. Fragmentation was stopped by adding 1 µl of Ambion stop solution. Hybridization mix was prepared by adding to the RNA mixture 31.5 µl of 20× SSC (412), 6.3 µl of 10% SDS (412), 20 µl of Agilent Control Target mix and RNAse-free water to a total volume of 210 µl. The microarray, hybridisation-mixture and cover-slide were joined and placed in the hybridisation chamber according to the manufacturers' instructions. The hybridisation was carried out at 62.5^{o}C in a rotation oven (Agilent, Palo Alto, Ca., USA) speed 4, for 16 h. Slides were washed at room temperature in a 50 ml Greiner tube containing 50 ml 2× SSC, 0.3% SDS for 10 min and at 38˚C in 0.1× SSC, 0.3% SDS for 10 min. SDS was completely removed by washing the slides in 0.06× SSPE (412) for 5 min, followed by a quick dry with compressed nitrogen. All wash-steps were performed with the aid of a Stuart Rotator SB3 (Barloworld Scientific, Stone, UK), set at speed 40rpm.

Data were extracted from microarray images using the Agilent Feature Extraction software, version 7.5 (www.agilent.com). Data normalisation was performed using a set of R based scripts (http://www.r-project.org/) as previously described (**Chapter 4**). Further microarray analysis was performed in a custom designed relational database which runs under the MySQL database management system (http://www.mysql.com/) using a series of custom made R scripts as previously described (**Chapter 4**).

PCR amplifications

Sequences of the primers used in this study, their specificity and the annealing temperature are listed in the Table 1. For the groups of SRB with only one specific primer, a universal bacterial primer was paired with the group-specific primer. PCR was performed using GOTaq PCR amplification kit (Promega, Madison, USA). Amplification specificity with degenerated primers was optimized by adding Bovine Serum Albumin (BSA)(Amersham Lifesciences, Amersham, UK) to a final concentration of 0.02%. PCR included following steps: initial denaturing at 95^{o}C for 3 min, 30 cycles of denaturing at 95^{o}C for 30 sec, annealing for 30 sec at varying temperatures (Table 1), and elongation at 72^{o}C for 40 sec, and finally an elongation step at 72^{o}C for 5 min. PCR amplifications were performed in Whatman Biometra T1 (Whatman Biometra, Gőttingen,Germany).

Denaturant gradient gel electrophoresis

The diversity of the selected SRB groups was studied by denaturing gradient gel electrophoresis (DGGE). Part of the SSU rRNA gene coding DNA, corresponding to the variable regions V3-V5 was amplified in a nested PCR performed on 1:100 dilutions of the SRB group specific PCR amplicons. Primers 5'-CCT ACG GGA GGC AGC AG-3'

containing 40-bp long GC tail on the 5' end, and 5'-ATT ACC GCG GCT GCT GG-3' were used to generate PCR products suitable for DGGE separation (342). A gradient of denaturing compounds ranging from 35% to 55% (100% denaturing acrylamide is defined as 7 M urea and 40% formamide) was used and DGGE was performed as previously described (114). DGGE gels were scanned and analysed by using Bionumerics software (http://www.applied-maths.com).

Table 1 SSU rRNA gene targeting primers used in this study. Specificity, sequence of the primer, annealing temperature used in PCR assay and the reference are listed.

Specificity	Sequence 5' to 3'	Tm, °C	Reference
most Bacteria	AGA GTT TGA TCC TGG CTC AG	67	(240)
Thermodesulfovibrio spp.	AGC GAT TCC GGG TTC AC	67	(282)
Thermodesulfobacteriaceae	ATT CAC GGC GGC ATG C	67	(282)
Desulfobulbus spp.	CGC GTA GAT AAC CTG TCY TCA TG GTA GKA CGT GTG TAG CCC TGG TC	66	(101)
Desulfomonile spp.	CGG GGT RTG GAG TAA AGT GG CGA CTT CTG GTG CAG TCA RC	62	(282)
Desulfobacterium spp.	CTA ATR CCG GAT RAA GTC AG ATT CTC ARG ATG TCA AGT CTG	64	(101)
Desulfobacter spp.	GAT AAT CTG CCT TCA AGC CTG G CYY YYY GCR RAG TCG STG CCC T	60	(101)
Desulfococcus–Desulfonema–Desulfosarcina	GAT CAG CCA CAC TGG RAC TGA CA GGG GCA GTA TCT TYA GAG TYC	65	(101)
Desulfovibrionales and Desulfuromonadales	GRG YCY GCG TYY CAT TAG C SYC CGR CAY CTA GYR TYC ATC	61	(101)
Bilophila wadsworthia	GAA TAT TGC GCA ATG GC CGT GTG AAT AAT GCG AGG G	60	(313)
Syntrophobacteraceae + *Syntrophobacteraceae*	ACG GGT AGC TGG TCT GAG GCC CAC GCA CTT CTG GTA	60	(282)
Desulfotomaculum spp.	GTA ACG CGT GGA TAA CCT CCT TCC TCC GTT TTG TCA	63	(460)
Desulfobacca acetoxidans	TAC GAG AAA GCC CGG CTT TTA GGC CAG CGA CAT CTG	58	(282)
Desulfosporosinus spp.	CGA TTA TGG ATG GAC CCG	63	(282)
Archaeoglobus spp.	CTA TCC GGC TGG GAC TA	52	(282)
most Bacteria and Archaea	GGY TAC CTT GTT ACG ACT T	52	(240)

Cloning and sequencing

Four clone libraries were made from Desulfovibrionales / Desulfuromonadales-targeted PCR products for healthy subjects and UC patients from both geographic locations. The clone libraries were made in the following manner; equal amounts of PCR products obtained in Desulfovibrionales / Desulfuromonadales-specific PCR from samples belonging to the same group (healthy Barcelona, UC Barcelona, healthy Cork or UC Cork) were joined. Joined PCR products were purified by the use of Qiaquick PCR purification kit (Qiagen, Hilden, Germany). Purified fragmnents were cloned into *E. coli* JM109 using the Promega pGEM-T vector system as previously described (537). Insert size was checked by using pGEM-T specific primers T7 (5' AAT ACG ACT CAC TAT AGG) and SP6 (5' ATT TAG GTG ACA CTA TAG) according to the manufacturer´s protocol. The diversity among clones, with the correct insert size, was checked by restriction fragment length polymorphism (RFLP)analysis with a mix of *Msp* I, *Cfo* I and *Alu* I enzymes. Representative clones of each restriction pattern were selected for sequencing. Plasmid inserts of the unique clones were amplified by the use of pGEM-T targeting primers M13-20 forward (5' TGT AAA ACG ACG GCC AGT) and M13 reverse (5' CAG GAA ACA GCT ATG ACC) and were sequenced by Greenomics Wageningen UR, Netherlands (www.greenomics.com). Sequencing results representing approximately 700 nucleotides long stretch of the SSU rRNA gene were compared to the sequences deposited in the NCBI DNA database (www.ncbi.nlm.nih.gov/BLAST).

Statistical analysis

The diversity of microbial community assessed by both DGGE and HITChip analysis was expressed as Simpson's reciprocal index of diversity (1/D) (440) that was calculated using the equation $\lambda = 1 / \Sigma$ Pi^2 where Pi is the proportion of i^{th} taxon. The proportion of each taxon was calculated as the height of individual peak in proportion to the sum of all peak heights in the densitometric curve for DGGE gels (162), and as proportion of each probe signal compared to the total signal for the HITChip analysis. Simpson's reciprocal index of diversity takes into account both the number of taxons present in a sample as well as their abundance in the community. A higher value of the Simpson's index corresponds to a more diverse community.

Principal component analysis (PCA) was performed using Canoco for Windows 4.5 (271). The microbiota of each sample was explained by a data matrix that consisted of an average signal intensity for 131 bacterial groups that are targeted by the HITChip and defined on the bases of the SSU rRNA sequence similarity (**Chapter 4**). Data matrices were organised in Microsoft Excel and converted into a Canoco-compatible format using the WCanoImp utility. PCA was performed focusing on inter-species correlation and diagrams were plotted by using CanoDraw.

Monte Carlo Permutation Procedure (MCPP) (212), was performed to access statistical significance for the variation in large datasets in relation to the health status. Alternatively, for datasets with two variables a Student's t-test was performed, and calculated p values were two tailed. P values lower than 0.05 were considered significant and were marked with an asterix in figures, while p values lower than 0.01 were considered as highly significant and were marked with a double asterix.

Results

Analysis of the total bacterial diversity in relation to UC

The faecal microbiota composition of 12 UC patients was assessed for both the active (relapse) and inactive (remission) period of the disease and was compared to that of 12 healthy subjects. The microbiota composition was described by the average hybridisation signal intensity of 131 SSU rRNA-based, phylogenetic groups targeted by the HITChip, which has shown to be semi-quantitative and comparable between subjects (**Chapter 4**). Projecting the faecal microbiota composition on the two principal components, which have the highest explanatory potential, showed a clear separation between UC patients and healthy subjects (Fig. 1). The separation was carried out by the specific grouping of the microbiota of healthy subjects, as the vast majority of those samples clustered within the left lower panel of the PCA plot. This analysis showed that the microbiota of healthy subjects, despite inter-individual variation, has a more uniform composition than that of UC patients. The health status could explain 12.1% of the variation in the dataset when the microbiota of relapse sample was analysed, and 15.1% when the remission sample was analysed. The composition of the faecal microbiota of UC patients in both analysed sets of samples had distinct composition when compared to the faecal microbiota composition of healthy subjects, and, as assessed by MCCP, this difference was found to be significant (p=0.002).

To define the groups that discriminated the microbiota of UC patients and healthy subjects, a Student's T test was performed on the average hybridisation signals of each of 131 phylogenetic groups. The microbiota of UC patients in both relapse and remission was compared to that of healthy volunteers. A common pattern of difference was observed for 17 phylogenetic groups (Table 2), while additional 15 phylogenetic groups appeared to be significantly different between the microbiota of healthy and UC patients in either relapse or remission (Table 3). The vast majority of the observed significant differences was caused by the decrease of several members of Firmicutes phylum.

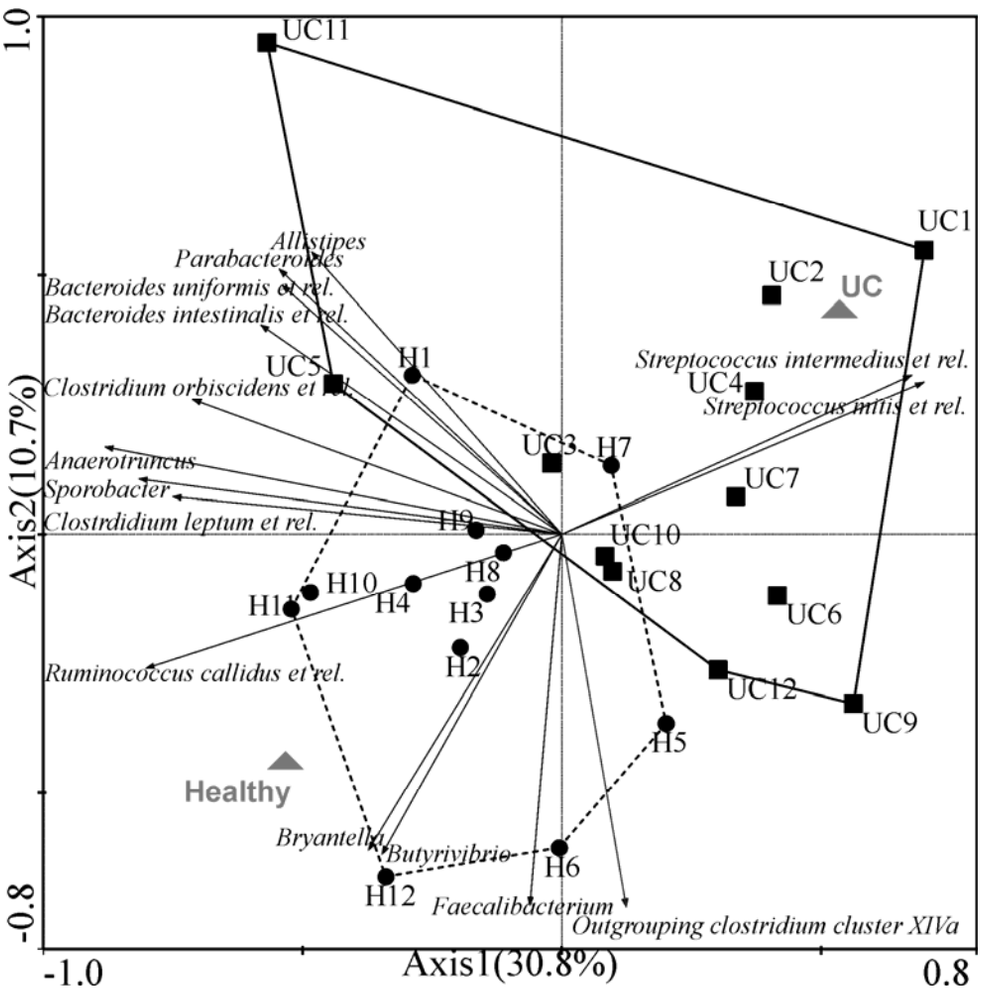


Figure 1 PCA triplot of the microbiota composition expressed by average hybridisation signals for 131 phylogenetic groups for 12 UC patients (■) and 12 healthy subjects (●). Microbial groups that contributed at least 40% to the explanatory axis used in the plot are presented as vectors, and health status (UC or healthy) are centroides of the plot. Percentage values at the axes indicate contribution of the principle components to the explanation of total variance in the dataset. UC patients are encoded by UC1-UC12, while healthy control subjects are denominated by H1-H12. MCPP revealed that health status correlated to the variation in microbiota composition with a p-value of 0.002.

Only three groups belonging to the Firmicutes phylum were significantly increased as compared to the healthy individuals, but only in relapse microbiota of the UC patients (Table 3). Detailed, phylotype-level, analysis of the diversity of these groups showed that the presence of *Peptostreptococcus anaerobius* and/or *Peptostreptococcus micros* in faecal samples of five UC patients but not at all in the healthy volunteers was supporting the significant increase of *Peptostreptococcus micros*-like organisms. The significantly different level of *Clostridium difficile*-like organisms could be explained by presence of *C. difficile* in faeces of six out of twelve UC patients and none of healthy controls, but also by a significant increase in abundance of two uncultured phylotypes related to *C. difficile*. Both phylotypes, uncultured bacterium clone LCLC21 (AF499843) and uncultured bacterium clone M364

(AY916153), were previously detected in the analysis of mucosa-associated microbiota of healthy subjects (129, 505). Finally, the detailed analysis of *Eubacterium limosum*-like organisms showed that none of the targeted phylotypes could be identified in faecal samples of UC patients or healthy subjects. Nevertheless, the hybridisation patterns obtained for the faecal samples of some of the UC patients suggested the presence of some, not yet identified *Eubacterium limosum*-like phylotypes, the abundance of which is increased in these subjects.

Table 2 SSU rRNA-based groups for which the relative abundance was found to be significantly different between healthy subjects and UC patients, during both relapse and remission. In addition to the group name, its physiological position is defined by the higher taxonomic group, the trend of the change observed for UC patients, and p value is provided.

Higher taxonomic group	Group	Change	P value
Actinobacteria	*Bifidobacterium*	Increase	0.037
Bacilli	*Aneurinibacillus*	Decrease	0.002
Clostridium cluster IV	*Anaerotruncus colihominis*-like	Decrease	0.004
	Eubacterium siraeum-like	Decrease	0.003
	Clostridium cellulose-like	Decrease	0.040
	Clostridium leptum-like.	Decrease	0.025
	Clostridium orbiscindens-like	Decrease	0.016
	Oscillospira guillermondii-like.	Decrease	0.029
	Ruminococcus callidus-like	Decrease	0.004
	Sporobacter termitidis-like.	Decrease	0.006
Clostridium cluster XIVa	*Bryantella formatexigens*-like	Decrease	0.023
	Butyrivibrio crossotus-like	Decrease	0.027
	Lachnobacillus bovis-like	Decrease	0.044
	Lachnospira pectinoschiza-like	Decrease	0.014
	Roseburia intestinalis-like	Decrease	0.016
Uncultured Firmicutes	Uncultured Clostridiales I	Decrease	0.002
	Uncultured Clostridiales II	Decrease	0.009

The microbiota of UC patients appeared to be less diverse, in both relapse and remission, when compared to the healthy subjects (Fig. 2). These differences were found to be significant as obtained p values were 0.012 and <0.001 for relapse and remission faecal samples, respectively. The diversity of the faecal microbiota of UC patients in remission was notably lower than in relapse and this difference reached an almost significant level (p=0.062).

Table 3 SSU rRNA-based groups for which the relative abundance was found to be significantly different between healthy subjects and UC patients, during either relapse (gray filled cells) or remission (empty cells). In addition to the group name, its physiological position is defined by the higher taxonomic group, the trend of the change observed for UC patients, and p value is provided.

Higher taxonomic group	Group	Change	P value
Clostridium cluster XI	*Clostridium difficile*-like.	Increase	0.016
Clostridium cluster XIII	*Peptostreptococcus micros*-like	Increase	0.044
Clostridium cluster XV	*Eubacterium limosum*-like	Increase	0.031
Bacteroidetes	*Prevotella oralis*-like	Decrease	0.039
	Prevotella tannerae-like	Decrease	0.040
Clostridium cluster IV	*Faecalibacterium prausnitzii*-like	Decrease	0.024
	Ruminococcus bromii-like	Decrease	0.021
	Subdoligranulum variable-like	Decrease	0.045
Clostridium cluster XIVa	*Clostridium sphenoides*-like	Decrease	0.008
	Clostridium symbiosum-like	Decrease	0.016
	Coprococcus eutactus-like	Decrease	0.029
	Eubacterium rectale-like	Decrease	0.042
	Outgrouping uncultured bacteria	Decrease	0.040
Clostridium cluster XVIII	*Coprobacillus catenaformis*-like	Decrease	0.047
Uncultured Firmicutes	Unculutured Mollicutes	Decrease	0.044

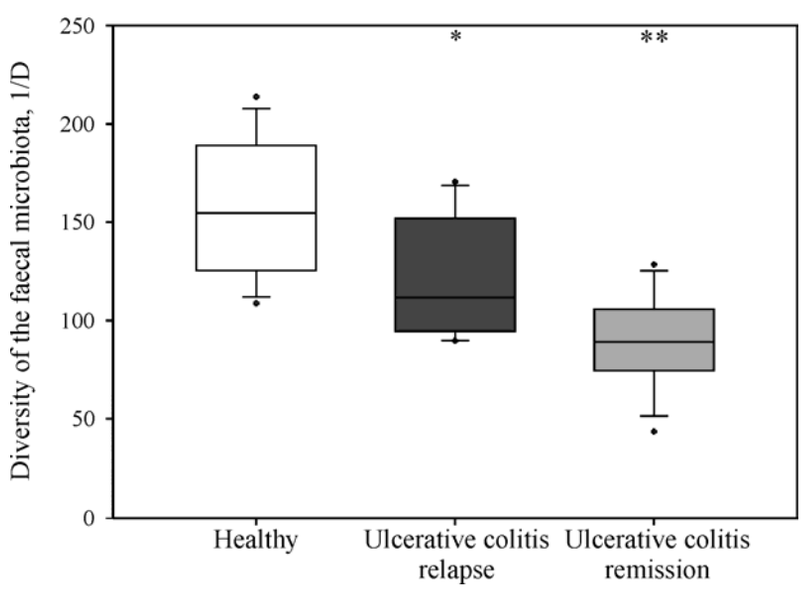


Figure 2 The diversity of the total microbiota expressed as Simpson index of the hybridisation profiles on the HITChip, represented as box and whiskers plots. The box extends from 25th percentile to 75th percentile, with a line at the median. The whiskers are showing the highest and the lowest value after exclusion of outliers, which are presented by dots. Double asterix indicates highly significant difference in similarity when compared to unrelated individuals ($p<0.01$), asterix indicates significant difference ($P<0.05$).

This reduction of the total bacterial diversity could be mainly attributed to differences in the members of the Firmicutes phylum (Fig. 3). The reduced diversity of different Firmicutes in relation to UC appeared to be the major factor controlling its low abundance. *Clostridium* cluster IV (also known as *Clostridium leptum* group) was significantly less diverse in both relapse ($p=0.003$) and remission ($p<0.001$). In addition, *Clostridium* cluster XIVa (also known as *Clostridium coccoides* group) and uncultured Firmicutes were less diverse in faeces of UC patients, collected during remission, when compared to healthy controls (p values 0.002 and 0.004, respectively). Even though the diversity of the microbiota has notably decreased between relapse and remission, this decrease was not significant in any of the analysed phylogenetic groups.

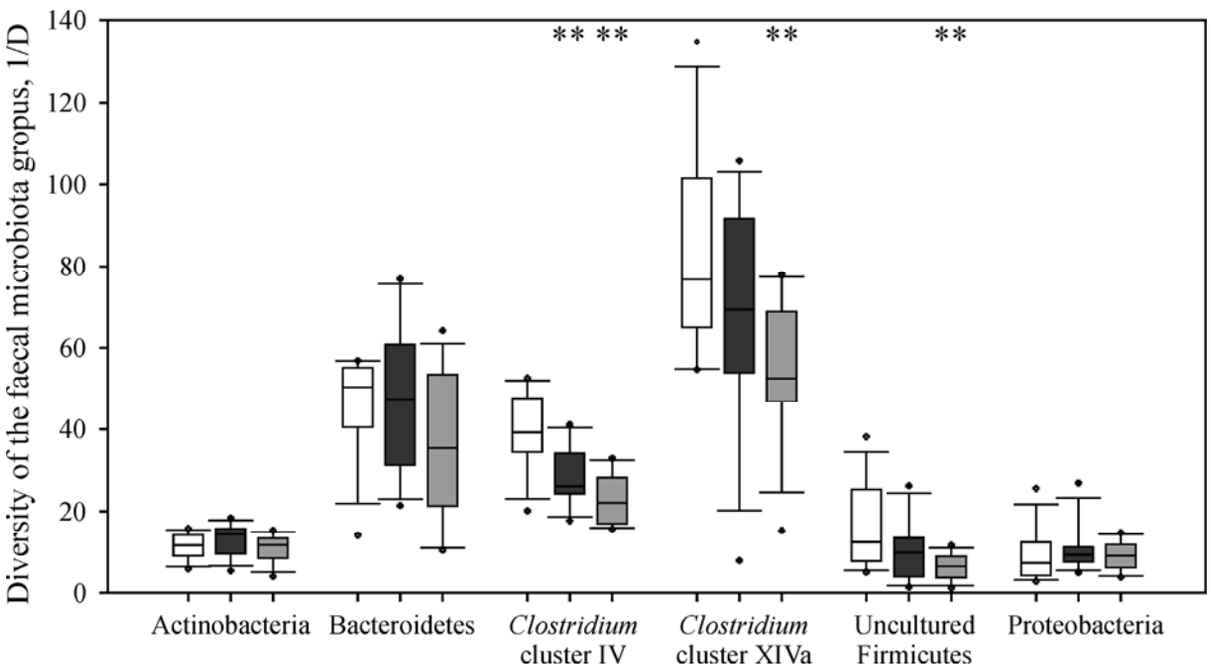


Figure 3 The diversity of three phylum level groups of the faecal microbiota and three subgroups of phylum Firmicutes expressed as Simpson's index and represented as box and whiskers plots (see Fig. 2). The results obtained on samples fromhealthy subjects, UC patients in relapse and in remission are represented by empty, dark filled, and light filled box, respectively. Double asterix indicates highly significant difference in similarity when compared to unrelated individuals ($p<0.01$), asterix indicates significant difference ($P<0.05$)

Analysis of SRB diversity in relation to UC

Thirteen groups of SRB, that fully cover the diversity of this group, were targeted by a series of specific PCR amplifications. Amplicons were obtained from all analysed samples when members of the orders Desulfovibrionales and Desulfuromonadales were targeted. *Desulfobacter*-targeting PCR gave product in 23 out of 24 analysed samples, while a PCR

product of appropriate size was obtained from seven samples in the *Desulfotomaculum*-specific PCR, on five samples in *Desulfosporosinus*- specific PCR, and on three samples for Syntrophobacteraceae-specific PCR. Other SRB-targeted PCRs failed on all analysed samples. No correlation with health status and success of the PCR was observed. A species-specific PCR assay that targets *Bilophila wadsworthia* enabled the detection of this microbe in 11 healthy subjects and in 8 UC subjects. To examine the diversity of the widely distributed SRB groups, DGGE was applied on the PCR products obtained with specific primers targeting Desulfovibrionales / Desulfuromonadales, and the genus *Desulfobacter*.. Profiles obtained with PCR amplicons targeting Desulfovibrionales / Desulfuromonadales, appeared to be more diverse in some of UC patients, which allowed for the formation of a UC-specific cluster of DGGE profiles (Fig. 4). The diversity index was calculated for the obtained DGGE patterns, and the mean value was 1.86 (SD 0.69) and 6.92 (SD 4.67) for healthy subjects and UC patients, respectively, which was found to be significantly different (p=0.001).

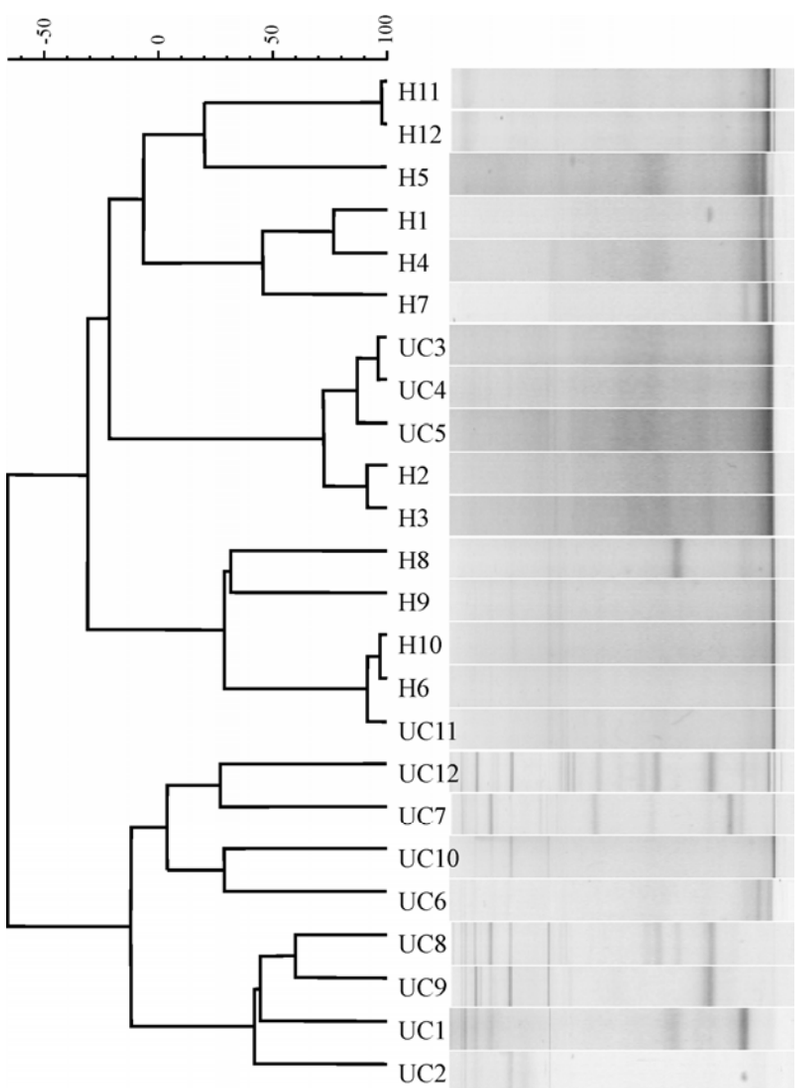


Figure 4 Ward cluster based on Jaccard distance between the DGGE profiles obtained with Desulfovibrionales / Desulfuromonadales – targeting primers on faecal DNA of 12 UC patients and 12 corresponding controls. UC patients are encoded by UC1-UC12, and healthy control subjects are denominated by H1-H12.

The Desulfovibrionales / Desulfuromonadales-targeting PCR products were subject to DNA sequence analysis, resulting in 30 clones with a distinct sequence, defined as less than 98% sequence similarity to other clones within the same clone library.

Table 4 Clones of SSU rRNA gene with the percentage of similarity to the closest cultivated relative recovered and their phylogenetic position.

Clone codes	Closest cultured relative	Phylum/SSU rRNA group
adhufecHB3, adhufecHC4, ucfecDB11, ucfecDC11	99% *Desulfovibrio piger*	Delta-Proteobacteria
adhufecHB1, adhufecHC5, ucfecDB4, ucfecDC4	99% *Bilophila wadsworthia*	
adhufecHB2	97% *Bilophila wadsworthia*	
adhufecHC2	91% *Desulfovibrio intestinalis*	
UcfecDB10	94% *Desulfovibrio alcholovorans*	
UcfecDB12	93% *Bilophila wadsworthia*	
UcfecDC3	95% *Bilophila wadsworthia*	
UcfecDC5	92% *Bacteroides capillosus*	*Clostridium* cluster IV
UcfecDC10	92% *Bacteroides capillosus*	
UcfecDB8	96% *Clostridium barlettii*	*Clostridium* cluster XI
UcfecDB9	99% *Dorea longicatena*	*Clostridium* cluster XIVa
UcfecDC6	98% *Eubacterium hallii*	
UcfecDC9	99% *Ruminococcus gnavus*	
UcfecDB3	98% *Ruminococcus obeum*	
UcfecDB6	95% *Ruminococcus shinkii*	
UcfecDB7	95% *Ruminococcus obeum*	
UcfecDB13	92% *Ruminococcus hansenii*	
UcfecDC2	95% *Ruminococcus obeum*	
UcfecDB5	99% *Clostridium ramosum*	*Clostridium* cluster XVIII
UcfecDB2	98% *Streptococcus pseudopneumoniae*	Bacilli
adhufecHC3, ucfecDB1, ucfecDC8	99% TM7 phylum oral clone DR034	TM7 phylum
adhufecHC1, ucfecDC7	99% oral clone BS003	

Three of these distinct sequences originated from the healthy Barcelona clone library, five from the healthy Cork, twelve from the UC Barcelona and ten from the UC Cork clone library. However, only 13 sequences (43%) could be identified as members of the Desulfovibrionales order, while the other amplified sequences clustered within Firmicutes and the TM7 phylum candidate of uncultured bacteria (Table 4). The specifically amplified sequences corresponded to two already established human intestinal inhabitants, *Desulfovibrio piger* and *Bilophila wadsworthia*, and five uncultured members of the Desulfovibrionales order. The unspecific amplification almost selectively occurred in UC samples (15 out of 17 distinct clone sequences), suggesting that the presence of other

phylotypes than those belonging to Desulfovibrionales order, caused the observed differences in the DGGE profiles between UC and healthy subjects.

Discussion

In this study a range of SSU rRNA-based techniques was applied to reveal the relation between UC and the diversity of faecal bacteria. The composition of the faecal microbiota of UC patients, both in relapse and remission, was found to be significantly different from that of healthy volunteers ($p=0.002$). This difference was supported by changes in more than 30 microbial groups targeted by the HITChip. Most of the discriminative groups clustered within the Firmicutes phylum, and were affected by a significant decrease in both diversity and relative abundance. Firmicutes are one of the most abundant and diverse groups of the human gastrointestinal microbiota, which principally groups strict anaerobes, and includes *Ruminococcus* spp. *Coprococcus* spp., *Butirivibrio* spp. and numerous *Clostridium* and *Eubacterium* species. The decreased proportion of anaerobes has already been associated with UC (293), and probably corresponds to the decrease of Firmicutes detected in our study.

