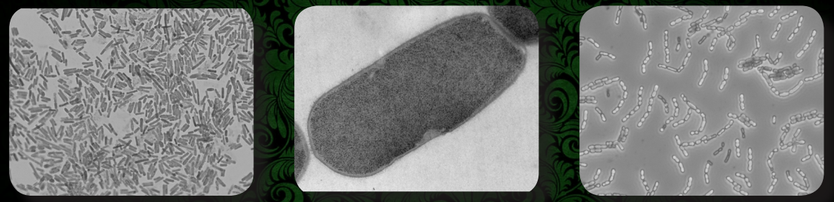

The genus *Romboutsia* - genomic and functional characterization

Jacoline Gerritsen

The genus *Romboutsia*

Genomic and functional characterization of novel bacteria dedicated to life in the intestinal tract

Jacoline Gerritsen

The Bonsai

For hundreds of years the image of a tree has been used as a visual representation of (evolutionary) relationships among organisms ('The Tree of Life'). The leaves of the tree represent individual species, while the splitting branches represent divergence events. Also relationships among bacterial species are often visualized in tree-shaped diagrams.

It is not just a tree that decorates the cover of this thesis, it is a bonsai. The Japanese word 'bonsai' is used to describe a tree which is planted in a shallow container. The bonsai metaphorically depicts the bacterial tree of life, since a number of techniques used for the care of bonsai are similar to those applied in bacterial classification.

Selective removal of parts of a tree, known as pruning, is performed to maintain and refine the shape of the tree. In addition, pruning of leaves (species) is sometimes needed to restrain excessive growth. Furthermore, branches are wired to bend and reposition them into the desired shape. This is similar to the creation of higher-level groups in the bacterial tree of life to support the formation of small branches and new leaves on specific locations in the tree. Grafting, the placement of new growing material into a prepared area on the trunk, resembles the reclassification of bacterial taxa.

Deciding which branches or leaves should stay or be removed can be challenging since there are no rules but only guidelines to support decision making, as is also the case for the classification of bacteria. So in the end, the shapes of the tree is strongly dependent on choices that have been made by the artist/scientist in the creative process.

The genus *Romboutsia*

Genomic and functional characterization of novel bacteria dedicated to life in the intestinal tract

Jacoline Gerritsen

Thesis committee

Promotors
Prof. Dr H. Smidt
Personal chair at the Laboratory of Microbiology
Wageningen University

Prof. Dr W.M. de Vos
Professor of Microbiology
Wageningen University

Co-promotor
Prof. Dr G.T. Rijkers
Professor of Medical Biology and Life Sciences
University College Roosevelt, Utrecht University

Other members
Prof. Dr E.J.M. Feskens, Wageningen University
Prof. Dr I.C.W. Arts, Maastricht University
Dr K. Scott, Rowett Institute of Nutrition and Health, Aberdeen, Scotland
Dr E.E. Vaughan, Sensus, Roosendaal

This research was conducted under the auspices of the Graduate School VLAG (Advanced studies in Food Technology, Agrobiotechnology, Nutrition and Health Sciences).

The genus *Romboutsia*

Genomic and functional characterization of novel bacteria dedicated to life in the intestinal tract

Jacoline Gerritsen

Thesis

submitted in fulfilment of the requirements for the degree of doctor
at Wageningen University
by the authority of the Rector Magnificus
Prof. Dr M.J. Kropff,
in the presence of the
Thesis Committee appointed by the Academic Board
to be defended in public
on Friday 23 January 2015
at 11 a.m. in the Aula.

Jacoline Gerritsen
The genus *Romboutsia* - Genomic and functional characterization of novel bacteria dedicated to life in the intestinal tract.
280 pages

PhD thesis, Wageningen University, Wageningen, NL (2015)
With references, with summaries in English and Dutch

ISBN 978-94-6257-242-3

Table of Contents

1 CHAPTER

General introduction and thesis outline

Intestinal microbes: they outnumber us!

A thesis that focusses on specific members of the bacterial communities residing in the mammalian intestinal tract has to start with a general introduction on intestinal microbiota, the microbial populations living in the intestinal tract. So, let's start with some numbers. The human body contains at least tenfold more microbial cells (10^{14}) than human cells [124]. Microbes colonize every surface that is exposed to the external environment. Of these surfaces, the intestinal tract is by far the most densely colonized. Furthermore, the microbes residing in the intestinal tract add ~150 times more genes to our total gene pool than our own human genome [46, 112]. Maybe these numbers have been quoted too often already, but they still illustrate very nicely that humans, like other mammals, are not single-species organisms, but in fact they constitute very complex ecosystems. The complex network of host-microbe and microbe-microbe interactions is tremendously important for host health, as will be discussed in **Chapter 2.** The microbes in our intestinal tract are our first line of defence against other (pathogenic) microbes and in this sense they supplement our immune system. Microbes are in constant interaction and competition with each other for nutrients and attachment sites [4, 65]. In addition, they are constantly communicating with each other, for example by the secretion of anti-microbial compounds such as short chain fatty acids (SCFA) and bacteriocins [22, 28]. In addition, they are not only communicating with each other, but also with their host. Especially in the intestinal tract microbes have the opportunity to communicate with their host, as the intestinal tract contains most of the body's immune cells, hormone producing cells, and the body's nervous system is embedded throughout the lining of the intestinal tract [87, 98, 149]. Last but certainly not least, the intestinal microbes contribute significantly to the metabolic capacities of the host by producing a wide array of compounds that are important nutrients for the host, such as SCFA and vitamins [39, 103, 161]. Metaphorically speaking, we are maybe nothing more than walking bioreactors; microbes outnumber us by at least one order of magnitude, and maybe the only way to control them is by how and what we feed them.

It is important to mention that (micro)organisms from all three domains of life are found in the intestinal tract: bacteria, archaea, and eukaryotes. The eukaryotes found in the intestinal tract consist of fungi, protozoa, helminths and others [125]. In addition, viruses are found abundantly in the intestinal tract, including bacteriophages, which are viruses that specifically infect bacteria [1, 10]. Future studies will provide much more insight in the role of each of these (micro)organisms in the intestinal ecosystem [153]. However, this thesis will focus only on bacterial members of the intestinal microbiota.

The intestinal tract contains distinct microenvironments

Factors such as nutrient availability, pH, redox potential, host secretions, and peristalsis strongly influence the composition of bacterial communities in the intestinal tract as

will be discussed next. However, the intestinal tract is not simply one big hollow tube that starts after the stomach and ends with the anal canal, but it is in fact divided into several distinct segments. Each segment has its own functionality, which is reflected in the anatomy of that segment. Although the anatomy of the intestinal tract varies greatly between different animals, for the purposes of this thesis, only the mammalian intestinal tract will be discussed in the next paragraphs, with a specific focus on the human intestinal tract.

Along the longitudinal axis, the intestinal tract can roughly be divided into two segments, the small intestine and the large intestine (Figure 1). The small intestine can be further subdivided into the duodenum, jejunum, and ileum. These are regions where most of the enzymatic digestion takes place, and host secretions provide a strong selective pressure on the microbes in these niches [6, 124, 154]. Increasing numbers of microbes have been found along the length of the small intestine [6, 27]. Gastric acid from the stomach ensures a low pH in the small intestine that gradually increases along the intestinal tract, reaching neutral pH at the terminal ileum [35]. Bile and pancreatic secretions containing digestive enzymes enter the intestinal tract in the duodenum. The small intestinal epithelium is extensively folded into crypts, and the luminal surface is covered with finger-like projections (villi) that increase the surface area available for nutrient absorption. The jejunum and ileum are regions where active and passive nutrient absorption takes place by the host. In these regions, the microbes are in strong competition with the host for these nutrients. It is observed that the microbes in the small intestine are adapted to this dynamic environment by their ability to rapidly take in and convert simple carbohydrates [166]. Most of the bile is reabsorbed in the ileum together with vitamin B12, an important vitamin for the host that can only be produced by microbial activities. One of the main characteristics of the ileum is the presence of Peyer's patches, which are organized lymphoid nodules that contain large numbers of lymphocytes and other immune cells [9]. This indicates that the ileum is an important region for immune-system related host-microbe interactions.

After the small intestine, the large intestine is the next segment along the intestinal tract, which can be further divided in the caecum (in some mammals, including humans, this includes an appendix), ascending colon, transverse colon, descending colon, sigmoid colon and rectum (Figure 1). The ileocaecal (IC) valve separates the ileum from the caecum and has two major functions, control of bowel flow towards the caecum and prevention of caecal content reflux in the terminal ileum (caecoileal reflux). Transit time is relatively short in the small intestine due to antegrade peristaltic movements, in humans 30 min to 3.5 h [54]. Transit time has an influence on microbial community composition, because intestinal stasis can result in small intestinal bowel overgrowth (SIBO) [122]. Vice versa, it has been observed that specific microbes can accelerate small intestinal transit time [56]. At this point in time it is still debatable whether the ileal microbiota represents a unique

community. The microbes in the ileum may originate from the caecum via leakage through the IC valve [166]. Contrarily, others have proposed that the ileum serves as inoculum for the colon [83]. The role of the IC valve in health and disease remains largely unknown. It has been hypothesized that by preventing caecoileal reflux, the IC valve protects against the development of bacterial overgrowth. Recently it was observed that patients with SIBO had significantly lower ileocaecal junction pressures as compared to individuals without SIBO and healthy controls [119].

The large intestine is the segment where most of the fermentation by microbes takes place [39, 88] and high microbial densities are reached (10^{10}-10^{12} cells/ml) [6, 27]. From the host's point of view, this is the region where water and any nutrient produced by the microbes are reabsorbed, including essential vitamins. Morphologically the large intestine is much wider than the small intestine and it does not contain villi like the small intestine. After passage through the colon, the undigested part of food is collected in the rectum with microbes and other bodily excretions (e.g. mucus and bile) until it leaves the body through the anal canal.

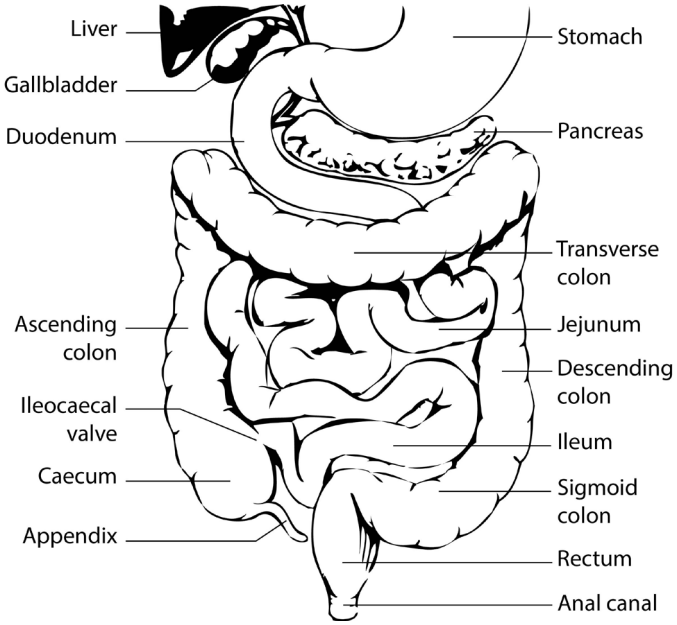

Figure 1. Schematic representation of the anatomy of the human (gastro-)intestinal tract.

Along the radial axis of the intestinal tract, there is a spatial organization in composition of different microbial communities from the lumen to the mucosa. It has been known for some time now that mucosa-associated bacterial communities differ from those in the luminal content [32, 77, 169]. One spatializing factor is the mucus layer that covers the lining of the intestinal tract and physically separates the microbes from the

epithelial cells. Defective functioning of the mucus layers has been associated with diseases such as inflammatory bowel disease (IBD), cystic fibrosis and infectious diseases [62, 69]. Mucus, which is a host secretion, is produced by goblet cells that are distributed unevenly along the intestinal tract [69]. Some commensal microbes interact intimately with the mucus layer and form mucosa-associated bacterial communities [34, 60]. These mucosa-associated microbes thrive on host-provided nutrients, such as mucin and dead epithelial cells, releasing mucosa-associated carbohydrates [99]. In addition, the mucosa-associated microbes are also the first microbes to come in contact with other host-secreted molecules such as hormones, secretory immunoglobulins, or antimicrobial peptides [99, 149]. The properties of the mucus layer differ between the various segments of the intestinal tract. In the colon, for example the mucus layer is composed of two distinct layers of which the one close to the epithelial cells is usually impermeable for microbes [61], whereas in the ileum there is only one layer that appears permeable for microbes [60]. Different mucosa-associated microbial communities can be found in the various segments of the intestinal tract [142]. Another spatializing factor is redox potential. The presence of oxygen close to the epithelial cells prevents strictly anaerobic microbes from living too close to the epithelium. However, the gradients and dynamics of redox potential in the intestinal tract are incompletely understood [18].

Altogether, the presence of distinct microenvironments in the intestinal tract results in a heterogeneous distribution of different microbial communities across and along the intestinal tract. Although this has been acknowledged for decades [124], many regions within the small intestine are still largely unexplored, mainly due to their inaccessibility [6]. However, since more and more is known about the diverse array of host-microbe interactions, it is time we know more about the presence and spread of specific microbial populations throughout the intestinal tract. Spatial distribution is an important factor to take into consideration when describing intestinal microbiota composition and function in order to better understand their niche-specialization and their potential impact on host physiology and health.

So, which bacteria live in our intestinal tract?

When we talk about bacteria, we mainly talk about bacterial species. For a crash-course on taxonomy and related terms see Box 1. Now some more numbers! It is estimated that more than 1000 species-level phylotypes can be found in the intestinal tract of the total human population [112, 117]. While the species richness in any given individual is much lower (~160 species per individual), the diversity of the species that reside in the total (human) population within the intestinal tract is high. Another metaphor: the bacterial tree of life in our intestinal tract is a big tree with only few primary branches, but many small twigs. With respect to microbial composition in the intestinal tract, it is mainly the twigs that differ between individuals. The biggest branches are formed by microbes belonging to just

two phyla, the *Bacteroidetes* and the *Firmicutes*. In addition, also members of other phyla, such as *Proteobacteria*, *Actinobacteria*, *Fusobacteria*, *Spirochaetes*, *Verrucomicrobia* and *Lentisphaerae,* have been detected [117, 167]. The *Firmicutes* is currently the largest bacterial phylum containing more than 200 genera. The majority of the *Firmicutes* detected in the intestinal tract belong to the class *Clostridia*. This thesis will focus on specific members of the intestinal microbiota that belong to the family *Peptostreptococcaceae* within the order *Clostridia*.

Box 1: A journey into classification, taxonomy, phylogeny, and systematics

Why do we keep all our socks in one drawer, and all our underwear in another? Why don't we throw our entire cutlery set into one big drawer, but do we separate them in forks, spoons, knives, teaspoons, etc.? Why? Because it is a way to organize our belongings, not only so that we can find them back easily, but also to make sure other people can easily follow the same strategy. When you think about this you will realize that we all have the inherent need to separate and classify the things around us in recognizable defined groups. Any kind of organisational scheme is a **classification**, whether it concerns sorting cutlery by shape, M&Ms by colour, or bacteria by their properties. Humans seem unable to resist the urge to classify; it is simply easier to think about a few groups of things than about lots of things individually. Also in science, classification is one of the most basic activities implemented in all its fields. In fact, we make a science out of it and we call it **taxonomy**: the science of classification.

In the field of biology, the term taxonomy is used for the classification of organisms in an ordered system. Taxonomy is traditionally divided into three separate but interrelated parts: identification, classification, and nomenclature (naming of things). Since there is a great diversity of organisms, biologists use the characteristics of organisms to describe the different organisms, to identify new ones and to classify them in distinct groups (taxa) and give them appropriate names. Once biologists have examined a number of characteristics of different organisms, they attempt to classify organisms based on their similarities and dissimilarities. Just like in any other field of biology, taxonomy should not simply be overlooked in microbiology. One could say that taxonomy is the science that makes communication possible among scientists, physicians, and all people who need to know and use microbiological information. Accurate and standardized names are a prerequisite for effective communication.

The main goal of taxonomy nowadays is to produce a formal system for naming and classifying species that illustrates their evolutionary relatedness. The term **phylogeny** (or phylogenetics) is used when evolutionary relatedness is explored among groups of organisms. Taxonomy is more and more guided by phylogeny, but it is a distinct discipline. To make things even more complicated, the degree to which taxonomy depends on phylogeny depends on the school of taxonomy.

The term **systematics** is often used synonymously with taxonomy. However, systematics is the study of diversification of living things and encompasses much more than taxonomy (and phylogeny). Systematics includes identification, classification, nomenclature (so taxonomy) and phylogeny, but also characterization, culture preservation and biodiversity.

The current hierarchical classification system

There are multiple ways in which a group of organisms can be classified. To date, we still use the hierarchical system designed by Carolus Linnaeus in the 18th century for classification. This is a rank-based classification system, which means it consists of a hierarchic series of ranks. This hierarchical classification system is based on the principle that members of higher-level groups share fewer characteristics than those in lower-level groups. Therefore, ranks are assigned based on dissimilarity, but the assessment of ranks is rather subjective [73]. There are seven main taxonomic ranks; kingdom, phylum, class, order, family, genus, and species. The basic ranks of the system are the species and the genus, according to the binomial system introduced by Linnaeus. Each distinct species is assigned to a genus using a binary name, consisting of a generic name and a species epithet (e.g. *Homo sapiens* or *Escherichia coli*).

In addition to the seven main taxonomic ranks, a rank named 'domain', positioned above kingdom, has become popular in recent years. Currently, the three-domain system, proposed in 1990, is the most widely accepted system [160]. In this system there is one eukaryotic domain (*Eukarya* or *Eukaryota*) and two prokaryotic domains: *Bacteria* (previously known as *Eubacteria*) and *Archaea* (previously known as *Archaebacteria*).

Composition and functionality: single organism *vs* complex communities

There are several approaches one can use to characterize organisms and to study their functional potential (Figure 2). Once an organism is isolated and growing in mono-culture, the organism can be identified and classified using approaches that will be discussed later on and an example of this approach will be discussed in **Chapter 4**. The taxonomic position of a certain organism in the tree of life can provide some insight into its functionalities by looking at its relatives. However, this should be done with great care since relatives can differ quite a lot from each other as will also be discussed later in this chapter. Whole genome sequence analysis of a single strain provides specific details on that particular strain (examples of this approach will be discussed in **Chapter 5** and **Chapter 6**). By looking at gene presence/absence, predictions can be made on potential functionalities of an organism. For example, it can be predicted which metabolic pathways are present in an organism, which is indicative for which substrates may be utilized by an organism and which metabolites may be produced. Sequencing technologies have advanced so much in the last few years that it is nowadays straightforward and easily affordable to sequence

large sets of complete microbial genomes [130]. A genome sequence can also provide insight in mechanisms by which an organism is adapted to a certain environment. For example, concerning the intestinal tract these mechanisms might be related to interaction with the host, or competition with other microbes. Also pathogenicity may be predicted from the genome based on the presence of certain virulence related genes. Gene presence/absence can therefore provide insight in the genetic potential of a particular organism.

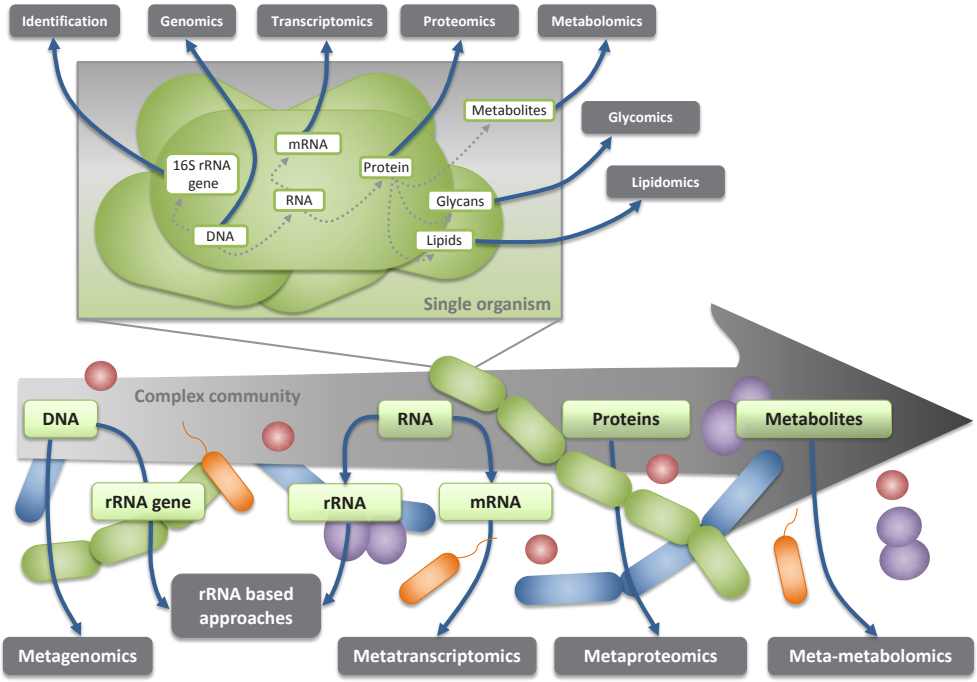

Figure 2. Overview of (similar) methods that can be used to study composition and function of single organisms or complex communities.

To study the conditions at which certain activities are expressed, one can look at gene expression profiles by transcriptome analysis and an example of this approach will be discussed in **Chapter 5**. In the past whole transcriptome analyses were predominantly done by microarray analysis. However, sequencing of RNA sequence libraries (RNA-seq) has now become more and more affordable and also more popular [86, 152]. While microarrays rely heavily on knowledge on the genome of an organism, RNA-seq does not have that limitation. Transcriptome analysis will provide a snapshot of presence and quantity of RNAs at a given moment in time, which can be used to gain insight into what an organism is aiming to achieve. A proteomic approach will tell more about what an organism really managed to achieve by looking at which proteins were produced in a certain condition.

There are many methods available to study proteins, but protein microarrays and mass spectrometry-based methods are nowadays the main methods used for large-scale studies on proteins [11, 106]. However, proteomics is much more complicated to study than genomics or even transcriptomics, because one gene or one transcript can result in more than one protein. Moreover, many proteins may have gone through a wide range of post-translational modifications, even though this is not as much an issue in bacteria as it is in eukaryotes, where post-translational modification of proteins is a rule rather than an exception. Finally, a metabolomics approach provides insight into which small-molecule metabolites are produced by the cells and an example of this approach will be discussed in **Chapter 5** [151]. Metabolites are the intermediates and end products of metabolism. Separation methods that can be used are gas chromatography (GC), high performance liquid chromatography (HPLC), or capillary electrophoresis (CE), which can be coupled to mass spectrometry for detection [40, 42]. Nuclear magnetic resonance (NMR) spectrometry is another method widely used for metabolomics [84, 133]. With these methods one can also look specifically at the whole set of glycans (glycomics) or lipids (lipidomics) produced by a microbe [24, 78, 79, 94]. However, the main limitations of these approaches are related to the enormous diversity of compounds within a sample (complexity) and the identification of individual compounds.

The laboratory is probably the only environment where microbes are living in isolation. If you think it is difficult to predict the functional contribution of a single organism, imagine how hard this will be for a complex community. However, most of the approaches that can be followed for single organisms can now also be applied to whole communities (Figure 2), mainly thanks to advances in (predominantly sequence) technologies.

When we study the structure and function of complex microbial communities there are several questions we can ask, like 'who is present?', 'with how many are they present?' and 'what are they doing?'. To answer the first two questions, both culture-dependent and culture-independent approaches can be used to study community composition and diversity. One can 'simply' try to isolate all the microbes present in a community and identify and characterize them using the methods described for single organisms. However, this is very labour-intensive and for many microbes *in vitro* cultivating has not yet been achieved [2, 53, 139]. Complex communities can now also be characterized using culture-independent molecular methods, as was also done in the work described in **Chapter 3** and **Chapter 7**. Molecular fingerprinting techniques have been developed, like temperature or denaturing gradient gel electrophoresis (TGGE or DGGE), and terminal-restriction fragment length polymorphism (T-RFLP), which are in general based on sequence-dependent separation of PCR-amplified 16S ribosomal RNA (rRNA) gene fragments. These methods have been very informative for studying changes in community composition and diversity over time [168]. However, a major breakthrough came with the advances in

sequence technologies that made it possible to get more information per sample (the amount of sequence information that can be delivered per sample is still increasing) at a decreasing price [3]. These new sequence technologies are highly multiplexed, thereby allowing for simultaneous sequencing and analysis of multiple samples at the same time. Although culture-independent methods provide comprehensive information on community structure, they will tell us little about community functionality and the methods cannot be used to answer the question 'what are they doing?'. To this end, functional microbiomic approaches are available [114, 150]. Metagenomic approaches provide a first step in assessing the functional potential of a (complex) community, by sequencing all the genetic information in environmental samples [46, 95, 127]. Metatranscriptomic approaches can complement DNA-based metagenomic analyses by identifying gene expression patterns of active members within a microbial community [45]. Similarly metaproteomic and meta-metabolomic approaches can provide insight in proteins and metabolites, respectively, produced *in situ* by a complex community [42, 50, 150, 163]. However, the aforementioned limitations of these techniques regarding the study of single organisms are even bigger regarding the study of complex communities.

Large-scale (functional) microbiomic studies are needed to fully understand the impact of the intestinal microbes on the host. This means that larger numbers of samples need to be analysed, deeper sequencing is needed, longer sequence reads are preferred for better taxonomic or genomic assignment, and more extensive comparative analyses are needed. Integration of this broad array of microbiomic approaches will help us to understand the functional contribution of specific intestinal microbes to host physiology and pathology. However, more importantly, they will help to understand how we can influence host physiology and pathology by steering the functionality of the ecosystem.

Culture-independent approaches are not the Holy Grail

Although the last few years a lot of emphasis has been on the advantages of whole community analysis using molecular approaches, we have also become increasingly aware of its limitations. To date, it is still difficult to connect functionality in an ecosystem to specific sets of genes and to individual microbial species, and vice versa [93]. For each sequence-based method it is necessary to compare sequences to reference databases in order to link sequences to specific organisms and/or to assign a function to them. However, these reference databases are still far from comprehensive and big gaps are still present, mainly because we are lacking available genome sequences for cultured strains or even missing cultured representatives for specific groups or functions [74]. It turns out we can only learn so much by studying complex communities as a whole and there is a growing need to learn more about the individuals [4]. We need genomic and functional analyses with single organisms as well as simple and defined communities in order to get an overall insight in the genetic and functional potential of specific members of a

microbial community and to unravel their interactions with other microbes and their host [38]. Hence, studies based on single organisms and whole-community analysis can complement each other. Also in intestinal microbiota-related studies these limitations have become increasingly evident, and there is an increasing need for culture isolates of intestinal origin. So, we have to go out and isolate! This is the only way in which we can fully understand the functional contribution of specific intestinal microbes to host physiology and health.

Prokaryotic taxonomy

In the current hierarchical classification system used by (micro)biologists, species is the most important unit for classification of new isolates. Taxonomists working with higher organisms apply a different definition of the term 'species' than microbiologists. For higher organisms, species are groups of interbreeding or potentially interbreeding natural populations that are reproductively isolated from other groups. Although this definition may be satisfactory for organisms capable of sexual reproduction, especially for animals, it fails completely in the case of bacteria [19, 30, 33, 73]. Bacteria divide asexually and for the most part do not show regionalism ('everything is everywhere'). A basic principle in taxonomy that we all seem to agree on, is that members of higher-level groups share fewer characteristics than those in lower-level groups. Ideally, organisms should be classified by their evolutionary relationships. Currently, most microbiologists follow a cladistic approach for bacterial taxonomy, assuming that classification corresponds to phylogenetic descent (a clade consisting of an ancestor and all its descendants), and that all valid taxa are monophyletic. Although it is questionable whether microbial evolution can be framed in a concept similar to that used by biologists, species delineation in microbiology does have a practical use, because it helps us to discuss about important groups of microbes. For example, medical treatments, and responses to food contaminations and bioterrorism are largely based on species-level identification of the microbes involved.

Throughout the years, different approaches have been used to identify and classify organisms [73, 126, 144]. Classical approaches make use of morphological, physiological and metabolic characteristics of organisms (Table 1). With this approach, microbes are grouped based on easily-observed phenotypic characteristics (e.g. cell morphology, Gram-staining, motility, structural features) and on distinguishing physiological features (e.g. carbon source utilization patterns, growth characteristics). Since analytic tools have been available for characterisation of biochemical properties of cells, microbes are also being grouped based on chemotaxonomic characteristics such as cellular fatty acid or polar lipid composition (Table 1). With the development of molecular techniques, characterization based on genotypic characteristics has now advanced into the field of taxonomy as well (Table 1). Since its development in 1970's, nucleic acid sequencing

methods have evolved into tools widely used by the whole scientific community. Also for the field of prokaryotic taxonomy, classification of microbes based on genetic characteristics has been revolutionary [120, 144]. Sequence analyses of genes encoding highly conserved molecules are nowadays widely used to classify and identify microbes, but also to follow large-scale evolutionary changes and to study phylogeny. Firstly, microbes can be grouped based on nucleic base composition (G+C content). Secondly, nucleic acid hybridization methods can be used to determine the similarity of strains at a genomic level. DNA-DNA hybridization (DDH) is currently still the gold standard for species delineation using a 70% threshold for DDH [12, 157], but this value should not be used as a strict species boundary [147]. Thirdly, sequencing of conserved protein-encoding genes, typically housekeeping genes, can provide high resolution information on the relatedness of strains. With the advances in sequencing technologies, sequencing of whole genomes is now also advancing as routine practice for the description of novel taxa. However, how genomics can be integrated in the taxonomy and systematics of prokaryotes is still up for discussion [17, 66, 73, 90, 118].

Table 1. Overview of characteristics that are used for the identification and classification of microbes

Classical characteristics: phenotypic characterization
Morphological characteristics
Cell shape, cell size, cell organization, colony morphology, ultrastructural characteristics, staining behaviour, mechanisms of motility (e.g. pili and flagella), endospore morphology and location, cellular inclusions, colour
Physiological and metabolic characteristics
General nutritional type, carbon and nitrogen sources, energy sources, mechanisms of energy conversion, fermentation products, growth temperature optimum and range, pH optimum and growth range, luminescence, salt requirements and tolerance, osmotic tolerance, oxygen relationships, photosynthetic pigments, sensitivity to metabolic inhibitors and antibiotics
Biochemical composition of the cell: chemotaxonomic characterization
Fatty acid composition, polar lipid composition, respiratory lipoquinones, amino acid composition of the peptidoglycan of the cell wall (Gram-positive bacteria), presence and size of mycolic acids, polyamine pattern, whole cell proteins
Molecular characteristics: genotypic characterization
Nucleic base composition (G+C content), nucleic acid hybridization (DNA-DNA or DNA-RNA), nucleic acid sequences (ranging from the sequences of conserved protein-encoding genes such as ribosomal RNA genes to complete genomes)

Currently a polyphasic taxonomic approach is the most widely accepted method used to describe novel taxa [144]. This approach is based on the assumption that all possible valid data should be employed in determining classification. This approach includes phenotypic, chemotaxonomic and genotypic data, as well as phylogenetic information. The polyphasic approach has proven its value in bacterial taxonomy, as is also evident from the work presented in **Chapter 4**. However, this approach for classification is very labour-intensive and can therefore not cope with the growing number of novel strains that are isolated each day. In addition, this approach cannot be applied in culture-

independent studies in which genetic information is the only characteristic available. To this end, other approaches are necessary.

There is no official classification of prokaryotes, because taxonomy remains a matter of scientific judgement and general agreement. However, the most widely accepted classification is the "Taxonomic Outline of the Prokaryotes" initiated in the early 1990s. There are generally accepted rules and guidelines for identification, classification and nomenclature (see Box 2 for details on this last topic). The accepted reference for the identification of bacteria is commonly referred to as "Bergey's", officially known as "The Bergey's Manual of Systematic Bacteriology". All new validly described bacterial taxa and taxonomic revisions can be found in the regularly updated List of Prokaryotic Names with Standing in Nomenclature (LPSN) [102].

Box 2: In the end it is all about giving it a name

The International Code of Nomenclature of Bacteria, or in short Bacterial Code, regulates the scientific names for prokaryotes, both for *Bacteria* and *Archaea* [75]. The base-line for bacterial names is the Approved Lists of Bacterial Names, which had a starting point in 1980. At that time it was decided to make a completely new start for nomenclature of bacteria. New bacterial names are reviewed by the International Committee on Systematics of Prokaryotes (ICSP) and published in the International Journal of Systematic and Evolutionary Microbiology (IJSEM). This includes the subjective changes in names whenever novel insights stimulate taxonomists to change the rank of taxa (as done in **Chapter 4**). Taxa above the rank of class are not covered by the Rules of the Bacterial Code (1990 Revision).

The Bacterial Code states that naming is based on nomenclatural types defined as "*that element of the taxon with which the name is permanently associated*". The type of a species is a designated strain, the type strain. A type strain is defined as follows: "*A type strain is made up of living cultures of an organism which are descended from a strain designated as the nomenclatural type. The strain should have been maintained in pure culture and should agree closely in its characters with those in the original description*". For the name of a (novel) species to be validly published, the following criteria have to be full-filled: 1) the name has to be published in the appropriate journal together with a description of the species, and 2) a culture of the type strain should be deposited permanently in established culture collections from which it would be readily available.

16S rRNA gene based classification

One molecule in particular has become THE marker used for taxonomy and phylogeny: the small subunit (SSU) ribosomal RNA (rRNA) molecule (for bacteria this is the 16S rRNA molecule). The use of the SSU rRNA molecule as phylogenetic marker dates back to 1977 [100, 159]. This molecule has been selected as a suitable phylogenetic marker mainly

because it is ubiquitously distributed and orthologous among all forms of life, functionally constant, genetically stable, of a size that is ideal for sequence analysis, and possesses independently evolving domains within the molecule. The use of the SSU rRNA gene as phylogenetic marker in combination with the advancements in sequence technologies has been a major breakthrough for the study of bacterial community composition and diversity [100]. Whole communities can now be characterized at once without the need for cultivation. Although there are a few issues to consider regarding 16S rRNA-gene based taxonomy (Box 3), the 16S rRNA gene currently is still the gene to study once you have isolated a novel strain.

As previously mentioned, DDH is still widely accepted as the gold standard for species delineation using a threshold of 70% DNA-DNA reassociation. It has been found that strains with more than 70% DNA-DNA reassociation usually have more that 97% 16S rRNA gene sequence identity [137]. Based on this observation, 97% 16S rRNA gene sequence identity is the threshold above which DDH experiments are recommended to determine relatedness between different strains [144]. However, more recent work has indicated that a 16S rRNA gene sequence identity of 98.2-99.0% may be a more appropriate threshold above which other methods, such as DDH, should be used for species delineation [89, 136]. Although 16S rRNA gene sequencing is nowadays widely used for bacterial classification, it has low resolution power at the species-level and in some cases even at the genus-level [59]. There are, for example, numerous cases where strains share more than 99% 16S rRNA gene sequence identity, but where classification into different species can be justified based on phenotypic/biochemical characteristics and DNA homology [59]. With the access to the genome sequences of microbes that contain multiple copies of the 16S rRNA gene, one came to realize that intrageneric heterogeneity can be substantial, varying from 0.06% to 9.73% sequence heterogeneity [104, 141].

Box 3: 16S rRNA gene-based taxonomy (or more accurate taxonomies)

Although the 16S rRNA gene is a relatively simple marker, only ~1,500 base pairs in size, there are several issues that arise once it is being used as THE phylogenetic marker.

Usually new 16S rRNA gene sequences are submitted to a public database; GenBank (United States), EMBL (United Kingdom) and DDBJ (Japan) collaborate to collect all sequence data reported, including 16S rRNA gene sequences. Currently there are several publically available databases dedicated to the collection of 16S rRNA gene sequences (mainly derived from the synchronized GenBank/EMBL/DDBJ collection). These include the Ribosomal Database Project (RDP) database [20], the SILVA database [111, 113], the EzTaxon database [68], and the Greengenes database [26], although the latter one is currently not updated anymore. Each database is maintained by its own team of researchers that tries to make sense of a similar collection of sequences. However, these collections contain millions of sequences and are increasing in size by the day, mainly by the addition of big

datasets of environmental sequences. One additional database to mention is the database from "The All-Species Living Tree" Project (LTP). The LTP database provides curated entries of high quality and manually checked alignments of all sequenced type strains [164].

As might be expected, they are different approaches for obtaining a hierarchical classification system, which results in different bacterial taxonomies. Currently each 16S rRNA gene-based bacterial classification system takes into account the Bergey's taxonomy supplemented with updated information from the List of Prokaryotic Names with Standing in Nomenclature (LPSN) on the publication of new taxa and taxonomic revisions [102]. However, the analysis of environmental communities has led to the discovery of an increasing number of taxa not covered by Bergey's or LPSN. For example, there are several phyla which have no known cultivable representative, and many sequences remain unclassified as candidate divisions [52, 53]. Assessment of candidate divisions involves a lot of manual curation of the sequences. A big issue with respect to this is how to assess which sequences represent true taxa, because not all sequences are of the same quality. Another complicating factor is the presence of chimera's, which are hybrid sequences derived from multiple parent sequences that represents a major artefact in PCR-based analyses. Different taxonomies exist, because there are different classification systems that use alternative approaches to define taxonomic groupings and to deal with candidate divisions. The taxonomy provided by NCBI is maybe the most widely consulted taxonomy. However, one must keep in mind that this taxonomy is kept up to date through user submissions that might not always be accurate. Currently the two main taxonomies that have strong underlying 16S rRNA gene-based classification systems are the RDP [155] and SILVA taxonomy [113]. However, there are some significant differences between the two. RDP uses the Bergey's taxonomy as a roadmap; therefore it is maybe the more conservative of the two as it is heavily influenced by historical groupings. SILVA on the other hand applies a tree-based approach using a seed tree constructed with 16S rRNA gene sequence of a representative set of organisms to which new sequences are added. In addition, differences arise from the fact that different sequence alignment methods are used that will result in different classifications. Most discrepancies between the different classifications systems exist within the candidate divisions. In order to all agree on one 16S rRNA gene-based taxonomy at a certain point, we need to fill in the gaps in 16S rRNA gene sequence databases. On this point, the "sequencing orphan species" (SOS) initiative has been very valuable. This initiative set out to sequence the 16S rRNA genes of all type strains that were not represented in the public databases yet, or re-sequence those that were of low sequence quality. This initiative has resulted in the addition of 351 type strain 16S rRNA gene sequences [165]. Future efforts, including the isolation of novel representative strains, will help to improve bacterial taxonomy.

A short family history: *Peptostreptococcaceae*

And now we move on to a taxonomic challenge in practice: the order *Clostridiales*. In fact, the current order *Clostridiales* is a text-book example of how a taxonomic group has turned into a large heterogeneous collection of organisms over the years [115]. Recently a nomenclatural issue concerning priority of name was resolved by renaming the order to *Eubacteriales* (**Chapter 4**). However, another issue that is more difficult to solve, originates from the fact that many taxa within this order do not comply with the current assumption in cladistics-based taxonomy that taxa should be monophyletic. You do not have to know much about taxonomy to notice that this group is taxonomically problematic; or less scientifically said a big mess. This includes members of the current family *Peptostreptococcaceae*. Many species within this family carry the genus name *Clostridium*, which implies a taxonomic position within the genus *Clostridium*. However, nothing could be further from the truth. There are many taxa with similar taxonomic issues, but the genus *Clostridium* is currently (still) one of the largest bacterial genera (it contains ~200 species) and the taxonomic issues poses problems not only for taxonomists, but also for other fields of life sciences. This includes the medical field, because numerous pathogens carry the genus name *Clostridium*.

To understand where the issues with the taxonomy of the genus *Clostridium* originate, we have to start with a bit of history. The name *Clostridium* is derived from the Greek noun '*kloster*' meaning 'spindle'. Adam Prazmowski is credited as being the first person ever to isolate and characterize a *Clostridium* species, namely *Clostridium butyricum,* the current type strain of the genus [109]. In his Ph.D. thesis Adam Prazmowski reintroduced the name *Clostridium* based on a previous designation by Trécul [31]. In the century that past since that time, many bacterial species have been assigned to the genus. With the current knowledge on the heterogeneity in morphology, physiology, and genetics of these species, it is difficult to understand how this single genus has grown so big. However, historically this should not surprise anybody. An organism only had to fulfil a few criteria to be placed in the genus *Clostridium*: (1) presence of endospores, (2) strictly anaerobic metabolism, and (3) inability to reduce sulfate to sulfite [47]. Later a Gram-positive cell wall structure was added as a fourth criterion. This classification resulted in an enormous collection of anaerobic Gram-positive rods capable of forming endospores. In addition to this broad classification scheme, a further issue was the fact that new species were added to the group just because the closest-relative was classified in the genus. All together these are the ingredients for challenging taxonomy; one short look at the group of organisms shows already that this broad classification scheme does not justify a genus-level grouping [115, 138].

For many of the species currently included in the genus *Clostridium* their genus assignment is debatable. The criteria previously used to distinguish *Clostridium* species from others do not seem to apply to all current members of the genus [47, 115]. Spores are

only rarely demonstrated in some species (e.g. *Clostridium perfringens* and *Clostridium ramosum*), some species are aerotolerant and can grow on agar media exposed to air (e.g. *Clostridium tertium* and *Clostridium histolyticum*), and some clostridia consistently stain Gram-negative (e.g. *Clostridium ramosum* and *Clostridium clostridioforme*). But most importantly, not all taxonomically validly described *Clostridium* species form phylogenetically coherent clusters. 16S rRNA gene sequenced based efforts have made it clear that *Clostridium* species were/are scattered throughout the order *Clostridiales*. Some of the species fall in range of other established genera or even families as is the case for the *Clostridium* species that fall within the family *Peptostreptococcaceae* [115].

Early attempts at reclassification of the genus *Clostridium* were proposed in the mid-1990s. Almost half of currently described *Clostridium* species, all members of a single clade, were assigned to one group referred to as group I, or cluster I, and later as *Clostridium* sensu stricto [21, 63, 158]. In addition to cluster I, Collins *et al.* proposed the grouping of the other *Clostridium* species into distinct phylogenetic clusters based on large scale comparisons of 16S rRNA gene sequences [21]. In the latest edition of Bergey's, which was published in 2009, clostridial classification was extensively revised in line with those proposed clusters and more than 50 bacteria previously placed in the genus *Clostridium* have been reassigned to other taxonomic groups [115]. For the family *Peptostreptococcaceae* this meant that at this point in time, the family contained only three genera: the type genus *Peptostreptococcus* and the genera *Filifactor* and *Tepidibacter* [36], although it was recognized that many *Clostridium* species were also part of this family. Despite the many (mainly family-level) reassignments, many of the species still retained their *Clostridium* genus name. For many taxonomically problematic groups reclassifications at a genus-level are still missing. Within the genus *Clostridium,* only ten species have been reclassified to other genera since the latest revision of the taxonomic roadmap published in the latest edition of Bergey's (Table 2), half of which has been done in line with the research presented in this thesis (**Chapter 4**). Therefore most of the species are still listed as *Clostridium* species in the LPSN and sequence databases (Box 3). For the family *Peptostreptococcaceae* this means that, although novel genera have been added to the family based on a novel species description (*Sporacetigenium* [15] and *Anaerosphaera* [146]), there is still a big need for taxonomic reassignment with corresponding changes in species names for the *Clostridium* (and *Eubacterium*) species that fall within this family.

It is a problem that novel species are still being added to the genus *Clostridium*. Since the latest edition of Bergey's, 15 species have been validly added to the genus *Clostridium* (Table 3). For some of them this might be legitimate because they fall within the clade of *Clostridium* cluster I, even though also this clade is probably still too big to be a genus-level clade. However, for others this has not been correct because they clearly belong to other genera. The last species that has been incorrectly added to the genus *Clostridium* was

published in 2009, not soon after the latest edition of Bergey's was published, indicating that the addition of species to the genus *Clostridium* is no longer allowed in cases where it is obvious that they belong to other taxa. However, novel taxa are constantly proposed outside the IJSEM journal. Although only some of these incorrect taxa will be taken up by the validation lists as validly described taxa, their names can be found in scientific literature and will be used by others who are not aware of these invalid names.

Table 2. Overview of species that have been reclassified to genera outside the genus *Clostridium* since the last edition of Bergey's (following taxonomic assessment and with official change in species name).

Old classification	Cluster	Reclassification	Reference
Clostridium coccoides	XIVa	*Blautia coccoides*	82
Clostridium proteoclasticum	XIVa	*Butyrivibrio proteoclasticus*	92
Clostridium lentocellum	XIVb	*Cellulosilyticum lentocellum*	13
Clostridium orbiscindens	IV	*Flavonifractor plautii*	14
Clostridium hathewayi	XIVa	*Hungatella effluvii*	67
Clostridium lituseburense	XI	*Romboutsia lituseburensis*	44
Clostridium bartlettii	XI	*Intestinibacter bartlettii*	44
Clostridium glycolicum	XI	*Terrisporobacter glycolicus*	44
Clostridium mayombei	XI	*Terrisporobacter mayombei*	44
Clostridium irregulare	XI	*Asaccharospora irregularis*	44

In many instances taxonomic assignments have caused confusion, even in scientific literature. The problem becomes more and more evident in large culture-independent studies, because these studies are largely 16S rRNA gene-sequence driven. In order to give taxonomic assignment to the thousands of sequences obtained in these studies, each researcher makes its own interpretation on how to assign taxonomy to these sequences, predominantly driven by the chosen classification system (Box 3). For the family *Peptostreptococcaceae* this means that many of the sequences that we assume to belong to the family *Peptostreptococcaceae,* are assigned by others to family-level groups like '*Clostridium* Cluster XI', 'Unclassified *Clostridiales*', or sometimes even '*Clostridiales* Family XI: *Incertae Sedis*', the latter being an example of incorrect assignment due to confusing taxa names. Because 16S rRNA gene sequences of family *Peptostreptococcaceae* species have been classified to different groups, their overall abundance and diversity in intestinal and environmental samples has been underestimated.

As Garrity *et al.* [43] already stated: *'Yet there is considerable reluctance among many contemporary systematic prokaryotic biologists to place new species and genera into higher taxa, especially at the intermediate levels (family, order, and class) because of uncertainty of phylogenetic models'*. The very limited number of proposed taxonomic revisions for these ranks supports this statement. This is mainly due to the lack of rules for delineating higher taxa. In practice, there is considerable subjectivity in assigning species to a genus or higher level groups, and taxonomists may disagree about the composition of

genera and higher level groups because these are in fact subjective concepts. However, one of the consequences is that species are assigned to a class or domain, without proper assignment to the intermediate levels, which is highly confusing for database driven taxonomic assignment. Examples for this are the genera *Proteocatella* [107] and *Acetoanaerobium* [134], that are not assigned to a family by the LSPN, but are assigned to the family *Peptostreptococcaceae* by some of the 16S rRNA gene sequence databases, and to other family-level taxa by others. However, for a positive point of view, this problem is now widely acknowledged and in the future taxonomic revisions can be expected to solve this problem. And it might very well be that in the (near) future somebody is brave enough to raise the family *Peptostreptococcaceae* to another level.

Table 3. Novel species that have been added to the genus *Clostridium* since the latest edition of Bergey's. This includes species that should have been assigned to other genera since they do not belong to Cluster I as defined by Collins et al. [21].

Species	Cluster	Reference
Clostridium amylolyticum	I	135
Clostridium algoriphilum	I	129
Clostridium arbusti	I	64
Clostridium cavendishii	I	8
Clostridium hydrogeniformans	I	8
Clostridium sulfidigenes	I	123
Clostridium tagluense	I	140
Clostridium chromiireducens	I	57
Clostridium alficaecis	I	162
Clostridium bornimense	I	48
Clostridium huakuii	I	121
Clostridium caenicola	III	132
Clostridium clariflavum	III	132
Clostridium sufflavum	III	97
Clostridium lavalense	XIVa	29

It is time to meet the family members: the *Peptostreptococcaceae*

The family *Peptostreptococcaceae* currently contains a wide array of genera and species (Figure 3). Members of the family *Peptostreptococcaceae* are ubiquitously present in very diverse (anaerobic) habitats, such as soil, water, sewage, and aquatic sediments. However, many of them have also been detected in the intestinal tract of humans and other mammals [116]. The family *Peptostreptococcaceae* is most known for its members that are associated with compromised health and disease. Some of the members are (opportunistic) pathogens with a clearly documented history of causing disease. Most of the pathogenic *Peptostreptococcaceae* species owe their virulence to their toxigenicity, meaning that they kill (host) cells by creating pores in the membranes of the cells. Well-studied for its pathogenic properties is *Clostridium difficile*, a major nosocomial pathogen and the main causative agent of antibiotic-associated diarrhoea [105]. Pathogenic *C. difficile*

strains produce the potent cytotoxins toxin A (TcdA) and toxin B (TcdB), and some strains produce an additional binary toxin CDT [108]. Asymptomatic colonization by *C. difficile* is common [55]. First of all, not all strains of *C. difficile* are toxigenic. Secondly, people in good health usually do not suffer from infections by *C. difficile*, because *C. difficile* will normally only cause intestinal damage and disease in individuals that have a perturbation of their intestinal microbiota related to compromised health. Other pathogenic members of the family *Peptostreptococcaceae* are *Clostridium sordellii* and *Clostridium bifermentans* [108]. Also for these two species it must be noted that not all strains are toxigenic. In addition to the *Clostridium* species within the family *Peptostreptococcaceae*, also some of the *Peptostreptococcus* species have been associated with disease. *Peptostreptococcus stomatis* and *Peptostreptococcus anaerobius*, the latter being the type strain of the genus *Peptostreptococcus* and the family *Peptostreptococcaceae*, have shown to be involved in (systemic) infections such as endocarditis and infections of the genitourinary and gastrointestinal tracts, and they have been implicated to play a role in oral diseases [72]. Recently *Peptostreptococcus* genus members were also found to be enriched in the intestinal microbiota of colorectal cancer patients [16, 156].

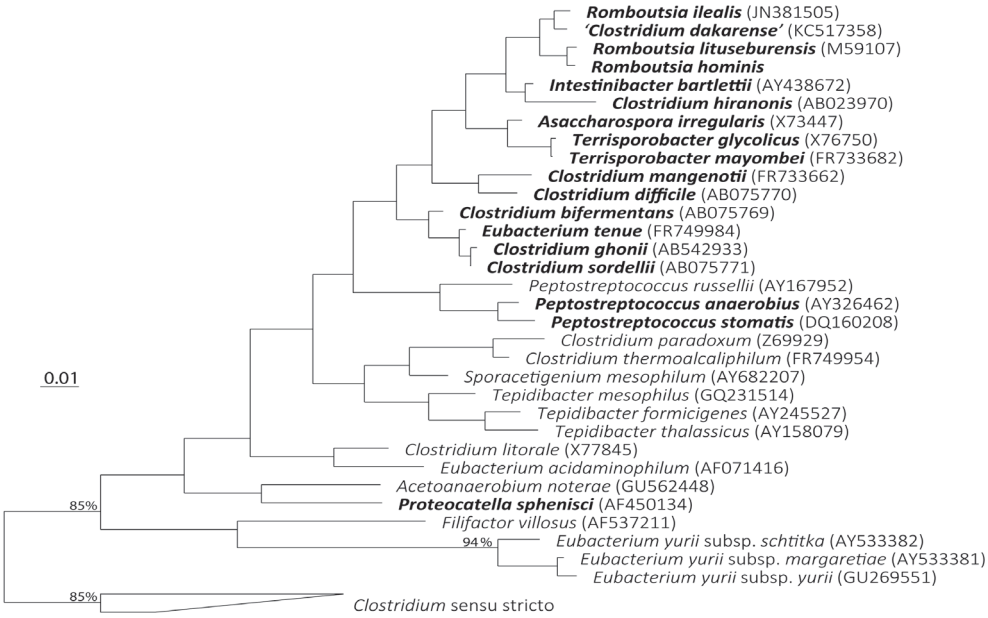

Figure 3. Neighbour-joining tree of the family *Peptostreptococcaceae* based on 16S rRNA gene sequence data. Species that have been detected in the intestinal tracts of humans and/or mammals are indicated in bold. The 16S rRNA gene sequences were aligned using the SINA aligner [110]. The phylogenetic tree was reconstructed using ARB [85] using the "The all-species living tree" project dataset release LTPs115 [164]. Only bootstrap values >70% are shown at branch nodes. Bar represents 1% sequence divergence. GenBank accession numbers are given in parentheses. The 16S rRNA gene sequences of several *Clostridium* sensu stricto species were used as an outgroup.

Just like in every family, a few black sheep are not a good reflection of the capabilities of the other family members. To date, the presence and diversity of *Peptostreptococcaceae* species in the intestinal tract have been underestimated. This is partly due to classification issues using modern day high-throughput sequencing and database driven classification, as has been discussed above. However, specific family members have repeatedly been reported as commensal members of the intestinal microbiota as will be shown in the following chapters. Little is known on the role of commensal *Peptostreptococcaceae* in the intestinal tract, and this family has not received much attention with respect to its potential symbiotic capabilities. In this thesis the characterization of members of the family *Peptostreptococcaceae* will be described in addition to the isolation of novel members. Furthermore, their functional potential and their possible role in the intestinal tract will be discussed.

Microbes in health and disease: "the good, the bad and the deadly"

Without doubt we can conclude that microbes play an important role in human disease, as they can be highly pathogenic and cause serious diseases, or can even be lethal. The existence of 'bad' and 'deadly' microbes is known by the general public, but the fact that there are also 'good' ones is not so commonly accepted. As will be discussed in **Chapter 2**, the microbes that are part of the complex microbial communities and reside in the intestinal tract can contribute positively to the health of the host. Only recently it became clear that diseases are not only caused by the presence of certain microbes, but also by their absence. The number of diseases in which the absence of specific microbes or the absence of specific microbial activities has been associated with the aetiology of the disease is increasing (as will be discussed in **Chapter 2**). These insights have led to the general believe that specific microbes can have beneficial effects on human health. For these beneficial microbes the term 'probiotic' was introduced. Currently, the 13-year-old FAO/WHO definition of probiotics, which was slightly improved recently, remains the most widely used: "live microorganisms that when administered in adequate amounts confer a health benefit on the host" [37, 51]. However, its meaning has seen several different interpretations over the years, and the exact definition is still a constant source of debate at conferences and in scientific literature. For example, the inclusion of 'live' in the definition is ambiguous since there are examples where microorganisms have beneficial effects without being alive [80, 143, 148]. In addition, in some cases it is difficult to determine viability [25].

Probiotics have their roots in fermented food products. More specifically in fermented dairy products (e.g. yoghurt), sour products made from fresh milk by the metabolic activity of lactic acid forming bacteria. Fermented milks have been consumed by humans for centuries. However, it is Ilya Ilyich Mechnikov (1845-1916), also known as Élie Metchnikoff, who is credited for being the grandfather of modern day probiotics.

1

At the turn of the twentieth century, the Russian immunologist studied Russian and Bulgarian farmers with long life spans. He proposed a link between the life span and their diet, consisting of daily consumption of large amounts of yoghurt. While working together with a group of researchers from the Louis Pasteur Institute in Paris, Metchnikoff proposed that lactic acid bacteria, when present in sufficient numbers within the intestinal tract, could prolong life by preventing deterioration related to old age [91]. He recognised the possibility of modifying the microbial populations in the human intestinal tract by replacing the harmful microbes by those that are more useful. His studies were based on the observations of Stamen Girgorov (1878-1945), a Bulgarian physician and microbiologist. He discovered in 1905 that a certain strain of *Bacillus* known as 'Bacterium bulgaricum – Grigoroff' or 'Bacillus bulgaricus (Grigoroff)', is responsible for the souring of milk, resulting in yoghurt. This organism is probably the same that became known as *Lactobacillus bulgaricus*, later renamed as *L. delbrueckii* subsp. *bulgaricus*. Around the same time, Henry Tissier, a French paediatrician, observed that the stools of children suffering from diarrhoea contained a low number of peculiar Y-shaped 'bifid' bacteria that were abundant in healthy children[145]. Metchnikoff and Tissier were thus among the first who made scientific suggestions about the potential beneficial use of microbes, specifically bifidobacteria, in relation to disease. The origin of the term 'probiotic' is credited to the German bacteriologist and food scientist Werner Kollath. He used the term 'Probiotika', which is derived from Greek and means 'for life', in his publication entitled "Nutrition and the tooth system" 71. Kollath used it to describe organic and inorganic supplements that are necessary to improve health in patients suffering from malnutrition resulting from excessive consumption of highly refined food. So, at that time, the term was not used in connection to microbes. In 1965, the term probiotics was first used in contrast to the term antibiotics: "substances secreted by one organism which stimulate the growth of another" 81. Later, Parker was the first to use the term in the context of having a beneficial effect on the host 101. In 1989, Fuller was the first to put it in the context of live microbes by removing the word substances from the definition and replacing it with "live microbial feed supplements" 41. Since then, the definition has evolved over the years and we have for sure not settled on a final definition yet.

The association between probiotics and fermented (dairy) products still stands today, as several bacterial strains that are currently on the probiotic market, are isolated from fermented (dairy) products. However, it has recently been proposed to refrain from using the term 'probiotics' for microbes that are solely used in fermentation of milk and have no other health-promoting effect 51. Since the focus is nowadays moving more into the direction of health benefits instead of viable cell counts, an important target of current probiotic research is to find novel effective strains. Decades of research have led to the realisation that microbes of intestinal origin may also be considerably beneficial to the host 70. However, we now realize that microbes cannot simply be classified as either 'good',

'bad' or 'deadly', since this turns out to be a false trilemma. Although these qualifications of microbes are generally done at the species-level, the characteristics that make them fall in either category can differ at the strain-level. As mentioned above, it is known that not all strains within a species are pathogenic. The number of examples, where non-pathogenic strains of generally accepted pathogenic species have been found to protect against infection by respective virulent strains, is growing. For example, the species *Escherichia coli* is well-known for its pathogenic capacity to cause diarrhoea and extra-intestinal diseases [23]. Based on these characteristics members of this species are generally accepted to be 'bad' microbes. However, it has become clear that the genetic diversity of a given species can be vast, and both pathogenic and non-pathogenic members of a certain species can be part of the commensal microbiota [76]. This notion is exemplified by an *E. coli* strain, Nissle 1917, that is currently available on the market as a probiotic strain. A German professor named Alfred Nissle used this strain to treat patients with infectious diarrhoea in an era when antibiotics were not available yet. He isolated the strain from the faeces of a soldier, who was the only one of his unit not suffering from infectious diarrhoea, and later the strain was named after him and the year the strain was isolated. Nowadays *E. coli* strain Nissle 1917 is considered to be one of the 'good' ones and is used for the prevention of infectious diarrhoea and the management of inflammatory bowel disease (IBD) [5, 58].

The same applies to the class *Clostridia*, which in general has a notorious reputation and members of this class have been considered to be 'bad', if not 'deadly'. However, also within this class, examples of potentially beneficial counterparts of pathogenic strains can be found. An example comes from *Clostridium difficile*, a well-known toxigenic pathogen. Colonisation by non-pathogenic strains of *C. difficile* has been demonstrated to protect against the colonisation by toxigenic strains [7, 96, 131]. The fact that some strains within a species are pathogenic does not mean that it is impossible for other strains of that same species to have beneficial properties. There is growing evidence that these non-pathogenic strains can be beneficial by similar mechanisms as 'traditional' probiotics. For example, some strains of the species *Clostridium butyricum* have pathogenic properties due to their ability to produce toxins [108]. However, not all *C. butyricum* have the capacity to produce toxins, and some atoxigenic *C. butyricum* strains have even shown to have probiotic potential, mainly attributed to their ability to produce butyrate (a short-chain fatty acid that is beneficial to the host) and their interaction with the host immune system [49]. In Japan, a particular strain of *C. butyricum* (MIYAIRI) has been used already for years for the prevention of antibiotic-associated diarrhoea [128].

Aims and outline of the thesis

The research described in this thesis aimed to provide a better understanding of the *Peptostreptococcaceae* residing in the intestinal tract of both animals and humans. To

date, their presence and role in the intestinal ecosystem has been underestimated, partly due to classification issues using modern day high-throughput sequencing and database driven classification. The research in this thesis presents an example of how studies on single organisms and whole-community analyses can complement each other in helping us to fully understand the impact of intestinal microbes on (human) health.