Several groups of *Clostridium* cluster XIVa group have displayed reduced relative abundance in faecal samples of UC patients, while the diversity of this group was significantly reduced only in the set of samples that were collected during remission. Previously, it has been shown that the abundance of members of this group is reduced in UC patients (443, 463), while in another study such a reduction was only observed after three month period of remission (38). The vast majority of organisms that belong to *Clostridium* cluster XIVa has not yet been cultivated and the knowledge on their physiology is very limited (**Chapter 3**). More is known about groups that have cultured representatives such as members of *Roseburia* genus (125, 126). *Roseburia* spp. are, in addition to *Feaealibacterium prausnitzii*–like organisms, the major butyrate producers in the human intestine (205). Both groups of butyrate-producers were found to be reduced in UC patients, although the reduction of *Feacalibacterium prausnitzii*–like organisms was detected only in faecal samples collected during remission. It has been suggested that impaired butyrate oxidation is linked to UC (396), although *in vivo* measurements did not confirm this hypothesis (441). However, it was recently shown that the intestinal contents of UC patients produce reduced levels of butyrate when compared to healthy volunteers (303). The decrease in the relative abundance of the major butyrate producers in faecal samples of UC patients, reported in this study, could explain this observed reduction in butyrate production. Butyrate is thought to be one the most relevant microbial metabolites as it provides energy for intestinal epithelial cells and promotes their growth (99) although the beneficial role of butyrate on the human colonic epithelium is being reconsidered (422). Moreover, it has been reported that the active butyrate producer *Fusobacterium varium*, isolated from an UC patient, can induce colonic ulcers and

inflammatory infiltration in mouse models (355). Furthermore, the analysis of the mucosa-associated bacteria of a 12-year old UC patient showed an increased incidence of a *Feacalibacterium prausnitzii*–like butyrate producing bacterium (503). Finally, in another study we have shown that increased levels of *Roseburia* spp. characterise the intestinal contents of irritable bowel syndrome (**Chapter 7**), that are, similar to UC patients in remission, characterised by low-grade inflammation and functional disturbances such as diarrhoea and/or constipation (188). These findings call for a more detailed analysis of the butyrate impact on intestinal epithelial cells and more specifically in relation to UC.

A significantly higher proportion of members of *Bifidobacterium* genus were detected in faecal samples of UC patients than in that of healthy controls. Previous studies, showed that the proportion of Bifidobacteria in the total microbiota of UC patients is 9.7% (38), while the enumeration of this group in healthy adults using the same technique has reached a remarkably lower value of 4.4% (268). In a recent study, the relative proportion of bifidobacteria was also found to be two-fold higher in UC patients when compared to healthy subjects although its significance was not addressed (443). Remarkably, the relative abundance of *Bifidobacterium animalis* was found to increase in colitis mice with the development of colitis (51). In contrast to these findings, cultivation of mucosa-associated bacteria in UC patients and healthy subjects showed that the bifidobacterial counts are significantly reduced in UC (293). However, it has been reported that the mucosa-associated microbiota of UC and other patients may differ from that recovered from faeces, which could explain the difference in observations (542). Furthermore, molecular-based enumeration showed that the abundance of faecal bacteria in UC patients is reduced from 1.16×10^{11} cells per gram, characteristic for healthy subjects, to 3.51×10^{10} cells per gram (38). Therefore, it is possible that the absolute decrease of Bifidobacteria may correspond to a relative increase of this group.

The increased incidence of *Peptostreptococcus micros*-like organisms in the microbiota of UC patients in relapse, was also shown in a previous cultivation based study, where the occurrence of *Peptostreptococcus* spp. appeared to be specifically associated with UC patients (293). These results, coupled with the fact that peptostreptococci are among the major oral H_2S producers (366), implies that this group might be responsible for the overproduction of H_2S in some of the UC patients. *In vitro* studies have shown that 5-aminosalicylic acid treatment induces the reduction of H_2S production in faecal slurries of UC patients, which could not be related to SRB (374). The fact that an increased number of *Peptostreptococcus* spp. was only detected in relapse might be related to the fact that patients taking part in this study were treated with 5-aminosalicylic acid.

Clostridium difficile detection was previously associated with the subset of UC patients during relapse (178, 343). Our results confirm this finding as both the diversity and relative abundance of *Clostridium difficile*-like organisms were found to be significantly

increased during relapse of UC patients. *C. difficile* was detected in six out of twelve analysed UC patients. The fact that patients recruited for this study were negative on presence on *C. dificile* exotoxins A and B, and the fact that *C. difficile* was detected in faeces of six UC patients might seem contradicting. However, it is an established fact that this organism can persist in the gastrointestinal contents in low numbers until an external factor, such as antibiotics, causes perturbations of the microbiota, when *C. difficile* multiplies and exhibits its pathogenic properties (357). Interestingly, the presence of *C. difficile* and *Peptostreptococcus* spp. appeared to be mutually exclusive, as 7 out of 9 subjects harboured only one of the two organisms (data not shown). In addition to *C. difficile* itself, two other, closely related uncultured phylotypes were found to have increased abundance and incidence among relapsing UC patients. Their impact on the UC remains to be determined.

In addition to the analysis of the total microbiota, we have assessed the diversity of SRB in relation to UC. Based on the reported results and numerous other results, which were not presented in this paper, no link between UC and the diversity of this group of bacteria could be established. It has already been shown that the distribution and abundance, of approximately 10^6-10^7 cells per gram, of *Desulfovibrio* spp. does not differ between healthy subjects and UC patients (154). Such low abundance can explain the fact that none of the SRB species was detected with the analysis of the total faecal microbiota on the HITChip. However, some of older studies report a higher abundance of SRB of up to 10^{10} cells per gram of faeces (171) and suggested a correlation between the distribution of SRB and UC (170). Furthermore, the distribution of *Desulfovibrio piger* was previously found to be associated with UC (281). Even though our sequencing results have confirmed the presence of *Desulfovibrio piger* in faecal samples of both healthy subjects and UC patients, we could not confirm its association with UC. Another member of Desulfovibrionales that was identified in this study is *Bilophila wadsworthia*. It was shown that cysteine and taurine have a greater effect than sulphate on the H_2S production in human faeces (273). *B. wadsworthia*, is a described taurine degrader (30) and was therefore targeted by a specific PCR assay. Based on the findings that this bacterium was more frequently detected among healthy subjects and the fact that *B. wadsworthia* was previously detected by SSU rRNA-gene sequencing of the microbiota of healthy adults (129, 192, 193), this microbe could not be associated to UC. Five other, not yet cultivated phylotypes that belong to Desulfovibrionales, were found to reside in the human intestine. Given that two of these phylotypes have already been detected in other studies of the gastrointestinal microbiota of healthy humans (129), it likely that they represent members of the normal microbiota. Together with the fact that the increased protein intake causes higher H_2S production in faeces (168, 297), while the increased dietary intake of meat and sulphur, but not sulphate, significantly increases the relapse incidence in UC (227) one could speculate that other bacteria, in addition to SRB, are related to the overproduction of

H_2S in UC patients. Possible candidates could be *Peptostreptococcus* spp. and other not yet identified species.

In this study, it was also shown that the microbiota of healthy controls is more diverse when compared to the microbiota of UC patients in relapse and even more in remission. The reduced diversity of the microbiota has been previously associated with inflamed mucosa samples of UC patients (356, 427). The decreased microbial diversity in faecal samples could be anticipated for the relapse, as the reduced diversity was earlier associated to the microbiota of diarrhoeal stools (298). However, it cannot be ruled out that the reduction of the microbial diversity is not as strongly pronounced as detected in this study since it was shown that a quantitatively relevant fraction of the microbiota of UC patients is comprised of yet unknown organisms (443). Aspecific amplification by Desulfovibrionales / Desulfomondales –targeting primers revealed the presence of members of the TM7 phylum. TM7 phylum represents a newly recognised, widely distributed group of yet uncultured bacteria (217), which were previously detected in the intestine of mice (410) and ruminants (465). Very little is known about the functionality of these filamentous bacteria, which are capable of both aerobic and anaerobic growth (475). Although, there is a lack of supporting evidence that these bacteria have an increased abundance in intestinal samples of UC patients, their incidental detection suggests that other microbes, which cannot be detected using currently available isolation procedures and SSU rRNA probes/primers, might be relevant for UC development.

In conclusion, the thorough analysis of the faecal microbiota composition of UC patients using phylogenetic microarrays showed numerous deviations, both in relapse and remission. These deviations included the reduction in diversity and relative abundance of the Firmicutes phylum members. The almost exclusive occurrence of either *Peptostreptococcus* spp. or *Clostridium difficile*-like organisms in faeces of UC patients, and the fact that healthy subjects have a more uniform gastrointestinal microbiota composition, suggests a multi-factorial character of UC. This could explain some contradicting reports concerning the microbiota of UC patients, including the fact that in this study we could not confirm the previously suggested association of SRB with UC. Therefore, the high throughput and comprehensive analysis of the microbiota of a large number of UC patients is the only reasonable approach for revealing the links between the microbiota and UC.

Acknowledgements

This study was carried out with financial support from the Commission of the European Communities, specifically the RTD programs Quality of Life and Management of Living Resources, QLK1-2000-00108, Microbe Diagnostics, coordinated by Michaël Blaut (Dife, Germany) and QLK1-2000-00563, PROGID coordinated by Fergus Shanahan (University

College Cork, Ireland). The study does not necessarily reflect the Commission's views and in no way anticipates its future policy in this area.

CHAPTER 2

General Discussion

Any scientific understanding starts with careful observation and description. For more than 100 years the human gastrointestinal microbiota has been studied from different perspectives, one of which relates to the description of its diversity. Describing the diversity of a microbial ecosystem is only the first, yet indispensable, step in advancing its understanding. Reliable insight into the diversity of the gastrointestinal microbiota is essential for providing a reference framework to study its dynamics in time and space, analyse its functions, and characterise host-microbe interactions. The work presented in this thesis summarises the hitherto gathered knowledge on the diversity of the human gastrointestinal microbiota that served as a basis for developing an effective and high throughput phylogenetic microarray. This newly developed HITChip (Human Intestinal Tract Chip) advanced the comprehensive and high throughput analysis of this ecosystem that provided novel perspectives of the diversity of the microbial consortia within the human gastrointestinal tract.

Diversity of the gastrointestinal microbiota – state of art

Less than 2 years ago, Bäckhed and co-authors stated in a seminal review paper that out of 55 known bacterial phyla only 8 can be found in the human intestine (23). Nowadays, both of these figures should be reconsidered. Depending on the used SSU rRNA gene database, the number of existing bacterial phyla ranges from 31 to 88 (111) while the results presented in this thesis indicate that members of at least 10 bacterial phyla are inhabitants of the human gastrointestinal tract (Fig. 1). The fact that these reconsiderations are needed, argues for the reliability of the present SSU rRNA gene databases, and points out that the description of the diversity of this ecosystem has not yet been brought to an end.

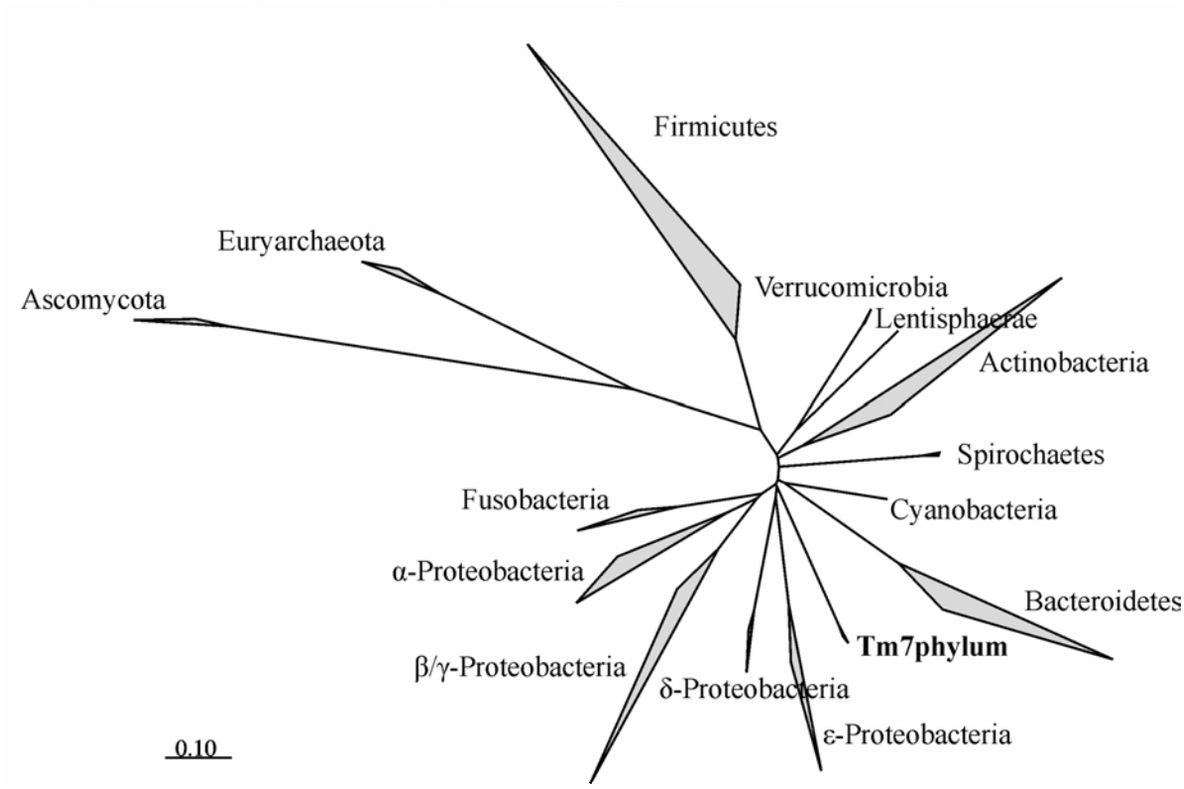


Figure 1 SSU rRNA-based phylogenetic tree representing the diversity of the human gastrointestinal microbiota. The reference bar indicates 10% sequence divergence.

The recent revival of the interest in the human gastrointestinal microbiota triggered numerous studies and resulted in an exceptionally dynamic field of research. In **Chapter 2** we have summarised all fully characterised isolates of human gastrointestinal origin and reported the description of 442 bacterial, 3 archaeal, and 17 eukaryotic species. The value of the provided measure of the species richness is certainly not in the figures themselves, as these are already outdated now due to the recent description of another bacteria species, *Enterococcus caccae* (75), and almost certain suffer from the omission of some older isolates. Furthermore, it is common knowledge that cultivation-based techniques underestimate the

true diversity of the human gastrointestinal microbiota. Nevertheless, this first systematic integration of results presented in almost 300 scientific publications has revealed that the number of known, fully characterised gastrointestinal inhabitants is within the range of the most frequently cited diversity prediction of this ecosystem (between 400 and 500 species). The developed database of these cultured species might also act as a valuable source of information, especially since it provides potential guidelines for the reclassification of some intestinal isolates.

The integration of the results generated in cultivation-dependent and –independent studies presented in **Chapter 3** have shown that number of presently reported gastrointestinal phylotypes exceeds the recent estimation of 1,000 organisms. However, the portrayal of the diversity of the human gastrointestinal microbiota given in **Chapter 3** is certainly not complete, as it omits the members of a phylum candidate TM7, for the first time, revealed in human intestinal samples (see **Chapter 8**). TM7 phylum represents a newly recognised, widely distributed group of yet uncultured bacteria (217). Members of this group had previously been detected in a number of ecosystems including the human oral cavity (244, 363), the mouse intestine (410), the rumen (465) and the beetle intestine (132). Therefore, its detection in the human intestinal tract does not seem to represent an environmental contamination, although its abundance and function in this ecosystem remains unknown. Taking into account this first report of the TM7 phylum, it appears that four out of ten presently known intestinal bacterial phyla have been detected within the last decade as Cyanobacteria were reported in 2006 (129), Verrucomicrobia in 2004 (110), whilst the only hitherto known human intestinal member of Lentisphaera was isolated in 2003 (83, 540). The reason for the delayed detection of these organisms by cultivation-based techniques are the special nutritional demands of those microbes, while the fact that they have escaped detection by SSU rRNA-based techniques might be explained by their relative low abundance and the limitations of the “universal” PCR primers. The design of SSU rRNA-based primers is largely dependent on the sequence knowledge, which is exponentially increasing, and have shown that many previously designed primers/probes do not have the claimed specificity or universality (215).

In **Chapter 3** we have predicted that diversity of the human gastrointestinal microbiota will include thousands of species. Based on the present pace of detecting novel phyla, the presence of other, not yet acknowledged, distantly-related groups can be anticipated. However, newly detected SSU rRNA gene sequences rarely represent distinct branches in the phylogenetic tree as the majority of the recovered sequences clusters within three phyla – Firmicutes, Bacteroidetes and Actinobacteria. Although the completion of the description of the gastrointestinal is a long- term objective, examples of novel isolates, such as *Bacteroides dorei* (27), which SSU rRNA gene sequence has been reported in nine cultivation independent studies, or *Clostridium glycyrrhizinilyticum* (408), which SSU rRNA

gene sequences has been reported twice, are suggesting that, the complete description of the intestinal microbiota is a feasible goal. The fact that the proportion of novel phylotypes reported in some studies appears discouraging (129), should be taken with caution as the number of detected (novel) phylotypes largely depends on the phylotype cut-off value. As previously reported, within the same clone library, the number of distinct phylotypes ranges from 148 to 643 when the cut-off value is changed from 97% to 99% of SSU rRNA gene sequence similarity (496). The choice of the cut-off value seems to be a rather ambiguous issue that depends on the author's decision. It is known that the sequence of different copies of a SSU rRNA in the same genome is rarely below 99%, although it can substantially vary (1). Moreover, this variation increases between different strains of the same species but it rarely reaches values of 97% (254, 403). Hence, in this thesis a cut-off value of 98% was chosen, which is often applied (192, 461) and compromises between three widely used values. In any case, it is clear that SSU rRNA gene sequences alone will not provide absolute answers about the diversity of the human gastrointestinal microbiota. This insight can only be reached after isolation and characterisation of all members of this ecosystem, which, at this stage, is a highly ambitious goal. Systematic isolation might not be the wisest approach since it has, up to now, revealed presence of numerous members of Proteobacteria, which, however, are often very rare, and present in low numbers (81, 340). Although it cannot be ruled out that even such organisms are relevant for the numerous functions of microbiota as whole, more valuable information about this complex ecosystem would be achieved with targeted isolation and characterisation of microbial groups of particular interest. An example of the cultivation that targeted butyrate producers (29, 128) enabled the isolation of *Roseburia* spp. (126) and the reclassification of *Fusobacterium prausntizii* into the novel genus of *Faecalibacterium* (127). These findings appeared to be highly relevant, as those organisms are abundant in the faeces (205, 462), and - as shown in this thesis - alerted in relation to health status (see **Chapters 7** & **8**). Other microbial groups are of particular interest to the ecology of the gastrointestinal microbiota include those that change or remain the same in time or in response to dietary changes, those that are specific for individuals of different health status, age or geographic origin, and those that have specific or strong interactions with the human host. Because of the remarkable complexity of the gastrointestinal microbiota and the strong inter-individual variation in its composition, the identification of such microbial groups is only possible with high throughput analysis of large amount of samples. In **Chapter 4** we have presented a strategy for the development of tool that enables this kind of analysis as is discussed below.

Novel tool for studying complex ecosystems

Chapter 4 describes a generic design strategy that was used for the development of the HITChip a versatile microarray platform that can provide phylogenetic profiling and relative quantification of members of the human gastrointestinal microbiota. Since a large proportion of inhabitants of this ecosystem has only been detected by SSU rRNA gene sequencing (**Chapter 3**), the presented phylogenetic microarray design strategy was based on this gene, or more specifically, the sequence of its two hyper-variable regions. Unlike other available algorithms for microarray probe design (136, 276), in the presented approach, only the sequence of the target organism is taken into account, while non-targeted sequences are included when defining the phylogenetic specificity of probes. The prime reason for introducing this, conceptually new approach was the fact that an accurate design of universal or specific sequences within the SSU rRNA gene is not feasible at the time when the databases of SSU rRNA gene sequences are exponentially increasing and the description of microbial ecosystems is still ongoing. Due to the recent discovery of numerous novel distinct SSU rRNA gene sequences, it appeared that none of the presently available "universal" bacterial primer pairs targets more than 74% of bacterial SSU rRNA gene sequences (215). Moreover, some of the group-specific primers do not have the claimed specificity when applied on complex ecosystem, as shown for the Desulfovibrionales primers in **Chapter 8**. Therefore, the sequence of such probes/primers should be reconsidered with the increasing knowledge on the microbial diversity. On the contrary, probes that are designed using the approach described in **Chapter 4** do not require such reconsideration since no specificity claim was set during their design. However, the phylogenetic specificity of each probe is not an absolute entity, and its accuracy increases with the size of the SSU rRNA sequence target set database. Another important feature that is unique for the presented probe design strategy is that it allows easy and systematic addition of probes that target newly discovered members of targeted ecosystem. Having in mind recently published results of sequencing of thousands of SSU rRNA genes of human intestinal origin (172, 275), this might be one of the most relevant characteristic of the newly presented probe design strategy.

The application of the HITChip showed that results obtained with this technique are in very good agreement with and in some cases superior to other SSU rRNA-based techniques such as Fluorescent In Situ Hybridisation (FISH) using specific labelled rRNA probes (**Chapter 4**) or Denaturing Gradient Gel Electrophoresis (DGGE) of rRNA amplicons (**Chapter 5**). The HITChip appeared to be an accurate tool that can be used for comprehensive profiling of the human gastrointestinal microbiota, which is upgraded with phylogenetic identification and relative quantification performed at different levels of phylogenetic resolution. Therefore, this novel phylgenetic microarray was used to address relevant questions concerning the composition of the human gastrointestinal microbiota in health and disease. The first studies in which the HITChip was employed could be considered

as pilot studies that - at the same time - provided novel perspectives about the human gastrointestinal ecosystem.

Novel perspectives from high throughput analysis

Analysis of the dynamics of the human gastrointestinal microbiota has, for the first time, shown that different subpopulations do not follow the same pattern of change during temporal variations (**Chapter 6**). While members of the Firmicutes were found to actively respond to environmental stimuli, Bacterodetes and Actinobacteria have exhibited an astonishing level of stability within a life span of four years. This finding allowed us to speculate that the human gastrointestinal microbiota differs from other ecosystems, such as bioreactors, in which unchanged metabolic activity can be preserved, despite notable changes in the community composition (148), and that due to the constant and intimate interactions between the human host and microbes a subset of the microbiota seems to resist to the environmental triggering. The molecular mechanisms of the communication between host and intestinal microbes are beginning to be understood (188, 385), and our results suggest the presence of specific receptors that may recognise genotype-specific *Bifidobacterium* and *Bacteroides* spp. Furthermore, we have shown that changes of the microbiota should be considered as fluctuations rather than shifts, as about 10% of the microbial species occasionally accompany the core microbiota of a person, which remains the same.

One of the explanations for the fact that those findings are only now reported is the fact that that this long-term analysis has not been previously addressed. Moreover, the direct coupling of microbiota profiling and phylogenetic analysis facilitated the efficient data extraction. This advantage of phylogenetic microarrays can be valuable for assessing other relevant information, such as difference between active and inactive bacteria by analysing DNA- and RNA-based profiles of the microbiota. Other fingerprinting techniques, such as DGGE, have been previously applied for such analysis (470, 537), although for identification of the differential groups observed in DGGE patterns, construction of clone libraries is required, which makes analysis laborious and time consuming. Assessment of those profiles for a faecal sample of an individual showed a more than two-fold difference between the hybridisation signals of more than 20 phylogenetic groups (Fig. 2). The microbial landscape as assessed by either DNA or RNA was found to be tremendously different, as Actinobacteria were composing 1% of the total bacteria (based on the DNA profile) and 18% of the active bacteria (based on the RNA profile), whereas Bacteroidetes were found to comprise 30% of the total (based on the DNA profile) and only 17% of the active bacteria (based on the RNA profile).

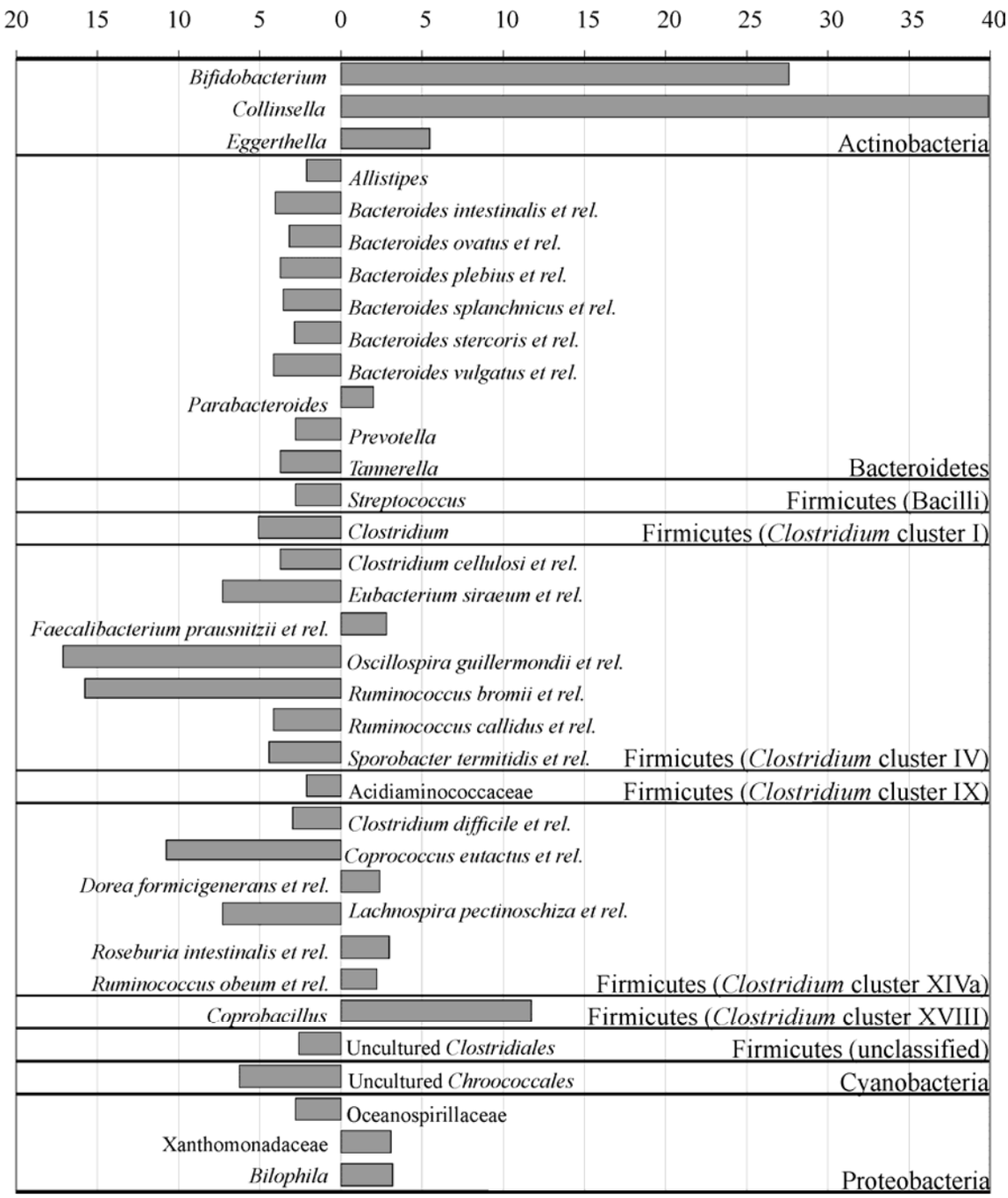


Figure 2 Ratio of the average hybridisation signals obtained by the analysis of faecal DNA and RNA. The orientation of bars depends on the obtained hybridisation intensities, left oriented bars represent phylogenetic groups that are less active (higher hybridisation signals with DNA), while right oriented bars represent active phylogenetic groups (higher hybridisation signal with RNA).

In a previous, DGGE-based study, *Collinsella* spp. and *Bifidobacterium* spp. were identified as the most active (470), which is in agreement with our findings, as those groups showed a remarkably increased signal in hybridisation with RNA as compared to the hybridisation with DNA (40- and 27-fold difference, respectively). The fact that the difference of the other 18 phylogenetic groups was not detected using DGGE, illustrates the

lower sensitivity of this technique in comparison to the HITChip. In addition, it should be noted that quantification of faecal bacteria by FISH and HITChip showed the largest disparity for Actinobacteria and Bacteroidetes (**Chapter 4**). Although the difference between RNA- and DNA-based HITChip profiles is hitherto assessed based on the analysis of a sole faecal sample, it can be speculated that this difference in expression but also the effect of the number of SSU rRNA genes in the genome, are the major factors for inducing differences in quantification by different techniques. It is to be expected that analysis with phylogenetic microarrays might help in identification and defining of such relevant factors.

Defining the faecal microbiota diversity in relation to irritable bowel syndrome (IBS) and ulcerative colitis (UC) was another goal of this thesis. The aetiology of UC and IBS is unknown and complex, and includes both environmental and genetic factors (41, 184, 187, 451). IBS, similarly to UC in remission, is characterised by low-grade inflammation and functional disturbances in the intestine and both pathologies are strongly associated with intestinal bacteria (188). The results of the high throughput profiling of faecal samples showed that the microbiota of both IBS and UC patients has a distinct composition when compared to corresponding controls of healthy individuals (**Chapter 7** & **8**). A general observation is that members of Firmicutes are the most dramatically affected in both aberrations, which is in line with our finding that this taxon is stronger affected than the other microbial groups by environmental changes (**Chapter 6**). However, Firmicutes are also the most diverse and abundant taxonomic group within the gastrointestinal microbiota, primarily consisting of uncultured organisms, the functionality of which is unknown. Therefore, the accurate rationalisation of the observed trends is possible for only few microbial groups.

The increased incidence of *Peptostrepotcoccus* spp. detection in some UC patients was for the first time correlated to their well-established overproduction of H_2S (273), pointing to the contribution of other H_2S producers than the canonical sulphate-reducing bacteria. Furthermore, the reduced abundance of *Roseburia* spp. in UC patients could be correlated with the recently shown reduction of butyrate levels in faeces of UC patients (303). In contrast, this group was found to be increased in diarrhoea-predominant IBS patients, which adds to the controverse about butyrate's effect on the human health (290, 422). While butyrate was, for long time, considered as an important microbial metabolite that promotes intestinal cell growth (395), some novel findings indicated that butyrate can also have a necrotic effect on the colonic cell lines (307).

Another novel finding of this work in relation to intestinal disorders concerns the report of increased heterogeneity among the microbial composition of both IBS and UC patients. In case of IBS this finding was in line with the systematisation of IBS patients into three distinct groups based on their clinical symptoms. Although it has been suggested that UC is also heterogeneous disease (416), in **Chapter 8**, we have presented the first biological evidence of its heterogeneity. These observations stress the need for high-throughput and

comprehensive microbial profiling of large amount of samples, which can be accomplished using phylogenetic microarrays.

Finally, on top of the analysis of samples collected in human trials, we have monitored the microbial composition during preparation and use of an *in vitro* model of the human large intestine (**Chapter 5**). Defining the microbiota of *in vitro* models is a relevant issue because of their wide application (291), the results of which seems to be appreciated by a wide scientific audience. Up to now the PubMed browser (www.ncbi.nlm.nih.gov) retrieved more than 600 scientific publications that described results obtained with *in vitro* models of the human intestinal tract. The use of *in vitro* models has certain advantages when compared to human trials, which includes more controlled and standardised experiments and sampling procedures (131). However, results obtained using different *in vitro* systems, have raised questions about the accuracy with which those models can mimic the microbiota of the gastrointestinal part they are intending to resemble. These data include the identification of a very simple subpopulation of the glucose-utilisers within the complex gastrointestinal microbiota (130), or the omission of the only human intestinal obligate mucin degrader *Akkermansia muciniphila* (110) during *in vitro* fermentation on mucin-based medium, despite the presence of this bacterium in the faecal inoculuum (270). The results presented in **Chapter 5** show that, although the microbial ecosystem in the TIM-2 *in vitro* model is considerably diverse, there are elements in operating the model that introduce changes in the microbial composition. Those changes cause deviations in the relative abundance of some microbial groups, when compared to the *in vivo* situation, which might have an impact on the accuracy of the results obtained with *in vitro* models. The identification of a reduced number of fastidious anaerobes can be used as a guideline for adjustments in the standard operating procedures to improve simulating the diversity of healthy adults. Finally, the awareness about the presence of differences is, as such, an advantage, which should be used for adequate data interpretation as any *in vitro* model represents an approximation of the real system and the limits of this approximation should be precisely defined.

The data presented in **Chapter 5** revealed another important feature of the analysis with the HITChip. This feature concerns easy combination of data from different studies, as the microbial composition of the TIM-2 *in vitro* model was compared with the microbiota of healthy adults already analysed in another study (**Chapter 6**). With the growing size of the HITChip database, novel questions can be raised and answered by simple comparisons of previously analysed samples. The influence of geographic origin or difference in the microbiota of babies and adults (Fig. 3) is one of the questions that can be readily answered by comparing selections of more than 250 microbiota profiles that are at this moment are present in the HITChip database.

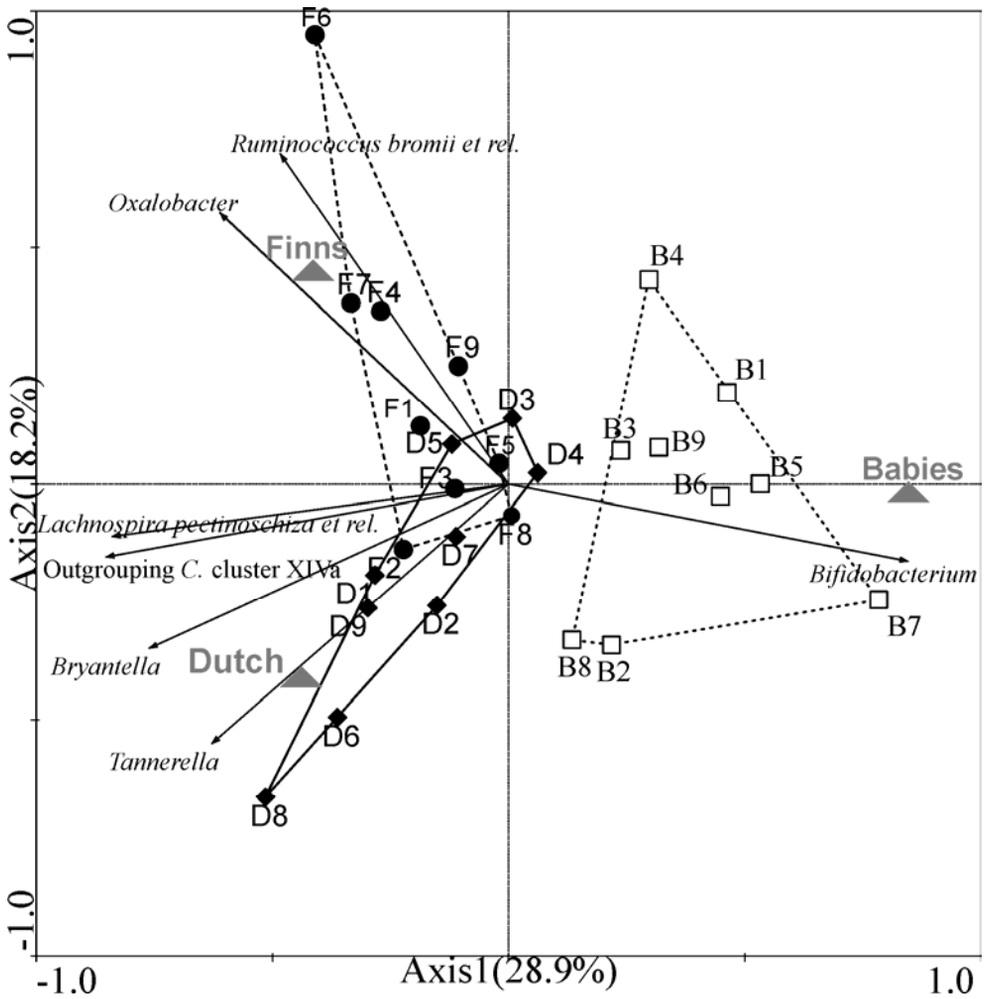


Figure 3 PCA triplot containing depicting the microbiota of Dutch healthy adults (◆D1-D9), Finnish healthy adults (●F1-F9), and Italian babies (□B1-B9). Percentage values at the axes indicate contribution of the principle components to the explanation of total variance in the dataset, while phylogenetic groups, named after cultivated representative, that contributed at least 70% to the explanatory axis used in the plot are presented as vectors, while centroides of the plot are describing age and geographic origin of the subjects.

Wider perspectives

In this thesis the phylogenetic microarray - HITChip - has been instrumental in benchmarking the intestinal microbiota composition of diseased individuals (**Chapters 7** & **8**), while analysis of the microbiota of healthy humans provided phylogenetic insight into the intestinal microbiota's fractions that are shaped by diet, stochastic processes and specific host interactions (**Chapter 5**). As those results only illustrate the power of the studies that employ phylogenetic microarrays, it can be claimed that the phylogenetic microarray-based analysis of the microbial diversity within the human gastrointestinal tract can and will provide novel insights into this ecosystem and will considerably improve our knowledge about it. However, as any other approach, phylogenetic microarrays display only one dimension, in particular

case the diversity, which should be complemented with other presently available approaches to build the complete portrayal of this highly relevant ecosystem.

The HITChip enables simultaneous analysis of over 1,000 phylotypes, which makes it the most comprehensive hitherto available tool for the analysis of the diversity of the human gastrointestinal microbiota. However, the targeted 1,000 phylotypes represent only a fraction of the total microbial diversity, as thousands of novel distinct SSU rRNA sequences are expected to be found in samples of human gastrointestinal origin (**Chapter 3**). Despite the fact that the generated phylogenetic profiles provide diagnostic information even about unknown phylotypes (**Chapter 4**), the further improvement of the HITChip, and other similar phylogenetic microarrays (359), is dependent on the discovery of novel gastrointestinal inhabitants. Therefore, the accuracy and comprehension of phylogenetic microarray analysis will largely benefit from the reports of novel SSU rRNA sequences from human gastrointestinal clone libraries (52, 129, 270, 275), but also from cultivation of the novel gastrointestinal inhabitants (75).

Fingerprinting techniques, such as DGGE, have played a major role in the molecular revolution of our view on complex ecosystems. Their application has revealed many relevant observations of the human gastrointestinal microbiota, including the conclusion that the microbiota is subject-specific, affected by genotype, and stable in time (536, 537). However, the high-resolution profiling that can be obtained using phylogenetic microarrays has several advantages over the well-established fingerprinting by DGGE (346)(**Chapter 5** & **6**). Therefore, it is to be expected that phylogenetic microarrays will prevail over other fingerprinting techniques in future diagnostic surveys of the intestinal and other ecosystems.

It may be considered that the capacity of clone library and phylogenetic microarray approaches is limited by the fact that they detect only the dominant groups. However, this is only correct only up to some extent, as the use of specific PCR approaches can enhance the specificity of clone libraries and phylogenetic microarrays, and therefore, enable gaining insight into less dominant populations (**Chapter 6**)(194, 417). Therefore, the only major disadvantage of these two approaches is that they enable only relative quantification of an ecosystem. A powerful technique that can provide the lacking information is quantitative PCR (qPCR), which can detect and enumerate subdominant populations with sensitivity as little as 10^3 cells per gram of faeces (~0.000001% of the total community) (304), while numerous group- and species-specific qPCR assays enable quantification of signification proportion of the human gastrointestinal inhabitants (305, 306, 393).

Another SSU rRNA-based technique that allows for quantification of an ecosystem is FISH, which is the only technique that can accurately express quantitative information in cell numbers. This becomes particularly relevant when analyses of uncultivated organisms are concerned, as quantification in DNA based assays is biased by variable number of SSU rRNA gene copies, as discussed in **Chapter 4**. Applications of FISH have enabled the first insights

into the morphology and physiology of uncultured organisms, for instance members of TM7 phylum candidate (217, 475). In addition to the quantification of microbial groups in complex ecosystems, which has been exceptionally fruitful (268, 339, 443), FISH technique has significant other potential. Its combination with fluorescence-activated cell sorting is a promising approach for determining relevant data on the complex ecosystems, such as analysis of alive, injured and dead subsets of intestinal bacteria (39, 490). The development of functional gene targeted CPRINC-FISH (245) or RING-FISH (545) in combination with cell sorting have a great potential for detectionion of microbial subpopulations that are responsible for relevant conversions. Similar to targeted cultivation (128), this approach would enable identification of important microbial species. Furthermore, functional genes have shown to have similar discriminative power as housekeeping genes, such as SSU rRNA, which can be used for species identification and detection of novel phylotypes of particular guild (498). Therefore, it can be speculated that relevant functional genes will form the basis for future generations of microarrays that couple analysis of diversity and activity.

Finally, functional analyses, which were not in the scope of the work presented in this thesis, are the ultimate goal of the research on the human gastrointestinal microbiota. Reductionist approaches, which analyse simplified microbial communities in animal or *in vitro* models, have revolutionised the perception of the microbial impact on the metabolism of host. The pioneering work of J. Gordon and co-workers, which has shown that microbial colonisation improves nutritional and defensive functions of host (210, 211), was further expanded by including other microbial species (108, 450), but also by analysing impact of the host on the microbe (102, 450). Furthermore, reductionist approaches of analysing different cultured cells provide insight into new genes and functions of individual organisms, which is a milestone for the global approaches, such as metagenomics. Although metagenomic analysis have shown their power for revealing the coding potential of the microbiome, the genome of the entire microbial ecosystem (172, 302), further functional metagenomics studies, such as metaproteomics, have been reported but show the limitations in our predictive capacity (252). Hence, there is a great need to integrate all available reductionistic and global, cultivation- and molecular-based approaches to finally describe and understand our micro companions throughout our journey on planet Earth.

Bibliography

1. **Acinas SG, Marcelino LA, Klepac-Ceraj V & Polz MF** (2004) Divergence and redundancy of 16S rRNA sequences in genomes with multiple rrn operons. *J Bacteriol* 186: 2629-2635.
2. **Adlercreutz H, Pulkkinen MO, Hamalainen EK & Korpela JT** (1984) Studies on the role of intestinal bacteria in metabolism of synthetic and natural steroid hormones. *J Steroid Biochem* 20: 217-229.
3. **Akehurst R & Kaltenthaler E** (2001) Treatment of irritable bowel syndrome: a review of randomised controlled trials. *Gut* 48: 272-282.
4. **Albert MJ, Mathan VI & Baker SJ** (1980) Vitamin B_{12} synthesis by human small intestinal bacteria. *Nature* 238: 781-782.
5. **Aleksic S, Heinzerling F & Bockemuhl J** (1996) Human infection caused by *Salmonellae* of subspecies II to VI in Germany, 1977-1992. *Zentralbl Bakteriol* 183: 391-398.
6. **Aleksic S, Steigerwalt AG, Bockemuhl J, Huntley-Carter GP & Brenner DJ** (1987) *Yersinia rohdei* sp. nov. isolated from human and dog feces and surface water. *Int J Syst Bacteriol* 37: 327-332.
7. **Allen A** (1981) Structure and function of gastrointestinal mucus. *Physiology of the gastrointestinal tract.* (Johnson LR, eds). pp. 617-619. Raven Press, New York.
8. **Allison MJ, Cook HM, Milne DB, Gallagher S & Clayman RV** (1986) Oxalate degradation by gastrointestinal bacteria from humans. *J Nutr* 116: 455-460.
9. **Allison MJ, Dawson KA, Mayberry WR & Foss JG** (1985) *Oxalobacter formigenes* gen. nov., sp. nov.: oxalate-degrading anaerobes that inhabit the gastrointestinal tract. *Arch Microbiol* 141: 1-7.
10. **Alston JM** (1928) An investigation of streptococci isolated from the alimentary tract of man and certain animals. *J Bacteriol* 16: 397-407.
11. **Altschul SF, Gish W, Miller W, Meyers EW & Lipman DJ** (1990) Basic Local Alignment Search Tool. *J Mol Biol* 215: 403.
12. **Amann RI, Ludwig W & Schleifer KH** (1995) Phylogenetic identification and *in situ* detection of individual microbial cells without cultivation. *Microbiol Rev* 59: 143-169.
13. **Anderson HW** (1917) Yeast-like fungi of the human intestinal tract. *J Infect Dis* 21: 341-386.
14. **Andresen V & Camilleri M** (2006) Irritable bowel syndrome: recent and novel therapeutic approaches 66: 1073-1088.
15. **Andrewes FW & Horder TJ** (1906) A study of the streptococci pathogenic for man. *Lancet* 2: 708-713.
16. **Anonymous** (1981) The pioneers of pediatric medicine: Teodor Escherich. *Eur J Pediatr* 137: 131.
17. **Aranda-Michel J & Giannella RA** (1999) Acute diarrhea: a practical review. *The American Journal of Medicine* 106: 670.
18. **Archer JR, Romero S, Ritchie AE, Hamacher ME, Steiner BM, Bryner JH & Schell RF** (1988) Characterization of an unclassified microaerophilic bacterium associated with gastroenteritis. *J Clin Microbiol* 26: 101-105.
19. **Aslam M, Batten JJ, Florin TH, Sidebotham RL & Baron JH** (1992) Hydrogen sulphide induced damage to the colonic mucosal barrier in the rat. *Gut* 33: S69.
20. **Attene-Ramos MS, Wagner ED, Plewa MJ & Gaskins HR** (2006) Evidence that hydrogen sulfide is a genotoxic agent. *Mol Cancer Res* 4: 9-14.
21. **Atuma C, Strugala V, Allen A & Holm L** (2001) The adherent gastrointestinal mucus gel layer: thickness and physical state *in vivo*. *Am J Physiol Gastrointest Liver Physiol* 280: G922-929.
22. **Avgustin G, Wallace RJ & Flint HJ** (1997) Phenotypic diversity among ruminal isolates of *Prevotella ruminicola*: proposal of *Prevotella brevis* sp. nov., *Prevotella bryantii* sp. nov., and *Prevotella albensis* sp. nov. and redefinition of *Prevotella ruminicola*. *Int J Syst Bacteriol* 47: 284-288.
23. **Bäckhed F, Ley RE, Sonnenburg JL, Peterson DA & Gordon JI** (2005) Host-bacterial mutualism in the human intestine. *Science* 307: 1915-1920.
24. **Baker SJ, Jacob E & Bowden GH** (2000) Crohn disease arthropathy: antigens in synovial fluid share epitopes with strains of two species of viridans streptococci. *Scand J Gastroenterol* 35: 287-292.
25. **Bakir MA, Kitahara M, Sakamoto M, Matsumoto M & Benno Y** (2005) *Bacteroides finegoldii* sp. nov., isolated from human faeces. *Int J Syst Evol Microbiol*.
26. **Bakir MA, Kitahara M, Sakamoto M, Matsumoto M & Benno Y** (2005) *Bacteroides intestinalis* sp. nov., isolated from human feces. *Int J Syst Evol Microbiol*: in press.
27. **Bakir MA, Sakamoto M, Kitahara M, Matsumoto M & Benno Y** (2006) *Bacteroides dorei* sp. nov., isolated from human faeces. *Int J Syst Evol Microbiol* 56: 1639-1643.
28. **Balsari A, Ceccarelli A, Dubini F, Fesce E & Poli G** (1982) The fecal microbial population in the irritable bowel syndrome. *Microbiologica* 5: 185-194.
29. **Barcenilla A, Pryde SE, Martin JC, Duncan SH, Stewart CS, Henderson C & Flint HJ** (2000) Phylogenetic relationships of butyrate-producing bacteria from the human gut. *Appl Environ Microbiol* 66: 1654-1661.
30. **Baron EJ** (1997) *Bilophila wadsworthia*: a unique Gram-negative anaerobic rod. *Anaerobe* 3: 83-86.

31. **Barrett SLR, Cookson BT, Carlson LC, Bernard KA & Coyle MB** (2001) Diversity within reference strains of *Corynebacterium matruchotii* includes *Corynebacterium durum* and a novel organism. *J Clin Microbiol* 39: 943-948.
32. **Barros RR, Carvalho MDGS, Peralta JM, Facklam RR & Teixeira LM** (2001) Phenotypic and genotypic characterization of pediococcus strains isolated from human clinical sources. *J Clin Microbiol* 39: 1241-1246.
33. **Batchelor MD** (1919) Aerobic spore-bearing bacteria in the intestinal tract of children. *J Bacteriol* 4: 23-3415.
34. **Bayliss CE, Bradley HK, Jones VA & Hunter JO** (1986) Some aspects of colonic microbial activity in irritable bowel syndrome associated with food intolerance. *Ann Ist Super Sanita* 22: 959-963.
35. **Bayston R, Leung TS & Spitz L** (1984) Faecal flora in neonates with oesophageal atresia. *Arch Dis Child* 59: 126-130.
36. **Begley M, Gahan CGM & Hill C** (2005) The interaction between bacteria and bile. *FEMS Microbiol Rev* 29: 625-651.
37. **Behrens S, Ruhland C, Inacio J, Huber H, Fonseca A, Spencer-Martins I, Fuchs BM & Amann R** (2003) *In situ* accessibility of small-subunit rRNA of members of the domains Bacteria, Archaea, and Eucarya to Cy3-labeled oligonucleotide probes. *Appl Environ Microbiol* 69: 1748-1758.
38. **Ben-Amor K** (2004) Microbial eco-physiology of the human intestinal tract: a flow cytometric approach. In *Laboratory of Microbiology*. PhD Wageningen: Wageningen University, pp. 166.
39. **Ben-Amor K, Heilig HG, Smidt H, Vaughan EE, Abee T & de Vos WM** (2005) Genetic diversity of viable, injured, and dead fecal bacteria assessed by fluorescence-activated cell sorting and 16S rRNA gene analysis. *Appl Environ Microbiol* 71: 4679-4689.
40. **Bengmark S** (1998) Ecological control of the gastrointestinal tract. The role of probiotic flora. *Gut* 42: 2-7.
41. **Bengtson MB, Ronning T, Vatn MH & Harris JR** (2006) Irritable bowel syndrome in twins: genes and environment. *Gut* 55: 1754-1759.
42. **Benno Y, Endo K, Takeo M, Namba Y, Komori T & Mitsuoka T** (1989) Comparison of fecal microflora of elderly persons in rural and urban areas of Japan. *Appl Environ Microbiol* 55: 1100-1105.
43. **Benno Y, Sawada K & Mitsuoka T** (1984) The intestinal microflora of infants: composition of fecal flora in breast-fed and bottle-fed infants. *Microbiol Immunol* 28: 975-986.
44. **Benno Y, Sawada K & Mitsuoka T** (1985) The intestinal microflora of infants: fecal flora of infants with vitamin K deficiency. *Microbiol Immunol* 29: 243-250.
45. **Benno Y, Sawada K & Mitsuoka T** (1986) Comparison of the fecal microflora in rural Japanese and urban Canadians. *Microbiol Immunol* 30: 521-532.
46. **Bernalier A, Willems A, Leclerc M, Rochet V & Collins MD** (1996) *Ruminococcus hydrogenotrophicus* sp. nov., a new H_2/CO_2-utilizing acetogenic bacterium isolated from human feces. *Arch Microbiol* 166: 176-183.
47. **Bertilsson S, Cavanaugh CM & Polz MF** (2002) Sequencing-independent method to generate oligonucleotide probes targeting a variable region in bacterial 16S rRNA by PCR with deatachable primers. *Appl Environ Microbiol* 68: 6077-6086.
48. **Biasoli SM, Tosello EM & Magaró MH** (2002) Adherence of *Candida* strains isolated from the human gastrointestinal tract. *Mycoses* 45: 465-469.
49. **Biavati B & Mattarelli P** (1991) *Bifidobacterium ruminantium* sp. nov. and *Bifidobacterium merycicum* sp. nov. from the rumens of cattle. *Int J Syst Bacteroil* 41: 163-168.
50. **Biavati B, Vescovo M, Torriani S & Bottazzi V** (2000) *Bifidobacteria*: history, ecology, physiology and applications. *Ann Microbiol* 50: 117-131.
51. **Bibiloni R, Simon MA, Albright C, Sartor B & Tannock GW** (2005) Analysis of the large bowel microbiota of colitic mice using PCR/DGGE. *Lett Appl Microbiol* 41: 45-51.
52. **Bik EM, Eckburg PB, Gill SR, Nelson KE, Purdom EA, Francois F, Perez-Perez G, Blaser MJ & Relman DA** (2006) Molecular analysis of the bacterial microbiota in the human stomach. *Proc Natl Acad Sci* 103: 732-737.
53. **Bjorkroth KJ, Schillinger U, Geisen R, Weiss N, Hoste B, Holzapfel WH, Korkeala HJ & Vandamme P** (2002) Taxonomic study of *Weissella confusa* and description of *Weissella cibaria* sp. nov., detected in food and clinical samples. *Int J Syst Evol Microbiol* 52: 141-148.
54. **Blanchard AP, Kaiser RJ & Hood LE** (1996) High-density oligonucleotide arrays. *Biosens Bioelectron* 11: 687–690.
55. **Bodrossy L & Sessitsch A** (2004) Oligonucleotide microarrays in microbial diagnostics. *Curr Opin Microbiol* 7: 245.
56. **Bodrossy L, Stralis-Pavese N, Murrell JC, Radajewski S, Weilharter A & Sessitsch A** (2003) Development and validation of a diagnostic microbial microarray for methanotrophs. *Environ Microbiol* 5: 566-582.

57. **Bolstad BM, Irizarry RA, Astrand M & Speed TP** (2003) A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. *Bioinformatics* 19: 185-193.
58. **Bond JH, Jr., Engel RR & Levitt MD** (1971) Factors influencing pulmonary methane excretion in man: an indirect method of studying the in situ metabolism of the methane-producing colonic bacteria. *J Exp Med* 133: 572-588.
59. **Bonnet R** (2002) Phylogenetic diversity of the human gut microflora based on small-subunit rRNA. In *School of Food Biosciences*. Vol. Ph.D. Reading: University of Reading.
60. **Bonnet R, Suau A, Doré J, Gibson GR & Collins MD** (2002) Differences in rDNA libraries of faecal bacteria derived from 10- and 25-cycle PCRs. *Int J Syst Evol Microbiol* 52: 757-763.
61. **Boyer EW, Ingle MB & Mercer GD** (1973) *Bacillus alcalophilus* subsp. *halodurans* subsp. nov.: an alkaline-amylase-producing, alkalophilic organism. *Int J Syst Bacteriol* 23: 238-242.
62. **Bray GA** (1998) Obesity: a time bomb to be defused. *The Lancet* 352: 160-161.
63. **Breitbart M, Hewson I, Felts B, Mahaffy JM, Nulton J, Salamon P & Rohwer F** (2003) Metagenomic analyses of an uncultured viral community from human feces. *J Bacteriol* 185: 6220-6223.
64. **Brenner DJ, Grimont PA, Steigerwalt AG, Fanning GR, Ageron E & Riddle CF** (1993) Classification of citrobacteria by DNA hybridization: designation of *Citrobacter farmeri* sp. nov., *Citrobacter youngae* sp. nov., *Citrobacter braakii* sp. nov., *Citrobacter werkmanii* sp. nov., *Citrobacter sedlakii* sp. nov., and three unnamed *Citrobacter* genomospecies. *Int J Syst Bacteriol* 43: 645-658.
65. **Brenner DJ, Hickman-Brenner FW, Lee JV, Steigerwalt AG, Fanning GR, Hollis DG, Farmer JJ, 3rd, Weaver RE, Joseph SW & Seidler RJ** (1983) *Vibrio furnissii* (formerly aerogenic biogroup of *Vibrio fluvialis*), a new species isolated from human feces and the environment. *Int J Syst Bacteriol* 18: 816-824.
66. **Brenner DJ, McWhorter AC, Kai A, Steigerwalt AG & Farmer JJ, 3rd** (1986) *Enterobacter asburiae* sp. nov., a new species found in clinical specimens, and reassignment of *Erwinia dissolvens* and *Erwinia nimipressuralis* to the genus *Enterobacter* as *Enterobacter dissolvens* comb. nov. and *Enterobacter nimipressuralis* comb. nov. *J Clin Microbiol* 23: 1114-1120.
67. **Brenner DJ, O'Hara CM, Grimont PAD, Janda JM, Falsen E, Aldova E, Ageron E, Schindler J, Abbott SL & Steigerwalt AG** (1999) Biochemical identification of *Citrobacter* species defined by DNA hybridization and description of *Citrobacter gillenii* sp. nov. (formerly *Citrobacter* genomospecies 10) and *Citrobacter murliniae* sp. nov. (formerly *Citrobacter* genomospecies 11). *J Clin Microbiol* 37: 2619-2624.
68. **Brouns F, Kettlitz B & Arrigoni E** (2002) Resistant starch and "the butyrate revolution". *Trends Food Sci Technol* 13: 251-261.
69. **Brown DW & Moore WEC** (1960) Distribution of *Butyrivibrio fibrisolvens* in nature. *J Dairy Sci* 43: 1570-1574.
70. **Bucher C & von Graevenitz A** (1982) Evaluation of three differential media for detection of *Enterobacter agglomerans* (*Erwinia herbicola*). *J Clin Microbiol* 15: 1164-1166.
71. **Buchholz-Cleven BEE, Rattunde B & Straub KL** (1997) Screening for genetic diversity of isolates of anaerobic Fe(II)-oxidizing bacteria using DGGE and whole cell hybridization. *Syst Appl Microbiol* 20: 301-309.
72. **Carlsson J & Gothefors L** (1975) Transmission of *Lactobacillus jensenii* and *Lactobacillus acidophilus* from mother to child at time of delivery. *J Clin Microbiol* 1: 124-128.
73. **Carnahan AM, Chakraborty T, Fanning GR, Verma D, Ali A, Janda JM & Joseph SW** (1991) *Aeromonas trota* sp. nov., an ampicillin-susceptible species isolated from clinical specimens. *J Clin Microbiol* 29: 1206-1210.
74. **Carter RL** (1989) On the partitioning of squared Euclidean distance and its application in cluster analysis. *Phychometrika* 54: 9-23.
75. **Carvalho MdGS, Shewmaker PL, Steigerwalt AG, Morey RE, Sampson AJ, Joyce K, Barrett TJ, Teixeira LM & Facklam RR** (2006) *Enterococcus caccae* sp. nov., isolated from human stools. *Int J Syst Evol Microbiol* 56: 1505-1508.
76. **Castellani A & Chalmers A** (1919) *Manual of tropical medicine.* Balliére, Tyndall and Cox, London.
77. **Cato EP, Salmon CW & Holdeman LV** (1974) *Eubacterium cylindroides* (Rocchi) Holdeman and Moore: Emended description and designation of neotype strain. *Int J Syst Bacteriol* 24: 256-259.
78. **Cebra JJ** (1999) Influences of microbiota on intestinal immune system development. *Am J Clin Nutr* 69: 1046S-1051.
79. **Ceelen L, Decostere A, Verschraegen G, Ducatelle R & Haesebrouck F** (2005) Prevalence of *Helicobacter pullorum* among patients with gastrointestinal disease and clinically healthy persons. *J Clin Microbiol* 43: 2984-2986.
80. **Chee M, Yang R, Hubbell E, Berno A, Huang XC, Stern D, Winkler J, Lockhart DJ, Morris MS & Fodor SPA** (1996) Accessing genetic information with high-density DNA arrays. *Science* 274: 610-614.

81. **Chiesa C, Pacifico L, Nanni F, Renzi AM & Ravagnan G** (1993) *Yersinia pseudotuberculosis* in Italy. Attempted recovery from 37,666 samples. *Microbiol Immunol* 37: 391-394.
82. **Child MW, Kennedy A, Walker AW, Bahrami B, Macfarlane S & Macfarlane GT** (2006) Studies on the effect of system retention time on bacterial populations colonizing a three-stage continuous culture model of the human large gut using FISH techniques. *FEMS Microbiol Ecol* 55: 299-310.
83. **Cho JC, Vergin KL, Morris RM & Giovannoni SJ** (2004) *Lentisphaera araneosa* gen. nov., sp. nov, a transparent exopolymer producing marine bacterium, and the description of a novel bacterial phylum, Lentisphaerae. *Environ Microbiol* 6: 611-621.
84. **Cleveland WS, Grosse E & Shyu WM** (1992) Local regression models. *Statistical Models in S*. (Chambers JM & Hastie TJ, eds)Wadsworth & Brooks, Cole.
85. **Coenye T, Vancanneyt M, Cnockaert MC, Falsen E, Swings J & Vandamme P** (2003) *Kerstersia gyiorum* gen. nov., sp. nov., a novel *Alcaligenes faecalis*-like organism isolated from human clinical samples, and reclassification of *Alcaligenes denitrificans* Ruger and Tan 1983 as *Achromobacter denitrificans* comb. nov. *Int J Syst Evol Microbiol* 53: 1825-1831.
86. **Cole JR, Chai B, Marsh TL, Farris RJ, Wang Q, Kulam SA, Chandra S, McGarrell DM, Schmidt TM, Garrity GM & Tiedje JM** (2003) The Ribosomal Database Project (RDP-II): previewing a new autoaligner that allows regular updates and the new prokaryotic taxonomy. *Nucl Acids Res* 31: 442-443.
87. **Collins M, Lawson P, Willems A, Cordoba J, Fernandez-Garayzabal J, Garcia P, Cai J, Hippe H & Farrow J** (1994) The phylogeny of the genus *Clostridium*: proposal of five new genera and eleven new species combinations. *Int J Syst Bacteriol* 44: 812-826.
88. **Collins MD, Bernard KA, Hutson RA, Sjoden B, Nyberg A & Falsen E** (1999) *Corynebacterium sundsvallense* sp. nov., from human clinical specimens. *Int J Syst Bacteriol* 49.
89. **Collins MD, Jones D, Farrow JAE, Kilpper-Bälz R & Schleifer KH** (1984) *Enterococcus avium* nom. rev., comb. nov.; *E. casseliflavus* nom. rev., comb. nov.; *E. durans* nom. rev., comb. nov.; *E. gallinarum* comb. nov.; and *E. malodoratus* sp. nov. *Int J Syst Bacteriol* 34.
90. **Collins MD & Lawson PA** (2000) The genus *Abiotrophia* (Kawamura et al.) is not monophyletic: proposal of *Granulicatella* gen. nov., *Granulicatella adiacens* comb. nov., *Granulicatella elegans* comb. nov. and *Granulicatella balaenopterae* comb. nov. *Int J Syst Evol Microbiol* 50: 365-369.
91. **Collins MD & Shah HN** (1986) Reclassification of *Bacteroides praeacutus* Tissier (Holdeman and Moore) in a new genus, *Tissierella*, as *Tissierella praeacuta* comb. nov. *Int J Syst Bacteriol* 36: 461-463.
92. **Collins MD & Wallbanks S** (1992) Comparative sequence analyses of the 16S rRNA genes of *Lactobacillus minutus, Lactobacillus rimae* and *Streptococcus parvulus*: proposal for the creation of a new genus *Atopobium*. *FEMS Microbiol Lett* 74: 235-240.
93. **Conte MP, Schippa S, Zamboni I, Penta M, Chiarini F, Seganti L, Osborn J, Falconieri P, Borrelli O & Cucchiara S** (2006) Gut-associated bacterial microbiota in paediatric patients with inflammatory bowel disease. *Gut* 55: 1760-1767.
94. **Conway JM, Ingwersen LA & Moshfegh AJ** (2004) Accuracy of dietary recall using the USDA five-step multiple-pass method in men: An observational validation study. *J Am Diet Assoc* 104: 595.
95. **Cooke JV & Keith HR** (1927) A type of urea-splitting bacterium gound in the human intestinal tract. *J Bacteriol* 13: 315-319.
96. **Croft DN** (1970) Body iron loss and cell loss from epithelia. *Proc R Soc Med* 63: 1221-1224.
97. **Croft DN & Cotton PB** (1973) Gastro-intestinal cell loss in man. Its measurement and significance. *Digestion* 8: 144-160.
98. **Cummings JH & Macfarlane GT** (1991) The control and consequences of bacterial fermentation in the human colon. *J Appl Bacteriol* 7: 443-459.
99. **Cummings JH & Macfarlane GT** (1997) Colonic microflora: Nutrition and health. *Nutrition* 13: 476.
100. **Dal Bello F & Hertel C** (2006) Oral cavity as natural reservoir for intestinal lactobacilli. *Syst Appl Microbiol* 29: 69.
101. **Daly K, Sharp RJ & McCarthy AJ** (2000) Development of oligonucleotide probes and PCR primers for detecting phylogenetic subgroups of sulfate-reducing bacteria. *Microbiology* 146: 1693-1705.
102. **de Vries MC** (2006) Analyzing global gene expression of Lactobacillus plantarum in the human gastrointestinal tract. In *Laboratory of Microbiology*. PhD Wageningen: Wageningen University, pp. 160.
103. **de Vries MC, Vaughan EE, Kleerebezem M & de Vos WM** (2004) Optimising single cell activity assessment of *Lactobacillus plantarum* by fluorescent in situ hybridisation as affected by growth. *J Microbiol Meth* 59: 109.
104. **Decroos K, Vanhemmens S, Cattoir S, Boon N & Verstraete W** (2005) Isolation and characterisation of an equol-producing mixed microbial culture from a human faecal sample and its activity under gastrointestinal conditions. *Arch Microbiol* 183: 45-55.
105. **Del Dot T, Osawa R & Stackebrandt E** (1996) *Phascolarctobacterium faecium* gen. nov., spec. nov., a novel taxon of the *Sporomusa* group of bacteria. *Syst Appl Microbiol* 16: 380-384.