Chapter 2 provides a review of the literature describing the composition of the intestinal microbiota and gives an overview of diseases related to changes in composition and diversity of the intestinal microbiota. In addition, this chapter provides insight in how probiotics can be used to modulate the intestinal microbiota, focussed on influencing the composition and diversity. An example of the latter is provided in **Chapter 3** where a multispecies probiotic mixture (Ecologic® 641) was demonstrated to modulate the abundance of specific species in the ileum of rats. These changes were correlated to an improved health in an animal model for acute pancreatitis. Interestingly the abundance of one specific phylotype was associated with an improved disease outcome. In **Chapter 4** the isolation of this specific phylotype from a rat ileal sample is described in combination with its characterization for taxonomic purposes. In this chapter the taxonomy of several closely related members of the family *Peptostreptococcaceae* was revised in order to provide a valid systematic name to the isolate, *Romboutsia ilealis* CRIB^T. **Chapter 5** focuses more in-depth on *R. ilealis* CRIB^T based on genomic insights, with specific attention to its carbohydrate metabolism. It shows how this microbe is adapted to a life in the small intestine. In addition, it provides an example of a study in which genomic and functional characterization of a single organism can provide ideas for how stimulation of the abundance and/or activity of this specific organism may be achieved by dietary intervention when present in a complex community. **Chapter 6** describes the research performed to gain more insight in current members of the genus *Romboutsia.* A novel *Romboutsia* species was isolated from human ileostoma effluent, *Romboutsia hominis* FRIFI. A comparative genomics analysis was done for all known *Romboutsia* species in order to gain more insight in their functional contributions as (potential) intestinal microbe. Finally, **Chapter 7** gives a summary of this thesis in combination with a general discussion of the observations made in the different experimental chapters. The chapter also contains unpublished data on the work that has been done on the characterization of *R. ilealis* CRIB^T as a potential probiotic strain. Furthermore, this chapter provides preliminary results from a study that was aimed to examine the differences in microbiota composition in various regions of the human small intestine in comparison to faecal microbiota composition, in addition to changes in microbiota composition as result of a synbiotic intervention.

The overall aim of this thesis therefore is to identify and to characterize specific groups of commensal bacteria in the intestinal tract that may potentially be exploited for their probiotic properties.

References

1. Abeles S.R. and D.T. Pride (2014). Molecular bases and role of viruses in the human microbiome. *J Mol Biol*.
2. Amann R.I., W. Ludwig and K.H. Schleifer (1995). Phylogenetic identification and in situ detection of individual microbial cells without cultivation. *Microbiol Rev*; 59 (1): 143-169.
3. Bahassi E.M. and P.J. Stambrook (2014). Next-generation sequencing technologies: breaking the sound barrier of human genetics. *Mutagenesis*; 29 (5): 303-310.
4. Ballal S.A., C.A. Gallini, N. Segata, C. Huttenhower and W.S. Garrett (2011). Host and gut microbiota symbiotic factors: lessons from inflammatory bowel disease and successful symbionts. *Cell Microbiol*; 13 (4): 508-517.
5. Behnsen J., E. Deriu, M. Sassone-Corsi and M. Raffatellu (2013). Probiotics: properties, examples, and specific applications. *Cold Spring Harb Perspect Med*; 3 (3): a010074.
6. Booijink C.C., E.G. Zoetendal, M. Kleerebezem and W.M. de Vos (2007). Microbial communities in the human small intestine: coupling diversity to metagenomics. *Future microbiology*; 2 (3): 285-295.
7. Borriello S.P. and F.E. Barclay (1985). Protection of hamsters against *Clostridium difficile* ileocaecitis by prior colonisation with non-pathogenic strains. *J Med Microbiol*; 19 (3): 339-350.
8. Bowman K.S., R.E. Dupre, F.A. Rainey and W.M. Moe (2010). *Clostridium hydrogeniformans* sp. nov. and *Clostridium cavendishii* sp. nov., hydrogen-producing bacteria from chlorinated solvent-contaminated groundwater. *Int J Syst Evol Microbiol*; 60 (2): 358-363.
9. Brandtzaeg P., H. Kiyono, R. Pabst and M.W. Russell (2008). Terminology: nomenclature of mucosa-associated lymphoid tissue. *Mucosal Immunol*; 1 (1): 31-37.
10. Breitbart M., M. Haynes, S. Kelley, F. Angly, R.A. Edwards, B. Felts, J.M. Mahaffy, *et al.* (2008). Viral diversity and dynamics in an infant gut. *Res Microbiol*; 159 (5): 367-373.
11. Breker M. and M. Schuldiner (2014). The emergence of proteome-wide technologies: systematic analysis of proteins comes of age. *Nat Rev Mol Cell Biol*; 15 (7): 453-464.
12. Brenner D.J. (1973). Deoxyribonucleic acid reassociation in the taxonomy of enteric bacteria. *Int J Syst Bacteriol*; 23 (4): 298-307.
13. Cai S. and X. Dong (2010). *Cellulosilyticum ruminicola* gen. nov., sp. nov., isolated from the rumen of yak, and reclassification of *Clostridium lentocellum* as *Cellulosilyticum lentocellum* comb. nov. *Int J Syst Evol Microbiol*; 60 (4): 845-849.
14. Carlier J.P., M. Bedora-Faure, G. K'Ouas, C. Alauzet and F. Mory (2010). Proposal to unify *Clostridium orbiscindens* Winter et al. 1991 and *Eubacterium plautii* (Seguin 1928) Hofstad and Aasjord 1982, with description of *Flavonifractor plautii* gen. nov., comb. nov., and reassignment of *Bacteroides capillosus* to *Pseudoflavonifractor capillosus* gen. nov., comb. nov. *Int J Syst Evol Microbiol*; 60 (3): 585-590.
15. Chen S., L. Song and X. Dong (2006). *Sporacetigenium mesophilum* gen. nov., sp. nov., isolated from an anaerobic digester treating municipal solid waste and sewage. *Int J Syst Evol Microbiol*; 56 (4): 721-725.
16. Chen W., F. Liu, Z. Ling, X. Tong and C. Xiang (2012). Human intestinal lumen and mucosa-associated microbiota in patients with colorectal cancer. *PloS one*; 7 (6): e39743.
17. Chun J. and F.A. Rainey (2014). Integrating genomics into the taxonomy and systematics of the *Bacteria* and *Archaea*. *Int J Syst Evol Microbiol*; 64 (2): 316-324.
18. Circu M.L. and T.Y. Aw (2012). Intestinal redox biology and oxidative stress. *Semin Cell Dev Biol*; 23 (7): 729-737.
19. Cohan F.M. (2002). What are bacterial species? *Annu Rev Microbiol*; 56: 457-487.
20. Cole J.R., Q. Wang, J.A. Fish, B. Chai, D.M. McGarrell, Y. Sun, C.T. Brown, *et al.* (2014). Ribosomal Database Project: data and tools for high throughput rRNA analysis. *Nucleic Acids Res*; 42 (Database issue): D633-642.
21. Collins M.D., P.A. Lawson, A. Willems, J.J. Cordoba, J. Fernandez-Garayzabal, P. Garcia, J. Cai, *et al.* (1994). The phylogeny of the genus *Clostridium*: proposal of five new genera and eleven new species combinations. *Int J Syst Bacteriol*; 44 (4): 812-826.
22. Corr S.C., C. Hill and C.G. Gahan (2009). Understanding the mechanisms by which probiotics inhibit gastrointestinal pathogens. *Adv Food Nutr Res*; 56: 1-15.
23. Croxen M.A., R.J. Law, R. Scholz, K.M. Keeney, M. Wlodarska and B.B. Finlay (2013). Recent advances in understanding enteric pathogenic *Escherichia coli*. *Clin Microbiol Rev*; 26 (4): 822-880.
24. Cummings R.D. and J.M. Pierce (2014). The challenge and promise of glycomics. *Chem Biol*; 21 (1): 1-15.
25. Davis C. (2014). Enumeration of probiotic strains: Review of culture-dependent and alternative techniques to quantify viable bacteria. *J Microbiol Methods*; 103: 9-17.
26. DeSantis T.Z., P. Hugenholtz, N. Larsen, M. Rojas, E.L. Brodie, K. Keller, T. Huber, *et al.* (2006). Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. *Appl Environ Microbiol*; 72 (7): 5069-5072.

27. Dethlefsen L., P.B. Eckburg, E.M. Bik and D.A. Relman (2006). Assembly of the human intestinal microbiota. *Trends Ecol Evol*; 21 (9): 517-523.
28. Dobson A., P.D. Cotter, R.P. Ross and C. Hill (2012). Bacteriocin production: a probiotic trait? *Appl Environ Microbiol*; 78 (1): 1-6.
29. Domingo M.C., A. Huletsky, M. Boissinot, M.C. Helie, A. Bernal, K.A. Bernard, M.L. Grayson, *et al.* (2009). *Clostridium lavalense* sp. nov., a glycopeptide-resistant species isolated from human faeces. *Int J Syst Evol Microbiol*; 59 (3): 498-503.
30. Doolittle W.F. and O. Zhaxybayeva (2009). On the origin of prokaryotic species. *Genome Res*; 19 (5): 744-756.
31. Dürre P. (2001). From Pandora's box to cornucopia: clostridia – a historical perspective. In: *Clostridia: biotechnology and medical applications*. pp 1-17. Edited by H. Bahl and P. Dürre. Weinheim, Germany: Wiley-VCH Verlag GmbH.
32. Eckburg P.B., E.M. Bik, C.N. Bernstein, E. Purdom, L. Dethlefsen, M. Sargent, S.R. Gill, *et al.* (2005). Diversity of the human intestinal microbial flora. *Science*; 308 (5728): 1635-1638.
33. Ereshefsky M. (2010). Microbiology and the species problem. *Biol Philos*; 25 (4): 553-568.
34. Ermund A., J.K. Gustafsson, G.C. Hansson and A.V. Keita (2013). Mucus properties and goblet cell quantification in mouse, rat and human ileal Peyer's patches. *PloS one*; 8 (12): e83688.
35. Evans D.F., G. Pye, R. Bramley, A.G. Clark, T.J. Dyson and J.D. Hardcastle (1988). Measurement of gastrointestinal pH profiles in normal ambulant human subjects. *Gut*; 29 (8): 1035-1041.
36. Ezaki T. (2009). Family VII. *Peptostreptococcaceae* fam. nov. In: *Bergey's Manual of Systematic Bacteriology*. 2nd Edn, 3: The *Firmicutes*, pp 1008. Edited by P. De Vos, G.M. Garrity, D. Jones, N.R. Krieg, *et al.* New York: Springer.
37. FAO/WHO (2001).Food and Agricultural Organization of the United Nations and World Health Organization. Health and nutritional properties of probiotics in food including powder milk with live lactic acid bacteria. In. Cordoba, Argentina.
38. Flint H.J., S.H. Duncan, K.P. Scott and P. Louis (2007). Interactions and competition within the microbial community of the human colon: links between diet and health. *Environ Microbiol*; 9 (5): 1101-1111.
39. Flint H.J., K.P. Scott, S.H. Duncan, P. Louis and E. Forano (2012). Microbial degradation of complex carbohydrates in the gut. *Gut Microbes*; 3 (4): 289-306.
40. Fuhrer T. and N. Zamboni (2014). High-throughput discovery metabolomics. *Curr Opin Biotechnol*; 31C: 73-78.
41. Fuller R. (1989). Probiotics in man and animals. *J Appl Bacteriol*; 66 (5): 365-378.
42. Gao P. and G. Xu (2014). Mass-spectrometry-based microbial metabolomics: recent developments and applications. *Anal Bioanal Chem*.
43. Garrity G.M., J.A. Bell and T.G. Lilburn (2004). Taxonomic outline of the prokaryotes Bergey's Manual of Systematic Baceteriology. Second edition, release 5.0. New York: Springer-Verlag.
44. Gerritsen J., S. Fuentes, W. Grievink, L. van Niftrik, B.J. Tindall, H.M. Timmerman, G.T. Rijkers, *et al.* (2014). Characterization of *Romboutsia ilealis* gen. nov., sp. nov., isolated from the gastro-intestinal tract of a rat, and proposal for the reclassification of five closely related members of the genus *Clostridium* into the genera *Romboutsia* gen. nov., *Intestinibacter* gen. nov., *Terrisporobacter* gen. nov. and *Asaccharospora* gen. nov. *Int J Syst Evol Microbiol*; 64 (5): 1600-1616.
45. Gilbert J.A. and M. Hughes (2011). Gene expression profiling: metatranscriptomics. *Methods Mol Biol*; 733: 195-205.
46. Gill S.R., M. Pop, R.T. Deboy, P.B. Eckburg, P.J. Turnbaugh, B.S. Samuel, J.I. Gordon, *et al.* (2006). Metagenomic analysis of the human distal gut microbiome. *Science*; 312 (5778): 1355-1359.
47. Gottschalk G., J.R. Andreesen and H. Hippe (1981). The genus *Clostridium* (Nonmedical Aspects). In: *The prokaryotes*. Vol. 2, pp 1767-1803. Edited by M.P. Starr, H. Stolp, H.G. Trüper, A. Balows, *et al.* Berlin, Heidelberg, New York: Springer.
48. Hahnke S., J. Striesow, M. Elvert, X.P. Mollar and M. Klocke (2014). *Clostridium bornimense* sp. nov., isolated from a mesophilic, two-phase, laboratory-scale biogas reactor. *Int J Syst Evol Microbiol*; 64 (8): 2792-2797.
49. Hayashi A., T. Sato, N. Kamada, Y. Mikami, K. Matsuoka, T. Hisamatsu, T. Hibi, *et al.* (2013). A single strain of Clostridium butyricum induces intestinal IL-10-producing macrophages to suppress acute experimental colitis in mice. *Cell Host Microbe*; 13 (6): 711-722.
50. Hettich R.L., C. Pan, K. Chourey and R.J. Giannone (2013). Metaproteomics: harnessing the power of high performance mass spectrometry to identify the suite of proteins that control metabolic activities in microbial communities. *Anal Chem*; 85 (9): 4203-4214.
51. Hill C., F. Guarner, G. Reid, G.R. Gibson, D.J. Merenstein, B. Pot, L. Morelli, *et al.* (2014). Expert consensus document. The International Scientific Association for Probiotics and Prebiotics consensus statement on the scope and appropriate use of the term probiotic. *Nat Rev Gastroenterol Hepatol*; 11 (8): 506-514.
52. Hugenholtz P., B.M. Goebel and N.R. Pace (1998). Impact of culture-independent studies on the emerging

phylogenetic view of bacterial diversity. *J Bacteriology*; 180 (18): 4765-4774.
53. Hugenholtz P., C. Pitulle, K.L. Hershberger and N.R. Pace (1998). Novel division level bacterial diversity in a Yellowstone hot spring. *J Bacteriology*; 180 (2): 366-376.
54. Hung G.U., C.C. Tsai and W.Y. Lin (2006). Development of a new method for small bowel transit study. *Ann Nucl Med*; 20 (6): 387-392.
55. Hung Y.P., J.C. Lee, H.J. Lin, H.C. Liu, Y.H. Wu, P.J. Tsai and W.C. Ko (2014). Clinical impact of *Clostridium difficile* colonization. *J Microbiol Immunol Infect*.
56. Husebye E., P.M. Hellstrom, F. Sundler, J. Chen and T. Midtvedt (2001). Influence of microbial species on small intestinal myoelectric activity and transit in germ-free rats. *Am J Physiol Gastrointest Liver Physiol*; 280 (3): G368-380.
57. Inglett K.S., H.S. Bae, H.C. Aldrich, K. Hatfield and A.V. Ogram (2011). *Clostridium chromiireducens* sp. nov., isolated from Cr(VI)-contaminated soil. *Int J Syst Evol Microbiol*; 61 (11): 2626-2631.
58. Jacobi C.A. and P. Malfertheiner (2011). Escherichia coli Nissle 1917 (Mutaflor): new insights into an old probiotic bacterium. *Dig Dis*; 29 (6): 600-607.
59. Janda J.M. and S.L. Abbott (2007). 16S rRNA gene sequencing for bacterial identification in the diagnostic laboratory: pluses, perils, and pitfalls. *J Clin Microbiol*; 45 (9): 2761-2764.
60. Johansson M.E., D. Ambort, T. Pelaseyed, A. Schutte, J.K. Gustafsson, A. Ermund, D.B. Subramani, *et al.* (2011). Composition and functional role of the mucus layers in the intestine. *Cell Mol Life Sci*; 68 (22): 3635-3641.
61. Johansson M.E., J.M. Larsson and G.C. Hansson (2011). The two mucus layers of colon are organized by the MUC2 mucin, whereas the outer layer is a legislator of host-microbial interactions. *Proc Natl Acad Sci USA*; 108 Suppl 1: 4659-4665.
62. Johansson M.E., H. Sjovall and G.C. Hansson (2013). The gastrointestinal mucus system in health and disease. *Nat Rev Gastroenterol Hepatol*; 10 (6): 352-361.
63. Johnson J.L. and B.S. Francis (1975). Taxonomy of the *Clostridia*: ribosomal ribonucleic acid homologies among the species. *J Gen Microbiol*; 88 (2): 229-244.
64. Jung M.Y., I.S. Park, W. Kim, H.L. Kim, W.K. Paek and Y.H. Chang (2010). *Clostridium arbusti* sp. nov., an anaerobic bacterium isolated from pear orchard soil. *Int J Syst Evol Microbiol*; 60 (9): 2231-2235.
65. Kaiko G.E. and T.S. Stappenbeck (2014). Host-microbe interactions shaping the gastrointestinal environment. *Trends Immunol*.
66. Kämpfer P. and S.P. Glaeser (2012). Prokaryotic taxonomy in the sequencing era - the polyphasic approach revisited. *Environ Microbiol*; 14 (2): 291-317.
67. Kaur S., M. Yawar, P.A. Kumar and K. Suresh (2014). *Hungatella effluvii* gen. nov., sp. nov., an obligately anaerobic bacterium isolated from an effluent treatment plant, and reclassification of *Clostridium hathewayi* as *Hungatella hathewayi* gen. nov., comb. nov. *Int J Syst Evol Microbiol*; 64 (3): 710-718.
68. Kim O.S., Y.J. Cho, K. Lee, S.H. Yoon, M. Kim, H. Na, S.C. Park, *et al.* (2012). Introducing EzTaxon-e: a prokaryotic 16S rRNA gene sequence database with phylotypes that represent uncultured species. *Int J Syst Evol Microbiol*; 62 (3): 716-721.
69. Kim Y.S. and S.B. Ho (2010). Intestinal goblet cells and mucins in health and disease: recent insights and progress. *Curr Gastroenterol Rep*; 12 (5): 319-330.
70. Kleerebezem M. and E.E. Vaughan (2009). Probiotic and gut lactobacilli and bifidobacteria: molecular approaches to study diversity and activity. *Annu Rev Microbiol*; 63: 269-290.
71. Kollath W. (1953). [Nutrition and the tooth system; general review with special reference to vitamins]. *Dtsch Zahnarztl Z*; 8 (11): Suppl 7-16.
72. Kononen E., A. Bryk, P. Niemi and A. Kanervo-Nordstrom (2007). Antimicrobial susceptibilities of *Peptostreptococcus anaerobius* and the newly described *Peptostreptococcus stomatis* isolated from various human sources. *Antimicrob Agents Chemother*; 51 (6): 2205-2207.
73. Konstantinidis K. and E. Stackebrandt (2013). Defining Taxonomic Ranks. In: *The Prokaryotes*. pp 229-254. Edited by E. Rosenberg, E. DeLong, S. Lory, E. Stackebrandt, *et al.*: Springer Berlin Heidelberg.
74. Kyrpides N.C., P. Hugenholtz, J.A. Eisen, T. Woyke, M. Goker, C.T. Parker, R. Amann, *et al.* (2014). Genomic encyclopedia of bacteria and archaea: sequencing a myriad of type strains. *PLoS Biol*; 12 (8): e1001920.
75. Lapage S.P., P.H.A. Sneath, E.F. Lessel, V.B.D. Skerman, H.P.R. Seeliger and W.A. Clark (1992). International Code of Nomenclature of Bacteria, and Statutes of the International Committee on Systematic Bacteriology, and Statutes of the Bacteriology and Applied Microbiology Section of the International Union of Microbiological Societies. Bacteriological code, 1990 revision. Washington, D.C.: ASM Press.
76. Leimbach A., J. Hacker and U. Dobrindt (2013). *E. coli* as an all-rounder: the thin line between commensalism and pathogenicity. *Curr Top Microbiol Immunol*; 358: 3-32.

77. Lepage P., P. Seksik, M. Sutren, M.F. de la Cochetiere, R. Jian, P. Marteau and J. Dore (2005). Biodiversity of the mucosa-associated microbiota is stable along the distal digestive tract in healthy individuals and patients with IBD. *Inflamm Bowel Dis*; 11 (5): 473-480.
78. Li L., J. Han, Z. Wang, J. Liu, J. Wei, S. Xiong and Z. Zhao (2014). Mass spectrometry methodology in lipid analysis. *Int J Mol Sci*; 15 (6): 10492-10507.
79. Li M., L. Yang, Y. Bai and H. Liu (2014). Analytical methods in lipidomics and their applications. *Anal Chem*; 86 (1): 161-175.
80. Li N., W.M. Russell, M. Douglas-escobar, N. Hauser, M. Lopez and J. Neu (2009). Live and heat-killed *Lactobacillus rhamnosus* GG: effects on proinflammatory and anti-inflammatory cytokines/chemokines in gastrostomy-fed infant rats. *Pediatr Res*; 66 (2): 203-207.
81. Lilly D.M. and R.H. Stillwell (1965). Probiotics: growth-promoting factors produced by microorganisms. *Science*; 147 (3659): 747-748.
82. Liu C., S.M. Finegold, Y. Song and P.A. Lawson (2008). Reclassification of *Clostridium coccoides*, *Ruminococcus hansenii*, *Ruminococcus hydrogenotrophicus*, *Ruminococcus luti*, *Ruminococcus productus* and *Ruminococcus schinkii* as *Blautia coccoides* gen. nov., comb. nov., *Blautia hansenii* comb. nov., *Blautia hydrogenotrophica* comb. nov., *Blautia luti* comb. nov., *Blautia producta* comb. nov., *Blautia schinkii* comb. nov. and description of *Blautia wexlerae* sp. nov., isolated from human faeces. *Int J Syst Evol Microbiol*; 58 (8): 1896-1902.
83. Looft T., H.K. Allen, B.L. Cantarel, U.Y. Levine, D.O. Bayles, D.P. Alt, B. Henrissat, *et al.* (2014). Bacteria, phages and pigs: the effects of in-feed antibiotics on the microbiome at different gut locations. *ISME J*; 8 (8): 1566-1576.
84. Ludwig C. and M.R. Viant (2010). Two-dimensional J-resolved NMR spectroscopy: review of a key methodology in the metabolomics toolbox. *Phytochem Anal*; 21 (1): 22-32.
85. Ludwig W., O. Strunk, R. Westram, L. Richter, H. Meier, Yadhukumar, A. Buchner, *et al.* (2004). ARB: a software environment for sequence data. *Nucleic Acids Res*; 32 (4): 1363-1371.
86. Mane S.P., C. Evans, K.L. Cooper, O.R. Crasta, O. Folkerts, S.K. Hutchison, T.T. Harkins, *et al.* (2009). Transcriptome sequencing of the Microarray Quality Control (MAQC) RNA reference samples using next generation sequencing. *BMC Genomics*; 10: 264.
87. Mann E.R. and X. Li (2014). Intestinal antigen-presenting cells in mucosal immune homeostasis: crosstalk between dendritic cells, macrophages and B-cells. *World J Gastroenterol*; 20 (29): 9653-9664.
88. McNeil N.I. (1984). The contribution of the large intestine to energy supplies in man. *Am J Clin Nutr*; 39 (2): 338-342.
89. Meier-Kolthoff J.P., M. Goker, C. Sproer and H.P. Klenk (2013). When should a DDH experiment be mandatory in microbial taxonomy? *Arch Microbiol*; 195 (6): 413-418.
90. Meier-Kolthoff J.P., H.P. Klenk and M. Goker (2014). Taxonomic use of DNA G+C content and DNA-DNA hybridization in the genomic age. *Int J Syst Evol Microbiol*; 64 (2): 352-356.
91. Metchnikoff É. (1908). The prolongation of life: optimistic studies. New York & London,: G. P. Putnam's sons.
92. Moon C.D., D.M. Pacheco, W.J. Kelly, S.C. Leahy, D. Li, J. Kopecny and G.T. Attwood (2008). Reclassification of *Clostridium proteoclasticum* as *Butyrivibrio proteoclasticus* comb. nov., a butyrate-producing ruminal bacterium. *Int J Syst Evol Microbiol*; 58 (9): 2041-2045.
93. Morales S.E. and W.E. Holben (2011). Linking bacterial identities and ecosystem processes: can 'omic' analyses be more than the sum of their parts? *FEMS Microbiol Ecol*; 75 (1): 2-16.
94. Morelle W. and J.C. Michalski (2005). Glycomics and mass spectrometry. *Curr Pharm Des*; 11 (20): 2615-2645.
95. Morgan X.C. and C. Huttenhower (2014). Meta'omic analytic techniques for studying the intestinal microbiome. *Gastroenterology*; 146 (6): 1437-1448 e1431.
96. Nagaro K.J., S.T. Phillips, A.K. Cheknis, S.P. Sambol, W.E. Zukowski, S. Johnson and D.N. Gerding (2013). Nontoxigenic *Clostridium difficile* protects hamsters against challenge with historic and epidemic strains of toxigenic BI/NAP1/027 *C. difficile*. *Antimicrob Agents Chemother*; 57 (11): 5266-5270.
97. Nishiyama T., A. Ueki, N. Kaku and K. Ueki (2009). *Clostridium sufflavum* sp. nov., isolated from a methanogenic reactor treating cattle waste. *Int J Syst Evol Microbiol*; 59 (5): 981-986.
98. O'Hara A.M. and F. Shanahan (2006). The gut flora as a forgotten organ. *EMBO Rep*; 7 (7): 688-693.
99. Ouwerkerk J.P., W.M. de Vos and C. Belzer (2013). Glycobiome: bacteria and mucus at the epithelial interface. *Best Pract Res Clin Gastroenterol*; 27 (1): 25-38.
100. Pace N.R., J. Sapp and N. Goldenfeld (2012). Phylogeny and beyond: scientific, historical, and conceptual significance of the first tree of life. *Proc Natl Acad Sci USA*; 109 (4): 1011-1018.
101. Parker R.B. (1974). Probiotics, the other half of the antibiotic story. *Anim Nutr Health*; (29): 4-8.
102. Parte A.C. (2014). LPSN-list of prokaryotic names with standing in nomenclature. *Nucleic Acids Res*; 42 (Database issue): D613-616.

103. Patterson E., J.F. Cryan, G.F. Fitzgerald, R.P. Ross, T.G. Dinan and C. Stanton (2014). Gut microbiota, the pharmabiotics they produce and host health. *Proc Nutr Soc*: 1-13.
104. Pei A.Y., W.E. Oberdorf, C.W. Nossa, A. Agarwal, P. Chokshi, E.A. Gerz, Z. Jin, *et al.* (2010). Diversity of 16S rRNA genes within individual prokaryotic genomes. *Appl Environ Microbiol*; 76 (12): 3886-3897.
105. Peniche A.G., T.C. Savidge and S.M. Dann (2013). Recent insights into *Clostridium difficile* pathogenesis. *Curr Opin Infect Dis*; 26 (5): 447-453.
106. Petriz B.A. and O.L. Franco (2014). Application of cutting-edge proteomics technologies for elucidating host-bacteria interactions. *Adv Protein Chem Struct Biol*; 95: 1-24.
107. Pikuta E.V., R.B. Hoover, D. Marsic, W.B. Whitman, B. Lupa, J. Tang and P. Krader (2009). *Proteocatella sphenisci* gen. nov., sp. nov., a psychrotolerant, spore-forming anaerobe isolated from penguin guano. *Int J Syst Evol Microbiol*; 59 (9): 2302-2307.
108. Popoff M.R. and P. Bouvet (2013). Genetic characteristics of toxigenic *Clostridia* and toxin gene evolution. *Toxicon*; 75: 63-89.
109. Prazmowski A. (1880). Untersuchungen über die Entwickelungsgeschichte und Fermentwirkung einiger Bacterien-Arten. Leipzig, Germany: Inaugural dissertation.
110. Pruesse E., J. Peplies and F.O. Glockner (2012). SINA: accurate high-throughput multiple sequence alignment of ribosomal RNA genes. *Bioinformatics*; 28 (14): 1823-1829.
111. Pruesse E., C. Quast, K. Knittel, B.M. Fuchs, W. Ludwig, J. Peplies and F.O. Glockner (2007). SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. *Nucleic Acids Res*; 35 (21): 7188-7196.
112. Qin J., R. Li, J. Raes, M. Arumugam, K.S. Burgdorf, C. Manichanh, T. Nielsen, *et al.* (2010). A human gut microbial gene catalogue established by metagenomic sequencing. *Nature*; 464 (7285): 59-65.
113. Quast C., E. Pruesse, P. Yilmaz, J. Gerken, T. Schweer, P. Yarza, J. Peplies, *et al.* (2013). The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res*; 41 (Database issue): D590-596.
114. Raes J. and P. Bork (2008). Molecular eco-systems biology: towards an understanding of community function. *Nat Rev Microbiol*; 6 (9): 693-699.
115. Rainey F.A., B.J. Hollen and A. Small (2009). Genus 1. *Clostridium*. In: *Bergey's Manual of Systematic Bacteriology*. 2nd Edn, Vol. 3: The *Firmicutes*, pp 738-828. Edited by P. De Vos, G.M. Garrity, D. Jones, N.R. Krieg, *et al.* New York: Springer.
116. Rajilic-Stojanovic M. and W.M. de Vos (2014). The first 1000 cultured species of the human gastrointestinal microbiota. *FEMS Microbiol Rev*.
117. Rajilic-Stojanovic M., H. Smidt and W.M. de Vos (2007). Diversity of the human gastrointestinal tract microbiota revisited. *Environ Microbiol*; 9 (9): 2125-2136.
118. Ramasamy D., A.K. Mishra, J.C. Lagier, R. Padhmanabhan, M. Rossi, E. Sentausa, D. Raoult, *et al.* (2014). A polyphasic strategy incorporating genomic data for the taxonomic description of novel bacterial species. *Int J Syst Evol Microbiol*; 64 (2): 384-391.
119. Roland B.C., M.M. Ciarleglio, J.O. Clarke, J.R. Semler, E. Tomakin, G.E. Mullin and P.J. Pasricha (2014). Low ileocecal valve pressure is significantly associated with small intestinal bacterial overgrowth (SIBO). *Dig Dis Sci*; 59 (6): 1269-1277.
120. Rosselló-Mora R. and R. Amann (2001). The species concept for prokaryotes. *FEMS Microbiol Rev*; 25 (1): 39-67.
121. Ruan Z., Y. Wang, C. Zhang, J. Song, Y. Zhai, Y. Zhuang, H. Wang, *et al.* (2014). *Clostridium huakuii* sp. nov., a novel anaerobic acetogenic bacterium isolated from methanogenic consortia. *Int J Syst Evol Microbiol*.
122. Sachdev A.H. and M. Pimentel (2013). Gastrointestinal bacterial overgrowth: pathogenesis and clinical significance. *Ther Adv Chronic Dis*; 4 (5): 223-231.
123. Sallam A. and A. Steinbuchel (2009). *Clostridium sulfidigenes* sp. nov., a mesophilic, proteolytic, thiosulfate- and sulfur-reducing bacterium isolated from pond sediment. *Int J Syst Evol Microbiol*; 59 (7): 1661-1665.
124. Savage D.C. (1977). Microbial ecology of the gastrointestinal tract. *Annu Rev Microbiol*; 31: 107-133.
125. Scanlan P.D. and J.R. Marchesi (2008). Micro-eukaryotic diversity of the human distal gut microbiota: qualitative assessment using culture-dependent and -independent analysis of faeces. *ISME J*; 2 (12): 1183-1193.
126. Schleifer K.H. (2009). Classification of *Bacteria* and *Archaea*: past, present and future. *Syst Appl Microbiol*; 32 (8): 533-542.
127. Schloss P.D. and J. Handelsman (2005). Metagenomics for studying unculturable microorganisms: cutting the Gordian knot. *Genome Biol*; 6 (8): 229.
128. Seki H., M. Shiohara, T. Matsumura, N. Miyagawa, M. Tanaka, A. Komiyama and S. Kurata (2003). Prevention of antibiotic-associated diarrhea in children by *Clostridium butyricum* MIYAIRI. *Pediatr Int*; 45 (1): 86-90.

129. Shcherbakova V.A., N.A. Chuvilskaya, E.M. Rivkina, S.A. Pecheritsyna, K.S. Laurinavichius, N.E. Suzina, G.A. Osipov, *et al.* (2010). *Clostridium algoriphilum* sp. nov. In *List of new names and new combinations previously effectively, but not validly, published*, Validation List no. 134. *Int J Syst Evol Microbiol*; 60 (7): 1477-1479.
130. Shendure J. and H. Ji (2008). Next-generation DNA sequencing. *Nat Biotechnol*; 26 (10): 1135-1145.
131. Shim J.K., S. Johnson, M.H. Samore, D.Z. Bliss and D.N. Gerding (1998). Primary symptomless colonisation by *Clostridium difficile* and decreased risk of subsequent diarrhoea. *Lancet*; 351 (9103): 633-636.
132. Shiratori H., K. Sasaya, H. Ohiwa, H. Ikeno, S. Ayame, N. Kataoka, A. Miya, *et al.* (2009). *Clostridium clariflavum* sp. nov. and *Clostridium caenicola* sp. nov., moderately thermophilic, cellulose-/cellobiose-digesting bacteria isolated from methanogenic sludge. *Int J Syst Evol Microbiol*; 59 (7): 1764-1770.
133. Simmler C., J.G. Napolitano, J.B. McAlpine, S.N. Chen and G.F. Pauli (2014). Universal quantitative NMR analysis of complex natural samples. *Curr Opin Biotechnol*; 25: 51-59.
134. Sleat R., R.A. Mah and R. Robinson (1985). *Acetoanaerobium noterae* gen. nov., sp. nov.: an anaerobic bacterium that forms acetate from H2 and CO2. *Int J Syst Bacteriol*; 35 (1): 10-15.
135. Song L. and X. Dong (2008). *Clostridium amylolyticum* sp. nov., isolated from H2-producing UASB granules. *Int J Syst Evol Microbiol*; 58 (9): 2132-2135.
136. Stackebrandt E. and J. Ebers (2006). Taxonomic parameters revisited: tarnished gold standards. *Microbial Today*; 33: 152-155.
137. Stackebrandt E. and B.M. Goebel (1994). Taxonomic note: a place for DNA-DNA reassociation and 16S rRNA sequence analysis in the present species definition in bacteriology. *Int J Syst Bacteriol*; 44 (4): 846-849.
138. Stackebrandt E., I. Kramer, J. Swiderski and H. Hippe (1999). Phylogenetic basis for a taxonomic dissection of the genus *Clostridium*. *FEMS Immunol Med Microbiol*; 24 (3): 253-258.
139. Staley J.T. and A. Konopka (1985). Measurement of in situ activities of nonphotosynthetic microorganisms in aquatic and terrestrial habitats. *Annu Rev Microbiol*; 39: 321-346.
140. Suetin S.V., V.A. Shcherbakova, N.A. Chuvilskaya, E.M. Rivkina, N.E. Suzina, A.M. Lysenko and D.A. Gilichinsky (2009). *Clostridium tagluense* sp. nov., a psychrotolerant, anaerobic, spore-forming bacterium from permafrost. *Int J Syst Evol Microbiol*; 59 (6): 1421-1426.
141. Sun D.L., X. Jiang, Q.L. Wu and N.Y. Zhou (2013). Intragenomic heterogeneity of 16S rRNA genes causes overestimation of prokaryotic diversity. *Appl Environ Microbiol*; 79 (19): 5962-5969.
142. Swidsinski A., V. Loening-Baucke, H. Lochs and L.P. Hale (2005). Spatial organization of bacterial flora in normal and inflamed intestine: a fluorescence in situ hybridization study in mice. *World J Gastroenterol*; 11 (8): 1131-1140.
143. Taverniti V. and S. Guglielmetti (2011). The immunomodulatory properties of probiotic microorganisms beyond their viability (ghost probiotics: proposal of paraprobiotic concept). *Genes Nutr*; 6 (3): 261-274.
144. Tindall B.J., R. Rosselló-Mora, H.J. Busse, W. Ludwig and P. Kämpfer (2010). Notes on the characterization of prokaryote strains for taxonomic purposes. *Int J Syst Evol Microbiol*; 60 (1): 249-266.
145. Tissier H. (1906). Traitement des infections intestinales par la méthode de la flore bactérienne de l'intestin. *C R Soc Biol*; 60 (359-361).
146. Ueki A., K. Abe, D. Suzuki, N. Kaku, K. Watanabe and K. Ueki (2009). *Anaerosphaera aminiphila* gen. nov., sp. nov., a glutamate-degrading, Gram-positive anaerobic coccus isolated from a methanogenic reactor treating cattle waste. *Int J Syst Evol Microbiol*; 59 (12): 3161-3167.
147. Ursing J.B., R.A. Rosselló-Mora, E. García-Valdés and J. Lalucat (1995). Taxonomic note: a pragmatic approach to the nomenclature of phenotypically similar genomic groups. *Int J Syst Bacteriol*; 45 (3): 604.
148. van Baarlen P., F.J. Troost, S. van Hemert, C. van der Meer, W.M. de Vos, P.J. de Groot, G.J. Hooiveld, *et al.* (2009). Differential NF-kappaB pathways induction by Lactobacillus plantarum in the duodenum of healthy humans correlating with immune tolerance. *Proc Natl Acad Sci USA*; 106 (7): 2371-2376.
149. Van den Abbeele P., T. Van de Wiele, W. Verstraete and S. Possemiers (2011). The host selects mucosal and luminal associations of coevolved gut microorganisms: a novel concept. *FEMS Microbiol Rev*; 35 (4): 681-704.
150. Van den Bogert B., M.M. Leimena, W.M. De Vos, E.G. Zoetendal and M. Kleerebezem (2011). Functional intestinal metagenomics. In: *Handbook of molecular microbial ecology II: metagenomics in different habitats*. pp 170-190. Edited by F.J. De Bruijn. Willey-Blackwell.
151. Viant M.R. (2008). Recent developments in environmental metabolomics. *Mol Biosyst*; 4 (10): 980-986.
152. Vikman P., J. Fadista and N. Oskolkov (2014). RNA sequencing: current and prospective uses in metabolic research. *J Mol Endocrinol*; 53 (2): R93-R101.
153. Virgin H.W. (2014). The virome in mammalian physiology and disease. *Cell*; 157 (1): 142-150.
154. Walter J. and R. Ley (2011). The human gut microbiome: ecology and recent evolutionary changes. *Annu Rev Microbiol*; 65: 411-429.

155. Wang Q., G.M. Garrity, J.M. Tiedje and J.R. Cole (2007). Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Appl Environ Microbiol* ; 73 (16): 5261-5267.
156. Wang T., G. Cai, Y. Qiu, N. Fei, M. Zhang, X. Pang, W. Jia*, et al.* (2012). Structural segregation of gut microbiota between colorectal cancer patients and healthy volunteers. *ISME J*; 6 (2): 320-329.
157. Wayne L.G., D.J. Brenner, R.R. Colwell, P.A.D. Grimont, O. Kandler, M.I. Krichevsky, L.H. Moore*, et al.* (1987). International Committee on Systematic Bacteriology. Report of the ad hoc committee on reconciliation of approaches to bacterial systematics *Int J Syst Bacteriol*; 37 (4): 463-464.
158. Wiegel J., R. Tanner and F.A. Rainey (2006). An introduction to the family *Clostridiaceae*. In: *The Prokaryotes: A Handbook on the Biology of Bacteria, Bacteria: Firmicutes, Cyanobacteria*. 3rd Edn, 4, pp 654-678. Edited by M. Dworkin, S. Falkow, E.I. Rosenberg, K. Schleifer*, et al.* New York: Springer.
159. Woese C.R. and G.E. Fox (1977). Phylogenetic structure of the prokaryotic domain: the primary kingdoms. *Proc Natl Acad Sci USA*; 74 (11): 5088-5090.
160. Woese C.R., O. Kandler and M.L. Wheelis (1990). Towards a natural system of organisms: proposal for the domains *Archaea*, *Bacteria*, and *Eucarya*. *Proc Natl Acad Sci USA*; 87 (12): 4576-4579.
161. Wong J.M., R. de Souza, C.W. Kendall, A. Emam and D.J. Jenkins (2006). Colonic health: fermentation and short chain fatty acids. *J Clin Gastroenterol*; 40 (3): 235-243.
162. Wu Y.F., H. Zheng, Q.L. Wu, H. Yang and S.J. Liu (2014). *Clostridium algifaecis* sp. nov, a novel anaerobic bacterial species from decomposing algal scum. *Int J Syst Evol Microbiol*.
163. Xie G., S. Zhang, X. Zheng and W. Jia (2013). Metabolomics approaches for characterizing metabolic interactions between host and its commensal microbes. *Electrophoresis*; 34 (19): 2787-2798.
164. Yarza P., M. Richter, J. Peplies, J. Euzeby, R. Amann, K.H. Schleifer, W. Ludwig*, et al.* (2008). The All-Species Living Tree project: a 16S rRNA-based phylogenetic tree of all sequenced type strains. *Syst Appl Microbiol*; 31 (4): 241-250.
165. Yarza P., C. Sproer, J. Swiderski, N. Mrotzek, S. Spring, B.J. Tindall, S. Gronow*, et al.* (2013). Sequencing orphan species initiative (SOS): Filling the gaps in the 16S rRNA gene sequence database for all species with validly published names. *Syst Appl Microbiol*; 36 (1): 69-73.
166. Zoetendal E.G., J. Raes, B. van den Bogert, M. Arumugam, C.C. Booijink, F.J. Troost, P. Bork*, et al.* (2012). The human small intestinal microbiota is driven by rapid uptake and conversion of simple carbohydrates. *ISME J*; 6 (7): 1415-1426.
167. Zoetendal E.G., M. Rajilic-Stojanovic and W.M. de Vos (2008). High-throughput diversity and functionality analysis of the gastrointestinal tract microbiota. *Gut*; 57 (11): 1605-1615.
168. Zoetendal E.G., E.E. Vaughan and W.M. de Vos (2006). A microbial world within us. *Mol Microbiol* 59 (6): 1639-1650.
169. Zoetendal E.G., A. von Wright, T. Vilpponen-Salmela, K. Ben-Amor, A.D. Akkermans and W.M. de Vos (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 (7): 3401-3407.

CHAPTER 2

Intestinal microbiota in human health and disease: The impact of probiotics

Jacoline Gerritsen, Hauke Smidt, Ger T. Rijkers and Willem M. de Vos

Genes & Nutrition, 6: 209-240 (2011)

Abstract

The complex communities of microorganisms that colonize the human gastrointestinal tract play an important role in human health. The development of culture-independent molecular techniques has provided new insights in the composition and diversity of the intestinal microbiota. Here, we summarize the present state of the art on the intestinal microbiota with specific attention for the application of high-throughput functional microbiomic approaches to determine the contribution of the intestinal microbiota to human health. Moreover, we review the association between dysbiosis of the microbiota and both intestinal and extra-intestinal diseases. Finally, we discuss the potential of probiotic microorganism to modulate the intestinal microbiota and thereby contribute to health and well-being. The effects of probiotic consumption on the intestinal microbiota are addressed, as well as the development of tailor-made probiotics designed for specific aberrations that are associated with microbial dysbiosis.

Introduction

It is known for over three decades that the human body contains tenfold more microbial cells (10^{14}) than human cells [190]. These microorganisms colonize practically every surface of the human body that is exposed to the external environment, including the skin, oral cavity, respiratory, urogenital and gastrointestinal tract. Of these body sites, the gastrointestinal (GI) tract is by far the most densely colonized organ. The complex community of microorganisms residing in or passing through the GI tract is referred to as the intestinal microbiota.

The intestinal microbiota plays a role in metabolic, nutritional, physiological and immunological processes in the human body. It exerts important metabolic activities by extracting energy from otherwise indigestible dietary polysaccharides such as resistant starch and dietary fibres. These metabolic activities also lead to the production of important nutrients, such as short-chain fatty acids (SCFAs), vitamins (e.g. vitamin K, vitamin B12 and folic acid) and amino acids, which humans are unable to produce themselves [70, 246]. In addition, the intestinal microbiota participates in the defence against pathogens by mechanisms such as colonization resistance and production of antimicrobial compounds. Furthermore, the intestinal microbiota is involved in the development, maturation and maintenance of the GI sensory and motoric functions, the intestinal barrier and the mucosal immune system. These are just a few examples of the functional contributions of the intestinal microbiota to human health, a subject that is regularly reviewed [12, 26, 158, 197, 255].

In recent years a sharp increase is seen in the number of publications addressing the intestinal microbiota. They have provided various lines of evidence supporting a close link between the intestinal microbiota and human health. This review aims to summarize the current knowledge on the composition and diversity of the intestinal microbiota. In addition, it is discussed how new molecular approaches have provided novel insights towards the phylogenetic and functional characterization of the intestinal microbiota. Furthermore, recent insights on the link between the intestinal microbiota and human health are provided. Finally, an overview is presented of ways to modulate the intestinal microbiota with specific attention for the use of probiotics, defined as 'live microorganisms which, when administered in adequate amounts, confer a health benefit on the host' [52].

Main text

Microbial diversity in the GI tract

The GI tract is a complex and dynamic ecosystem containing a diverse collection of microorganisms. These microorganisms are either resident members of the intestinal microbiota or transient passengers introduced from the environment, for example by the regularly influx of microorganisms by the intake of food.

Compositional diversity of the intestinal microbiota

The intestinal microbiota can be described in richness ('who is present') and evenness ('with how many are they present') that together form the ecological term of diversity. If applied at the species-level, richness describes the number of species present in a specific ecosystem, not taking into account their relative abundance. This contrasts with evenness, which represents the relative abundance of each species in a specific ecosystem. These definitions are used to describe the microbial diversity in the GI tract.

Up till recently, conventional culture-based methods were used to assess the intestinal microbial diversity. Over 400 bacterial species have been successfully isolated, cultured and characterized from the human GI tract [172]. However, these culture-based methods have proven to be inadequate in determining the true microbial diversity of the intestinal microbiota since a large fraction of the microbiota remains uncultivated. For a more accurate analysis of the compositional diversity of the intestinal microbiota culture-independent approaches have been developed and it has been revealed that the human intestinal microbiota is an even more complex ecosystem than previously expected. Most of these techniques target the highly conserved 16S ribosomal RNA (rRNA) gene sequences of bacterial and archaeal microorganisms. Molecular techniques that are used to study the diversity of the intestinal microbiota include quantitative polymerase chain reaction (qPCR), temperature or denaturing gradient gel electrophoresis (TGGE or DGGE), terminal-restriction fragment length polymorphism (T-RFLP) and fluorescent *in situ* hybridization (FISH). The latest developments in high-throughput technologies, such as next generation sequencing and phylogenetic micro-arrays, now allow more in-depth analysis of the complete phylogenetic diversity of the intestinal microbiota [225, 255]. Moving beyond the analysis of the variation in the sequence of a single marker gene, it is currently also possible to characterize the complete genetic material obtained from environmental samples such as the GI tract. With the aid of large-scale sequencing approaches these so called metagenomes can be studied and so far several metagenomic inventories of the intestinal microbiota have been reported (Table 1).

Table 1. Overview of metagenomic studies of the human intestinal microbiota.

Nationality of individuals	Number of individuals	Sequencing technology	Total length of sequences obtained (Gb)	Reference
American	2	Sanger	0.2	66
Japanese	13	Sanger	0.727	111
American	18	454 FLX Titanium	2.14	220
European (Danish or Spanish)	124	Solexa (Illumina)	576.7	171 (MetaHIT)
European	20	Sanger	2.6	Genescope
French	49	SoliD	200	INRA

Since the first application of culture-independent methods to determine diversity, it has been shown that the composition of the intestinal microbiota varies substantially between individuals [252]. At least part of this diversity can be attributed to genetic differences between hosts. A positive relation between similarity in dominant faecal microbial communities and genetic relatedness of the hosts has been observed [207, 220, 253]. It is estimated that more than 1000 species-level phylotypes can be found in the GI tract of the total human population [171, 172]. However, the phylogenetic diversity in one individual is much lower, since the intestinal microbiota of each individual only consists of approximately 160 different bacterial species [171]. This estimation is based on metagenomic analysis using the number of non-redundant genes contained by an average-sized genome. Despite the high species richness and inter-individual variability of the intestinal microbiota, a limited number of bacterial phylotypes is more prevalent amongst individuals and might therefore represent a shared phylogenetic core [171, 215]. However, the estimation of the size of the phylogenetic core is dependent on the minimal relative abundance of a given phylotype that can be detected by the molecular approaches deployed. Recent analysis of metagenomic data indicated that there is a high variability in relative abundance (evenness) of core phylotypes between individuals (12- to 2,200-fold difference) [171]. Altogether, these results demonstrate that an accurate estimation of the size of the phylogenetic core is still difficult to make as this is highly dependent on the depth of the analysis.

The vast majority of all microbial cells in the human GI tract are bacteria. At the phylum-level, both culture-dependent and independent studies have demonstrated that the majority of the intestinal bacteria belong to two phyla, the *Bacteroidetes* and the *Firmicutes* [129]. The phylum *Bacteroidetes* consists of three classes, of which the class *Bacteroidetes*, containing the well-known genera *Bacteroides* and *Prevotella*, is probably the most well-studied. The *Firmicutes* is currently the largest bacterial phylum containing more than 200 genera. The majority of the *Firmicutes* detected in the GI tract fall primarily into two main groups, the *Clostridium coccoides* group (also known as *Clostridium* cluster XIVa) and the *Clostridium leptum* group (also referred to as *Clostridium* cluster IV) [35, 129]. Both groups contain members of the genera *Clostridium*, *Eubacterium* and *Ruminococcus* that are taxonomically polyphyletic. In addition to the two phyla *Bacteroidetes* and *Firmicutes*, also members of other phyla, such as *Proteobacteria*, *Actinobacteria*, *Fusobacteria*, *Spirochaetes*, *Verrucomicrobia* and *Lentisphaerae,* have been detected [172, 255].

Although bacteria dominate the GI tract ecosystem, species from the archaeal domain can also be found in the GI tract, with the methanogens *Methanobrevibacter smithii* and *Methanosphaera stadtmanae* being by far the most dominant archaeal groups [66, 141]. While it was previously assumed that these methanogens were only present in a minor fraction of healthy subjects, application of new DNA isolation methods has led to the observation that they are in fact highly prevalent [46, 182]. In addition to bacteria and archaea, eukaryotic

microorganisms can also be members of the intestinal microbiota. Culture-independent analysis of the fungal diversity in the GI tract has demonstrated that the majority of the phylotypes belonged to the two fungal phyla *Ascomycota* (which includes the genera *Candida* and *Saccharomyces*) and *Basidiomycota* [161, 191].

Microbial diversity along the GI tract

Host physiology and intestinal microbiota are intimately connected. This is evident from the fact that each distinct anatomical region along the GI tract is characterized by its own physicochemical conditions and that these changing conditions exert a selective pressure on the microbiota. The physicochemical conditions that influence the composition of the intestinal microbiota include intestinal motility, pH, redox potential, nutrient supplies, host secretions (e.g. hydrochloric acid, digestive enzymes, bile and mucus), and the presence of an intact ileocaecal valve [21]. Thus, the GI tract harbours many distinct niches, each containing a different microbial ecosystem that varies according to the location within the GI tract. This is already demonstrated by the fact that the microbial density increases along the GI tract. Per gram of intestinal content the microbial density increases from 10^1-10^4 microbial cells in the stomach and duodenum, 10^4-10^8 cells in the jejunum and ileum, to 10^{10}-10^{12} cells in the colon and faeces [21, 43]. Despite the fact that it is well-known that the intestinal microbiota is not homogeneously distributed within the GI tract, it is still largely unknown how the diversity varies in the different niches along the GI tract ('who is present where').

By far the most detailed knowledge is available with respect to the microbial composition of faeces. This is mainly because faecal material can be collected non-invasively and contains a large biomass of microbial cells. However, as it is increasingly acknowledged that the composition of microbiota differs significantly in the different niches, more efforts are undertaken to determine the spatio-temporal dynamics of the microbial diversity along the whole GI tract [257]. The large intestine has a rather uniform composition of luminal intestinal microorganisms, and faecal material seems to represent the colonic microbiota composition best [50]. In contrast, there is only limited insight in the composition of the microbiota that resides in the small intestine. Especially the lower part of the small intestine, the ileum, has received minimal attention, mainly due to sampling difficulties caused by the inaccessibility of this region [21]. The composition of the small intestinal microbiota is largely influenced by a combination of gastric acid, bile and pancreatic secretions that enter the GI tract in the duodenum, and which together create a harsh environment for most microorganisms [21]. Hence, compared to other regions, few microorganisms are able to inhabit the upper part of the GI tract. In addition, the antegrade peristaltic movements as part of the migrating motor complex (MMC) ensure a relatively short passage time through the small intestine (3-5 h) by pushing the microbiota towards the large intestine, thus leaving limited time for microorganisms to

replicate and increase in numbers [21]. The short passage time allows transitioning bacteria to retain viability. Furthermore, cellular enzymes such as glutamate decarboxylase and bile resistance systems offer protection against the low pH and bile salts encountered in this upper part of the GI tract, respectively [7, 140].

The small intestine is the part of the GI tract where most of the host enzymatic digestion of the food occurs. The products of these digestive activities are absorbed in more distal parts of the small intestine, the jejunum and especially the ileum. The conditions in the ileum are more favourable for microbial growth compared to the proximal part of the small intestine, as for example the pH is less acidic and bile acids are reabsorbed. Therefore, the number of microorganisms in the ileum can be higher compared to the duodenum [21].

Most of the knowledge about the small intestinal microbiota has been derived from studies with ileal biopsies collected during surgical intervention [2, 14, 238, 240, 245] or from samples obtained from elderly individuals at autopsy [74]. In addition, ileal effluent from ileostomy patients has been used to study the diversity of the luminal microbiota of the human ileum [20, 73]. It was shown that the composition of the microbiota in ileostomy effluent clearly differs from that of the faecal microbiota. Compared to faecal microbiota, ileostomy effluent microbiota is less diverse and less stable, since large fluctuations in ileal microbiota profiles per individual were observed over time [20]. One of the main findings of this study by Booijink and colleagues was that ileostomy effluent showed a higher relative abundance of species within the orders *Lactobacillales* and *Clostridiales*, especially *Veillonella*- and *Streptococcus*-related phylotypes [20]. In addition, species belonging to *Clostridium* cluster I were detected in high levels, in contrast with the reduced levels of species belonging to the *Bacteroidetes* and *Clostridium* clusters III, IV and XIVa. More recently, it was demonstrated that the microbiota composition of ileostomy effluent, which is characterized by an abundance in *Streptococcus* and *Veillonella* species, is more similar to the proximal small intestinal microbiota and clearly differs from that of the ileum [254].

In addition to the variation in microbial composition along the GI tract, the microbiota present in the intestinal lumen also differs significantly from that attached to and imbedded in the intestinal mucus layer. Since mucosa-associated microorganisms live in close contact with host cells, it is likely they execute different functions within the GI ecosystem compared to luminal microorganisms. Several studies have reported a significant difference in dominant microbial community composition between colonic biopsies and faecal samples in humans [50, 114, 257]. It should be kept in mind, however, that in these studies colonic biopsies were obtained from humans undergoing standard colonoscopy, which in general is preceded by a laxative preparation in order to clean the GI tract. The influence of this procedure on the luminal and mucosa-associated microbiota is still largely unknown [124].

Animal models could provide a means to study both the microbial composition along the GI tract as well as the difference in luminal and mucosal-associated microbiota, without the need for physiological alterations during sampling. It has been demonstrated in rodents that intestinal microorganisms are able to survive and even proliferate in the outer loose mucus layer since the glycans present in this layer are accessible as energy source for these microorganisms [84, 97]. In contrast, the inner stratified firmly attached mucus layer probably prevents the intestinal bacteria from coming in contact with the colonic epithelial cells [85]. The organization of the mucus layers varies between the different parts of the GI tract, as it has been observed that the mucus layers in the stomach and the colon are well defined, in contrast to the small intestine where the mucus is less evenly distributed [6, 84]. Most likely, such differences in mucus-layer organization will be associated with variation in the mucosa-associated microbiota along the GI tract. A recent study in mice has shown that the dominant microbiota composition of proximal colonic mucosa-associated and faecal microbiota are very similar to each other, but differ both significantly from distal colonic mucosa-associated samples [241]. In addition, the study demonstrated that the region-specific mucosa-associated microbiota determines the region-specific expression of host genes, in this case of genes encoding Toll-like receptors (TLRs).

Due to the application of culture-independent molecular approaches, our knowledge of the intestinal microbiota has been advanced significantly [256]. Yet, a complete description of the microbial diversity along the human GI tract cannot be given at this moment. Future research should include more samples from the various distinct niches along the GI tract, which nowadays can be collected using minimally invasive methods and which can be deeply analysed using high-throughput technologies

Functional diversity of the intestinal microbiota

Recently, the collective genome of the human intestinal microbiota (the human intestinal microbiome) was estimated to contain 3.3 million microbial genes, which is ~150 times more genes than the human genome [171]. The presence of this wide array of genes in addition to our own genome, suggests that a profound influence of intestinal microorganisms on the human body can be expected. This means that meaningful information related to human health does not only originate from insights in the compositional diversity ('who is present', 'with how many are they present' and 'who is present where'), but can also be derived from knowledge on the function of the microbiota ('what are they doing'). The extent to which the intestinal microbiota is able to expand the metabolic, nutritional, physiological and immunological functions the host is able to perform, is still largely unknown. To address this question, metagenomic studies can provide information on the diversity of the genes encoded by the intestinal microbiota. Recently, it was calculated that almost 40% of the microbial genes present in each human individual were shared with at least half of the human individuals in the studied cohort.

These data provide evidence for the existence of a functional core (core microbiome) [171]. Since functional redundancy within members of the intestinal microbiota exists, there is the possibility that the phylogenetic core does not fully correspond to the functional core [255]. The functional core may contain shared metabolic functions (e.g. degradation of sugar monomers, production of vitamins or butyrate formation) as well as sequential pathways which would, respectively, restrict or expand functional diversity irrespective of phylogenetic diversity.

A main focus of current research is to understand the functional contribution of the human intestinal microbiota to the host. Function-driven metagenomics is a first step in assessing the functional capacity of the intestinal microbiota. A prediction of the functional capacity can originate from the metagenome by comparing the assembled sequences to reference databases, such as the COG (Clusters of Orthologous Groups) and KEGG (Kyoto Encyclopedia of Genes and Genomes) databases. Moreover, function-driven metagenomics can be applied to assign a function to predicted gene products and can even contribute to gene discovery [37, 216]. The first metagenomic studies have demonstrated that, compared to the human genome, the human intestinal microbiome is highly enriched for COG and KEGG categories involved in metabolism [66, 111, 220]. Pathways involved in metabolism of energy, amino acids, nucleotides, carbohydrates, cofactors and vitamins, terpenoids and polyketides, and the biosynthesis of secondary metabolites are highly represented in the human microbiome. These pathways not only allow the microbes to generate energy, to grow and proliferate but also to influence the host. Some of the metabolites are being taken away from the host while other ones are provided (e.g. SCFA, vitamins, gases). Overall, the (metabolic) interaction between microbes and host is beneficial for both parties. Future studies should provide data to further establish and detail the functional contribution of the intestinal microbiota to the metabolic capacity of the host.