106. **Deplancke B, Finster K, Graham WV, Collier CT, Thurmond JE & Gaskins HR** (2003) Gastrointestinal and microbial responses to sulfate-supplemented drinking water in mice. *Exp Biol Med* 228: 424-433.
107. **Deplancke B & Gaskins HR** (2003) Hydrogen sulfide induces serum-independent cell cycle entry in nontransformed rat intestinal epithelial cells. *FASEB J* 17: 1310-1312.
108. **Derrien M** (2007) Mucin utilisation and host interactions of the novel intestinal microbe *Akkermansia muciniphila*. In *Laboratory of Microbiology*. PhD Wageningen: Wageningen University, pp. 164.
109. **Derrien M, Adawi D, Ahrné S, Jeppsson B, Molin G, Osman N, Stsepetova J, Vaughan EE, de Vos WM & Ouwehand AC** (2004) The intestinal mucosa as a habitat of the gut microbiota and a rational target for probiotic functionality and safety. *Microbial Ecol Health Dis* 16: 137-144.
110. **Derrien M, Vaughan EE, Plugge CM & de Vos WM** (2004) *Akkermansia muciniphila* gen. nov., sp. nov., a human intestinal mucin-degrading bacterium. *Int J Syst Evol Microbiol* 54: 1469-1476.
111. **DeSantis TZ, Hugenholtz P, Larsen N, Rojas M, Brodie EL, Keller K, Huber T, Dalevi D, Hu P & Andersen GL** (2006) Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. *Appl Environ Microbiol* 72: 5069-5072.
112. **Dethlefsen L, Eckburg PB, Bik EM & Relman DA** (2006) Assembly of the human intestinal microbiota. *Trends Ecol Evol* 21: 517-523.
113. **Dicks LMT, Fantuzzi L, Gonzalez FC, Du Toit M & Dellaglio F** (1993) *Leuconostoc argentinum* sp. nov., isolated from Argentine raw milk. *Int J Syst Bacteroil* 43: 347-351.
114. **Dilly O, Bloem J, Vos A & Munch JC** (2004) Bacterial diversity in agricultural soils during litter decomposition. *Appl Environ Microbiol* 70: 468-474.
115. **Distaso A** (1911) Sur les microbes protéolytiques de la flore intestinale de l'homme et des animaux. *Zbl Bakt Parasit* 59: 97-103.
116. **Distaso A** (1912) Contribution à l'étude sur l'intoxication intestinale. *Zbl Bakt Parasit* 62.
117. **Dixon WJ** (1950) Analysis of extreme values. *Ann Math Stat* 21: 488-506.
118. **Dolina M & Peduzzi R** (1993) Population genetics of human, animal, and environmental *Yersinia* strains. *Appl Environ Microbiol* 59: 442-450.
119. **Doré J & Sghir A** (1998) Design and evaluation of a 16S rRNA-targeted oligonucleotide probe for specific detection and quantitation of human fecal *Bacteroides* populations. *Syst Appl Microbiol* 21: 65-71.
120. **Downes J, Munson M & Wade WG** (2003) *Dialister invisus* sp. nov., isolated from the human oral cavity. *Int J Syst Evol Microbiol* 53: 1937-1940.
121. **Drancourt M, Bollet C, Carta A & Rousselier P** (2001) Phylogenetic analyses of *Klebsiella* species delineate *Klebsiella* and *Raoultella* gen. nov., with description of *Raoultella ornithinolytica* comb. nov., *Raoultella terrigena* comb. nov. and *Raoultella planticola* comb. nov. *Int J Syst Evol Microbiol* 51: 925-932.
122. **Drasar BS & Barrow PA** (1985) *Intestinal microbiology.* American Society for Microbiology, Washington, D. C.
123. **Drasar BS & Hill MJ** (1974) Composition of the gut flora. *Human intestinal flora*. (Drasar BS & Hill MJ, eds). pp. 26-35. Academic Press Inc. Ltd., London.
124. **Drasar BS & Hill MJ** (1974) The metabolic activities of the gut bacteria. *Human intestinal flora*. (Drasar BS & Hill MJ, eds). pp. 26-35. Academic Press Inc. Ltd., London.
125. **Duncan SH, Aminov RI, Scott KP, Louis P, Stanton TB & Flint HJ** (2006) Proposal of *Roseburia faecis* sp. nov., *Roseburia hominis* sp. nov. and *Roseburia inulinivorans* sp. nov., based on isolates from human faeces. *Int J Syst Evol Microbiol* 56: 2437-2441.
126. **Duncan SH, Hold GL, Barcenilla A, Stewart CS & Flint HJ** (2002) *Roseburia intestinalis* sp. nov., a novel saccharolytic, butyrate-producing bacterium from human faeces. *Int J Syst Evol Microbiol* 52: 1615-1620.
127. **Duncan SH, Hold GL, Harmsen H, Stewart CS & Flint HJ** (2002) Growth requirements and fermentation products of *Fusobacterium prausnitzii*, and a proposal to reclassify it as *Faecalibacterium prausnitzii* gen. nov., comb. nov. *Int J Syst Evol Microbiol* 52: 2141-2146.
128. **Duncan SH, Louis P & Flint HJ** (2004) Lactate-utilizing bacteria, isolated from human feces, that produce butyrate as a major fermentation product. *Appl Environ Microbiol* 70: 5810-5817.
129. **Eckburg PB, Bik EM, Bernstein CN, Purdom E, Dethlefsen L, Sargent M, Gill SR, Nelson KE & Relman DA** (2005) Diversity of the human intestinal microbial flora. *Science* 308: 1635-1638.
130. **Egert M, de Graaf AA, Maathuis A, de Waard P, Plugge CM, Smidt H, Deutz NE, Dijkema C, de Vos WM & Venema K** (2007) Identification of glucose-fermenting bacteria present in an in vitro model of the human intestine by RNA-stable isotope probing. *FEMS Microbiol Ecol* in press.
131. **Egert M, de Graaf AA, Smidt H, de Vos WM & Venema K** (2006) Beyond diversity: functional microbiomics of the human colon. *Trends in Microbiology* 14: 86.
132. **Egert M, Wagner B, Lemke T, Brune A & Friedrich MW** (2003) Microbial community structure in midgut and hindgut of the humus-feeding larva of *Pachnoda ephippiata* (Coleoptera: Scarabaeidae). *Appl Environ Microbiol* 69: 6659-6668.

133. **Eggerth AH** (1935) The Gram-positive non-spore-bearing anaerobic bacilli of human feces. *J Bacteriol* 30: 277-299.
134. **Eggerth AH & Gagnon BH** (1933) The bacteroides of human feces. *J Bacteriol* 25: 389-413.
135. **Elomari M, Coroler L, Verhille S, Izard D & Leclerc H** (1997) *Pseudomonas monteilii* sp. nov., isolated from clinical specimens. *Int J Syst Bacteriol* 47: 846-852.
136. **Emrich SJ, Lowe M & Delcher AL** (2003) PROBEmer: a web-based software tool for selecting optimal DNA oligos. *Nucl Acids Res* 31: 3746-3750.
137. **Engberg J, On SLW, Harrington CS & Gerner-Smidt P** (2000) Prevalence of *Campylobacter*, *Arcobacter*, *Helicobacter*, and *Sutterella* spp. in human fecal samples as estimated by a reevaluation of isolation methods for campylobacters. *J Clin Microbiol* 38: 286-291.
138. **Ewing WH** (1962) The tribe *Proteae*: its nomenclature and taxonomy. *Int Bull Bacteriol Nomencl Taxon*: 93-102.
139. **Ewing WH, McWhorter AC, Escobar MR & Lubin AH** (1965) *Edwardsiella*, a new genus of *Enterobacteriaceae* based on a new species *E. tarda*. *Int Bull Bacteriol Nomencl Taxon* 15: 33-38.
140. **Ezaki T, Kawamura Y, Li N, Li ZY, Zhao L & Shu S** (2001) Proposal of the genera *Anaerococcus* gen. nov., *Peptoniphilus* gen. nov. and *Gallicola* gen. nov. for members of the genus *Peptostreptococcus*. *Int J Syst Evol Microbiol* 51: 1521-1528.
141. **Ezaki T, Li N, Hashimoto Y, Miura H & Yamamoto H** (1994) 16S ribosomal DNA sequences of anaerobic cocci and proposal of *Ruminococcus hansenii* comb. nov. and *Ruminococcus productus* comb. nov. *Int J Syst Bacteriol* 1.
142. **Ezaki T, Liu S-L, Hashimoto Y & Yabuuchi E** (1990) *Peptostreptococcus hydrogenalis* sp. nov. from human fecal and vaginal flora. *Int J Syst Bacteriol* 40: 305-306.
143. **Fanaro S, Chierici R, Guerrini P & Vigi V** (2003) Intestinal microflora in early infancy: composition and development. *Acta Paediatr* 92: 48-55.
144. **Farmer JJ, 3rd, Fanning GR, Davis BR, O'Hara CM, Riddle C, Hickman-Brenner FW, Asbury MA, Lowery VA & Brenner DJ** (1985) *Escherichia fergusonii* and *Enterobacter taylorae*, two new species of *Enterobacteriaceae* isolated from clinical specimens. *J Clin Microbiol* 21: 77-81.
145. **Faure M, Moënnoz D, Montigon F, Fay LB, Breuillé D, Finot PA, Olivier B & Boza J** (2002) Development of a rapid and convenient method to purify mucins and determine their *in vivo* synthesis rate in rats. *Anal Biochem* 307: 244-251.
146. **Favier CF, Vaughan EE, de Vos WM & Akkermans ADL** (2002) Molecular monitoring of succession of bacterial communities in human neonates. *Appl Environ Microbiol* 68: 219-226.
147. **Felis GE, Torriani S & Dellaglio F** (2004) The status of species *Lactobacillus rogosae* Holdeman and Moore 1974. Request for an oppinon. *Int J Syst Evol Microbiol* 54: 1903-1904.
148. **Fernandez AS, Hashsham SA, Dollhopf SL, Raskin L, Glagoleva O, Dazzo FB, Hickey RF, Criddle CS & Tiedje JM** (2000) Flexible community structure correlates with stable community function in methanogenic bioreactor communities perturbed by glucose. *Appl Environ Microbiol* 66: 4058-4067.
149. **Finegold SM** (1969) Intestinal bacteria. The role they play in normal physiology, pathologic physiology, and infection. *Calif Med* 110: 455-459.
150. **Finegold SM, Howard RA & Vera LS** (1974) Effect of diet on human intestinal fecal flora: comparison of Japanese and American diets. *Am J Clin Nutr* 27: 1456-1469.
151. **Finegold SM, Lawson PA, Vaisanen ML, Molitoris DR, Song Y, Liu C & Collins MD** (2004) *Anaerofustis stercorihominis* gen. nov., sp. nov., from human faeces. *Anaerobe* 10: 41-45.
152. **Finegold SM, Vaisanen M, Molitoris DR, Tomzynski TJ, Song Y, Liu C, Collins MD & Lawson PA** (2003) *Cetobacterium somerae* sp. nov. from human feces and emended description of the genus *Cetobacterium*. *Syst Appl Microbiol* 26: 177-181.
153. **Finegold SM, Vera LS, Paul TS, Harvey AE, Shirley ML & Ronald LP** (1977) Fecal microbial flora in Seventh Day Adventist populations and control subjects. *Am J Clin Nutr* 30: 1781-1792.
154. **Fite A, Macfarlane GT, Cummings JH, Hopkins MJ, Kong SC, Furrie E & Macfarlane S** (2004) Identification and quantitation of mucosal and faecal desulfovibrios using real time polymerase chain reaction. *Gut* 53: 523-529.
155. **Flügge C** (1886) *Die Mikroorganismen.*, Leipzig.
156. **Fogarty LR & Voytek MA** (2005) Comparison of *Bacteroides-Prevotella* 16S rRNA genetic markers for fecal samples from different animal species. *Appl Environ Microbiol* 71: 5999-6007.
157. **Fogel GB, Collins CR, Li J & Brunk CF** (1999) Prokaryotic genome size and SSU rDNA copy number: estimation of microbial relative abundance from a mixed population. *Microb Ecol* 38: 93-113.
158. **Fox JG, Chien CC, Dewhirst FE, Paster BJ, Shen Z, Melito PL, Woodward DL & Rodgers FG** (2000) *Helicobacter canadensis* sp. nov. isolated from humans with diarrhea as an example of an emerging pathogen. *J Clin Microbiol* 38: 2546-2549.

159. **Franks AH, Harmsen HJM, Raangs GC, Jansen GJ, Schut F & Welling GW** (1998) Variations of bacterial populations in human feces measured by fluorescent in situ hybridization with group-specific 16S rRNA-targeted oligonucleotide probes. *Appl Environ Microbiol* 64: 3336-3345.
160. **Freitas RA** (1999) *Navigational alimentography.* Landes Bioscience, Georgetown.
161. **Freitas RA** (1999) *Overall cellular structure.* Landes Bioscience, Georgetown.
162. **Fromin N, Hamelin J, STarnawski S, Roesti D, Jourdain-Miserez K, Forestier N, Teyssier-Cuvelle S, Gillet F, Aragno M & Rossi P** (2002) Statistical analysis of denturing gel electrophoresis (DGE) fingerprinting patterns. *Environ Microbiol* 4: 634-643.
163. **Furrie E** (2006) A molecular revolution in the study of intestinal microflora. *Gut* 55: 141-143.
164. **Garner MR, Flint JF & Russell JB** (2002) *Allisonella histaminiformans* gen. nov., sp. nov. a novel bacterium that produces histamine, utilizes histidine as its sole energy source, and could play a role in bovine and equine laminitis. *Syst Appl Microbiol* 25: 298-506.
165. **Garrity GM & Holt JG** (2001) *Bergey's Manual of Systematic Bacteriology, 2nd ed.* Springer-Verlag, New York.
166. **Gelsomino R, Vancanneyt M, Cogan TM & Swings J** (2003) Effect of raw-milk cheese consumption on the enterococcal flora of human feces. *Appl Environ Microbiol* 69: 312-319.
167. **Gentry TJ, Wickham GS, Schadt CW, He Z & Zhou J** (2006) Microarray applications in microbial ecology research. *Microb Ecol* 52: 159-175.
168. **Geypens B, Claus D, Evenepoel P, Hiele M, Maes B, Peeters M, Rutgeerts P & Ghoos Y** (1997) Influence of dietary protein supplements on the formation of bacterial metabolites in the colon. *Gut* 41: 70-76.
169. **Gibson GR, Cummings JH & Macfarlane GT** (1988) Use of a three-stage continuous culture system to study the effect of mucin on dissimilatory sulfate reduction and methanogenesis by mixed populations of human gut bacteria. *Appl Environ Microbiol* 54: 2750-2755.
170. **Gibson GR, Cummings JH & Macfarlane GT** (1991) Growth and activities of sulphate reducing bacteria in gut contents of healthy subjects and patients with ulcerative colitis. *FEMS Microbiol Ecol* 86: 103-112.
171. **Gibson GR, Macfarlane GT & Cummings JH** (1988) Occurrence of sulphate-reducing bacteria in human faeces and the relationship of dissimilatory sulphate reduction to methanogenesis in the large gut. *J Appl Bacteroil* 65: 103-111.
172. **Gill SR, Pop M, Deboy RT, Eckburg PB, Turnbaugh PJ, Samuel BS, Gordon JI, Relman DA, Fraser-Liggett CM & Nelson KE** (2006) Metagenomic analysis of the human distal gut microbiome. *Science* 312: 1355-1359.
173. **Giordano C, de Pascale C, Balestrieri C, Cittadini D & Crescenzi A** (1968) Incorporation of urea ^{15}N in amino acids of patients with chronic renal failure on low nitrogen diet. *Am J Clin Nutr* 21: 394-404.
174. **Giuliano M, Barza M, Jacobus NV & Gorbach SL** (1987) Effect of broad-spectrum parenteral antibiotics on composition of intestinal microflora of humans. *Antimicrob Agents Chemother* 31: 202-206.
175. **Gmur R, Munson MA & Wade WG** (2006) Genotypic and phenotypic characterization of fusobacteria from Chinese and European patients with inflammatory periodontal diseases. *Syst Appl Microbiol* 29: 120.
176. **Goossens H, Pot B, Vlaes L, Van den Borre C, Van den Abbeele R, Van Naelten C, Levy J, Cogniau H, Marbehant P & Verhoef J** (1990) Characterization and description of "*Campylobacter upsaliensis*" isolated from human feces. *J Clin Microbiol* 28: 1039-1046.
177. **Gophna U, Sommerfeld K, Gophna S, Doolittle WF & Veldhuyzen van Zanten SJO** (2006) Differences between tissue-associated intestinal microfloras of patients with Crohn's disease and ulcerative colitis. *J Clin Microbiol* 44: 4136-4141.
178. **Greenfield C, Aguilar Ramirez JR, Pounder RE, Williams T, Danvers M, Marper SR & Noone P** (1983) *Clostridium difficile* and inflammatory bowel disease. *Gut* 24: 713-717.
179. **Grimont PA, Grimont F, Farmer JJ, 3rd & Asbury MA** (1981) *Cedecea davisae* gen. nov., sp. nov. and *Cedecea lapagei* sp. nov., new *Enterobacteriaceae* from clinical specimens. *Int J Syst Bacteriol* 31: 317-326.
180. **Grimont PAD & Grimont F** (1978) The Genus *Serratia*. *Ann Rev Microbiol* 32: 221-248.
181. **Gronlund MM, Lentonen O-P, Eerola E & Kero P** (1999) Fecal microflora in healthy infants born by different methods of delivery: permanent changes in intestinal flora after cesarian section. *J Pediatr Gastroenterol Nutr* 28: 19-25.
182. **Guschin DY, Mobarry BK, Proudnikov D, Stahl DA, Rittmann BE & Mirzabekov AD** (1997) Oligonucleotide microchips as genosensors for determinative and environmental studies in microbiology. *Appl Environ Microbiol* 63: 2397-2402.
183. **Guyton AC** (1985) *Anatomy and Physiology.* Saunders College Pub, Philadelphia.
184. **Gwee KA** (2005) Irritable bowel syndrome in developing countries - a disorder of civilization or colonization? *Neurogastroent Motil* 27: 317-324.
185. **Haenel H** (1970) Human normal and abnormal gastrointestinal flora. *Am J Clin Nutr* 23: 1433-1439.

186. **Haffajee AD, Japlit M, Bogren A, Kent RLJ, Goodson JM & Socransky SS** (2005) Differences in the subgingival microbiota of Swedish and USA subjects who were periodontally healthy or exhibited minimal periodontal disease. *J Clin Periodontol* 32: 33-39.
187. **Halfvarson J, Bodin L, Tysk C, Lindberg E & Jarnerot G** (2003) Inflammatory bowel disease in a Swedish twin cohort: a long-term follow-up of concordance and clinical characteristics. *Gastroenterology* 124: 1767.
188. **Haller D** (2006) Intestinal epithelial cell signalling and host-derived negative regulators under chronic inflammation: to be or not to be activated determines the balance towards commensal bacteria. *Neurogastroenterol Motil* 18: 184-199.
189. **Häne BG, Jäger K & Drexler HG** (1993) The Pearson product-moment correlation coefficient is better suited for identification of DNA fingerprint profiles than band matching algorithms. *Electrophoresis* 14.
190. **Hauschild AHW & Holdeman LV** (1974) *Clostridium celatum* sp. nov., isolated from normal human feces. *Int J Syst Bacteriol* 24: 478-481.
191. **Hayashi H, Sakamoto M & Benno Y** (2002) Fecal microbial diversity in a strict vegetarian as determined by molcular analysis and cultivation. *Microbiol Immunol* 46: 819-831.
192. **Hayashi H, Sakamoto M & Benno Y** (2002) Phylogenetic analysis of the human gut microbiota using 16S rDNA clone libraties and strictly anaerobic culture-based methods. *Microbiol Immunol* 46: 535-548.
193. **Hayashi H, Sakamoto M & Benno Y** (2003) Molecular analysis of fecal microbiota in elderly individuals using 16S rDNA library and T-RFLP. *Microbiol Immunol* 47: 557-570.
194. **Heilig HG, Zoetendal EG, Vaughan EE, Marteau P, Akkermans ADL & de Vos WM** (2002) Molecular diversity of *Lactobacillus* spp. and other lactic acid bacteria in the human intestine as determined by specific amplification of 16S ribosomal DNA. *Appl Environ Microbiol* 68: 114-123.
195. **Held GA, Grinstein G & Tu Y** (2003) Modeling of DNA microarray data by using physical properties of hybridization. *Proc Natl Acad Sci* 100: 7575-7580.
196. **Heuer H, Hartung K, Wieland G, Kramer I & Smalla K** (1999) Polynuclotide probes that target a hypervariable region of 16S rRNA genes to identify bacterial isolates corresponding to bands of community fingerprints. *Appl Environ Microbiol* 65: 1045-1049.
197. **Heyndrickx M, Vandemeulebroecke K, Scheldeman P, Kersters K, de Vos P, Logan NA, Aziz AM, Ali N & Berkeley RC** (1996) A polyphasic reassessment of the genus *Paenibacillus*, reclassification of *Bacillus lautus* (Nakamura 1984) as *Paenibacillus lautus* comb. nov. and of *Bacillus peoriae* (Montefusco et al. 1993) as *Paenibacillus peoriae* comb. nov., and emended descriptions of *P. lautus* and of *P. peoriae*. *Int J Syst Bacteriol* 46: 988-1003.
198. **Hickman-Brenner FW, Huntley-Carter GP, Fanning GR, Brenner DJ & Farmer JJ, 3rd** (1985) *Koserella trabulsii*, a new genus and species of *Enterobacteriaceae* formerly known as Enteric Group 45. *J Clin Microbiol* 21: 39-42.
199. **Hickman-Brenner FW, Huntley-Carter GP, Saitoh Y, Steigerwalt AG, Farmer JJ, 3rd & Brenner DJ** (1984) *Moellerella wisconsensis*, a new genus and species of *Enterobacteriaceae* found in human stool specimens. *J Clin Microbiol* 19: 460-463.
200. **Hickman-Brenner FW, MacDonald KL, Steigerwalt AG, Fanning GR, Brenner DJ & Farmer JJ, 3rd** (1987) *Aeromonas veronii*, a new ornithine decarboxylase-positive species that may cause diarrhea. *J Clin Microbiol* 25: 900-906.
201. **Hickman-Brenner FW, Vohra MP, Huntley-Carter GP, Fanning GR, Lowery VA, Brenner DJ & Farmer JJ, 3rd** (1985) *Leminorella*, a new genus of *Enterobacteriaceae*: identification of *Leminorella grimontii* sp. nov. and *Leminorella richardii* sp. nov. found in clinical specimens. *J Clin Microbiol* 21: 234-239.
202. **Hickman FW** (1982) Identification of *Vibrio hollisae* sp. nov. from patients with diarrhea. *J Clin Microbiol* 15: 395-401.
203. **Himelbloom BH & Canale-Parola E** (1989) *Clostridium methylpentosum* sp. nov.: a ring-shaped intestinal bacterium that ferments only methylpentoses and pentoses. *Arch Microbiol* 151: 287-293.
204. **Hold GL, Pryde SE, Russell VJ, Furrie E & Flint HJ** (2002) Assessment of microbial diversity in human colonic samples by 16S rDNA sequence analysis. *FEMS Microbiol Ecol* 39: 33-39.
205. **Hold GL, Schwiertz A, Aminov RI, Blaut M & Flint HJ** (2003) Oligonucleotide probes that detect quantitatively significant groups of butyrate-producing bacteria in human feces. *Appl Environ Microbiol* 69: 4320-4324.
206. **Holdeman LV, Good IJ & Moore WEC** (1976) Human fecal flora: variation in bacterial composition within individuals and a possible effect of emotional stress. *Appl Environ Microbiol* 31: 359-375.
207. **Holdeman LV & Moore WEC** (1974) New Genus, *Coprococcus*, twelve new species, and emended description of four previously described species of bacteria from human feces. *Int J Syst Bacteriol* 24: 260-277.

208. **Hollis DG, Hickman FW, Fanning GR, Farmer JJ, 3rd, Weaver RE & Brenner DJ** (1981) *Tatumella ptyseos* gen. nov., sp. nov., a member of the family *Enterobacteriaceae* found in clinical specimens. *J Clin Microbiol* 14: 79-88.
209. **Holmstrom K, Collins MD, Moller T, Falsen E & Lawson PA** (2004) *Subdoligranulum variabile* gen. nov., sp. nov. from human feces. *Anaerobe* 10: 197-203.
210. **Hooper LV & Gordon JI** (2001) Commensal host-bacterial relationships in the gut. *Science* 292: 1115-1118.
211. **Hooper LV, Xu J, Falk PG, Midtvedt T & Gordon JI** (1999) A molecular sensor that allows a gut commensal to control its nutrient foundation in a competitive ecosystem. *Proc Natl Acad Sci* 96: 9833-9838.
212. **Hope ACA** (1968) A simplified Monte Carlo significance test procedure. *J R Stat Soc* 30B: 582-598.
213. **Hopkins MJ, Englyst HN, Macfarlane S, Furrie E, Macfarlane GT & McBain AJ** (2003) Degradation of cross-linked and non-cross-linked arabinoxylans by the intestinal microbiota in children. *Appl Environ Microbiol* 69: 6354-6360.
214. **Hopkins MJ, Sharp R & Macfarlane GT** (2001) Age and disease related changes in intestinal bacterial populations assessed by cell culture, 16S rRNA abundance, and community cellular fatty acid profiles. *Gut* 48: 198-205.
215. **Horz HP, Vianna ME, Gomes BPFA & Conrads G** (2005) Evaluation of universal probes and primer sets for assessing total bacterial load in clinical samples: general implications and practical use in endodontic antimicrobial therapy. *J Clin Microbiol* 43: 5332–5337.
216. **Hugenholtz P, Pitulle C, Hershberger KL & Pace NR** (1998) Novel division level bacterial diversity in a Yellowstone hot spring. *J Bacteriol* 180: 366-376.
217. **Hugenholtz P, Tyson GW, Webb RI, Wagner AM & Blackall LL** (2001) Investigation of candidate division TM7, a recently recognized major lineage of the domain bacteria with no known pure-culture representatives. *Appl Environ Microbiol* 67: 411-419.
218. **Hughes JB, Hellmann JJ, Ricketts TH & Bohannan BJM** (2001) Counting the uncountable: statistical approaches to estimating microbial diversity. *Appl Environ Microbiol* 67: 4399-4406.
219. **Hungate RE** (1947) Studies on cellulose fermentation. The culture and isolation of cellulose decomposing bacteria from the rumen of cattle. *J Bacteriol* 53: 631-645.
220. **Hungate RE** (1969) *Methods in microbiology.* Academic press, London.
221. **Jackson AA** (1995) Salvage of urea-nitrogen and protein requirements. *Proc Nutr Soc* 54: 535-547.
222. **Jensen NS & Canale-Parola E** (1986) *Bacteroides pectinophilus* sp. nov. and *Bacteroides galacturonicus* sp. nov.: two pectinolytic bacteria from the human intestinal tract. *Appl Environ Microbiol* 52: 880-887.
223. **Jian W & Dong X** (2002) Transfer of *Bifidobacterium inopinatum* and *Bifidobacterium denticolens* to *Scardovia inopinata* gen. nov., comb. nov., and *Parascardovia denticolens* gen. nov., comb. nov., respectively. *Int J Syst Evol Microbiol* 52: 809-812.
224. **Johnson AS, Tarr CL, Brown BH, Jr., Birkhead KM & Farmer JJ, 3rd** (2005) First case of septicemia due to a strain belonging to enteric group 58 (*Enterobacteriaceae*) and its designation as *Averyella dalhousiensis* gen. nov., sp. nov., based on analysis of strains from 20 additional cases. *J Clin Microbiol* 43: 5195-5201.
225. **Johnson JL, Moore WEC & Moore LVH** (1986) *Bacteroides caccae* sp. nov., *Bacteroides merdae* sp. nov., and *Bacteroides stercoris* sp. nov. isolated from human feces. *Int J Syst Bacteriol* 36: 499-501.
226. **Jonard PP, Rambaud JC, Dive C, Vaerman JP, Galian A & Delacroix DL** (1984) Secretion of immunoglobulins and plasma proteins from the jejunal mucosa. Transport rate and origin of polymeric immunoglobulin A. *J Clin Invest* 74: 525-535.
227. **Jowett SL, Seal CJ, Pearce MS, Phillips E, Gregory W, Barton JR & Welfare MR** (2004) Influence of dietary factors on the clinical course of ulcerative colitis: a prospective cohort study. *Gut* 53: 1479-1484.
228. **Justesen T, Nielsen OH & Krasilnikoff PA** (1984) Normal cultivable microflora in upper jejunal fluid in children without gastrointestinal disorders. *J Pediatr Gastroenterol Nutr* 3: 683-686.
229. **Kageyama A & Benno Y** (2000) *Catenibacterium mitsuokai* gen. nov., sp. nov., a Gram-positive anaerobic bacterium isolated from human faeces. *Int J Syst Evol Microbiol* 50: 1595-1599.
230. **Kageyama A & Benno Y** (2000) *Coprobacillus catenaformis* gen. nov., sp. nov., a new genus and species isolated from human feces. *Microbiol Immunol* 44: 23-28.
231. **Kageyama A & Benno Y** (2000) Emendation of genus *Collinsella* and proposal of *Collinsella stercoris* sp. nov. and *Collinsella intestinalis* sp. nov. *Int J Syst Evol Microbiol* 50: 1767-1774.
232. **Kageyama A & Benno Y** (2000) Phylogenic and phenotypic characterization of some *Eubacterium*-like isolates from human feces: description of *Solobacterium moorei* gen. nov., sp. nov. *Microbiol Immunol* 44: 223-227.
233. **Kageyama A, Benno Y & Nakase T** (1999) Phylogenetic and phenotypic evidence for the transfer of *Eubacterium aerofaciens* to the genus *Collinsella* as *Collinsella aerofaciens* gen. nov., comb. nov. *Int J Syst Bacteriol* 49: 557-565.