Metagenomic studies provide only insight in the genetic potential of the intestinal microbiota and do not demonstrate its true functional contribution to the maintenance of health and well-being [255]. In order to obtain insights in the *in situ* expression of genes encoded by the intestinal microbiome, other functional microbiomic approaches, such as metatranscriptomics, metaproteomics and metabolomics are required. A recent example of a metatranscriptomic approach to study the intestinal microbiota is provided by the study performed by Booijink and colleagues [20]. These authors were able to demonstrate that the gene expression of the human faecal microbiota is subject-specific and enriched for genes involved in (carbohydrate) metabolism. Gosalbes and colleagues also applied a metatranscriptomic approach to study the functionality of the faecal microbiota of healthy volunteers [69]. Remarkably, more rRNA genes were observed than protein-encoding genes. Analysis of the latter showed a uniform functional pattern in carbohydrate metabolism, energy production and synthesis of cellular components as

well as regulatory elements (small RNAs). More specific information has been derived from the metatranscriptomic analysis of bifidobacteria in early life that revealed marked differences between breast-fed and formula-fed infants. Moreover, the specific expression of genes involved in the degradation of human-derived sugars and vitamins such as folic acid biosynthesis testify for the health impacting function of intestinal bifidobacteria [100]. Furthermore, metaproteomics approaches have been applied to investigate faecal samples obtained from human infants [101] and adults [178, 231]. In human adults it was demonstrated that the faecal metaproteome is enriched in proteins belonging to the COG categories involved in translation, energy production and carbohydrate metabolism [231]. Compared to metatranscriptomics and metaproteomics approaches, metabolomic approaches have up to this date been applied more frequently, using NMR spectroscopy and mass spectroscopy in conjunction with computational multivariate analysis [156]. A recent study by Martin and colleagues demonstrates that metabolic profiling can be used for studying nutrient-microbiota relationships by examining the effects of dietary intervention on the presence of faecal metabolites [130]. A variety of systemic diseases such as hypertension [76] and diabetes [47] appear to be directly influenced by microbial metabolism in model animals and human [98]. The metabolic pathways that are involved in drug metabolism are also influenced by the intestinal microbiota in an *in vitro* system [8].

Altogether, functional microbiomic approaches can be applied to examine microbial gene expression and to establish the effects of microbial gene products on the host. However, up to this date it is difficult to connect functionality to the presence of individual microbial species in the human GI tract. In order to link specific sets of genes to the presence of distinct microbial species, complete microbial genome sequences will be needed. Several independent research consortia have taken up the effort to sequence the genomes of hundreds of bacterial strains, which together will form a catalogue of reference genomes from the human microbiota. Recently, the initial sequencing of 178 reference genomes was reported and the first results of comparative genomic analysis of these sequences provided important insight into the inter-strain diversity of bacterial genomes [154]. Large-scale functional microbiomic analyses are needed to fully understand the impact of the human microbiome on the host. This means that a larger number of samples, deeper sequencing, longer sequence reads and more extensive comparative analyses are needed. Integration of all these microbiomic approaches will help to define the functional contribution of each individual microbial phylotype in the human GI tract to the health status of the host.

Changes in composition and diversity of the intestinal microbiota are related to disease

The type and number of microbial species that persist and colonize the GI tract is not determined by chance, but by a combination of factors including but not limited to the

inflammatory state of the host, diet, host genetics, and environmental factors [22, 26, 71, 150]. This means that the host itself influences the composition of the intestinal microbiota. However, the relative impact of these factors on the intestinal microbiota is still largely unknown.

The intestinal microbiota and the host have co-evolved [117]. Human evolution has taken place amidst a world of microorganisms. Symbiotic microorganisms have occupied the niches offered by the gastrointestinal tract and probably adapted to the local circumstances. This in turn may have influenced human evolution in terms of metabolic and nutritional requirements. Ultimately, man depends on its intestinal microbiota for a number of vital functions and thus these intestinal microorganisms may contribute to health. It is, however, difficult to describe the precise impact of the intestinal microbiota on human health and the involvement in human disease.

Perturbation of the microbiota composition, also known as dysbiosis, has been recognized in various diseases, of which many are associated with the GI-tract. However, before dysbiosis can be established the composition of a healthy 'normal' microbiota has to be defined. Yet, the definition of a healthy microbiota is not easy to give. From an operational point of view it could be stated that a healthy intestinal microbiota is the microbiota composition as it can be found in healthy individuals. For practical reasons, the phylogenetic characterization of the microbiota of diseased individuals in comparison with apparently healthy individuals is at this moment the main approach to study changes in composition of the intestinal microbiota in relation to disease. However, since there are substantial inter-individual and intra-individual variations in the composition of the intestinal microbiota, it is difficult to establish the precise relationships between human health and the presence and relative abundance of specific microbial communities. In the future, specific changes in compositional diversity, or even functional diversity, may be applied as biomarkers for health or specific diseases. It must be noted, however, that it is questionable whether changes in phylogenetic composition are really cause or consequence of a given disease.

A role for the intestinal microbiota in the pathogenesis of several diseases and disorders has been suggested. Intensively studied examples for which dysbiosis of the intestinal microbiota has been described, include inflammatory bowel disease (IBD), irritable bowel syndrome (IBS) and obesity, which will be discussed in-depth in this part of the review.

The microbial composition at different stages of life and its relation to health

The intestinal microbiota of healthy adult individuals is relatively stable over time [36, 61, 229, 252]. However, intra-individual fluctuations occur due to environmental changes and pathological events. In addition, substantial changes in the composition of intestinal microbiota occur at both ends of life, in infants and elderly

individuals [217, 223]. Since alterations in the microbial composition are recognized to be of influence on human health, the interest in the development and composition of the microbiota of infant and elderly humans has significantly increased in the last years.

It is widely accepted that microbial colonization of the GI tract starts during and directly after birth when neonates are exposed to bacteria that are derived from the mother and the surrounding environment [1, 123]. Yet, the human fetal environment is not completely microbiologically sterile and there are indications that non-pathological *in utero* exposure of the fetus to intestinal bacteria or bacterial DNA frequently occurs [168, 188]. In addition the isolation of bacteria from the meconium (the first stool of the neonate), umbilical cord blood and amniotic fluid of healthy neonates has been reported [82, 83]. Postnatal colonization of the GI tract is highly variable amongst neonates and is influenced by several factors including mode of delivery, type of infant feeding, gestational age, infant hospitalization and antibiotic use [167]. It is, however, still unclear how each of these factors exactly influences the infant microbial diversity and how this is related to health. A disturbed development of the infant microbiota has been associated with the development of disease later in life [223]. For example, associations have been made between dysbiosis in infants and the later development of childhood obesity [33, 88] and atopic and allergic diseases [18, 89, 166, 200, 239].

Several culture-independent studies have shown that there is a large inter-individual variability amongst infants in the development of the microbiota [54, 164, 167, 176]. In addition, it has been demonstrated that the infant microbiota is highly dynamic and develops in a step-wise fashion with an increase in diversity over time [164, 176]. An important stage in the colonization of the GI tract of infants is the period in which the infants feed on the milk they receive either by breastfeeding or by infant-formula feeding. During this period, the faecal microbiota of infants consists mainly of bifidobacteria [176, 177]. Some bifidobacteria are highly adapted to the digestion of the oligosaccharides present in human milk [251]. The infant intestinal microbiota contains a relatively low diversity in *Bifidobacterium* populations; *B. breve*, *B. bifidum* and *B. longum* subsp. *infantis* are the most common *Bifidobacterium* species [176]. Compared to breast-fed infants, the intestinal microbiota of formula-fed infants is characterized by less diverse *Bifidobacterium* populations [176] and more complex communities of *Clostridia*, *Enterobacteriaceae, Bacteroides* and *Enterococcus* [72, 167]. The introduction of solid food (weaning) marks an increase in microbial diversity and changes in the microbial composition towards an adult microbiota [104]. For example, dominant *Bifidobacterium* populations change; *B. adolescentis*, *B. catenulatum* and *B. longum* subsp. *longum* are more abundantly present in the adult microbiota [134]. The successive shifts of different microbial communities within the first years of life ultimately result in the development of an adult-like microbiota.

In the elderly (usually defined as people over the age of 65), there are major physiological changes that have an impact on the composition and the functionality of

the intestinal microbiota [217, 247]. Many elderly humans suffer from decreased intestinal motility, which can result in prolonged intestinal transit time and faecal retention. Age-related changes, such as decreased senses for smell and taste, dental decay and swallowing difficulties can lead to narrowing of the nutritional intake and even malnutrition. In addition, the age-related gradual deterioration of the immune system (immunosenescence) is associated with changes in intestinal microbiota composition [194]. Furthermore the increased use of laxatives, antibiotics and other medication in elderly individuals will affect intestinal microbiota composition.

Culture-independent studies have demonstrated that the composition of the intestinal microbiota significantly changes with age [13, 129, 148, 258]. Recently, high-throughput methods have been applied to study the changes in the intestinal microbiota of elderly individuals. Biagi and colleagues have used the HITChip, a phylogenetic microarray specifically designed to study the human GI tract microbiota, to compare the intestinal microbiota composition of young adults to that of elderly individuals and centenarians. It was demonstrated that especially the microbiota of centenarians showed significant differences compared to microbiota composition of the other two age groups [15]. The microbiota of centenarians was characterized by low species-level diversity, specific changes in *Firmicutes* subpopulations, enrichment in *Proteobacteria* and a decrease in bifidobacteria. In addition, high-throughput next generation sequencing has been used by Claesson and colleagues to study the composition, variability and temporal stability of the intestinal microbiota of the elderly [29]. They observed that the faecal microbiota of elderly individuals was relatively stable over a three-month period in the majority of the subjects. However, compared to younger control subjects the microbiota of the elderly was characterized by a high inter-individual variation in microbiota composition, also at phylum level. The relative abundance of the *Firmicutes* varied between 8% and 80%, whereas the *Bacteroidetes* levels varied between 14% and 92%. Furthermore, it was found that in the majority of the elderly subjects the microbiota was characterized by a higher *Bacteroidetes/Firmicutes* ratio compared to that observed in younger adults. In addition, distinct differences were seen in *Proteobacteria, Actinobacteria* and *Clostridium* populations between young and older adults. With respect to human health, associations have been found between microbiota composition and frailty in elderly individuals [227]. In frail elderly persons a significant reduction in the number of lactobacilli, *Faecalibacterium prausnitzii* and *Bacteroides/Prevotella* groups was seen. In contrast, the number of *Enterobacteriaceae* was significantly higher. The number of studies that have focused on the age-related differences in intestinal microbiota composition is still limited. In general they suggest that maintenance of (microbial) homeostasis in the GI tract is essential for healthy ageing.

Table 2. Overview of human studies that demonstrate an association between IBS and compositional dysbiosis of the intestinal microbiota determined with culture-independent methods.

Study material	Population	Analytical methods	Reference
Faeces (3 time points)	27 IBS patients 22 healthy individuals	qPCR	126*
Faeces (3 time points)	26 IBS patients 25 healthy individuals	Conventional culturing DGGE Clone library sequencing (16S)	135*
Biopsies: inflamed and non-inflamed tissue (ileum, ascending / sigmoid colon)	20 CD patients 20 UC patients 20 self-limiting colitis patients 20 IBS patients 20 healthy individuals	FISH	211
Faeces (2 time points)	16 IBS patients 16 healthy individuals	DGGE TRAC	136*
Faeces	24 IBS patients 23 healthy individuals	G+C based profiling Clone library sequencing (16S) qPCR	94*
Duodenal biopsies Faeces	41 IBS patients 26 healthy individuals	FISH qPCR	96#
Faeces	10 (+2) IBS (only IBS-D) 23 healthy individuals	G+C based profiling Clone library sequencing (16S) qPCR	110*
Faeces (3 time points)	20 IBS patients 15 healthy individuals	qPCR	122*
Colonic biopsies Faeces	10 IBS patients (only IBS-D) 10 healthy individuals	Conventional culturing qPCR	25
Colonic biopsies Faeces	47 IBS patients 33 healthy individuals	DGGE	30
Duodenal biopsies Faeces	37 IBS patients 20 healthy individuals	DGGE Clone library sequencing (16S) qPCR	95#
Faeces	44 IBS patients	qPCR	125*
Faeces	26 IBS patients 26 healthy individuals	Conventional culturing qPCR HPLC	213

All studies have applied Rome II or III criteria to recruit their subjects and categorize them in IBS subtypes. Studies that have used subjects from the same cohort are indicated by * and #.
DGGE, denaturing gradient gel electrophoresis; FISH, fluorescence *in-situ* hybridization; HPLC, high-performance liquid chromatography; qPCR, quantitative polymerase chain reaction; TRAC, transcript analysis with the aid of affinity capture.

Microbial diversity and IBS

IBS is a functional bowel disorder which is characterized by recurrent abdominal pain or discomfort, irregular bowel movements and disordered stool patterns such as constipation or diarrhoea [120]. The occurrence of these symptoms, however, can vary from person to person. The aetiology of IBS is probably complex and still not well understood. Several

factors are thought to be involved in IBS and may include altered GI motility, visceral hypersensitivity, low-grade inflammation, and psychosocial (anxiety and depression), genetic and dietary factors [27, 93].

Several studies, using both culture-dependent and independent methods, have demonstrated an association between IBS and dysbiosis of the intestinal microbiota (Table 2). In general, faecal material has been used to study dysbiosis in IBS patients. However, more recently also duodenal [95, 96] and colonic [25, 30] biopsies have been used to study the mucosa-associated microbiota. Most studies aimed to show changes in the predominant microbial communities, however, in some cases the focus was more on specific microbial groups. Quantitative differences in microbiota composition, even for specific microbial groups, have been observed in IBD patients compared to healthy individuals. However, the results of the various studies are inconsistent and no consensus has been reached on the association between specific microbial groups and IBS notably as the power of the studies was low and the depth of the analysis limited [181]. At the functional level, some studies have demonstrated altered colonic fermentation patterns and increased gas production in IBS patients [105, 146, 213, 219]. However, these results have not yet been confirmed in other studies at the molecular level.

In most of the studies IBS patients have been classified into different subtypes based on Rome II criteria for IBS: diarrhoea-predominant IBS (IBS-D), constipation-predominant IBS (IBS-C) or a mixed type of IBS with alternating stool patterns (IBS-A). Distinct changes in microbiota composition have been observed in the different IBS subtypes compared to healthy individuals [122, 126, 136]. It appears that the intestinal microbiota of IBS-D patients deviates the most from that of healthy individuals [25, 110]. These data demonstrate the relevance of clinical subtyping of IBS patients when analysing the intestinal microbiota. So far, however, the results from the studies which have applied IBS subtyping have also not shown uniform changes in microbial composition [181].

Microbial diversity and IBD

IBD is a collective name for chronic inflammatory disorders of the GI tract, of which Crohn's disease (CD) and ulcerative colitis (UC) are the most prevalent forms. These are both chronic and relapsing diseases that affect the intestinal mucosa. For both CD and UC the exact aetiology is still not clear, however, it is has been suggested that an aberrant immune response directed against intestinal microbial antigens is involved [71, 187, 203]. CD affects the whole GI tract and is characterized by discontinuous inflammation of the epithelial lining and deep ulcers. UC on the other hand is restricted to the colon and the rectum and is characterized by a continuous mucosal inflammation and superficial ulcers.

Table 3. Overview of human studies that demonstrate an association between IBD and compositional dysbiosis of the intestinal microbiota determined with culture-independent methods.

Study material	Population	Analytical methods	Reference
Biopsies (terminal ileum, colon)	12 CD patients (active disease) 12 UC patients (active disease) 14 non-IBD controls	FISH	103
Biopsies: inflamed and non-inflamed tissue (ileum, ascending / sigmoid colon)	28 self-limiting colitis patients 104 indeterminate colitis patients 119 UC patients 54 CD patients 40 non-IBD controls	Conventional culturing qPCR FISH	210
Faeces	8 CD patients (active disease) 9 CD patients (in remission) 16 healthy individuals	Dot-blot hybridization TGGE	198
Faeces	4 CD patients 4 healthy controls	Clone library sequencing (16S)	127
Colonic biopsies: inflamed tissue	26 CD patients (active disease) 31 UC patients (active disease) 15 inflammatory controls 31 non-inflammatory controls	SSCP Clone library sequencing (16S) qPCR	162
Rectal biopsies: inflamed and non-inflamed tissue	4 CD patients (active disease) 2 CD patients (in remission) 14 UC patients (active disease) 19 UC patients (in remission) 14 non-IBD controls	FISH	151
Biopsies: inflamed and non-inflamed tissue (ileum, ascending / sigmoid colon)	20 CD patients 20 UC patients 20 self-limiting colitis patients 20 IBS patients 20 Non-IBD controls	FISH	211
Biopsies: inflamed and non-inflamed tissue (ileum, ascending / transverse / descending colon, rectum)	20 CD patients (active disease) 15 UC patients (active disease) 14 non-IBD controls	Clone library sequencing (16S) DGGE qPCR	16
Biopsies: inflamed and non-inflamed tissue (ileum, transverse / sigmoid colon, rectum)	6 CD patients 5 UC patients 5 non-IBD controls	Clone library sequencing (16S)	67
Faeces	6 CD patients (in remission) 6 healthy individuals	Clone library screening (metagenome: 16S) FISH / flow cytometry	128*
Biopsies: inflamed and non-inflamed tissue (ileum, ascending / transverse / descending / sigmoid colon, rectum)	19 CD patients 2 UC patients 1 ischemic colitis patient 15 non-IBD controls	DGGE 16S rRNA gene sequence analysis	131
Faeces (several time points)	16 CD patients 18 healthy individuals	DGGE	193

Study material	Population	Analytical methods	Reference
Faeces	13 CD patients (active disease) 13 UC patients (active disease) 5 infectious colitis patients 13 healthy individuals	FISH / flow cytometry	205#
Faeces	29 UC patients (active disease) 12 UC patients (in remission) 46 healthy individuals	T-RFLP	3
Ileal biopsies	13 CD patients (ileum) 8 CD patients (colon) 7 non-IBD controls	Clone library sequencing (16S) qPCR FISH	14
Biopsies (small intestine and colon)	68 CD patients 61 UC patients 61 non-IBD controls	Clone library sequencing (16S) qPCR	60
Biopsies (caecum, colon, rectum)	13 CD patients 19 UC patients 15 healthy individuals	RISA Conventional culturing	108
Faeces	17 CD patients (active disease) 17 CD patients (in remission) 20 healthy controls	T-RFLP	4
Faeces	10 twin pairs with CD 8 healthy twin pairs	G+C profiling Clone library sequencing (16S) T-RFLP	45†
Faeces (several time points)	16 UC patients (in remission) 8 healthy controls	DGGE	132
Colonic biopsies Faeces	15 CD patients (active disease) 8 CD patients (in remission) 44 UC patients (active disease) 29 UC patients (in remission)	Conventional culturing qPCR FISH HPLC	212
Rectal biopsies: inflamed and non-inflamed tissue	9 UC patients (active disease) 11 non-IBD controls	T-RFLP	157
Faeces	22 CD patients (active disease) 10 CD patients (in remission) 12 UC patients (active disease) 4 UC patients (in remission) 8 infectious colitis patients 27 healthy individuals	qPCR	204#
Biopsies (terminal ileum, ascending / transverse / descending colon, rectum)	10 twin pairs with CD 8 healthy twin pairs	T-RFLP Clone library sequencing (16S) qPCR	245†
Faeces	6 CD patients (in remission) 6 healthy individuals	qPCR Phylogenetic microarray (16S)	91*
Faeces	4 CD patients (in remission) 21 UC patients (in remission) 14 healthy individuals	Clone library sequencing (metagenome)	171
Faeces	68 CD patients (in remission) 84 unaffected relatives 55 healthy individuals	DGGE qPCR	86

Study material	Population	Analytical methods	Reference
Faeces	16 CD patients (active disease) 16 healthy individuals	qPCR Phylogenetic microarray (16S)	143
Biopsies: inflamed and non-inflamed tissue (ileum, ascending / transverse / descending / sigmoid colon, rectum)	12 CD patients (active disease) 6 UC patients (active disease) 5 non-IBD controls	qPCR Clone library sequencing (16S)	237

Most of the studies used the Crohn's disease activity index (CDAI; for CD) and/or the clinical activity index (CAI; for UC patients) to assess disease activity in the subjects and to define active disease or remission. Studies that have used subjects from the same cohort are indicated by *, # and †.

DGGE, denaturing gradient gel electrophoresis; FISH, fluorescence *in-situ* hybridization; HPLC, high-performance liquid chromatography ; qPCR, quantitative polymerase chain reaction; RISA, ribosomal intergenic spacer analysis; SSCP, Single strand conformation polymorphism; T-RFLP, terminal-restriction fragment length polymorphism; TGGE, temperature gradient gel electrophoresis.

During the last decade, numerous culture-independent studies have compared the intestinal microbiota composition of IBD patients with that of healthy individuals (Table 3). There is increasing evidence that dysbiosis of the intestinal microbiota has a role in the pathogenesis of IBD. Up to this date, however, the phylum-level changes observed in IBD patients have not always been consistent. In general, an overall decrease in microbial diversity and stability of the intestinal microbiota has been observed in IBD patients [71]. In addition, a decrease in specific members of the *Firmicutes* has been reported in IBD patients, which in some cases coincided with an increase in *Bacteroidetes* and facultative anaerobes such as *Enterobacteriaceae* [71]. Significant differences exist in the microbiota composition of CD patients compared to UC patients [60, 205]. Recently Joossens and colleagues identified a set of five bacterial species that characterized the predominant dysbiosis in CD patients compared to unaffected relatives and healthy individuals [86]. These five species are *Dialister invisus*, an uncharacterised species of *Clostridium* cluster XIVa, *Faecalibacterium prausnitzii*, *Bifidobacterium adolescentis* and *Ruminococcus gnavus*. Of these species, *F. prausnitzii* has been associated with prolongation of remission in CD [202,204] (see also below), while bifidobacteria in general have shown to have beneficial effects on health (see above). Most interestingly, the unaffected relatives of CD patients also have a different composition of their predominant microbiota compared to healthy individuals in general. The impact of these observations on IBD diagnostics and aetiology now has to be addressed.

The role of several different microorganisms in the aetiology of IBD has been studied in more detail. Adherent-invasive *Escherichia coli* [39, 40] and *Mycobacterium avium* subspecies *paratuberculosis* [179] are two prime suspects that have been implicated to be involved in CD pathogenesis. However, a causal relation has not yet been demonstrated. Recently, the presence of two species belonging to the family *Enterobacteriaceae*, *Klebsiella pneumoniae* and *Proteus mirabilis*, was correlated with the development of colitis in a mouse model [65]. However, despite the considerable amount of studies concerning this subject, the evidence that specific microorganisms can induce intestinal inflammation

and cause IBD is inconclusive. In addition to the identification of potential pathogenic bacteria, other bacterial species have been suggested to protect against IBD. For example, it has been shown that the relative abundance of *F. prausnitzii*, a commensal bacterium with anti-inflammatory properties, is significantly decreased in CD patients compared to healthy individuals [202].

High-throughput metagenomic studies can provide more insight in the composition and diversity of the intestinal microbiota of IBD patients. IBD is one of the first diseases that have been the subject of metagenomic investigation [171]. Based on the relative abundance of 155 microbial species (present in at least one individual at a genome coverage of ≥1% in this study population), it was possible to separate patients from healthy individuals, and UC from CD patients [171]. The next step is to compare the IBD subpopulations to healthy individuals at microbial gene-level. On average 25% fewer genes could be detected in the faecal samples of IBD patients compared to individuals not suffering from IBD [171]. These results suggest that the microbiota of IBD patients has a lower functional diversity compared to healthy individuals. The intestinal microbiota in IBD patients produce reduced amounts of SCFA, in particular butyrate, while sulphate reduction (by sulphate-reducing bacteria) is increased [53]. In the near future, metagenomic studies like these will provide more insight in the shifts in functionality which characterizes the differences between IBD patients and healthy individuals.

The observed compositional and functional changes in IBD patients suggest that the intestinal microbiota plays an important role in the aetiology and pathogenesis of IBD. However, up to this date it is still unclear whether dysbiosis is a direct cause for the inflammation in IBD, or merely the result of a disturbed environment in the GI tract. In the latter case, a role for the intestinal microbiota in disease maintenance and severity is possible and will have to be explored in the future.

Microbial diversity and obesity

Obesity is a complex disease characterized by excess body fat accumulation. It has been associated with phylum-level changes in the composition of the intestinal microbiota (Table 4). An increase in the relative abundance of *Firmicutes* and a reduction in the level of *Bacteroidetes* has been observed in both obese mice (*ob*/*ob*) [116] and humans [118]. However, since the original publication, a series of studies have failed to confirm the study of Ley and colleagues and shown variable results with respect to the compositional changes in the microbiota of obese humans [33, 49, 88, 153, 183, 184, 196, 250]. Altogether these data suggest that instead of phylum-level changes, more subtle changes in the composition of the intestinal microbiota are associated with the development of obesity. Recently, Turnbaugh and colleagues have observed a reduced compositional microbial diversity in obese individuals compared to lean individuals [220].

Table 4. Overview of human studies that demonstrate an association between obesity and compositional dysbiosis of the intestinal microbiota determined with culture-independent methods.

Study material	Population	Analytical methods	Key findings	Reference
Faeces (3 time points)	12 obese individuals (on diet) 2 normal-weight individuals	Clone library sequencing (16S)	Obese individuals compared to lean: ↓ *Bacteriodetes* ↑ *Firmicutes*	118
Faeces (3 time points)	19 obese individuals (on diet)	FISH GC	Obese individuals on diet of decreased carbohydrate intake: ↓ *Roseburia* species ↓ *Eubacterium rectale* subgroup of cluster XIVa ↓ bifidobacteria	48*
Faeces (2 time points)	18 obese pregnant women 36 normal-weight pregnant women	FISH / flow cytometry qPCR	Overweighed pregnant women: ↑ *Bacteroides* ↑ *Clostridium* ↑ *Staphylococcus*	33#
Faeces (3 time points)	23 overweight/obese individuals (on diet) 14 non-obese individuals	FISH	During weight-loss diet: ↔ *Bacteriodetes* ↓ butyrate-producing *Firmicutes*	49*
Faeces (2 time points)	25 overweight/obese children 24 normal-weight children (prospective study)	FISH / flow cytometry qPCR	Intestinal microbiota during infancy preceding overweight during childhood: ↓ bifidobacteria ↑ *Staphylococcus aureus*	88
Faeces	20 obese individuals 9 individuals with anorexia nervosa 20 normal-weight individuals	qPCR	Obese individuals: ↓ *Bacteriodetes* ↑ *Lactobacillus* Anorexic individuals: ↑ *Methanobrevibacter smithii*	5
Faeces (2 time points)	39 overweight/obese adolescents (on diet and physical activity)	FISH / flow cytometry	Obese individuals: ↑ *C. histolyticum* ↑ *E. rectale-C. coccoides* Upon calorie restricted diet: ↑ *C. histolyticum* ↑ *C. lituseburense* ↑ *E. rectale-C. coccoides* ↑ *Bacteroides-Prevotella* group	153†
Faeces (2 time points)	36 overweight/obese adolescents (on diet and physical activity)	qPCR	Obese adolescents on diet with a high weight-loss: ↑ Total bacteria ↑ *B. fragilis* group ↑ *C. leptum* group ↑ *B. catenulatum* group ↓ *C. coccoides* group ↓ *Lactobacillus* group	184†

Study material	Population	Analytical methods	Key findings	Reference
Faeces (2 time points)	31 monozygotic twin pairs 23 dizygotic twin pairs 46 mothers of twin pairs	Sanger sequencing (16S) 454 pyrosequencing (metagenome)	Most obesity-associated genes are from: - *Actinobacteria* - *Firmicutes* Most lean-enriched genes are from: - *Bacteroidetes*	220
Faeces	3 obese individuals 3 individuals with a gastric-bypass 3 normal-weight individuals	Clone library sequencing (16S) 454 pyrosequencing (16S) qPCR	Obese individuals: ↑ H(2)-producing *Prevotellaceae* ↑ H(2)-utilizing methanogenic *Archaea*	250
Faeces	15 obese Indian adolescents 13 non-obese Indian adolescents	qPCR	Obese children: ↔ *Bacteroides-Prevotella* ↔ *Bifidobacterium* species ↔ *L. acidophilus* ↔ *E. rectale* ↑ *F. prausnitzii*	11
Faeces (2 time points)	16 infants of overweight women 26 infants of normal-weight women	FISH / flow cytometry qPCR	Infants of overweight mothers: ↑ *Bacteroides* ↑ *Staphylococcus*	34[#]
Faeces	33 obese individuals 35 overweight individuals 30 normal-weight individuals	qPCR GC	Obese individuals compared to lean: ↑ *Bacteriodetes* ↓ *Firmicutes*	196
Faeces	16 overweight pregnant women 34 normal-weight pregnant women	qPCR	Overweight pregnant women: ↓ *Bifidobacterium* species ↓ *Bacteroides* ↑ *Staphylococcus* ↑ *Enterobacteriaceae* ↑ *E. coli*	183

All studies have used the body mass index (BMI) to define normal-weight, overweight and obesity. Studies that have used subjects from the same cohort are indicated by *, [#] and [†].
FISH, fluorescence *in-situ* hybridization; GC; gas chromatography; qPCR, quantitative polymerase chain reaction.

It is evident that (excessive) food intake has an influence on body (over)weight. Recently, a direct link between intestinal microbiota composition and body weight has been suggested. One of the first publications that provides evidence for this link is the publication by Backhed and colleagues for which they colonized germ-free mice with the microbiota of conventionally raised mice [9]. They observed an increase of body fat content of the colonized germ-free mice despite reduced food intake, which was suggested to be caused by the introduction of intestinal microbial communities. In a later study it was demonstrated that the absence of intestinal microorganisms protected germ-free mice against the development of obesity after being fed a high-fat, sugar-

rich diet [10]. Furthermore, it was demonstrated that colonization of germ-free mice with the microbiota of obese mice induced a significant greater increase in body fat weight compared to germ-free mice colonized with the microbiota of lean mice [221]. In addition, these experiments in germ-free mice have demonstrated that the intestinal microbiota is involved in the regulation of fat storage. It was shown that introduction of an intestinal microbiota resulted in an increase in metabolic rate, modulation of *de novo* lipogenesis and an increase in the uptake of monosaccharides from the intestine [9]. Based on these results it has been hypothesized that obese individuals are more efficient in converting food into usable energy and in storing this energy in fat than lean individuals [221]. As discussed above, the intestinal microbiota has a crucial role in the digestion of food, in particular the metabolism of poly- and oligosaccharides and the production of SCFAs that provide the host with additional amounts of energy. Altered representation of bacterial genes and metabolic pathways, including those involved in nutrient harvest, were found to be related to obesity [220]. The results from this study demonstrate that major insights in the differences between various physiological states of the host (in this case obese *vs.* lean) can be obtained by studying the functional microbial diversity in addition to phylogenetic diversity. In line with this conclusion is the observation that the amount of SCFAs produced by the intestinal microbiota rather than the changes in the composition of the microbiota are important in the development of obesity [196].

As for the IBD and IBS, which were discussed above, also for obesity the question remains whether dysbiosis of the intestinal microbiota is a direct cause for obesity or whether it reflects a disturbed host environment. It needs to be established whether the changes in the intestinal microbial communities in obese individuals are not merely an adaptation to a change in the host's diet. Some of the studies that have shown an altered composition of the intestinal microbiota in obese individuals, have also incorporated analysis of the effect of diet change on the observed dysbiosis (Table 4) [48, 49, 118, 153, 184]. Little is known, however, about the influence of dietary change on microbiota composition in humans. A recent study demonstrated rapid and reversible changes in the relative abundance of specific dominant bacterial groups after dietary changes [236]. Most striking was the strong increase in the relative abundance of *Ruminococcus bromii* and *Eubacterium rectale* phylotypes as result of a diet rich in resistant starch. It was suggested that indigestible dietary polysaccharides can substantially change the composition of the intestinal microbiota, however, it is likely that this depends on the initial composition of the intestinal microbiota. Interestingly, *R. bromii* and *E. rectale* were identified as key degraders of starch in an *in vitro* model of the human colon, using 16S rRNA-based stable isotope probing [109]. Recent studies in mice show that the influence of the diet (high-fat *vs.* standard chow or low-fat) on the composition of the intestinal microbiota is independent of genetic disposition for obesity [75, 149].

In addition to obesity, it has also been suggested that the intestinal microbiota is

involved in obesity-associated metabolic disorders, such as type 2 diabetes metabolic endotoxemia, low-grade inflammation and adiposity [24, 115, 235]. In a recent study a high-throughput sequencing approach was used to demonstrate that type 2 diabetes, a metabolic disease primarily caused by obesity-linked insulin resistance, is associated with changes in the composition of the intestinal microbiota [113]. The relative abundance of *Firmicutes* was significantly lower in diabetic patients compared to non-diabetic persons. On the other hand the *Bacteroidetes* and *Proteobacteria* were present in higher abundance. In addition, the *Bacteroidetes*/*Firmicutes* and *Bacteroides-Prevotella*/*C.coccoides-E.rectale* ratios were positively and significantly correlated with plasma glucose levels.

Microbial diversity and other human diseases

In addition to IBD, IBS and obesity, the intestinal microbiota has also been suggested to be involved in several other (chronic) diseases and disorders. Associations have been described between intestinal microbial dysbiosis and intestinal diseases such as coeliac disease, colorectal cancer, pouchitis and necrotizing enterocolitis (NEC) (Table 5). The most recent data show that the intestinal microbiota of coeliac disease patients displays a greater diversity than healthy controls with higher numbers of *Bacteroides/ Prevotella* [41, 195]. *Bifidobacterium, Clostridium histolyticum*, *Clostridium lituseburense* and *Faecalibacterium prausnitzii* were less abundant in coeliac disease patients [41]. Also in the case of colorectal cancer the bacterial diversity and richness was observed to be higher in patients compared to healthy controls [199]. In addition, the intestinal microbiota composition of colorectal cancer patients differs from that of healthy controls, however no consistent pattern has yet been observed.

The mucosal and faecal microbiota of UC pouchitis patients contained more *Clostridium* and *Eubacterium* and fewer *Lactobacillus* and Streptococcus genera compared to the microbiota of healthy pouches from familial adenomatous polyposis (FAP) patients [119, 249]. Luminal samples of UC pouchitis patients contained more *Firmicutes* and *Verrucomicrobia* and fewer *Bacteroidetes* and *Proteobacteria* compared to FAP patients.

The overall microbiota profiles of premature infants with necrotizing enterocolitis (NEC) were not distinguishable from that of control subjects, but 16S rRNA gene sequence analysis detected *Citrobacter*-like sequences and an increased frequency of *Enterococcus*-like sequences [147].

Intestinal microbial dysbiosis has also been observed in extra-intestinal diseases such as atopic and allergic diseases, autism, Type 2 diabetes and rheumatoid arthritis (Table 6). In children who develop an allergic disease later in life, a reduced diversity of faecal microbiota was already observed at 1 week of age [242] (and Niers et al., personal communication). During the first 2 months of life, they were less often colonized with lactobacilli group I (*L. rhamnosus, L. casei, L. paracasei*), *Bifidobacterium adolescentis* and *Clostridium difficile* [89, 200].

Table 5. Overview of human studies that demonstrate an association between intestinal disease and compositional dysbiosis of the intestinal microbiota.

Study material	Population	Analytical methods	Reference
Coeliac disease			
Faeces	26 coeliac patients (no diet, active disease) 23 children without gluten intolerance	Conventional culturing FISH	31
Duodenal biopsies	20 coeliac patients (no diet, active disease) 10 coeliac patients (gluten-free diet, symptom-free) 8 children without gluten intolerance	FISH / flow cytometry	152
Faeces	10 coeliac patients (no diet, active disease) 10 children without gluten intolerance	DGGE	185
Duodenal biopsies Faeces	30 coeliac patients (no diet) 18 coeliac patients (gluten-free diet) 30 children without gluten intolerance	qPCR	32
Faeces	24 coeliac patients (no diet, active disease) 18 coeliac patients (gluten-free diet, symptom-free) 20 children without gluten intolerance	FISH / flow cytometry	41
Duodenal biopsies	20 coeliac patients (active disease / symptom-free) 10 children without gluten intolerance	TGGE	195
Colorectal cancer			
Faeces	18 patients with polyps 32 individuals with high-risk for colon cancer 38 individuals with low-risk for colon cancer	Conventional culturing	145
Faeces	13 patients at high risk for sigmoid colon cancer 14 healthy individuals	Conventional culturing	90
Faeces (3 time points)	20 colon cancer patients 20 polypectomized patients 20 healthy individuals	DGGE	192
Colorectal biopsies	21 individuals with adenomas 23 individuals without adenomas	T-RFLP Clone library sequencing (16S) FISH	199
Pouchitis			
Pouch biopsies Ileostomy effluent Faeces	12 patients with pouchitis 14 patients with indeterminable pouchitis 23 patients without pouchitis 20 ileostomy patients 9 Healthy individuals	Conventional culturing	160
Pouch effluent	5 patients with pouchitis 9 patients without pouchitis	Conventional culturing	180
Pouch effluent	UC patients: 8 patients with healthy pouches 9 patients, no active pouchitis for at least one year 9 patients, no active pouchitis for at least 6 weeks 11 patients with pouchitis, on antibiotic treatment 8 patients with pouchitis FAP patients: 5 patients with healthy pouches	Conventional culturing	159

Study material	Population	Analytical methods	Reference
Pouch effluent	9 patients with pouchitis (UC) 13 patients with healthy pouches (UC)	Conventional culturing	79
Ileum biopsies Pouch biopsies Pouch effluent	5 patients with pouchitis (UC) 15 patients with healthy pouches (UC) 13 healthy individuals	LH-PCR Clone library sequencing (16S)	106
Pouch effluent	5 patients with pouchitis (UC) 15 patients with healthy pouches (UC)	T-RFLP Clone library sequencing (16S)	119
Pouch contents Pouch biopsies	9 patients with pouchitis (UC) 3 patients with healthy pouches (UC) 7 patients with healthy pouches (FAP)	T-RFLP Clone library sequencing (16S)	249
Necrotizing Enterocolitis			
Faeces	10 preterm infants with NEC 10 preterm infants without NEC	T-RFLP Clone library sequencing (16S)	242
Faeces (several time points)	6 preterm infants with NEC or suspected sepsis 6 preterm control infants	DGGE 454 pyrosequencing (16S)	147

The intestinal diseases IBD, IBS and obesity are discussed separately in the article.
DGGE, denaturing gradient gel electrophoresis; FAP, familial anastomosis polyposis; FISH, fluorescence *in-situ* hybridization; LH-PCR, length heterogeneity polymerase chain reaction; qPCR, quantitative polymerase chain reaction; T-RFLP, terminal-restriction fragment length polymorphism; TGGE, temperature gradient gel electrophoresis; UC, ulcerative colitis.

The number of *Clostridium* species found in the stools of children with autism was greater than in the stools of control children, specifically of the *C. histolyticum* group (*Clostridium* clusters I and II) [57, 165]. *Bacteroidetes* was found at high levels in the severely autistic children while populations of the *Bifidobacterium* genus were reduced [56].

Firmicutes and *Clostridia* are reduced in type 2 diabetes [113]. Furthermore, the *Bacteroidetes*/*Firmicutes* ratio as well as the ratio of the *Bacteroides/Prevotella* group to the *C. coccoides-E. rectale* group were observed to be correlated with plasma glucose concentration. In a Chinese population of diabetes patients, reduced populations of bifidobacteria were found [248].

In comparison to patients with fibromyalgia, patients with rheumatoid arthritis had significantly less bifidobacteria and bacteria of the *Bacteroides/Prevotella* group, *Bacteroides fragilis* subgroup, and *Eubacterium rectale/Clostridium coccoides* group [222].

Almost all of the diseases and disorders mentioned above are largely undefined and have a heterogeneous aetiology, which makes it difficult to relate changes in microbiota composition and diversity to disease. Again, also for all these diseases the causality argument of the observed microbiota changes is unresolved. Ultimately, causality and knowledge of the underlying mechanisms will be crucial for a full understanding of the role of the intestinal microbiota in the aetiology of specific diseases.

Table 6. Overview of human studies that demonstrate an association between extra-intestinal disease and compositional dysbiosis of the intestinal microbiota.

Study material	Population	Analytical methods	Reference
Atopic and allergic diseases			
Faeces	27 allergic children 36 non-allergic children	Conventional culturing	17
Faeces (5 time points)	18 infants who developed allergy 26 infants who remained non-allergic	Conventional culturing	18
Faeces (2 time points)	76 infants at high risk for atopic disease	Conventional culturing FISH	89
Faeces (2/3 time points)	27 infants with atopic dermatitis 10 infants without atopic dermatitis	Conventional culturing FISH	99
Faeces	30 children with atopic dermatitis 68 children without atopic dermatitis	Conventional culturing	243
Faeces	957 infants	qPCR	166
Faeces	20 allergic children 20 non-allergic children	DGGE	208
Faeces (3 time points)	10 allergic infants 16 non-allergic infants	qPCR	209
Faeces	37 infants with atopic dermatitis 24 infants without atopic dermatitis	TGGE FISH / flow cytometry	68
Faeces	15 infants who developed atopic dermatitis 20 infants who remained without atopic dermatitis	T-RFLP TGGE	239
Faeces (3 time points)	16 infants who developed allergy 31 infants who remained non-allergic	qPCR	200
Autism			
Faeces Stomach contents Small intestine contents	13 autistic children 8 non-autistic children	Conventional culturing 16S rRNA gene sequencing	57
Faeces	15 autistic children 8 non-autistic children	qPCR	206
Faeces	58 autistic children 12 non-autistic siblings 10 non-autistic children	FISH	165
Faeces	33 autistic children 7 non-autistic siblings 8 non-autistic children	454 pyrosequencing (16S)	65
Diabetes type 2			
Faeces	16 type 2 diabetic patients 12 non-diabetic individuals	DGGE qPCR	248
Faeces	18 type 2 diabetic patients 18 non-diabetic individuals	qPCR 454 pyrosequencing (16S)	113
Rheumatoid Arthritis			
Faeces	51 patients with early rheumatoid arthritis 50 patients with fibromyalgia	FISH / flow cytometry	222

DGGE, denaturing gradient gel electrophoresis; FISH, fluorescence *in-situ* hybridization; qPCR, quantitative polymerase chain reaction; T-RFLP, terminal-restriction fragment length polymorphism; TGGE, temperature gradient gel electrophoresis.

Modulation of the intestinal microbiota

Since it is known that the intestinal microbiota plays an important role in human health and disease, manipulation of these microorganisms by antibiotics, probiotics, prebiotics and synbiotics are attractive approaches to improve and maintain health [64, 170].

Antibiotics are widely used as antimicrobial agents to treat bacterial infections caused by pathogenic microorganisms. In general, however, antibiotics (even narrow-spectrum antibiotics) do not only affect pathogens, but also commensal intestinal microbial communities. This can result in dysbiosis of the intestinal microbiota, subsequently leading to intestinal problems, such as antibiotic-associated diarrhoea (AAD) [137]. The antibiotic-induced disturbances in microbiota composition can be temporary, returning to its original composition within two months, but recently also medium and long-term disturbances in (specific) microbial communities have been described [44, 80, 81, 107]. An additional problem of the widespread antibiotic use, is the increased prevalence of antibiotic resistance resulting from the transfer of antibiotic resistance genes between microorganisms [81].

The intestinal microbiota can be modulated in a more biological manner by the use of probiotics. According to the definition formulated by the World Health Organization (WHO) probiotics are "live microorganisms which, when administered in adequate amounts, confer a health benefit on the host" [52]. Moreover, prebiotics are used to manipulate the microbiota composition in the GI tract. The definition of prebiotics is even more generic than the one of probiotics: "non-digestible food ingredients that, when consumed in sufficient amounts, selectively stimulate the growth and/or activity(ies) of one or a limited number of microbial genus(era)/species in the gut microbiota that confer(s) health benefits to the host" [174]. Mixtures of both probiotics and prebiotics are referred to as synbiotics.

The final part of this review will focus on probiotics as a way to modulate the intestinal microbiota. The opportunities for probiotic intervention in maintaining and restoring health are increasingly being acknowledged, and the field of probiotic research has grown significantly during the past few years [64, 155, 173].

An introduction to probiotics

In line with the very generic definition of probiotics, many different microorganisms have been studied for their potential use as a probiotic, in relation with a wide range of biological or clinical effects. Most of the microorganisms that have been studied are naturally present in the human GI tract. At this moment the most commonly used probiotic microorganisms belong mainly to the bacterial genera *Lactobacillus* and *Bifidobacterium* [19, 102]. These two genera contain a large number of species and strains of which many are being used as probiotic strains. In addition to these genera, the probiotic market contains members from some additional lactic acid bacterial genera,

such as *Streptococcus* and *Enterococcus*, and members from the genera *Bacillus* and *Propionibacterium*. Furthermore, some gram-negative bacteria (e.g. *E. coli* Nissle 1917) and yeast (e.g. *Saccharomyces*) are being used as probiotic microorganisms [64, 77, 78].

Numerous health-beneficial effects have been attributed to probiotic micro-organisms [78, 155]. In general these health benefits can be categorized into three levels of probiotic action [173]. First of all, probiotic microorganisms can act directly within the GI tract (level 1), for example by direct interaction with the intestinal microbiota or by enzymatic activities. Secondly, they can interact directly with the intestinal mucus layer and epithelium (level 2), thereby influencing the intestinal barrier function and the mucosal immune system. Thirdly, probiotics can have effects outside the GI tract (level 3), for example on the systemic immune system and other organs, such as the liver and the brain. Although *in vivo* data are emerging, most of the mechanistic studies with probiotic microorganisms have been performed *in vitro*, *ex vivo* or with the aid of animal models. The *in vitro* activity of a given probiotic does not necessarily correlate with the efficacy of the intended clinical *in vivo*. In addition, it is important to note that each probiotic strain has its own specific properties. The health benefits that can be attributed to one probiotic strain cannot be extrapolated to other probiotic strains or mixtures of strains. Even closely related microbial strains of the same species may have different physiological effects [121, 138, 139, 201, 226, 232, 233].

Nowadays, multistrain or multispecies probiotic mixtures, which contain more than one probiotic strain, are becoming increasingly popular [28, 218]. Compared to single strain probiotics, probiotic mixtures have the possible advantage that the properties of single strains may have additive or even synergistic effects when put in a mixture together with other probiotics strains, which can result in higher efficacy. Another potential advantage of probiotic mixtures may be that compared to a single strain probiotic a wider range of health benefits could be accomplished. In contrast, mixing of probiotic strains might also result in reduced efficacy since individual strains may have opposite effects or even inhibit each other. There are, however, a limited number of *in vivo* studies available that compare the effects of a probiotic mixture to those of the individual strains [28]. This means that the evidence for the hypothesis that probiotic mixtures are more effective than the individual strains is still limited.

The influence of probiotics on composition and diversity of the intestinal microbiota

Modulation of the intestinal microbiota (part of level 1 probiotic action) is one of the potential health-beneficial effects of probiotics. The mechanisms by which probiotic microorganisms are able to modify the intestinal microbiota include reduction of luminal pH, competition for nutrients, secretion of antimicrobial compounds (bacteriocins), prevention of bacterial adhesion and evasion of epithelial cells, and induction of the production of antimicrobial compounds (defensins) by the host [59, 155]. By these mechanisms

probiotics can not only potentially modulate the intestinal microbiota composition, but also prevent pathogenic bacterial overgrowth.

Up to this date, many studies have been performed that examine the effects of probiotics on the composition and diversity of the intestinal microbiota, both in diseased and healthy individuals. For a given disease, the desired outcome of probiotic intervention is the modulation of the intestinal microbiota in such a way that a healthy microbiota composition is achieved. However, also other parameters have been addressed such as stabilization of the microbiota as in the case of IBS and a multispecies probiotic, determined with the use of a phylogenetic microarray [87]. The interpretation of the effects of probiotics on the intestinal microbiota composition in healthy individuals are, however, more difficult to interpret (Table 7). Those studies do provide information on the effects of probiotics on the intestinal microbiota without a potential bias caused by disease effects. However, this does not imply that in a diseased situation these probiotic products will have the same influence on the intestinal microbiota.

Until recently, in most of the probiotic studies conventional culture-based methods have been used to study the influence of probiotics on the intestinal microbiota. However, since a few years culture-independent methods are now also being applied in probiotic research (Table 7). In general, demonstrating the colonization of the supplemented probiotic microorganism(s) has been the primary aim of most studies in healthy individuals. In most cases a transient colonization of the probiotic microorganism(s) has been observed. It is still questionable, however, whether probiotic strains would need to colonize in order to be effective or whether transient presence would also suffice to exert health-beneficial effects.

The probiotic studies performed in humans have almost exclusively examined the effect of probiotic administration on the composition of the faecal microbiota, whereas other niches of the GI tract have hardly been studied thus far (Table 7). As already indicated, even major local changes in microbiota composition in specific niches of the GI tract might not be reflected in the faeces. This means that there is still a major gap in knowledge on the influence of probiotic microorganisms on the intestinal microbiota. In addition, the influence of probiotic microorganisms on mucosa-associated intestinal microbiota is also not well studied. However, these interactions are possibly of key importance in relation to disease pathogenesis, since mucosa-associated microorganisms are in more close contact with the intestinal barrier and immune system. One of the few examples of a study on the *in vivo* effects of probiotics on the human host is a recent study by Van Baarlen and colleagues [224]. The authors examined the influence of a probiotic microorganism on human duodenal mucosal gene expression and they showed that changes in gene expression patterns, especially in the NF-κβ dependent pathways, induced by *Lactobacillus plantarum* WCFS1 could be linked to the establishment of immunotolerance in human adults.

Table 7. Details of studies performed to examine the effects of probiotic intervention on intestinal microbiota composition of healthy subjects determined with culture-independent methods.

Population	Study groups (based on treatment)	Study material	Analytical methods	Key findings	Reference
Animals					
Healthy rats (n = 30)	Probiotic: *B. lactis* BI and *S. thermophilus* Prebiotic: FOS Placebo: only carrier material	Caecum (tissue and contents)	Conventional culturing DGGE	Both prebiotic and probiotic group: ↓ *Clostridia* ↓ *Bacteroides* ↓ total anaerobes Prebiotic-treated group: ↓ coliforms ↑ *Bifidobacterium* Probiotic-treated group: ↑ diversity ↑ coliforms	144
Healthy mice (n = 16)	Probiotic: *L. casei* Probiotic: *L. plantarum* Probiotic: mixture of *L. casei* and *L. plantarum* Control: no treatment	Faeces Intestinal tissue	DGGE T-RFLP Clone-library sequencing	Mixture-treated group: - No significant effect on dominant microbiota composition - Shifts in the diversity of *Lactobacillus* species	62
Healthy fish (red tilapia) (n = 12)	Probiotic: diet containing *Pediococcus acidilactici* Placebo: normal diet	Intestinal contents	Conventional culturing DGGE	Probiotic-treated group: ↓ species richness and diversity - transiently colonization by *P. acidilactici*	55
Humans					
Healthy adults (n = 10)	Probiotic: milk powder containing *L. rhamnosus* DR20	Faeces	Conventional culturing FISH DGGE	Probiotic-treated group: - No significant effect on dominant microbiota composition	214
Healthy adults (n = 30)	Probiotic: *B. animalis* subsp. *lactis* Bb-12 Prebiotic: GOS Synbiotic: GOS and *B. animalis* subsp. *lactis* Bb-12	Faeces	DGGE	All groups: - No qualitative changes in faecal *Bifidobacterium* communities Probiotic/synbiotic-treated groups: - transiently colonization by *B. animalis* subsp. *lactis* Bb-12	189
Healthy children (n = 26)	Probiotic: Yoghurt containing *L. delbrueckii* subsp. *bulgaricus, S. thermophilus* and *L. paracasei* A Placebo: pasteurized yoghurt	Faeces	Conventional culturing RAPD-PCR DGGE	Probiotic-treated group: - No significant effect on dominant microbiota composition - GI survival and transiently colonization by *L. paracasei*	133
Preterm infants (n = 69)	Probiotic: *B. animalis* subsp. *lactis* Bb-12 Placebo: only carrier material	Faeces	Conventional culturing FISH	Probiotic-treated group: ↑ *Bifidobacterium*	142

Population	Study groups (based on treatment)	Study material	Analytical methods	Key findings	Reference
Healthy adults (n = 12)	Probiotic: yoghurt containing *L. delbrueckii* subsp. b*ulgaricus, S. thermophilus* and *B. animalis* DN-173 010 Probiotic: *B. animalis* DN-173 010 (lyophilised)	Faeces	Conventional culturing Colony immuno-blotting DGGE FISH	Both probiotic-treated groups: - No significant effect on dominant microbiota composition - GI survival and transiently colonization by *B. animalis*	175
Healthy adults (n = 30)	Prebiotic: lactulose Probiotic: *S. boulardii* Synbiotic: lactulose and *S. boulardii* Placebo: maltodextrin	Faeces	DGGE Group-specific qPCR	Prebiotic-treated group: ↑ *B. adolescentis* ↑ *Bifidobacterium* Probiotic/synbiotic-treated group: - no changes	228
Healthy elderly (n = 55)	Probiotic: fermented oat drink containing *B. longum* 46 and *B. longum* 2C Probiotic: fermented oat drink containing *B. animalis* subsp. *lactis* Bb-12 Placebo: only fermented oat drink	Faeces	Species-specific qPCR	Probiotic-treated group (*B. longum*): ↑ *B. adolescentis* ↑ *B. catenulatum* Probiotic-treated group (*B. animalis)* ↑ *B. animalis*	163
Healthy adults (n = 14)	Probiotic: encapsulated *L. rhamnosus* R11 and *L. acidophilus* R52	Faeces	Conventional culturing qPCR	No significant effect on dominant microbiota composition GI survival and transiently colonization by *L. rhamnosus*	58
Healthy adults (n = 79)	Probiotic: yoghurt containing *L. delbrueckii* subsp. *bulgaricus* and *S. thermophilus* Placebo: pasteurized yoghurt Control: no yoghurt	Faeces	DGGE qPCR	Probiotic-treated group: ↑ lactic acid bacteria ↑ *C. perfringens* Yoghurt-receiving groups: ↓ *Bacteroides* group	63
Healthy adults on antibiotic treatment (n = 40)	Probiotic: *B. animalis* subsp. *lactis* BI-04, *B. animalis* subsp. *lactis* Bi-07, *L. acidophilus* NCFM, *L. paracasei* Lpc-37, *B. bifidum* Bb-02 and maltodextran Placebo: only maltodextran	Faeces	Conventional culturing T-RFLP	Probiotic-treated group: A more rapid return to pre-antibiotic microbiota composition ↑ *Enterobacteriaceae* ↑ *Bifidobacterium*	51
Healthy elderly (n = 66)	Probiotic: fermented oat drink containing *B. longum* 46 and *B. longum* 2C Placebo: non-fermented oat drink	Faeces	Conventional culturing species-specific qPCR	Probiotic-treated group: Significant change in *Bifidobacterium* communities (↑ *B. catenulatum*, *B. bifidum* and *B. breve*)	112

DGGE, denaturing gradient gel electrophoresis; FISH, fluorescence *in-situ* hybridization; qPCR, quantitative polymerase chain reaction; RAPD-PCR, random amplification of polymorphic DNA polymerase chain reaction; T-RFLP, terminal-restriction fragment length polymorphism.

In contrast with most human probiotic studies, animal studies have focused on the spatial influence of probiotics on the intestinal microbiota (Table 7). However, to which extent these results reflect the human situation has to be determined. Administration of a given probiotic strain will result in the (temporarily) increase of that strain the GI tract, but may also change the overall composition of the intestinal microbiota. Indeed, the results of relevant experiments performed thus far demonstrate that probiotic-induced changes in microbiota composition are not restricted to the administered species. Which probiotic microorganisms are able to influence the relative abundance of which specific intestinal microorganisms are questions that are currently under study. It should be realized, however, that a change in composition or diversity of the intestinal microbiota by probiotic intervention is not a health benefit by itself.

As discussed previously, dysbiosis of the intestinal microbiota has been associated with a growing number of (intestinal) diseases. Since modulation of the composition of intestinal microbiota by probiotics was demonstrated to be possible, probiotic intervention has the potential to counterbalance intestinal dysbiosis and thus restore health. The effectiveness of probiotic intervention has been studied in a number of human diseases, including IBD (CD, UC and pouchitis), IBS, constipation, diarrhoea (including AAD), colon cancer, cardiovascular disease, NEC, allergic diseases, obesity and metabolic disorders and these have been the subject of systematic reviews as well as Cochrane reviews [64, 78, 169, 186, 244]. With the possible exception of NEC and pouchitis, variable clinical effects are found. One, and probably the most important reason for the variable clinical effects is the variation in probiotic species and strains which are being used. On top of that, there is a lack of standardized methods for the study of the intestinal microbiota (e.g. sample collection, sample storage and analysis methods), which makes it almost impossible to directly compare findings from different groups. Apart from the large variety of probiotic species and strains, also different dosages of probiotic microorganisms are used, or combinations of probiotic species and strains, or prebiotic supplements can be added. In addition, the populations of interest can be relatively heterogeneous since health and disease are not always well defined. At the same time, host-dependent factors (e.g. host genotype) may have an influence on the effectiveness of intestinal microbiota modulation by probiotics. Finally, most clinical studies have included only a small number of patients and used short-term intervention periods. All of this, in combination with the fact that the intestinal microbiota composition is diverse and maybe even unique for each individual, makes it problematic to observe general changes in microbiota composition as result of probiotic intervention.

In the early days of probiotic research it was thought that decreased intestinal microbial diversity could be a direct cause of gastrointestinal disease. In such a concept, probiotic intervention should be aimed at increasing this diversity, which would be sufficient to resolve the clinical problem. For some diseases such as IBD, there is indeed evidence for

a decreased diversity [45, 157, 171]. By contrast, a recent culture-independent study shows a higher richness and diversity of bacteria in the faeces of autistic individuals compared to healthy controls [56]. Nowadays, it is recognized that the interaction between intestinal microbiota and the host is more complex than just a high or low microbial diversity. Thus, no general statements can be made on the role of microbial diversity in health and disease, since different microbe-host interactions are involved in the pathophysiology of different diseases. Knowledge of the molecular and physiological mechanisms behind specific diseases and aberrations that are associated with microbial dysbiosis will contribute to the development of tailor-made probiotics designed for specific interventions.

Application of high-throughput molecular approaches in probiotic research

New insights in the potential effect of probiotic intervention on the intestinal microbiota can be obtained by application of high-throughput molecular approaches in probiotic research. An example is provided by a study in which the effectiveness of daily probiotic supplementation of *Lactobacillus rhamnosus* GG (LGG) on preventing the development of early markers of asthma in a human clinical study was examined [23]. The probiotic bacterium LGG is one of the most widely used probiotic microorganisms and has been used in a large number of clinical trials. An explanation for its probiotic properties has recently been provided by its genomic characterization revealing the presence of mucus binding pili in LGG that are assumed to interact with the host [92]. A phylogenetic microarray analysis was used to study the effect of LGG abundance on the bacterial community structure of stool samples of six-month-old infants [38]. Since the researchers were blinded to the treatment of the infants (probiotic or placebo), the effect of LGG administration on LGG abundance and intestinal microbiota composition could not be examined. However, cluster analysis of the microarray data demonstrated that LGG abundance was associated with a distinct community composition. Communities with high relative abundance of LGG showed an increased relative abundance of a large number of bacterial taxa and the majority of these taxa were phylogenetically clustered. In addition there was a significant difference in evenness of the intestinal microbiota between samples containing a low or high abundance of LGG. It was hypothesized that a possible mechanism of the probiotic action of LGG is the stimulation of a stable, even and functionally redundant microbiota and facilitating the colonization by other beneficial microorganisms [92]. Whether the ability of pili of LGG to bind to intestinal mucosal surfaces is important in this respect remains to be determined.

Recently, high-throughput metagenomic sequencing was used to relate the effect of probiotic intervention to microbiota composition [230]. It was demonstrated that consumption of a fermented milk product supplemented with *Bifidobacterium animalis* subsp. *lactis* (*B*FMP) induced some specific metabolic shifts in an ulcerative colitis mouse model. In addition, it was shown that the immune status of the mice had an effect on the

shifts in the composition of the intestinal microbiota. Moreover, subsets of mice could be identified based on microbiota composition that clustered together corresponding to effectiveness of the *B*FMP treatment. These results reinforce the notion that the composition of the endogenous intestinal microbiota plays an important role in the host response to the probiotic intervention, thereby influencing the effectiveness of probiotic intervention. In this study it was also observed that *B*FMP consumption resulted in a metabolic shift; a decreased caecal pH and alteration in short chain fatty acid profiles. It must be noted, however, that these beneficial effects cannot be directly linked to the activity of this specific *Bifidobacterium* strain since a non-fermented milk product was used as control product. Overall, the data support the hypothesis that probiotics are not only able to influence the composition, but also the metabolic activity of the intestinal microbiota [42]. Both of these effects need to be studied separately to get a complete picture of the influence of probiotic intervention on the intestinal microbiota. This is also emphasized in a recent study in which the effect of a specific synbiotic product on the intestinal microbiota was examined [234]. This study showed no influence on the composition of dominant faecal microbiota, but significant changes in faecal metabolic profiles were observed. These results suggest that synbiotic intervention is able to affect the metabolic activity of the intestinal microbiota while maintaining microbiota composition with respect to its predominant components.

Conclusions

Knowledge on the composition and diversity of a healthy microbiota and on how changes in the intestinal microbiota lead to or are associated with disease, is far from complete. More research is needed to examine the species and strain diversity in the GI tract, the diversity of microbial genes (microbiome) in the GI tract and the activity of these genes (mRNA, protein and metabolite production). For future probiotic research it is important to determine the level of compositional and functional microbial dysbiosis in relevant target populations and identify potential members of the healthy microbiota to counteract the dysbiosis. Understanding the molecular mechanisms of action attributed to commensal and pathogenic bacteria will contribute to better designed probiotic products. In the future this knowledge can be applied in the development of tailor-made probiotics designed for clearly characterized target populations.

Acknowledgements

This work was partly funded by Agentschap NL, project FND-07013.

References

1. Adlerberth I. and A.E. Wold (2009). Establishment of the gut microbiota in Western infants. *Acta Paediatr*; 98 (2): 229-238.
2. Ahmed S., G.T. Macfarlane, A. Fite, A.J. McBain, P. Gilbert and S. Macfarlane (2007). Mucosa-associated bacterial diversity in relation to human terminal ileum and colonic biopsy samples. *Appl Environ Microbiol*; 73 (22): 7435-7442.
3. Andoh A., S. Sakata, Y. Koizumi, K. Mitsuyama, Y. Fujiyama and Y. Benno (2007). Terminal restriction fragment length polymorphism analysis of the diversity of fecal microbiota in patients with ulcerative colitis. *Inflamm Bowel Diseases*; 13 (8): 955-962.
4. Andoh A., T. Tsujikawa, M. Sasaki, K. Mitsuyama, Y. Suzuki, T. Matsui, T. Matsumoto*, et al.* (2008). Fecal microbiota profile of Crohn's disease determined by terminal restriction fragment length polymorphism (t-rflp) analysis. *Aliment Pharmacol Ther*; 29 (75-82).
5. Armougom F., M. Henry, B. Vialettes, D. Raccah and D. Raoult (2009). Monitoring bacterial community of human gut microbiota reveals an increase in *Lactobacillus* in obese patients and Methanogens in anorexic patients. *PLoS One*; 4 (9): e7125.
6. Atuma C., V. Strugala, A. Allen and L. Holm (2001). The adherent gastrointestinal mucus gel layer: thickness and physical state in vivo. *Am J Physiol Gastroint Liver Physiol*; 280 (5): G922-929.
7. Audia J.P., C.C. Webb and J.W. Foster (2001). Breaking through the acid barrier: an orchestrated response to proton stress by enteric bacteria. *Int J Med Microbiol*; 291 (2): 97-106.
8. Aura A.M., I. Mattila, T. Hyotylainen, P. Gopalacharyulu, C. Bounsaythip, M. Oresic and K.M. Oksman-Caldentey (2011). Drug metabolome of the Simvastatin formed by human intestinal microbiota in vitro. *Mol Biosyst*; 7 (2): 437-446.
9. Backhed F., H. Ding, T. Wang, L.V. Hooper, G.Y. Koh, A. Nagy, C.F. Semenkovich*, et al.* (2004). The gut microbiota as an environmental factor that regulates fat storage. *Proc Natl Acad Sci USA*; 101 (44): 15718-15723.
10. Backhed F., J.K. Manchester, C.F. Semenkovich and J.I. Gordon (2007). Mechanisms underlying the resistance to diet-induced obesity in germ-free mice. *Proc Natl Acad Sci USA*; 104 (3): 979-984.
11. Balamurugan R., G. George, J. Kabeerdoss, J. Hepsiba, A.M. Chandragunasekaran and B.S. Ramakrishna (2010). Quantitative differences in intestinal *Faecalibacterium prausnitzii* in obese Indian children. *Br J Nutr*; 103 (3): 335-338.
12. Barbara G., V. Stanghellini, G. Brandi, C. Cremon, G. Di Nardo, R. De Giorgio and R. Corinaldesi (2005). Interactions between commensal bacteria and gut sensorimotor function in health and disease. *Am J Gastroenterol*; 100 (11): 2560-2568.
13. Bartosch S., A. Fite, G.T. Macfarlane and M.E. McMurdo (2004). Characterization of bacterial communities in feces from healthy elderly volunteers and hospitalized elderly patients by using real-time PCR and effects of antibiotic treatment on the fecal microbiota. *Appl Environ Microbiol*; 70 (6): 3575-3581.
14. Baumgart M., B. Dogan, M. Rishniw, G. Weitzman, B. Bosworth, R. Yantiss, R.H. Orsi*, et al.* (2007). Culture independent analysis of ileal mucosa reveals a selective increase in invasive *Escherichia coli* of novel phylogeny relative to depletion of *Clostridiales* in Crohn's disease involving the ileum. *ISME J*; 1 (5): 403-418.
15. Biagi E., L. Nylund, M. Candela, R. Ostan, L. Bucci, E. Pini, J. Nikkila*, et al.* (2010). Through ageing, and beyond: gut microbiota and inflammatory status in seniors and centenarians. *PLoS One*; 5 (5): e10667.
16. Bibiloni R., M. Mangold, K.L. Madsen, R.N. Fedorak and G.W. Tannock (2006). The bacteriology of biopsies differs between newly diagnosed, untreated, Crohn's disease and ulcerative colitis patients. *J Med Microbiol*; 55 (Pt 8): 1141-1149.
17. Bjorksten B., P. Naaber, E. Sepp and M. Mikelsaar (1999). The intestinal microflora in allergic Estonian and Swedish 2-year-old children. *Clin Exp Allergy*; 29 (3): 342-346.
18. Bjorksten B., E. Sepp, K. Julge, T. Voor and M. Mikelsaar (2001). Allergy development and the intestinal microflora during the first year of life. *J Allergy Clin Immunol*; 108 (4): 516-520.
19. Boesten R.J. and W.M. de Vos (2008). Interactomics in the human intestine: Lactobacilli and Bifidobacteria make a difference. *J Clin Gastroenterol*; 42 Suppl 3 Pt 2: S163-167.
20. Booijink C.C., S. El-Aidy, M. Rajilic-Stojanovic, H.G. Heilig, F.J. Troost, H. Smidt, *et al.* (2010). High temporal and inter-individual variation detected in the human ileal microbiota. *Environ Microbiol*; 12 (12): 3213-3227.
21. Booijink C.C., E.G. Zoetendal, M. Kleerebezem and W.M. de Vos (2007). Microbial communities in the human small intestine: coupling diversity to metagenomics. *Future Microbiol*; 2: 285-295.
22. Buddington R.K. and P.T. Sangild (2011). Development of the mammalian gastrointestinal tract, the resident

microbiota, and the role of diet in early life. *J Anim Sci*; 89: 1506-1519.
23. Cabana M.D., M. McKean, A.R. Wong, C. Chao and A.B. Caughey (2007). Examining the hygiene hypothesis: the Trial of Infant Probiotic Supplementation. *Paediatr Perinat Epidemiol*; 21 Suppl 3: 23-28.
24. Cani P.D. and N.M. Delzenne (2009). The role of the gut microbiota in energy metabolism and metabolic disease. *Curr Pharm Des*; 15 (13): 1546-1558.
25. Carroll I.M., Y.H. Chang, J. Park, R.B. Sartor and Y. Ringel (2010). Luminal and mucosal-associated intestinal microbiota in patients with diarrhea-predominant irritable bowel syndrome. *Gut Pathog*; 2 (1): 19.
26. Cerf-Bensussan N. and V. Gaboriau-Routhiau (2010). The immune system and the gut microbiota: friends or foes? *Nat Rev Immunol*; 10 (10): 735-744.
27. Chang J.Y. and N.J. Talley (2011). An update on irritable bowel syndrome: from diagnosis to emerging therapies. *Curr Opin Gastroenterol*; 27 (1): 72-78.
28. Chapman C.M., G.R. Gibson and I. Rowland (2011). Health benefits of probiotics: are mixtures more effective than single strains? *Eur J Nutr*; 50 (1): 1-17.
29. Claesson M.J., S. Cusack, O. O'Sullivan, R. Greene-Diniz, H. de Weerd, E. Flannery, J.R. Marchesi, *et al.* (2010). Composition, variability, and temporal stability of the intestinal microbiota of the elderly. *Proc Natl Acad Sci USA*; 108 (Suppl 1):4586-4591.
30. Codling C., L. O'Mahony, F. Shanahan, E.M. Quigley and J.R. Marchesi (2010). A molecular analysis of fecal and mucosal bacterial communities in irritable bowel syndrome. *Dig Dis Sci*; 55 (2): 392-397.
31. Collado M.C., M. Calabuig and Y. Sanz (2007). Differences between the fecal microbiota of coeliac infants and healthy controls. *Curr Issues Intest Microbiol*; 8 (1): 9-14.
32. Collado M.C., E. Donat, C. Ribes-Koninckx, M. Calabuig and Y. Sanz (2008). Imbalances in faecal and duodenal *Bifidobacterium* species composition in active and non-active coeliac disease. *BMC Microbiol*; 8: 232.
33. Collado M.C., E. Isolauri, K. Laitinen and S. Salminen (2008). Distinct composition of gut microbiota during pregnancy in overweight and normal-weight women. *Am J Clin Nutr*; 88 (4): 894-899.
34. Collado M.C., E. Isolauri, K. Laitinen and S. Salminen (2010). Effect of mother's weight on infant's microbiota acquisition, composition, and activity during early infancy: a prospective follow-up study initiated in early pregnancy. *Am J Clin Nutr*; 92 (5): 1023-1030.
35. Collins M.D., P.A. Lawson, A. Willems, J.J. Cordoba, J. Fernandez-Garayzabal, P. Garcia, J. Cai, *et al.* (1994). The phylogeny of the genus *Clostridium*: proposal of five new genera and eleven new species combinations. *Int J Syst Bacteriol*; 44 (4): 812-826.
36. Costello E.K., C.L. Lauber, M. Hamady, N. Fierer, J.I. Gordon and R. Knight (2009). Bacterial community variation in human body habitats across space and time. *Science*; 326 (5960): 1694-1697.
37. Cowan D., Q. Meyer, W. Stafford, S. Muyanga, R. Cameron and P. Wittwer (2005). Metagenomic gene discovery: past, present and future. *Trends Biotechnol*; 23 (6): 321-329.
38. Cox M.J., Y.J. Huang, K.E. Fujimura, J.T. Liu, M. McKean, H.A. Boushey, M.R. Segal, *et al.* (2010). *Lactobacillus casei* abundance is associated with profound shifts in the infant gut microbiome. *PLoS One*; 5 (1): e8745.
39. Darfeuille-Michaud A. (2002). Adherent-invasive *Escherichia coli*: a putative new *E. coli* pathotype associated with Crohn's disease.*Int J Med Microbiol*; 292 (3-4): 185-193.
40. Darfeuille-Michaud A., J. Boudeau, P. Bulois, C. Neut, A.L. Glasser, N. Barnich, M.A. Bringer, *et al.* (2004). High prevalence of adherent-invasive *Escherichia coli* associated with ileal mucosa in Crohn's disease. *Gastroenterology*; 127 (2): 412-421.
41. De Palma G., I. Nadal, M. Medina, E. Donat, C. Ribes-Koninckx, M. Calabuig and Y. Sanz (2010). Intestinal dysbiosis and reduced immunoglobulin-coated bacteria associated with coeliac disease in children. *BMC Microbiol*; 10: 63.
42. De Preter V., H.M. Hamer, K. Windey and K. Verbeke (2010). The impact of pre- and/or probiotics on human colonic metabolism: Does it affect human health? *Mol Nutr Food Res*; 55 (1): 46-57.
43. Dethlefsen L., P.B. Eckburg, E.M. Bik and D.A. Relman (2006). Assembly of the human intestinal microbiota. *Trends Ecol Evol*; 21 (9): 517-523.
44. Dethlefsen L., S. Huse, M.L. Sogin and D.A. Relman (2008). The pervasive effects of an antibiotic on the human gut microbiota, as revealed by deep 16S rRNA sequencing. *PLoS Biol*; 6 (11): e280.
45. Dicksved J., J. Halfvarson, M. Rosenquist, G. Jarnerot, C. Tysk, J. Apajalahti, L. Engstrand, *et al.* (2008). Molecular analysis of the gut microbiota of identical twins with Crohn's disease. *ISME J*; 2 (7): 716-727.
46. Dridi B., M. Henry, A. El Khechine, D. Raoult and M. Drancourt (2009). High prevalence of *Methanobrevibacter smithii* and *Methanosphaera stadtmanae* detected in the human gut using an improved DNA detection protocol. *PLoS One*; 4 (9): e7063.
47. Dumas M.E., S.P. Wilder, M.T. Bihoreau, R.H. Barton, J.F. Fearnside, K. Argoud, L. D'Amato, *et al.* (2007). Direct

quantitative trait locus mapping of mammalian metabolic phenotypes in diabetic and normoglycemic rat models. *Nat Genet*; 39 (5): 666-672.

48. Duncan S.H., A. Belenguer, G. Holtrop, A.M. Johnstone, H.J. Flint and G.E. Lobley (2007). Reduced dietary intake of carbohydrates by obese subjects results in decreased concentrations of butyrate and butyrate-producing bacteria in feces. *Appl Environ Microbiol*; 73 (4): 1073-1078.
49. Duncan S.H., G.E. Lobley, G. Holtrop, J. Ince, A.M. Johnstone, P. Louis and H.J. Flint (2008). Human colonic microbiota associated with diet, obesity and weight loss. *Int J Obes (Lond)*; 32 (11): 1720-1724.
50. Eckburg P.B., E.M. Bik, C.N. Bernstein, E. Purdom, L. Dethlefsen, M. Sargent, S.R. Gill, *et al.* (2005). Diversity of the human intestinal microbial flora. *Science*; 308 (5728): 1635-1638.
51. Engelbrektson A., J.R. Korzenik, A. Pittler, M.E. Sanders, T.R. Klaenhammer, G. Leyer and C.L. Kitts (2009). Probiotics to minimize the disruption of faecal microbiota in healthy subjects undergoing antibiotic therapy. *J Med Microbiol*; 58 (Pt 5): 663-670.
52. FAO/WHO (2002). Working Group Report on Drafting Guidelines for the Evaluation of Probiotics in Food. *London, Ontario, Canada*.
53. Fava F. and S. Danese (2011). Intestinal microbiota in inflammatory bowel disease: Friend of foe? *World J Gastroenterol*; 17 (5): 557-566.
54. Favier C.F., E.E. Vaughan, W.M. De Vos and A.D. Akkermans (2002). Molecular monitoring of succession of bacterial communities in human neonates. *Appl Environ Microbiol*; 68 (1): 219-226.
55. Ferguson R.M., D.L. Merrifield, G.M. Harper, M.D. Rawling, S. Mustafa, S. Picchietti, J.L. Balcazar, *et al.* (2010). The effect of *Pediococcus acidilactici* on the gut microbiota and immune status of on-growing red tilapia (*Oreochromis niloticus*). *J Appl Microbiol*; 109 (3): 851-862.
56. Finegold S.M., S.E. Dowd, V. Gontcharova, C. Liu, K.E. Henley, R.D. Wolcott, E. Youn, *et al.* (2010). Pyrosequencing study of fecal microflora of autistic and control children. *Anaerobe*; 16 (4): 444-453.
57. Finegold S.M., D. Molitoris, Y. Song, C. Liu, M.L. Vaisanen, E. Bolte, M. McTeague, *et al.* (2002). Gastrointestinal microflora studies in late-onset autism. *Clin Infect Dis*; 35 (Suppl 1): S6-S16.
58. Firmesse O., A. Mogenet, J.L. Bresson, G. Corthier and J.P. Furet (2008). *Lactobacillus rhamnosus* R11 consumed in a food supplement survived human digestive transit without modifying microbiota equilibrium as assessed by real-time polymerase chain reaction. *J Mol Microbiol Biotechnol*; 14 (1-3): 90-99.
59. Fooks L.J. and G.R. Gibson (2002). Probiotics as modulators of the gut flora. *Br J Nutr*; 88 Suppl 1: S39-49.
60. Frank D.N., A.L. St Amand, R.A. Feldman, E.C. Boedeker, N. Harpaz and N.R. Pace (2007). Molecular-phylogenetic characterization of microbial community imbalances in human inflammatory bowel diseases. *Proc Natl Acad Sci USA*; 104 (34): 13780-13785.
61. Franks A.H., H.J. Harmsen, G.C. Raangs, G.J. Jansen, F. Schut and G.W. Welling (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 (9): 3336-3345.
62. Fuentes S., M. Egert, M. Jimenez-Valera, A. Ramos-Cormenzana, A. Ruiz-Bravo, H. Smidt and M. Monteoliva-Sanchez (2008). Administration of *Lactobacillus casei* and *Lactobacillus plantarum* affects the diversity of murine intestinal lactobacilli, but not the overall bacterial community structure. *Res Microbiol*; 159 (4): 237-243.
63. Garcia-Albiach R., M.J. Pozuelo de Felipe, S. Angulo, M.I. Morosini, D. Bravo, F. Baquero and R. del Campo (2008). Molecular analysis of yogurt containing *Lactobacillus delbrueckii* subsp. *bulgaricus* and *Streptococcus thermophilus* in human intestinal microbiota. *Am J Clin Nutr*; 87 (1): 91-96.
64. Gareau M.G., P.M. Sherman and W.A. Walker (2010). Probiotics and the gut microbiota in intestinal health and disease. *Nat Rev Gastroenterol Hepatol*; 7 (9): 503-514.
65. Garrett W.S., C.A. Gallini, T. Yatsunenko, M. Michaud, A. DuBois, M.L. Delaney, S. Punit, *et al.* (2010). *Enterobacteriaceae* act in concert with the gut microbiota to induce spontaneous and maternally transmitted colitis. *Cell Host Microbe*; 8 (3): 292-300.
66. Gill S.R., M. Pop, R.T. Deboy, P.B. Eckburg, P.J. Turnbaugh, B.S. Samuel, J.I. Gordon, *et al.* (2006). Metagenomic analysis of the human distal gut microbiome. *Science*; 312 (5778): 1355-1359.
67. Gophna U., K. Sommerfeld, S. Gophna, W.F. Doolittle and S.J. Veldhuyzen van Zanten (2006). Differences between tissue-associated intestinal microfloras of patients with Crohn's disease and ulcerative colitis. *J Clin Microbiol*; 44 (11): 4136-4141.
68. Gore C., K. Munro, C. Lay, R. Bibiloni, J. Morris, A. Woodcock, A. Custovic, *et al.* (2008). *Bifidobacterium pseudocatenulatum* is associated with atopic eczema: a nested case-control study investigating the fecal microbiota of infants. *J Allergy Clin Immunol*; 121 (1): 135-140.
69. Gosalbes M.J., A. Durban, M. Pignatelli, J.J. Abellan, N. Jimenez-Hernandez, A.E. Perez-Cobas, A. Latorre, *et al.* (2011).

Metatranscriptomic approach to analyze the functional human gut microbiota. *PLoS One*; 6 (3): e17447.
70. Hamer H.M., D. Jonkers, K. Venema, S. Vanhoutvin, F.J. Troost and R.J. Brummer (2008). Review article: the role of butyrate on colonic function. *Aliment Pharmacol Ther*; 27 (2): 104-119.
71. Hansen J., A. Gulati and R.B. Sartor (2010). The role of mucosal immunity and host genetics in defining intestinal commensal bacteria. *Curr Opin Gastroenterol*; 26 (6): 564-571.
72. Harmsen H.J., A.C. Wildeboer-Veloo, G.C. Raangs, A.A. Wagendorp, N. Klijn, J.G. Bindels and G.W. Welling (2000). Analysis of intestinal flora development in breast-fed and formula-fed infants by using molecular identification and detection methods. *J Pediatr Gastroenterol Nutr*; 30 (1): 61-67.
73. Hartman A.L., D.M. Lough, D.K. Barupal, O. Fiehn, T. Fishbein, M. Zasloff and J.A. Eisen (2009). Human gut microbiome adopts an alternative state following small bowel transplantation. *Proc Natl Acad Sci USA*; 106 (40): 17187-17192.
74. Hayashi H., R. Takahashi, T. Nishi, M. Sakamoto and Y. Benno (2005). Molecular analysis of jejunal, ileal, caecal and recto-sigmoidal human colonic microbiota using 16S rRNA gene libraries and terminal restriction fragment length polymorphism. *J Med Microbiol*; 54 (Pt 11): 1093-1101.
75. Hildebrandt M.A., C. Hoffmann, S.A. Sherrill-Mix, S.A. Keilbaugh, M. Hamady, Y.Y. Chen, R. Knight, *et al.* (2009). High-fat diet determines the composition of the murine gut microbiome independently of obesity. *Gastroenterology*; 137 (5): 1716-1724 e1711-1712.
76. Holmes E., R.L. Loo, J. Stamler, M. Bictash, I.K. Yap, Q. Chan, T. Ebbels, *et al.* (2008). Human metabolic phenotype diversity and its association with diet and blood pressure. *Nature*; 453 (7193): 396-400.
77. Holzapfel W.H., P. Haberer, J. Snel, U. Schillinger and J.H. Huis in't Veld (1998). Overview of gut flora and probiotics. *Int J Food Microbiol*; 41 (2): 85-101.
78. Iannitti T. and B. Palmieri (2010). Therapeutical use of probiotic formulations in clinical practice. *Clin Nutr*; 29 (6): 701-725.
79. Iwaya A., T. Iiai, H. Okamoto, Y. Ajioka, T. Yamamoto, T. Asahara, K. Nomoto, *et al.* (2006). Change in the bacterial flora of pouchitis. *Hepato-gastroenterology*; 53 (67): 55-59.
80. Jernberg C., S. Lofmark, C. Edlund and J.K. Jansson (2007). Long-term ecological impacts of antibiotic administration on the human intestinal microbiota. *ISME J*; 1 (1): 56-66.
81. Jernberg C., S. Lofmark, C. Edlund and J.K. Jansson (2010). Long-term impacts of antibiotic exposure on the human intestinal microbiota. *Microbiology*; 156 (Pt 11): 3216-3223.
82. Jimenez E., L. Fernandez, M.L. Marin, R. Martin, J.M. Odriozola, C. Nueno-Palop, *et al.* (2005). Isolation of commensal bacteria from umbilical cord blood of healthy neonates born by cesarean section. *Curr Microbiol*; 51 (4): 270-274.
83. Jimenez E., M.L. Marin, R. Martin, J.M. Odriozola, M. Olivares, J. Xaus, L. Fernandez, *et al.* (2008). Is meconium from healthy newborns actually sterile? *Res Microbiol*; 159 (3): 187-193.
84. Johansson M.E., J.M. Holmen Larsson and G.C. Hansson (2010). The two mucus layers of colon are organized by the MUC2 mucin, whereas the outer layer is a legislator of host-microbial interactions. *Proc Natl Acad Sci USA*; 108 (Suppl 1): 4659-4665.
85. Johansson M.E., M. Phillipson, J. Petersson, A. Velcich, L. Holm and G.C. Hansson (2008). The inner of the two Muc2 mucin-dependent mucus layers in colon is devoid of bacteria. *Proc Natl Acad Sci USA*; 105 (39): 15064-15069.
86. Joossens M., G. Huys, M. Cnockaert, V. De Preter, K. Verbeke, P. Rutgeerts, P. Vandamme, *et al.* (2011). Dysbiosis of the faecal microbiota in patients with Crohn's disease and their unaffected relatives. *Gut*; 60 (5): 631-637.
87. Kajander K., E. Myllyluoma, M. Rajilic-Stojanovic, S. Kyronpalo, M. Rasmussen, S. Jarvenpaa, E.G. Zoetendal, *et al.* (2008). Clinical trial: multispecies probiotic supplementation alleviates the symptoms of irritable bowel syndrome and stabilizes intestinal microbiota. *Aliment Pharmacol Ther*; 27 (1): 48-57.
88. Kalliomaki M., M.C. Collado, S. Salminen and E. Isolauri (2008). Early differences in fecal microbiota composition in children may predict overweight. *Am J Clin Nutr*; 87 (3): 534-538.
89. Kalliomaki M., P. Kirjavainen, E. Eerola, P. Kero, S. Salminen and E. Isolauri (2001). Distinct patterns of neonatal gut microflora in infants in whom atopy was and was not developing. *J Allergy Clin Immunol*; 107 (1): 129-134.
90. Kanazawa K., F. Konishi, T. Mitsuoka, A. Terada, K. Itoh, S. Narushima, M. Kumemura, *et al.* (1996). Factors influencing the development of sigmoid colon cancer. Bacteriologic and biochemical studies. *Cancer*; 77 (8 Suppl): 1701-1706.
91. Kang S., S.E. Denman, M. Morrison, Z. Yu, J. Dore, M. Leclerc and C.S. McSweeney (2010). Dysbiosis of fecal microbiota in Crohn's disease patients as revealed by a custom phylogenetic microarray. *Inflamm Bowel Dis*; 16 (12): 2034-2042.
92. Kankainen M., L. Paulin, S. Tynkkynen, I. von Ossowski, J. Reunanen, P. Partanen, R. Satokari, *et al.* (2009). Comparative genomic analysis of *Lactobacillus rhamnosus* GG reveals pili containing a human- mucus binding protein. *Proc Natl Acad Sci USA*; 106 (40): 17193-17198.
93. Karantanos T., T. Markoutsaki, M. Gazouli, N.P. Anagnou and D.G. Karamanolis (2010). Current insights in to the pathophysiology of Irritable Bowel Syndrome. *Gut Pathog*; 2 (1): 3.

94. Kassinen A., L. Krogius-Kurikka, H. Makivuokko, T. Rinttila, L. Paulin, J. Corander, E. Malinen, *et al.* (2007). The fecal microbiota of irritable bowel syndrome patients differs significantly from that of healthy subjects. *Gastroenterology*; 133 (1): 24-33.
95. Kerckhoffs A.P., K. Ben-Amor, M. Samsom, M.E. van der Rest, J. de Vogel, J. Knol and L.M. Akkermans (2010). Molecular analysis of faecal and duodenal samples reveals significantly higher prevalence and numbers of Pseudomonas aeruginosa in IBS. *J Med Microbiol*; 60: 236-245.
96. Kerckhoffs A.P., M. Samsom, M.E. van der Rest, J. de Vogel, J. Knol, K. Ben-Amor and L.M. Akkermans (2009). Lower Bifidobacteria counts in both duodenal mucosa-associated and fecal microbiota in irritable bowel syndrome patients. *World J Gastroenterol*; 15 (23): 2887-2892.
97. Kim Y.S. and S.B. Ho (2010). Intestinal goblet cells and mucins in health and disease: recent insights and progress. *Curr Gastroenterol Rep*; 12 (5): 319-330.
98. Kinross J.M., A.W. Darzi and J.K. Nicholson (2011). Gut microbiome-host interactions in health and disease. *Genome Med*; 3 (3): 14.
99. Kirjavainen P.V., E. Apostolou, T. Arvola, S.J. Salminen, G.R. Gibson and E. Isolauri (2001). Characterizing the composition of intestinal microflora as a prospective treatment target in infant allergic disease. *FEMS Immunol Med Microbiol*; 32 (1): 1-7.
100. Klaassens E.S., R.J. Boesten, M. Haarman, J. Knol, F.H. Schuren, E.E. Vaughan and W.M. de Vos (2009). Mixed-species genomic microarray analysis of fecal samples reveals differential transcriptional responses of bifidobacteria in breast- and formula-fed infants. *Appl Environ Microbiol*; 75 (9): 2668-2676.
101. Klaassens E.S., W.M. de Vos and E.E. Vaughan (2007). Metaproteomics approach to study the functionality of the microbiota in the human infant gastrointestinal tract. *Appl Environ Microbiol*; 73 (4): 1388-1392.
102. Kleerebezem M. and E.E. Vaughan (2009). Probiotic and gut lactobacilli and bifidobacteria: molecular approaches to study diversity and activity. *Annu Rev Microbiol*; 63: 269-290.
103. Kleessen B., A.J. Kroesen, H.J. Buhr and M. Blaut (2002). Mucosal and invading bacteria in patients with inflammatory bowel disease compared with controls. *Scand J Gastroenterol*; 37 (9): 1034-1041.
104. Koenig J.E., A. Spor, N. Scalfone, A.D. Fricker, J. Stombaugh, R. Knight, L.T. Angenent, *et al.* (2010). Succession of microbial consortia in the developing infant gut microbiome. *Proc Natl Acad Sci USA*; 108 (Suppl 1): 4578-4585.
105. Koide A., T. Yamaguchi, T. Odaka, H. Koyama, T. Tsuyuguchi, H. Kitahara, M. Ohto, *et al.* (2000). Quantitative analysis of bowel gas using plain abdominal radiograph in patients with irritable bowel syndrome. *Am J Gastroenterol*; 95 (7): 1735-1741.
106. Komanduri S., P.M. Gillevet, M. Sikaroodi, E. Mutlu and A. Keshavarzian (2007). Dysbiosis in pouchitis: evidence of unique microfloral patterns in pouch inflammation. *Clin Gastroenterol Hepatol*; 5 (3): 352-360.
107. Koning C.J., D. Jonkers, H. Smidt, F. Rombouts, H.J. Pennings, E. Wouters, E. Stobberingh, *et al.* (2008). The effect of a multispecies probiotic on the composition of the faecal microbiota and bowel habits in chronic obstructive pulmonary disease patients treated with antibiotics. *Br J Nutr*; 103 (10): 1452-1460.
108. Kotlowski R., C.N. Bernstein, S. Sepehri and D.O. Krause (2007). High prevalence of *Escherichia coli* belonging to the B2+D phylogenetic group in inflammatory bowel disease. *Gut*; 56 (5): 669-675.
109. Kovatcheva-Datchary P., M. Egert, A. Maathuis, M. Rajilic-Stojanovic, A.A. de Graaf, H. Smidt, W.M. de Vos, *et al.* (2009). Linking phylogenetic identities of bacteria to starch fermentation in an in vitro model of the large intestine by RNA-based stable isotope probing. *Environ Microbiol*; 11 (4): 914-926.
110. Krogius-Kurikka L., A. Lyra, E. Malinen, J. Aarnikunnas, J. Tuimala, L. Paulin, H. Makivuokko, *et al.* (2009). Microbial community analysis reveals high level phylogenetic alterations in the overall gastrointestinal microbiota of diarrhoea-predominant irritable bowel syndrome sufferers. *BMC Gastroenterol*; 9: 95.
111. Kurokawa K., T. Itoh, T. Kuwahara, K. Oshima, H. Toh, A. Toyoda, H. Takami, *et al.* (2007). Comparative metagenomics revealed commonly enriched gene sets in human gut microbiomes. *DNA Res*; 14 (4): 169-181.
112. Lahtinen S.J., L. Tammela, J. Korpela, R. Parhiala, H. Ahokoski, H. Mykkanen and S.J. Salminen (2009). Probiotics modulate the *Bifidobacterium* microbiota of elderly nursing home residents. *Age*; 31 (1): 59-66.
113. Larsen N., F.K. Vogensen, F.W. van den Berg, D.S. Nielsen, A.S. Andreasen, B.K. Pedersen, W.A. Al-Soud, *et al.* (2010). Gut microbiota in human adults with type 2 diabetes differs from non-diabetic adults. *PLoS One*; 5 (2): e9085.
114. Lepage P., P. Seksik, M. Sutren, M.F. de la Cochetiere, R. Jian, P. Marteau and J. Dore (2005). Biodiversity of the mucosa-associated microbiota is stable along the distal digestive tract in healthy individuals and patients with IBD. *Inflamm Bowel Dis*; 11 (5): 473-480.
115. Ley R.E. (2010). Obesity and the human microbiome. *Curr Opin Gastroenterol*; 26 (1): 5-11.
116. Ley R.E., F. Backhed, P. Turnbaugh, C.A. Lozupone, R.D. Knight and J.I. Gordon (2005). Obesity alters gut microbial ecology. *Proc Natl Acad Sci USA*; 102 (31): 11070-11075.

117. Ley R.E., C.A. Lozupone, M. Hamady, R. Knight and J.I. Gordon (2008). Worlds within worlds: evolution of the vertebrate gut microbiota. *Nat Rev Microbiol*; 6 (10): 776-788.
118. Ley R.E., P.J. Turnbaugh, S. Klein and J.I. Gordon (2006). Microbial ecology: human gut microbes associated with obesity. *Nature*; 444 (7122): 1022-1023.
119. Lim M., J.D. Adams, M. Wilcox, P. Finan, P. Sagar and D. Burke (2009). An assessment of bacterial dysbiosis in pouchitis using terminal restriction fragment length polymorphisms of 16S ribosomal DNA from pouch effluent microbiota. *Dis Colon Rectum*; 52 (8): 1492-1500.
120. Longstreth G.F., W.G. Thompson, W.D. Chey, L.A. Houghton, F. Mearin and R.C. Spiller (2006). Functional bowel disorders. *Gastroenterology*; 130 (5): 1480-1491.
121. Lopez P., M. Gueimonde, A. Margolles and A. Suarez (2010). Distinct *Bifidobacterium* strains drive different immune responses in vitro. *Int J Food Microbiol*; 138 (1-2): 157-165.
122. Lyra A., T. Rinttila, J. Nikkila, L. Krogius-Kurikka, K. Kajander, E. Malinen, J. Matto, *et al.* (2009). Diarrhoea-predominant irritable bowel syndrome distinguishable by 16S rRNA gene phylotype quantification. *World J Gastroenterol*; 15 (47): 5936-5945.
123. Mackie R.I., A. Sghir and H.R. Gaskins (1999). Developmental microbial ecology of the neonatal gastrointestinal tract. *Am J Clin Nutr*; 69 (5): 1035S-1045S.
124. Mai V., B. Greenwald, J.G. Morris, Jr., J.P. Raufman and O.C. Stine (2006). Effect of bowel preparation and colonoscopy on post-procedure intestinal microbiota composition. *Gut*; 55 (12): 1822-1823.
125. Malinen E., L. Krogius-Kurikka, A. Lyra, J. Nikkila, A. Jaaskelainen, T. Rinttila, T. Vilpponen-Salmela, *et al.* (2010). Association of symptoms with gastrointestinal microbiota in irritable bowel syndrome. *World J Gastroenterol*; 16 (36): 4532-4540.
126. Malinen E., T. Rinttila, K. Kajander, J. Matto, A. Kassinen, L. Krogius, M. Saarela, *et al.* (2005). Analysis of the fecal microbiota of irritable bowel syndrome patients and healthy controls with real-time PCR. *Am J Gastroenterol*; 100 (2): 373-382.
127. Mangin I., R. Bonnet, P. Seksik, L. Rigottier-Gois, M. Sutren, Y. Bouhnik, C. Neut, *et al.* (2004). Molecular inventory of faecal microflora in patients with Crohn's disease. *FEMS Microbiol Ecol*; 50 (1): 25-36.
128. Manichanh C., L. Rigottier-Gois, E. Bonnaud, K. Gloux, E. Pelletier, L. Frangeul, R. Nalin, *et al.* (2006). Reduced diversity of faecal microbiota in Crohn's disease revealed by a metagenomic approach. *Gut*; 55 (2): 205-211.
129. Mariat D., O. Firmesse, F. Levenez, V. Guimaraes, H. Sokol, J. Dore, G. Corthier, *et al.* (2009). The *Firmicutes*/*Bacteroidetes* ratio of the human microbiota changes with age. *BMC Microbiol*; 9: 123.
130. Martin F.P., N. Sprenger, I. Montoliu, S. Rezzi, S. Kochhar and J.K. Nicholson (2010). Dietary modulation of gut functional ecology studied by fecal metabonomics. *J Proteome Res*; 9 (10): 5284-5295.
131. Martinez-Medina M., X. Aldeguer, F. Gonzalez-Huix, D. Acero and L.J. Garcia-Gil (2006). Abnormal microbiota composition in the ileocolonic mucosa of Crohn's disease patients as revealed by polymerase chain reaction-denaturing gradient gel electrophoresis. *Inflamm Bowel Dis*; 12 (12): 1136-1145.
132. Martinez C., M. Antolin, J. Santos, A. Torrejon, F. Casellas, N. Borruel, F. Guarner, *et al.* (2008). Unstable composition of the fecal microbiota in ulcerative colitis during clinical remission. *Am J Gastroenterol*; 103 (3): 643-648.
133. Marzotto M., C. Maffeis, T. Paternoster, R. Ferrario, L. Rizzotti, M. Pellegrino, F. Dellaglio, *et al.* (2006). *Lactobacillus paracasei* A survives gastrointestinal passage and affects the fecal microbiota of healthy infants. *Res Microbiol*; 157 (9): 857-866.
134. Matsuki T., K. Watanabe, J. Fujimoto, Y. Kado, T. Takada, K. Matsumoto and R. Tanaka (2004). Quantitative PCR with 16S rRNA-gene-targeted species-specific primers for analysis of human intestinal bifidobacteria. *Appl Environ Microbiol*; 70 (1): 167-173.
135. Matto J., L. Maunuksela, K. Kajander, A. Palva, R. Korpela, A. Kassinen and M. Saarela (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 (2): 213-222.
136. Maukonen J., R. Satokari, J. Matto, H. Soderlund, T. Mattila-Sandholm and M. Saarela (2006). Prevalence and temporal stability of selected clostridial groups in irritable bowel syndrome in relation to predominant faecal bacteria. *J Med Microbiol*; 55 (Pt 5): 625-633.
137. McFarland L.V. (1998). Epidemiology, risk factors and treatments for antibiotic-associated diarrhea. *Dig Dis*; 16 (5): 292-307.
138. Medina M., E. Izquierdo, S. Ennahar and Y. Sanz (2007). Differential immunomodulatory properties of *Bifidobacterium longum* strains: relevance to probiotic selection and clinical applications. *Clin Exp Immunol*; 150 (3): 531-538.
139. Meijerink M., S. van Hemert, N. Taverne, M. Wels, P. de Vos, P.A. Bron, H.F. Savelkoul, *et al.* (2010). Identification of genetic loci in *Lactobacillus plantarum* that modulate the immune response of dendritic cells using comparative

genome hybridization. *PLoS One*; 5 (5): e10632.
140. Merritt M.E. and J.R. Donaldson (2009). Effect of bile salts on the DNA and membrane integrity of enteric bacteria. *J Med Microbiol*; 58 (Pt 12): 1533-1541.
141. Mihajlovski A., M. Alric and J.F. Brugere (2008). A putative new order of methanogenic Archaea inhabiting the human gut, as revealed by molecular analyses of the mcrA gene. *Res Microbiol*; 159 (7-8): 516-521.
142. Mohan R., C. Koebnick, J. Schildt, S. Schmidt, M. Mueller, M. Possner, M. Radke, *et al.* (2006). Effects of *Bifidobacterium lactis* Bb12 supplementation on intestinal microbiota of preterm infants: a double-blind, placebo-controlled, randomized study. *J Clin Microbiol*; 44 (11): 4025-4031.
143. Mondot S., S. Kang, J.P. Furet, D. Aguirre de Carcer, C. McSweeney, M. Morrison, P. Marteau, *et al.* (2011). Highlighting new phylogenetic specificities of Crohn's disease microbiota. *Inflamm Bowel Dis*; 17 (1): 185-192.
144. Montesi A., R. Garcia-Albiach, M.J. Pozuelo, C. Pintado, I. Goni and R. Rotger (2005). Molecular and microbiological analysis of caecal microbiota in rats fed with diets supplemented either with prebiotics or probiotics. *Int J Food Microbiol*; 98 (3): 281-289.
145. Moore W.E. and L.H. Moore (1995). Intestinal floras of populations that have a high risk of colon cancer. *Appl Environ Microbiol*; 61 (9): 3202-3207.
146. Mortensen P.B., J.R. Andersen, S. Arffmann and E. Krag (1987). Short-chain fatty acids and the irritable bowel syndrome: the effect of wheat bran. *Scand J Gastroenterol*; 22 (2): 185-192.
147. Mshvildadze M., J. Neu, J. Shuster, D. Theriaque, N. Li and V. Mai (2010). Intestinal microbial ecology in premature infants assessed with non-culture-based techniques. *J Pediatr*; 156 (1): 20-25.
148. Mueller S., K. Saunier, C. Hanisch, E. Norin, L. Alm, T. Midtvedt, A. Cresci, *et al.* (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 (2): 1027-1033.
149. Murphy E.F., P.D. Cotter, S. Healy, T.M. Marques, O. O'Sullivan, F. Fouhy, S.F. Clarke, *et al.* (2010). Composition and energy harvesting capacity of the gut microbiota: relationship to diet, obesity and time in mouse models. *Gut*; 59 (12): 1635-1642.
150. Musso G., R. Gambino and M. Cassader (2010). Obesity, diabetes, and gut microbiota: the hygiene hypothesis expanded? *Diabetes Care*; 33 (10): 2277-2284.
151. Mylonaki M., N.B. Rayment, D.S. Rampton, B.N. Hudspith and J. Brostoff (2005). Molecular characterization of rectal mucosa-associated bacterial flora in inflammatory bowel disease. *Inflamm Bowel Dis*; 11 (5): 481-487.
152. Nadal I., E. Donat, C. Ribes-Koninckx, M. Calabuig and Y. Sanz (2007). Imbalance in the composition of the duodenal microbiota of children with coeliac disease. *J Med Microbiol*; 56 (Pt 12): 1669-1674.
153. Nadal I., A. Santacruz, A. Marcos, J. Warnberg, M. Garagorri, L.A. Moreno, M. Martin-Matillas, *et al.* (2009). Shifts in clostridia, bacteroides and immunoglobulin-coating fecal bacteria associated with weight loss in obese adolescents. *Int J Obes (Lond)*; 33 (7): 758-767.
154. Nelson K.E., G.M. Weinstock, S.K. Highlander, K.C. Worley, H.H. Creasy, J.R. Wortman, D.B. Rusch, *et al.* (2010). A catalog of reference genomes from the human microbiome. *Science*; 328 (5981): 994-999.
155. Ng S.C., A.L. Hart, M.A. Kamm, A.J. Stagg and S.C. Knight (2009). Mechanisms of action of probiotics: recent advances. *Inflamm Bowel Dis*; 15 (2): 300-310.
156. Nicholson J.K. and J.C. Lindon (2008). Systems biology: Metabonomics. *Nature*; 455 (7216): 1054-1056.
157. Nishikawa J., T. Kudo, S. Sakata, Y. Benno and T. Sugiyama (2009). Diversity of mucosa-associated microbiota in active and inactive ulcerative colitis. *Scand J Gastroenterol*; 44 (2): 180-186.
158. O'Hara A.M. and F. Shanahan (2006). The gut flora as a forgotten organ. *EMBO reports*; 7 (7): 688-693.
159. Ohge H., J.K. Furne, J. Springfield, D.A. Rothenberger, R.D. Madoff and M.D. Levitt (2005). Association between fecal hydrogen sulfide production and pouchitis. *Dis Colon Rectum*; 48 (3): 469-475.
160. Onderdonk A.B., A.M. Dvorak, R.L. Cisneros, R.S. McLeod, D. Antionoli, W. Silen, J.E. Blair, *et al.* (1992). Microbiologic assessment of tissue biopsy samples from ileal pouch patients. *J Clin Microbiol*; 30 (2): 312-317.
161. Ott S.J., T. Kuhbacher, M. Musfeldt, P. Rosenstiel, S. Hellmig, A. Rehman, O. Drews, *et al.* (2008). Fungi and inflammatory bowel diseases: Alterations of composition and diversity. *Scand J Gastroenterol*; 43 (7): 831-841.
162. Ott S.J., M. Musfeldt, D.F. Wenderoth, J. Hampe, O. Brant, U.R. Folsch, K.N. Timmis, *et al.* (2004). Reduction in diversity of the colonic mucosa associated bacterial microflora in patients with active inflammatory bowel disease. *Gut*; 53 (5): 685-693.
163. Ouwehand A.C., N. Bergsma, R. Parhiala, S. Lahtinen, M. Gueimonde, H. Finne-Soveri, T. Strandberg, *et al.* (2008). Bifidobacterium microbiota and parameters of immune function in elderly subjects. *FEMS Immunol Med Microbiol*; 53 (1): 18-25.
164. Palmer C., E.M. Bik, D.B. DiGiulio, D.A. Relman and P.O. Brown (2007). Development of the human infant intestinal

microbiota. *PLoS Biol*; 5 (7): e177.
165. Parracho H.M., M.O. Bingham, G.R. Gibson and A.L. McCartney (2005). Differences between the gut microflora of children with autistic spectrum disorders and that of healthy children. *J Med Microbiol*; 54 (Pt 10): 987-991.
166. Penders J., C. Thijs, P.A. van den Brandt, I. Kummeling, B. Snijders, F. Stelma, *et al.* (2007). Gut microbiota composition and development of atopic manifestations in infancy: the KOALA Birth Cohort Study. *Gut*; 56 (5): 661-667.
167. Penders J., C. Thijs, C. Vink, F.F. Stelma, B. Snijders, I. Kummeling, P.A. van den Brandt, *et al.* (2006). Factors influencing the composition of the intestinal microbiota in early infancy. *Pediatrics*; 118 (2): 511-521.
168. Pettker C.M., I.A. Buhimschi, L.K. Magloire, A.K. Sfakianaki, B.D. Hamar and C.S. Buhimschi (2007). Value of placental microbial evaluation in diagnosing intra-amniotic infection. *Obstet Gynecol*; 109 (3): 739-749.
169. Pham M., D.A. Lemberg and A.S. Day (2008). Probiotics: sorting the evidence from the myths. *Med J Aust*; 188 (5): 304-308.
170. Preidis G.A. and J. Versalovic (2009). Targeting the human microbiome with antibiotics, probiotics, and prebiotics: gastroenterology enters the metagenomics era. *Gastroenterology*; 136 (6): 2015-2031.
171. Qin J., R. Li, J. Raes, M. Arumugam, K.S. Burgdorf, C. Manichanh, T. Nielsen, *et al.* (2010). A human gut microbial gene catalogue established by metagenomic sequencing. *Nature*; 464 (7285): 59-65.
172. Rajilic-Stojanovic M., H. Smidt and W.M. de Vos (2007). Diversity of the human gastrointestinal tract microbiota revisited. *Environ Microbiol*; 9 (9): 2125-2136.
173. Rijkers G.T., S. Bengmark, P. Enck, D. Haller, U. Herz, M. Kalliomaki, S. Kudo, *et al.* (2010). Guidance for substantiating the evidence for beneficial effects of probiotics: current status and recommendations for future research. *J Nutr*; 140 (3): 671S-676S.
174. Roberfroid M., G.R. Gibson, L. Hoyles, A.L. McCartney, R. Rastall, I. Rowland, D. Wolvers, *et al.* (2010). Prebiotic effects: metabolic and health benefits. *Br J Nutr*; 104 Suppl 2: S1-63.
175. Rochet V., L. Rigottier-Gois, A. Ledaire, C. Andrieux, M. Sutren, S. Rabot, A. Mogenet, *et al.* (2008). Survival of *Bifidobacterium animalis* DN-173 010 in the faecal microbiota after administration in lyophilised form or in fermented product - a randomised study in healthy adults. *J Mol Microbiol Biotechnol*; 14 (1-3): 128-136.
176. Roger L.C., A. Costabile, D.T. Holland, L. Hoyles and A.L. McCartney (2010). Examination of faecal *Bifidobacterium* populations in breast- and formula-fed infants during the first 18 months of life. *Microbiology*; 3329: 3329-3341.
177. Roger L.C. and A.L. McCartney (2010). Longitudinal investigation of the faecal microbiota of healthy full-term infants using fluorescence in situ hybridization and denaturing gradient gel electrophoresis. *Microbiology*; 156 (Pt 11): 3317-3328.
178. Rooijers K., C. Kolmeder, C. Juste, J. Dore, M. de Been, S. Boeren, P. Galan, *et al.* (2011). An Iterative Workflow for Mining the Human Intestinal Metaproteome. *BMC genomics*; 12 (1): 6.
179. Rosenfeld G. and B. Bressler (2010). *Mycobacterium avium paratuberculosis* and the etiology of Crohn's disease: a review of the controversy from the clinician's perspective. *Can J Gastroenterol*; 24 (10): 619-624.
180. Ruseler-van Embden J.G., W.R. Schouten and L.M. van Lieshout (1994). Pouchitis: result of microbial imbalance? *Gut*; 35 (5): 658-664.
181. Salonen A., W.M. de Vos and A. Palva (2010). Gastrointestinal microbiota in irritable bowel syndrome: present state and perspectives. *Microbiology*; 156 (Pt 11): 3205-3215.
182. Salonen A., J. Nikkila, J. Jalanka-Tuovinen, O. Immonen, M. Rajilic-Stojanovic, R.A. Kekkonen, A. Palva, *et al.* (2010). Comparative analysis of fecal DNA extraction methods with phylogenetic microarray: effective recovery of bacterial and archaeal DNA using mechanical cell lysis. *J Microbiol Methods*; 81 (2): 127-134.
183. Santacruz A., M.C. Collado, L. Garcia-Valdes, M.T. Segura, J.A. Martin-Lagos, T. Anjos, M. Marti-Romero, *et al.* (2010). Gut microbiota composition is associated with body weight, weight gain and biochemical parameters in pregnant women. *Br J Nutr*; 104 (1): 83-92.
184. Santacruz A., A. Marcos, J. Warnberg, A. Marti, M. Martin-Matillas, C. Campoy, L.A. Moreno, *et al.* (2009). Interplay between weight loss and gut microbiota composition in overweight adolescents. *Obesity*; 17 (10): 1906-1915.
185. Sanz Y., E. Sanchez, M. Marzotto, M. Calabuig, S. Torriani and F. Dellaglio (2007). Differences in faecal bacterial communities in coeliac and healthy children as detected by PCR and denaturing gradient gel electrophoresis. *FEMS Immunol Med Microbiol*; 51 (3): 562-568.
186. Sanz Y., A. Santacruz and P. Gauffin (2010). Gut microbiota in obesity and metabolic disorders. *Proc Nutr Soc*; 69 (3): 434-441.
187. Sartor R.B. (2008). Microbial influences in inflammatory bowel diseases. *Gastroenterology*; 134 (2): 577-594.
188. Satokari R., T. Gronroos, K. Laitinen, S. Salminen and E. Isolauri (2009). Bifidobacterium and Lactobacillus DNA in the human placenta. *Lett Appl Microbiol*; 48 (1): 8-12.
189. Satokari R.M., E.E. Vaughan, A.D. Akkermans, M. Saarela and W.M. De Vos (2001). Polymerase chain reaction and

denaturing gradient gel electrophoresis monitoring of fecal bifidobacterium populations in a prebiotic and probiotic feeding trial. *Syst Appl Microbiol*; 24 (2): 227-231.
190. Savage D.C. (1977). Microbial ecology of the gastrointestinal tract. *Annu Rev Microbiol*; 31: 107-133.
191. Scanlan P.D. and J.R. Marchesi (2008). Micro-eukaryotic diversity of the human distal gut microbiota: qualitative assessment using culture-dependent and -independent analysis of faeces. *ISME J*; 2 (12): 1183-1193.
192. Scanlan P.D., F. Shanahan, Y. Clune, J.K. Collins, G.C. O'Sullivan, M. O'Riordan, E. Holmes, *et al.* (2008). Culture-independent analysis of the gut microbiota in colorectal cancer and polyposis. *Environ Microbiol*; 10 (3): 789-798.
193. Scanlan P.D., F. Shanahan, C. O'Mahony and J.R. Marchesi (2006). Culture-independent analyses of temporal variation of the dominant fecal microbiota and targeted bacterial subgroups in Crohn's disease. *J Clin Microbiol*; 44 (11): 3980-3988.
194. Schiffrin E.J., J.E. Morley, A. Donnet-Hughes and Y. Guigoz (2010). The inflammatory status of the elderly: the intestinal contribution. *Mutat Res*; 690 (1-2): 50-56.
195. Schippa S., V. Iebba, M. Barbato, G. Di Nardo, V. Totino, M.P. Checchi, C. Longhi, *et al.* (2010). A distinctive 'microbial signature' in coeliac pediatric patients. *BMC Microbiol*; 10: 175.
196. Schwiertz A., D. Taras, K. Schafer, S. Beijer, N.A. Bos, C. Donus and P.D. Hardt (2010). Microbiota and SCFA in lean and overweight healthy subjects. *Obesity*; 18 (1): 190-195.
197. Sekirov I., S.L. Russell, L.C. Antunes and B.B. Finlay (2010). Gut microbiota in health and disease. *Physiol Rev*; 90 (3): 859-904.
198. Seksik P., L. Rigottier-Gois, G. Gramet, M. Sutren, P. Pochart, P. Marteau, R. Jian, *et al.* (2003). Alterations of the dominant faecal bacterial groups in patients with Crohn's disease of the colon. *Gut*; 52 (2): 237-242.
199. Shen X.J., J.F. Rawls, T. Randall, L. Burcal, C.N. Mpande, N. Jenkins, B. Jovov, *et al.* (2010). Molecular characterization of mucosal adherent bacteria and associations with colorectal adenomas. *Gut microbes*; 1 (3): 138-147.
200. Sjogren Y.M., M.C. Jenmalm, M.F. Bottcher, B. Bjorksten and E. Sverremark-Ekstrom (2009). Altered early infant gut microbiota in children developing allergy up to 5 years of age. *Clin Exp Allergy*; 39 (4): 518-526.
201. Snel J., Y.M. Vissers, B.A. Smit, J.M. Jongen, E.T. van der Meulen, R. Zwijsen, J. Ruinemans-Koerts, *et al.* (2010). Strain-specific immunomodulatory effects of *Lactobacillus plantarum* strains on birch-pollen-allergic subjects out of season. *Clin Exp Allergy*; 41 (2): 232-242.
202. Sokol H., B. Pigneur, L. Watterlot, O. Lakhdari, L.G. Bermudez-Humaran, J.J. Gratadoux, S. Blugeon, *et al.* (2008). *Faecalibacterium prausnitzii* is an anti-inflammatory commensal bacterium identified by gut microbiota analysis of Crohn disease patients. *Proc Natl Acad Sci USA*; 105 (43): 16731-16736.
203. Sokol H. and P. Seksik (2010). The intestinal microbiota in inflammatory bowel diseases: time to connect with the host. *Curr Opin Gastroenterol*; 26 (4): 327-331.
204. Sokol H., P. Seksik, J.P. Furet, O. Firmesse, I. Nion-Larmurier, L. Beaugerie, J. Cosnes, *et al.* (2009). Low counts of *Faecalibacterium prausnitzii* in colitis microbiota. *Inflamm Bowel Dis*; 15 (8): 1183-1189.
205. Sokol H., P. Seksik, L. Rigottier-Gois, C. Lay, P. Lepage, I. Podglajen, P. Marteau, *et al.* (2006). Specificities of the fecal microbiota in inflammatory bowel disease. *Inflamm Bowel Dis*; 12 (2): 106-111.
206. Song Y., C. Liu and S.M. Finegold (2004). Real-time PCR quantitation of clostridia in feces of autistic children. *Appl Environ Microbiol*; 70 (11): 6459-6465.
207. Stewart J.A., V.S. Chadwick and A. Murray (2005). Investigations into the influence of host genetics on the predominant eubacteria in the faecal microflora of children. *J Med Microbiol*; 54 (Pt 12): 1239-1242.
208. Stsepetova J., E. Sepp, K. Julge, E. Vaughan, M. Mikelsaar and W.M. de Vos (2007). Molecularly assessed shifts of *Bifidobacterium* ssp. and less diverse microbial communities are characteristic of 5-year-old allergic children. *FEMS Immunol Med Microbiol*; 51 (2): 260-269.
209. Suzuki S., N. Shimojo, Y. Tajiri, M. Kumemura and Y. Kohno (2007). Differences in the composition of intestinal *Bifidobacterium* species and the development of allergic diseases in infants in rural Japan. *Clin Exp Allergy*; 37 (4): 506-511.
210. Swidsinski A., A. Ladhoff, A. Pernthaler, S. Swidsinski, V. Loening-Baucke, M. Ortner, J. Weber, *et al.* (2002). Mucosal flora in inflammatory bowel disease. *Gastroenterology*; 122 (1): 44-54.
211. Swidsinski A., J. Weber, V. Loening-Baucke, L.P. Hale and H. Lochs (2005). Spatial organization and composition of the mucosal flora in patients with inflammatory bowel disease. *J Clin Microbiol*; 43 (7): 3380-3389.
212. Takaishi H., T. Matsuki, A. Nakazawa, T. Takada, S. Kado, T. Asahara, N. Kamada, *et al.* (2008). Imbalance in intestinal microflora constitution could be involved in the pathogenesis of inflammatory bowel disease. *Int J Med Microbiol*; 298 (5-6): 463-472.
213. Tana C., Y. Umesaki, A. Imaoka, T. Handa, M. Kanazawa and S. Fukudo (2010). Altered profiles of intestinal microbiota and organic acids may be the origin of symptoms in irritable bowel syndrome. *Neurogastroenterol Motil*; 22 (5): 512-

519, e114-515.
214. Tannock G.W., K. Munro, H.J. Harmsen, G.W. Welling, J. Smart and P.K. Gopal (2000). Analysis of the fecal microflora of human subjects consuming a probiotic product containing *Lactobacillus rhamnosus* DR20. *Appl Environ Microbiol*; 66 (6): 2578-2588.
215. Tap J., S. Mondot, F. Levenez, E. Pelletier, C. Caron, J.P. Furet, E. Ugarte*, et al.* (2009). Towards the human intestinal microbiota phylogenetic core. *Environ Microbiol*; 11 (10): 2574-2584.
216. Tasse L., J. Bercovici, S. Pizzut-Serin, P. Robe, J. Tap, C. Klopp, B.L. Cantarel*, et al.* (2010). Functional metagenomics to mine the human gut microbiome for dietary fiber catabolic enzymes. *Genome Res*; 20 (11): 1605-1612.
217. Tiihonen K., A.C. Ouwehand and N. Rautonen (2010). Human intestinal microbiota and healthy ageing. *Ageing Res Rev*; 9 (2): 107-116.
218. Timmerman H.M., C.J. Koning, L. Mulder, F.M. Rombouts and A.C. Beynen (2004). Monostrain, multistrain and multispecies probiotics - A comparison of functionality and efficacy. *Int J Food Microbiol*; 96 (3): 219-233.
219. Treem W.R., N. Ahsan, G. Kastoff and J.S. Hyams (1996). Fecal short-chain fatty acids in patients with diarrhea-predominant irritable bowel syndrome: in vitro studies of carbohydrate fermentation. *J Pediatr Gastroenterol Nutr*; 23 (3): 280-286.
220. Turnbaugh P.J., M. Hamady, T. Yatsunenko, B.L. Cantarel, A. Duncan, R.E. Ley, M.L. Sogin*, et al.* (2009). A core gut microbiome in obese and lean twins. *Nature*; 457 (7228): 480-484.
221. Turnbaugh P.J., R.E. Ley, M.A. Mahowald, V. Magrini, E.R. Mardis and J.I. Gordon (2006). An obesity-associated gut microbiome with increased capacity for energy harvest. *Nature*; 444 (7122): 1027-1031.
222. Vaahtovuo J., E. Munukka, M. Korkeamaki, R. Luukkainen and P. Toivanen (2008). Fecal microbiota in early rheumatoid arthritis. *The Journal of rheumatology*; 35 (8): 1500-1505.
223. Vael C. and K. Desager (2009). The importance of the development of the intestinal microbiota in infancy. *Curr Opin Pediatr*; 21 (6): 794-800.
224. Van Baarlen P., F.J. Troost, S. van Hemert, C. van der Meer, W.M. de Vos, P.J. de Groot, G.J. Hooiveld*, et al.* (2009). Differential NF-kappaB pathways induction by *Lactobacillus plantarum* in the duodenum of healthy humans correlating with immune tolerance. *Proc Natl Acad Sci USA*; 106 (7): 2371-2376.
225. Van den Bogert B., W.M. de Vos, E.G. Zoetendal and M. Kleerebezem (2011). Microarray analysis and barcoded pyrosequencing provide consistent microbial profiles depending on the source of human intestinal samples. *Appl Environ Microbiol*; 77 (6): 2071-2080.
226. Van Hemert S., M. Meijerink, D. Molenaar, P.A. Bron, P. de Vos, M. Kleerebezem, J.M. Wells*, et al.* (2010). Identification of *Lactobacillus plantarum* genes modulating the cytokine response of human peripheral blood mononuclear cells. *BMC Microbiol*; 10: 293.
227. Van Tongeren S.P., J.P. Slaets, H.J. Harmsen and G.W. Welling (2005). Fecal microbiota composition and frailty. *Appl Environ Microbiol*; 71 (10): 6438-6442.
228. Vanhoutte T., V. De Preter, E. De Brandt, K. Verbeke, J. Swings and G. Huys (2006). Molecular monitoring of the fecal microbiota of healthy human subjects during administration of lactulose and *Saccharomyces boulardii*. *Appl Environ Microbiol*; 72 (9): 5990-5997.
229. Vanhoutte T., G. Huys, E. Brandt and J. Swings (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 (3): 437-446.
230. Veiga P., C.A. Gallini, C. Beal, M. Michaud, M.L. Delaney, A. DuBois, A. Khlebnikov*, et al.* (2010). *Bifidobacterium animalis* subsp. lactis fermented milk product reduces inflammation by altering a niche for colitogenic microbes. *Proc Natl Acad Sci USA*; 107 (42): 18132-18137.
231. Verberkmoes N.C., A.L. Russell, M. Shah, A. Godzik, M. Rosenquist, J. Halfvarson, M.G. Lefsrud*, et al.* (2009). Shotgun metaproteomics of the human distal gut microbiota. *ISME J*; 3 (2): 179-189.
232. Vissers Y.M., J. Snel, P.F. Zuurendonk, M. Kleerebezem, H.J. Wichers and H.F. Savelkoul (2011). *Lactobacillus* strains differentially modulate cytokine production by hPBMC from pollen-allergic patients. *FEMS Immunol Med Microbiol*; 61 (1): 28-40.
233. Vissers Y.M., J. Snel, P.F. Zuurendonk, B.A. Smit, H.J. Wichers and H.F. Savelkoul (2010). Differential effects of *Lactobacillus acidophilus* and *Lactobacillus plantarum* strains on cytokine induction in human peripheral blood mononuclear cells. *FEMS Immunol Med Microbiol*; 59 (1): 60-70.
234. Vitali B., M. Ndagijimana, F. Cruciani, P. Carnevali, M. Candela, M.E. Guerzoni and P. Brigidi (2010). Impact of a synbiotic food on the gut microbial ecology and metabolic profiles. *BMC Microbiol*; 10: 4.
235. Vrieze A., F. Holleman, E.G. Zoetendal, W.M. de Vos, J.B. Hoekstra and M. Nieuwdorp (2010). The environment within: how gut microbiota may influence metabolism and body composition. *Diabetologia*; 53 (4): 606-613.