234. **Kageyama A, Benno Y & Nakase T** (1999) Phylogenetic evidence for the transfer of *Eubacterium lentum* to the genus *Eggerthella* as *Eggerthella lenta* gen. nov., comb. nov. *Int J Syst Bacteriol* 49: 1725-1732.
235. **Kajander K, Hatakka K, Poussa T, Färkkilä M & Korpela R** (2005) A probiotic mixture alleviates symptoms in irritable bowel syndrome patients: a controlled 6-month intervention. *Aliment Pharmacol Ther* 22: 387-394.
236. **Kajander K, Kyrönpalo S, Rasmussen M, Myllyluoma E, Ristimäki A, Vapaatalo H & Korpela R** (2006) A probiotic combination effectively reduces abdominal pain and bloating in irritable bowel syndrome [abstract]. 27.-29.9.2006, Tallinn, Estonia. In *SSA GutImpact: the 2nd Platform Meeting on Foods for Intestinal health* Tallinn, Estonia, pp. 47.
237. **Kalliomaki M, Kirjavainen P, Eerola E, Kero P, Salminen S & Isolauri E** (2001) Distinct patterns of neonatal gut microflora in infants in whom atopy was and was not developing. *J Allergy Clin Immunol* 107: 129-134.
238. **Kalmbach S, Manz W, Wecke J & Szewzyk U** (1999) *Aquabacterium* gen. nov., with description of *Aquabacterium citratiphilum* sp. nov., *Aquabacterium parvum* sp. nov. and *Aquabacterium commune* sp. nov., three *in situ* dominant bacterial species from the Berlin drinking water system. *Int J Syst Bacteriol* 49.
239. **Kanazawa K, Konishi F, Mitsuoka T, Terada A, Itoh K, Narushima S, Kumemura M & Kimura H** (1996) Factors influencing the development of sigmoid colon cancer. Bacteriologic and biochemical studies. *Cancer* 77: 1701-1706.
240. **Kane MD, Poulsen LK & Stahl DA** (1993) Monitoring the enrichment and isolation of sulfate-reducing bacteria by using oligonucleotide hybridization probes designed from environmentally derived 16S rRNA sequences. *Appl Environ Microbiol* 59: 682-686.
241. **Kaneuchi C, Miyazato, T., Shinjo, T., Mitsuoka, T.** (1979) Taxonomic study of helically coiled, sporeforming anaerobes isolated from the intestines of humans and other animals: *Clostridium cocleatum* sp. nov. and *Clostridium spiroforme* sp. nov. *Int J Syst Bacteriol* 29: 1-12.
242. **Kaneuchi C, Watanabe K, Terada A, Benno Y & Mitsuoka T** (1976) Taxonomic study of *Bacteroides clostridiiformis* subsp. *clostridiiformis* (Burri and Ankersmit) Holdeman and Moore and of related organisms: Proposal of *Clostridium clostridiiformis* (Burri and Ankersmit) comb. nov. and *Clostridium symbiosum* (Stevens) comb. nov. *Int J Syst Bacteriol* 26: 195-204.
243. **Kaufmann P, Pfefferkorn A, Teuber M & Meile L** (1997) Identification and quantification of *Bifidobacterium* species isolated from food with genus-specific 16S rRNA-targeted probes by colony hybridization and PCR. *Appl Environ Microbiol* 63: 1268-1273.
244. **Kazor CE, Mitchell PM, Lee AM, Stokes LN, Loesche WJ, Dewhirst FE & Paster BJ** (2003) Diversity of bacterial populations on the tongue dorsa of patients with halitosis and healthy patients. *J Clin Microbiol* 41: 558-563.
245. **Kenzaka T, Tamaki S, Yamaguchi N, Tani K & Nasu M** (2005) Recognition of individual genes in diverse microorganisms by cycling primed *in situ* amplification. *Appl Environ Microbiol* 71: 7236-7244.
246. **Kilpper-Balz R & Schleifer KH** (1988) Transfer of *Streptococcus morbillorum* to the genus *Gemella* as *Gemella morbillorum* comb. nov. *Int J Syst Bacteriol* 38: 442-443.
247. **King TS, Elia M & Hunter JO** (1998) Abnormal colonic fermentation in irritable bowel syndrome. *The Lancet* 352: 1187.
248. **Kirjavainen PV, Apostolou E, Arvola T, Salminen SJ, Gibson GR & Isolauri E** (2001) Characterizing the composition of intestinal microflora as a prospective treatment target in infant allergic disease. *FEMS Immunology and Medical Microbiology* 32: 1.
249. **Kitahara M, Sakamoto M, Ike M, Sakata S & Benno Y** (2005) *Bacteroides plebeius* sp. nov. and *Bacteroides coprocola* sp. nov., isolated from human faeces. *Int J Syst Evol Microbiol* 55: 2143-2147.
250. **Kitahara M, Takamine F, Imamura T & Benno Y** (2000) Assignment of *Eubacterium* sp. VPI 12708 and related strains with high bile acid 7alpha-dehydroxylating activity to *Clostridium scindens* and proposal of *Clostridium hylemonae* sp. nov., isolated from human faeces. *Int J Syst Evol Microbiol* 50: 971-978.
251. **Kitahara M, Takamine F, Imamura T & Benno Y** (2001) *Clostridium hiranonis* sp. nov., a human intestinal bacterium with bile acid 7alpha-dehydroxylating activity. *Int J Syst Evol Microbiol* 51: 39-44.
252. **Klaassens ES, de Vos WM & Vaughan EE** (2006) A metaproteomics approach to study the functionality of the microbiota in the human infant gastrointestinal tract. *Appl Environ Microbiol*: AEM.01921-01906.
253. **Klappenbach JA, Dunbar JM & Schmidt TM** (2000) rRNA operon copy number reflects ecological strategies of bacteria. *Appl Environ Microbiol* 66: 1328-1333.
254. **Klappenbach JA, Saxman PR, Cole JR & Schmidt TM** (2001) rrndb: the ribosomal RNA operon copy number database. *Nucl Acids Res* 29: 181-184.
255. **Kleessen B** (2000) Culture-based knowledge on biodiversity, development and stability of human gastrointestinal microflora. *Microb Ecol Health Dis* 12: 53-63.

256. **Knorr M** (1922) Über die fusispirillare Symbiose, die Gattung *Fusobacterium* (K. B. Lehman) und *Spirillum* sputigenum. Die Gattung *Fusobacterium*. *Zbl Bakt Parasit* 87: 536-545.
257. **Koide AA** (2000) Quantitative analysis of bowel gas using plain abdominal radiograph in patients with irritable bowel syndrome. *Am J Gastroenterol* 95: 1735-1741.
258. **Konstantinov SR, Smidt H & Akkermans ADL** (2005) Ecology and activity of clostridia in the intestine of mammals. *Handbook of Clostridia*. (Dürre P, eds)CRC Press, Boca Raton, Fl.
259. **Kumar A & Sharma VK** (1978) *Enterobacteria* of emerging pathogenic significance from clinical cases in man and animals and detection of toads and wall lizards as their reservoirs. *Antonie Van Leeuwenhoek* 44: 219-228.
260. **Lambert-Zechovsky N, Bingen E, Aujard Y & Mathieu H** (1985) Impact of cefotaxime on the fecal flora in children. *Infection* 13: S140-144.
261. **Lamendella R, Domingo JW, Oerther DB, Vogel JR & Stoeckel DM** (2006) Assessment of fecal pollution sources in a small northern-plains watershed using PCR and phylogenetic analyses of Bacteroidetes 16S rRNA gene. *FEMS Microbiol Ecol* In press.
262. **Lane DJ** (1991) 16S/23S rRNA sequencing. *Nucleic acid techniques in bacterial systematics*. (Stackebrandt E & Goodfellow M, eds)J. Wiley & Sons, Chichester, U. K.
263. **Larkin JE, Frank BC, Gavras H, Sultana R & Quackenbush J** (2005) Independence and reproducibility across microarray platforms. *Nat Methods* 2: 337-344.
264. **Lauer E** (1990) *Bifidobacterium gallicum* sp. nov. isolated from human faeces. *Int J Syst Bacteriol* 40: 100-102.
265. **Lauer E, Helming C & Kandler O** (1980) Heterogeneity of the species *Lactobacillus acidophilus* (Moro) Hansen and Moquot as revealed by biochemical characteristics and DNA-DNA hybridisation. *Zentralbl Bakteriol Hyg* C1: 150-168.
266. **Lawson AJ, On SLW, Logan JMJ & Stanley J** (2001) *Campylobacter hominis* sp. nov., from the human gastrointestinal tract. *Int J Syst Evol Microbiol* 51: 651-660.
267. **Lawson PA, Song Y, Liu C, Molitoris DR, Vaisanen ML, Collins MD & Finegold SM** (2004) *Anaerotruncus colihominis* gen. nov., sp. nov., from human faeces. *Int J Syst Appl Microbiol* 54: 413-417.
268. **Lay C, Rigottier-Gois L, Holmstrom K, Rajilic M, Vaughan EE, de Vos WM, Collins MD, Thiel R, Namsolleck P, Blaut M & Doré J** (2005) Colonic microbiota signatures across five Northern European countries. *Appl Environ Microbiol* 71: 4153-4155.
269. **Lee JV, Shread P, Furniss AL & Bryant TN** (1981) Taxonomy and description of *Vibrio fluvialis* sp. nov. (synonym group F vibrios, group EF6). *J Appl Bacteriol* 50: 73-94.
270. **Leitch EC, Walker AW, Duncan SH, Holtrop G & Flint HJ** (2007) Selective colonization of insoluble substrates by human faecal bacteria. *Environ Microbiol* 9: 667-679.
271. **Lepš J & Šmilauer P** (2003) *Multivariate analysis of ecological data using CANOCO.* Cambridge University Press, Cambridge.
272. **Lerche M & Reuter G** (1961) Isolierung und Differenzierung anaerober Lacotbacilleae aus Darm erwachsener Menschen. *Zentralbl Bakteriol* 182: 324-356.
273. **Levine J, Ellis CJ, Furne JK, Springfield J & Levitt MD** (1998) Fecal hydrogen sufide production in ulcerative colitis. *Am J Gastroenterol* 93: 83-87.
274. **Ley RE, Peterson DA & Gordon JI** (2006) Ecological and evolutionary forces shaping microbial diversity in the human intestine. *Cell* 124: 837.
275. **Ley RE, Turnbaugh PJ, Klein S & Gordon JI** (2006) Microbial ecology: Human gut microbes associated with obesity. *Nature* 444: 1022-1023.
276. **Li X, He Z & Zhou J** (2005) Selection of optimal oligonucleotide probes for microarrays using multiple criteria, global alignment and parameter estimation. *Nucl Acids Res* 33: 6114-6123.
277. **Lichtenberger LM** (1995) The hydrophobic barrier properties of gastrointestinal mucus. *Annu Rev Physiol* 57: 565-583.
278. **Littel KJ & Hartman PA** (1983) Fluorogenic selective and differential medium for isolation of fecal streptococci. *Appl Environ Microbiol* 45: 622-627.
279. **Lloyd-Still JD & Shwachman H** (1975) Duodenal microflora: a prospective study in pediatric gastrointestinal disorders. *Am J Dig Dis* 20: 708-715.
280. **Lodinova-Zadnikova R, Prokesova L, Tlaskalova H, Kocourkova I, Zizkam J & Stranak Z** (2004) Influence of oral colonization with probiotic *E. coli* strain after birth on frequency of recurrent infections, allergy and development of some immunologic parameters. Long-term studies. *Ceska Gynekol* 69: 91-97.
281. **Loubinoux J, Bronowicki J-P, Pereira IAC, Mougenel J-L & Le Faou A** (2002) Sulfate-reducing bacteria in human feces and their association with inflammatory bowel diseases. *FEMS Microbiol Ecol* 40.
282. **Loy A** (2003) DNA Microarray technology for biodiversity inventories of sulfate reducing prokaryotes. In *Abteilung Mikrobielle Ökologie, Institut für Ökologie und Naturschutz* Wien: Universität Wien, pp. 34, 35.

283. **Loy A & Bodrossy L** (2006) Highly parallel microbial diagnostics using oligonucleotide microarrays. *Clin Chim Acta* 363: 106.
284. **Loy A, Lehner A, Lee N, Adamczyk J, Meier H, Ernst J, Schleifer K-H & Wagner M** (2002) Oligonucleotide microarray for 16S rRNA gene-based detection of all recognized lineages of sulfate-reducing prokaryotes in the environment. *Appl Environ Microbiol* 68: 5064-5081.
285. **Ludwig W & Klenk HP** (2001) Overview: A phylogenetic backbone and taxonomic framework for procaryotic systematics. *Bergey's Manual of Systematic Bacteriology*. Vol. 1. (Garrity GM, eds). pp. 49–66. Springer, New York.
286. **Ludwig W, Strunk O, Westram R, Richter L, Meier H, Yadhukumar, Buchner A, Lai T, Steppi S, Jobb G, Forster W, Brettske I, Gerber S, Ginhart AW, Gross O, Grumann S, Hermann S, Jost R, Konig A, Liss T, Lussmann R, May M, Nonhoff B, Reichel B, Strehlow R, Stamatakis A, Stuckmann N, Vilbig A, Lenke M, Ludwig T, Bode A & Schleifer K-H** (2004) ARB: a software environment for sequence data. *Nucl Acids Res* 32: 1363-1371.
287. **MacCallum WG & Hastings TW** (1899) A case of acute endocarditis caused by *Micrococcus zymogenes* (nov. spec.), with a description of the microorganism. *J Exp Med* 4.
288. **Macfarlane GT, Cummings JH & Allison C** (1986) Protein degradation by human intestinal bacteria. *J Gen Microbiol* 132: 1647-1656.
289. **Macfarlane GT, Gibson GR & Cummings JH** (1992) Comparison of fermentation reactions in different regions of the human colon. *J Appl Bacteriol* 72: 57-64.
290. **Macfarlane GT, Gibson GR & Macfarlane S** (1994) Short chain fatty acid and lactate production by human intestinal bacteria grown in batch and continuous culture. *Short chain fatty acids*. (Binder HJ, Cummings JH & Soergel KH, eds)Kluwer Publishing, London.
291. **Macfarlane GT & Macfarlane S** (2007) Models for intestinal fermentation: association between food components, delivery systems, bioavailability and functional interactions in the gut. *Curr Opin Biotechnol*.
292. **Macfarlane GT, Macfarlane S & Gibson GR** (1998) Validation of a three-stage compound continuous culture system for investigating the effect of retention time on the ecology and metabolism of bacteria in the human colon. *Microb Ecol* 35: 180-187.
293. **Macfarlane S, Furrie E, Cummings JH & Macfarlane GT** (2004) Chemotaxonomic analysis of bacterial populations colonizing the rectal mucosa in patients with ulcerative colitis. *Clin Infect Dis* 38: 1690-1699.
294. **Macfarlane S, Woodmansey EJ & Macfarlane GT** (2005) Colonization of mucin by human intestinal bacteria and establishment of biofilm communities in a two-stage continuous culture system. *Appl Environ Microbiol* 71: 7483-7492.
295. **MacKay WG, Williams CL, McMillan M, Ndip RN, Shepherd AJ & Weaver LT** (2003) Evaluation of protocol using gene capture and PCR for detection of *Helicobacter pylori* DNA in feces. *J Clin Microbiol* 41: 4589-4593.
296. **Mackie RI, Aminov RI, Hu W, Klieve AV, Ouwerkerk D, Sundset MA & Kamagata Y** (2003) Ecology of uncultivated *Oscillospira* species in the rumen of cattle, sheep, and reindeer as assessed by microscopy and molecular approaches. *Appl Environ Microbiol* 69: 6808-6815.
297. **Magee EA, Richardson CJ, Hughes R & Cummings JH** (2000) Contribution of dietary protein to sulfide production in the large intestine: an in vitro and a controlled feeding study in humans. *Am J Clin Nutr* 72: 1488-1494.
298. **Mai V, Braden CR, Heckendorf J, Pironis B & Hirshon JM** (2006) Monitoring of stool microbiota in subjects with diarrhea indicates distortions in composition. *J Clin Microbiol* 44: 4550-4552.
299. **Malinen E, Rinttilä T, Kajander K, Mättö J, Kassinen A, Krogius L, Saarela M, Korpela R & Palva A** (2005) Analysis of the fecal microbiota of irritable bowel syndrome patients and healthy controls with real-time PCR. *Am J Gastroenterol* 100: 373-382.
300. **Malnick H** (1997) *Anaerobiospirillum thomasii* sp. nov., an anaerobic spiral bacterium isolated from the feces of cats and dogs and from diarrheal feces of humans, and emendation of the genus *Anaerobiospirillum. Int J Syst Bacteriol* 47: 381-384.
301. **Mangin I, Bonnet R, Seksik P, Rigottier-Gois L, Sutren M, Bouhnik Y, Neut C, Collins MD, Colombel J-F, Marteau P & Doré J** (2004) Molecular inventory of faecal microflora in patients with Crohn's disease. *FEMS Microbiol Ecol* 50: 25-36.
302. **Manichanh C, Rigottier-Gois L, Bonnaud E, Gloux K, Pelletier E, Frangeul L, Nalin R, Jarrin C, Chardon P, Marteau P, Roca J & Doré J** (2006) Reduced diversity of faecal microbiota in Crohn's disease revealed by a metagenomic approach. *Gut* 55: 205-211.
303. **Marchesi JR, Holmes E, Khan F, Kochhar S, Scanlan P, Shanahan F, Wilson ID & Wang Y** (2007) Rapid and noninvasive metabonomic characterization of inflammatory bowel disease. *J Proteome Res* 6: 546-551.

304. **Matsuda K, Tsuji H, Asahara T, Kado Y & Nomoto K** (2007) Sensitive quantitative detection of commensal bacteria by rRNA-targeted reverse transcription-PCR. *Appl Environ Microbiol* 73: 32-39.
305. **Matsuki T, Watanabe K, Fujimoto J, Miyamoto Y, Takada T, Matsumoto K, Oyaizu H & Ryuichiro T** (2002) Development of 16S rRNA-gene-targeted group-specific primers for the detection and identification of predominant bacteria in human feces. *Appl Environ Microbiol* 68: 5445-5451.
306. **Matsuki T, Watanabe K, Fujimoto J, Takada T & Tanaka R** (2004) Use of 16S rRNA gene-targeted group-specific primers for real-time PCR analysis of predominant bacteria in human feces. *Appl Environ Microbiol* 70: 7220-7228.
307. **Matsumoto T, Hayasak T, Nishimura Y, Nakamura M, Takeda T, Tabuchi Y, Obinata M, Hanawa T & Yamada H** (2006) Butyrate induces necrotic cell death in murine colonic epithelial cell MCE301. *Biol Pharm Bull* 29: 2041-2045.
308. **Mättö J, Maunuksela L, Kajander K, Palva A, Korpela R, Kassinen A & Saarela M** (2005) Composition and temporal stability of gastrointestinal microbiota in irritable bowel syndrome - a longitudinal study in IBS and control subjects. *FEMS Immunol Med Microbiol* 43: 213-222.
309. **Maukonen J, Satokari R, Mättö J, Söderlund H, Mattila-Sandholm T & Saarela M** (2006) Prevalence and temporal stability of selected clostridial groups in irritable bowel syndrome in relation to predominant faecal bacteria. *J Med Microbiol* 55: 625-633.
310. **Maxwell PR, Rink E, Kumar D & Mendall MA** (2002) Antibiotics increase functional abdominal symptoms. *Am J Gastroenterol* 97: 104-108.
311. **McCartney AL, Wenzhi W & Tannock GW** (1996) Molecular analysis of the composition of the bifidobacterial and lactobacillus microflora of humans. *Appl Environ Microbiol* 62: 4608-4613.
312. **McGowan CC, Cover TL & Blaser MJ** (1996) *Helicobacter pylori* and gastric acid: Biological and therapeutic implications. *Gastroenterology* 110: 926.
313. **McOrist A, Warhurst M, McOrist S & Bird A** (2001) Colonic infection by *Bilophila wadsworthia* in pigs. *J Clin Microbiol* 39: 1577-1579.
314. **McWhorter AC, Haddock RL, Nocon FA, Steigerwalt AG, Brenner DJ, Aleksic S, Bockemuhl J & Farmer JJ, 3rd** (1991) *Trabulsiella guamensis*, a new genus and species of the family *Enterobacteriaceae* that resembles *Salmonella* subgroups 4 and 5. *J Clin Microbiol* 29: 1480-1485.
315. **Melito PL, Munro C, Chipman PR, Woodward DL, Booth TF & Rodgers FG** (2001) *Helicobacter winghamensis* sp. nov., a novel *Helicobacter* sp. isolated from patients with gastroenteritis. *J Clin Microbiol* 39: 2412-2417.
316. **Metchnikoff E** (1908) Etude sur la flore intestinale. *Ann d l'Inst Pasteur* 22.
317. **Metges CC** (2000) Contribution of microbial amino acids to amino acid homeostasis of the host. *J Nutr* 130: 1857S-1864.
318. **Metges CC, El-Khoury AE, Henneman L, Petzke KJ, Grant I, Bedri S, Pereira PP, Ajami AM, Fuller MF & Young VR** (1999) Availability of intestinal microbial lysine for whole body lysine homeostasis in human subjects. *Am J Physiol Endocrinol Metab* 277: E597-607.
319. **Mikosza ASJ, La T, de Boer WB & Hampson DJ** (2001) Comparative prevalences of *Brachyspira aalborgi* and *Brachyspira* (*Serpulina*) *pilosicoli* as etiologic agents of histologically identified intestinal spirochetosis in Australia. *J Clin Microbiol* 39: 347-350.
320. **Miller TL & Wolin MJ** (1985) *Methanosphaera stadtmaniae* gen. nov., sp. nov.: a species that forms methane by reducing methanol with hydrogen. *Arch Microbiol* 141: 116-122.
321. **Miller TL & Wolin MJ** (1986) Methanogens in human and animal tracts. *Syst Appl Microbiol* 7: 223-229.
322. **Minekus M, Smeets-Peeters M, Bernalier A, Marol-Bonnin S, Havenaar R, Marteau P, Alric M, Fonty G & Huis in't Veld JH** (1999) A computer-controlled system to simulate conditions of the large intestine with peristaltic mixing, water absorption and absorption of fermentation products. *Appl Microbiol Biotechnol* 53: 108-114.
323. **Mitsuoka T** (1969) Vergleichende Untersuchungen über die Laktobazillen aus den Faeces von Menschen, Schweinen und Hühnern. *Zbl Bakt Parasit* 210.
324. **Mitsuoka T** (1992) The human gastrointestinal tract. *The lactic acid bacteria in health and disease*. Vol. 1. (Wood BJB, eds). pp. 69-114. Elsevier Science Publishers, Ltd., Essex.
325. **Mitsuoka T** (1992) Intestinal flora and aging. *Nutr Rev* 50: 438-446.
326. **Mitsuoka T, Hayakawa K & Kimura N** (1975) The fecal flora of man. III. Communication: The composition of *Lactobacillus* flora of different age groups. *Zentralbl Bakteriol* 232: 499-501.
327. **Mitsuoka T, Terada A, Watanabe K & Uchida K** (1974) *Bacteroides multiacidus*, a new species from the feces of humans and pigs. *Int J Syst Bacteriol* 24: 35-41.
328. **Mohan R, Namsolleck P, Lawson PA, Osterhoff M, Collins MD, Alpert C-A & Blaut M** (2006) *Clostridium asparagiforme* sp. nov., isolated from a human faecal sample. *Syst Appl Microbiol* 29: 292.
329. **Moir RJ** (1951) The seasonal variation in the ruminal microorganisms of grazing sheep. *Austr J Agric Res* 2: 322.

330. **Molenda JR, Johnson WG, Fishbein M, Wentz B, Mehlman IJ & Dadisman TA, Jr.** (1972) Vibrio parahaemolyticus Gastroenteritis in Maryland: Laboratory Aspects. *Appl Environ Microbiol* 24: 444-448.
331. **Moore WEC, Cato EP & Holdeman LV** (1972) *Ruminococcus bromii* sp. n. and emendation of the description of *Ruminococcus* Sijpestein. *Int J Syst Bacteriol* 22: 78-80.
332. **Moore WEC & Holdeman LV** (1972) Identification of anaerobic bacteria. *Am J Clin Nutr* 25: 1306-1313.
333. **Moore WEC & Holdeman LV** (1974) Human fecal flora: the normal flora of 20 Japanese-Hawaiians. *Appl Microbiol* 27: 961-979.
334. **Moore WEC & Holdeman LV** (1974) Special problems associated with the isolation and identification of intestinal bacteria in fecal flora studies. *Am J Clin Nutr* 27: 1450-1455.
335. **Moore WEC, Johnson JL & Holdeman LV** (1976) Emendation of Bacteroidaceae and *Butyrivibrio* and descriptions of *Desulfomonas* gen. nov. and ten new species in the genera *Desulfomonas*, *Butyrivibrio*, *Eubacterium*, *Clostridium*, and *Ruminococcus*. *Int J Syst Bacteriol* 26: 238-252.
336. **Moore WEC & Moore LH** (1995) Intestinal floras of populations that have a high risk of colon cancer. *Appl Environ Microbiol* 61: 3202-3207.
337. **Morgan JAW & Winstanley C** (1997) Microbial biomarkers. *Modern soil microbiology*. (van Elsas JD, Trevors JT & Wellington EMH, eds). pp. 331-352. Marcel Dekker Inc., New York.
338. **Moro E** (1900) Ueber den *Bacillus acidophilus*. *Jahrb f Kinderh* 52: 38-55.
339. **Mueller S, Saunier K, Hanisch C, Norin E, Alm L, Midtvedt T, Cresci A, Silvi S, Orpianesi C, Verdenelli MC, Clavel T, Koebnick C, Zunft H-JF, Doré J & Blaut M** (2006) Differences in fecal microbiota in different European study populations in relation to age, gender, and country: a cross-sectional study. *Appl Environ Microbiol* 72: 1027-1033.
340. **Müller HE** (1986) Occurrence and pathogenic role of *Morganella-Proteus-Providencia* group bacteria in human feces. *J Clin Microbiol* 23: 404-405.
341. **Murdoch DA & Shah HN** (1999) Reclassification of *Peptostreptococcus magnus* (Prevot 1933) Holdeman and Moore 1972 as *Finegoldia magna* comb. nov. and *Peptostreptococcus micros* (Prevot 1933) Smith 1957 as *Micromonas micros* comb. nov. *Anaerobe* 5: 555-559.
342. **Muyzer G, de Waal EC & Uitterlinden AG** (1993) Profiling of complex microbial populations by denaturing gradient gel electrophoresis analysis of polymerase chain reaction-amplified genes coding for 16S rRNA. *Appl Environ Microbiol* 59: 695-700.
343. **Mylonaki M, Langmead L, Pantes A, Johnson F & Rampton DS** (2004) Enteric infection in relapse of inflammatory bowel disease: importance of microbiological examination of stool. *Eur J Gastroenterol Hepatol* 16: 775-778.
344. **Nakamura LK** (1984) *Bacillus amylolyticus* sp. nov., nom. rev., *Bacillus lautus* sp. nov., nom. rev., *Bacillus pabuli* sp. nov., nom. rev., and *Bacillus validus* sp. nov., nom. rev. *Int J Syst Bacteriol* 34: 224-226.
345. **Nakazawa F, Sato M, Poco SE, Hashimura T, Ikeda T, Kalfas S, Sundqvist G & Hoshino H** (2000) Description of *Mogibacterium pumilum* gen. nov., sp. nov. and *Mogibacterium vescum* gen. nov., sp. nov., and reclassification of *Eubacterium timidum* (Holdeman et al. 1980) as *Mogibacterium timidum* gen. nov., comb. nov. *Int J Syst Evol Microbiol* 50: 679-688.
346. **Neufeld JD, Mohn WW & de Lorenzo V** (2006) Composition of microbial communities in hexachlorocyclohexane (HCH) contaminated soils from Spain revealed with a habitat-specific microarray. *Environ Microbiol* 8: 126-140.
347. **Nguyen H-K & Southern EM** (2000) Minimising the secondary structure of DNA targets by incorporation of a modified deoxynucleoside: implications for nucleic acid analysis by hybridisation. *Nucl Acids Res* 28: 3904-3909.
348. **Nicholson JK, Holmes E & Wilson ID** (2005) Gut microorganisms, mammalian metabolism and personalized health care. *Nat Rev Microbiol* 3: 431-438.
349. **Nottingham PM & Hungate RE** (1968) Isolation of methanogenic bacteria from feces of man. *J Bacteriol* 96: 2178-2179.
350. **Noverr MC & Huffnagle GB** (2004) Does the microbiota regulate immune responses outside the gut? *Trends Microbiol* 12: 562-568.
351. **Noverr MC & Huffnagle GB** (2005) The 'microflora hypothesis' of allergic diseases. *Clin Exp Allergy* 35: 1511-1520.
352. **Nübel U, Engelen B, Felske A, Snaidr J, Wieshuber A, Amann RI, Ludwig W & Backhaus H** (1996) Sequence heterogeneities of genes encoding 16S rRNAs in *Paenibacillus polymyxa* detected by temperature gradient gel electrophoresis. *J Bacteriol* 178: 5636-5643.
353. **O'Hara AM & Shanahan F** (2006) The gut flora as a forgotten organ. *EMBO Rep* 7: 688-693.
354. **O'Mahony L, McCarthy J, Kelly P, Hurley G, Luo F, Chen K, O'Sullivan GC, Kiely B, Collins JK, Shanahan F & Quigley EMM** (2005) Lactobacillus and bifidobacterium in irritable bowel syndrome: Symptom responses and relationship to cytokine profiles. *Gastroenterology* 128: 541.

355. **Ohkusa T, Okayasu I, Ogihara T, Morita K, Ogawa M & Sato N** (2003) Induction of experimental ulcerative colitis by *Fusobacterium varium* isolated from colonic mucosa of patients with ulcerative colitis. *Gut* 52: 79-83.
356. **Ott SJ, Musfeldt M, Wenderoth DF, Hampe J, Brant O, Folsch UR, Timmis KN & Schreiber S** (2004) Reduction in diversity of the colonic mucosa associated bacterial microflora in patients with active inflammatory bowel disease. *Gut* 53: 685-693.
357. **Ozaki E, Kato H, Kita H, Karasawa T, Maegawa T, Koino Y, Matsumoto K, Takada T, Nomoto K, Tanaka R & Nakamura S** (2004) *Clostridium difficile* colonization in healthy adults: transient colonization and correlation with enterococcal colonization. *J Med Microbiol* 53: 167-172.
358. **Paisley JW, Mirrett S, Lauer BA, Roe M & Reller LB** (1982) Dark-field microscopy of human feces for presumptive diagnosis of *Campylobacter fetus* subsp. *jejuni* enteritis. *J Clin Microbiol* 15: 61-63.
359. **Palmer C, Bik EM, Eisen MB, Eckburg PB, Sana TR, Wolber PK, Relman DA & Brown PO** (2006) Rapid quantitative profiling of complex microbial populations. *Nucl Acids Res* 34: e5.
360. **Parker W** (1997) Re: What amount of mucus does the average person produce daily?
361. **Parracho HMRT, Bingham MO, Gibson GR & McCartney AL** (2005) Differences between the gut microflora of children with autistic spectrum disorders and that of healthy children. *J Med Microbiol* 54: 987-991.
362. **Passini** (1905) Studien uber faulnisserregende anaerobe Bakterien des normalen menschlichen Darmes und ihre Bedeutung. *Z Hyg Infektionskr* 49.
363. **Paster BJ, Boches SK, Galvin JL, Ericson RE, Lau CN, Levanos VA, Sahasrabudhe A & Dewhirst FE** (2001) Bacterial diversity in human subgingival plaque. *J Bacteriol* 183: 3770-3783.
364. **Penders J, Thijs C, Vink C, Stelma FF, Snijders B, Kummeling I, van den Brandt PA & Stobberingh EE** (2006) Factors influencing the composition of the intestinal microbiota in early infancy. *Pediatrics* 118: 511-521.
365. **Penner JL & Hennessy JN** (1980) Passive hemagglutination technique for serotyping *Campylobacter fetus* subsp. *jejuni* on the basis of soluble heat-stable antigens. *J Clin Microbiol* 12: 732-737.
366. **Persson S, Edlund M, Claesson R & Carlsson J** (1990) The formation of hydrogen sulfide and methyl mercaptan by oral bacteria. *Oral Microbiol Immunol* 5: 195-201.
367. **Peters LB, Wesselink PR & van Winkelhoff AJ** (2002) Combinations of bacterial species in endodontic infections. *Int Endod J* 35: 698-702.
368. **Petit L, Gibert M & Popoff MR** (1999) *Clostridium perfringens*: toxinotype and genotype. *Trends Microbiol* 7: 104.
369. **Petruschky J** (1896) *Bacillus faecalis alcaligenes* (n.sp.). *Zentr Bakt Parasiten* 19: 187-191.
370. **Phillips ML** (2006) Interdomain interactions: dissecting animal–bacterial symbioses. *BioScience* 56: 376-381.
371. **Piechaud D, Piechaud M & Second L** (1956) *Varietes proteolytiques* de *Moraxella lwofti* et de *Moraxella glucidolytica* (*Bact. anltratum*). *Ann Inst Pasteur* 90: 517-522.
372. **Pimentel M, Chow EJ & Lin HC** (2000) Eradication of small intestinal bacterial overgrowth reduces symptoms of irritable bowel syndrome. *Am J Gastroenterol* 95: 3503-3506.
373. **Pimentel M, Chow EJ & Lin HC** (2003) Normalization of lactulose breath testing correlates with symptom improvement in irritable bowel syndrome. a double-blind, randomized, placebo-controlled study. *Am J Gastroenterol* 98: 412-419.
374. **Pitcher MCL, Beatty ER & Cummings JH** (2000) The contribution of sulphate reducing bacteria and 5-aminosalicylic acid to faecal suphide in patients with ulcerative colitis. *Gut* 46: 64-72.
375. **Pitcher MCL & Cummings JH** (1996) Hydrogen sulphide: a bacterial toxin in ulcerative colitis? *Gut* 39: 1-4.
376. **Podschun R** (1991) Isolation of *Klebsiella terrigena* from human feces: biochemical reactions, capsule types, and antibiotic sensitivity. *Zentralbl Bakteriol* 275: 73-78.
377. **Podschun R, Acktun H, Okpara J, Linderkamp O, Ullmann U & Borneff-Lipp M** (1998) Isolation of *Klebsiella planticola* from newborns in a neonatal ward. *J Clin Microbiol* 36: 2331-2332.
378. **Possemiers S, Verthe K, Uyttendaele S & Verstraete W** (2004) PCR-DGGE-based quantification of stability of the microbial community in a simulator of the human intestinal microbial ecosystem. *FEMS Microbiol Ecol* 49: 495-507.
379. **Poyart C, Quesne G & Trieu-Cuot P** (2002) Taxonomic dissection of the *Streptococcus bovis* group by analysis of manganese-dependent superoxide dismutase gene (sodA) sequences: reclassification of '*Streptococcus infantarius* subsp. *coli*' as *Streptococcus lutetiensis* sp. nov. and of *Streptococcus bovis* biotype II.2 as *Streptococcus pasteurianus* sp. nov. *Int J Syst Evol Microbiol* 52: 1247-1255.
380. **Pozhitkov A, Noble PA, Domazet-Loso T, Nolte AW, Sonnenberg R, Staehler P, Beier M & Tautz D** (2006) Tests of rRNA hybridization to microarrays suggest that hybridization characteristics of oligonucleotide probes for species discrimination cannot be predicted. *Nucl Acids Res* 34: e66-.