236. Walker A.W., J. Ince, S.H. Duncan, L.M. Webster, G. Holtrop, X. Ze, D. Brown, *et al.* (2010). Dominant and diet-responsive groups of bacteria within the human colonic microbiota. *ISME J*; 5 (2); 220-230.
237. Walker A.W., J.D. Sanderson, C. Churcher, G.C. Parkes, B.N. Hudspith, N. Rayment, J. Brostoff, *et al.* (2011). High-throughput clone library analysis of the mucosa-associated microbiota reveals dysbiosis and differences between inflamed and non-inflamed regions of the intestine in inflammatory bowel disease. *BMC Microbiol*; 11 (1): 7.
238. Wang M., S. Ahrne, B. Jeppsson and G. Molin (2005). Comparison of bacterial diversity along the human intestinal tract by direct cloning and sequencing of 16S rRNA genes. *FEMS Microbiol Ecol*; 54 (2): 219-231.
239. Wang M., C. Karlsson, C. Olsson, I. Adlerberth, A.E. Wold, D.P. Strachan, P.M. Martricardi, *et al.* (2008). Reduced diversity in the early fecal microbiota of infants with atopic eczema. *J Allergy Clin Immunol*; 121 (1): 129-134.
240. Wang X., S.P. Heazlewood, D.O. Krause and T.H. Florin (2003). Molecular characterization of the microbial species that colonize human ileal and colonic mucosa by using 16S rDNA sequence analysis. *J Appl Microbiol*; 95 (3): 508-520.
241. Wang Y., S. Devkota, M.W. Musch, B. Jabri, C. Nagler, D.A. Antonopoulos, A. Chervonsky, *et al.* (2010). Regional mucosa-associated microbiota determine physiological expression of TLR2 and TLR4 in murine colon. *PLoS One*; 5 (10): e13607.
242. Wang Y., J.D. Hoenig, K.J. Malin, S. Qamar, E.O. Petrof, J. Sun, D.A. Antonopoulos, *et al.* (2009). 16S rRNA gene-based analysis of fecal microbiota from preterm infants with and without necrotizing enterocolitis. *ISME J*; 3 (8): 944-954.
243. Watanabe S., Y. Narisawa, S. Arase, H. Okamatsu, T. Ikenaga, Y. Tajiri and M. Kumemura (2003). Differences in fecal microflora between patients with atopic dermatitis and healthy control subjects. *J Allergy Clin Immunol*; 111 (3): 587-591.
244. Weichselbaum E. (2010). Potential benefits of probiotics - main findings of an in-depth review. *Br J Community Nurs*; 15 (3): 110, 112, 114.
245. Willing B., J. Halfvarson, J. Dicksved, M. Rosenquist, G. Jarnerot, L. Engstrand, C. Tysk, *et al.* (2009). Twin studies reveal specific imbalances in the mucosa-associated microbiota of patients with ileal Crohn's disease. *Inflamm Bowel Dis*; 15 (5): 653-660.
246. Wong J.M., R. de Souza, C.W. Kendall, A. Emam and D.J. Jenkins (2006). Colonic health: fermentation and short chain fatty acids. *J Clin Gastroenterol*; 40 (3): 235-243.
247. Woodmansey E.J. (2007). Intestinal bacteria and ageing. *J Appl Microbiol*; 102 (5): 1178-1186.
248. Wu X., C. Ma, L. Han, M. Nawaz, F. Gao, X. Zhang, P. Yu, *et al.* (2010). Molecular characterisation of the faecal microbiota in patients with type II diabetes. *Curr Microbiol*; 61 (1): 69-78.
249. Zella G.C., E.J. Hait, T. Glavan, D. Gevers, D.V. Ward, C.L. Kitts and J.R. Korzenik (2010). Distinct microbiome in pouchitis compared to healthy pouches in ulcerative colitis and familial adenomatous polyposis. *Inflamm Bowel Dis*;17 (5): 1092-1100.
250. Zhang H., J.K. DiBaise, A. Zuccolo, D. Kudrna, M. Braidotti, Y. Yu, P. Parameswaran, *et al.* (2009). Human gut microbiota in obesity and after gastric bypass. *Proc Natl Acad Sci USA*; 106 (7): 2365-2370.
251. Zivkovic A.M., J.B. German, C.B. Lebrilla and D.A. Mills (2010). Microbes and Health Sackler Colloquium: Human milk glycobiome and its impact on the infant gastrointestinal microbiota. *Proc Natl Acad Sci USA*; 108 (Suppl 1): 4653-4658.
252. Zoetendal E.G., A.D. Akkermans and W.M. De Vos (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 (10): 3854-3859.
253. Zoetendal E.G., A.D.L. Akkermans, W.M. Akkermans-Van Vliet, J.A.G.M. de Visser and W.M. de Vos (2001). The host genotype affects the bacterial community in the human gastrointestinal tract. *Microbial Ecol Health Dis*; 13 (3): 129-134.
254. Zoetendal E.G., J. Raes, B. Van den Bogert, M. Arumugam, C.C.G.M. Booijink, F.J. Troost, *et al.* (2012). The human small intestinal microbiota is driven by rapid uptake and conversion of simple carbohydrates. *ISME J*; 6: 14-15-1426.
255. Zoetendal E.G., M. Rajilic-Stojanovic and W.M. de Vos (2008). High-throughput diversity and functionality analysis of the gastrointestinal tract microbiota. *Gut*; 57 (11): 1605-1615.
256. Zoetendal E.G., E.E. Vaughan and W.M. de Vos (2006). A microbial world within us. *Mol Microbiol*; 59 (6): 1639-1650.
257. Zoetendal E.G., A. von Wright, T. Vilpponen-Salmela, K. Ben-Amor, A.D. Akkermans and W.M. de Vos (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 (7): 3401-3407.
258. Zwielehner J., K. Liszt, M. Handschur, C. Lassl, A. Lapin and A.G. Haslberger (2009). Combined PCR-DGGE fingerprinting and quantitative-PCR indicates shifts in fecal population sizes and diversity of *Bacteroides*, bifidobacteria and *Clostridium* cluster IV in institutionalized elderly. *Exp Gerontol*; 44 (6-7): 440-446.

3 CHAPTER

Correlation between protection against sepsis by probiotic therapy and stimulation of a novel bacterial phylotype

Jacoline Gerritsen*, Harro M. Timmerman*, Susana Fuentes, L. Paul van Minnen, Henk Panneman, Sergey R. Konstantinov, Frans M. Rombouts, Hein G. Gooszen, Louis M. A. Akkermans, Hauke Smidt and Ger T. Rijkers

* These authors contributed equally to this work.

Applied and Environmental Microbiology, 77: 7749-7756 (2011)

Abstract

Prophylactic probiotic therapy has shown beneficial effects in an experimental rat model for acute pancreatitis on the health status of the animals. Mechanisms by which probiotic therapy interfere with severity of acute pancreatitis and associated sepsis, however, are poorly understood. The aims of this study were to identify the probiotic-induced changes in the intestinal microbiota and to correlate these changes to disease outcome. Duodenum and ileum samples were obtained from healthy and diseased rats subjected to pancreatitis for seven days and prophylactically treated with either a multispecies probiotic mixture or placebo. Intestinal microbiota was characterized by terminal-restriction fragment length polymorphism (T-RFLP) analyses of PCR-amplified 16S rRNA gene fragments. These analyses showed that during acute pancreatitis the host-specific ileal microbiota was replaced by an 'acute pancreatitis associated microbiota'. This replacement was not reversed by administration of the probiotic mixture. An increase, however, was observed in the relative abundance of a novel bacterial phylotype most closely related to *Clostridium lituseburense* and referred to as commensal rat ileum bacterium (CRIB). Specific primers targeting the CRIB 16S rRNA gene sequence were developed to detect this phylotype by quantitative PCR. An ileal abundance of CRIB 16S rRNA genes of more than 7.5% of the total bacterial 16S rRNA gene pool was correlated with reduced duodenal bacterial overgrowth, reduced bacterial translocation to remote organs, improved pancreas pathology, and reduced pro-inflammatory cytokine levels in plasma. Our current findings and future studies involving this uncharacterized bacterial phylotype will contribute to unraveling one of the potential mechanisms of probiotic therapy.

Introduction

Systemic inflammatory response syndrome (SIRS), sepsis, and multiple organ dysfunction syndrome (MODS) are major causes of mortality in severely ill patients. The role of the intestinal microbiota in initiation and propagation of critical illness has been increasingly acknowledged during the past decades. In several disease models it has been demonstrated that intestinal barrier dysfunction allows gut-derived bacteria to translocate to extra-intestinal organs and cause sepsis and secondary infectious complications resulting in high mortality rates [8].

Severe acute pancreatitis is an example of a critical illness that is characterized by intestinal barrier dysfunction. Acute pancreatitis usually has a mild and self-limiting clinical course; however, in about 20 to 30% of the patients it develops into a serious disease with SIRS and MODS. In these patients, infection of necrotic pancreatic tissue with gut-derived bacteria occurs in 30 to 70% of the cases [5]. These infectious complications are frequently the cause of mortality in patients with acute pancreatitis, with a mortality rate up to 50% [1, 4, 20].

We have previously shown in an experimental rat model for acute pancreatitis that prophylactic administration of a mixture of probiotic bacterial species resulted in reduced bacterial translocation to the pancreas and other extra-intestinal organs, improved disease outcome, and reduced late mortality of acute pancreatitis [26]. Although the effects of the probiotic therapy were highly significant, not all treated animals were protected from the adverse consequences of experimental acute pancreatitis. Furthermore, the underlying mechanism(s) by which probiotics ameliorate the severity of acute pancreatitis are incompletely understood.

In order to get a better insight in the role of gut microbiota in the process of bacterial translocation during experimental pancreatitis, a detailed analysis of the bacterial communities present in the various parts of the intestine was made. The aims of this study were to identify the probiotic-induced changes in the gut microbiota and to correlate these changes to disease outcome. As described here, molecular analysis of intestinal samples of rats treated with probiotics have led to a new hypothesis on how probiotics can improve the clinical course of acute pancreatitis.

Materials and methods

Animals

Male specific-pathogen-free Sprague-Dawley rats, 250 to 350 g (Harlan, Horst, The Netherlands), were kept under stable housing conditions with a 12-hour light/dark cycle and free access to water and food (RMH 1110; Hope Farms, Woerden, The Netherlands) throughout the experiment. All animals were allowed to adjust to these conditions for one week prior to surgery. Rats were randomized between two experimental groups: 17 rats were included in the group prophylactically treated with probiotics, and 21 rats were

assigned to the placebo group. Both groups were subjected to the surgical procedures described below. In addition six rats, which did not receive any surgical intervention or treatment, were used as healthy controls. All animal procedures were performed in accordance with institutional guidelines and with approval from the institutional animal care committee of the University Medical Center, Utrecht, The Netherlands.

Probiotics and placebo

The multispecies probiotic mixture consisted of equal amounts of six different viable, freeze-dried probiotic strains, *Bifidobacterium bifidum* (W23), *Bifidobacterium animalis* subsp. *lactis* (W52), *Lactobacillus acidophilus* (W70), *Lactobacillus casei* (W56), *Lactobacillus salivarius* (W24), and *Lactococcus lactis* (W58), blended in a carrier material consisting of maize starch, maltodextrins, and a mineral mix (Ecologic® 641, Winclove Bio Industries, Amsterdam, The Netherlands). The placebo product consisted of carrier material only. Directly before administration, both the freeze-dried placebo product and the probiotic mixture were reconstituted in sterile water for 15 min at 37°C. A single probiotic dose of 1.0 mL contained a total of about 5×10^9 colony forming units (CFU). Probiotics or placebo were administered intragastrically through a permanent gastric cannula once daily, starting five days prior to induction of acute pancreatitis, and twice daily for six days after induction of acute pancreatitis (see Figure S1 in the supplemental material).

Surgical procedures

Surgical procedures were performed as described previously [26]. Briefly, all procedures were performed under general anesthesia using a combination of 2% isoflurane gas and 0.3 mL 10% buprenorphine intramuscular (Temgesic; Reckitt Benckiser Healthcare Ltd., Hull, UK). At the start of the experiment, a permanent gastric cannula was fitted by tunneling a silicone cannula subcutaneously from the abdominal wall to the back of the animal. The gastric end of the cannula was inserted into the stomach through a puncture within a purse-string suture on the greater curvature. Animals were allowed to recover for three days prior to the start of daily probiotics or placebo administration. Five days after starting daily administration of probiotics or placebo, acute pancreatitis was induced as described previously, with minor adaptations [21]. The common bile duct was clamped and 0.5 mL sterilized glycodeoxycholic acid in 10 mM glycylglycine-NaOH-buffered solution (pH 8.0, 37°C; chemicals obtained from Sigma-Aldrich Chemie B.V., Zwijndrecht, The Netherlands) was infused, after which hepato-duodenal bile flow was restored. Next, the right jugular vein was cannulated for continuous intravenous infusion of cerulein (5 µg/kg per hour, for 6 h; chemicals obtained from Sigma-Aldrich Chemie B.V.).

Tissue and fluid samples

Seven days after induction of pancreatitis, surviving rats were anesthetized to allow

aseptic removal of tissue samples and sampling of peritoneal fluid and blood. After sample collection, rats were euthanized by blood loss. Mesenteric lymph nodes (MLNs), liver, spleen, pancreas, and duodenum were removed for microbiological analysis. After carefully removing all pancreatic and mesenteric tissue from the proximal duodenum and distal part of the terminal ileum, sections of approximately 2 cm were excised at both locations. Both segments were then transferred to a sterile vial, snap-frozen in liquid nitrogen and stored at -80°C. A portion of pancreatic tissue was fixed in 4% formalin and analyzed histopathologically, using standard hematoxylin and eosin (H&E) staining. Histopathological severity of acute pancreatitis was assessed based on a scoring system modified from Schmidt et al. [21] as previously described [26]. Several aspects of (peri)pancreatic histopathology were assessed (peritonitis, edema, ductal pathology, inflammatory infiltrate, acinar cells, acinar dilatation, and hemorrhagic changes). These aspects were scored on a scale, varying from maximal 2 to maximal 5 points per aspect.

All tissue samples were weighed and processed immediately for quantitative and qualitative culturing of aerobic and anaerobic microorganisms on appropriate media as described previously [26]. Tissue samples were homogenized in cysteine broth and cultured in 10-fold dilution series. The samples were cultured aerobically on blood agar, MacConkey agar and Columbia Colistin Nalidixic Acid agar, micro-aerobically on Man-Rogosa-Sharpe-agar, and anaerobically on Schaedler agar. Bacterial counts are presented in CFU/g tissue. Threshold detection level of bacterial growth was $>10^2$ CFU/g.

Multiplex cytokine assays

Plasma cytokine levels of interleukin-1 alpha (IL-1α), IL-6, IL-12p70, IL-18, interferon-gamma (IFN-γ), CXCL1 (growth related oncogene; GRO/KC) and CCL2 (monocyte chemoattractant protein-1; MCP-1), all from Linco Research, Inc. (St. Louis, MO, USA), and IL-1β, IL-2, IL-10, tumor necrosis factor alpha (TNF-α), all from Bio-Rad Laboratories (Hercules, CA, USA), were analyzed using rat cytokine multiplex assays according to manufacturer's instructions.

DNA extraction

Duodenum and ileum samples were taken from the freezer (-20°C) and a 1 cm section from the center of the sample was excised on a sterile field. A longitudinal incision over the full 1 cm section was made and the sample was left to thaw on a piece of sterile aluminum foil placed directly on a cooled metal plate at 4°C. After thawing, the intestinal content including the mucosa was scraped off with the back of a sterile surgical blade and collected in a 2 mL screw cap tube containing a buffer solution of 6 M guanidine thyocyanate, 0.6% Tween-20 (v/v), 10 mM EDTA, 50 mM Tris-HCl (pH 6.5) and 2 g of zirconia beads (<0.1 mM; Biospec, Bartlesville, OK, USA). For DNA extraction from pure bacterial cultures and the complete probiotic mixture, approximately 10^8 CFU were used.

Cells were physically disrupted by shaking for 4 min in a MiniBeadbeater-96™ (Biospec), followed by heating for 5 min at 90°C. Samples were centrifuged for 1 min at 10,000 × g and 200 µL of the supernatant was subsequently used for DNA isolation. DNA isolation was essentially performed as described previously [7], except for omitting the final washing step with acetone.

T-RFLP analysis

Terminal-restriction fragment length polymorphism (T-RFLP) analysis was performed essentially as described previously [15]. For PCR amplification of 16S rRNA gene fragments the universal primers 8F and 926R (Table 1), with fluorescent dyes 6-FAM and NED, respectively, at the 5' end, were used. Reaction mixtures (15 µL) contained PCR buffer (Applied Biosystems, Foster City, CA, USA), 62.5 µM of each deoxynucleoside triphosphate (dNTP; Applied Biosystems), 1.5 mM $MgCl_2$, 0.5 µM of each primer, 0.5 U of *Taq* DNA polymerase (Promega, Madison, WI, USA) and 1 µL of DNA sample. DNA amplification was performed with a 9700 thermal cycler (Perkin-Elmer, Norwalk, CT, USA) using the following program: 94°C for 5 min, followed by 35 cycles of 30 s at 94°C, 45 s at 56°C and 2 min at 72°C, and a final extension for 5 min at 72°C. PCR products were digested with MspI and HinP1I, with both enzymes in the same reaction (New England Biolabs, Ipswich, MA, USA). Resulting fragments were size-separated using an ABI 3100 genetic analyzer (Applied Biosystems) equipped with 36 cm capillaries using POP4 gel matrix. A custom size standard with ROX-labeled fragments, MapMarker 30–1,000 (Bioventures, Murfreesboro, TN, USA) was added to each sample prior to electrophoresis for accurate sizing of terminal-restriction fragments (T-RFs). After electrophoresis, sizes of fluorescently labeled T-RFs were determined using GeneScan software (Applied Biosystems).

Observed T-RFs were assigned to microbial taxa using the MCPC database (Dr. Van Haeringen Laboratorium B.V., Wageningen, The Netherlands), which was built using *in silico* T-RF predictions from data that were extracted from prokaryote sections of the EMBL sequence database (release 82, March 2005) using Patscan [9]. For each of the extracted sequences the T-RF size using the aforementioned combination of primers and restriction enzymes was calculated for both primers, and calculated values were validated for a number of pure cultures including *E. coli*, *Clostridium perfringens,* and *Staphylococcus aureus* (data not shown). The relative peak intensity of each T-RF (referred to as relative abundance) was defined as the height of a specific peak as percentage of the total sum of all peaks heights for a given sample. Relative abundance was calculated for all T-RFs between 50 and 600 bp long and with peak heights of >50 fluorescence units (FU), which was well above the background noise level (10-20 FU) of non-template control reactions. This procedure was followed for both primers; the data obtained with primer 8F was used in the figures.

Table 1. Primers and their targets used in this study

Primer	Primer sequence (5'-3')	Target	Reference
8F	AGA GTT TGA TCC TGG CTC (AG)	Universal 16S rRNA gene	14
926R	CCG TCA ATT CCT TTR AGT TT	Universal 16S rRNA gene	17
T7	TAA TAC GAC TCA CTA TAG G	pGEM-T vector specific	Promega
SP6	GAT TTA GGT GAC ACT ATA G	pGEM-T vector specific	Promega
519R	GWA TTA CCG CGG CKG CTG	Universal 16S rRNA gene	14
533F	GTG CCA GCA GCC GCG GTA A	Universal 16S rRNA gene	27
Bact1369F	CGG TGA ATA CGT TCY CGG	Universal 16S rRNA gene	24
Prok1492R	GGW TAC CTT GTT ACG ACT T	Universal 16S rRNA gene	24
CRIB-61F	GTC GAG CGA TTT ACT TCG GTA	CRIB-specific 16S rRNA gene	This study
CRIB-235R	GGG TCC ATC CTG TAC CGC AAA	CRIB-specific 16S rRNA gene	This study

Identification of the 16S rRNA gene sequence of CRIB

PCR reaction mixtures (50 µL) were prepared using *Taq* DNA polymerase kit from Invitrogen (Gaithersburg, MD, USA) and contained 0.5 µL of *Taq* DNA polymerase (1.25 U), 20 mM Tris-HCl (pH 8.5), 50 mM KCl, 3.0 mM $MgCl_2$, 5 pmol of the primers 8F and 926R (Table 1), 200 µM of each dNTP, and 1 µL of template DNA. Amplification was performed with a T1 thermocycler (Whatman Biometra, Göttingen, Germany) using the following program: 94°C for 5 min, followed by 35 cycles of 30 s at 94°C, 20 s at 56°C and 40 s at 68°C, and a final extension for 7 min at 68°C. Aliquots (5 µL) were analyzed by agarose electrophoresis to check for product size and quantity.

Products were purified with the QIAquick PCR purification kit (Westburg, Leusden, The Netherlands) according to manufacturer's instructions, and cloned into the pGEM-T easy vector system (Promega) using competent *E. coli* JM109 as a host. Colonies of ampicillin-resistant transformants were transferred with a sterile toothpick to 15 µL of Tris-EDTA buffer and boiled for 15 min at 95°C. PCR was performed using vector-specific primers T7 and SP6 (Table 1) to check the sizes of the inserts using the cell lysate as a template. PCR products of ~0.95 kb were purified as described above, and digested at 37°C for 90 min with CfoI, which is an isoschizomer of HinP1I, and MspI (Boehringer, Mannheim, Germany). One clone, containing an insert yielding a fragment of approximately 450 bp after restriction analysis, was grown in Luria Broth liquid medium (5 mL) with ampicillin (100 µg/mL) and used for further 16S rRNA gene sequence analysis. Plasmid DNA was isolated using the Wizard Plus purification system (Promega) and sequenced by using the Sequenase (T7) sequencing kit (Amersham Life Sciences, Slough, UK) according to manufacturer's specifications using primers T7, SP6, 519R and 533F (Table 1) labeled with IRD-800. Sequences were automatically analyzed on a LiCor DNA Sequencer 4000L (LiCor, Lincoln, NB, USA) and manually assembled.

Phylogenetic analysis and primer design

The CRIB 16S rRNA gene sequence was aligned with published 16S rRNA sequences (see

Figure S2 in the supplemental material) using the ARB software package and release 90 of the of the ARB-SILVA reference database [16, 18]. Different primer sets were developed using the ProbeDesign and ProbeMatch functions in ARB, and tested for specificity by conventional PCR. The primer set CRIB-61F/CRIB-235R (Table 1) was further tested by qPCR. For phylogenetic analyses as presented here, the sequence was realigned using the SINA WebAligner (http://www.arb-silva.de/), and compared to published sequences using ARB-Silva release 102.

qPCR analysis

Quantitative PCR (qPCR) was performed using the iQ5 Real-Time PCR Detection System (Bio-Rad). Reaction mixtures (25 µL) contained 12.5 µL iQ SYBR Green Supermix (Biorad), 0.2 µM of each primer and 5 µL of template DNA. All reactions were performed in triplicate. Standard curves were generated from a dilution series of 16S rRNA gene fragments amplified from the target sequence (10^8 to 10^0 copies/µL). The universal primer set Bact1369/Prok1492 (Table 1) was used for quantification of total bacterial 16S rRNA gene copies using the following program: 3 min at 95°C, followed by 40 cycles of 15 s at 95°C, 30 s at 56°C and 30 s at 72°C. Primers CRIB-61F/CRIB-235R were used to determine the relative abundance of CRIB 16S rRNA gene copies in the samples. Amplification conditions included 3 min at 95°C, followed by 40 cycles of 20 s at 95°C and 1 min at 68°C. Gradient PCR was initially used to determine the optimal annealing temperature, with amplification efficiency, range of linearity, and lowest detectable concentration as criteria. The predicted specificity of the CRIB-specific qPCR primer set was confirmed using nearly full-length 16S rRNA gene amplicons that were obtained, using primer pair 8F/Prok1492 (Table 1), from genomic DNA of closely related bacterial species: *Clostridium lituseburense* (DSM 797^T), *Clostridium irregulare* (DSM 2635^T), *Clostridium hiranonis* (DSM 13275^T), and *Clostridium bartlettii* (DSM 16795^T).

Relative abundance of CRIB was calculated by dividing CRIB 16S rRNA gene copies amplified using the primer set CRIB-61F/CRIB-235R by total 16S rRNA gene copies amplified using the primer set Bact1369/Prok1492 per µL of isolated DNA.

Statistical analysis

T-RFLP profiles were compared by visual inspection and Principal Component Analysis (PCA) using VHL Analysis software (Dr. Van Haeringen Laboratorium B.V.) that was custom-built in collaboration with Dalicon B.V. (Wageningen, The Netherlands) using IDL (RSI, Boulder, CO, USA). The number of T-RFs are presented as mean ± standard error of the mean (SEM).

The non-parametric Kruskal-Wallis test, followed by *post hoc* Mann-Whitney *U* tests with Bonferroni correction, were used to test for statistical differences in relative abundance of CRIB between the different independent treatment groups. The intraclass correlation

coefficient (ICC) was calculated to compare the T-RFLP to qPCR analysis. The correlation between two variables was described with Spearman's rank correlation coefficients and corresponding *P* values. Differences in bacterial counts, histopathological scores, and plasma cytokine levels between different groups of animals were analyzed using the Mann-Whitney *U* test. Differences with *P* values <0.05 were considered statistically significant.

Nucleotide sequence accession numbers

The partial 16S rRNA gene sequence of CRIB has been submitted to the GenBank/EMBL/DBBJ databases under accession number HQ224563.

Results

Acute pancreatitis induces changes in duodenal and ileal microbiota

To study the changes in duodenal and ileal microbiota as result of acute pancreatitis, duodenal and ileal samples were studied by T-RFLP analysis of PCR-amplified 16S rRNA gene fragments.

Previously we have demonstrated that there was a small increase in the total bacterial counts in the duodenum and a significant increase in culturable opportunistic pathogens seven days after induction of acute pancreatitis [26]. In the present study, T-RFLP analysis of duodenal microbiota demonstrated that the number of T-RFs almost doubled in the diseased animals (seven days after induction of acute pancreatitis) compared to healthy control animals (25 ± 3.5 *vs.* 41 ± 8.4; healthy control *vs*. placebo, respectively). T-RFLP analysis of the bacterial communities in the terminal ileum, however, showed that induction of acute pancreatitis resulted only in a modest increase in the number of T-RFs in diseased animals compared to healthy controls (60.3 ± 7.2 *vs* 69.0 ± 7.5; healthy control *vs.* placebo, respectively).

The composition of the ileal microbiota, as evaluated by PCA, was altered in the diseased animals compared to the healthy controls. Each healthy rat displayed a unique microbiota fingerprint, visible as a characteristic scattering of the samples in the two-dimensional space of the first principal components (Figure 1). In contrast, the samples of the diseased animals, with one exception, show a strong clustering demonstrating that the diseased animals apparently acquired a rather uniform 'acute pancreatitis associated microbiota' independent of which treatment was administered.

Probiotics stimulate a not described bacterium in the terminal ileum

T-RFLP analysis demonstrated that T-RFs of identical length as predicted for the probiotic bacteria could also be detected in placebo-treated animals indicating that the endogenous microbiota shares similar T-RFs with the probiotic bacteria (see Figure S3 and S4 in the supplemental material).

Probiotic-induced changes in the duodenal microbiota could not be detected based on T-RFLP analysis (Figure S3). In contrast, in the ileal samples a significant increase of an unknown bacterial phylotype (T-RF 457 bp, primer 8F, MspI and HinP1I digested) was observed in the probiotic-treated animals compared to the placebo-treated animals (Figure 2 and Figure S4). This T-RF could also be detected in a number of healthy control animals. It could be excluded that this bacterial phylotype was administered inadvertently, because it was absent in both the multispecies probiotic mixture and the carrier material (placebo) (data not shown).

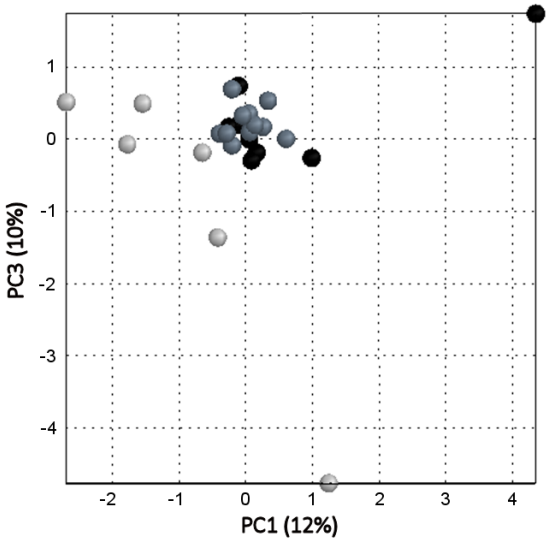

Figure 1. Correlations between the T-RFLP profiles obtained from the ileal samples mapped onto the first and third principle component of PCA. Healthy rats are indicated with light gray dots. Diseased rats are indicated by black dots (placebo-treated animals) or dark gray dots (probiotic-treated animals).

In order to further characterize this phylotype, a 16S rRNA gene clone library was generated from ileal DNA samples showing the 457 bp T-RF. Clones were screened by RF analysis using enzymes recognizing the same restriction sites as those applied for T-RFLP analysis, and a representative clone containing an insert leading to a fragment of approximately 450 bp was selected for further sequence analysis. The cloned 16S rRNA gene fragment, yielding a predicted T-RF of 457 bp, showed less than 97% sequence identity to the 16S rRNA gene sequences of bacterial isolates (Table 2), suggesting that it might represent a not previously described bacterial species, and which in the following is referred to as 'commensal rat ileum bacterium' (CRIB). Higher sequence similarities were observed with environmental sequences mostly retrieved from mammalian intestinal samples, including mouse and human (Figure 3).

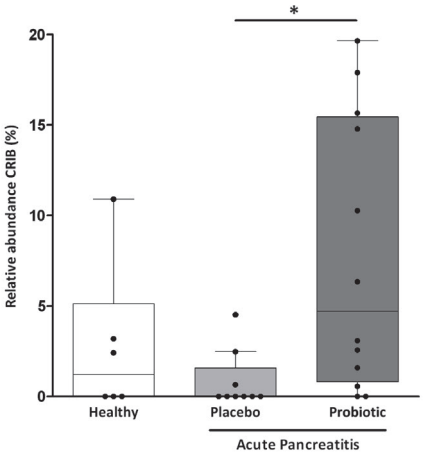

Figure 2. Effect of pancreatitis and treatment with either placebo or the probiotic mixture on the relative ileal abundance of T-RF 457, corresponding to the yet uncharacterized bacterial phylotype referred to as CRIB. Indicated are healthy control (white bar) and diseased animals, treated either with placebo (light gray bar) or the probiotic mixture (dark gray bar), respectively. Box-and-whisker diagrams represent median with interquartile range (boxes) and minimum and maximum non-outlier values (whiskers). An asterisk indicates a significant differences (P <0.05).

Based on the 16S rRNA gene sequence of the CRIB clone, a primer set was designed for the specific quantitative detection of CRIB in environmental samples by qPCR. In order to demonstrate that these primers indeed specifically detect the same 16S rRNA gene sequences that give rise to the 457 bp T-RF, qPCR was performed on the ileal DNA samples and compared to the T-RFLP results. An overall good agreement between the two techniques for calculation of the relative abundance of CRIB was demonstrated by the highly significant correlation between the two datasets (ICC = 0.893, P < 0.001; see Figure S5 in the supplemental material).

Table 2. 16S rRNA gene sequence identities of the CRIB clone with published sequences

GenBank accession no.	Description	% identity*
EU089965	*Clostridium maritimum* strain G12	96%
EU089964	*Clostridium ruminantium* strain LA1	96%
M59107	*Clostridium lituseburense* (DSM 797)	96%
AB538434	*Clostridium bifermentans* (JCM 7832)	95%
AB550230	*Clostridium sordellii* (JCM 3814)	94%
AF320283	*Clostridium bifermentans* (DSM 13560)	94%
AB075771	*Clostridium sordellii* (ATCC 9714)	94%
NR_027573	*Clostridium bartlettii* (DSM 16795)	94%
X73451	*Clostridium ghonii* (NCIMB 10636)	94%
X73450	*Clostridium difficile* (DSM 11209)	94%
NR_029249	*Clostridium irregulare* strain 6VI (DSM 2635)	93%
NR_028611	*Clostridium hiranonis* strain TO-931 (DSM 13275)	92%

* The percentage of identity was determined by comparison of the partial 16S rRNA gene sequence of CRIB with sequences present in the database using the BLAST tool from NCBI.

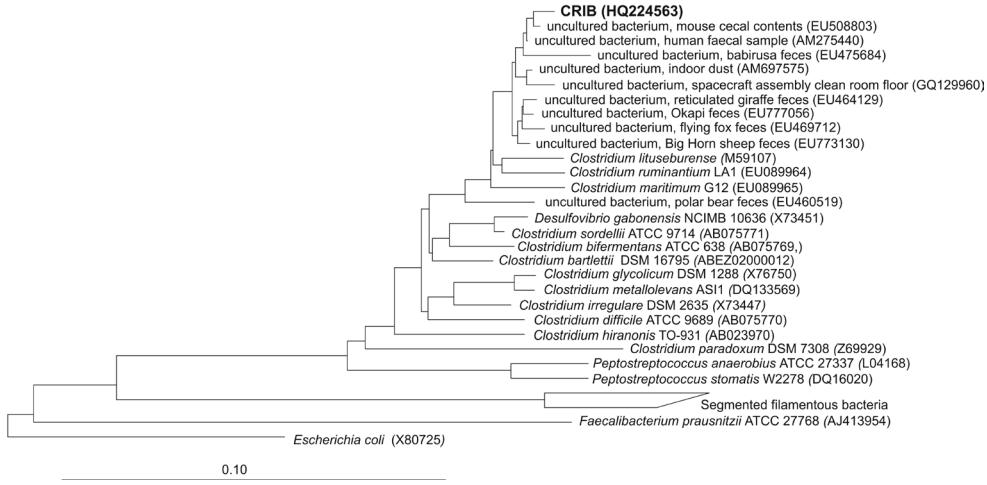

Figure 3. Neighbor-joining tree based on the 16S rRNA gene sequence of CRIB and other related *Clostridia*. *E. coli* was included as an outgroup. Alignment and phylogenetic analysis were performed with the ARB software [16]. Tree was calculated for *E. coli* positions 55 to 926. The reference bar indicates 10% sequence divergence. GenBank accession numbers are given in parantheses.

Relative ileal abundance of CRIB is associated with decreased disease severity

Different measures of disease outcome were analyzed in order to assess disease severity, including duodenal bacterial overgrowth, bacterial translocation to remote organs and pancreas pathology. The relative abundance of CRIB in the ileum of the diseased rats was inversely correlated with the degree of bacterial overgrowth in the duodenum (Figure 4a). Total bacterial counts were lower in the duodenum of animals with high relative ileal abundance of CRIB. In addition, bacterial infection of the MLNs (Figure 4b), as well as the liver, spleen and pancreas (see Figure S6 in the supplemental material) was also inversely correlated with the relative ileal abundance of CRIB. Furthermore, histopathological evaluation of pancreatic tissue obtained seven days after induction of acute pancreatitis, demonstrated that pancreas pathology was significantly and inversely correlated with the relative ileal abundance of CRIB (Figure 4c). Acinar cell pathology and inflammatory cell infiltrate were the aspects which contributed most to this correlation (Figure S6). It should be noted that the relative ileal numbers of the administered probiotic strains, or endogenous strains with identical T-RFs, did not correlate with the clinical and histological severity of the pancreatitis.

The data presented above suggest that the clinical effects of intervention with probiotics are mediated by CRIB. In order to substantiate this association, the diseased animals were divided in two groups with either a low (<7.5%) or a high (>7.5%) relative ileal abundance of CRIB, based on their average relative ileal abundance of CRIB as determined by qPCR analysis. A significantly lower total number of duodenal bacteria was detected

in the animals with >7.5% CRIB compared to the animals with <7.5% CRIB (P <0.05). In addition, a relative abundance of CRIB above 7.5% was also significantly correlated (P <0.05) with less bacterial infection of the MLNs and lower total histopathological scores of the pancreas.

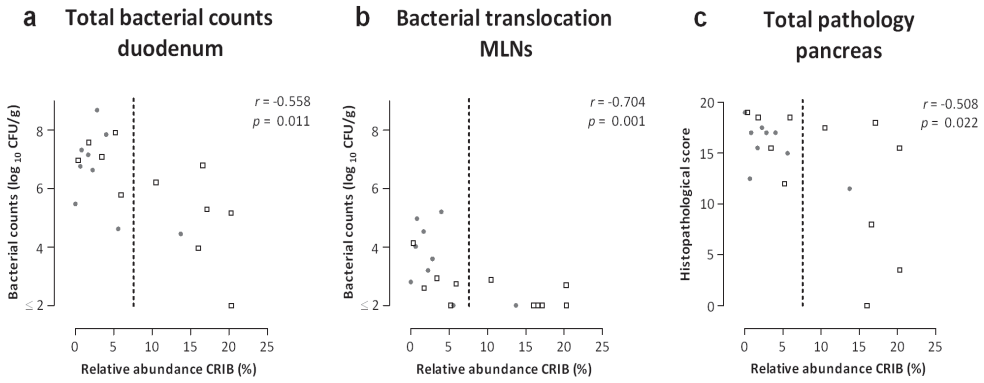

Figure 4. Associations between different measures of disease outcome and the relative ileal abundance of CRIB in diseased rats after seven days of acute pancreatitis. Animals were treated with either placebo (gray dots) or probiotics (quares). Indicated are (a) total bacterial counts in the duodenum, (b) bacterial translocation to the MLNs and (c) total pathology of the pancreas. The relative abundance of CRIB was determined by qPCR analysis. Bacterial counts are expressed in $\log_{10}$ CFU per gram sample ($\log_{10}$ CFU/g). Pearson's correlation coefficients are provided, with corresponding P values. The dotted lines indicate the division between the samples with a low (<7.5%) or high (>7.5%) relative ileal abundance of CRIB (>7.5%).

Relative ileal abundance of CRIB is associated with altered cytokine levels during acute pancreatitis

To determine the possible association between high relative abundance of CRIB in the ileum and cytokine plasma levels, a cytokine multiplex assay was performed on the plasma samples obtained seven days after induction of acute pancreatitis. In a previous study, we identified high plasma levels of IL-1β, IL-6, IL-10, CXCL1 and TNF-α as strong predictors of mortality and bacteremia during the course of acute pancreatitis in rats, and it was found that treatment with a multispecies probiotic mixture causes a mild reduction of these cytokines (manuscript in preparation). In line with this observation, the animals with high relative ileal abundance of CRIB (>7.5%) showed significantly lower plasma levels of IL-1β and TNF-α and a tendency to reduced levels of IL-10 and CXCL1 (Figure 5). In addition, elevated plasma levels of IL-18, a pro-inflammatory cytokine, which is significantly increased in acute pancreatitis patients with local or systemic complications [19], were only detected in animals having a low relative ileal abundance of CRIB (<7.5%; Figure 5b).

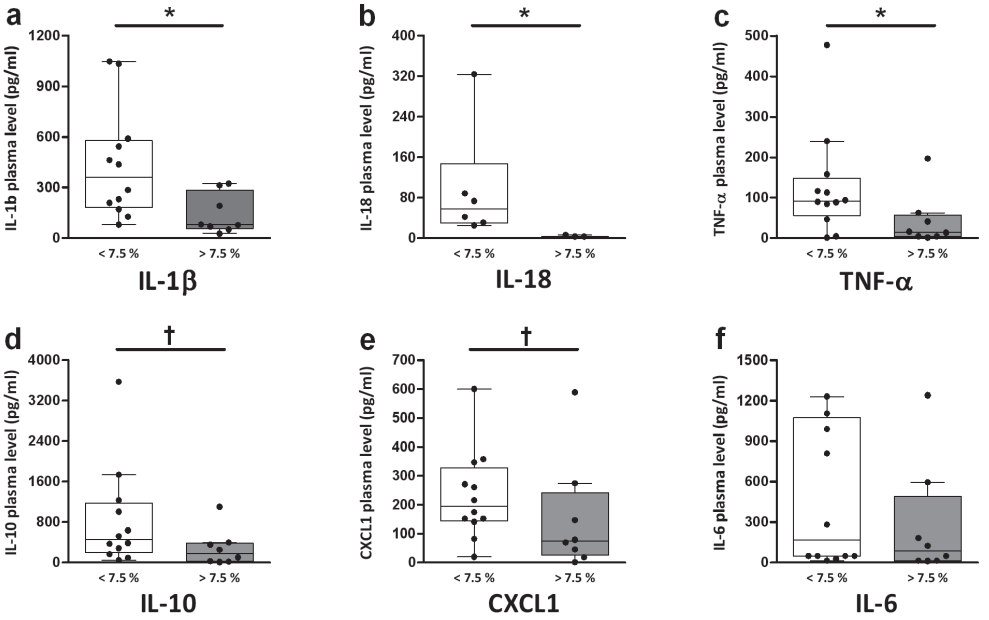

Figure 5. Effect of low (<7.5%; white bars) or high (>7.5%; gray bars) relative ileal abundance of CRIB in diseased rats on plasma cytokine levels of (a) IL-1β, (b) IL-18, (c) TNF-α, (d) IL-10, (e) CXCL1, and (f) IL-6. The relative abundance of CRIB was determined by qPCR analysis. Box-and-whisker diagrams represent median with interquartile range (boxes) and minimum and maximum non-outlier values (whiskers). Significant differences: *, $P < 0.05$; †, $P < 0.10$.

Discussion

A not previously described bacterial phylotype (CRIB) was identified in this study to be associated with reduced severity of pancreatitis and associated sepsis. A higher than average relative ileal abundance of CRIB (i.e. >7.5%) was significantly correlated with decreased duodenal bacterial overgrowth, reduced bacterial translocation to remote organs, reduced infection of pancreatic necrosis, and improved pancreas histology. In addition, high relative abundance of this bacterial phylotype was associated with less severe immune responses during acute pancreatitis as demonstrated by lower plasma levels of pro-inflammatory cytokines. Moreover, we demonstrated that there is an association between the presence of CRIB in the ileum of rats and the administration of a multispecies probiotic mixture. Together, these results suggest that effects of this multispecies probiotic mixture (Ecologic® 641) are mediated by stimulation of a not previously described gut commensal bacterium (CRIB), which protects the host from severe sepsis. The bacterial species most closely related to this not previously described bacterial phylotype is *Clostridium lituseburense*. For the detection of this bacterial phylotype, a specific and quantitative PCR (qPCR) assay was developed using 16S rRNA gene-targeted primers. This qPCR can be used in future studies to detect this phylotype in clinically relevant samples.

Dysbiosis of the small intestinal microbiota is often seen in critically ill patients and can be a cause for the development of sepsis. For example, small intestinal bacterial overgrowth can be detected in patients with hepatic cirrhosis and is associated with systemic endotoxemia [3, 28]. In the case of acute pancreatitis, the occurrence of small intestinal bacterial overgrowth is a pathological event during the course of acute pancreatitis due to impaired intestinal motility and is quantitatively and qualitatively correlated with bacterial translocation and infection of pancreatic necrosis [25, 26]. In addition, it has recently been demonstrated in a rat model that during acute pancreatitis bacterial translocation occurs from the small intestine rather than from the colon [10]. The duodenal bacterial counts found in this study confirm our previous findings that acute pancreatitis leads to duodenal bacterial overgrowth. Despite the occurrence of bacterial overgrowth in the duodenum, there was no change in composition of the duodenal microbiota. In contrast, the composition of the ileal microbiota was considerably changed upon induction of acute pancreatitis. PCA revealed that each healthy rat displayed a unique ileal microbiota fingerprint, as previously observed for ileal microbiota of healthy humans [6]. In contrast, increased similarity of the T-RFLP profiles of diseased animals indicated that these animals acquired an 'acute pancreatitis associated microbiota' independent of which treatment was deployed. Apparently, pancreatitis induces such a strong modification of the ileal microbiota that the unique individual microbial inhabitants are replaced by other, potentially pathogenic bacteria.

Previous studies have demonstrated that although the effects of this specific probiotic mixture were highly significant, not all probiotic-treated animals were protected from the adverse effects of experimental acute pancreatitis [26]. Here we show that one bacterial phylotype (referred to as CRIB) was significantly more abundant in the ileum of animals prophylactically treated with the probiotic mixture compared to the placebo-treated animals. However, both T-RFLP and qPCR analysis showed high inter-individual variation in the relative abundance of CRIB in the terminal ileum, which suggests that this bacterium is not an equally dominant member of the normal microbiota in every animal. This might explain why the administration of the multispecies probiotic mixture could not prevent pancreatitis-associated infectious complications in every animal. In fact, in animals that reacted poorly to probiotic treatment (e.g. the animals with high pancreatic bacterial counts), CRIB was only present in relatively low numbers (<7.5%). The chosen threshold value was set at 7.5%, based on the average relative abundance of CRIB in the total group of diseased animals. The relatively small sample size prevented a more accurate estimate of the threshold by ROC analysis. Still, the data point towards an important role of CRIB in the effectiveness of probiotic treatment. As the present study had not been designed to provide mechanistic insight in this association, it is currently unknown if and how probiotic administration can lead to an increase in relative abundance of CRIB, as well as a decrease in disease severity. The causality of the observed association between CRIB

and protection against severe sepsis will be addressed in future experiments.

The terminal ileum is the main site for interaction of commensal microbiota with the host immune system. This is reflected by the fact that the Peyer's patches, organized lymphoid tissue important for immune surveillance and initiating of immune responses in the gut, are mainly located in the ileum [13]. In our study we demonstrated that CRIB can be a dominant member of the ileal microbiota in rats, reaching up to 20% of the total microbiota in some animals. The mechanisms by which CRIB confers protective effects may include modulation of the mucosal immune system. This is suggested by the observation that the relative abundance of CRIB was correlated with altered plasma cytokine levels during acute pancreatitis.

Recently important roles in counterbalancing dysbiosis and regulation of immune responses have been suggested for several specific members of the gut microbiota. *Faecalibacterium prausnitzii* has been described to have anti-inflammatory properties in both *in vitro* and *in vivo* models [22], and underrepresentation of this bacterium in the intestinal microbiota was suggested to be involved in IBD pathogenesis in both animals and humans [23]. Furthermore, segmented filamentous bacteria (SFB) were shown to have a crucial role in maturation of T cell responses in the gut [11, 12]. *Faecalibacterium prausnitzii* and SFBs are members of the same taxonomic class as CRIB, the *Clostridia*, however, they belong to different subgroups (Figure 3). In addition, a recent study has demonstrated a significant induction of colonic regulatory T cells as result of the colonization by a specific mix of indigenous *Clostridium* species in a mouse experiment [2]. In that study, early inoculation of these *Clostridium* species resulted in the amelioration of the effects of experimental colitis and a lower IgE production Altogether, these studies demonstrate that specific members of the *Clostridia* can have health-promoting effects.

We anticipate that the findings reported here will be of key importance in unraveling one of the potential mechanisms of probiotic action and further understanding of the relation between gut microbiota and the host. Future experiments, including efforts towards the isolation of CRIB, will evaluate whether CRIB also plays a role in human disease.

Acknowledgements

We thank W. Akkermans-van Vliet (Wageningen University), W. Harmsen, W. de Jager, F. Lutgendorff, A. Verheem and M.R. Visser (all from UMC Utrecht) for expert technical assistance and advice. Part of this study was supported by Senter-Novem, an agency of the Dutch Ministry of Economic Affairs (grant number: TSGE-3109). Susana Fuentes was supported by a fellowship of the Fundación Alfonso Martín Escudero (FAME).

References

1. Ammori B.J. (2003). Role of the gut in the course of severe acute pancreatitis. *Pancreas*; 26 (2): 122-129.
2. Atarashi K., T. Tanoue, T. Shima, A. Imaoka, T. Kuwahara, Y. Momose, G. Cheng, *et al.* (2011). Induction of colonic regulatory T cells by indigenous *Clostridium* species. *Science*; 331 (6015): 337-341.
3. Bauer T.M., H. Schwacha, B. Steinbrückner, F.E. Brinkmann, A.K. Ditzen, J.J. Aponte, K. Pelz, *et al.* (2002). Small intestinal bacterial overgrowth in human cirrhosis is associated with systemic endotoxemia. *The American journal of gastroenterology*; 97 (9): 2364-2370.
4. Beger H.G., B. Rau and R. Isenmann (2003). Natural history of necrotizing pancreatitis. *Pancreatology*; 3 (2): 93-101.
5. Beger H.G. and B.M. Rau (2007). Severe acute pancreatitis: clinical course and management. *World journal of gastroenterology : WJG*; 13 (38): 5043-5051.
6. Booijink C.C., S. El-Aidy, M. Rajilić-Stojanović, H.G. Heilig, F.J. Troost, H. Smidt, M. Kleerebezem, *et al.* (2010). High temporal and inter-individual variation detected in the human ileal microbiota. *Environ Microbiol*; 12 (12): 3213-3227.
7. Carter M.J. and I.D. Milton (1993). An inexpensive and simple method for DNA purifications on silica particles. *Nucleic Acids Res*; 21 (4): 1044.
8. Clark J.A. and C.M. Coopersmith (2007). Intestinal crosstalk: a new paradigm for understanding the gut as the "motor" of critical illness. *Shock*; 28 (4): 384-393.
9. Dsouza M., N. Larsen and R. Overbeek (1997). Searching for patterns in genomic data. *Trends Genet*; 13 (12): 497-498.
10. Fritz S., T. Hackert, W. Hartwig, F. Rossmanith, O. Strobel, L. Schneider, K. Will-Schweiger, *et al.* (2010). Bacterial translocation and infected pancreatic necrosis in acute necrotizing pancreatitis derives from small bowel rather than from colon. *American journal of surgery*; 200 (1): 111-117.
11. Gaboriau-Routhiau V., S. Rakotobe, E. Lécuyer, I. Mulder, A. Lan, C. Bridonneau, V. Rochet, *et al.* (2009). The key role of segmented filamentous bacteria in the coordinated maturation of gut helper T cell responses. *Immunity*; 31 (4): 677-689.
12. Ivanov I.I., K. Atarashi, N. Manel, E.L. Brodie, T. Shima, U. Karaoz, D. Wei, *et al.* (2009). Induction of intestinal Th17 cells by segmented filamentous bacteria. *Cell*; 139 (3): 485-498.
13. Iwasaki A. (2007). Mucosal dendritic cells. *Annu Rev of Immunol*; 25: 381-418.
14. Lane D.J. (1991). 16S/23S rRNA sequencing. In: *Nucleic Acid Techniques in Bacterial Systematics*. pp 115-175. Edited by E. Stackebrandt and M. Goodfellow. New York: John Wiley and Sons, Inc.
15. Liu W.T., T.L. Marsh, H. Cheng and L.J. Forney (1997). Characterization of microbial diversity by determining terminal restriction fragment length polymorphisms of genes encoding 16S rRNA. *Appl Environ Microbiol*; 63 (11): 4516-4522.
16. Ludwig W., O. Strunk, R. Westram, L. Richter, H. Meier, Yadhukumar, A. Buchner, *et al.* (2004). ARB: a software environment for sequence data. *Nucleic Acids Res*; 32 (4): 1363-1371.
17. Muyzer G., A. Teske, C.O. Wirsen and H.W. Jannasch (1995). Phylogenetic relationships of *Thiomicrospira* species and their identification in deep-sea hydrothermal vent samples by denaturing gradient gel electrophoresis of 16S rDNA fragments. *Arch Microbiol*; 164 (3): 165-172.
18. Pruesse E., C. Quast, K. Knittel, B.M. Fuchs, W. Ludwig, J. Peplies and F.O. Glöckner (2007). SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. *Nucleic Acids Res*; 35 (21): 7188-7196.
19. Rau B., K. Baumgart, A.S. Paszkowski, J.M. Mayer and H.G. Beger (2001). Clinical relevance of caspase-1 activated cytokines in acute pancreatitis: high correlation of serum interleukin-18 with pancreatic necrosis and systemic complications. *Critical care medicine*; 29 (8): 1556-1562.
20. Schmid S.W., W. Uhl, H. Friess, P. Malfertheiner and M.W. Büchler (1999). The role of infection in acute pancreatitis. *Gut*; 45 (2): 311-316.
21. Schmidt J., D.W. Rattner, K. Lewandrowski, C.C. Compton, U. Mandavilli, W.T. Knoefel and A.L. Warshaw (1992). A better model of acute pancreatitis for evaluating therapy. *Ann Surg*; 215 (1): 44-56.
22. Sokol H., B. Pigneur, L. Watterlot, O. Lakhdari, L.G. Bermúdez-Humarán, J.J. Gratadoux, S. Blugeon, *et al.* (2008). *Faecalibacterium prausnitzii* is an anti-inflammatory commensal bacterium identified by gut microbiota analysis of Crohn disease patients. *Proceedings of the National Academy of Sciences of the United States of America*; 105 (43): 16731-16736.
23. Sokol H., P. Seksik, J.P. Furet, O. Firmesse, I. Nion-Larmurier, L. Beaugerie, J. Cosnes, *et al.* (2009). Low counts of *Faecalibacterium prausnitzii* in colitis microbiota. *Inflammatory bowel diseases*; 15 (8): 1183-1189.

24. Suzuki M.T., L.T. Taylor and E.F. DeLong (2000). Quantitative analysis of small-subunit rRNA genes in mixed microbial populations via 5'-nuclease assays. *Appl Environ Microbiol*; 66 (11): 4605-4614.
25. van Felius I.D., L.M.A. Akkermans, K. Bosscha, A. Verheem, W. Harmsen, M.R. Visser and H.G. Gooszen (2003). Interdigestive small bowel motility and duodenal bacterial overgrowth in experimental acute pancreatitis. *Neurogastroenterology and Motility*; 15 (3): 267-276.
26. van Minnen L.P., H.M. Timmerman, F. Lutgendorff, A. Verheem, W. Harmsen, S.R. Konstantinov, H. Smidt, *et al.* (2007). Modification of intestinal flora with multispecies probiotics reduces bacterial translocation and improves clinical course in a rat model of acute pancreatitis. *Surgery*; 141 (4): 470-480.
27. Weisburg W.G., S.M. Barns, D.A. Pelletier and D.J. Lane (1991). 16S ribosomal DNA amplification for phylogenetic study. *J Bacteriol*; 173 (2): 697-703.
28. Yang C.Y., C.S. Chang and G.H. Chen (1998). Small-intestinal bacterial overgrowth in patients with liver cirrhosis, diagnosed with glucose H2 or CH4 breath tests. *Scandinavian journal of gastroenterology*; 33 (8): 867-871.

Supplemental material

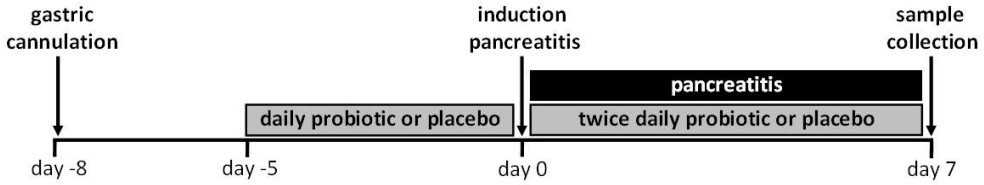

Figure S1. Experimental design. Eight days prior to induction of acute pancreatitis, a permanent gastric cannula was inserted. Probiotics or placebo were administered intragastrically through the permanent gastric cannula once daily, starting five days prior to induction of acute pancreatitis, and twice daily from day zero until six days after induction of acute pancreatitis. Seven days after induction of acute pancreatitis, surviving rats were anesthetized to allow aseptic removal of tissue and fluid samples. After sample collection, rats were euthanized by blood loss.

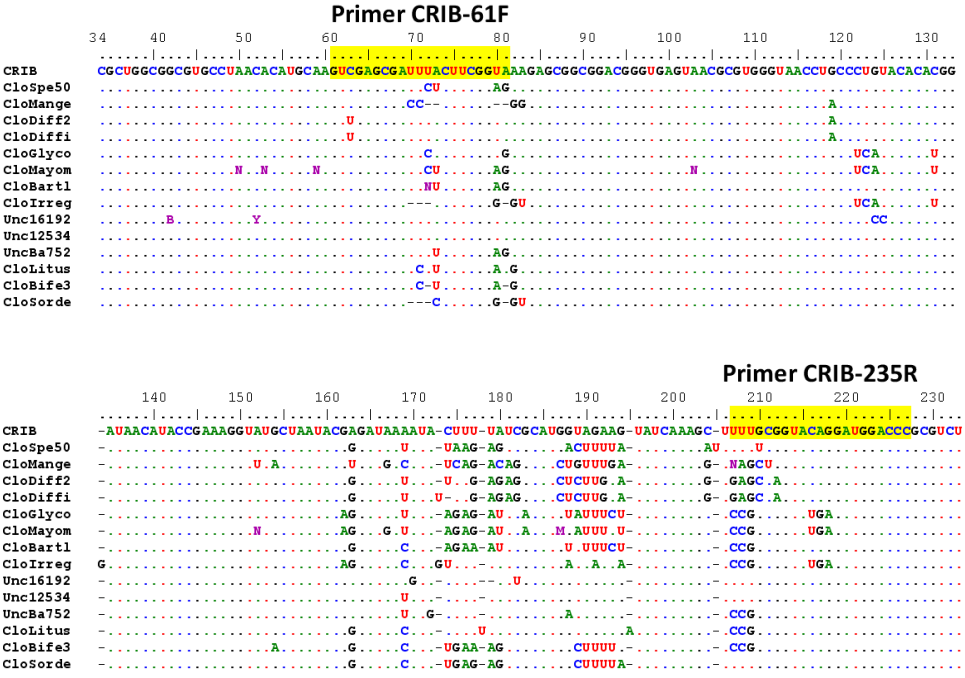

Figure S2. Alignment of the 16S rRNA gene sequence of CRIB with selected reference sequences of closely related *Clostridium* spp and environmental clones. Nucleotides identical to the CRIB sequence are indicated by dots; gaps are shown as dashes. The positions of the primers CRIB-61F and CRIB-235R, developed to specifically detect CRIB, are indicated in yellow.

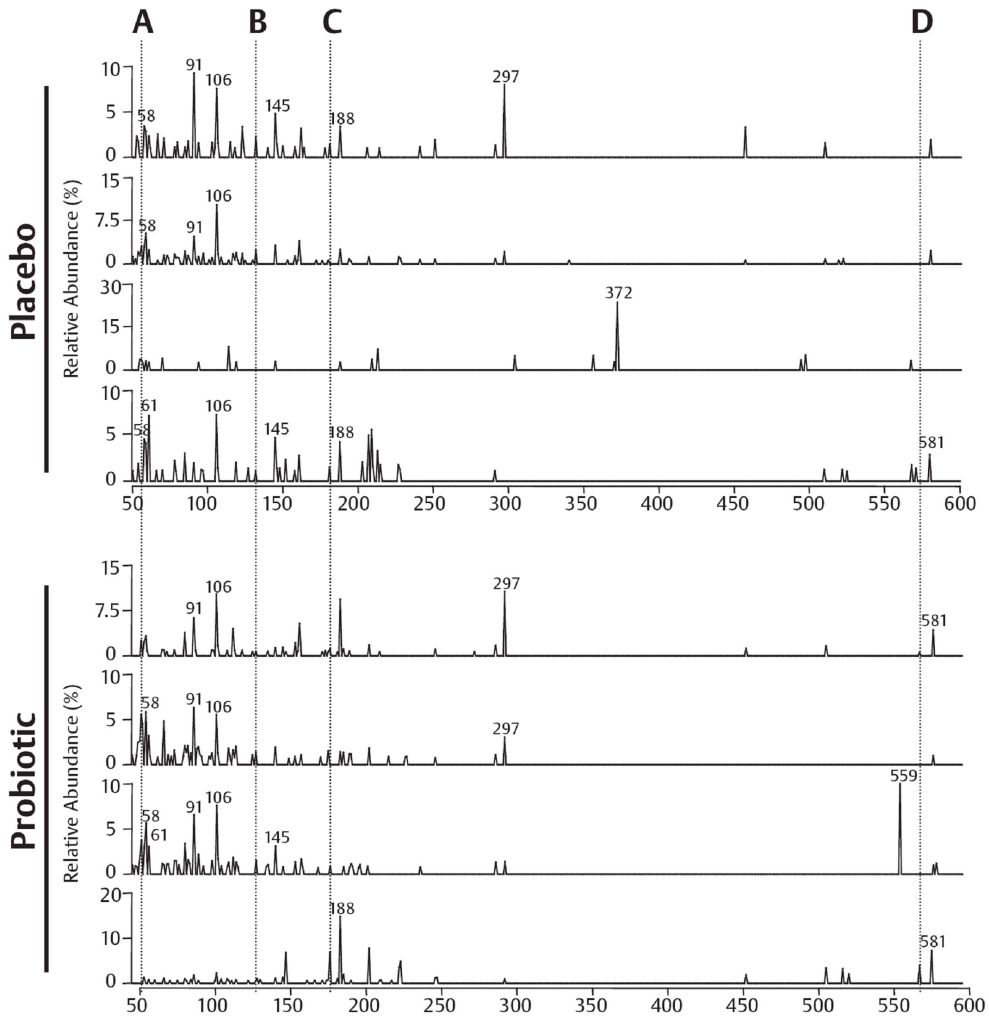

Figure S3. T-RFLP patterns of the duodenal microbiota of 8 representative diseased rats, treated with either placebo (upper 4 traces) or the probiotic mixture (lower 4 traces). Duodenal samples were obtained 7 days after induction of pancreatitis. Major TR-F peaks with corresponding bacterial species are: 58 - *Lactobacillus* spp., 61 - *Myxococcus* spp., 91 - *Prevotella* spp., 106 - unknown, 145 - *L. thermophilus*, 188 - *L. johnsonii*, 372 - *Escherichia coli*, 559 - *Streptococcus* spp. and 581 - *Pediococcus pentosaceus*. TR-F peaks matching with the 6 applied probiotic strains, as determined experimentally (data not shown), are indicated with dotted lines and can be divided into four groups: A - *Lactococcus lactis*, B - *Bifidobacterium bifidum* and *B. lactis*, C -*Lactobacillus acidophilus* and D - *L. casei* and *L. salivarius*.

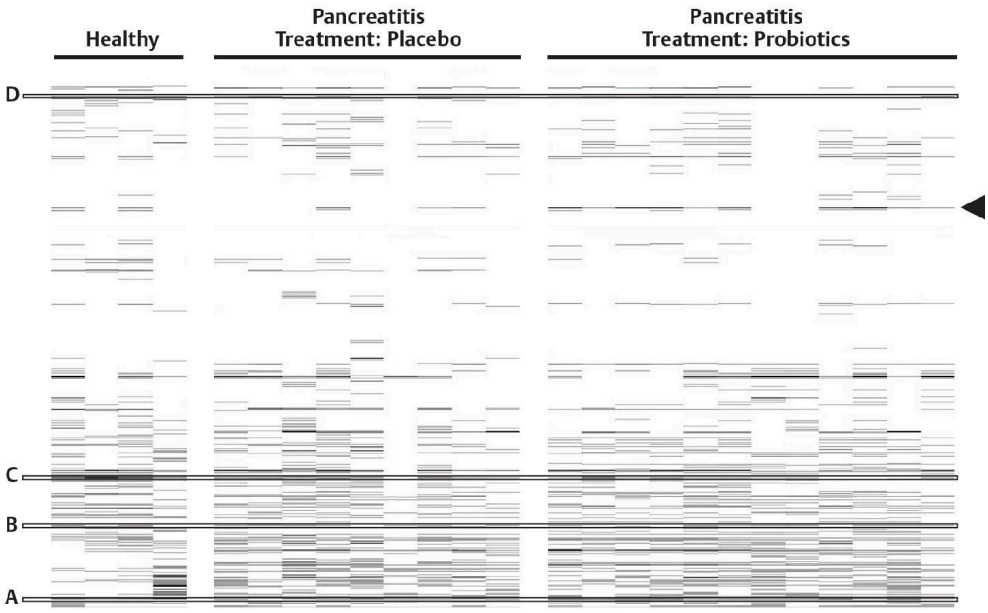

Figure S4. Representative T-RFLP patterns of the ileal microbiota of healthy control and diseased animals, treated with either placebo or the probiotic mixture. Ileal samples from the diseased animals were obtained 7 days after induction of pancreatitis. The arrowhead indicates the position of T-RF 457, corresponding to a yet uncharacterized bacterial phylotype (referred to as commensal rat ileum bacterium; CRIB). T-RF peaks matching with the 6 applied probiotic strains are indicated with thin boxes and can be divided into four groups: A - *Lactococcus lactis*, B - *Bifidobacterium bifidum* and *B. lactis*, C -*Lactobacillus acidophilus* and D - *L. casei* and *L. salivarius*.

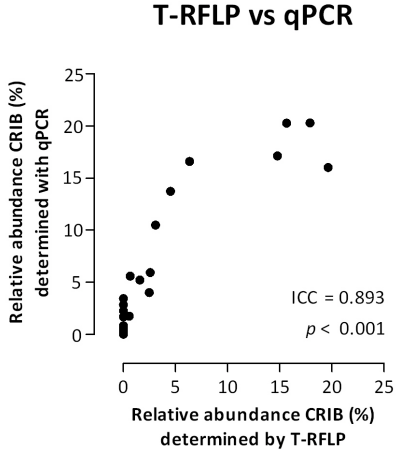

Figure S5. Correlation between the relative abundance of CRIB determined by T-RFLP analysis (relative peak intensity of T-RF 457 bp) and qPCR analysis (relative abundance of 16S rRNA gene copies amplified with CRIB-specific primers). The intraclass correlation coefficient (ICC) is provided, with corresponding *P* value.

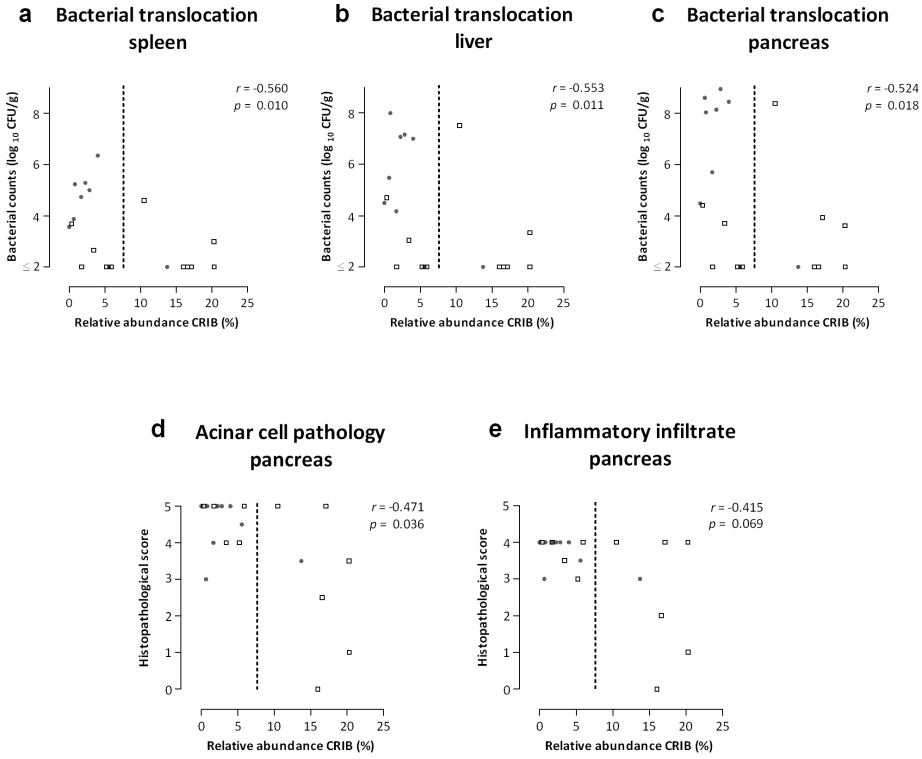

Figure S6. Associations between different measures of disease outcome and the relative ileal abundance of CRIB in diseased rats after seven days of acute pancreatitis. Animals were treated with either placebo (gray dots) or probiotics (black-lined squares). Indicated are bacterial translocation to the (**a**) spleen, (**b**) liver and (**c**) pancreas, and histopathological scores of the pancreas of (**d**) acinar cell pathology and (**e**) inflammatory infiltrate. Relative abundance of CRIB was determined by qPCR analysis. Bacterial counts are expressed in $\log_{10}$ colony forming units per gram sample ($\log_{10}$ CFU/g). Spearman's correlation coefficients are provided, with corresponding *P* values. The dotted lines indicate the division between samples containing a low (<7.5%) or high (>7.5%) relative ileal abundance of CRIB (>7.5%).

Table S1. Overview of experimental data

			Relative abundance CRIB (%)		Microbiology: bacterial counts (Log_{10} CFU/g)					Histopathological score			Plasma cytokine levels (pg/ml)					
Rat #	Treatment	Death by acute pancreatitis (hours after induction)	T-RFLP analysis	qPCR analysis	Duodenum	MLN	Spleen	Liver	Pancreas	Total pathology	Acinar cell pathology	Inflammatory infiltrate	IL1B	IL6	IL10	IL18	TNFa	CXCL1
1	healthy		10.90															
2	healthy		0.00															
3	healthy		2.42															
4	healthy		0.00															
5	healthy		3.20															
6	healthy		0.00															
7	placebo		Technical failure	Technical failure	5.59	2	2	2	2	13.5	4.5	4						
8	probiotic		14.78	17.15	5.28	2	2	2	3.93	18	5	4	69.14	123.82	104.52	Technical failure	16.14	69.59
9	probiotic		1.59	5.21	7.91	2	2	2	2	12	4	3	464.65	26.94	283.69	Technical failure	46.61	140.83
10	placebo		Technical failure	Technical failure	3.71	2	2	2	2	15.5	5	4						
11	probiotic		17.90	20.31	2	2	2	2	2	3.5	1	1	51.81	14.42	10.79	Technical failure	4.77	45.2
12	placebo	Yes (36 hours)																
13	probiotic		15.66	20.29	5.16	2.7	2.98	3.34	3.62	15.5	3.5	4	78.88	14.42	397.39	Technical failure	4.77	146.43
14	placebo	Yes (48 hours)																
15	probiotic	Yes (12 hours)																
16	placebo		0.00	0.67	6.76	4.02	3.87	5.48	8.61	12.5	3	3	80.94	48.48	384.45	41.97	158.07	357.6
17	probiotic	Yes (36 hours)																
18	placebo		0.00	0.83	7.31	4.97	5.23	8	8.04	17	5	4	1034.82	14655.5	1235.27	323.76	84.86	174.39
19	placebo		0.00	0.02	5.47	2.81	3.56	4.49	4.48	19	5	4	127.75	1229.56	369.91	25.18	88.82	271.39
20	placebo	Yes (160 hours)																
21	probiotic		0.56	1.74	7.58	2.6	2	2	2	18.5	5	4	171.42	808.83	1736.59	30.9	113.24	152.13
22	placebo	Yes (12 hours)																
23	placebo	Yes (12 hours)																
24	probiotic		6.35	16.6	6.79	2	2	2	2	8	2.5	2	26.73	1240.8	28.89	3.26	1.67	1.75
25	probiotic		0.00	0.37	6.96	4.14	3.69	4.7	4.41	19	5	4	231.39	988.37	167.74	73.38	93.74	216.16
26	probiotic		0.00	3.44	7.08	2.93	2.65	3.04	3.71	15.5	4	3.5	210.2	1104.5	96.43	88.46	1.67	19.67
27	probiotic		10.27	Technical failure	3.92	2.98	2	2.54	2.93	12.5	3.5	3	313.96	49.02	1103.53	Technical failure	196.6	588.92
28	probiotic		19.65	16.03	3.97	2	2	2	2	0	0	0	81.85	12.41	355.32	6.52	62.81	79.31
29	placebo		4.52	13.73	4.45	2	2	2	2	11.5	3.5	3	192.2	594.93	19.81	3.26	13.86	18.69
30	probiotic		Technical failure	Technical failure	4.4	3.24	4.12	2.74	3.22	16.5	4.5	4						
31	placebo	Yes (96 hours)																
32	placebo	Yes (12 hours)																
33	probiotic	Yes (12 hours)																
34	placebo	Yes (24 hours)																
35	probiotic		3.09	10.49	6.21	2.88	4.6	7.51	8.38	17.5	5	4	324.05	181.77	256.1	Technical failure	41.47	274.02
36	probiotic	Yes (36 hours)																
37	placebo		0.00	1.68	7.15	4.54	4.73	4.17	5.7	15.5	4	4	543.63	281.72	634.72	Technical failure	116.53	151.84
38	probiotic		2.57	5.93	5.78	2.74	2	2	2	18.5	5	4	590.14	49.02	44.32	Technical failure	4.77	81.94
39	placebo		0.00	2.84	8.68	3.6	5	7.16	8.95	17	5	4	286.31	14.42	521.28	Technical failure	89.93	260.48
40	placebo		0.66	5.58	4.62	2	2	2	2	15	4.5	3.5						
41	placebo		2.48	4.01	7.84	5.21	6.35	7	8.45	17	5	4	437.07	49.02	1005.22	Technical failure	240.03	346.77
42	placebo		0.00	2.25	6.63	3.2	5.29	7.07	8.15	17.5	5	4	1047.97	49.02	3573.03	Technical failure	477.15	600.14
43	placebo	Yes (36 hours)																
44	placebo	Yes (160 hours)																

4 CHAPTER

Characterization of *Romboutsia ilealis* gen. nov., sp. nov., isolated from the gastro-intestinal tract of a rat and proposal for the reclassification of five closely related members of the genus *Clostridium* into the genera *Romboutsia* gen. nov., *Intestinibacter* gen. nov., *Terrisporobacter* gen. nov. and *Asaccharospora* gen. nov.

Jacoline Gerritsen, Susana Fuentes, Wieke Grievink, Laura van Niftrik, Brian J. Tindall, Harro M. Timmerman, Ger T. Rijkers and Hauke Smidt

International Journal of Systematic and Evolutionary Microbiology, 64: 1600-1616 (2014)

Abstract

A Gram-positive staining, rod-shaped, non-motile, spore-forming obligately anaerobic bacterium, designated CRIB^T, was isolated from the gastro-intestinal tract of a rat and characterized. The major cellular fatty acids of strain CRIB^T were saturated and unsaturated straight chain C_{12}-C_{19} fatty acids, with $C_{16:0}$ being the predominant fatty acid. The polar lipid profile comprised six glycolipids, four phospholipids and one lipid that did not stain with any of the specific spray reagents used. The only quinone was MK-6. The predominating cell wall sugars were glucose and galactose. The peptidoglycan type of strain CRIB^T was A1δ lanthionine-direct. The genomic DNA G+C content of strain CRIB^T was 28.1 mol%. On the basis of 16S rRNA gene sequence similarity, strain CRIB^T was most closely related to a number of species of the genus *Clostridium*, including *Clostridium lituseburense* (97.2%), *Clostridium glycolicum* (96.2%), *Clostridium mayombei* (96.2%), *Clostridium bartlettii* (96.0%) and *Clostridium irregulare* (95.5%). All these species show very low 16S rRNA gene sequence similarity (<85%) to the type strain of *Clostridium butyricum*, the type species of the genus *Clostridium*. DNA-DNA hybridization with closely related reference strains indicated reassociation values below 32%. On the basis of phenotypic and genetic studies, a novel genus, *Romboutsia* gen. nov., is proposed. The novel isolate CRIB^T (=DSM 25109^T =NIZO 4048^T) is proposed as the type strain of the type species, *Romboutsia ilealis* gen. nov., sp. nov. of the proposed novel genus. It is proposed that *C. lituseburense* is transferred to this genus as *Romboutsia lituseburensis* comb. nov. Furthermore, the reclassification into novel genera is proposed for *C. bartlettii* as *Intestinibacter bartlettii* comb. nov. (type species of the genus), *C. glycolicum* as *Terrisporobacter glycolicus* comb. nov. (type species of the genus), *C. mayombei* as *Terrisporobacter mayombei* comb. nov., and *C. irregulare* as *Asaccharospora irregularis* comb. nov. (type species of the genus), on the basis of additional data collected in this study. In addition, an emendation of the species *Peptostreptococcus anaerobius* and the order *Eubacteriales* is provided.

The gastro-intestinal tract of both humans and animals contains an enormous diversity in microbial species (**Chapter 2** [20, 44]), of which many still remain to be cultured and characterized. Here we describe a novel isolate obtained from the gastro-intestinal tract of rats. This isolate is the first representative of a novel bacterial phylotype (referred to as 'CRIB'), which was recently reported by Gerritsen *et al.* to be associated with probiotic-induced changes in the gut microbiota of rats in an experimental model for acute pancreatitis (**Chapter 3** [21]). To determine the taxonomic position of this isolate, it was subjected to further characterization in accordance with the recommendations provided by Kämpfer *et al.* [33] and Tindall *et al.* [80].

Strain CRIB^T was isolated from an ileal digesta sample obtained from a healthy Sprague-Dawley rat that was selected based on high relative abundance of the target phylotype as determined using quantitative PCR (qPCR) and the primer set CRIB-61F/CRIB-235R as described previously (**Chapter 3** [21]). Isolation was performed by serial dilution of the sample in liquid anoxic bicarbonate-buffered medium, further referred to as CRIB medium. CRIB medium consisted of a basal bicarbonate-buffered medium [72] supplemented with (per litre distilled water): 30 g bacteriological peptone, 5 g yeast extract, 5 g beef extract, 4 g glucose, 1 g cellobiose, 1 g maltose, 1 g soluble starch, 0.5 g L-cysteine hydrochloride, 0.4 g bile salts, 0.25 mg hemin, 0.0001% (v/v) vitamin K_1 and 0.5% (v/v) clarified, sterile rumen fluid. The final pH of the medium was 7.0. The increase in relative abundance of the target phylotype was followed in the subsequent serial dilutions using qPCR as described previously (**Chapter 3** [21]). After primary isolation, accomplished by repeated rounds of fast transfers (within 24 hours), the strain was purified by repeated inoculation and sub-culturing on solid CRIB medium (containing 1.5% w/v Bacto™ agar, BD, Sparks, MD, USA) using the anaerobic roll tube method. The purified strain was cultured routinely in liquid CRIB medium at 37°C (pH 7.0), and stored as a glycerol suspension (25% v/v) at –80°C. Unless indicated otherwise, morphological, physiological, molecular and chemotaxonomic studies were performed with cells grown on liquid anoxic basal peptone-yeast extract (PY) medium [27] supplemented with 0.5% (w/v) glucose at 37°C (pH 7.0), further referred to as PYG medium. Solid PYG or CRIB medium was prepared by supplementation of liquid medium with 0.8% gelrite (Carl Roth GmbH + Co. KG, Karlsruhe, Germany).

Total DNA was extracted for phylogenetic analysis of strain CRIB^T using the FastDNA SPIN Kit for Soil (MP Biomedicals, Illkirch Cedex, France). The 16S rRNA gene was amplified by PCR using the universal primers 8F [41] and 1492R [74], and PCR products were purified using the High Pure PCR Cleanup Micro Kit (Roche Diagnostics GmbH, Mannheim, Germany). In order to obtain an almost complete 16S rRNA gene sequence of strain CRIB^T, purified PCR products were cloned in *Escherichia coli* using the pGEM-T easy vector system (Promega, Madison, WI, USA) according to the manufacturer's instructions. Plasmid DNA was isolated

from two transformant cultures using the QIAprep Spin Miniprep kit (QIAGEN GmbH, Hilden, Germany) and used as template for sequence analysis with an ABI 3730XL DNA analyser (BaseClear, Leiden, The Netherlands) using insert-flanking T7 and SP6 promoter-targeted primers (Promega). In order to merge the sequences of the two clones and the different primers and to identify sequencing errors in individual reads, sequences were aligned using tools available in the DNASTAR package (DNASTAR, Madison, WI, USA), and the alignment was corrected manually. An almost complete 16S rRNA gene sequence (1466 bp) of strain CRIB^T was obtained. The 16S rRNA gene sequence of strain CRIB^T and other members of the family *Peptostreptococcaceae* were aligned using the SINA aligner (http://www.arb-silva.de/aligner/) [60]. Phylogenetic trees were constructed using MEGA 6 software [75]. Clustering was determined with the neighbour-joining method and bootstrap values were calculated based on 1000 replications. Tree topology was also confirmed using maximum likelihood and maximum parsimony methods. For pairwise 16S rRNA gene nucleotide sequence alignments, the EzTaxon-e server (http://eztaxon-e.ezbiocloud.net/) [37] was used to generate the pairwise similarity values.

Pairwise nucleotide sequence similarities of the 16S rRNA gene indicated that the closest known relatives of strain CRIB^T are *Clostridium lituseburense* (97.2%), *Clostridium glycolicum* (96.2%), *Clostridium mayombei* (96.2%), *Clostridium bartlettii* (96.0%), and *Clostridium irregulare* (95.5%). However, further phylogenetic analyses indicated that strain CRIB^T together with all the species listed above were all members of the family *Peptostreptococcaceae* Ezaki 2010 [15] (Figure 1 and see Figures S1 and S2 in the supplemental material). It is recognized that species belonging to the genus *Clostridium* form a large and taxonomically heterogeneous group. Many species were originally assigned to this genus on the basis of a relatively small set of phenotypic characteristics such as Gram-positive staining, rod-shape, anaerobic growth and the ability to form spores [23, 62]. However, this has resulted in the expansion of the genus *Clostridium* to more than 200 species, forming a very heterogeneous taxon that is long overdue for taxonomic re-arrangements [9, 62, 71, 87, 89]. This is clearly reflected by the fact that many of the current species of the genus *Clostridium* show less than 90% 16S rRNA gene sequence similarity to the type strain of the type species of the genus, *Clostridium butyricum* [9, 62]. Strain CRIB^T and its closest known relatives also have low 16S rRNA gene sequence similarities (<85%) to the type strain of *C. butyricum*. It has been well-known for decades that there are problems with the taxonomic structure of the genus *Clostridium*, and important steps have already been undertaken to clarify some of these issues. Collins and colleagues have proposed the grouping of the presently known species of the genus *Clostridium* into a number of distinct phylogenetic clusters based on a large-scale comparison of 16S rRNA gene sequences [9]. These clusters also include species from a range of other genera, indicating that the taxonomy of the genus *Clostridium* is not unambiguous. It

has been proposed that only species classified into cluster I are true members of the genus *Clostridium* (*Clostridium sensu stricto*) [87]. However, this is problematic since this also includes the genus *Sarcina*, and this name would have priority over the genus name *Clostridium* if all members of cluster I were to be placed in a single genus [88]. Some of the species previously classified in the genus *Clostridium* have been transferred to novel genera within the family *Clostridiaceae* and to novel families within the *Firmicutes* [86]. However, many species of the genus *Clostridium* still remain misclassified if one considers that the genus *Clostridium* should be restricted as proposed by Collins *et al.* (1994). This also holds true for members of *Clostridium* cluster XI [9], including the closest known relatives of strain CRIB^T, which although formally not currently assigned to the family *Peptostreptococcaceae* Ezaki 2010 [14] group with the members of this family based on the 16S rRNA gene sequence analysis (Figure 1); the classification of which will have to be further revised in the future.

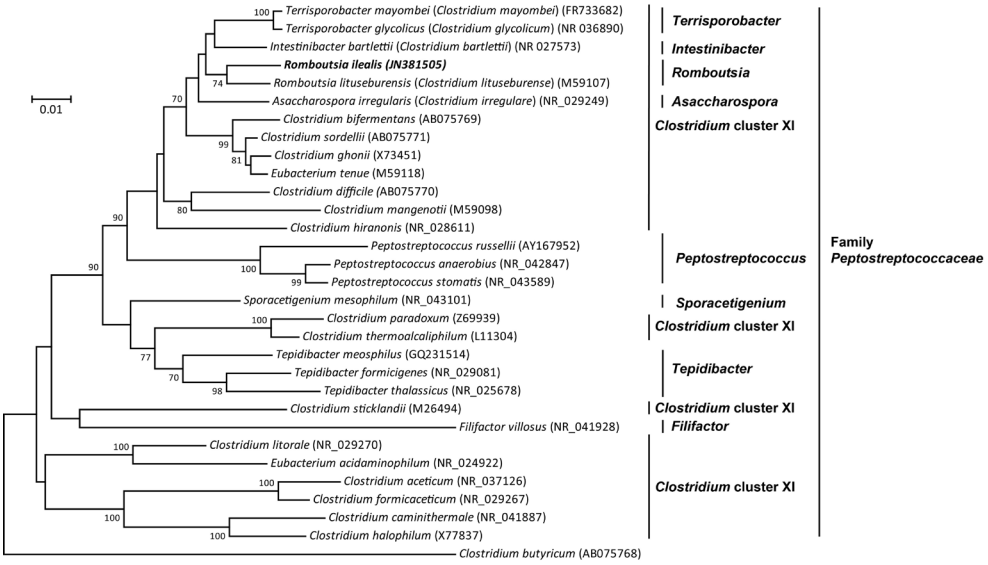

Figure 1. Neighbour-joining tree based on 16S rRNA gene sequence data, showing the phylogenetic position of strain CRIB^T and other (misclassified) members of the family *Peptostreptococcaceae.* The 16S rRNA gene sequences were aligned using the SINA aligner (http://www.arb-silva.de/aligner/). The phylogenetic tree was constructed using MEGA 6 software [75] with Kimura's two-parameter model as substitution model. Only bootstrap values ≥70% are shown at branch nodes. Bar, 1% sequence divergence. GenBank accession numbers are given in parentheses. The 16S rRNA gene sequence from *Clostridium butyricum* ATCC 19398^T (accession nr. AB075768), the type strain of the type species of the genus *Clostridium,* was used as an outgroup.

In order to substantiate that strain CRIB^T represents a novel species, the closest known relatives of strain CRIB^T defined by >95% 16S rRNA gene sequence similarity, were included for further taxonomic characterization. Strain CRIB^T and the type strains of five species of the genus *Clostridium* (*C. lituseburense* DSM 797^T, *C. bartlettii* DSM 16795^T, *C. glycolicum*

DSM 1288^T, C. *mayombei* DSM 6539^T, and *C. irregulare* DSM 2635^T) were also compared to the type strain of the type species of the type genus of the family *Peptostreptococcaceae*, *Peptostreptococcus anaerobius* DSM 2949^T, in order to investigate whether these species should be included in the genus *Peptostreptococcus*. This study represents a step further in the clarification of the taxonomy at the genus level of a small number of members of the class *Clostridia*.

For further support the description of strain CRIB^T as representative of a novel species, genotypic characterization was carried out by the Leibniz-Institut DSMZ (Deutsche SammLung vor Mikroorganismen und Zellkulturen, Braunschweig, Germany). Cells of strain CRIB^T were cultured in liquid CRIB medium for 15 h at 37°C for analysis of the genomic DNA G+C content and for DNA-DNA hybridizations (DDH). For DDH experiments, the type strains of three phylogenetically related and presumably misclassified species of the genus *Clostridium* [*C. lituseburense* (*Romboutsia lituseburensis* comb. nov.) DSM 797^T, *C. bartlettii* (*Intestinibacter bartlettii* comb. nov.) DSM 16795^T, and *C. irregulare* (*Asaccharospora irregularis* comb. nov.) DSM 2635^T] were obtained from the Leibniz-Institut DSMZ and cultured in the media suggested in the Leibniz-Institut DSMZ online catalogue (www.dsmz.de). Cells were disrupted by using a French pressure cell (Thermo Spectronic), and the DNA in the crude cell lysate was purified by chromatography on hydroxyapatite as described by Cashion *et al.* [6]. The genomic DNA G+C content was determined by HPLC according to the method of Mesbah *et al.* [49]. The genomic DNA G+C content of strain CRIB^T was determined to be 28.1 mol%, which demonstrates strain CRIB^T is a bacterium with a low G+C content. DDH experiments were carried out at 61°C by the liquid renaturation method as described previously [10] and as modified by Huss *et al.* [28]. Strain CRIB^T showed low DNA-DNA relatedness (mean percent reassociation ±SD, n=2) to the following type strains: 15.5±0.8% with *C. lituseburense* (*R. lituseburensis* comb. nov.) DSM 797^T, 20.4±3.5% with *C. bartlettii* (*I. bartlettii* comb. nov.) DSM 16795^T, and 18.1±1.3% with *C. irregulare* (*A. irregularis* comb. nov.) DSM 2635^T. These reassociation values were well below the cut-off point of 70% for species delineation that was recommended by Wayne *et al.* [83] and thereby confirm that strain CRIB^T represents a novel species. This is also consistent with the 16S rRNA gene sequence similarity values where more recent work has indicated that a 16S rRNA gene sequence similarity of 98.2-99.0% may be a more appropriate threshold above which DDH experiments should be carried out than the previously used value of 97% [48, 70].

Cell morphology of strain CRIB^T was examined using a phase-contrast microscope (DM2000, Leica Microsystems, Rijswijk, The Netherlands) at 1000× magnification, with cells grown for 24 h or 48 h at 37°C in liquid CRIB or PYG medium (see Figures S3a and S3b in the supplemental material). Gram staining was performed using a 4-step Gram stain kit (BD)

on cells from active cultures. Endospore staining was performed according to Schaeffer-Fulton's method [68] using 5% (w/v) malachite green and 2% safranin as counterstain. Stained cells were examined using a light microscope at 1000× magnification (Primo Star, Carl Zeiss MicroImaging GmbH, Jena, Germany). Sporulating cells that were visualized with the malachite green endospore staining were seen occasionally after prolonged incubation (see Figure S4 in the supplemental material). However, no free spores were observed, and no viable cells could be recovered from sporulating cultures exposed to 80°C for 10 min. The use of media that have previously demonstrated to be able to support the sporulation of species belonging to the class *Clostridia* [4, 13, 27, 65] did not result in an increased induction of sporulation in strain CRIB^T. Although the process of sporulation appears to be initiated in strain CRIB^T, it was not completed under the experimental conditions examined since free mature spores were not observed. This was distinct from *C. lituseburense* (*R.* lituseburensis comb. nov.) DSM 797^T that was able to sporulate in most of the media used and high numbers of free mature spores were observed (Figure S4).

Transmission electron microscopy was carried out as previously described by van Niftrik *et al.* [81]. Cells of strain CRIB^T were cryofixed by high-pressure freezing (Leica EMHPF, Leica Microsystems, Vienna, Austria, now Wohlwend), and freeze-substituted in acetone containing 2% osmium tetroxide or 2% osmium tetroxide, 0.2% uranyl acetate and 1% water [82] in an automatic freeze-substitution unit (AFS, Leica Microsystems). The cells were then embedded in Epon resin [51], sectioned using an ultramicrotome (Reichert Ultracut E, Leica Microsystems) and post-stained with 20% uranyl acetate in 70% methanol for 4 min and Reynolds' lead citrate staining [63] for two min. The ultrathin (70 nm) sections were investigated at 60-120 kV in a transmission electron microscope (Tecnai12, FEI Company, Eindhoven, The Netherlands). Images were recorded using a charge-coupled device camera (MegaView II, Olympus, Münster, Germany) and the AnalySIS software (Olympus). Transmission electron micrographs confirmed the typical cell wall morphology of Gram-positive staining cells (see Figures S3c and S3d in the supplemental material). No flagella or other cell wall extensions were observed under these conditions with these methods.

Colony morphology was examined after 24 h and 72 h of growth at 37°C under anoxic conditions on solid PYG or CRIB medium or commercial chocolate agar supplemented with PolyViteX (PVX; bioMérieux, Marcy l'Etoile, France).

Cell motility of strain CRIB^T was examined by detection of turbidity throughout stab-inoculated tubes containing semi-solid PYG medium [43] after 72 h of growth at 37°C. In addition, motility of CRIB was compared to that of representatives of three closely related (misclassified) species of the genus *Clostridium*: *C. lituseburense* (*R. lituseburensis* comb.

nov.) DSM 797^T, *C. bartlettii* (*I. bartlettii* comb. nov.) DSM 16795^T, and *C. irregulare* (*A. irregularis* comb. nov.) DSM 2635^T. Cell motility was examined after 48 h of growth at 37°C in semi-solid SIM medium. Semi-solid SIM medium consisted of (per litre distilled water): 20 g tryptone, 6.1 g peptone, 0.2 g ferrous ammonium sulfate, 0.2 g sodium thiosulfate, and 0.4% (w/v) gelrite. Both in semi-solid PYG and SIM medium, growth of strain CRIB^T was only observed in a distinct zone directly along the stab, indicating that CRIB is non-motile. This was distinct from *C. lituseburense* (*R .lituseburensis* comb. nov.) DSM 797^T that was demonstrated to be motile with this method, since growth was observed extending from the stab line (Table 1).

Table 1. Differential characteristics of strain CRIB^T and the type strains of the five species proposed to be reclassified in the novel genera *Romboutsia* gen. nov., *Intestinibacter* gen. nov., *Terrisporobacter* gen. nov. and *Asaccharospora* gen. nov., compared to those of the type species of the genus *Clostridium*, *C. butyricum*.

Characteristics	1	2	3	4	5	6	7
Cell size:							
width (µm)	1.0 - 2.0	1.4 - 1.7[ǂ,a]	1.0 - 1.5[b]	0.3 - 1.3[c]	1[d]	0.8 - 1.6[a]	0.5 – 1.7[a]
length (µm)	1.0 - 5.3	3.1 - 6.3[a]	5.0 - 50[b]	1.8 - 15.4[c]	2 - 6[d]	3.5 - 12.6[a]	2.4 – 7.6[a]
Motility	-	+[*]	-[*]	+[e]	+[d]	+[*]	+/-[a]
Bile resistance (20%)	-	-[a]	+[b]	-[e]	ND	-[a]	ND
Gelatin hydrolysis	-	+[*]	-[*]	-[e]	ND	+[*]	-/w[f]
Acid produced from (API systems):							
D-arabinose	w	-[*]	-	-[e]	-[d]	-	+/-[a]
D-fructose	-[†]	+[*]	+[*]	+[c,e]	+[d]	-	+[a]
L-fucose	+	-	-	ND	-[d]	-	ND
D-galactose	w	-[*]	-	-[c,e]	-[d]	-	+[a]
Gentiobiose	-	-	+[*]	ND	ND	-	ND
D-glucose	+	+	+[*]	+[c,e]	+[d]	-	+[f]
D-maltose	-[†]	+[*]	+[*]	+[c,e]	+[d]	-	+[a]
D-mannitol	-	-[*]	+[*]	-[c,e]	-[d]	-	-/+[a]
D-raffinose	+	-[*]	-	-[c,e]	-[d]	-	+[a]
D-ribose	-	+/-[*]	+[*]	-[c]	-[d]	-	+[a]
D-sorbitol	-	-[*]	+	+[c,e]	+[d]	-	-[a]
D-sucrose	+	+[*]	+[*]	-[c,e]	-[d]	-	+[a]
D-tagatose	-	-	+	ND	ND	-	ND
Xylose	-	-[*]	-[*]	+[c,e]	+[d]	-	+[a]
H_2 production in PYG	+	w[*]	w	ND	ND	-[*]	+[a]
Products from PYG[§]	A, F, l (p)	A, F, l, p (b, iv)[*,‖]	A, F, iv, l (ib)[*,¶]	A, iv, ib (p, f, l)[c]	A, iv (p, ib)[d]	A, iv, f, l, p[*]	B, A, F (l)[a]
Polar lipids[#]	6 GL, 4 PL, L	5 GL, 5 PL	7 GL, 4 PL, L	7 GL, 7 PL, L	6 GL, 6 PL	4 GL, 4 PL	**
Predominant cellular fatty acids	$C_{16:0}$	$C_{16:0}$ $C_{17:0}$ $C_{18:1}$ ω7c	iso-$C_{15:0}$ $C_{16:0}$ DMA	$C_{16:0}$	$C_{16:0}$	$C_{16:0}$ $C_{16:1}$ ω7c	$C_{16:0}$[f] $C_{18:0}$[f]
DNA G+C content (mol%)	28.1	27[a]	29.8[b]	29[c,e]	25.6[d]	ND	27-28[a]

Strains: 1, *Romboutsia ilealis* gen. nov., sp. nov. CRIB^T; 2, *C. lituseburense* (*R. lituseburensis* comb. nov.) DSM 797^T; 3, *C. bartlettii* (*I. bartlettii* comb. nov.) DSM 16795^T; 4, *C. glycolicum* (*T. glycolicus* comb. nov.) DSM 1288^T; 5, *C. mayombei* (*T. mayombei* comb. nov.) DSM 6539^T; 6, *C. irregulare* (*A. irregularis* comb. nov.) DSM 2635^T; 7, *C. butyricum* DSM 10702^T. +, positive; w, moderately positive; -, negative; ND, no data available. For strain 1, 2, 5 and 6 the data were obtained in this study, unless indicated otherwise. With the API 50 CH and API 20 A systems (bioMérieux), strain 1, 2, 5 and 6 were negative for D-adonitol, amidon (starch), amygdalin, L-arabinose, D-arabitol, L-arabitol, arbutin, catalase activity, D-cellobiose, dulcitol, erythritol. D-fucose, glycerol, glycogen, indole formation, inositol, inulin†, D-lactose†, D-lyxose, D-mannose, D-melezitose, D-melibiose†, methyl-α-D-glucopyranoside, methyl-α-D-mannopyranoside, methyl-ß-D-xylopyranoside, potassium 2-ketogluconate, potassium 5-ketogluconate, potassium gluconate, L-rhamnose, salicin, L-sorbose, D-trehalose, D-turanose, urease activity, xylitol, D-xylose and L-xylose.

* Result the same as that reported in the literature.

† Compared to the absence of acid production in the API systems, there was weak growth of strain CRIB^T on this substrate when grown in liquid PY medium.

‡ Data were taken from: *a*, Rainey et al. [62]; b, Song *et al.* [69]; c, Chamkha *et al.* [7]; d, Kane *et al.* [34]; e, Gaston and Stadtman [19]; f, Biebl and Spröer [2]

§ Products (listed in the order usually detected); A, acetic acid; B, butyric acid; F, formic acid; IB, iso-butyric acid, IV, iso-valeric acid, L, lactic acid; P; propionic acid. Upper case letters indicate major components, lower case minor components (< than 20% of total measured metabolic end product production). Products in parentheses are not detected uniformly.

|| Butyric acid and iso-valeric acid are produced in the presence of casamino acids.

¶ Iso-butyric acid is produced in the presence of casamino acids and absence of glucose.

GL, Glycolipid; PL, phospholipid; L, lipid.

** The polar lipids of *C. butyricum* comprise phosphatidylglycerol, diradyl (i.e. diacyl glycerol and 1-*O*-alk-1′-enyl- 2-acyl glycerol) derivatives of phosphatidylglycerol, diphosphatidylglycerol, phosphatidylethanolamine, a glycerol acetal derivative of the plasmalogen form of phosphatidylethanolamine and a phosphatidylglycerol acetal derivative of the plasmalogen form of phosphatidylethanolamine [22, 32].

Substrate utilization properties of strain CRIB^T were compared to those of three closely related (misclassified) species of the genus *Clostridium* using the API 50 CH and API 20 A systems (bioMérieux) according to the manufacturer's instructions except that liquid PY medium was used for inoculation and the strips were incubated anoxically. *C. lituseburense* (*R. lituseburensis* comb. nov.) DSM 797^T, *C. bartlettii* (*I. bartlettii* comb. nov.) DSM 16795^T, and *C. irregulare* (*A. irregularis* comb. nov.) DSM 2635^T were used for comparison.

In addition to the API tests, substrate utilization of strain CRIB^T was confirmed by adding one of the following compounds to liquid PY medium in culture bottles to a final concentration of 0.5% (w/v) incubated under anoxic conditions: D-adonitol, D-arabinose, L-arabinose, D-arabitol, L-arabitol, D-cellobiose, erythritol, D-fructose, D-fucose, L-fucose, D-galactose, D-glucose, glycerol, glycogen, inositol, inulin, D-lactose, D-maltose, D-mannitol, D-mannose, D-melibiose, methyl-α-D-glucopyranoside, methyl-α-D-mannopyranoside, D-raffinose, L-rhamnose, D-ribose, D-sucrose, D-sorbitol, L-sorbose, D-trehalose, D-turanose, xylitol and L-xylose. In addition to these substrates, which were all also previously examined in the API 50 CH and API 20 A systems, utilization of cellulose, soluble starch and D-xylose (final concentration of 0.5% (w/v) for all substrates) were also examined. Growth was determined spectrophotometrically by measuring optical density at 600 nm. Growth on a substrate was defined by comparing the increase in OD_{600nm} in liquid PY medium with additional substrate to the increase in OD_{600nm} in liquid PY medium

lacking additional substrate. A more than 2-fold higher increase in OD_{600nm} was considered to reflect good growth, a 1.5- to 2-fold higher increase was considered to reflect moderate growth, and a 1- to 1.5-fold higher increase than the control was considered to reflect weak growth. Acid formation was observed by measuring the pH of the media at regular intervals until 30 days of incubation. The carbohydrate utilization pattern of strain CRIB^T was clearly different from that of *C. lituseburense* (*R. lituseburensis* comb. nov.) DSM 797^T and the representatives of the other (misclassified) species of the genus *Clostridium* (Table 1). Strain CRIB^T is able to utilize D-arabinose, L-fucose, D-galactose and D-raffinose, which are all carbohydrates that the other type strains are not able to utilize.

The metabolic end products produced during growth in liquid PYG medium (including short chain fatty acids) were analysed by HPLC using a Metacarb 67H column (Varian, Middelburg, The Netherlands). H_2 production was measured by GC using a Shimadzu GC 14B (Shimadzu Scientific Instruments Inc., Kyoto, Japan) fitted with a Molsieve 13x column (Varian). CO_2 production was measured by GC using a Shimadzu GC 2014 (Shimadzu Scientific Instruments Inc.) fitted with a CP-PoraPLOT Q column (Varian). The metabolic end products of strain CRIB^T were compared to those of the three closely related and presumably misclassified species of the genus *Clostridium*. Acetate was the major end product formed in all species. However, compared to *C. lituseburense* (*R. lituseburensis* comb. nov.) DSM 797^T and the other type strains of (misclassified) species of the genus *Clostridium*, strain CRIB^T did not produce butyrate, iso-butyrate, or iso-valerate.

Growth characteristics of strain CRIB^T were determined at various temperatures (20–50°C, in increments of 5°C) and pH values (5.0–9.0, in increments of 0.5 pH units). Tolerance to NaCl was tested at different salt concentrations [0-3% (w/v), in increments of 0.5%, and 1–7% (w/v), in increments of 1%]. Tolerance to bile salts (Difco™ Oxgall, BD) was tested at different bile salt concentrations [0-25% (v/v), in increments of 5%]. The growth of strain CRIB^T in the presence of Tween-80 was examined using different Tween-80 concentrations [0-0.04% (v/v), in increments of 0.01%]. The influence of shaking during incubation of liquid cultures on the growth of strain CRIB^T was determined by comparing the growth without or with shaking at 100 rpm. Growth was determined spectrophotometrically by measuring optical density at 600 nm. An increase in OD_{600nm} of >0.2 was considered to reflect growth.

The oxygen tolerance of strain CRIB^T was examined after 72 h of growth at 37°C in solid PYG medium supplemented with 0.05% (w/v) sodium thioglycolate.

Sulfate, thiosulfate and sulfite (30mM) were tested as electron acceptors in liquid PY and PYG medium.

MIC-values (µg/mL) of several antimicrobial agents (clindamycin, penicillin G and metronidazole) against strain CRIB^T were determined using Etest gradient strips (bioMérieux).

Detailed physiological characteristics of strain CRIB^T are provided in the species description of *Romboutsia ilealis* sp. nov.

Chemotaxonomic characterization of strain CRIB^T included analysis of peptidoglycan structure, cell wall sugars, and quinone composition and was carried out by the Leibniz-Institut DSMZ on cells of strain CRIB^T cultured in liquid CRIB medium for 15 h at 37°C.

The peptidoglycan of strain CRIB^T was isolated and purified from wet cell biomass according to the methods previously described by Schleifer [66]. Peptidoglycan preparations were obtained after disruption of cells by shaking with glass beads and subsequent trypsin digestion. The peptidoglycans were hydrolysed in 4N HCl (16 h at 100°C). The amino acids and peptides in the cell-wall hydrolysates were analysed by two-dimensional TLC on cellulose plates using a previously described solvent system [64]. After derivatization [46], the molar ratios of amino acids were determined by GC. The *N*-heptafluorobutyryl amino acid isobutyl esters obtained by derivatization were subjected to GC-MS to determine the identity of the components not identified by previous analyses.

For cell wall sugar analysis the peptidoglycan of strain CRIB^T was hydrolysed in 1N H_2SO_4 (2 h at 100°C). H_2SO_4 was removed by adding 20% *N*,*N*-dioctylmethylamine in chloroform according to Whiton *et al.* [85] and the sugars in the hydrolysate were analysed by TLC on cellulose plates according to Staneck & Roberts [73].

Respiratory lipoquinones were extracted from lyophilized biomass of strain CRIB^T using the two-stage method described by Tindall [78, 79] using methanol/hexane, followed by phase separation into hexane. UV-absorbing bands were removed from the plate and further analysed by HPLC.

To support the (re)classification into novel genera, analysis of the polar lipid profiles and cellular fatty acid composition was done for the following strains: *Romboutsia ilealis* gen. nov., sp. nov. CRIB^T, *C. lituseburense* (*R. lituseburensis* comb. nov.) DSM 797^T, *C. bartlettii* (*I. bartlettii*) DSM 16795^T, *C. glycolicum* (*T. glycolicus* comb. nov.) DSM 1288^T, *C. mayombei* (*T. mayombei* comb. nov.) DSM 6539^T, *C. irregulare* (*A. irregularis* comb. nov.) DSM 2635^T, and for comparison *P. anaerobius* DSM 2949^T. Cultivation of the strains and subsequent analyses were performed by the Leibniz-Institut DSMZ. All strains were grown under identical conditions in liquid DSM medium 104b. This medium consisted of (per litre

distilled water): 5 g trypticase peptone, 5 g peptone from meat (pepsin-digested), 10 g yeast extract, 5 glucose, 1 mg resazurin, 40 mL salts solution [27] and 0.5 g L-cysteine hydrochloride. Cells were harvested in mid-exponential to end-exponential phase. Polar lipids were extracted from lyophilized biomass using the two-stage method described by Tindall [78, 79] and separated by two-dimensional silica gel TLC. All polar lipids were detected by spraying the plates with 5% ethanolic molybdophosphoric acid followed by heating, while head groups were detected using specific staining reagents as described previously Tindall [78, 79]. Since chromatography conditions were identical, comparison of the TLC plates was possibly and spots were labelled according to their staining behaviour and R_f value. For determination of cellular fatty acid composition, fatty acid methyl esters were obtained from fresh cells by saponification, methylation and extraction using minor modifications of previously described methods [40, 50]. Fatty acid methyl esters mixtures were separated by GC and analysed using the Sherlock Microbial Identification System (MIS) as previously described [1]. Peaks were automatically integrated, and fatty acid identification (using the MOORE6 peak naming database) and relative concentrations were calculated. In the case of anaerobes that synthesize plasmalogens, cleavage of these compounds gives rise to the corresponding aldehydes and dimethyl acetal derivatives [17, 18, 52].

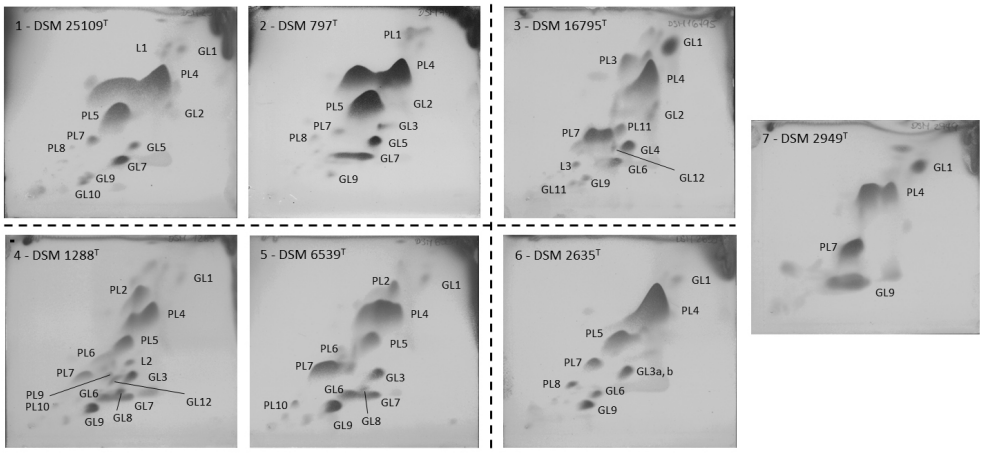

Figure 2. Two-dimensional thin-layer chromatograms showing total polar lipid profiles of strain CRIB^T and the type strains of the five species proposed to be reclassified in the novel genera *Romboutsia* gen. nov., *Intestinibacter* gen. nov., *Terrisporobacter* gen. nov. and *Asaccharospora* gen. nov. Profiles are compared to the polar lipid profile derived from the type strain of the type species of the type genus of the family *Peptostreptococcaceae*, *P. anaerobius* DSM 2949^T. Strains: 1, *Romboutsia ilealis* gen. nov., sp. nov. CRIB^T; 2, *C. lituseburense* (*R. lituseburensis* comb. nov.) DSM 797^T; 3, *C. bartlettii* (*I. bartlettii* comb. nov.) DSM 16795^T; 4, *C. glycolicum* (*T. glycolicus* comb. nov.) DSM 1288^T; 5, *C. mayombei* (*T. mayombei* comb. nov.) DSM 6539^T; 6, *C. irregulare* (*A. irregularis* comb. nov.) DSM 2635^T; 7, *P .anaerobius* DSM 2949^T. Separation of the polar lipids was performed by two dimensional silica gel TLC, with chloroform : methanol : water (65 : 25 : 4, by vol) in the first direction, and chloroform : methanol : acetic acid : water (80 : 12 : 15 : 4, by vol) in the second direction. Plates were stained with 5% ethanolic molybdophosphoric acid to show all lipids. Polar lipids were labelled according to their staining behaviour and their R_f values. GL, glycolipid; PL, phospholipid; L, lipid. PL4 is most probably identified as diphosphatidylglycerol and PL7 as phosphatidylglycerol.

Detailed results of the polar lipid and cellular fatty acid analyses can be found in Figure 2 and Table 2, respectively.

The polar lipid profiles of all species examined were dominated by glycolipids and phospholipids (Figure 2). In some species, lipids were also detected that did not react with any of the specific spray reagents used. The polar lipid profiles of strain CRIB^T and the type strains of the five (misclassified) species of the genus *Clostridium* all clearly differed from that of *P. anaerobius,* the type species of the type genus of the family *Peptostreptococcaceae*. Based on published information on a limited number of species within *Clostridium* cluster I (including members of the type species, *C. butyricum*) the polar lipid composition is clearly different from that of the strains examined here. In particular the wide diversity of glycolipids documented here is not present in those strains assigned to *Clostridium* cluster I. In addition, the presence of phosphatidylethanolamine in all the strains of *Clostridium* cluster I examined in previous studies [22, 24-26, 30-32, 54, 76, 77] clearly distinguishes them from the strains examined here that do not synthesize phosphatidylethanolamine, which was supported by the absence of phospholipids with free amino groups in the polar lipid analyses. Care should, however, be taken with regard to the conclusions drawn in some of these publications because links are made to some strains that in the past were incorrectly identified. The potential to synthesize phosphatidylethanolamine was further investigated by examining the genomes available for the presence of *psd* genes encoding the enzyme (phosphatidylserine decarboxylase) responsible for the formation of phosphatidylethanolamine via decarboxylation of phosphatidylserine [3, 36]. The protein sequence ABK62661 extracted from the genome of *Clostridium novyi* strain NT (CP000382) was used to 'seed' a BLAST search of the available genomes of members of the family *Peptostreptococcaceae* held on the Joint Genome Institute (JGI) Integrated Microbial Genomes (IMG) website (http://img.jgi.doe.gov/) on 10 September 2013. No matches were found in any of the genomes, indicating that homologues of this enzyme do not appear to be present in the genomes, which would also explain the absence of the corresponding polar lipid (phosphatidylethanolamine). Based on currently available data it would appear that all members of *Clostridium* cluster I [9, 29] produce phosphatidylethanolamine, whereas none of the strains examined here produce this phospholipid.

With respect to the cellular fatty acids, the profiles of strain CRIB^T and the type strains of the five (misclassified) species of the genus *Clostridium* were compared to the type species of the type genus of the family *Peptostreptococcaceae*, *P. anaerobius* DSM 2949^T. Other studies have reported cellular fatty acid profiles of *P. anaerobius* and other species belonging to the genus *Peptostreptococcus* [16, 84]. However, since there are large discrepancies in the data reported in these studies and because results were carried

out under different growth conditions, we performed an additional analysis of the cellular fatty acid profile of *P. anaerobius*. The profile of *P. anaerobius* presented in this study was determined under identical growth conditions as strain CRIB^T and the type strains of the five (misclassified) species of the genus *Clostridium*. We found that the cellular fatty acid profile of *P. anaerobius* is characterized by an almost complete absence of unsaturated straight chain fatty acids, which is clearly distinct from the cellular fatty acid profiles of strain CRIB^T and the type strains of the five (misclassified) species of the genus *Clostridium* (Table 2).

Table 2. Cellular fatty acid profiles of strain CRIB^T and the type strains of the five species proposed to be reclassified in the novel genera *Romboutsia* gen. nov., *Intestinibacter* gen. nov., *Terrisporobacter* gen. nov. and *Asaccharospora* gen. nov. Profiles are compared to the cellular fatty acid profile derived from the type strain of the type species of the type genus of the family *Peptostreptococcaceae*, *Peptostreptococcus anaerobius*.

Fatty acid	1	2	3	4	5	6	7
Saturated straight chains							
$C_{12:0}$	1.56	0.55	0.21	1.15	1.28	1.46	2.21
$C_{14:0}$	3.83	1.54	1.09	2.05	3.44	4.02	6.67
$C_{15:0}$	6.23	5.55	1.29	3.17	3.18	2.21	0.23
$C_{16:0}$	**31.34**	**10.72**	4.16	**15.61**	**29.38**	**26.94**	**9.62**
$C_{17:0}$	9.96	**16.81**	2.00	7.45	5.83	1.41	-
$C_{18:0}$	6.58	4.64	1.04	1.56	2.53	2.29	0.29
Unsaturated straight chains							
$C_{16:1}$ ω5c	-	-	-	0.69	0.51	1.50	-
$C_{16:1}$ ω7c	9.12	4.14	2.82	6.89	6.42	**17.59**	0.43
$C_{16:1}$ ω9c	6.45	4.04	1.94	5.09	6.40	7.95	-
$C_{17:1}$ ω7c	3.26	7.07	1.23	3.85	2.27	0.70	-
$C_{18:1}$ ω9c	3.41	8.10	1.01	1.36	1.48	2.49	0.15
$C_{19:1}$ ω12c	-	2.05	-	-	-	-	-
Branched chains							
$C_{12:0}$ ISO	-	-	0.18	0.43	0.71	2.03	4.71
$C_{13:0}$ ANTEISO	-	-	0.35	0.27	0.32	2.36	8.45
$C_{14:0}$ ISO	-	-	3.23	1.96	2.09	0.74	**11.16**
$C_{15:0}$ ISO	-	-	**12.19**	2.25	0.98	2.51	0.50
$C_{15:0}$ ANTEISO	-	-	7.69	1.24	1.04	1.69	7.79
$C_{16:0}$ ISO	-	-	4.66	3.16	3.04	5.33	6.46
$C_{17:0}$ ISO	-	-	3.72	1.52	0.71	4.19	0.20
$C_{17:0}$ ANTEISO	-	-	3.40	0.97	0.76	3.02	0.66
Dimethyl acetals (DMA) + corresponding aldehydes							
$C_{14:0}$ DMA	-	-	0.88	0.54	0.52	-	3.46
$C_{15:0}$ ISO DMA	-	-	3.53	0.28	-	-	0.20
$C_{16:0}$ ALDE	-	-	1.95	7.06	1.34	-	3.93
$C_{16:0}$ DMA	-	-	**10.80**	9.30	8.24	0.61	9.79
(Summed $C_{16:0}$ ALDE + DMA*)	-	-	(12.75)	(16.36)	(9.58)	(0.61)	(13.72)
$C_{17:0}$ DMA	-	-	2.65	2.09	1.18	0.23	0.03
$C_{17:0}$ ANTEISO DMA	-	-	4.45	0.32	-	-	0.29
Hydroxyl							
$C_{13:0}$ ISO 3OH	-	-	-	-	-	-	3.50

Fatty acid	1	2	3	4	5	6	7
Cyclopropane							
C_{19} cyclopropane 9,10	0.84	1.62	-	-	-	-	-
Unknown							
Unknown 7.87	-	-	-	-	-	-	7.87
Unknown 16.107	-	-	6.06	1.14	0.71	-	3.74
Unknown 17.103	-	-	6.74	0.43	-	-	0.07
Summed features†							
Summed feature 4‡		(1.00/0.55)		(1.15/0.56)			
$C_{15:1}$ ω8c	1.09	1.55	0.50	1.71	1.13	0.34	-
Summed feature 5§							
$C_{15:0}$ DMA	-	-	2.50	2.71	1.77	-	0.19
Summed feature 7‖	-	7.93	-	0.83	-	-	-
Summed feature 8¶	5.28	8.93	1.82	3.06	2.51	1.08	-
Summed feature 10#							
$C_{18:1}$ ω7c	8.34	**10.66**	1.05	2.86	3.98	2.49	-
Summed feature 13**	-	-	0.24	-	-	-	3.20
Total	97.29	95.90	95.38	93.00	93.75	95.18	95.80

Strains: 1, *Romboutsia ilealis* gen. nov., sp. nov. CRIB^T; 2, *C. lituseburense* (*R. lituseburensis* comb. nov.) DSM 797^T; 3, *C. bartlettii* (*I. bartlettii* comb. nov.) DSM 16795^T; 4, *C. glycolicum* (*T. glycolicus* comb. nov.) DSM 1288^T; 5, *C. mayombei* (*T. mayombei* comb. nov.) DSM 6539^T; 6, *C. irregulare* (*A. irregularis* comb. nov.) DSM 2635^T; 7, *P. anaerobius* DSM 2949^T. Fatty acid methyl esters (and other components that included aldehyde and dimethyl acetal cleavage products of plasmalogen containing lipids) were separated by GC and detected by flame ionisation, using the MIDI Sherlock Microbial Identification System (MIS) and the Anaerobic Bacteria Library (MOORE6) for peak identification. Data are presented as percentages of the total fatty acid content. Only fatty acids with an abundance >1.5% in at least one of the strains, are shown. For each strain the predominant fatty acid(s) (≥10%) are indicated in bold type. All strains were grown in DSM medium 104b at 37°C and cells were harvested in mid-exponential to end-exponential phase. DSM medium 104b does not contain either Tween (80) or animal serum, both of which may be sources of 18:1ω9c in strains grown in the presence of these materials. Dimethyl acetals and aldehydes are cleavage products of plasmalogens [17, 18, 52]. In the absence of experimental data on the effect of the hydrolysis/methylation conditions used in the MIDI system it is not clear whether the presence of both compounds with the same chain lengths is due to incomplete conversion of plasmalogens to the corresponding dimethyl acetals, or an equilibrium in the ratio of aldehydes/dimethyl acetals released from the plasmalogens. Consequently where both the aldehyde and dimethyl acetal of the same chain length occur together they have also been listed as the sum of the aldehydes/dimethyl acetals. It should be also noted that while plasmalogens appear to be synthesized by an oxygen-dependent pathway in eukaryotes they appear to be synthesized in anaerobic prokaryotes by an oxygen-independent mechanism [24].