381. **Pribram E** (1929) A Contribution to the classification of microorganisms. *J Bacteriol* 18: 361-394.
382. **Prigent-Delecourt L, Coffin B, Colombel JF, Dehennin JP, Vaerman JP & Rambaud JC** (1995) Secretion of immunoglobulins and plasma proteins from the colonic mucosa: an *in vivo* study in man. *Clin Exp Immunol* 99: 221-225.
383. **Probert HM, Apajalahti JHA, Rautonen N, Stowell J & Gibson GR** (2004) Polydextrose, lactitol, and fructo-oligosaccharide fermentation by colonic bacteria in a three-stage continuous culture system. *Appl Environ Microbiol* 70: 4505-4511.
384. **Proskuriakova NB, Zhigun EA & Shevchuk VB** (1982) *Acinetobacter colcoaceticus*: its properties, identification key and a study of its role in acute intestinal diseases. *Zh Mikrobiol Epidemiol Immunobiol* 3: 100-103.
385. **Rakoff-Nahoum S, Paglino J, Eslami-Varzaneh F, Edberg S & Medzhitov R** (2004) Recognition of commensal microflora by toll-like receptors is required for intestinal homeostasis. *Cell* 118: 229.
386. **Ramos CP, Falsen E, Alvarez N, Akervall E, Sjoden B & Collins MD** (1997) *Actinomyces graevenitzii* sp. nov., isolated from human clinical specimens. *Int J Syst Bacteriol* 47: 885-888.
387. **Ramotar K, Conly JM, Chubb H & Louie TJ** (1984) Production of menaquinones by intestinal anaerobes. *J Infect Dis* 150: 213-218.
388. **Rathinavelu S, Zavros Y & Merchant JL** (2003) *Acinetobacter lwoffii* infection and gastritis. *Microbes Infect* 5: 651.
389. **Rautio M, Eerola E, Väisänen-Tunkelrott M, Molitoris D, Lawson P, Collins MD & Jousimies-Somer H** (2003) Reclassification of *Bacteroides putredinis* (Weinberg et al., 1937) in a new genus *Alistipes* gen. nov., as *Alistipes putredinis* comb. nov., and description of *Alistipes finegoldii* sp. nov., from human sources. *Syst Appl Microbiol* 26: 182-188.
390. **Reuter G** (1963) Vergleichenden Untersuchung uber die Bifidus-Flora im Sauglings und Erwachsenenstuhl. *Zentralbl Bakteriol Parasitenkd Infektionskr Hyg Abt 1* 191: 486-507.
391. **Reuter G** (2001) The *Lactobacillus* and *Bifidobacterium* microflora of the human intestine: composition and succession. *Curr Issues Intest Microbiol* 2: 43-53.
392. **Rhodes MW & Kator H** (1999) Sorbitol-fermenting bifidobacteria as indicators of diffuse human faecal pollution in estuarine watersheds. *J Appl Microbiol* 87: 528-535.
393. **Rinttilä T, Kassinen A, Malinen E, Krogius L & Palva A** (2004) Development of an extensive set of 16S rDNA-targeted primers for quantification of pathogenic and indigenous bacteria in faecal samples by real-time PCR. *J Appl Microbiol* 97: 1166-1177.
394. **Robredo B, Singh KV, Torres C & Murray BE** (2000) Streptogramin resistance and shared pulsed-field gel electrophoresis patterns in vanA-containing *Enterococcus faecium* and *Enterococcus hirae* isolated from humans and animals in Spain. *Microb Drug Resist* 6: 305-311.
395. **Roediger WEW** (1980) Role of anaerobic bacteria in the metabolic welfare of the colonic mucosa in man. *Gut* 21: 793-798.
396. **Roediger WEW, Moore J & Babidge W** (1997) Colonic sulfide in pathogenesis and treatment of ulcerative colitis. *Dig Dis Sci* 42: 1571-1579.
397. **Rogosa M** (1965) The Genus *Veillonella* IV. Serological groupings, and genus and species emendations. *J Bacteriol* 90: 704-709.
398. **Rogosa M** (1984) *Bergey's Manual of Systematic Bacteriology.* Williams & Wilkins, Baltimore.
399. **Rogosa M, Mitchell JA & Wiseman RF** (1951) A selective medium for the isolation and enumeration of oral and fecal lactobacilli. *J Bacteriol* 62.
400. **Ron EZ** (2006) Host specificity of septicemic *Escherichia coli*: human and avian pathogens. *Curr Opin Microbiol* 9: 28.
401. **Roos S, Engstrand L & Jonsson H** (2005) *Lactobacillus gastricus* sp. nov., *Lactobacillus antri* sp. nov., *Lactobacillus kalixensis* sp. nov. and *Lactobacillus ultunensis* sp. nov., isolated from human stomach mucosa. *Int J Syst Evol Microbiol* 55: 77-82.
402. **Ruoff KL** (2002) Miscellaneous catalase-negative, Gram-positive cocci: Emerging opportunists. *J Clin Microbiol* 40: 1129-1133.
403. **Sacchi CT, Alber D, Dull P, Mothershed EA, Whitney AM, Barnett GA, Popovic T & Mayer LW** (2005) High level of sequence diversity in the 16S rRNA genes of *Haemophilus influenzae* isolates is useful for molecular subtyping. *J Clin Microbiol* 43: 3734-3742.
404. **Sahel J & Sarles H** (1979) Modifications of pure human pancreatic juice induced by chronic alcohol consumption. *Dig Dis Sci* 24: 897-905.
405. **Sajjan US, Corey M, Karmali MA & Forstner JF** (1992) Binding of *Pseudomonas cepacia* to normal human intestinal mucin and respiratory mucin from patients with cystic fibrosis. *J Clin Invest* 89: 648-656.

406. **Sakamoto M & Benno Y** (2006) Reclassification of *Bacteroides distasonis*, *Bacteroides goldsteinii* and *Bacteroides merdae* as *Parabacteroides distasonis* gen. nov., comb. nov., *Parabacteroides goldsteinii* comb. nov. and *Parabacteroides merdae* comb. nov. *Int J Syst Evol Microbiol* 56: 1599-1605.
407. **Sakamoto M, Suzuki M, Huang Y, Umeda M, Ishikawa I & Benno Y** (2004) *Prevotella shahii* sp. nov. and *Prevotella salivae* sp. nov., isolated from the human oral cavity. *Int J Syst Evol Microbiol* 54: 877-883.
408. **Sakuma K, Kitahara M, Kibe R, Sakamoto M & Benno Y** (2006) *Clostridium glycyrrhizinilyticum* sp. nov., a glycyrrhizin-hydrolysing bacterium isolated from human faeces. *Microbiol Immunol* 50: 481-485.
409. **Salyers AA** (1984) *Bacteroides* of the human lower intestinal tract. *Ann Rev Microbiol* 38: 293-313.
410. **Salzman NH, de Jong H, Paterson Y, Harmsen HJM, Welling GW & Bos NA** (2002) Analysis of 16S libraries of mouse gastrointestinal microflora reveals a large new group of mouse intestinal bacteria. *Microbiology* 148: 3651-3660.
411. **Sambrook J, Fritsch EF & Maniatis T** (1989) Calculating melting temperatures for perfectly matched hybrids between oligonucleotides and their target sequences. *Molecular cloning - a laboratory manual*. Vol. 2eds). pp. 11.46. Cold Spring Habour Laboratory Press, New York.
412. **Sambrook J, Fritsch EF & Maniatis T** (1989) Preparation of reagents and bufferes used in molecular cloning. *Molecular cloning - a laboratory manual*. Vol. 3eds). pp. B.13. Cold Spring Habour Laboratory Press, New York.
413. **Sanborn AG** (1931) The faecal flora of adults, with particular attention to individual differences and their relationship to the effects of various diets. *J Infect Dis* 48.
414. **Sanguinetti CJ, Dias Neto E & Simpson AJG** (1994) Rapid silver staining and recovery of PCR products separated on polyacrylamide gels. *BioTechniques* 17: 915-919.
415. **SantaLucia J, Jr.** (1998) A unified view of polymer, dumbbell, and oligonucleotide DNA nearest-neighbor thermodynamics. *Proc Natl Acad Sci* 95: 1460-1465.
416. **Sartor RB** (2006) Mechanisms of disease: pathogenesis of Crohn's disease and ulcerative colitis. *Nat Clin Pract Gastroenterol Hepatol* 3: 390-407.
417. **Satokari RM, Vaughan EE, Akkermans ADL, Saarela M & de Vos WM** (2001) Bifidobacterial diversity in human feces detected by genus-specific PCR and denaturing gradient gel electrophoresis. *Appl Environ Microbiol* 67: 504-513.
418. **Savage DC** (1977) Microbial ecology of the gastrointestinal tract. *Ann Rev Microbiol* 31: 107-133.
419. **Scardovi V & Crociani F** (1974) *Bifidobacterium catenulatum*, *Bifidobacterium dentium*, and *Bifidobacterium angulatum*: three new species and their deoxyribonucleic acid homology relationships. *Int J Syst Bacteriol* 24: 6-20.
420. **Scardovi V, Trovatelli LD, Biavati B & Zani G** (1979) *Bifidobacterium cuniculi*, *BIfidobacterium choerinum*, *Bifidobacterium boum*, and *Bifidobacterium pseudocatenulatum*: four new species and their deoxyribonucleic acid homology relationships. *Int J Syst Bacteriol* 29: 291-311.
421. **Schell MA, Karmirantzou M, Snel B, Vilanova D, Berger B, Pessi G, Zwahlen M-C, Desiere F, Bork P, Delley M, Pridmore RD & Arigoni F** (2002) The genome sequence of *Bifidobacterium longum* reflects its adaptation to the human gastrointestinal tract. *Proc Natl Acad Sci* 99: 14422-14427.
422. **Scheppach W & Weiler F** (2004) The butyrate story: old wine in new bottles? *Curr Opin Clin Nutr Metab Care* 7: 563-567.
423. **Schicho R, Krueger D, Zeller F, Von Weyhern CWH, Frieling T, Kimura H, Ishii I, De Giorgio R, Campi B & Schemann M** (2006) Hydrogen sulfide is a novel prosecretory neuromodulator in the Guinea-pig and human colon. *Gastroenterology* 131: 1542.
424. **Schlegel L, Grimont F, Collins M, Regnault B, Grimont P & Bouvet A** (2000) *Streptococcus infantarius* sp. nov., *Streptococcus infantarius* subsp. *infantarius* subsp. nov. and *Streptococcus infantarius* subsp. *coli* subsp. nov., isolated from humans and food. *Int J Syst Evol Microbiol* 50: 1425-1434.
425. **Schwiertz A, Hold GL, Duncan HS, Gruhl B, Collins MD, Lawson PA, Flint HJ & Blaut M** (2002) *Anaerostipes caccae* gen. nov., sp. nov., a new saccharolytic, acetate-utilising, butyrate-producing bacterium from human faeces. *Syst Appl Microbiol* 25: 46-51.
426. **Sebald M & Hauser D** (1995) Pasteur, oxygen and the anaerobes revised. *Anaerobe* 1: 11-16.
427. **Sepehri S, Kotlowski R, Bernstein CN & Krause DO** (2007) Microbial diversity of inflamed and noninflamed gut biopsy tissues in inflammatory bowel disease. *Inflamm Bowel Dis* in press.
428. **Shah HN & Collins MD** (1982) Reclassification of *Bacteroides hypermegas* (Harrison and Hansen) in a new genus *Megamonas*, as *Megamonas hypermegas* comb. nov. *Zentralbl Bakteriol Parasitenkd Infektionskr Hyg Abt 1* C3: 394-398.
429. **Shah HN & Collins MD** (1982) Reclassification of *Bacteroides multiacidus* (Mitsuoka, Terada, Watanabe and Uchida) in a new genus *Mitsuokella*, as *Mitsuokella multiacidus* comb. nov. *Zentralbl Bakteriol Parasitenkd Infektionskr Hyg Abt 1*: 491-494.

430. **Shah HN & Collins MD** (1986) Reclassification of *Bacteroides furcosus* Veillon and Zuber (Hauduroy, Ehringer, Urbain, Guillot and Magrou) in a new genus *Anaerorhabdus*, as *Anaerorhabdus furcosus* comb. nov. *Syst Appl Microbiol* 8.
431. **Shah HN & Collins MD** (1990) *Prevotella*, a new genus to include *Bacteroides melaninogenicus* and related species formerly classified in the genus *Bacteroides*. *Int J Syst Bacteriol* 40: 205-208.
432. **Shen J, Zhang B, Wei G, Pang X, Wei H, Li M, Zhang Y, Jia W & Zhao L** (2006) Molecular profiling of the *Clostridium leptum* subgroup in human fecal microflora by PCR-denaturing gradient gel electrophoresis and clone library analysis. *Appl Environ Microbiol* 72: 5232-5238.
433. **Sherman JM** (1938) The enterococci and related streptococci. *J Bacteriol* 35: 81-93.
434. **Shida O, Takagi H, Kadowaki K & Komagata K** (1996) Proposal for two new genera, *Brevibacillus* gen. nov. and *Aneurinibacillus* gen. nov. *Int J Syst Bacteriol* 46: 939-946.
435. **Shida O, Takagi H, Kadowaki K, Yano H, Abe M, Udaka S & Komagata K** (1994) *Bacillus aneurinolyticus* sp. nov., nom. rev. *Int J Syst Bacteriol* 44: 143-150.
436. **Simmering R, Taras D, Schwiertz A, Blay GL, Gruhl B, Lawson PA, Collins MD & Blaut M** (2002) *Ruminococcus luti* sp. nov., isolated from a human faecal sample. *Syst Appl Microbiol* 25: 189-193.
437. **Simon GL & Gorbach SL** (1984) Intestinal flora in health and disease. *Gastroenterology* 86: 174-193.
438. **Simon GL & Gorbach SL** (1986) The human intestinal microflora. *Dig Dis Sci* 31: 147S-162S.
439. **Simor AE & Wilcox L** (1987) Enteritis associated with *Campylobacter laridis*. *J Clin Microbiol* 25: 10-12.
440. **Simpson EH** (1949) Measurements of diversity. *Nature* 163: 688.
441. **Simpson EJ, Chapman MAS, Dawson J, Berry D, Macdonald IA & Cole A** (2000) *In vivo* measurement of colonic butyrate metabolism in patients with quiescent ulcerative colitis. *Gut* 46: 73-77.
442. **Smoot LM, Smoot JC, Smidt H, Noble PA, Konneke M, McMurry ZA & Stahl DA** (2005) DNA microarrays as salivary diagnostic tools for characterizing the oral cavity's microbial community. *Adv Dent Res* 18: 6-11.
443. **Sokol H, Seksik P, Rigottier-Gois L, Lay C, Lepage P, Podglajen I, Marteau P & Doré J** (2006) Specificities of the fecal microbiota in inflammatory bowel disease. *Inflamm Bowel Dis* 12: 106-111.
444. **Song Y, Kononen E, Rautio M, Liu C, Bryk A, Eerola E & Finegold SM** (2006) *Alistipes onderdonkii* sp. nov. and *Alistipes shahii* sp. nov., of human origin. *Int J Syst Evol Microbiol* 56: 1985-1990.
445. **Song Y, Liu C, Lee J, Bolanos M, Vaisanen M-L & Finegold SM** (2005) "*Bacteroides goldsteinii* sp. nov." isolated from clinical specimens of human intestinal origin. *J Clin Microbiol* 43: 4522-4527.
446. **Song Y, Liu C, Molitoris D, Tomzynsk TJ, Mc Teague M, Read E & Finegold SM** (2002) Use of 16S-23S rRNA spacer-region (SR)-PCR for identification of intestinal clostridia. *Syst Appl Microbiol* 25: 528-535.
447. **Song Y, Liu C, Molitoris DR, Tomzynski TJ, Lawson PA, Collins MD & Finegold SM** (2003) *Clostridium bolteae* sp. nov., isolated from human sources. *Syst Appl Microbiol* 26: 84-89.
448. **Song YL, Liu CX, McTeague M & Finegold SM** (2004) "*Bacteroides nordii*" sp. nov. and "*Bacteroides salyersae*" sp. nov. isolated from clinical specimens of human intestinal origin. *J Clin Microbiol* 42: 5565-5570.
449. **Song YL, Liu CX, McTeague M, Summanen P & Finegold SM** (2004) *Clostridium bartlettii* sp. nov., isolated from human faeces. *Anaerobe* 10: 179.
450. **Sonnenburg JL, Chen CT & Gordon JI** (2006) Genomic and metabolic studies of the impact of probiotics on a model gut symbiont and host. *PLoS Biol* 12: e413.
451. **Sood A, Midha V, Sood N, Bhatia AS & Avasthi G** (2003) Incidence and prevalence of ulcerative colitis in Punjab, North India. *Gut* 52: 1587-1590.
452. **Sprague LD & Neubauer H** (2005) *Yersinia aleksiciae* sp. nov. *Int J Syst Evol Microbiol* 55: 831-835.
453. **Stackebrandt E, Frederiksen W, Garrity GM, Grimont PAD, Kampfer P, Maiden MCJ, Nesme X, Rossello-Mora R, Swings J, Truper HG, Vauterin L, Ward AC & Whitman WB** (2002) Report of the ad hoc committee for the re-evaluation of the species definition in bacteriology. *Int J Syst Evol Microbiol* 52: 1043-1047.
454. **Stackebrandt E, Kramer I, Swiderski J & Hippe H** (1999) Phylogenetic basis for a taxonomic dissection of the genus *Clostridium*. *FEMS Immunol Med Microbiol* 24: 253.
455. **Stahl DA, Lane DJ, Olsen GJ & Pace NR** (1985) Characterization of a Yellowstone hot spring microbial community by 5S rRNA sequences. *Appl Environ Microbiol* 49: 1379-1384.
456. **Stanley J, Linton D, Burnens AP, Dewhirst FE, On SL, Porter A, Owen RJ & Costas M** (1994) *Helicobacter pullorum* sp. nov.-genotype and phenotype of a new species isolated from poultry and from human patients with gastroenteritis. *Microbiol* 140: 3441-3449.
457. **Starr SE, Thompson FS, Dowell VR, Jr. & Balows A** (1973) Micromethod system for identification of anaerobic bacteria. *Appl Environ Microbiol* 25: 713-717.
458. **Steer T, Collins MD, Gibson GR, Hippe H & Lawson PA** (2001) *Clostridium hathewayi* sp. nov., from human faeces. *Syst Appl Microbiol* 24: 353-357.

459. **Stewart JA, Chadwick VS & Murray A** (2005) Investigations into the influence of host genetics on the predominant eubacteria in the faecal microflora of children. *J Med Microbiol* 54: 1239-1242.
460. **Stubner S & Meuser K** (2000) Detection of *Desulfotomaculum* in an Italian rice paddy soil by 16S ribosomal nucleic acid analyses. *FEMS Microbiol Ecol* 34: 73-80.
461. **Suau A, Bonnet R, Sutren M, Godon J-J, Gibson GR, Collins MD & Doré J** (1999) Direct analysis of genes encoding 16S rRNA from complex communities reveals many novel molecular species within the human gut. *Appl Environ Microbiol* 65: 4799-4807.
462. **Suau A, Rochet V, Sghir A, Gramet G, Brewaeys S, Sutren M, Rigottier-Gois L & Doré J** (2001) *Fusobacterium prausnitzii* and related species represent a dominant group within the human fecal flora. *Syst Appl Microbiol* 24: 139-145.
463. **Swidsinski A, Weber J, Loening-Baucke V, Hale LP & Lochs H** (2005) Spatial organization and composition of the mucosal flora in patients with inflammatory bowel disease. *J Clin Microbiol* 43: 3380-3389.
464. **Tada A, Shiiba M, Yokoe H, Hanada N & Tanzawa H** (2004) Relationship between oral motor dysfunction and oral bacteria in bedridden elderly. *Oral Surg Oral Med Oral Pathol Oral Radiol Endod* 98: 184.
465. **Tajima K, Arai S, Ogata K, Nagamine T, Matsui H, Namakura M, Aminov RI & Benno Y** (2000) Rumen bacterial community transition during adaptation to high-grain diet. *Anaerobe* 6: 273-284.
466. **Tan PK, Downey TJ, Spitznagel EL, Jr., Xu P, Fu D, Dimitrov DS, Lempicki RA, Raaka BM & Cam MC** (2003) Evaluation of gene expression measurements from commercial microarray platforms. *Nucl Acids Res* 31: 5676-5684.
467. **Tannock GW** (1995) *Normal Microflora: an introduction to microbes inhabiting the human body.* Charpman & Hill, London.
468. **Tannock GW** (1999) Analysis of the intestinal microflora: a renaissance. *Antonie van Leeuwenhoek* 76: 265-278.
469. **Tannock GW** (2004) A special fondness for lactobacilli. *Appl Environ Microbiol* 70: 3189-3194.
470. **Tannock GW, Munro K, Bibiloni R, Simon MA, Hargreaves P, Gopal P, Harmsen H & Welling G** (2004) Impact of consumption of oligosaccharide-containing biscuits on the fecal microbiota of humans. *Appl Environ Microbiol* 70: 2129-2136.
471. **Tannock GW, Munro K, Harmsen HJM, Welling GW, Smart J & Gopal PK** (2000) Analysis of the fecal microflora of human subjects consuming a probiotic product containing *Lactobacillus rhamnosus* DR20. *Appl Environ Microbiol* 66: 2578-2588.
472. **Taras D, Simmering R, Collins MD, Lawson PA & Blaut M** (2002) Reclassification of *Eubacterium formicigenerans* Holdeman and Moore 1974 as *Dorea formicigenerans* gen. nov., comb. nov., and description of *Dorea longicatena* sp. nov., isolated from human faeces. *Int J Syst Evol Microbiol* 52: 423-428.
473. **Taylor GR, Kropp KD & Molina TC** (1985) Nine-year microflora study of an isolator-maintained immunodeficient child. *Appl Environ Microbiol* 50: 1349-1356.
474. **Thompson WG, Longstreth GF, Drossman DA, Heaton KW, Irvine EJ & Muller-Lissner SA** (1999) Functional bowel disorders and functional abdominal pain. *Gut* 45: ii43-47.
475. **Thomsen TR, Kjellerup BV, Nielsen JL, Hugenholtz P & Nielsen PH** (2002) *In situ* studies of the phylogeny and physiology of filamentous bacteria with attached growth. *Environ Microbiol* 4: 383-391.
476. **Tindall BJ, Grimont PAD, Garrity GM & Euzeby JP** (2005) Nomenclature and taxonomy of the genus *Salmonella. Int J Syst Evol Microbiol* 55: 521-524.
477. **Tissier MH** (1900) Recherches sur la flore intestinale normale et pathologique du nourrisson. In *Medicine* Paris: Université de Paris.
478. **Tissier MH** (1908) Recherches sur la flore intestinale normale des enfants ages dún an a cin. ans. *Ann D l'Inst Pasteur* 22: 189-207.
479. **Tortora GJ** (1989) *Principles of Human Anatomy.* Harper & Row Publisher, New York.
480. **Tullus K, Berglund B, Fryklund B, Kuhn I & Burman LG** (1988) Epidemiology of fecal strains of the family Enterobacteriaceae in 22 neonatal wards and influence of antibiotic policy. *J Clin Microbiol* 26: 1166-1170.
481. **Ursing J, Brenner DJ, Bercovier H, Fanning GR, Steigerwalt AG, Brault J & Mollaret HH** (1980) *Yersinia frederiksenii*: a new species of Enterobacteriaceae composed of rhamnose-positive strains (formerly called atypical *Yersinia enterocolitica* or *Yersinia enterocolitica*-like). *Curr Microbiol* 4: 213-217.
482. **Vallaeys T, Topp E, Muyzer G, Macheret V, Laguerre G, Rigeaud A & Soulas G** (1997) Evaluation of denaturing gradient gel electrophoresis in the detection of 16S rDNA sequence variation in Rhizobia and methanotrophs. FEMS Microbiol Ecol. 1997;24:279–285. *FEMS Microbiol Ecol* 24: 279-285.
483. **van de Wiele TR, Boon N, Possemiers S, Jacobs H & Verstraete W** (2004) Prebiotic effects of chicory inulin in the simulator of the human intestinal microbial ecosystem. *FEMS Microbiol Ecol* 51: 143-153.

484. **van der Waaij D** (1989) The ecology of the human intestine and its consequences for overgrowth by pathogens such as *Clostridium difficile*. *Ann Rev Microbiol* 43: 69-87.
485. **Van Eldere JR, De Pauw G & Eyssen HJ** (1987) Steroid sulfatase activity in a *Peptococcus niger* strain from the human intestinal microflora. *Appl Environ Microbiol* 53: 1655-1660.
486. **van Houte J & Gibbons RJ** (1966) Studies of the cultivable flora of normal human feces. *Antonie Van Leeuwenhoek* 32: 212-222.
487. **van Nuenen MHMC, Meyer PD & Venema K** (2003) The effect of various inulins and *Clostridium difficile* on the metabolic activity of the human colonic microbiota *in vitro*. *Microbial Ecol Health Dis* 15: 137-144.
488. **Vandamme P, Falsen E, Pot B, Hoste B, Kersters K & De Ley J** (1989) Identification of EF group 22 campylobacters from gastroenteritis cases as *Campylobacter concisus*. *J Clin Microbiol* 27: 1775-1781.
489. **Vanhoutte T, Huys G, Brandt E & Swings J** (2004) Temporal stability analysis of the microbiota in human feces by denaturing gradient gel electrophoresis using universal and group-specific 16S rRNA gene primers. *FEMS Microbiol Ecol* 48: 437-446.
490. **Vaughan EE, Heilig HGHJ, Ben-Amor K & de Vos WM** (2005) Diversity, vitality and activities of intestinal lactic acid bacteria and bifidobacteria assessed by molecular approaches. *FEMS Microbiology Reviews* 29: 477.
491. **Vaughan EE, Schut F, Heilig HG, Zoetendal EG, de Vos WM & Akkermans AD** (2000) A molecular view of the intestinal ecosystem. *Curr Issues Intest Microbiol* 1: 1-12.
492. **Veillon A & Zuber A** (1898) Recherches sur quelques microbes strictement anaerobies et leur role en pathologie. *Arch de med exper et d'anat path* 10: 517-545.
493. **Venema K, van Nuenen M, Smeets-Peeters M, Minekus M & Havenaar R** (2000) TNO's in vitro large intestinal model: an excellent screening tool for functional food and pharmaceutical research. *Nutrition/Ernährung* 24: 558-564.
494. **Venema K, van Nuenen M, van den Heuvel EG, Pool W & van der Vossen JMBM** (2003) The effect of lactulose on the composition of the intestinal microbiota and short-chain fatty acid production in human volunteers and a computer-controlled model of the proximal large intestine. *Microbial Ecol Health Dis* 15: 94-105.
495. **Venema K, Vermunt SHF & Brink EJ** (2005) D-tagatose increases butyrate production by the colonic microbiota in healthy men and women. *Microbial Ecol Health Dis* 17: 47-57.
496. **Venter JC, Remington K, Heidelberg JF, Halpern AL, Rusch D, Eisen JA, Wu D, Paulsen I, Nelson KE, Nelson W, Fouts DE, Levy S, Knap AH, Lomas MW, Nealson K, White O, Peterson J, Hoffman J, Parsons R, Baden-Tillson H, Pfannkoch C, Rogers Y-H & Smith HO** (2004) Environmental genome shotgun sequencing of the Sargasso Sea. *Science* 304: 66-74.
497. **Veron M & Chatelain R** (1973) Taxonomic study of the genus *Campylobacter*, Sebald and Veron and designation of the neotype strain for the type species, *Campylobacter fetus* (Smith and Taylor) Sebald and Veron. *Int J Syst Bacteriol* 23: 122-134.
498. **Wagner M, Loy A, Klein M, Lee N, Ramsing NB, Stahl DA, Friedrich MW & Jared RL** (2005) Functional marker genes for identification of sulfate-reducing prokaryotes. *Meth Enzym*. Vol. 397eds). pp. 469. Academic Press,
499. **Wagner M, Smidt H, Loy A & Zhou J** (2007) Unravelling microbial communities with DNA-microarrays: challenges and future directions. *Microb Ecol* Epub ahead of print.
500. **Waidmann M, Bechtold O, Frick J-s, Lehr H-a, Schubert S, Dobrindt U, Loeffler J, Bohn E & Autenrieth IB** (2003) *Bacteroides vulgatus* protects against *Escherichia coli*-induced colitis in gnotobiotic interleukin-2-deficient mice. *Gastroenterology* 125: 162.
501. **Wang M, Ahrne S, Antonsson M & Molin G** (2004) T-RFLP combined with principal component analysis and 16S rRNA gene sequencing: an effective strategy for comparison of fecal microbiota in infants of different ages. *J Microbiol Meth* 59: 53.
502. **Wang M, Ahrne S, Jeppsson B & Molin G** (2005) Comparison of bacterial diversity along the human intestinal tract by direct cloning and sequencing of 16S rRNA genes. *FEMS Microbiol Ecol* 54: 219.
503. **Wang M, Molin G, Ahrne S, Adawi D & Jeppsson B** (2007) High proportions of proinflammatory bacteria on the colonic mucosa in a young patient with ulcerative colitis as revealed by cloning and sequencing of 16S rRNA genes. *Dig Dis Sci* in press.
504. **Wang RF, Beggs ML, Robertson LH & Cerniglia CE** (2002) Design and evaluation of oligonucleotide-microarray method for the detection of human intestinal bacteria in fecal samples. *FEMS Microbiol Lett* 213: 175-182.
505. **Wang X, Heazlewood SP, Krause DO & Florin THJ** (2003) Molecular characterization of the microbial species that colonize human ileal and colonic mucosa by using 16S rDNA sequence analysis. *J Appl Microbiol* 95: 508-520.

506. **Wang Y, Zhang Z & Ramanan N** (1997) The actinomycete *Thermobispora bispora* contains two distinct types of transcriptionally active 16S rRNA genes. *J Bacteriol* 179: 3270-3276.
507. **Watanabe K, Kodama Y & Harayama S** (2001) Design and evaluation of PCR primers to amplify bacterial 16S ribosomal DNA fragments used for community fingerprinting. *J Microbiol Meth* 44: 253.
508. **Wathen-Grady HG, Britt LE, Strockbine NA & Wachsmuth IK** (1990) Characterization of *Shigella dysenteriae* serotypes 11, 12, and 13. *J Clin Microbiol* 28: 2580-2584.
509. **Wauters G, Janssens M, Steigerwalt AG & Brenner DJ** (1988) *Yersinia mollaretii* sp. nov. and *Yersinia bercovieri* sp. nov., formerly called *Yersinia enterocolitica* biogroups 3A and 3B. *Int J Syst Bacteriol* 38: 424-429.
510. **Wayne LG, Brenner DJ, Colwell RR, Grimont PAD, Kandler O, Krichevsky MI, Moore LH, Moore WEC, Murray RGE, Stackebrandt E, Starr MP & Truper HG** (1987) Report of the AD-HOC-Committee on reconciliation of approaches to bacterial systematics. *Int J System Bacteriol* 37: 463-464.
511. **Weis MotTRCBK** (2005) Standardizing global gene expression analysis between laboratories and across platforms. *Nat Methods* 2: 351-356.
512. **Weizman Z, Asli G & Alsheikh A** (2005) Effect of a probiotic infant formula on infections in child care centers: comparison of two probiotic agents. *Pediatrics* 115: 5-9.
513. **Welch & Nuttall** (1892) A gas producing bacillus (*Bacillus aerogenes capsulatus*, nov. spec.) capable of rapid development in the blood vessels after death. *Bull Johns Hopkins Hosp* 3.
514. **Werner H, Rintelen G & Kunstek-Santos H** (1975) A new butyric acid-producing bacteroides species: *B. splanchnicus* n. sp. *Zentralbl Bakteriol* 231: 133-144.
515. **Wexler HM, Reeves D, Summanen PH, Molitoris E, McTeague M, Duncan J, Wilson KH & Finegold SM** (1996) *Sutterella wadsworthensis* gen. nov., sp. nov., bile-resistant microaerophilic *Campylobacter gracilis*-like clinical isolates. *Int J Syst Bacteriol* 41: 252-258.
516. **Whiley RA, Beighton D, Winstanley TG, Fraser HY & Hardie JM** (1992) *Streptococcus intermedius*, *Streptococcus constellatus*, and *Streptococcus anginosus* (the *Streptococcus milleri* group): association with different body sites and clinical infections. *J Clin Microbiol* 30: 243-244.
517. **Whiley RA, Fraser HY, Douglas CW, Hardie JM, Williams AM & Collins MD** (1990) *Streptococcus parasanguis* sp. nov., an atypical viridans *Streptococcus* from human clinical specimens. *FEMS MIcrobiol Lett* 68: 115-122.
518. **Whitman WB, Coleman DC & Wiebe WJ** (1998) Prokaryotes: The unseen majority. *Proc Natl Acad Sci* 95: 6578-6583.
519. **Wick LM, Rouillard JM, Whittam TS, Gulari E, Tiedje JM & Hashsham SA** (2006) On-chip non-equilibrium dissociation curves and dissociation rate constants as methods to assess specificity of oligonucleotide probes. *Nucl Acids Res* 34: e26-.
520. **Willems A & Collins MD** (1994) Phylogenetic placement of *Sarcina ventriculi* and *Sarcina maxima* within group I *Clostridium*, a possible problem for future revision of the genus *Clostridium*. *Int J Syst Bacteriol* 44: 591-593.
521. **Willems A & Collins MD** (1996) Phylogenetic relationships of the genera *Acetobacterium* and *Eubacterium* sensu stricto and reclassification of *Eubacterium alactolyticum* as *Pseudoramibacter alactolyticus* gen. nov., comb. nov. *Int J Syst Bacteriol* 46: 1083-1087.
522. **Willems A, Moore WE, Weiss N & Collins MD** (1998) Phenotypic and phylogenetic characterization of some *Eubacterium*-like isolates containing a novel type B wall murein from human feces: description of *Holdemania filiformis* gen. nov., sp. nov. *Int J Syst Bacteriol* 47: 1201-1204.
523. **Wilson KH & Blitchington RB** (1996) Human colonic biota studied by ribosomal DNA sequence analysis. *Appl Environ Microbiol* 62: 2273-2278.
524. **Wilson KH, Wilson WJ, Radosevich JL, DeSantis TZ, Viswanathan VS, Kuczmarski TA & Andersen GL** (2002) High-density microarray of small-subunit ribosomal DNA probes. *Appl Environ Microbiol* 68: 2535-2541.
525. **Winter J, Popoff MR, Grimont P & Bokkenheuser VD** (1991) *Clostridium orbiscindens* sp. nov., a human intestinal bacterium capable of cleaving the flavonoid C-ring. *Int J Syst Bacteriol* 41: 355-357.
526. **Woese CR** (1987) Bacterial evolution. *Microbiol Rev* 51: 221-271.
527. **Woese CR, Magrum LJ, Gupta R, Siegel RB, Stahl DA, Kop J, Crawford N, Brosius R, Gutell R, Hogan JJ & Noller HF** (1980) Secondary structure model for bacterial 16S ribosomal RNA: phylogenetic, enzymatic and chemical evidence. *Nucl Acids Res* 8: 2275-2294.
528. **Wolin MJ, Miller TL, Collins MD & Lawson PA** (2003) Formate-dependent growth and homoacetogenic fermentation by a bacterium from human feces: Description of *Bryantella formatexigens* gen. nov., sp. nov. *Appl Environ Microbiol* 69: 6321-6326.
529. **Xiao JZ, Kondo S, Yanagisawa N, Takahashi N, Odamaki T, Iwabuchi N, Iwatsuki K, Kokubo S, Togashi H, Enomoto K & Enomoto T** (2006) Effect of probiotic *Bifidobacterium longum* BB536 in relieving clinical symptoms and modulating plasma cytokine levels of Japanese cedar pollinosis during

the pollen season. A randomized double-blind, placebo-controlled trial. *J Investig Allergol Clin Immunol* 16: 86-93.

530. **Xu J, Bjursell MK, Himrod J, Deng S, Carmichael LK, Chiang HC, Hooper LV & Gordon JI** (2003) A genomic view of the human-*Bacteroides thetaiotaomicron* symbiosis. *Science* 299: 2074-2076.
531. **Xu J & Gordon JI** (2003) Honor thy symbionts. *Proc Natl Acad Sci* 100: 10452-10459.
532. **Yanagita K, Manome A, Meng X-Y, Hanada S, Kanagawa T, Tsuchida T, Mackie RI & Kamagata Y** (2003) Flow cytometric sorting, phylogenetic analysis and in situ detection of *Oscillospira guillermondii*, a large, morphologically conspicuous but uncultured ruminal bacterium. *Int J Syst Evol Microbiol* 53: 1609-1614.
533. **Zhang T, Breitbart M, Lee WH, Run J-Q, Wei CL, Soh SWL, Hibberd ML, Liu ET, Rohwer F & Ruan Y** (2006) RNA viral community in human feces: prevalence of plant pathogenic viruses. *PLoS Biol* 4.
534. **Zhou J** (2003) Microarrays for bacterial detection and microbial community analysis. *Curr Opin Microbiol* 6: 288.
535. **Zinkevich V & Beech IB** (2000) Screening of sulfate-reducing bacteria in clonoscopy samples from healthy and colitic human gut mucosa. *FEMS Microbiol Ecol* 34: 147-155.
536. **Zoetendal EG, Akkermans AD, Akkermans-van Vliet WM, de Visser AJGM & de Vos WM** (2001) The host genotype affects the bacterial community in the human gastrointestinal tract. *Microb Ecol Health Dis* 13: 129-134.
537. **Zoetendal EG, Akkermans AD & de Vos WM** (1998) Temperature gradient gel electrophoresis analysis of 16S rRNA from human fecal samples reveals stable and host-specific communities of active bacteria. *Appl Environ Microbiol* 64: 3854-3859.
538. **Zoetendal EG, Collier CT, Koike S, Mackie RI & Gaskins HR** (2004) Molecular ecological analysis of the gastrointestinal microbiota: A review. *J Nutr* 134: 465-472.
539. **Zoetendal EG, Heilig HGHJ, Klaassens ES, Booijink CCGM, Kleerebezem M, Smidt H & de Vos WM** (2006) Isolation of DNA from bacterial samples of the human gastrointestinal tract. *Nat Protocols* 1: 870-873.
540. **Zoetendal EG, Plugge CM, Akkermans ADL & de Vos WM** (2003) *Victivallis vadensis* gen. nov., sp. nov., a sugar-fermenting anaerobe from human faeces. *Int J Syst Evol Microbiol* 53: 211-215.
541. **Zoetendal EG, Vaughan EE & de Vos WM** (2006) A microbial world within us. *Mol Microbiol* 59: 1639-1650.
542. **Zoetendal EG, von Wright A, Vilpponen-Salmela T, Ben-Amor K, Akkermans ADL & de Vos WM** (2002) Mucosa-associated bacteria in the human gastrointestinal tract are uniformly distributed along the colon and differ from the community recovered from feces. *Appl Environ Microbiol* 68: 3401-3407.
543. **Zubrzycki L & Spaulding EH** (1962) Studies on the stability of the normal human fecal flora. *J Bacteriol* 83: 968-974.
544. **Zuker M** (2003) Mfold web server for nucleic acid folding and hybridization prediction. *Nucl Acids Res* 31: 3406-3415.
545. **Zwirglmaier K, Ludwig W & Schleifer KH** (2004) Recognition of individual genes in a single bacterial cell by fluorescence in situ hybridization - RING-FISH. *Mol Microbiol* 51: 89-96.
546. **Button DK, Robertson BR** (2001) Determination of DNA content of aquatic bacteria by flow cytometry. Appl Environ Microbiol 67:1636-1645
547. **Rigottier-Gois L, Le Bourhis A, Gramet G, Rochet V & Doré J** (2003) Fluorescent hybridisation combined with flow cytometry and hybridisation of total RNA to analyse the composition of microbial communities in human faeces using 16S rRNA probes. FEMS Microbiol Ecol 43: 237-245
548. **Desantis TZ, Brodie EL, Moberg JP, Zubieta IX, Piceno YM & Andersen GL** (2007) High-density universal 16S rRNA microarray analysis reveals broader diversity than typical clone library when sampling the environment Microb Ecol. *in press*

Supplements

Samenvatting

Резиме

Acknowledgements

About the Author

Publications' List

Training and Supervision Plan

Samenvatting

Het laatste gedeelte van dit boekje is voor de groentjes, waarin ik op een populair wetenschappelijke manier zal uit probeer te leggen wat voor onderzoek is gedaan, en wat de belangrijkste resultaten zijn.

Laten we beginnen met de titel van dit boekje, die direct vertaald “de diversiteit van de humane darmflora – nieuwe inzichten uit grootschalige analyses” betekent. Ik begrijp dat dit nog niet veel duidelijk maakt. Allereerst is het belangrijk de betekenis van darmflora te verhelderen. Van de biologielessen op de lagere school herinneren we ons dat het laatste gedeelte van het spijsverteringskanaal, de dikke darm, dichtbevolkt is met micro-organismen. Er werd verteld dat deze micro-organismen een ondergeschoven rol spelen bij het opruimen van voedselresten, en dat de belangrijkste functie van de dikke darm zelf, wateropname is uit de laatste voedselresten, waarna ze het lichaam als feces verlaten. We kunnen ons afvragen of dit echt wel zo is, wanneer we bedenken dat de dikke darm een imposant orgaan is van 1.5 m in lengte en een dichtheid aan zenuwcellen heeft, die alleen door onze hersenen wordt geëvenaard. Waarom zou de natuur een orgaan ontwikkelen alleen voor het absorberen van in verhouding een kleine hoeveelheid water die nauwelijks zal bijdragen aan de 70% water waaruit ons lichaam bestaat? Dit gedeelte van het menselijke lichaam is veel belangrijker dan ons op school werd geleerd, en dat komt door de micro-organismen die er in leven en collectief de darmflora worden genoemd[*]. Mensen weten weinig over micro-organismen en dat komt vooral door hun grootte. Ondanks hun beperkte afmetingen, bevat de darm van een volwassene gemiddeld 1 kg van deze minuscule beestjes. De reden dat op school nagenoeg niets wordt onderwezen over deze fascinerende gemeenschap, is dat de wetenschap het beschrijven van de aanwezige soorten, en belangrijker, wat hun functie is in dit ecosysteem dat 1000 keer meer organismen bevat dan er mensen op de aardbol rond lopen, nog niet heeft voltooid.

De micro-organismen in de darm vervullen een rol die niet noodzakelijk gerelateerd is aan de vertering van voedsel. **Hoofdstuk 1** beschrijft een aantal van de belangrijkste functies, welke worden uitgevoerd door de darmflora, en duidelijk maakt waarom dit ecosysteem werd gekozen voor een diepgaande studie. Van de vele functies worden er hier twee uitgelicht. De darmflora maakt vitamine K aan, waarin K staat voor coagulatie en betrokken is bij het verdikken van bloed als reactie op verwondingen. Zo zou zonder vitamine K een wond niet helen. Hoewel planten ook vitamine K kunnen aanmaken, betrekken mensen het hoofdzakelijk van de darmflora. Daarnaast maakt de darmflora de 20 noodzakelijke aminozuren aan, welke de bouwstenen zijn voor de aanmaak van eiwitten. Hoewel van sommige van deze aminozuren wordt gezegd dat ze via dierlijke producten moeten worden

[*] Flora is natuurlijk niet het juiste woord, omdat flora voor planten staat zoals fauna voor dieren. Sinds kort wordt de term “biota” gebruikt, wat staat voor alles wat leeft en niet tot het planten of dierenrijk behoort.

ingenomen, kunnen veganisten en vastende gelovigen prima zonder. Het is zeer waarschijnlijk de darmflora die deze taak op zich neemt en voorziet in de aminozuurproductie.

De beschreven functies worden collectief toegeschreven aan de darmflora (wat tegenwoordig het vergeten orgaan van de mens wordt genoemd), terwijl eigenlijk maar weinig bekend is over de specifieke taken van de soorten afzonderlijk. Wat we wel weten is dat mensen met specifieke aandoeningen, zoals diarree, ulcerative colitis, allergieën en zelfs autisme, een verstoorde darmflora hebben. Dit is de reden waarom tegenwoordig het onderzoek zich wereldwijd richt op het verbinden van deze verstoorde darmflora met de genoemde ziektebeelden. Voordat we over verstoord kunt spreken, of wat verstoord betekent, moeten we eerst de normale darmflora definiëren.

De **hoofdstukken 2** en **3** beschrijven een literatuuronderzoek naar micro-organismen die specifiek zijn beschreven in de relatie tot de darmflora. We konden hiervoor putten uit twee bronnen van informatie die betrekking hebben op de manier van bestuderen. Enerzijds uit de klassieke en anderzijds de moleculaire microbiologie.

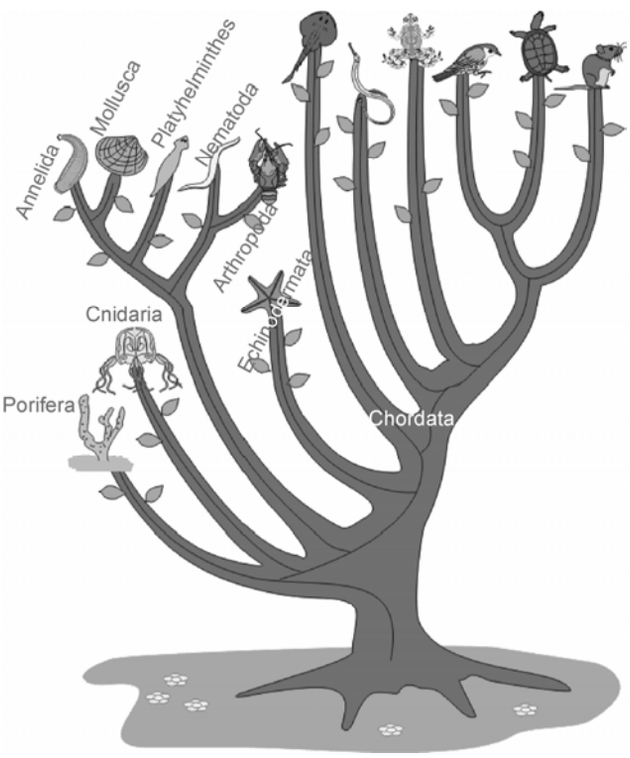
Klassieke microbiologie richt zich op het groeien van micro-organismen op verschillende voedingsbodems. Daarnaast worden bacteriën die op een specifiek medium groeien, verder gekarakteriseerd op basis van hun uiterlijke kenmerken zoals vorm en mnier van voortbeweging. In principe is dit een goede opzet, maar niet geschikt om de darmflora volledig te karakteriseren. De diversiteit van de darmflora is enorm en hiermee ook de behoefte aan verschillende combinaties van voedingsstoffen. Daarnaast gooit zuurstof voor de meeste darmflorasoorten roet in het eten, omdat dit voor hen bijzonder giftig is. Wanneer we dit doortrekken naar macrowereld; gaat de vergelijking tussen mens en vis goed op. Hoewel het beiden dieren zijn, kunnen ze niet in elkaars leefomgeving overleven. In **hoofdstuk 2** staat beschreven dat via de klassieke methode, tot nu toe, slechts een klein deel (~450 soorten) van de totale soortendiversiteit is beschreven.

In de moleculaire identificatie van de darmflora diversiteit wordt gebruik gemaakt van DNA, het erfelijke materiaal, en hiervan een gen dat genoeg informatie bevat om soorten afzonderlijk te identificeren. Laten we ons voorstellen dat er een gen is dat genoeg bureaucratische informatie bevat, zodat de nationaliteit en leeftijd van leden van een gemeenschap kan worden bepaald. Als ieder persoon een specifiek gen heeft dan zou dat er bijvoorbeeld zo uit kunnen zien:

```
Mijn naam is Petar Jovanovic ik ben geboren in 1950 en ik ben   Servisch.
Mijn naam is Luuk de Vries   ik ben geboren in 1980 en ik ben Nederlands.
Mijn naam is Joost Zegers    ik ben geboren in 1983 en ik ben Nederlands.
Mijn naam is Hans Schmidt    ik ben geboren in 1945 en ik ben      Duits.
```

Uit de informatie blijkt dat de groep een middelbare Serviër en Duitser bevat en twee jonge Nederlanders. We hoeven deze mensen niet te ontmoeten om de informatie uit de groep

te destilleren. Volgens hetzelfde principe, maar nu genoteerd als het 16S en 18S rRNA gen, is het mogelijk een micro-organisme te identificeren. Uit de analyse van de informatie opgesloten in dit gen, blijkt dat in de humane darm meer dan 1000 soorten aanwezig zijn, waarvan de meesten nog niet eerder gevonden zijn (**hoofdstuk 3**). Deze 1000 soorten representeren mogelijk maar een klein deel van de bacteriën die in de humane darm aanwezig kunnen zijn, omdat er duizenden of mogelijk tienduizenden verschillende micro-organismen zijn. De soorten kunnen worden ingedeeld in negen groepen, die wetenschappelijk stammen worden genoemd. Om een gevoel te krijgen voor de enorme verschillen tussen deze groepen, staat er als voorbeeld in figuur 1, een stamboom van negen stammen en hun vertegenwoordigers in de zichtbare wereld afgebeeld.



Figuur 1 Schematische voorstelling van een stamboom, welke negen stammen en hun vertegenwoordigers bevat. Ieder stam is vertegenwoordigd door een afbeelding van een soort behalve het fylum van de Chordata (Chordadieren), welke is vertegenwoordigd door zes soorten. Mensen en alle andere zoogdieren behoren ook tot dit stam. Figuur is gekopieerd van webpagina www.oum.ox.ac.uk

Dus in ieder van ons huist een universum van micro-organismen. Het is interessant dat dit universum, omdat de samenstelling van persoon tot persoon verschilt, voor ieder van ons specifiek is, en dat de darmflora alleen overeenkomsten vertoont tussen personen van dezelfde familie. Dit maakt een analyse van de darmflora tot een complexe bezigheid. Wanneer we bijvoorbeeld de verschillen in darmflora zijn tussen gezonde en autistische mensen willen achterhalen, dan is het niet voldoende om die van deze twee individuen te bestuderen, maar meer die van tientallen of liever honderden personen. Dit in ogenschouw genomen, en het feit dat de microflora (tien)duizenden soorten bevat, maakt duidelijk dat voor dit soort analyses gereedschap nodig is dat alle mogelijke micro-organismen in grote aantallen monsters en in korte tijd kan analyseren.

Onze gereedschapkist bevat nu de HITChip, waarvan de ontwikkeling beschreven staat in **hoofdstuk 4**. De HIT (Human Intestinal Tract=humane darm) Chip is ontwikkeld op basis van het gen dat wordt gebruikt voor de identificatie van micro-organismen (16S rRNA gen). Om het principe van de HITChip te begrijpen moeten we terug naar de zinnen op pagina 195. Het is duidelijk dat niet ieder gedeelte van de zin even informatief is. De grijze gedeelten zijn hetzelfde en daarom niet bruikbaar voor de identificatie van deze personen. Zo is de opbouw ook in de DNA sequentie van het 16S rRNA gen, dat stukken bevat die hetzelfde, en gedeelten die zeer verschillend zijn en gebruikt kunnen worden voor de identificatie van micro-organismen. Deze informatieve stukken in het voorbeeld op pagina 195 zijn analoog aan de volgende delen van de zin “1950” en “Servisch”, en zijn genoeg voor de identificatie op basis van leeftijd en nationaliteit. Na het analyseren van de DNA sequenties van duizenden organismen werd duidelijk dat het gebruik van slechts twee informatieve stukken genoeg is voor de volledige identificatie van alle aanwezige soorten micro-organismen. Deze delen werden gebruikt voor de ontwikkeling van de HITChip.

Om te begrijpen hoe deze stukken worden gebruikt, is het belangrijk te weten dat ieder gen als deel van het DNA, bestaat uit twee complementaire delen (zoals een origineel en zijn spiegelbeeld). Zo zal de zin er als volgt uit zien:

```
Mijn naam is Petar Jovanovic ik ben geboren in 1950 en ik ben   Servisch.
```

In realiteit, voelen twee complementaire delen DNA zich enorm tot elkaar aangetrokken, en binden ze via vele chemische bindingen. Het feit dat complementaire delen van DNA elkaar aantrekken vormt de basis van het principe van een DNA chip. Als we een complement van het woord “Servisch” nemen, dan zal dit iedere zin aantrekken dat het woord “Servisch” bevat. Het zal zo als “aas” worden gebruikt om zinnen te “vangen” die dit gespiegelde “Servisch” in zich hebben. Voor het “vangen” van de andere nationaliteiten moet een ander complementair “aas” worden gebruikt. Wanneer we nu een groep van zinnen van een onbekende gemeenschap hebben, die informatie bevat over de nationaliteit van de leden van de gemeenschap, dan hoeven we deze zelfs niet te lezen. Het is genoeg ze te markeren met een kleurstof en ze in de buurt te brengen van de “hengel” die is uitgerust met “aas” voor iedere nationaliteit. Iedere gekleurde zin die zijn spiegelbeeld vindt, bindt en kleurt zo de plaats waar het “aas” voor een specifieke nationaliteit aanwezig is, en kan zo de aanwezigheid van een lid van een specifieke nationaliteit in de geanalyseerde gemeenschap laten zien.

De HITChip kan worden gezien als een goed uitgeruste vishengel, welke complementen heeft voor de informatieve gedeelten van het 16S rRNA gen, waarmee de identificatie van meer dan duizend micro-organismen mogelijk is. Het grote voordeel van de chip is dat in één experiment informatie vrijkomt over de samenstelling van de darmflora en de hoeveelheden waarin ieder soort aanwezig is. Wanneer we bedenken dat in twee dagen ongeveer tien fecale monsters kunnen worden geanalyseerd, dan betekent het dat dit een high

throughput methode is, wat een grootschalige analyse mogelijk maakt. Voor het begrijpen van de darmflora, haar gedrag in gezonde personen en haar relatie tot ziekte, is dit een meer dan welkom gereedschap.

In de volgende hoofdstukken wordt de toepassing beschreven van de HITChip, en van de vele resultaten worden er hier een aantal uitgelicht. **Hoofdstuk 6** beschrijft de verandering die de darmflora ondergaat in de tijd. Een interessant aspect van deze studie is dat de darmflora van ieder persoon als een vingerafdruk gebruikt kan worden, omdat de unieke samenstelling, zelfs na een periode van negen jaar, slechts minimaal verandert. De darmflora is stabiel en persoonsgebonden kan zelfs in forensisch onderzoek worden toepast, wat daadwerkelijk is gedaan in een recente Nederlandse verkrachtingszaak. Een andere belangrijke ontdekking is dat sommige groepen bacteriën immuun zijn tegen iedere verandering in voedingspatroon, leefgewoonte, stress, reizen, enz. Het was al duidelijk dat het menselijke lichaam en bacteriën met elkaar “spreken”, en dat via ontvangers in de darmwand goede bacteriën worden “gevangen”. Wanneer een bacterie passeert die niet als zodanig wordt herkend, dan vernietigd ons lichaam hem, zoals dat ook met virussen gebeurt in een afweerreactie. Uit ons onderzoek blijkt dat het lichaam lievelingsbacteriën heeft die worden gekoesterd door de darm op een manier die nog niet helemaal duidelijk is. Deze bacteriën behoren waarschijnlijk tot het meest bruikbare deel van de darmflora, en hebben mogelijk, gezien hun aanwezigheid in yoghurtproducten, een positief effect op onze gezondheid.

In **hoofdstuk 8**, worden de verschillen in darmflora tussen gezonde personen en patiënten met ulcerative colitis (UC) beschreven. UC is een vervelende ziekte die zich manifesteert als een afweerreactie op micro-organismen. Zelfs de goede bacteriën, die normaal worden herkend door de eerder genoemde ontvangers, kunnen dan zo´n reactie opwekken. De symptomen van deze aandoening zijn pijnlijke buikkramp, een bloedende dikke darm, en diarree. UC komt vaker voor in westerse landen dan in ontwikkelingslanden, en neemt daarnaast nog steeds toe (zoals ook bijvoorbeeld allergieën en overgewicht), wat betekent dat een van de oorzaken van deze ziekte kan liggen in het type leefpatroon. Aan welk leefpatroon dit gelinked zou kunnen worden, werd onderzocht door de analyses van de HITChip te combineren met de uitkomsten van eerder gedaan onderzoek.

Dat bacteriën in de dikke darm van UC patiënten grote hoeveelheden waterstofsulfide produceren (gas dat ruikt naar rotte eieren en even giftig is als cyanide), was al bekend. In **hoofdstuk 8** beschrijven we dat de darmflora van UC patiënten grote hoeveelheden peptostreptococcen bevat – het type bacteriën dat eiwitten verbruikt en daarbij waterstofsulfide produceren. Op basis hiervan concluderen we dat het consumeren van te veel eiwit een mogelijke reden kan zijn voor het ontwikkelen van UC, en zich kan manifesteren op de volgende manier; Van nature kunnen we slechts 200 gram eiwit per dag consumeren, wat de maximale hoeveelheid is die kan worden verteerd. Wanneer we hiervan te veel binnen krijgen, meer dan de verteringscapaciteit, eindigen eiwitten intact in de dikke darm, waar

bacteriën dit gebruiken en grote hoeveelheden waterstofsulfide (giftig!) afgeven. Ondanks dat deze bacteriën bij normale eiwitinname een nuttige functie hebben in de darm, vergiftigen ze nu de gastheer. Dit vraagt om een afweerreactie die wordt opgestart tegen deze goede bacteriën, wat resulteert in de beschreven symptomen. Hoewel dit mechanisme de ontwikkeling van UC niet in iedere patiënt kan verklaren, is het belangrijk te weten dat verkeerde voeding fatale gevolgen kan hebben (waterstofsulfide wordt bijvoorbeeld als mogelijke oorzaak van darmkanker opgevoerd).

In ons onderzoek komt verder naar voren, dat bij mensen met een verstoorde vertering (IBS) de darmflora aanzienlijk verschilt van die van gezonde mensen (**hoofdstuk 7**). Daarnaast hebben we met de HITChip ontdekt, dat zelfs het meest moderne model van het menselijke spijsverteringskanaal niet dezelfde micro-organismen bevat als die, die in werkelijkheid in de humane darm leven (**hoofdstuk 4**). Dit resultaat en andere resultaten zijn samengevat en worden bediscussieerd in het laatste hoofdstuk van dit boekje – **hoofdstuk 9**. Laat ieder die nu geïntrigeerd is door deze samenvatting, beginnen met het echte werk. Lezen maar!

Резиме

Да бих оправдала звање доктора наука, у складу са народном пословицом „није знање знање знати, већ је знање знање дати“; у овом поглављу ћу покушати да пренесем својој породици и другарима знање које сам стекла током претходних пет година.

Да почнемо од наслова тезе, који директно преведен гласи: „Разноврсност људске цревне микрофлоре – нове перспективе проистекле из опсежних анализа“.

Да би остатак приче био разумљив, потребно је одмах разјаснити шта је то заправо цревна микрофлора. Можда се неко сећа да смо још у основној школи учили да је последњи део пробавног тракта, који се зове колон или дебело црево, густо насељен микроорганизмима, углавном бактеријама. У основној школи нас уче да те бактерије немају неку посебну функцију, сем што, себе ради, једу остатке несварене хране, и да дебело црево служи за апсорпцију вишка воде и за уклањање отпадних остатака. Овакво учење звучи нелогично ако се узме у обзир да је дебело црево прилично импресиван орган, дугачак је око метар и по, са највећом концентрацијом нервних ћелија у организму, наравно после мозга. И зашто би, заправо, природа радила на развијању једног таквог органа а да он служи само за уклањање мале количине воде? део људског тела је ипак значајнији него што нас уче у школи и то превасходно због микроорганизама који га насељавају, а који се колективно зову микрофлора*. Људи мало знају о микроорганизмима, о тим, како им само име каже, микро (сићушним) бићима, управо зато што су тако сићушна. Но иако су мали, не значи да их је мало – у људским цревима има преко 1 kg микроорганизама. Разлог што о њима не учимо у школи је што и најученији људи још увек раде на томе да опишу ко су и шта раде сва та силна бића којих у сваком човеку има хиљаду пута више него што има људи на Земљи!

Микроорганизми који насељавају наша црева нису било ко, то су одабране врсте, које живе у симбиози са људима и обављају разне корисне функције, које су и везане и невезане за варење. Неке од многих утврђених функција које обавља цревна микрофлора су побројане у **Поглављу 1**, које представља преглед литературе, са неамером да објасни зашто је баш овај екосистем (биолошка заједница) изабран за тему мог доктората. Функција је много, а ја ћу овде навести само две, мени интересантне. Цревна микрофлора производи витамин К, а ово „К“ означава коагулацију, што је стручан термин за згрушавање, у конкретном случају крви. Наиме, без овог витамина ниједна рана не би зарасла, и иако неке биљке синтетишу витамин К, за људе је његов основни извор управо цревна микрофлора. Даље, цревна микрофлора производи свих 20, човеку потребних, аминокиселина (аминокиселине су градивне јединке

* Микрофлора је распрострањен иако неправилан термин јер је термин „флора“ резервисан за биљке а „фауна“ за животиње. Термин „биота“ се односи на сав живи свет, а нарочито онај који се не може сврстати ни у биљке нити у животиње, што би било адекватно за микроорганизме.