† Summed features represent groups of two or more fatty acids that are grouped together for the purpose of evaluation by the MIDI system. In some cases peaks may be identified on the basis of their separate equivalent chain length (ECL) while in other cases the ECLs are almost identical, making an unambiguous identification difficult.

‡ Listed as summed feature 4 comprising $C_{15:1}$ ω8*c*, $C_{15:2}$ and/or an unknown $C_{15:2}$, however the ECL indicates that the main peak is $C_{15:1}$ ω8*c*.

§ Listed as summed feature 5 comprising $C_{15:0}$ DMA and/or $C_{14:0}$ 3OH, however the ECL indicates that the main peak is $C_{15:0}$ DMA.

|| Listed as summed feature 7 comprises $C_{17:1}$ ω9*c* and/or $C_{17:2,}$ the ECLs are sufficiently close together making an unambiguous identification difficult, however biochemical considerations suggest that the synthesis of a $C_{17:2}$ may require oxygen.

¶ Listed as summed feature 8 comprising $C_{17:1}$ ω8*c* and/or $C_{17:2}$, the ECLs are sufficiently close together making an unambiguous identification difficult, however biochemical considerations suggest that the synthesis of a $C_{17:2}$ may require oxygen.

Listed as summed feature 10 comprising $C_{18:1}$ ω7*c* and/or an unknown fatty acid, however the ECL indicates that the main peak is $C_{18:1}$ ω7*c*.

** Listed as summed feature 13 comprising $C_{14:0}$ 2OH and/or $C_{15:0}$ ANTEISO DMA, the ECLs are sufficiently close together making an unambiguous identification difficult; however, biochemical considerations suggest that the synthesis of a 2-OH fatty acid may require oxygen.

Based on the polar lipid and cellular fatty acid profiles we can conclude that strain CRIB^T and the type strains of the five (misclassified) species of the genus *Clostridium* do not belong to either the genus *Peptostreptococcus* or the genus *Clostridium* (i.e. members of cluster I as defined by Collins *et al.* [9]) and should therefore be transferred to novel genera. Based on a combination of polar lipid and cellular fatty acid analyses, strain CRIB^T and the type strains of the five (misclassified) species of the genus *Clostridium* could be divided into four groups, which we propose should be reclassified into four novel genera. The first group to be reclassified as a novel genus, and for which we propose the name *Romboutsia* gen. nov., consists of strain CRIB^T and *Romboutsia lituseburensis* comb. nov. (*C. lituseburense*) DSM 797^T. The genus *Romboutsia* is characterized by the predominance of straight-chain saturated and unsaturated fatty acids (mainly C_{16}-C_{17}) and the absence of branched-chain fatty acids, dimethyl acetals, and aldehydes. In addition, C_{19} cyclopropane 9,10 was detected in low abundance in both species within this group (Table 2). The second group, for which we propose the name *Intestinibacteri* gen. nov., comprises *Intestinibacter bartlettii* comb. nov. (*C. bartlettii*) DSM 16795^T. The genus *Intestinibacter* is, compared to the other proposed genera, characterized by a low abundance of C_{16} fatty acids. In turn, branched-chain saturated fatty acids (mainly C15) and dimethyl acetals (mainly C_{16}) predominate in this group. The third group, for which we propose the name *Terrisporobacter* gen. nov., consists of *Terrisporobacter glycolicus* comb. nov. (*C. glycolicum*) DSM 1288^T and *Terrisporobacter mayombei* comb. nov. (*C. mayombei*) DSM 6539^T. This group is characterized by the dominance of (straight and branched-chain) saturated and unsaturated fatty acids (mainly C_{16}-C_{17}). In addition, dimethyl acetals (mainly C_{16} and its corresponding aldehyde) predominate. The fourth group, for which the name *Asaccharospora* gen. nov. is proposed, comprises *Asaccharospora irregularis* comb. nov. (*C. irregulare*) DSM 2635^T. This group is characterized by the dominance of (straight and branched-chain) saturated and unsaturated fatty acids (mainly C_{16}). Dimethyl acetals and aldehydes are almost absent. Both the polar lipid profiles and the phylogenetic clustering based on the 16S rRNA gene sequence are consistent with the reclassification of strain CRIB^T and the type strains of the five (misclassified) species of the genus *Clostridium* into four groups that we propose are novel genera (Figure 1 and Figure 2).

In addition to large variability in phenotypic properties, many species belonging to *Clostridium* cluster XI show low 16S rRNA gene sequence similarity to one another (Figure 1). Based on these observations, it is very likely that the closely related and also presumably misclassified species *Clostridium bifermentans*, *Clostridium sordellii*, *Clostridium ghonii* and *Eubacterium tenue* will form a separate novel genus within the *Peptostreptococcaceae*. However, additional taxonomic characterization experiments (such as comparative cellular fatty acid and polar lipid analyses) are necessary to confirm this. Based on the data presented here the placement of *C. lituseburense*, *C. glycolicum*,

C. mayombei, *C. bartlettii* and *C. irregulare* in novel genera also has implications for the nomenclature of the important human pathogen *Clostridium difficile* where it would seem unavoidable that within the foreseeable future this species will also have to be placed in a novel genus. Drucker *et al.* [12] and Korachi *et al*. [39] reported on a single main phospholipid present in *C. difficile* (phosphatidylglycerol) and the absence of glycolipids, suggesting very clear differences from the taxa examined in this study. Yutin & Galperin have proposed that *C. difficile* be transferred to a novel genus, '*Peptoclostridium*' [89]. While they also suggest that *C. bartlettii*, *C. glycolicum*, *C. irregulare*, *C. lituseburense* and *C. mayombei* should be transferred to this genus, their paper seems to lack a formal proposal for the creation of new combinations. However, based on the data presented here there would be no inconsistency between their evaluation of the data and the inclusion of additional chemical data that we interpret as indicating the presence of several genera and not just one.

A similar situation arises with members of other taxa that share a high degree of 16S rRNA gene sequence similarity with *C. difficile*, including *Clostridium mangenottii* and *Clostridium hiranonis* that either are members of the same novel genus or may need to be placed in other novel genera. It is beyond the purpose of this article to reclassify all misclassified species of the genus *Clostridium*. Nevertheless, the transfer of the five (misclassified) species of the genus *Clostridium* as proposed here is another important step in resolving some of the taxonomic problems associated with the genus *Clostridium*.

Based on phenotypic and genetic characterization we conclude that strain CRIB^T a rod-shaped organism isolated from the gastro-intestinal tract of rats, clearly represents a novel species. In addition, we propose, based on phenotypic (including comparative cellular fatty acid and polar lipid analyses) and genetic considerations, that a number of (misclassified) species of the genus *Clostridium* be reclassified in four novel genera. We propose a novel genus, *Romboutsia* gen. nov., to include *Romboutsia lituseburensis comb. nov.* The novel isolate CRIB^T (=DSM 25109^T =NIZO 4048^T) is proposed as type strain of the type species, *Romboutsia ilealis* gen. nov., sp. nov. of the proposed novel genus. In addition, we propose a second novel genus, *Intestinibacter* gen. nov., to include *Intestinibacter bartlettii* comb. nov. (type species). Furthermore, we propose a third novel genus, *Terrisporobacter* gen. nov., to include *Terrisporobacter glycolicus* comb. nov. (type species) and *Terrisporobacter mayombei* comb. nov. We propose a fourth novel genus, *Asaccharospora*, to include *Asaccharospora irregularis* comb. nov. (type species).

Description of *Romboutsia* gen. nov.

Romboutsia (Rom.bout'si.a. N.L. fem. n. *Romboutsia* in honour of the Dutch microbiologist Frans M. Rombouts, in recognition of his contributions to food microbiology and probiotic research).

Members of the genus are anaerobic. Cells stain Gram-positive or Gram-variable. Cells are straight or slightly curved rods, which often form chains. All species are capable of spore formation, but free spore formation can be very sparse. The major end products of metabolism are acetate and formate. Cells contain both phospholipids and glycolipids. The major phospholipids are labelled PL4, PL5, PL7 and PL8 and the glycolipids comprise GL2, GL5, GL7 and GL9, the R_f values of which can be determined by reference to Figure 2 and also serve as reference points for future work on the elucidation of these structures. Phospholipids with free amino groups (e.g. phosphatidylethanolamine) are absent. The fatty acids present in the hydrolysis products of whole cells are straight chain saturated and unsaturated (Table 2). Dimethyl acetals (i.e. plasmalogens) are not present. The genomic DNA G+C content is 27-28 mol%. The type species is *Romboutsia ilealis*.

Description of *Romboutsia ilealis* sp. nov.

Romboutsia ilealis (i'le.a'lis N.L. fem. adj. *ilealis* pertaining to the ileum).

Cells are obligately anaerobic, non-motile, spore-forming rods. Typical cells are 1.0-2.0 µm × 1.0-5.3 µm in size and occur primarily in chains, however, single cells and pairs are observed as well. Cells stain Gram-positive; however, they stain Gram-negative as cultures reach stationary phase. Sporulation is seen occasionally after prolonged incubation. Growth occurs both on liquid and solid CRIB or PYG medium. Surface colonies on solid PYG and CRIB medium, and PVX agar plates, incubated under anoxic conditions for 24 h, are white or light grey, circular, 0.5-1 mm in diameter with a shiny and smooth surface, a moist texture, an entire margin and the elevation is flat to raised. After incubation for 72 h the colonies are larger (2-4 mm) and have translucent and undulate margins. Temperature range for growth is 30-45°C with an optimum temperature of 37°C. The pH range for growth is 6.5-8.0 with an optimum pH of 7.0-7.5. Growth occurs at NaCl concentrations of 0-1% (w/v). Growth is inhibited by 5-20% bile salts. Growth is stimulated by addition of 0.01% Tween 80 (v/v) to liquid PYG medium and shaking of the cultures at 100 rpm. Cultures in liquid PYG medium are turbid with a smooth sediment and a pH of 5.5-6.0 after incubation for a week. Abundant gas is produced in PYG deep agar cultures. Strain CRIB^T is able to grow mainly on carbohydrates. Weak growth is observed on yeast extract and bacterial peptone as the sole carbon source. In the API 50 CH and API 20 A systems, acid is produced from D-arabinose, L-fucose, D-galactose, D-glucose, D-raffinose and D-sucrose. When grown in liquid PY medium, weak growth is observed on D-fructose, inulin,

D-lactose, D-maltose and D-melibiose, in addition to moderate growth on D-arabinose and D-galactose and good growth on L-fucose, D-glucose, D-raffinose and D-sucrose. No growth is observed on N-acetylglucosamine, D-adonitol, amidon, amygdalin, L-arabinose, D-arabitol, L-arabitol, arbutin, D-cellobiose, cellulose, dulcitol, erythritol, D-fructose, D-fucose, gentiobiose, glycerol, glycogen, inositol, inulin, D-lactose, D-lyxose, D-maltose, D-mannitol, D-mannose, D-melezitose, D-melibiose, methyl-α-D-glucopyranoside, methyl-α-D-mannopyranoside, methyl-ß-D-xylopyranoside, potassium 2-ketogluconate, potassium 5-ketogluconate, potassium gluconate, L-rhamnose, D-ribose, salicin, soluble starch, D-sorbitol, L-sorbose, D-tagatose, D-trehalose, D-turanose, xylitol, D-xylose and L-xylose. The strain is negative for indole production, urease and catalase activity, gelatin and starch hydrolysis. Metabolic products produced from PYG are acetate, formate and lactate. Abundant H_2 and CO_2 are produced. Sulfite is reduced. Sulfate and thiosulfate are not reduced. The type strain is sensitive to clindamycin (MIC 0.125 µg/mL), penicillin G (MIC <0.016 µg mL^{-1}), and metronidazole (MIC <0.016 µg mL^{-1}). The major cellular fatty acids of strain CRIB^T are saturated and unsaturated straight chain C_{12}-C_{19} fatty acids, with $C_{16:0}$ being the predominant fatty acid. Details of the fatty acid profile are provided in Table 2, with percentage compositions serving as a guide to the relative abundance of the different compounds. The peptidoglycan of strain CRIB^T contains alanine and glutamic acid (molar ratio 1.4:1.0) and a di-*N*-heptafluorobutyryl-lanthionine-diisobutylester was detected by GC-MS indicating that lanthionine is present. The peptidoglycan of strain CRIB^T is of the A1σ lanthionine-direct type, which has, to our knowledge, not been previously described for species belonging to the class *Clostridia*. The cell wall sugars are glucose and galactose. The only respiratory lipoquinone present is the menaquinone MK-6. The polar lipids of strain CRIB^T comprise six glycolipids, four phospholipids and a lipid that did not stain with any of the specific spray reagents used (Figure 2). Phospholipids with free amino groups (e.g. phosphatidylethanolamine) are absent.

The type strain, CRIB^T (=DSM 25109^T =NIZO 4048^T), was isolated from the gastro-intestinal tract of a Sprague-Dawley rat at Wageningen University, The Netherlands. The genomic DNA G+C content of strain CRIB^T is 28.1 mol%. Strain CRIB^T shows low DNA-DNA relatedness (percentage reassociation) to the type strains *Romboutsia lituseburensis* (*Clostridium lituseburense*) DSM 797^T (15.5±0.8%), *Intestinibacter bartlettii* (*Clostridium bartlettii*) DSM 16795^T (20.4±3.5%) and *Asaccharospora irregularis* (*Clostridium irregulare*) DSM 2635^T (18.1±1.3%).

Description of *Romboutsia lituseburensis* (Laplanche and Saissac 1948) comb. nov.
Romboutsia lituseburensis (li.tus.e.bur.en'sis. L. n. *litus*,coast; L. n. *ebur* ivory; N.L. fem. adj. *lituseburensis* pertaining to Côte d'Ivoire).

Basonym: *Clostridium lituseburense* (Laplanche and Saissac in Prévot 1948) McClung and McCoy 1957 (Approved Lists 1980) [47, 58, 67].

The properties of *Romboutsia lituseburensis* are as given for *Clostridium lituseburense* by Holdeman *et al.* [27], with the following additions. With the API 50 CH and API 20 A systems, acid is produced from D-fructose, D-glucose, D-maltose, D-ribose and D-sucrose. The polar lipid profile comprises five glycolipids and five phospholipids (Figure 2). Phospholipids with free amino groups (e.g. phosphatidylethanolamine) are absent. The predominant cellular fatty acids are $C_{16:0}$, $C_{17:0}$ and $C_{18:1}$ ω7c. Details of the fatty acid profile are provided in Table 2, with percentage compositions serving as a guide to the relative abundance of the different compounds.

The type strain is A25K^T (=ATCC 25759^T =BCRC 14536^T =CCUG 18920^T =DSM 797^T =JCM 1404^T =NCIMB 10637^T =VTT E-021853).

Description of *Intestinibacter* gen. nov.
Intestinibacter (In.tes'ti.ni.bac'ter. L. neut. n. *intestinum* gut; N.L. masc. n. *bacte*r a rod; N.L. masc. n. *Intestinibacter* a rod from the gut).

Members of the genus are anaerobic. Cells stain Gram-positive or Gram-variable. Cells are straight or slightly curved non-motile rods, which often form chains. Capable of spore formation, but free spore formation can be very sparse. The major end products of metabolism are acetate and formate. Cells contain both phospholipids and glycolipids. The major phospholipids are labelled PL3, PL4, PL7 and P11 and the glycolipids comprise GL1, GL2, GL4, GL6, GL9, GL11 and GL12, the R_f values of which can be determined by reference to Figure 2 and also serve as reference points for future work on the elucidation of these structures. Phospholipids with free amino groups (e.g. phosphatidylethanolamine) are absent. The fatty acids present in the hydrolysis products of whole cells are both iso- and anteiso-branched as well straight chain saturated and unsaturated (Table 2). Dimethyl acetals (i.e. plasmalogens) are predominately straight chain with smaller amounts of iso- and anteiso-branched derivates. The genomic DNA G+C content is 29.8 mol%. The type species is *Intestinibacter bartlettii*.

Description of *Intestinibacter bartletii* (Song *et al.* 2004) comb. nov.

Intestinibacter bartlettii (bart.let'ti.i. N.L. masc. gen. n. *bartlettii* to honour John G. Bartlett, for his contributions to the role of intestinal microbiota in disease and to our knowledge of infectious diseases in general).

Basonym: *Clostridium bartlettii* Song et al. 2004 [69].

The properties of *Intestinibacter bartlettii* are as given for *Clostridium bartlettii* by Song *et al.* [69], with the following additions. With the API 50 CH and API 20 A systems, acid is produced from D-fructose, gentiobiose, D-glucose, D-maltose, D-mannitol, D-ribose, D-sorbitol, D-sucrose, and D-tagatose. Little or no H_2 is formed. With the API 50 CH and API 20 A systems, acid is produced from D-fructose, gentiobiose, D-glucose, D-maltose, D-mannitol, D-ribose, D-sorbitol, D-sucrose, and D-tagatose. The polar lipid profile comprises seven glycolipids, four phospholipids and one lipid that did not stain with any of the specific spray reagents used. Phospholipids with free amino groups (e.g. phosphatidylethanolamine) are absent. The predominant cellular fatty acids are iso-$C_{15:0}$ and $C_{16:0}$ DMA. Details of the fatty acid profile are provided in Table 2, with percentage compositions serving as a guide to the relative abundance of the different compounds.

The type strain is WAL 16138^T (=ATCC BAA-827^T =CCUG 48940^T =DSM 16795^T).

Description of *Terrisporobacter* gen. nov.

Terrisporobacter (Ter.ri.spo'ro.bac'ter. L. n. *terra*, soil; Gr. fem. n. *spora* a seed, and in biology a spore; N.L. masc. n. *bacter* a rod; N.L. masc. n. *Terrisporobacter* a spore-forming rod found in soil).

Members of the genus are anaerobic. Cells stain Gram-positive or Gram-variable. Cells are straight or slightly curved motile rods, occurring singly or in pairs. All species are capable of spore formation. The major end product of metabolism is acetate. Cells contain both phospholipids and glycolipids. The major phospholipids are labelled PL2, PL4, PL5, PL6, PL7 and PL10 and the glycolipids comprise GL1, GL3, GL6, GL7, GL8 and GL9, the R_f values of which can be determined by reference to Figure 2 and also serve as reference points for future work on the elucidation of these structures. Phospholipids with free amino groups (e.g. phosphatidylethanolamine) are absent. The fatty acids present in the hydrolysis products of whole cells are both straight chain saturated and unsaturated (which predominate) as well as iso- and anteiso-branched (Table 2). Dimethyl acetals (i.e. plasmalogens) include straight chain, iso- and anteiso-branched derivates. The genomic DNA G+C content is 25-30 mol%. The type species is *Terrisporobacter glycolicus.*

Description of *Terrisporobacter glycolicus* (Gaston and Stadtman 1963) comb. nov.
Terrisporobacter glycolicus (gly.co'li.cus. L. adj. suff. *-icus* related to, belonging to; N.L. masc. adj. *glycolicus* referring to the ability to ferment ethylene glycol).

Basonym: *Clostridium glycolicum* Gaston and Stadtman 1963 (Approved Lists 1980) [19, 67].

The properties of *Terrisporobacter glycolicus* are as given for *Clostridium glycolicum* by Holdeman *et al.* [27] and in the emendation of Chamkha *et al.* [7], with the following additions. The polar lipid profile comprises seven glycolipids, seven phospholipids and one lipid that did not stain with any of the specific spray reagents used. Phospholipids with free amino groups (e.g. phosphatidylethanolamine) are absent. The predominant cellular fatty acid is $C_{16:0}$. Details of the fatty acid profile are provided in Table 2, with percentage compositions serving as a guide to the relative abundance of the different compounds.

The type strain is ATCC 14880^T (=BCRC 14553^T =DSM 1288^T =JCM 1401^T =NCIMB 10632^T =NCTC 13026^T).

4

Description of *Terrisporobacter mayombei* (Kane *et al.* 1992) comb. nov.
Terrisporobacter mayombei (ma.yom.be.i. N.L. gen. n. *mayombei* of the Mayombe tropical rainforest, People's Republic of Congo).

Basonym: *Clostridium mayombei* Kane *et al.* 1992 [35].

Although the epithet *mayombensis* may be considered to be the proper orthographic form, this correction is no longer allowed as stated in the amendment to Rule 61 [minute 7 (1)] [11] of the International Code of Nomenclature of Bacteria [42].

The properties of *Terrisporobacter mayombei* are as given for *Clostridium mayombei* by Kane *et al.* [34], with the following additions. The polar lipid profile comprises six glycolipids and six phospholipids. Phospholipids with free amino groups (e.g. phosphatidylethanolamine) are absent. The predominant cellular fatty acid is $C_{16:0}$. Details of the fatty acid profile are provided in Table 2, with percentage compositions serving as a guide to the relative abundance of the different compounds.

The type strain is SFC-5^T (=ATCC 51428^T =DSM 6539^T).

Description of *Asaccharospora* gen. nov.

Asaccharospora (A.sac.cha.ro.spo'ra. Gr. pref. *a*, not; Gr. n. *sakchâr* sugar; Gr. fem. n. *spora* a seed, and in biology a spore; N.L. fem. n. *Asaccharospora* a spore-forming organism that is unable to ferment sugars).

Members of the genus are anaerobic. Cells stain Gram-positive or Gram-variable. Cells are straight or slightly curved motile rods, which often form chains. Cells may be quite filamentous. Spore formation is observed. Carbohydrates are not fermented. The major end product of metabolism is acetate. Cells contain both phospholipids and glycolipids. The major phospholipids are labelled PL4, PL5, PL7 and PL8 and the glycolipids comprise GL1, GL3a/b, GL6 and GL9, the R_f values of which can be determined by reference to Figure 2 and also serve as reference points for future work on the elucidation of these structures. Phospholipids with free amino groups (e.g. phosphatidylethanolamine) are absent. The fatty acids present in the hydrolysis products of whole cells are both straight chain saturated and unsaturated (which predominate) as well as iso- and anteiso-branched (Table 2). Dimethyl acetals (i.e. plasmalogens) are almost absent. The type species is *Asaccharospora irregularis*.

Description of *Asaccharospora irregularis* (Choukévitch 1911) comb. nov.

*Asaccharospora irregularis (*ir.re.gu.lar'is. L. fem. adj. *irregularis* irregular, referring to pleomorphic, irregular cells.)

Basonym: *Clostridium irregulare* (Choukévitch 1911) Prévot 1938 (Approved Lists 1980) [8, 57, 67].

The properties of *Asaccharospora irregularis* are as given for *Clostridium irregulare* by Holdeman *et al.* [27], with the following additions. With the API 50 CH and API 20 A systems, there is no acid is produced with any of the carbohydrates tested. The polar lipid profile comprises four glycolipids and four phospholipids. The predominant cellular fatty acids are $C_{16:0}$ and $C_{16:1}$ ω7c. Details of the fatty acid profile are provided in Table 2, with percentage compositions serving as a guide to the relative abundance of the different compounds.

The type strain is 6VI^T (=ATCC 25756^T =BCRC 14528^T =DSM 2635^T =JCM 1425^T =NCIMB 11830^T).

Based on the results presented here an emended description of *Peptostreptococcus anaerobius* would also be appropriate.

Emended description of *Peptostreptococcus anaerobius* (Natvig 1905) Kluyver and van Niel 1936 (Approved Lists 1980)

Peptostreptococcus anaerobius (an.a.e.ro'bi.us. Gr. pref. *an* not; Gr. n. *aer* air; Gr. n. *bios* life; N.L. adj. *anaerobius* not living in air, anaerobic).

Peptostreptococcus anaerobius (Natvig 1905) Kluyver and van Niel 1936 (Approved Lists 1980) [38, 53, 67].

The properties of the species are those given by Holdeman *et al.* [27] and Ezaki [14] (for the species *Peptostreptococcus anaerobius*) with the following additions. Respiratory lipoquinones have not been reported. Polar lipids comprise phospholipids and glycolipids. The phospholipids present are PL4 and PL7 and the only glycolipids comprise GL1 and GL9, the R_f values of which can be determined by reference to Figure 2 and also serve as reference points for future work on the elucidation of these structures. Phospholipids with free amino groups (e.g. phosphatidylethanolamine) are absent. The fatty acids present in the hydrolysis products of whole cells are predominantly saturated straight chain and iso- and anetiso-branched chain (Table 2). Dimethyl acetals (i.e. plasmalogens) are predominately even straight chain with smaller amounts of iso- and anteiso-branched derivates. Details of the fatty acid profile are provided in Table 2, with percentage compositions serving as a guide to the relative abundance of the different compounds. The fatty acid data cited by Ezaki [14] from Ezaki *et al.* [16] are not consistent with the results presented here and are not to be included in the emended description.

The type strain is Prévot 4372^T (=ATCC 27337^T =BCRC 10722^T =CCUG 7835^T =CIP 104411^T =DSM 2949^T =KCTC 5182^T =LMG 15865^T =NCTC 11460^T =VTT E-022078^T).

The present publication also has to deal with an additional nomenclatural issue. Based on information published in the second edition of *Bergey's Manual of Systematic Bacteriology*, the order *Clostridiales* Prévot 1953 [61] contains the family *Eubacteriaceae* Ludwig *et al.* 2010 [45], which in turn is based on the type genus *Eubacterium* Prévot 1938 (Approved Lists 1980). The use of the name *Clostridiales* Prévot 1953 (Approved Lists 1980) [59, 67] is in contradiction to Principles 1 and 8 as well as Rule 23a of the current Code of nomenclature dealing with prokaryotes that requires that the older name *Eubacteriales* Buchanan 1917 (Approved Lists 1980) [5, 67] is to be used for this order [42]. This would also require that an emended description of the order *Eubacteriales* Buchanan 1917 (Approved Lists 1980) is published. We therefore formally emend the circumscription and description of the order *Eubacteriales* Buchanan 1917 (Approved Lists 1980).

Emendation of the order *Eubacteriales* Buchanan 1917 (Approved Lists 1980)

The properties of the order are those given by Rainey [61] under the order name *Clostridiales* Prévot 1953 (Approved Lists 1980). The type genus is *Eubacterium* Prévot 1938 (Approved Lists 1980). The order contains the families *Caldicoprobacteraceae, Christensenellaceae, Clostridiaceae, Defluviitaleaceae, Eubacteriaceae, Gracilibacteraceae, Heliobacteriaceae, Lachnospiraceae, Peptococcaceae, Peptostreptococcaceae, Ruminococcaceae,* and *Syntrophomonadaceae*.

It should be noted that as long as the type genera of the order *Clostridiales* Prévot 1953 (Approved Lists 1980) [i.e. *Clostridium* Prazmowski 1880 [55](Approved Lists 1980) [55, 67]] and the order *Eubacteriales* Buchanan 1917 (Approved Lists 1980) [i.e. *Eubacterium* Prévot 1938 (Approved Lists 1980) [56, 67]] are placed in one order, then the name of the order *Eubacteriales* Buchanan 1917 (Approved Lists 1980) has priority. However, should these two genera be placed in separate orders, then the names *Clostridiales* Prévot 1953 (Approved Lists 1980) and *Eubacteriales* Buchanan 1917 (Approved Lists 1980) may be used for the two orders, but an emendation would be required.

Acknowledgements

The authors thank Dr. Peter Schumann, Dr. Susanne Verbarg, Dr. Cathrin Spröer and Anja Frühling (all from the Leibniz-Institut DSMZ) for their assistance in chemotaxonomic and genotypic characterization, Elly van Donselaar and Rob Mesman (from the Electron Microscopy group, Utrecht University) for operation of the high-pressure freezer, the technicians of the Laboratory for Medical Microbiology and Immunology of the St. Antonius Hospital (Nieuwegein) for their assistance in the assessment of colony morphology, and Dr. Caroline Plugge (Laboratory of Microbiology, Wageningen University) for her helpful comments and discussions during the preparation of the manuscript. This work was supported by grants from SenterNovem (Agentschap NL), an agency of the Dutch Ministry of Economic Affairs (grant numbers: TSGE-3109 and FND-07013). Laura van Niftrik was supported by a grant from the Netherlands Organization for Scientific Research (VENI grant number 863.09.009). All animal procedures were performed in accordance with institutional guidelines and with approval from the institutional animal care committee of the University Medical Center, Utrecht, The Netherlands.

References

1. Adachi K., A. Katsuta, S. Matsuda, X. Peng, N. Misawa, Y. Shizuri, R.M. Kroppenstedt, *et al.* (2007). *Smaragdicoccus niigatensis* gen. nov., sp. nov., a novel member of the suborder *Corynebacterineae*. *Int J Syst Evol Microbiol*; 57 (Pt 2): 297-301.
2. Biebl H. and C. Spröer (2002). Taxonomy of the glycerol fermenting clostridia and description of *Clostridium diolis* sp. nov. *Syst Appl Microbiol*; 25 (4): 491-497.
3. Borkenhagen L.F., E.P. Kennedy and L. Fielding (1961). Enzymatic formation and decarboxylation of phosphatidylserine. *J Biol Chem*; 236 (6): PC28-PC30.
4. Brown W.L., Z.J. Ordal and H.O. Halvorson (1957). Production and cleaning of spores of putrefactive anaerobe 3679. *Appl Microbiol*; 5 (3): 156-159.
5. Buchanan R.E. (1917). Studies on the Nomenclature and Classification of the Bacteria: III. The Families of the Eubacteriales. *J Bacteriol*; 2 (4): 347-350.
6. Cashion P., M.A. Holder-Franklin, J. McCully and M. Franklin (1977). A rapid method for the base ratio determination of bacterial DNA. *Anal Biochem*; 81 (2): 461-466.
7. Chamkha M., M. Labat, B.K. Patel and J.L. Garcia (2001). Isolation of a cinnamic acid-metabolizing *Clostridium glycolicum* strain from oil mill wastewaters and emendation of the species description. *Int J Syst Evol Microbiol*; 51 (Pt 6): 2049-2054.
8. Choukévitch J. (1911). Étude de la flore bactérienne du gros intestin du cheval. *Ann Inst Pasteur*; 25: 245-368.
9. Collins M.D., P.A. Lawson, A. Willems, J.J. Cordoba, J. Fernandez-Garayzabal, P. Garcia, J. Cai, *et al.* (1994). The phylogeny of the genus *Clostridium*: proposal of five new genera and eleven new species combinations. *Int J Syst Bacteriol*; 44 (4): 812-826.
10. De Ley J., H. Cattoir and A. Reynaerts (1970). The quantitative measurement of DNA hybridization from renaturation rates. *Eur J Biochem*; 12 (1): 133-142.
11. De Vos P. and H. Truper (2000). Judicial Commission of the International Committee on Systematic Bacteriology; IXth International (IUMS) Congress of Bacteriology and Applied Microbiology. *Int J Syst Evol Microbiol*; 50 (6): 2239-2244.
12. Drucker D.B., H.M. Wardle and V. Boote (1996). Phospholipid profiles of *Clostridium difficile*. *J Bacteriol*; 178 (19): 5844-5846.
13. Duncan C.L. and D.H. Strong (1968). Improved medium for sporulation of *Clostridium perfringens*. *Appl Microbiol*; 16 (1): 82-89.
14. Ezaki T. (2009). Genus I. *Peptostreptococcus*. In: *Bergey's Manual of Systematic Bacteriology*. 2nd Edn, 3: The *Firmicutes*, pp 1008-1009. Edited by P. De Vos, G.M. Garrity, D. Jones, N.R. Krieg, *et al.* New York: Springer.
15. Ezaki T. (2010). *Peptostreptococcaceae* fam. nov. In *List of new names and new combinations previously effectively, but not validly, published*, Validation List no. 132. *Int J Syst Bacteriol*; 60: 469-472.
16. Ezaki T., N. Yamamoto, K. Ninomiya, S. Suzuki and E. Yabuuchi (1983). Transfer of *Peptococcus indolicus*, *Peptococcus asaccharolyticus*, *Peptococcus prevotii*, and *Peptococcus magnus* to the genus *Peptostreptococcus* and proposal of *Peptostreptococcus tetradius* sp. nov. *Int J Syst Bacteriol*; 33 (4): 683-698.
17. Feulgen R. and K. Voit (1924). Über einen weitverbreiteten festen aldehyd. *Pflüger's Archiv für die gesamte Physiologie des Menschen und der Tiere*; 206 (1): 389-410.
18. Frosolono M.F. and M.M. Rapport (1969). Reactivity of plasmalogens: kinetics of acid-catalyzed hydrolysis. *J Lipid Res*; 10 (5): 504-506.
19. Gaston L.W. and E.R. Stadtman (1963). Fermentation of ethylene glycol by *Clostridium glycolicum*, sp. n. *J Bacteriol*; 85: 356-362.
20. Gerritsen J., H. Smidt, G.T. Rijkers and W.M. de Vos (2011). Intestinal microbiota in human health and disease: the impact of probiotics. *Genes Nutr*; 6 (3): 209-240.
21. Gerritsen J., H.M. Timmerman, S. Fuentes, L.P. van Minnen, H. Panneman, S.R. Konstantinov, F.M. Rombouts, *et al.* (2011). Correlation between protection against sepsis by probiotic therapy and stimulation of a novel bacterial phylotype. *Appl Environ Microbiol*; 77 (21): 7749-7756.
22. Goldfine H. and N.C. Johnston (2005). Membrane lipids of *Clostridia*. In: *Handbook on Clostridia*. pp 297-309. Edited by P. Dürre. Boca Raton, Florida: CRC Press.
23. Gottschalk G., J.R. Andreesen and H. Hippe (1981). The genus *Clostridium* (nonmedical aspects). In: *The Prokaryotes*. 1 Edn, 2, pp 1767-1803. Edited by M.P. Starr, H. Stolp, H.G. Trüper, A. Balows, *et al.* New York: Springer-Verlag.
24. Guan Z., N.C. Johnston, S. Aygun-Sunar, F. Daldal, C.R. Raetz and H. Goldfine (2011). Structural characterization of the polar lipids of *Clostridium novyi* NT. Further evidence for a novel anaerobic biosynthetic pathway to plasmalogens.

Biochim Biophys Acta; 1811 (3): 186-193.
25. Guan Z., N.C. Johnston, C.R. Raetz, E.A. Johnson and H. Goldfine (2012). Lipid diversity among botulinum neurotoxin-producing clostridia. *Microbiology*; 158 (Pt 10): 2577-2584.
26. Guan Z., B. Tian, A. Perfumo and H. Goldfine (2013). The polar lipids of *Clostridium psychrophilum*, an anaerobic psychrophile. *Biochim Biophys Acta*; 1831 (6): 1108-1112.
27. Holdeman L.V., E.P. Cato and W.E.C. Moore (1977). Anaerobe Laboratory Manual. 4th Edn. Blacksburg, VA: Virginia Polytechnic Institute and State University.
28. Huss V.A.R., H. Festl and K. Schleifer (1983). Studies on the spectrophotometric determination of DNA hybridization from renaturation rates. *Syst Appl Microbiol*; 4: 184-192.
29. Johnson J.L. and B.S. Francis (1975). Taxonomy of the *Clostridia*: ribosomal ribonucleic acid homologies among the species. *J Gen Microbiol*; 88 (2): 229-244.
30. Johnston N.C., S. Aygun-Sunar, Z. Guan, A.A. Ribeiro, F. Daldal, C.R. Raetz and H. Goldfine (2010). A phosphoethanolamine-modified glycosyl diradylglycerol in the polar lipids of *Clostridium tetani. J Lipid Res*; 51 (7): 1953-1961.
31. Johnston N.C., J.K. Baker and H. Goldfine (2004). Phospholipids of *Clostridium perfringens*: a reexamination. *FEMS Microbiol Lett*; 233 (1): 65-68.
32. Johnston N.C. and H. Goldfine (1983). Lipid composition in the classification of the butyric acid-producing clostridia. *J Gen Microbiol*; 129 (4): 1075-1081.
33. Kämpfer P., S. Buczolits, A. Albrecht, H.J. Busse and E. Stackebrandt (2003). Towards a standardized format for the description of a novel species (of an established genus): *Ochrobactrum gallinifaecis* sp. nov. *Int J Syst Evol Microbiol*; 53 (Pt 3): 893-896.
34. Kane M.D., A. Brauman and J.A. Breznak (1991). *Clostridium mayombei* sp. nov., an H2/CO2 acetogenic bacterium from the gut of the African soil-feeding termite, *Cubitermes speciosus. Arch Microbiol*; 156: 99-104.
35. Kane M.D., A. Brauman and J.A. Breznak (1992). *Clostridium mayombei* sp. nov. In *Validation of the publication of new names and new combinations previously effectively published outside the IJSB*, List no. 40. *Int J Syst Bacteriol*; 42: 191-192.
36. Kanfer J. and E.P. Kennedy (1964). Metabolism and function of bacterial lipids: II. Biosynthesis of phospholipids in *Escherichia coli*. *J Biol Chem*; 239: 1720-1726.
37. Kim O.S., Y.J. Cho, K. Lee, S.H. Yoon, M. Kim, H. Na, S.C. Park, *et al.* (2012). Introducing EzTaxon-e: a prokaryotic 16S rRNA gene sequence database with phylotypes that represent uncultured species. *Int J Syst Evol Microbiol*; 62 (Pt 3): 716-721.
38. Kluyver A.J. and C.B. van Niel (1936). Prospects for a natural system of classification of bacteria. *Zentralbl Bakterol Parasitenkd Infectinskr Hyg, Abt II*; 94: 109-115.
39. Korachi M., M. Rupnik, A.S. Blinkhorn, V. Boote and D.B. Drucker (2002). Comparison of polar lipid profiles of *Clostridium difficile* isolates from different geographical locations. *Anaerobe*; 8 (1): 35-39.
40. Kuykendall L.D., M.A. Roy, J.J. O'Neill and T.E. Devine (1988). Fatty acids, antibiotic resistance, and deoxyribonucleic acid homology groups of *Bradyrhizobium japonicum. Int J Syst Bacteriol*; 38: 358-361.
41. Lane D.J. (1991). 16S/23S rRNA sequencing. In: *Nucleic Acid Techniques in Bacterial Systematics*. pp 115-175. Edited by E. Stackebrandt and M. Goodfellow. New York: John Wiley and Sons, Inc.
42. Lapage S.P., P.H.A. Sneath, E.F. Lessel, V.B.D. Skerman, H.P.R. Seeliger and W.A. Clark (1992). International Code of Nomenclature of Bacteria, and Statutes of the International Committee on Systematic Bacteriology, and Statutes of the Bacteriology and Applied Microbiology Section of the International Union of Microbiological Societies. Bacteriological code, 1990 revision. Washington, D.C.: ASM Press.
43. Leifson E. (1960).Atlas of bacterial flagellation. In. New York: Academic Press Inc.
44. Ley R.E., C.A. Lozupone, M. Hamady, R. Knight and J.I. Gordon (2008). Worlds within worlds: evolution of the vertebrate gut microbiota. *Nat Rev Microbiol*; 6 (10): 776-788.
45. Ludwig W. (2010). *Eubacteriaceae* fam. nov. In *List of new names and new combinations previously effectively, but not validly, published*, Validation List no. 132. *Int J Syst Evol Microbiol*; (60): 469-472.
46. MacKenzie S.L. (1987). Gas chromatographic analysis of amino acids as the N-heptafluorobutyryl isobutyl esters. *J Assoc Off Anal Chem*; 70 (1): 151-160.
47. McClung L.S. and E. McCoy (1957). Genus II *Clostridium* Prazmowski 1880. In: *Bergey's Manual of Determinative Bacteriology*. 7th Edn, pp 634-693. Edited by R.S. Breed, E.G.D. Murray and N.R. Smith. Baltimore: The Williams & Wilkins Company.
48. Meier-Kolthoff J.P., M. Goker, C. Sproer and H.P. Klenk (2013). When should a DDH experiment be mandatory in microbial taxonomy? *Archives of Microbiology*; 195 (6): 413-418.

49. Mesbah M., U. Premachandran and W.B. Whitman (1989). Precise measurements of the G+C content of deoxyribonucleic acid by high-pressure liquid chromatography. *Int J Syst Bacteriol*; 39: 159-167.
50. Miller L.T. (1982). Single derivatization method for routine analysis of bacterial whole-cell fatty acid methyl esters, including hydroxy acids. *J Clin Microbiol*; 16 (3): 584-586.
51. Mollenhauer H.H. (1964). Plastic embedding mixtures for use in electron microscopy. *Stain Technol*; 39: 111-114.
52. Morrison W.R. and L.M. Smith (1964). Preparation of fatty acid methyl esters and dimethylacetals from lipids with boron fluoride-methanol. *J Lipid Res*; 5: 600-608.
53. Natvig H. (1905). Bakteriologische verhältnisse in weiblichen genitalsekreten. *Arch Gynekol*; 76 (3): 701-858.
54. Oulevey J., H. Bahl and T.O. W. (1986). Novel alkyl-1-enyl ether lipids isolated from *Clostridium acetobutylicum*. *Arch Microbiol*; 144 (2): 166-168.
55. Prazmowski A. (1880).Untersuchungen über die entwicklungsgeschichte und fermentwirkung einiger bacterien-arten. Inaug. Diss. Hugo Voigt, Leipzig. In: 1-58.
56. Prévot A.R. (1938). Études de systématique bactérienne. III. Invalidité du genre Bacteroides Castellani et Chalmers démembrement et reclassification. *Ann Inst Pasteur*; 60: 285-307.
57. Prévot A.R. (1938). Études de systématique bactérienne. IV Critique de la conception actuelle du genre *Clostridium*. *Ann Inst Pasteur*; 61: 72-91.
58. Prévot A.R. (1948). Étude des bactéries anaérobies d'Afrique occidentale française (Sénégal, Guinée, Côte d'Ivoire). *Ann Inst Pasteur*; 74: 157-170.
59. Prévot A.R. (1953). Dictionnaire des bactéries pathogènes. 2nd Edn. Paris: Masson.
60. Pruesse E., J. Peplies and F.O. Glockner (2012). SINA: accurate high-throughput multiple sequence alignment of ribosomal RNA genes. *Bioinformatics*; 28 (14): 1823-1829.
61. Rainey F.A. (2009). Order I. *Clostridiales*. In: *Bergey's Manual of Systematic Bacteriology*. 2nd Edn, 3: The *Firmicutes*, pp 736. Edited by P. De Vos, G.M. Garrity, D. Jones, N.R. Krieg, *et al.* New York: Springer.
62. Rainey F.A., B.J. Hollen and A. Small (2009). Genus 1. *Clostridium*. In: *Bergey's Manual of Systematic Bacteriology*. 2nd Edn, Vol. 3: The *Firmicutes*, pp 738-828. Edited by P. De Vos, G.M. Garrity, D. Jones, N.R. Krieg, *et al.* New York: Springer.
63. Reynolds E.S. (1963). The use of lead citrate at high pH as an electron-opaque stain in electron microscopy. *J Cell Biol*; 17: 208-212.
64. Rhuland L.E., E. Work, R.F. Denman and D.S. Hoare (1955). The behavior of the isomers of α,ε-diaminopimelic acid on paper chromatograms. *J Am Chem Soc*; 77 (18): 4844-4846.
65. Sacks L.E. and P.A. Thompson (1977). Increased spore yields of *Clostridium perfringens* in the presence of methylxanthines. *Appl Environ Microbiol*; 34 (2): 189-193.
66. Schleifer K. (1985). Analysis of the chemical composition and primary structure of murein. In: *Methods in Microbiology*. 18, pp 123-157. Edited by G. Gottschalk. London: Academic Press Inc.
67. Skerman V.B.D., V. McGowan and P.H.A. Sneath (1980). Approved Lists of Bacterial Names. *Int J Syst Bacteriol*; 30: 225-420.
68. Smibert R.M. and N.R. Krieg (1994). Phenotypic characterization. In: *Methods for General and Molecular Bacteriology*. pp 607-654. Edited by P. Gerhardt, R.G.E. Murray, W.A. Wood, N.R. Krieg. Washington, D.C.: American Society for Microbiology.
69. Song Y.L., C.X. Liu, M. McTeague, P. Summanen and S.M. Finegold (2004). *Clostridium bartlettii* sp. nov., isolated from human faeces. *Anaerobe*; 10 (3): 179-184.
70. Stackebrandt E. and J. Ebers (2006). Taxonomic parameters revisited: tarnished gold standards. *Microbial Today*; 33: 152-155.
71. Stackebrandt E., I. Kramer, J. Swiderski and H. Hippe (1999). Phylogenetic basis for a taxonomic dissection of the genus *Clostridium*. *FEMS Immunol Med Microbiol*; 24 (3): 253-258.
72. Stams A.J., J.B. Van Dijk, C. Dijkema and C.M. Plugge (1993). Growth of syntrophic propionate-oxidizing bacteria with fumarate in the absence of methanogenic bacteria. *Appl Environ Microbiol*; 59 (4): 1114-1119.
73. Staneck J.L. and G.D. Roberts (1974). Simplified approach to identification of aerobic actinomycetes by thin-layer chromatography. *Appl Microbiol*; 28 (2): 226-231.
74. Suzuki M.T., L.T. Taylor and E.F. DeLong (2000). Quantitative analysis of small-subunit rRNA genes in mixed microbial populations via 5'-nuclease assays. *Appl Environ Microbiol*; 66 (11): 4605-4614.
75. Tamura K., G. Stecher, D. Peterson, A. Filipski and S. Kumar (2013). MEGA6: Molecular Evolutionary Genetics Analysis Version 6.0. *Mol Biol Evol*; 30 (12): 2725-2729.
76. Thiele O.W., J. Oulevey and H. Bahl (1985). Neuartige alkenylether-lipide aus anaeroben bakterien. *Fette, Seifen, Anstrichmittel*; 87 (13-14): 551-556.

77. Tian B., Z. Guan and H. Goldfine (2013). An ethanolamine-phosphate modified glycolipid in *Clostridium acetobutylicum* that responds to membrane stress. *Biochim Biophys Acta*; 1831 (6): 1185-1190.
78. Tindall B.J. (1990). A comparative study of the lipid-composition of *Halobacterium saccharovorum* from various sources. *Syst Appl Microbiol*; 13 (2): 128-130.
79. Tindall B.J. (1990). Lipid composition of *Halobacterium lacusprofundi*. *FEMS Microbiol Lett*; 66: 199-202.
80. Tindall B.J., R. Rosselló-Móra, H.J. Busse, W. Ludwig and P. Kämpfer (2010). Notes on the characterization of prokaryote strains for taxonomic purposes. *Int J Syst Evol Microbiol*; 60 (Pt 1): 249-266.
81. van Niftrik L., W.J. Geerts, E.G. van Donselaar, B.M. Humbel, A. Yakushevska, A.J. Verkleij, M.S. Jetten, *et al.* (2008). Combined structural and chemical analysis of the anammoxosome: a membrane-bounded intracytoplasmic compartment in anammox bacteria. *Journal of structural biology*; 161 (3): 401-410.
82. Walther P. and A. Ziegler (2002). Freeze substitution of high-pressure frozen samples: the visibility of biological membranes is improved when the substitution medium contains water. *Journal of microscopy*; 208 (Pt 1): 3-10.
83. Wayne L.G., D.J. Brenner, R.R. Colwell, P.A.D. Grimont, O. Kandler, M.I. Krichevsky, L.H. Moore, *et al.* (1987). International Committee on Systematic Bacteriology. Report of the ad hoc committee on reconciliation of approaches to bacterial systematics *Int J Syst Bacteriol*; 37 (4): 463-464.
84. Whitehead T.R., M.A. Cotta, E. Falsen, E. Moore and P.A. Lawson (2011). *Peptostreptococcus russellii* sp. nov., isolated from a swine-manure storage pit. *Int J Syst Evol Microbiol*; 61 (Pt 8): 1875-1879.
85. Whiton R.S., P. Lau, S.L. Morgan, J. Gilbart and A. Fox (1985). Modifications in the alditol acetate method for analysis of muramic acid and other neutral and amino sugars by capillary gas chromatography-mass spectrometry with selected ion monitoring. *J Chromatogr*; 347 (1): 109-120.
86. Wiegel J. (2009). Family I. *Clostridiaceae*. In: *Bergey's Manual of Systematic Bacteriology*. 2nd Edn, 3: The *Firmicutes*, pp 736-738. Edited by P. De Vos, G.M. Garrity, D. Jones, N.R. Krieg, *et al.* New York: Springer.
87. Wiegel J., R. Tanner and F.A. Rainey (2006). An introduction to the family *Clostridiaceae*. In: *The prokaryotes: a handbook on the biology of bacteria; Bacteria: Firmicutes, Cyanobacteria*. 3rd Edn, 4, pp 654-678. Edited by M. Dworkin, S. Falkow, E. Rosenberg, K. Schleifer, *et al.* New York: Springer.
88. Willems A. and M.D. Collins (1994). Phylogenetic placement of *Sarcina ventriculi* and *Sarcina maxima* within group I *Clostridium*, a possible problem for future revision of the genus *Clostridium*. Request for an opinion. *Int J Syst Bacteriol*; 44 (3): 591-593.
89. Yutin N. and M.Y. Galperin (2013). A genomic update on clostridial phylogeny: Gram-negative spore formers and other misplaced clostridia. *Environ Microbiol*; 15 (10): 2631-2641.

Supplemental material

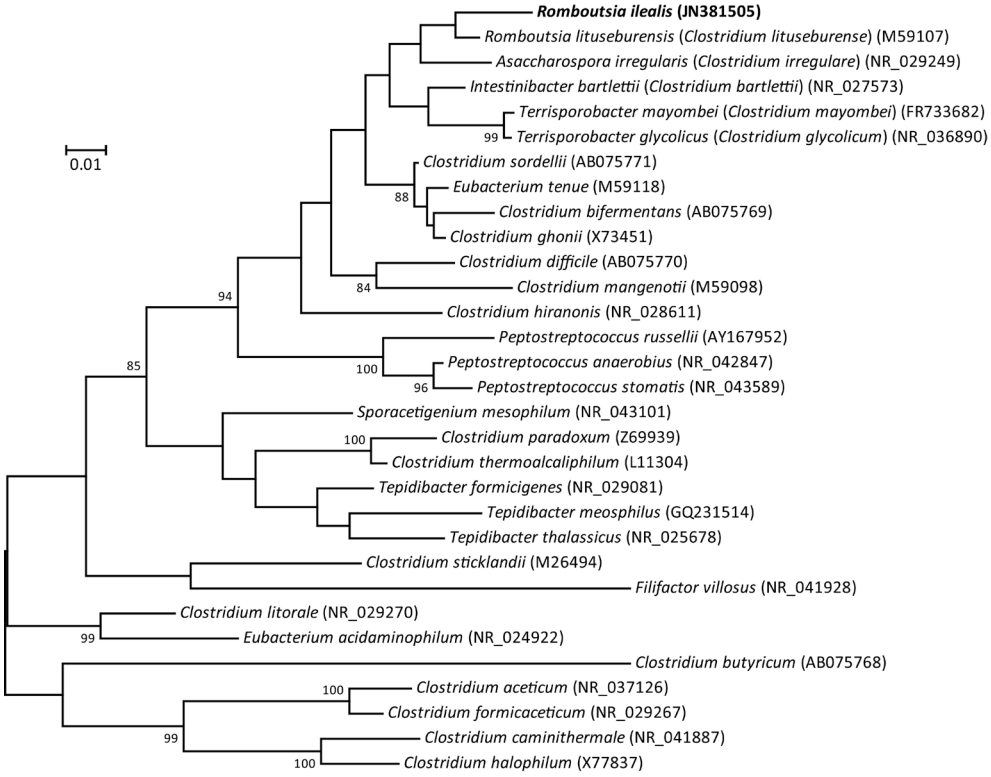

Figure S1. Maximum Likelihood tree based on 16S rRNA gene sequence data, showing the phylogenetic position of strain CRIB^T and other (misclassified) members of the family *Peptostreptococcaceae.* The 16S rRNA gene sequences were aligned using the SINA aligner (http://www.arb-silva.de/aligner/). Phylogenetic tree was constructed using MEGA 6 software [75] with Tamura-Nei's model as substitution model. Only bootstrap values ≥70% are shown at branch nodes. Bar, 1% sequence divergence. GenBank accession numbers are given in parentheses. The 16S rRNA gene sequence from *C. butyricum* ATCC 19398^T (accession no. AB075768), the type strain of the type species of the genus *Clostridium,* was included as well.

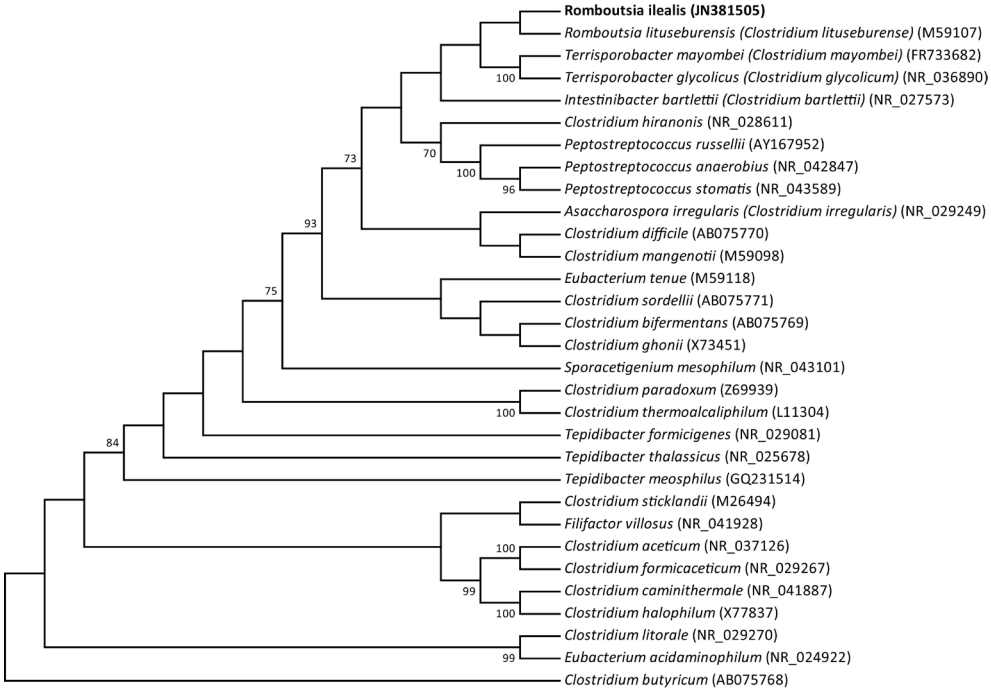

Figure S2. Maximum Parsimony tree based on 16S rRNA gene sequence data, showing the phylogenetic position of strain CRIB^T and other (misclassified) members of the family *Peptostreptococcaceae*. The 16S rRNA gene sequences were aligned using the SINA aligner (http://www.arb-silva.de/aligner/). Phylogenetic tree was constructed using MEGA 6 software [75]. Only bootstrap values ≥70% are shown at branch nodes. Bar, 1% sequence divergence. GenBank accession numbers are given in parentheses. The 16S rRNA gene sequence from *C. butyricum* ATCC 19398^T (accession no. AB075768), the type strain of the type species of the genus *Clostridium*, was included as well.

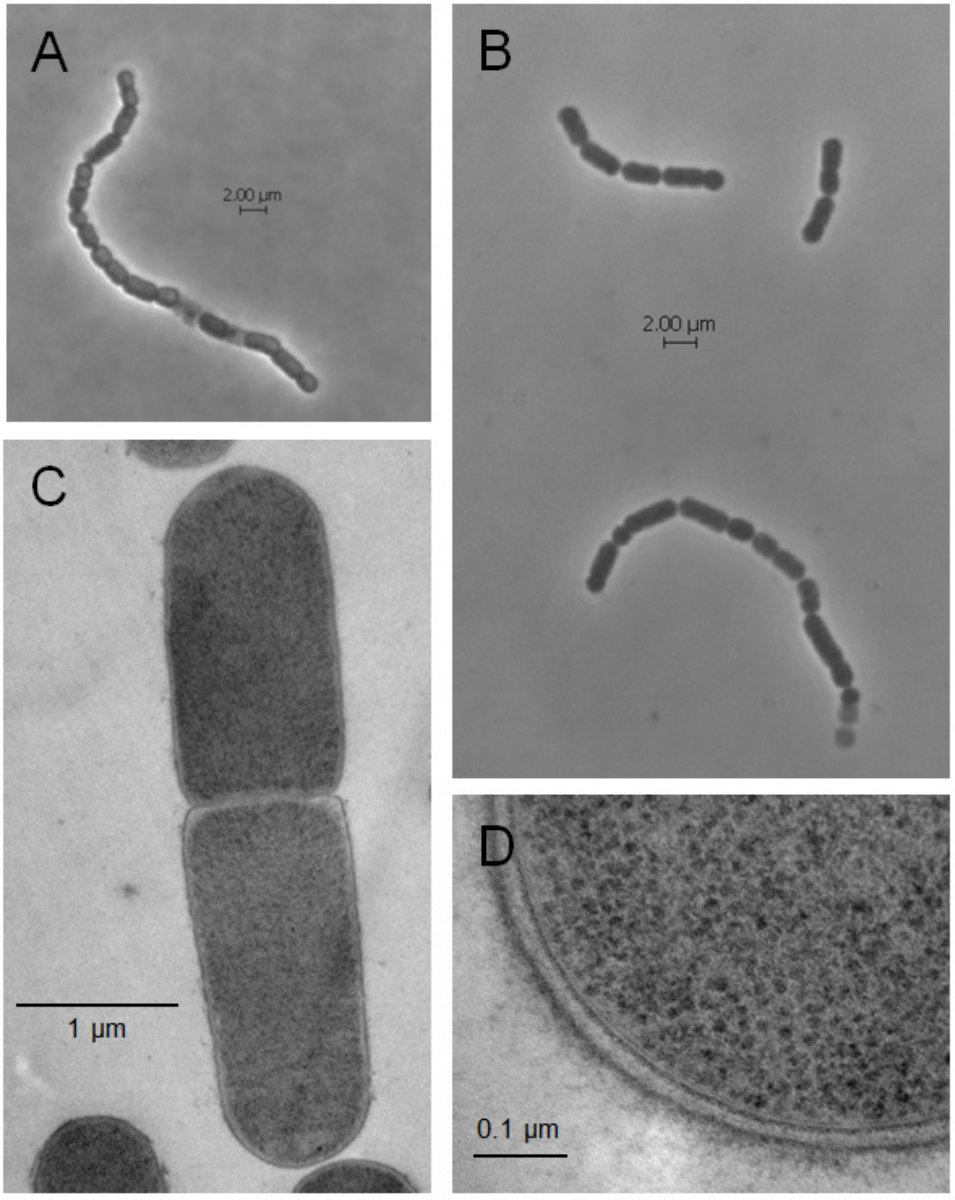

Figure S3. Cell morphology of strain CRIB^T examined by light and transmission electron microscopy. Phase-contrast micrograph of strain CRIB^T grown in liquid (**a**) CRIB medium (bar, 2 µm) and (**b**) PYG medium (bar, 2 µm) after 24 h incubation at 37°C. Transmission electron micrographs of (**c**) cryofixed, freeze-substituted (in acetone containing 2% osmium tetroxide), and Epon-embedded cells of strain CRIB^T grown in CRIB medium (bar, 1 µm) and (**d**) a close-up of the thick cell wall that confers on the typical cell wall morphology of Gram-positive staining cells (bar, 0.1 µm).