беланчевина). За неке од ових аминокиселина нутриционисти тврде да се морају уносити путем намирница животињског порекла (дакле меса, млека, јаја), али с обзиром да и наша цревна микрофлора може да произведе свих 20 аминокиселина, користећи друге остатке хране, јасно је како и вегетаријанци и монаси испосници могу да буду потпуно здрави и ухрањени без животињске хране. Ове и многе друге важне функције које обавља цревна микрофлора везују се за целу заједницу микроорганизама (коју неки аутори у последње време називају и заборављеним органом људског тела), а мало је познато о томе шта заправо ради свака од јединки тј. врста у тој заједници. Но, примећено је да је састав цревне микрофлоре поремећен код болесника који болују од дијареје, улцерозног колитиса али и алергије па чак и аутизама, и научници нашег времена (укључујући и мене) покушавају да објасне везе између болести и поремећаја цревне микрофлоре. Дефиниција поремећаја није ни могућа без дефиниције „нормалног“ тако да је један од задатака мог доктората био да разјасни шта је то заправо нормална цревна микрофлора.

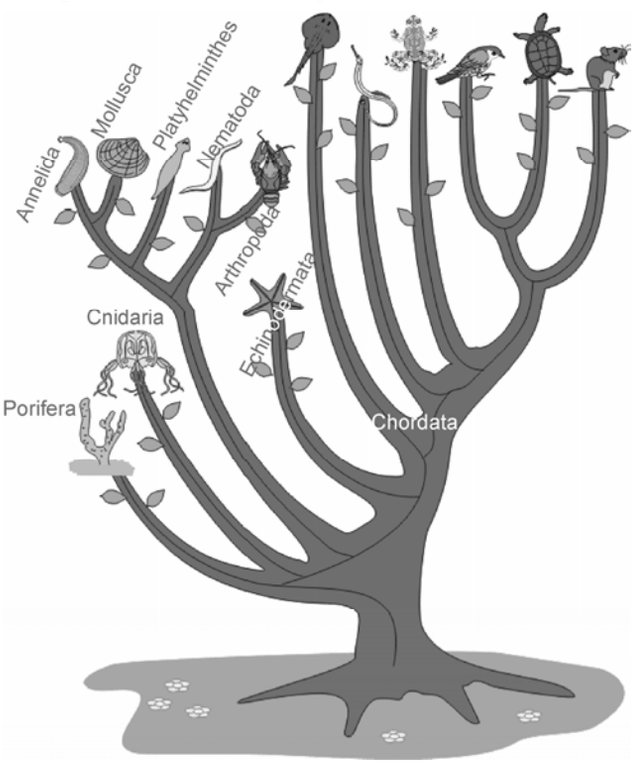
Поглавље 2 и **3** представљају комбинацију прегледа литературе, објављене током целог прошлог века и анализе података о разноврсности микроорганизама који чине цревну микрофлору, са намером да се направи што комплетнија слика састава овог екосистема. Разлог што је једна иста тема обрађена у два поглавља је тај што постоје два извора информација о разноликости цревне флоре: класичан и молекуларан.

Класичан начин анализе микробиолошких система подразумева узгајање миркоорганизама на различитим хранљивим подлогама, да би се на основу неких основних карактеристика (нпр. облика и покретљивости) вршила њихова идентификација. Овај, у принципу сасвим исправан концепт, се испоставио као неадекватан за анализу људске цревне микрофлоре. Разлози за то су многи, и укључују чињеницу да је цревна микрофлора многобројна и да садржи много врста, које расту у потпуно различитим условима (преведено на макро свет, људи и рибе не живе у истим условима), тако да је потребно направити десетине различитих хранљивих подлога, избегавати најкраћи сусрет са кисеоником, који је за већину цревних микроорганизама отров итд. Због тога је класична микробиологија успела да идентификује само део микроорганизама из овог система, како се испоставило у **поглављу 2**, преко 450 врста.

Молекуларна идентификација микроорганизама базира се на чињеници да постоји један ген, који ћемо посматрати као реченицу, што, у принципу, и није погрешно, која садржи довољно информација за идентификацију врсте. Ради једноставности направићу једну аналогију разумљивију обичном човеку. Рецимо да постоји ген који садржи један прилично бирократски запис и који нам омогућава да одредимо националност и годиште особа присутних у једној заједници. Пошто свака индивидуа садржи један такав ген (запис), онда скуп записа који се може добити од групе људи изгледа на пример овако:

```
Зовем се Петар Јовановић, рођен сам 1950, по националности сам      Србин.
Зовем се     Лук де фрис, рођен сам 1980, по националности сам Холанђанин.
Зовем се    Јоуст Зегерс, рођен сам 1983, по националности сам Холанђанин.
Зовем се       Ханс Шмит, рођен сам 1945, по националности сам      Немац.
```

Јасно је да се на основу оваквог записа може закључити да конкретну групу чине средовечни Србин и Немац, и два млађа Холанђанина. Дакле ове људе није потребно упознати да бисмо ово сазнали, довољно је од њих само извући овакву информацију. На врло сличном принципу, на основу записа у тзв. 16С или 18С рРНК гену, могуће је идентификовати врсту микроорганизма. Тако се анализом записа из овог гена показало да у људским цревима има преко 1000 врста, углавном потпуно непознатих микроорганизама (као што је представљено у **поглављу 3**). Ових 1000 врста представља само мали, за сада откривени део од укупне разноликости цревне микрофлоре, јер цревна микрофлора човека, вероватно садржи хиљаде ако не и десетине хиљада различитих организама. Микроорганизми из наших црева се могу сврстати у девет група, тзв. девет раздела живота из микро света, а да би било јасно колико је то разноврсан екосистем на слици 1 је представљено „породично стабло“ које садржи представнике девет раздела живота из нама видљивог света.



Слика 1 Шематски приказ „породичног стабла“ које садржи представнике девет раздела живота укључујући и хордате којима припадају сви сисари па и људи. Приказан је по један представник сваког од девет раздела изузев хордата који је представљен са шест врста. Слика ја преузета са интернет странице: www.oum.ox.ac.uk

Дакле, у сваком од нас живи читав један микро свет. Занимљиво је да је тај микро свет карактеристичан за сваког од нас, његов састав се мења од човека до човека, а сличнији је што су особе сродније. Због такве различитости у саставу цревне микрофлоре сваког од нас, готово је немогуће установити које су разлике између одређених група људи. Нпр. ако неко жели да сазна која је разлика у саставу цревне микрофлоре здравог човека и болесника (рецимо од аутистичара), није довољно да се само анализира цревна микрофлора једног здравог и једног аутистичног човека.

Напротив, потребно је вршити опсежне анализе на десетинама, ако не стотинама људи, а ако се узме у обзир колико је разноврсна микрофлора, онда је јасно да је за такве врсте анализа неопходан метод који може да обради што више узорака за кратко време и који може да детектује што више врста микроорганизама у исто време.

Један такав метод је развијен у оквиру мог доктората, а начин на који је развијен овај конкретни ХИТ чип је описан у **поглављу 4**. ХИТ чип је добио име на основу скраћенице од „Human Intestinal Tract Chip“, што директно преведено значи: „чип за људски пробавни тракт“. ХИТ чип је развијен на основу 16С рРНК гена – истог оног гена који се користи за молекуларно идентификовање микроорганизама. Да бих објаснила на који начин је развијен тај чип вратићу се на скуп реченица са претходне стране. Очигледно, није сваки део ових реченица једнако информативан; осенчени делови реченица су исти и, као такви, они су потпуно небитни за опис било ког од ова четири човека. Слично је и са записом из гена 16С рРНК јер он садржи делове која су потпуно идентични у свим микроорганизмима, али и информативне делове који су специфични за сваку врсту. Информативни делови гена 16С рРНК могу се упоредити са деловима реченице „1950“, или „Србин“, јер само је ова информација довољна да се одреди старосно доба и националност особе. Анализирајући 16С рРНК записе преко 1000 микроорганизама испоставило се да су за потпуну идентификацију врсте довољна два кратка записа из гена 16С рРНК, које су искоришћени за прављење ХИТ чипа.

Да би било разумљиво како такви записи могу бити искоришћени за прављење чипа (па макар он био и ДНК чип) битно је истаћи још једну чињеницу. Сваки ген, као део ДНК, се састоји из два дела која су међусобно комплементарна (ради сликовитости може се рећи да се односе један према другом као оригинал и слика у огледалу). Тиме, једна од реченица коју смо раније користили као аналогију за запис у гену, у својој пуној форми, када има два дела, изгледа овако:

У реалности, два комплементарна дела ДНК се привлаче, и кад су довољно близу она се повежу помоћу многобројних хемијских веза. Чињеница да се два комплементарна дела ДНК привлаче је у основи принципа на ком раде ДНК чипови и састоји се у следећем: уколико направимо реч која представља одраз речи „Србин“ он ће бити привлачан за сваки запис који садржи реч „Србин“ и због тога се може искористити као „мамац“ да би се из непознатог узорка „упецали“ сви записи који садрже ову реч. На сличан начин могу се направити „мамци“ за записе који садрже име било које нације. Тако ако желимо да установимо национални састав једне непознате заједнице, није потребно чак ни читати записе из бирократског гена, довољно је урадити само следеће: Након што издвојимо бирократске записе из неке непознате заједнице, потребно је да их обојимо и доведемо у близину једној „пецаљки

опремљеној „мамцима“ за све могуће нације. Обојени запис који нађе свој одраз у огледалу ће се спојити са одговарајућим „мамцем“, и на тај начин обојити место на којима је би закачен „мамац“ и дати нам на знање да у конкретном систему постоје представници конкретне нације (или конкретних нација).

Тиме се ХИТ чип може поистоветити са једном добро опремљеном пецаљком, која садржи комплементарну страну информативних делова из гена 16С рРНК који могу да се искористе за идентификацију преко 1000 микроорганизама из цревне микрофлоре. Велика предност оваквог чипа је што може, у једном јединому експерименту да детаљно опише изразито комплексан екосистем као што је цревна микрофлора и да поред идентификације организама, обезбеди податак о бројности сваког микроорганизма или групе микроорганизама. Ако се узме у обзир да експеримент који траје око два дана може да опише десетак узорака, јасно је да се ради о високо проточном систему који омогућава опсежне анализе, неопходне за разумевање цревне микрофлоре, њеног понашања код здравих људи и везе са болестима.

У наредним поглављима су презентовани резултати примене ХИТ чипа на анализу састава цревне микрофлоре. Од многих резултата опет ћу издвојити само неколико интересантних. **Поглавље 6** описује начин на који се мења цревна микрофлора са годинама а интересантно откриће овог истраживања је да је састав цревне микрофлоре нека врста отиска прста сваког од нас, јер се састав цревне флоре врло мало мења током година, и састав цревне микрофлоре је специфичан за сваког од нас, и после периода од девет година (утицај дужег периода од овог још нисмо проверили). Ова чињеница, слично чињеници да је отисак прста јединствен за сваку особу, може да се искористи и у сврху проналажења криминалаца и ово је заиста и учињено у једном случају силовања, овде у Холандији. Друго важно откриће представљено у **поглављу 6** је да су неке бактеријске група готово потпуно имуне на било какве промене у исхрани, начину живота, на присуство стреса, путовања и сл. Од раније је познато да из зидова наших црева вире мали „торњеви“ на којима се идентификују микроорганизми присутни у цревној микрофлори и сваки микроорганизам који не прође идентификацију на једном од „торњева“ бива уклоњен из система у одбрамбеној реакцији (сличну реакцију имамо кад се прехладимо). Оно што је први пут показано у овом докторату је да постоје и „миљеници“ у оквиру цревне микрофлоре и да људско тело на неки начин стимулише и контролише присуство тих конкретних група бактерија које, изгледа, представљају најважније саставне делове цревне микрофлоре. Занимљиво да су неке од ових „миљеница“, од раније познате као корисне за здравље људи те се и стављају у јогурте типа „pro viva”-е.

Наредно занимљиво откриће које ћу овде представити је описано у **поглављу 8**. У оквиру ове студије покушали смо да установимо по чему се разликује састав цревне микрофлоре здравих људи и људи који болују од улцерозног колитиса. Улцерозни

колитис је једна врло непријатна болест која се развија из непознатог разлога и која је специфична по томе што постоји јака одбрамбена реакција на неке од нормалних бактерија цревне микрофлоре (које бивају препознате на већ поменутим „торњевима“ присутним на зиду црева). Ово се код болесника манифестује, јаким болом, крварењем дебелог црева и проливом. Улцерозни колитис је једна од болести модерних цивилизација (као и алергије, гојазност или рак), што значи да се њен узрок може тражити у неким од тековина модерног стила живота. Која је то промена, ми смо покушали да откријемо и објаснимо на основу већ познатих чињеница, али и на основу нових открића представљених у овој докторској тези.

Од раније је познато да се у цревима људи који болују од улцерозног колитиса производе велике количине водоник-сулфида (то је гас који смрди на покварена јаја и који је отрован за људе колико и цијанид). У **поглављу 8** смо показали да цревна флора људи који болују од улцерозног колитиса има увећану концентрацију пептострептокока – групе бактерије које се хране беланчевинама из којих могу да ослободе водоник-сулфид. На основу овога можемо да закључимо да је конзумација превелике количине беланчевина један од разлога за појаву улцерозног колитиса, који се дешава на следећем принципу: природно је уређено да људи током дана уносе ограничену количину беланчевина (око 200 g), јер је то количина коју можемо да сваримо. Ако се претера (као код нас на свадбама и славама), онда вишак беланчевина (читај меса, млека, сира, јаја) доспе до дебелог црева где га ове цревне бактерије крену да једу и при том ослобађају велику количину водоник-сулфида (да подсетим једног ужасног отрова). И иако су те цревне бактерије у нормалним условима корисне за човека пошто га у конкретном тренутку трују, човечји одбрамбени систем ће напасти ове „добродошле“ бактерије и доводи до болести. Иако овај механизам не може да објасни разлоге за појаву улцерозног колитиса код сваког болесника, битно је знати да неправилна исхрана може да има и фаталне последице (јер је овај исти водоник сулфид повезан и са развојем рака дебелог црева).

Остала открића овог доктората укључују и чињеницу да се састав цревне микрофлоре људи који имају нестабилан рад пробавног система (који болују од синдрома изиритираног црева) разликује од састава цревне флоре здравих људи (**поглавље 7**), али и чињеницу да чак и најмодернији лабораторијски модел дебелог црева не садржи оне врсте бактерије које живе у нормалном човеку (**поглавље 4**). Ови и други резултати су сумирани и продискутовани у последњем поглављу доктората – **поглављу 9**, те се оном кога је прича из овог резимеа заинтригирала саветује да почне читање доктората управо са **поглављем 9**, а ако има неког ко је много амбициозан, он нек крене од почетка, од **поглавља 1**!

Acknowledgments

After five years of working on one project, now when its final outcome – this PhD thesis – is ready, I would like to thank people that have helped me to finish my ambitious PhD project, but also people who made my out of work hours “gezellig”. I have really enjoyed the previous five years of my life, and I completely agree with an observation of a former colleague of mine, Kriztina Gábor, who noticed that “Wageningen has a very fortunate climate; it helps to seed friendships, to grow scientific passion and to harvest good memories”.

First of all, my thanks go to Willem who was the most important person for growing my scientific passion here in Wageningen. Willem, it has been a great honour for me to do a PhD under your supervision. I have experienced you as an optimistic, creative, motivating and supportive supervisor, which made my PhD study easier and significantly improved the quality of the outcome of my research and writing. I want to thank you for all that, but also for believing in the HITChip project and for finding all kinds of alternative ways to financially support my research.

Hans, I was very fortunate to collaborate with you during my entire PhD study. You contributed a lot to the fact that I am finishing my PhD thesis, first of all by explaining me the principals of all molecular techniques during my first day in the Laboratory of Microbiology. You certainly have quite some teaching talent! I also have to thank you for the enormous amount of lab work you did, based on which a large part of this thesis is written. Besides all these (important) work related issues, I would like to thank you for your friendship during all these years. Thanks for translating (& writing) the Dutch summary of this thesis and finally, thanks for accepting to be my paranymph.

Muriel, as you know, when I came to NL for the first time, as a student, I was quite lonely and unhappy. Luckily, when I came back and started my PhD you had just started your PhD project within the same EU cluster, you were sitting in the same lab, and living in the same house like me. Because of this I was never lonely or unhappy! Starting with our trips with Jana and Cornelia in your small Peugeot, it was obvious that this longer stay in NL would be fun. Then we both became mothers within one month, so we could share experiences (both beautiful and those less nice) of motherhood with a lot of understanding for each other. I am happy that I had such a childish but mature friend during all this time. I guess it could not be different, but I am glad that you accepted to be by my side during my PhD defence, so thanks for becoming my paranymph.

Elaine, you are the first WUR person that I was in touch with, and you offered me the PhD position which resulted in this thesis. I want to thank you for recognising me as a good candidate, but also for picking me up upon my arrival to NL and lending me kitchenware and a radio, which made my survival during the first months in NL easier. Furthermore, I would like to thank you for the supervision and support during the first two years of my PhD study.

Hauke, I would like to thank you for your great input, especially during writing of several chapters of my thesis, but also during brainstorming about the HITChip, and your essential help for handling ARB. You were always very efficient (as a real German☺) but at the same time kind, supportive and full of understanding. Thank you for that.

Douwe, your participation in my research project was one of the most essential elements that moved it further (with you we got rid of the static friction). I am very grateful for the database construction, for writing R scripts, for having time and understanding for my questions, and for your interest and will to work on this project in your free time.

Erwin, as the Moleco's expert for the gastrointestinal microbiology you were very interested in my work and your creative ideas were very helpful in solving some technical issues of my HITChip experiments. Thank you for your interest, discussions, tips about propositions, and your contribution in writing our common publications. Also I want to thank you together with Mark and Thomas for the pleasurable lunch breaks and nice discussions.

I was fortunate to be involved in an EU project and to meet quite some nice researchers from all over Europe. For the good atmosphere in the project I am grateful to Pawel, Ralf, Christophe, Linonel, Kim, Bob, Paul, Michel, David and Prof. Michael Blaut. My short involvement in C-012 WCFS project resulted in a fruitful and enjoyable collaboration with Annette and Koen. I want to thank them for that. Furthermore, thanks to girls from Valio: Anu, Kajsa, Riinna and Riitta, it was a real pleasure to meet you and to work with you. Finally, thanks to Aat, Elaine, Max and Rober from Unilever for their interest in the HITChip and a nice collaboration which will, hopefully, result in a joined publication.

I would also like to thank my students, three nice girls: Selena, Liliana and Elena. I really enjoyed spending time and working with all three of you. Especially I want to thank Elena, who I enjoyed supervising but also discussing with, as she was very motivated, enthusiastic, bright and hardworking student. I also want to thank Elena and her professor Patrizia Brigidi, for inviting me to give a seminar in Bologna (my very first invited seminar), and for their hospitality and kindness during my visit to Bologna.

I would also like to specially thank some of many special Moleco peoples: Elientje, I am really happy that we met and became friends. I have learned many things from you, mostly about beer, but also about taking care of our planet, and being better person. I would like to thank you, together with Mark, Maaike and Vesy, for organising our hilarious Aquarius parties! Also, I have enjoyed organising the Japan PhD trip together with you, Elena, Jasper, Mark and David. David and Rocio, my paranymphs in law, I have to thank you together with Neslihan, for the nice atmosphere in our office, and for your understanding and support during my complaining sessions. Davy, special thanks for teaching me how to climb, but also for your (together with Muriel & Enora) flexibility about the babysitting schedule. Rocio, sharing your experience about finishing PhD has given me courage and support when I was "blue" in the last days of my PhD study. Also, I want to thank you for all you did for me

on the day of delivering the reading version of this book. Finally, thank you and Hans for your will to babysit in the evenings. Carien, your positive energy and kindness affect the whole group. Thanks for making such a nice atmosphere! Vesy, I was very lucky to meet you during my PhD study. I want to thank you for your friendship, nice chats, gaming, and all gezelligheid. Also thank you for making me your paranymph – this was a great honour!

Furthermore, I would like to thank all girls that organised and joined my bachelor's party: Carien, Danijela, Darja, Eline, Esther, Hoa, Lidia, Maaike, Mariana, Muriel, Nataša, Neslihan, Pauline, Rocio, Stephanie, Susana, and Vesela. Thanks for the nice and creative present; it will always remind me of that wonderful day and all of you!

It was a real pleasure to share Studentkaamer 2015 with David, Rocio, Neslihan, Kees, Sergey and for a short while with Rozalin. Peoples, thank you being such good roommates.

I want to thank all members of the dynamic Moleco group, for the nice atmosphere at work and outside work, during our drinks in Vlaam, our Christmas dinners, and outs in Texel and Zeeland. In addition to the people I have already mentioned, thanks to Anne, Anna, Carmelina, Diana, Farai, Hermien, Ineke, Jelena, Kaouther, Mahmut, Maria, Markus, Meta, Odette, Petia, Prapha, Rina, Sahar, Toshio, Wilma, and of course – Antoon – whose positive spirit is still very present in the Moleco group.

For their help with paperwork I would like to thank Francis and Nees, and Wim for his incredible skilfulness in solving any kind of computer related problem. Many other members of the Laboratory of Microbiology make this place very enjoyable and among them I would like to mention Miriam, and the organisers of the Labtrip 2005: Bart, René, and Suzanne.

Marjanne, Frank, Joost, Luuk and Florine, I would like to thank you for your hospitality, and more than that, for providing me the first home outside my parents' home. I have really enjoyed the "French style" you have in handling the garden, but also in living life.

Maudi and Jim, we zijn heel gelukkig dat Gavrilo zulke leuke pleegouders heeft. Ik wil jullie bedanken voor de flexibiliteit die jullie toonden als reactie op mijn wat vreemde gevoel voor tijd. Hartelijk bedankt dat jullie zo goed voor Gavrilo hebben gezorgd. Ik weet dat hij bij jullie in goede handen is en ik me volledig en onbezorgd op mijn werk kan storten.

Такође, желим да се захвалим Сузани и професору Јосипу Барасу који су ме увели у свет пробиотичких бактерија. Без такве једне ставке у својој биографији, и без Сузанине препоруке, данас сигурно не бих завршавала овај докторат. Сузо, теби још хвала и за пријатељство, негде између маме и другарице, пуно сам научила од тебе!

Овде, на другом крају Европе, земљаци значајно доприноси квалитету живота, између осталог „DHL" типом услуга за плазму, турску кафу, али и неке важније ствари. Свим тим другарима желим да се захвалим. Дарја, твоје и Касперово пријатељство нам много значи. Супер је што нам бањавате јер са мојим организационим способностима се вероватно никад не бисмо видели. Хвала што увек можемо да рачунамо на вас! Даље, не могу да не поменем моје „тетке из дијаспоре", Марину и Славицу, али и Ану,

Милену и Мишела, и Јелену. Једно специјално хвала је резервисано за Весну и Сашку. Хвала вам за све корисне савете који су ми помогли да лакше пребродим културолошке разлике и да научим да волим да живим у овим ниским земљама. Драго ми је да сам имала срећу да упознам две тако дивне, добронамерне и паметне жене.

Данијела, куме, другару најбољи, мислим да није испало тако лоше што ниси отишла на Оксфорд него у Гронинген. И иако то ниси урадила због мене, драго мије што си ми остала географски близу свих ових година. Доста се тога променило од како смо дошле у ове северне крајеве, али смо ја и ти и даље остале најбољи другари. Много ми значи што имам неког ко може да ме теши и ко ме скоро безусловно воли (иако се тако воле само деца, али ваљда си ти сувише мама у души ☺). Бонџићу, куме, желим да ти се захвалим за све наше занимљиве дискусије и литературу којом си ме снабдео, јер без тога, и поред утицаја Достојавског, ти ми можда данас не би био кум (то што си ме крстио је сигурно једна од најбитнијих ствари која ми се десила у ових пет година). Наташа, без тебе ова листа не би била потпуна, хвала за све лепе тренутке!

Другари са јужнијих меридијана нису ми ништа мање важни, тим пре што су ми остали другари и након пет година велике географске удаљености. Иако свако набрајање делује прилично безлично, ја се надам да свако од вас зна да ми његово пријатељство много значи. Дакле, хвала Мартини, Маријици, Томићу, Цањару, Жагару, Иванчици, Пуђи, Милени, Марији, Владани и Марини.

Даље, желим да се захвалим баки и тетка Бори, што нису заборавиле ни један мој рођендан. Најискреније хвала и Снежи за њену добру вољу да дође и чува Гаврила!

Млађане, и сам знаш да си ти први заслужан за то што сам завршила управо **овај** докторат, јер си баш ти послао моју пријаву за студентску праксу на Вилемову адресу. Наравно, хвала за ингениозну идеју за име мог чипа, за коју се испоставило да је прави хит. Хвала ти што си показао да си прави цар у задњих пола године, кад ме готово и није било код куће, и што си разбио све предрасуде о Балканцима као мушкарцима који немају разумевања за жену каријеристу. Даље, желим да остане, црно на бело, да сам ти захвална што си оставио породицу, другаре, посао и Пирот (врло, врло битно!) да бисмо били заједно. Ти и Гаврило поред мене сте дефинитивно најзаслужнији што сам срећна и задовољна на било којој географској ширини.

На крају бих желела да се захвалим мојим родитељима. Драгане и Вера, иако нисте били директно укључени у израду овог доктората, вама је дугујем захвалност за васпитање, за усађени систем вредности, за љубав, за подршку, и за још много тога без чега не бих била то што сам данас. Највише вам хвала, што сте ми омогућили да растем уз Марка, Милоша и Милицу. За мене је највећа сатисфакција овог доктората што знам да све вас, моју стару породицу, али и моју нову породицу чини поносном.

(An untitled impression of the author of this thesis by Milenko Kukavica, 2000)

About the Author

Mirjana Rajilić-Stojanović (maiden name Mirjana Rajilić) was born on the 9th of February 1978 in Sremska Mitrovica, Serbia (at that time Yugoslavia). Soon after her birth, Mirjana moved with her family to a town called Bugojno in central Bosnia, where she grew up and started her elementary education. After moving back to her birth town in 1990, she completed elementary education and continued studying a secondary school – gymnasium – orientation towards natural sciences.

In 1996 Mirjana started her studies at The University of Belgrade. Despite her original wish to study Pharmacy, she continued the education at The Faculty of Chemical Engineering and Metallurgy at The Department of Biochemical Engineering and Biotechnology. It appeared that the educational direction Mirjana took was actually the one where she could really show all of her learning potential with a lot of enthusiasm. The best proof that a "second best" choice turned into a great success is the fact that Mirjana got awarded as the best student of the whole generation '96/'97 at this faculty. Being a good student ensured her a scholarship during studies, and working position upon graduation in a local pharmaceutical company. Once more, pharmacy was her choice that was not meant to be as she continued her education in microbiology.

During her MSc thesis project Mirjana was involved in the defining process parameters for production of probiotic powder based on the enriched and fermented soy milk. The experience in working with probiotics was essential for getting her involved in the microbiological research, and enabled her a three months long student internship in The Laboratory of Microbiology, Wageningen University in 2001. The objective of her practical training was defining diversity of archaeal community of a bioreactor. Working on this project, she became familiar with molecular biology, and finally got acquainted with her future PhD project, which started in 2002 and resulted in this thesis.

Mirjana is presently working as a PostDoc in The Laboratory of Microbiology, Wageningen University, on a project that represents continuation of her PhD study research.

(An untitled impression about the author of this thesis by Danijela Jovanović, 2007)

Publications' List

Namsolleck P, Thiel R, Lawson PA, Holmstrøm K, **Rajilić** M, Vaughan EE, Rigottier-Gois L, Collins MD, de Vos WM & Blaut M (2004) Molecular methods for the analysis of gut microbiota. Microbial Ecology in Health and Disease 16, 71-85

Lay C, Rigottier-Gois L, Holmstrøm K, **Rajilić** M, Vaughan EE, de Vos WM, Collins MD, Thiel R, Namsolleck P, Blaut M & Doré J. (2005) Colonic microbiota signatures across five northern European countries. Applied and Environtal Microbiology 71: 4153-4155

Rajilić-Stojanović M & de Vos WM (2007) Diversity of the human gastrointestinal microbiota. Agro Food Industry Hi Tech. In press.

Rajilić-Stojanović M, Smidt H & de Vos WM (2007) Diversity of the human gastrointestinal microbiota. Environmental Microbiology. In press.

Rajilić-Stojanović M, HGHJ Heilig, Molenaar D, Smidt H & de Vos WM. A generic design strategy for phylogenetic microarrays based on SSU rRNA gene sequence diversity - the human gastrointestinal microbiota as a case. Submitted.

Rajilić-Stojanović M, Maathuis A, Heilig HGHJ, Venema K, de Vos WM & Smidt H. Evaluating the microbial diversity in an *in vitro* model of the human large intestine by phylogenetic microarray analysis. Submitted.

Surakka A, Kajander A, **Rajilić**-Stojanović M, Mikkola H, Hatakka K, Vapaatalo H, Zoetendal EG, de Vos WM, Korpela R & Tynk-kynen S. Galactooligosaccharide-containing yoghurt relieves the symptoms of constipation and promotes growth of bifidobacteria in elderly people. Submitted.

Rajilić-Stojanović M & de Vos WM Characterised microbes of the human gastrointestinal tract. In preparation.

Rajilić-Stojanović M, Heilig HGHJ, Zoetendal EG & de Vos WM. Dynamics of the adult gastrointestinal microbiota. In preparation.

Rajilić-Stojanović M, Heilig HGHJ, Guarner F, Shanahan F& de Vos WM. Diversity of the human gastrointestinal microbiota in relation to ulcerative colitis. In preparation.

Rajilić-Stojanović M, Biagi E, Kajander K, Smidt H, Korpela R, Zoetendal EG & de Vos WM. Phylogenetic microarray analysis reveals specific faecal microbiota composition in patients suffering from irritable bowel syndrome. In preparation.

Rajilić-Stojanović M, de Vos WM & Zoetendal EG. Accessing uncultivated microorganisms . In preparation.

Kajander K, Myllyluoma E, **Rajilić**-Stojanović M, Kyrönpalo S, Rasmussen M, Ristimäki A, Zoetendal EG, de Vos WM, Vapaatalo H & Korpela R. A probiotic combination reduces abdominal pain and bloating in irritable bowel syndrome. In preparation.

Training and Supervision Plan

Graduate School VLAG	Year	ETCS
Discipline specific activities		
Courses		
Safe handling with radioactive materials and sources, Larenstein, Velp	2002	1.4
ARB course, Laboratory of Microbiology	2003	0.3
Ecophysiology of the gastrointestinal tract, VLAG,Wageningen	2003	1.4
Bioinformatics technology 1, VLAG, Wageningen	2003	2.0
Food fermentation, VLAG, Wageningen	2004	1.0
Genetics and physiology of food associated microorganisms, VLAG, Wageningen	2004	1.0
Meetings		
NL array platform, Utrecht	2002	0.3
ID array symposium, Wageningen	2002	0.3
ID array symposium, Wageningen	2003	0.3
Darmendag, Lelystad	2002	0.3
Darmendag, Utrecht (poster presented)	2004	0.8
Darmendag, Maastricht (poster presented)	2005	0.8
COST meeting, Wädenswil, Switzerland	2002	0.7
COST meeting, Wageningen (oral presentation)	2006	1.1
PROEUHEALTH workshop, Taormina, Italy (poster presented)	2003	1.5
PROEUHEALTH workshop, Sitges, Spain (poster & oral presentation)	2004	1.5
FEMS 1st congress of European microbiologists, Ljubljana, Slovenia (poster presented)	2003	1.9
2nd GutImpact platform meeting, Tallinn, Estonia (oral presentation)	2006	1.3
General courses		
VLAG PhD week 2002	2002	1.1
Organising and supervising MSc thesis projects,OWI, Wageningen	2006	0.8
PhD scientific writing, Centa, Wageningen	2006	1.7
Career orientation, WGS, Wageningen	2006	1.5
Optionals		
Preparing PhD research proposal	2002	6.0
PhD/Postdoc meetings (Laboratory of Microbiology)	2002-6	4.2
Organising PhD trip to Japan	2005	2.5
Total		35.7

One European Credit Transfer System (ECTS) equals 28 hours of study

The research described in this thesis has been carried out with the financial support of the Commission of the European Communities, specific RTD program “Quality of Life and Management of Living Resources”, research project Microbe Diagnostics (QLK1-2000-00108).

Design and Layout: Mirjana Rajilić-Stojanović

Printed by: Grafisch bedrijf Ponsen & Looijen BV, Wageningen

Printing sponsors: Foundation “Gut Flora in Health and Disease”
Akkermansia Foundation
Unilever