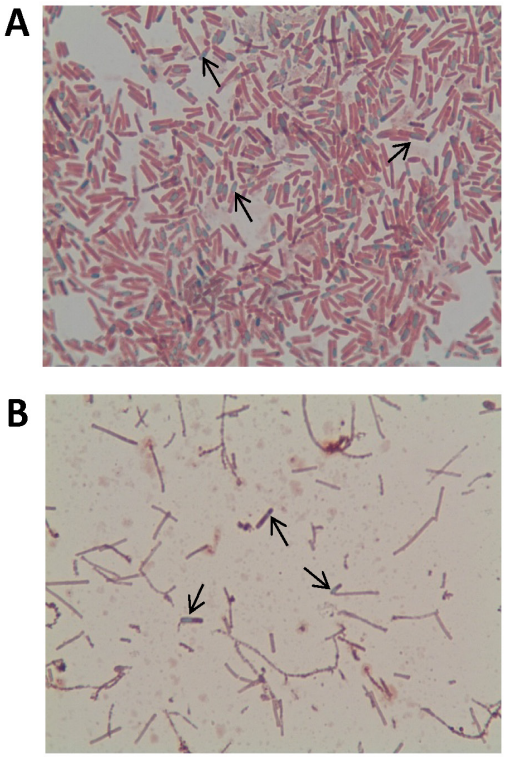

Figure S4. Microscopic pictures (1000× magnification) of spore-forming cultures stained with malachite green (5% w/v in water) and counterstained with safranin (2.5% w/v in ethanol). (**a**) *R. lituseburensis* (*C. lituseburense* comb. nov.) DSM 797^T was grown in Duncan-Strong medium at 37°C for 1 d. (**b**) Strain CRIB^T was grown in CRIB medium at 37°C for 4 d. The black arrows indicate the positions of (fore)spores.

5 CHAPTER

Genomic and functional analysis of *Romboutsia ilealis* CRIB^T reveals adaptation to the small intestine

Jacoline Gerritsen*, Bastian Hornung*, Bernadet Renckens, Sacha A. F. T. van Hijum, Ger T. Rijkers, Peter Schaap, Willem M. de Vos and Hauke Smidt

* These authors contributed equally to this work.

Abstract

The microbial communities in the small intestine are dependent on their capacity to rapidly import and ferment available carbohydrates. To maintain themselves in a complex, dynamic and highly competitive ecosystem, intestinal microbes have adapted or even specialized in foraging certain niche-specific substrates. Here we present the completely sequenced and annotated genome of *Romboutsia ilealis* CRIB^T, a natural and abundant inhabitant of the small intestinal tract of rats. *R. ilealis* CRIB^T possesses a single, circular chromosome of 2,581,778 bp that contains 2,351 predicted protein coding sequences, and a non-mobilizable plasmid of 6,145 bp carrying eight predicted protein coding sequences. Analysis of the genome revealed only limited ability to synthesize *de novo* amino acids and vitamins. However, multiple and partially redundant pathways for the utilization of a wide array of relatively simple carbohydrates could be identified. A whole-genome transcriptome analysis allowed pinpointing components of the key pathways involved in the degradation of glucose, L-fucose and fructo-oligosaccharide. These analyses reveal that *R. ilealis* CRIB^T is a flexible anaerobe that is adapted to a nutrient-rich environment in which carbohydrates and exogenous sources of amino acids and vitamins are abundantly available. Other features of ecological interest include the presence of urease and bile salt hydrolase encoding genes. This work shows how a combination of genome mining and functional analyses with single microbes can provide an overall insight in the genetic and functional potential of specific members of the intestinal microbiota.

Introduction

Intestinal microbes live in a complex and dynamic ecosystem, and in order to survive in this highly competitive environment they have developed close (symbiotic) associations with a diverse array of other intestinal microbes and with their host. This has left us with a complex network of host-microbe and microbe-microbe interactions in which the intestinal microbes and the host co-metabolise many substrates [4, 82]. In addition to competition for readily available carbohydrates in the diet, intestinal microbes are able to extract energy from dietary polysaccharides that are indigestible for the host [28]. Furthermore, intestinal microbes can utilize host-derived secretions (e.g. mucus) as substrates for metabolic processes [68]. In turn, the metabolic activities of the intestinal microbes result in the production of a wide array of compounds of which some are important nutrients for the host. For example, short chain fatty acids (SCFA), the main end products of bacterial fermentation in the gut, can be readily absorbed by the host and further metabolized as energy sources [26, 50]. All together the metabolic activity of the intestinal microbiota has a major impact on health of the host, and recent studies have indicated an important role for microbial activity in (human) diseases such as inflammatory bowel disease, irritable bowel syndrome and obesity [71] (see **Chapter 2** [33] as well).

The wide array of microbial genes present in the intestinal tract in addition to the host's own genome provides insight in the complex network of possible host-microbe and microbe-microbe interactions. To this end, it has been estimated that together the human intestinal microbes contribute ~150 times more genes than the human genome [70]. To be able to maintain themselves in an ecosystem such as the intestinal tract, microbes have adapted or even specialized in foraging certain niche-specific substrates. However, spatial and temporal heterogeneity in community composition and activity along the length of the intestinal tract is still poorly understood. In order to unravel the functional contribution of specific intestinal microbes to host physiology and pathology, we have to understand their metabolic capabilities. However, up to this date it is still difficult to connect functionality in an ecosystem to specific sets of genes and in turn to individual microbial species, and vice versa. A combination of genome mining and functional analyses with single microbes or simple and defined communities can provide an overall insight in the genetic and functional potential of specific members of the intestinal microbial community [37, 57, 105].

Little is known about the intestinal microbes adapted to the small intestine [14, 95, 106]. The small intestine is a nutrient-rich environment, and previous studies have shown that the microbial communities in the (human) small intestine are driven by the rapid uptake and conversion of simple carbohydrates [107]. In addition, community composition and activity in the small intestine is largely determined by host digestive fluids such as gastric acid, bile and pancreatic secretions. In-depth genomic analysis on *Streptococcus* isolates of small intestinal origin has shown that these microbes are adapted to a highly dynamic

environment [94]. Here we present the completely sequenced and annotated genome of *Romboutsia ilealis* CRIB^T which was recently isolated from the small intestinal tract of rats (**Chapter 4** [32]). It was found to be a natural inhabitant of the rat small intestine, specifically of the ileum, and a correlation with improved health status of the rats was observed in an experimental model of acute pancreatitis (**Chapter 3** [33]). We performed a genome-guided physiological analysis of the small intestinal inhabitant *R. ilealis* CRIB^T and provide a general overview of the metabolic capabilities and nutritional potential of this organism. These examinations revealed that *R. ilealis* CRIB^T is a flexible anaerobe which is adapted to degrade a wide range of relatively simple carbohydrates and contains potential mechanisms to survive in the competitive small intestinal environment, including bile salt hydrolase and urease activity.

Materials and methods

Strain and growth conditions

R. ilealis CRIB^T (DSM 25109) was routinely cultured in CRIB medium at 37°C as previously described (**Chapter 4** [32]). Genomic DNA was extracted from overnight grown cultures. Cells were pelleted by centrifugation at 9400 × g for 10 min and directly used for DNA extraction. DNA extraction was performed as previously described [94]. DNA quality and concentrations were determined by NanoDrop (Thermo Scientific) spectrophotometric analysis and by electrophoresis in a 1.0% (w/v) agarose gel. DNA was stored at 4°C until subsequent sequencing.

Genome sequencing and assembly

Genome sequencing of *R. ilealis* CRIB was done using 454 Titanium pyrosequencing technology (Roche 454 GS FLX), as well as Illumina (Genome Analyzer II and HiSeq2000) and PacBio sequencing (PacBio RS). The total sequence data amounted to 234,223 reads from the pyrosequencing, 34,807,392 Illumina paired-end reads (50 bp long), 9,886,736 Illumina mate-pair reads (4kb insert, 50 bp long) and 3,699 PacBio circular consensus sequencing (CCS) reads [46]. Mate-pair data was generated by BaseClear (Leiden, the Netherlands). All other data were generated by GATC Biotech (Konstanz, Germany).

For genome assembly, the size of the genome was predicted with KMERSPECTRUM-ANALYZER (downloaded at 15.08.2013) [104] The assembly was done in parallel using two different assemblers. Ray v2.2 [13] was used for the untrimmed Illumina paired-end dataset together with the Illumina mate-pair and PacBio CCS datasets. Default settings were used with the exception of increasing the k-mer length to 35 bp. In addition, Edena v3.130110 [38] was used on the Illumina paired-end and mate-pair datasets using default settings. The contigs obtained with both assemblers were merged using Zorro, the masked assembler (http://lge.ibi.unicamp.br/zorro/; release 13.04.2011), which relies on alignments with Nucmer [21] to determine duplications and misassembled contigs. The

merge was done with default settings, with input of one of the paired-end files of the Illumina paired-end dataset. Scaffolding was done with the merged contigs using Opera v1.2 [31] using default settings besides the usage of bowtie v1.0.0 [52], and the paired- end data as input, and afterwards further scaffolding was done with the mate-pair dataset using SSPACE v2.0 [11]. After the scaffolding step with opera, all contigs were discarded that either had a length of less than 100 bp or which mapped to another region of a bigger contig with 99% identity over 98% of its length. Default settings were used together with an expected insert size of 4,000 nucleotides and an error of 0.5. To close the gaps in the resulting assembly, GapFiller v1.11 [12] was used with the Illumina paired-end and mate-pair reads. Insert size for the Illumina paired-end data was estimated by mapping the raw Illumina paired-end reads back to both the initial assemblies using Bowtie2 v2.0.6 [51] and the CollectInsertSizeMetrics utility from the PicardTools package v1.94 (http://picard.sourceforge.net/). Default settings were used with exception of increasing the number of iterations to 40. Next, two further rounds of scaffolding using SSPACE v2.0 [11] and gapfilling using GapFiller v1.11 [12] were performed as described above. After the second round of gapfilling, all scaffolds with a length of less than 500 bp were discarded. As last step in the assembly, Pilon v1.4 (http://www.broadinstitute.org/software/pilon/) was used for quality assurance using Illumina paired-end and PacBio CSS reads mapped to the assembled genome with Bowtie2 v2.0.6 [51] using default settings. Afterwards a final round of gapfilling was performed using GapFiller v1.11 [12].

Assembly quality control

To control the quality of intermediate and final assemblies, intermediate steps were annotated, and all predicted proteins were compared to the predicted proteome of *R. ilealis* CRIB^T itself by BLASTP analysis. If no unlikely duplications (multiple co-located proteins with more than 95% identity, which are unlikely to be of biological origin, in contrast to e.g. transposases) as a result of one of the last steps in the assembly were identified, the assembly was used for further refinement. To exclude that any duplication was missed during the annotation, also a TBLASTX search of protein coding sequences against the genome was performed. Additional quality checks were done on duplications of single copy genes [49], completeness of rRNA operons and the presence of all tRNAs. To control the intermediate steps for possible loss or duplication of genomic material, the Mauve aligner [20] was used to compare the assemblies resulting from the different steps described above against each other, as well as the BLAST Ring Image Generator [1].

Genome annotation

Annotation was carried out with an in-house pipeline. Prodigal v2.5 was used for prediction of protein coding DNA sequences (CDS) [41], InterProScan 5RC7 for protein annotation [40], tRNAscan-SE v1.3.1 for prediction of tRNAs [59] and RNAmmer v1.2 for the

prediction of rRNAs [48]. Additional protein function predictions were derived via BLAST identifications against the UniRef50 [89] and Swissprot [93] databases (download August 2013). Afterwards the annotation was further enhanced by adding EC numbers via PRIAM version March 06, 2013 [17]. Non-coding RNAs were identified using rfam_scan.pl v1.04, on release 11.0 of the RFAM database [15]. CRISPRs were annotated using CRISPR Recognition Tool v1.1 [10]. CDS were assigned to clusters of orthologous groups (COGs) via bidirectional best hit [69] against the COG database [90] applying an e-value cut-off of 0.0001. A further step of automatic curation was performed, by weighting the annotation of the different associated domains, and penalizing uninformative functions (e.g. 'Domain of unknown function'), and prioritizing functions of interest (e.g. domains containing 'virus', 'phage', 'integrase' for phage related elements; similar procedure for different other functions). Pseudogenes were identified by manual curation of genes of interest.

Metabolic modelling

Pathway tools v18.0 [53] was used on the annotation to build a genome-scale metabolic model, which was manually curated with the built in curation tools. Afterwards a flux balance analysis (FBA) was performed with the integrated FBA tool. The necessary biomass components were obtained from related literature of genome-scale metabolic models for *Mycoplasma genitalium* [88] and *Staphylococcus aureus* [36]. Pathways for the production of essential cofactors were manually checked using the CoFactor database [27]. After testing a set of possible nutrients to satisfy the basic biomass production, also all carbohydrates previously tested *in vitro* were tested in the FBA, to verify whether they could serve as energy and carbon source.

Carbohydrate growth experiment

R. ilealis CRIB^T was grown in a basal bicarbonate-buffered medium [87] supplemented with 16 g/L yeast extract (BD) and an amino acids solution as used for the growth of *Clostridium difficile* [43]. In addition the medium was supplemented with either 0.5% (w/v) D-glucose (Fisher Scientific), L-fucose (Sigma-Alderich) or fructo-oligosaccharide (FOS) P06 (DP 2-4; Winclove Probiotics). The final pH of the medium was 7.0. Cultures were three times transferred on the respective carbohydrate before start of the experiment. For each condition triplicate cultures were set up. As control, basal medium (without an additional carbohydrate) was inoculated with cells preconditioned on each of the carbohydrates (one culture for each carbohydrate). For RNA-seq analysis, the cells were harvested in mid-exponential phase (OD_{600nm} = 0.25-0.55, ~8-10h incubation) (see Table S1 in the supplemental material). Cell pellets of 2 mL cultures were used for RNA purification. Cells were collected by centrifugation at 9,400 × g for 10 min. at 4°C and cell pellets were stored at -80°C until RNA purification.

Growth on mucin was examined by supplementation of the basal medium described

above with 0.25% (v/v) commercial hog gastric mucin (Type III; Sigma-Alderich), purified by ethanol precipitation as described previously [63].

For fermentation product analysis samples were obtained before inoculation, in mid-exponential phase (~8-10h incubation) and in stationary phase (24h incubation) (Table S1). Carbohydrate degradation and short-chain fatty acid production was determined by high-performance liquid chromatography (HPLC) using a Metacarb 67H column (Varian).

RNA purification and RNA-seq library preparation

Cells were enzymatically lysed using TE-buffer (100mM TRIS and 50mM EDTA, pH 8.0) supplemented with 160,000 U lysozyme and 100 U mutanolysin (Sigma-Alderich) during 30 min incubation at 37°C. Total RNA was isolated using the RNeasy Mini Kit (QIAGEN), and genomic DNA was removed by on-column DNase digestion step during RNA purification (DNase I, Roche). Yields and RNA-qualities were assessed using the Experion™ RNA StdSens Analysis Kit in combination with the Experion™ System (Bio-Rad Laboratories, Inc.). Depletion of rRNA was performed using the Rib-Zero™ Kit for bacteria (Epicentre Biotechnologies) according to the manufacturer's instructions. Success of the rRNA depletion step was checked using the Experion™ RNA StdSens Analysis Kit in combination with the Experion™ System. Library construction for whole-transcriptome sequencing (RNA-seq) was done by use of the ScriptSeq™ v2 RNA-seq Library Preparation Kit in combination with ScriptSeq™ Index PCR primers (Epicentre Biotechnologies) according to the manufacturer's instructions.

Whole-genome transcriptome analysis

The barcoded cDNA libraries were pooled and sent to GATC (Konstanz, Germany) where 150 bp sequencing was performed on one single lane using the Illumina HiSeq2500 platform in combination with the TruSeq Rapid SBS (200 cycles) and TruSeq Rapid SR Cluster Kits (Illumina).

rRNA reads were removed with SortMeRNA v1.9 [45] and all included databases. Adapters were trimmed with cutadapt v1.2.1 [62] using default settings except for an increased error value of 20% for the adapters. Quality trimming was performed with PRINSEQ Lite v0.20.0 [80] with a minimum sequence length of 40 bp and minimum quality of 30 on both ends of the read and as mean quality. All reads with non-IUPAC characters were discarded as were all reads containing more than three Ns. Reads were mapped to the genome with Bowtie2 v2.0.6 [51] using default settings. Details on the RNA-seq raw data analysis can be found in Table S2 in the supplemental material. BAM files were converted with SAMtools v0.1.18 [56] and genome coverage was calculated with BEDTools v2.17.0 [72]. For quality control the RNA-seq datasets were pooled per condition and assembled with Ray v2.3 [13] using a kmer-value of 71.

Gene expression abundance estimates and differential expression analysis was

performed using Cuffdiff v2.1.1 [92] with default settings. Differentially expressed genes were determined by pairwise comparison of a given condition to the other three conditions for a total of six pairwise comparisons. Genes were considered significantly differentially expressed when they showed a ≥1.5 log2(fold change) in any of the conditions with a false discovery rate (FDR)-corrected *P* value (q value) ≤0.05 (see Tables S3-S6 in the supplemental material). Principal component analysis was performed with Canoco 5.0 [91] on log-transformed gene transcript abundances using Hellinger standardization. Gene expression heatmaps were generated based on gene transcript abundances using R v3.1.0 and R-packages svDialogs and gplots.

Nucleotide sequence accession number

All related data has been deposited at the European Nucleotide Archive. The raw reads for the genome of *R. ilealis* CRIB^T can be accessed via the accession numbers ERR366773, ERX397233, ERX397242 and ERX339449. The assembly can be accessed under accession numbers LN555523-LN555524. The RNAseq data has been deposited under accession numbers ERS533849- ERS533861.

Results

General genome features

R. ilealis CRIB^T contains a single, circular chromosome of 2,581,778 bp and a plasmid of 6,145 bp (Table 1 and Figure 1). The chromosome contains 2,351 predicted protein CDS, of which 321 were annotated as hypothetical and for 91 only a domain of unknown function could be assigned. The plasmid carries eight predicted protein CDS, of which none has an obvious (metabolic) function. Furthermore, it appears to be a non-mobilizable plasmid, given that it lacks any mobilization associated genes. The overall G+C content of the genome is 27.9%, which is similar to the G+C content of 28.1 mol% previously determined for *R. ilealis* CRIB^T by HPLC methods (**Chapter 4** [32]).

With a total of 14 copies of the 16S rRNA gene, *R. ilealis* CRIB^T is among the species with the highest number of 16S rRNA gene copies reported up to this date [55]. Not all 16S rRNA gene containing operons in *R. ilealis* CRIB^T have the conserved 16S-23S-5S rRNA operon structure. This has been reported for other genomes containing multiple rRNA operons as the result of duplications [8, 81]. The individual rRNA genes share more than 98% identity to each other, except for one 5S rRNA, which shows a higher degree of identity to an rRNA gene from *C. difficile*. This latter 5S rRNA is located in an rRNA operon that contains two 5S genes, a structure which also has been observed in other microbes such as *Escherichia coli* [3]. The current assembly contains three gaps, all of which are located within rRNA operons. Based on the Illumina mate-pair data, one gap was estimated to have a length of 30 bp and is located within a 23S rRNA gene, another gap was estimated to have a length of 2,675 bp and is located between a 23S and 5S rRNA gene, and the

last gap is of unknown size, but contains most likely a 23S rRNA gene, as it is positioned between a 16S and 5S rRNA gene. As the gaps are located in highly similar ribosomal operons, it was not possible to design unique PCR primers to span the gaps.

A COG category could only be assigned to 1,647 of the CDS (70%) including 372 CDS (16%) assigned to the categories R (general function prediction only) and S (function unknown) (see Figure S1 in the supplemental material). This low coverage in function prediction is most likely due to the fact that the COG categorization has not been updated since 2003. With InterProScan a predicted function could be assigned to 82% of the CDS. Based on the InterPro and PRIAM classifications, an EC number could be assigned to more than 500 CDS.

Mobile genetic elements identified in the chromosome include different transposons, insertion sequence (IS) elements and phage-related proteins (Figure 1). The chromosome of *R. ilealis* CRIB^T contains one apparently complete prophage genome and several remnants. Overall, the amount of phage-related genes in the genome *R. ilealis* CRIB^T makes up 2.5% of the CDS in the whole genome. In comparison, the genome of *C. difficile,* that has been described as highly mobile, has 11% of its CDS consisting of mobile genetic elements, mainly in the form of transposons [84].

A gene cluster encoding a CRISPR-Cas system (clustered regularly interspaces short palindromic repeats and CRISPR-associated (Cas) proteins), which are known to be responsible for the prokaryotic adaptive immunity to phage infections [96, 103], was identified in the genome of *R. ilealis* CRIB^T (Figure 1). Based on the set of Cas proteins, the CRISPR-Cas system of *R. ilealis* CRIB^T can be classified to the type I-B according to the classification of Makarova *et al.* [60]. The set of 71 CRISPR repeat sequences found in the direct neighbourhood of the Cas genes is identical to CRISPR repeat sequences found in the genomes of multiple *C. difficile* strains.

Table 1. General features of the *R. ilealis* CRIB^T genome.

	Chromosome	Plasmid
Size (bp)	2,581,778	6,145
G+C content (%)	27.9	29.3
Protein CDS	2,351	8
Pseudogenes	12	0
Coding density	1.10	1.02
Average gene size (bp)	899	531
rRNA genes		
16S rRNA genes	14	0
23S rRNA genes	14*	0
5S rRNA genes	14	0
tRNAs	109	0
ncRNAs	28	0
CRISPR repeats	1*71	0

* An additional 23S rRNA gene sequence is expected in one of the gaps.

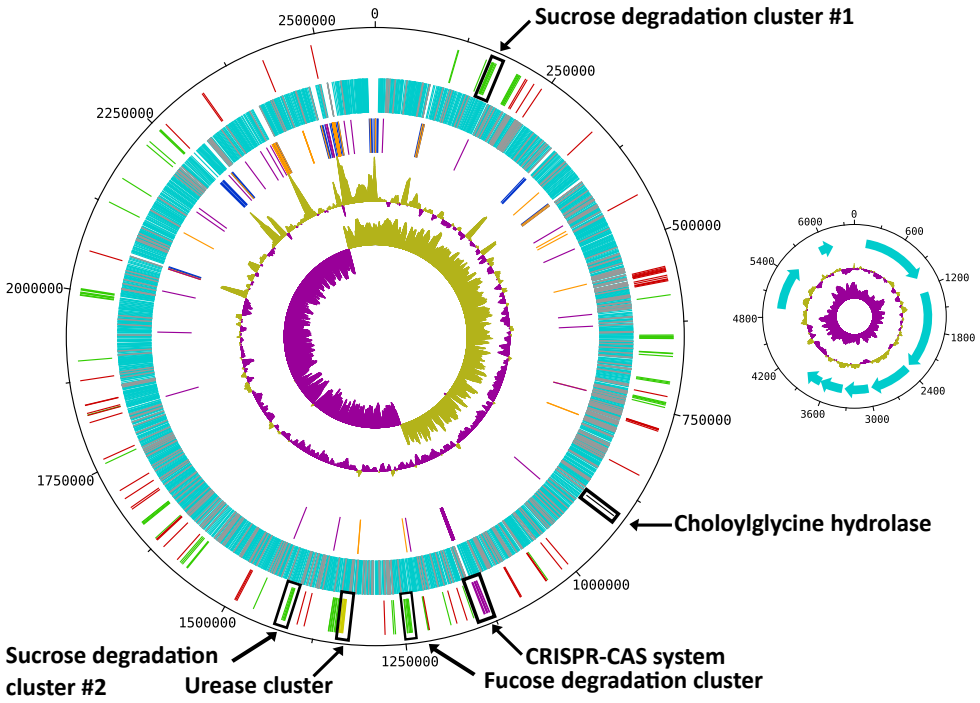

Figure 1. Circular map of the *R. ilealis* CRIB^T genome; both chromosome and non-mobilizable plasmid are shown. For the chromosome tracks from inside to outside are as follows: 1, GC skew; 2, G+C content; 3, RNAs [rRNAs (blue), tRNAs (orange) and ncRNAs (purple)]; 4, all predicted protein CDS [with predicted function (light-blue), hypothetical proteins and proteins to which only a domain of unknown function could be assigned (grey)]; 5, genes or gene clusters of interest [(mobile genetic elements (red), Cas proteins (pink), urease gene cluster (yellow), choloylglycine hydrolase (black), gene clusters involved in carbohydrate utilization (green)]. For the plasmid tracks from inside to outside are as follows: 1, GC skew; 2, G+C content; 3, all predicted CDS.

5

General metabolic pathways

Genome sequence analysis of *R. ilealis* CRIB^T revealed the presence of a complete set of enzymes for the glycolysis pathway. In line with the anaerobic lifestyle of the organism, the enzymes for the oxidative phase of the pentose phosphate pathway could not be detected. Furthermore, the genes that encode the enzymes involved in the tricarboxylic acid cycle were lacking. The metabolic model confirmed that *R. ilealis* CRIB^T is a mixed acid fermenter as was previously reported (**Chapter 4** [32]). End products of fermentation are a mixture of acetate, formate, lactate and ethanol, with the possibility of gas formation (CO_2 and H_2). In addition to ethanol, which can be produced during mixed acid fermentation, 1,2-propanediol was predicted to be produced via the L-fucose degradation pathway. No other solvents were predicted to be produced. The fermentation end products formate, acetate and lactate are produced as pyruvate derivatives. Propionate production was not predicted, however, low amounts of propionate production were observed repeatedly

during *in vitro* growth. Neither the succinate, nor the acrylate or the propanediol pathway, which are three well-known pathways for propionate production in the intestinal tract [73], could be identified in the genome of *R. ilealis* CRIB^T. It might be possible that propionate is produced together with formate from threonine fermentation by the activity of acetate kinase in the final step instead of a dedicated propionate kinase, as these enzymes have very similar catabolic mechanisms and carboxylic acid binding sites [42]. The ability to produce butyrate was suggested based on the annotation of a butyrate kinase and a phosphate butyryl transferase. However, the unambiguous function prediction of these enzymes based on gene homology is difficult to achieve as they share a high sequence identity with enzymes involved in acetate formation [44, 101]. Moreover, both genes were found in different locations in the genome whereas they are usually found together in one operon in genomes of canonical butyrate-producing clostridia like *C. acetobutylicum* [67]. Furthermore, the rest of the enzymes that constitute known pathways for butyrate production [100] are not present. In conclusion, a complete pathway for butyrate production, which is present in close relatives of *R. ilealis* CRIB^T (e.g. *C. difficile),* appears to be absent in *R. ilealis* CRIB^T, in line with the lack of butyrate production *in vitro*.

Fermentation is probably the main source of energy for *R. ilealis.* However, potential for anaerobic respiration was also suggested based on the presence of a sulfite reductase gene cluster (CRIB_1284-CRIB_1286) of the dissimilatory *asrC*-type [24]. Similar siroheme-dependent sulfite reductases are found in many close-relatives of *R. ilealis* such as *Intestinibacter bartlettii*, *Clostridium sordellii* and *C. difficile* [19]. Sulfite reduction by *R. ilealis* CRIB^T, and close relatives, has been previously demonstrated *in vitro* (**Chapter 4** [32]). For *R. ilealis* CRIB^T, increased growth yield and metabolite production was observed in the presence of sulfite (data not shown). In the intestinal tract sulfite is derived from food sources that contain sulfite as preservative, and it has been shown that neutrophils release sulfite as part of the host defence against microbes [64].

In addition to being sulfite reducers, some of the close relatives of *R. ilealis* (e.g. *C. difficile*) have been characterized as acetogens using the Wood-Ljungdahl pathway for autotrophic growth with CO_2 as electron acceptor [44]. *R. ilealis* CRIB^T also possesses the Wood-Ljungdahl cluster of genes, however, no homologue of known formate dehydrogenases encoding genes could be identified meaning that R. *ilealis* CRIB^T is likely not capable of acetogenesis.

Metabolism of growth factors and cofactors

Many genes encoding the enzymes required for amino acid biosynthesis appeared to be absent in *R. ilealis* CRIB^T. Complete pathways are present for the biosynthesis of aspartate, asparagine, glutamate, glutamine and cysteine, using carbon skeletons available from central metabolites or via conversion of other amino acids. Furthermore, a pathway for the production of selenocysteine, the 21st amino acid, was identified. Based on the

RFAM prediction not SECIS element encoding genes were identified. Two proteins were annotated as possible selenoproteins, but their exact function could not be determined.

It was predicted that L-alanine is produced from L-cysteine by cysteine desulfurase, an enzyme involved in the biosynthesis of iron-sulfur clusters, however, it is not known whether this pathway has a major role in generating the cellular supply of alanine. Within the pathways for the biosynthesis of glycine, lysine, methionine, phenylalanine, serine, threonine and tyrosine there were a few genes missing, possibly not yet identified in the current annotation. The pathways for biosynthesis of arginine, histidine, isoleucine, leucine, proline, tryptophan and valine were almost completely absent. The absence of genes to produce branched-chain amino acids (leucine, isoleucine and valine) is also reflected in the absence of branched chain fatty acids in the cell membrane of *R. ilealis*, which is characteristic for the genus *Romboutsia* (**Chapter 4** [32]). From these observations it can be concluded that *R. ilealis* depends on exogenous amino acids, peptides and/or proteins to fuel protein synthesis.

The dependency on an exogenous source of amino acids is reflected by the identification of multiple amino acid transporters, including an arginine/ornithine antiporter, multiple serine/threonine exchangers, a transporter for branched amino acids, and several amino acid symporters and permeases without predicted specificity. Interestingly, four genes predicted to encode amino acid permeases or exchangers were found adjacent to each other in the genome (CRIB_116-CRIB_119). They share >90% identity to each other on amino acid level suggesting a possible duplication event. In addition to amino acid transporters, one oligopeptide transporter was identified. Furthermore, numerous genes have been annotated as protease or peptidase. As most of them will be involved in intracellular processes (e.g. sporulation or germination) it is difficult to predict, whether they play a role in the use of exogenous protein or peptides to satisfy the need for amino acids.

R. ilealis CRIB^T appears to contain all genes for *de novo* purine and pyrimidine synthesis, as well as for the production of the coenzymes NAD and FAD via salvage pathways from niacin and riboflavin, respectively. While some organic cofactors can be produced by *R. ilealis* CRIB, including siroheme, it mainly relies on salvage pathways (e.g. for lipoic acid) or exogenous sources for the supply of precursors mainly in the form of vitamins (e.g. thiamin, riboflavin, niacin, pantothenate, pyridoxine, biotin, vitamin B12). The pathway for *de novo* folate biosynthesis is present via the pABA branch, however, the gene encoding dihydrofolate reductase, which is essential in both *de novo* and salvage pathways of tetrahydrofolate (THF), is lacking. For *de novo* vitamin B12 production, a gene cluster related to its synthesis could be detected, however, genes encoding several enzymes of the late cobyrinate biosynthesis are missing, based on which was concluded that *R. ilealis* CRIB^T is not able to synthesize vitamin B12 *de novo*.

Carbohydrate transport and metabolism

As previously reported *R. ilealis* CRIB^T is able to utilize a wide variety of carbohydrates (**Chapter 4** [32]). Previously good growth of *R. ilealis* on L-fucose, glucose, raffinose and sucrose has been described in addition to moderate growth on D-arabinose and D-galactose and weak growth on D-fructose, inulin, lactose, maltose and melibiose. Growth on L-fucose, fructose, galactose, glucose, lactose, maltose, melibiose, raffinose and sucrose was predicted from the genome-scale metabolic model as well. Besides the central glycolytic genes, which are present in a single genomic area, genes encoding the specific carbohydrate degradation enzymes are distributed throughout the genome in gene clusters together with their respective transporters and transcriptional regulator. No separate pathway for the use of D-arabinose could be predicted, however, the L-fucose degradation pathway is likely also used for D-arabinose utilization [54]. An arabinose transporter, similar to the maltose and sucrose transporters, could be identified. In addition to these carbohydrates, ribose utilization was predicted as well as a ribokinase was identified together with an ABC transporter system for ribose (CRIB_1680-CRIB_1684). This is in contrast, however, to the fact that *in vitro* no growth was observed on ribose. Furthermore, a gene cluster involved in the degradation of the host-derived carbohydrate sialic acid could be predicted (CRIB_613-CRIB_619) [2]. The structure of this gene cluster is similar to the one identified in *C. difficile* [66]. The ability to degrade the predominantly host-derived carbohydrates L-fucose and sialic acid suggest a role in the utilization of mucin, an abundant host-derived glycoprotein in the intestinal tract [22, 68]. However, no growth on mucin was observed (data not shown), which is in line with the lack of a predicted extracellular fucosidase and/or sialidase.

Other genes of ecological interest

A gene cluster encoding a urease, consisting of three subunits (*ureABC*), and a number of urease accessory genes were identified (CRIB_1381-CRIB_1388). The gene cluster identified in *R. ilealis* CRIB is very similar to the urease gene cluster in the genome of *C. sordellii* (see Figure S3 in the supplemental material), a species in which the urease activity is used to phenotypically distinguish *C. sordellii* strains from *C. bifermentans* strains [77]. Furthermore, a possible ammonium transporter (CRIB_1389) was identified in the genome of *R. ilealis* CRIB^T next to the urease gene cluster. Ureases are nickel-containing metalloenzymes that catalyse the hydrolysis of urea to ammonia and carbon dioxide, and thereby these enzymes allow microbes to use urea as nitrogen source by assimilation via glutamate. They are ubiquitous proteins occurring in diverse organisms [65]. In the intestinal environment, where urea is abundantly present [29], some bacteria use ureases to survive the acidic conditions in the upper part of the intestinal tract as urea hydrolysis leads to a local increase in pH [78].

Another gene of ecological interest is the predicted choloylglycine hydrolase. Proteins

within the choloylglycine hydrolase family are the bile salt hydrolases (BSHs), also known as conjugated bile acid hydrolases, that are widespread among Gram-positive and Gram-negative intestinal microbes [74]. They are involved in the hydrolysis of the amide linkage in conjugated bile salts, releasing primary bile acids. There is a large heterogeneity among BSHs, for example with respect to their substrate specificity. The BSH of *R. ilealis* CRIB^T was found to be the most similar to the one found in *C. butyricum*. Although the physiological advantages of BSHs for the microbes are not completely understood, it has been hypothesized that they constitute a mechanism to detoxify bile salts and thereby enhance bacterial colonization [19].

Metabolite and transcriptome analysis during growth on different carbohydrates

To study key pathways predicted to be involved in carbohydrate utilization and their regulation in more detail, a whole-genome transcriptome analysis was performed, focussing on four experimental conditions. Firstly, growth on glucose, a preferred substrate for many microbes present in the intestinal tract, was studied. Secondly the growth on fructans, oligo-/polysaccharides present in many food substances, was examined. Previously weak growth on inulin, a polysaccharide consisting of long chains of ß1→2 linked fructose units, was observed (**Chapter 4** [32]). For this study a shorter fructan (FOS P06, DP2-4) was chosen as growth on shorter fructans is ecologically likely more relevant for microbes living in the small intestine [107]. Thirdly, growth on L-fucose was examined as growth on this substrate was found to be unique for *R. ilealis* CRIB^T compared to other related microbes. Fourthly, *R. ilealis* CRIB^T was also grown in the basal medium in the absence of an additional carbon source for comparison (control condition). Sampling was done in mid-exponential phase (~8-10 h incubation; used for transcriptome analysis) and in stationary phase (24 h incubation) based on optical density and pH change measurements (growth characteristics of individual cultures can be found in Table S1), and sugar utilization and fermentation products were measured with HPLC (Table 2). The fact that in neither of the experimental conditions the supplied carbohydrates were depleted and metabolites were still produced during the sampling points at ~8-10 h and 24 h, confirmed that samples obtained for transcriptome analysis at ~8-10 h were taken during exponential growth. In the FOS cultures an accumulation of extra-cellular fructose was observed. As predicted from the metabolic model, growth on glucose resulted in the production of formate, acetate and lactate (Table 2).

Growth on FOS was slightly less than on glucose, however, after 24 h of growth the same fermentation products were observed in similar amounts (Table 2). Growth on L-fucose showed production of 1,2-propanediol instead of lactate. The fact that 1,2-propanediol was observed in one of the control cultures could be explained by the fact that an L-fucose grown culture was used as inoculum for this culture, leading to carry-over of minor amounts of metabolites.

Table 2. Fermentation end products of *R. ilealis* CRIB^T produced during growth on different carbohydrates (glucose, FOS or L-fucose) or in basal medium in the absence of an additional carbon source (control condition). Samples were obtained during mid-exponential phase (~8-10 h incubation; used for transcriptome analysis) and in stationary phase (24 h incubation). For the control cultures, fermentation products are shown for the individual cultures separating the carbohydrates used for preconditioning of the inoculum. For the three other conditions, values represent means of triplicate cultures with standard deviations. N.D., not detected.

	Formate		Acetate		Propionate		Lactate		1,2-propanediol	
	8-10 h	24 h	8-10 h	24 h	8-10 h	24 h	8-10 h	24 h	8-10 h	24 h
Control: basal medium										
(glucose inoc.)	3.2	7.7	2.0	6.2	2.0	2.2	N.D.	N.D.	N.D.	N.D.
(FOS inoc.)	4.5	9.2	2.4	7.4	2.4	2.9	N.D.	N.D.	N.D.	N.D.
(L-fucose inoc)	4.8	10.8	2.3	9.8	2.3	3.0	N.D.	N.D.	1.0	1.0
Basal medium + glucose (5% w/v)	4.4±1.2	28.2±4.3	1.0±0.9	16.3±2.2	1.0±0.9	1.3±0.1	N.D.	3.0±0.7	N.D.	N.D.
Basal medium + FOS (5% w/v)	4.7±0.6	27.3±2.5	1.4±0.0	17.7±1.4	1.4±0.0	1.6±0.1	N.D	2.5±0.3	N.D.	N.D.
Basal medium + L-fucose (5% w/v)	6.7±0.1	19.5±3.6	2.8±0.1	16.3±2.9	2.8±0.1	2.8±0.4	N.D.	N.D.	1.3±0.1	7.7±1.4

For the whole genome-transcriptome analysis of triplicate cultures grown in the four different conditions (i.e. a total of 12 cultures), a total of 159,250,634 reads (with size of 150 bp) were generated by RNA-seq (overview in Table S2). On average 13% of all reads were excluded as encoding rRNA, and of the remaining reads 20% was removed as adapters. On average, 93.27% of the remaining reads passed to the quality filtering and of these 97.9% could be mapped back to the genome.

Principal component analysis of the overall transcriptomes of the individual cultures showed that the cultures clustered by condition (Figure 2). Also for the control samples, the transcriptomes clustered together instead of with the cultures grown on the respective carbohydrates the inocula were preconditioned on, confirming that for pairwise comparison the individual transcriptomes could be grouped per condition.

To identify significantly differentially regulated genes, pairwise comparisons were done with cuffdiff using a cut off of ≥1.5 log2(fold-change) and q-value ≤0.05. During growth on glucose 157 genes were identified to be significantly upregulated compared to a least one of the other conditions (Table S3). Furthermore, growth on FOS and L-fucose resulted in the significant upregulation of 53 genes and 54 genes, respectively, compared to at least one of the other conditions (Table S4 and Table S5, respectively). During growth in the control condition (basal medium in the absence of an additional carbon source) this was the case for 80 genes (Table S6). Figure 3 shows a heat map of the significantly differentially regulated genes.

The gene cluster involved in glycolysis (CRIB_186-CRIB_191) was most abundantly expressed in the conditions that support the highest growth rates as determined by maximum cell density (glucose, followed by FOS; Figure 3). This was also reflected in

the fact that the differential expression of genes involved in replication (e.g. ribosomal proteins, proteins involved in cell wall biosynthesis and general cell division processes) were most strongly expressed during growth in the presence of glucose and to a lesser extent FOS. Other genes involved in the central sugar metabolic pathways (e.g. CRIB_1849, CRIB_140, CRIB_2223, and CRIB_105) were upregulated in these conditions, albeit not significantly differentially regulated. This suggests that these are more tightly regulated, probably because they are also involved in other processes than sugar degradation [18].

Altogether, the transcriptome of *R. ilealis* CRIB^T grown on FOS was very similar to its transcriptome when grown on glucose (Figure 2), with only 18 genes significantly upregulated during growth in the presence of FOS compared to glucose (Table S4). Apparent was the upregulation of the gene clusters involved in the transport and degradation of the respective sugars or their derivatives (Figure 3). In the presence of glucose the glucose-specific PTS system (CRIB_2017-CRIB_2018) was significantly upregulated, together with its associated transcriptional regulator (CRIB_2019). In the presence of FOS two clusters predicted to be involved in sucrose degradation (CRIB_148-CRIB_152 and CRIB_1458-1461) were significantly upregulated.

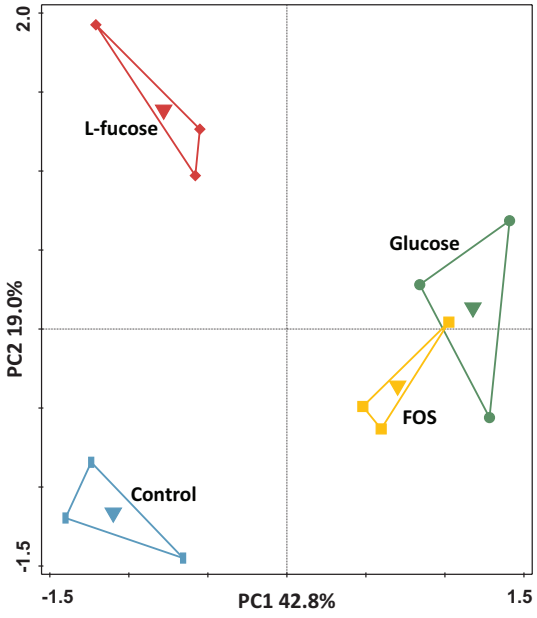

Figure 2. Principal component analysis of the transcriptomes of *R. ilealis* CRIB^T grown on different carbohydrates (glucose, FOS and L-fucose) or in the absence of an additional carbon source (control). First and second ordination axes are plotted, explaining 42.8% and 19.0% of the variability in the data set, respectively. Individual transcriptomes are color-coded by experimental condition: glucose (green circles), FOS (yellow squares), L-fucose (red diamonds) and control (blue rectangles). The experimental conditions were used as supplementary variables as well and could explain 62.9% of the variation.

gene cluster predicted to be involved in sucrose degradation (CRIB_1399-1400) was not significantly regulated during growth on FOS, however, it should be noted that these genes are located in a cluster related to melibiose metabolism and are most likely regulated by the transcriptional regulator in this cluster. In addition to the two sucrose degradation clusters, a transport cluster of unknown function (CRIB_1506-CRIB_1509) was upregulated during growth on FOS, albeit only significant compared to growth on glucose. During growth in the presence of L-fucose, the gene cluster predicted to be involved in L-fucose degradation (CRIB_1294-CRIB_1298) was significantly upregulated, including the gene encoding the corresponding transcriptional regulator (CRIB_1299). An overview of the main carbohydrate degradation pathways regulated in the different conditions is given in Figure 4.

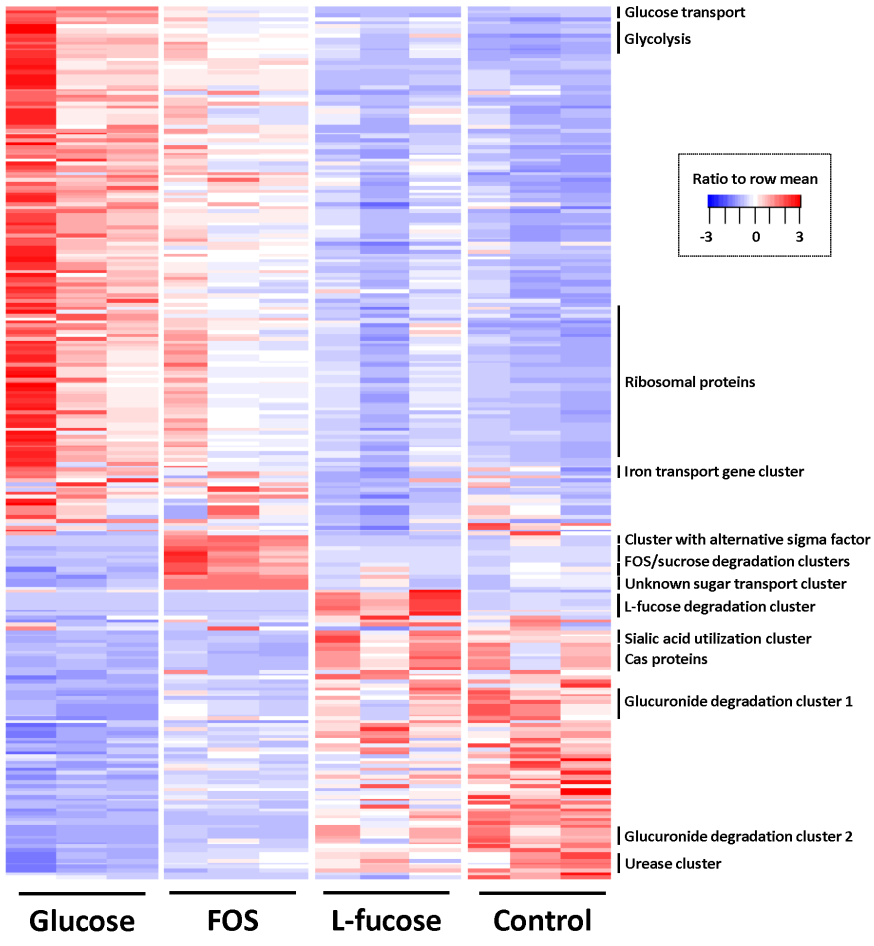

Figure 3. Heatmap of significantly differentially expressed genes in at least one of the four conditions (≥1.5 log2(fold change) and q value ≤0.05). Colour coding by ratio to row mean. Key gene clusters are indicated.

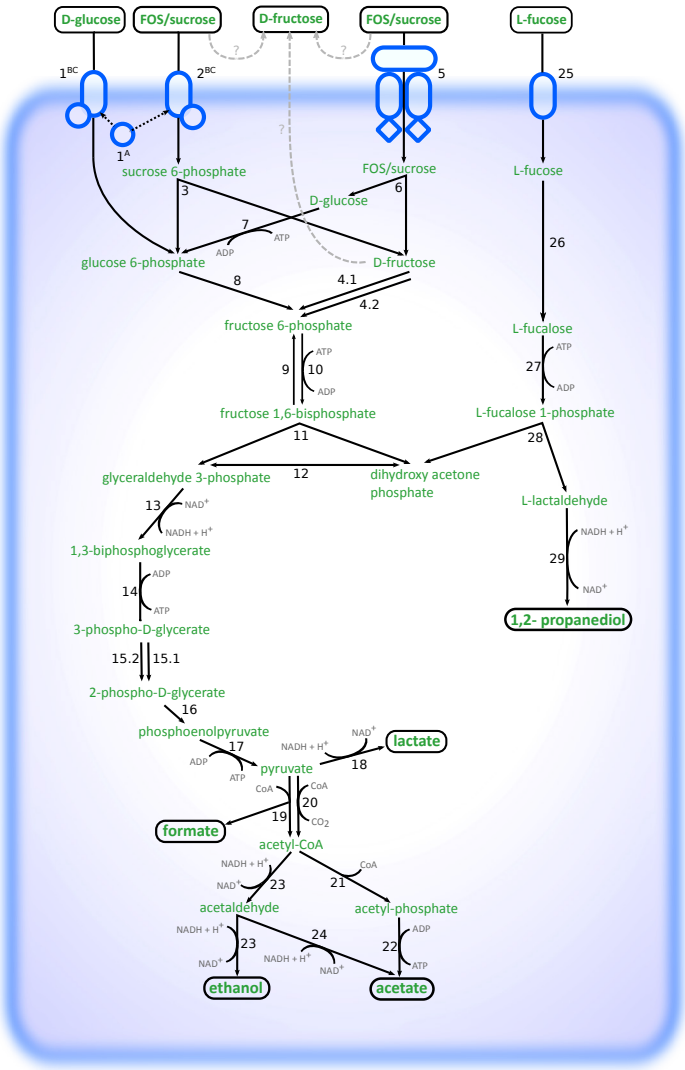

Figure 4. Schematic overview of the pathways involved in degradation of glucose, FOS and L-fucose in *R. ilealis* CRIB^{T}. 1^{A}; PTS system glucose-specific EIIA component (CRIB_2018); 1^{BC}, PTS system glucose-specific EIIBC component (CRIB_2017); 2^{BC}, PTS system sucrose-specific EIIBC component (CRIB_1461); 3, ß-fructofuranosidase with RDD family protein (CRIB_1459 and CRIB_1460); 4, fructokinase (CRIB_152 and CRIB_1458); 5, ABC-type transporter (CRIB_148-CRIB_150); 6, ß-fructofuranosidase (CRIB_151); 7, glucokinase (CRIB_1849); 8, glucose 6-phosphate isomerase (CRIB_140); 9, fructose 1,6-bisphosphatase (CRIB_45 and CRIB_2020); 10, 6-phosphofructokinase ; (CRIB_104); 11, fructose-bisphosphate aldolase (CRIB_2223); 12, triosephosphate isomerase (CRIB_189); 13, glyceraldehyde-3-phosphate dehydrogenase (CRIB_187); 14, phosphoglycerate kinase; 15, phosphoglycerate mutase (CRIB_1223) and 2,3-bisphosphoglycerate-independent phosphoglycerate mutase (CRIB_190); 16, enolase (CRIB_191); 17, pyruvate kinase (CRIB_105); 18, L-lactate dehydrogenase (CRIB_684); 19, formate acetyltransferase (CRIB_2141); 20, pyruvate-flavodoxin oxidoreductase (CRIB_2021); 21, phosphate acetyltransferase (CRIB_2171); 22, acetate kinase (CRIB_1927); 23, bifunctional aldehyde-alcohol dehydrogenase (CRIB_2231); 24, fatty aldehyde dehydrogenase (CRIB_2231); 25, L-fucose permease (CRIB_1294); 26, L-fucose isomerase (CRIB_1298); 27, L-fuculokinase (CRIB_1297); 28, L-fuculose phosphate aldolase (CRIB_1297); 29, lactaldehyde reductase (CRIB_1300).

During growth on glucose, L-lactate dehydrogenase (CRIB_684) was significantly upregulated, albeit not significant compared to growth on FOS. This enzyme catalyses the reduction of pyruvate resulting in the production of L-lactate and the reoxidation of the NADH formed during glycolysis. Only at time point 24 h lactate was observed (Table 2). This suggests that at time point ~8-10 h the cells were starting to regenerate NAD by upregulating the necessary genes. In the presence of L-fucose, NAD^+ regeneration is achieved via the reduction of lactaldehyde to 1,2-propanediol by lactaldehyde reductase (CRIB_1300), which was upregulated in the presence of L-fucose together with the L-fucose degradation gene cluster. In the spent medium of L-fucose grown cells, 1,2-propanediol was already seen at time point ~8-10 h whereas no lactate production was observed. Another way to regenerate NAD^+ is to reduce pyruvate to ethanol (Figure 4). In the presence of both glucose and FOS an upregulation was seen of the bifunctional aldehyde/alcohol dehydrogenase (CRIB_2231), which converts acetyl-CoA to ethanol. However, in none of the samples ethanol was measured by HPLC analysis. Instead of through fermentation, NAD^+ can also be regenerated via respiration. The anaerobic sulfite reductase gene cluster (CRIB_1284/1286) was not significantly regulated, however, also no additional sulfite was provided in the medium.

Remarkable was the significant upregulation of a gene cluster related to iron transport (CRIB_892-CRIB-898) during growth on glucose and FOS compared to growth on L-fucose. The significance of this gene cluster for carbohydrate utilization is not known. However, several enzymes could be identified in the genome of *R. ilealis* $CRIB^T$ that use different forms of iron as cofactor, for example the hydrogenases involved in hydrogen metabolism [16], several ferredoxins, and the L-threonine dehydratase (CRIB426) that was significantly upregulated during growth on L-fucose. As multiple transporters involved in the transport of iron compounds were predicted, it might also be that the uptake of iron provides a competitive advantage to other microbes that are dependent on iron for respiration and other metabolic processes [47].

During growth on FOS, a small gene cluster (CRIB_601-CRIB_603) that includes a gene encoding an alternative sigma factor was significantly upregulated. Interestingly, this was also apparent in the control culture that was inoculated with FOS-preconditioned cells. This suggests that in the presence of FOS (or its derivatives sucrose or fructose) transcription is also regulated by RNA polymerase promoter recognition.

In the intestinal environment *R. ilealis* CRIB will encounter a wide array of carbohydrates that are either continually or transiently present. Prioritization of carbohydrate utilization is partly achieved at a transcriptional level by the selective expression of genes. The primary mechanism by which bacteria regulate the utilization of non-preferred carbohydrates in the presence of preferred carbon sources is known as carbon catabolite repression (CCR), a hierarchical system for coordinating sugar metabolism [23]. The fact that, compared to glucose and FOS, L-fucose is utilized by a pathway that does not directly involve fructose-

1,6-bisphosphate, a key metabolite in the regulation of CCR of Gram-positive bacteria, made it possible to study CCR by either glucose or FOS. The transcriptome analysis suggests that some genes and operons in *R. ilealis* CRIB^T were indeed subject to CCR in response to the presence of glucose, a preferred substrate. For example, two gene clusters predicted to be involved in hexuronate metabolism (CRIB_649-CRIB_652 and CRIB_2244-CRIB_2249), pathways that make the use of D-glucuronate and D-galacturonates as sole carbon source possible, were significantly upregulated during growth in the presence of L-fucose compared to growth on glucose (Table S5). In addition, the gene cluster predicted to be involved in sialic acid utilization (CRIB_613-CRIB_616) was downregulated in the presence of glucose as well. Furthermore, when comparing the expression of the gene cluster involved in L-fucose degradation during growth on glucose relative to the growth in the absence of a carbon source (control condition), also this gene cluster appeared to be under CCR in the presence of glucose (Table S5). These results suggest that in *R. ilealis* CRIB^T multiple gene clusters that are involved in the use of alternative carbon sources are subject to CCR. The urease encoding gene cluster (CRIB_1381-CRIB_1388) was significantly upregulated when grown in the absence of an additional carbon source. The fact that this was significant compared to growth on glucose suggests possible CCR of the urease gene cluster. However, upregulation of this gene cluster in the absence of an exogenous carbon source might also be a possible mechanism.

Discussion

Previously it was observed by 16S rRNA gene sequence-based analysis that *R. ilealis* CRIB^T is a dominant member of the small intestine microbiota in rats, especially in the ileum (**Chapter 3** [33]). The genomic and transcriptional analysis of *R. ilealis* CRIB^T reported here provides new insights in the genetic and functional potential of this inhabitant of the small intestine. Genomic analysis revealed the presence of metabolic pathways for the utilization of a wide array of 'simple' carbohydrates in addition to a multitude of carbohydrate uptake systems that included a series of PTS systems, carbohydrate specific ABC transporters, permeases and symporters. Considering the habitat of *R. ilealis* CRIB^T, we chose to focus on key pathways involved in the utilization of specific diet- and host-derived carbon sources by whole-genome transcriptome analysis.

In the intestinal tract, the diet-derived carbohydrates that the host is unable to digest are important sources of energy for many microbes. Here we examined the growth of *R. ilealis* CRIB^T on FOS, a relatively simple oligosaccharide that is indigestible by the host. The transcriptome of *R. ilealis* CRIB^T grown on FOS was very similar to its transcriptome when grown on glucose, a monosaccharide used by the majority of microbes present in the intestinal tract. This is not surprising considering that glucose is in addition to fructose one of the two subunits present in FOS. Differential gene expression analysis demonstrated the apparent FOS-induced upregulation of two separate gene clusters

that were predicted to be involved in sucrose degradation. Both gene clusters contain an associated carbohydrate uptake system, one a PTS system, the other an ABC-type transporter. Both PTS systems and ABC transporters are reported to be involved in fructan utilization in other (intestinal) microbes, such as *Lactobacillus acidophilus* [6], *Lactobacillus paracasei* [34], *Lactobacillus plantarum* [79] and *Streptococcus pneumoniae* ([58]; for a review see [30]). Remarkable was the accumulation of fructose in the culture supernatant during growth of *R. ilealis* CRIB^T on FOS. A simple explanation could be the release of fructose or beta-fructofuranosidase activity after cell lysis. The fact that *R. ilealis* only grows weakly on D-fructose (**Chapter 4** [32]) and no transporter specific for fructose could be identified as in close relatives that are able to grow on D-fructose, could explain fructose accumulation. Another explanation could be extracellular degradation of FOS, followed by import of sucrose and/or glucose into the cell. Fructan degradation by extracellular enzymes is described for other (intestinal) microbes [97]. However, no extracellular fructansucrase or glucansucrase could be predicted. Furthermore, no new candidates for this activity could be identified via the differential gene expression analysis described here. One possible candidate could be, however, the predicted beta-fructofuranosidase present in the PTS system-containing sucrose degradation gene cluster. Next to the beta-fructofuranosidase-encoding gene, a gene was found to which no function could be assigned, but that is predicted to have a transmembrane region and a domain which could be involved in transport. Given that both loci overlap by a few nucleotides, and that the overlap is within a homopolymer region, it is possible that both loci form one protein due to ribosomal slippage on the homopolymer [85]. This could possibly lead to an external membrane-bound enzymatically active protein, which would explain the accumulation of fructose. Future studies with mutant strains might shed more light on the specific contribution of the two predicted sucrose degradation gene clusters to the degradation of FOS, or even longer fructans (inulin), in *R. ilealis* CRIB^T. In addition, future studies will also have to reveal whether FOS can stimulate the presence of *R. ilealis in vivo*.

It was previously reported that unlike other members of the family *Peptostreptococcaceae*, *R. ilealis* CRIB^T is able to grow on L-fucose, a predominantly host-derived carbon source (**Chapter 4** [32]). The whole-genome transcriptome analysis confirmed the presence of a functional L-fucose degradation pathway, similar to the pathways previously identified in other intestinal inhabitants such as *E. coli* [5], *Bacteroides thetaiotaomicron* [39] and *Roseburia inulinivorans* [83]. By gene sequence homology a similar pathway is found in *Clostridium perfringens* and the more closely related species *C. sordellii* (see Figure S2 in the supplemental material). L-fucose is a common sugar present within the intestinal environment, as it is a monosaccharide that is an abundant component of many N- and O-linked glycans and glycolipids produced by mammalian cells, including the fucosylated glycans that are found at the terminal positions of mucin glycoproteins [7]. It has been observed that fucosylated mucin glycoproteins are especially found in the (human)

ileum [75, 76]. For both intestinal commensals and pathogens the ability to utilize L-fucose has been demonstrated to provide a competitive advantage in the intestinal environment [39, 86]. In *R. ilealis* all enzymes for L-fucose degradation are present in one cluster, however, no fucosidase-encoding gene could be identified, which means that *R. ilealis* is not able to release L-fucose units from fucosylated glycans (e.g. mucin) by itself. Hence, in the intestinal environment *R. ilealis* is dependent on the free L-fucose monosaccharides released by other microbes. Furthermore, a gene cluster involved in degradation of sialic acid [2, 98, 99] was predicted from the genome, but no extracellular sialidase could be identified similar to *C. difficile* [66]. This suggests that also for sialic acid, a common residue found in mucin glycoproteins, *R. ilealis* CRIB^T seems to be dependent on the activity of other microbes.

Interesting was the identification of a urease gene cluster in *R. ilealis* CRIB^T, which appeared to be induced in carbon source limiting circumstances. Urea in the intestinal tract is derived from the breakdown of amino acids. *Helicobacter pylori* is a well-known example where urease activity contributes to the survival of the bacterium in the acidic environment of the stomach [61]. For some of the urease-positive bacteria, this enzyme has been shown to act as a virulence factor as it is responsible for urea hydrolysis that leads to increased pH and ammonia toxicity [78]. However, for intestinal bacteria ureases can probably function as colonization factors as well, as they contribute in general to acid resistance and thereby play a role in gastrointestinal survival. Urea is released into all parts of the intestinal tract via diffusion from the blood, but it has been reported that entry into the small intestine via pancreatic excretions and bile is a main route of entry [9]. We have not been able to demonstrate urease activity in *R. ilealis* CRIB^T during growth in the presence of urea (data not shown). However, different mechanisms for the expression of urease have been identified in other microbes: constitutive, inducible by urea, or controlled by nitrogen source availability [65]. For *C. perfringens* for example, the urease activity, which is plasmid borne, was shown to be only expressed in nitrogen-limiting conditions [25]. The increased urease gene expression by *R. ilealis* CRIB^T observed in the control condition, in the absence of an additional carbohydrate, suggests an alternative mechanism for regulation of urease gene expression.

For future in-depth analyses on the metabolism of *R. ilealis* CRIB^T, growth in a chemically defined (minimal) medium would be preferred over media containing undefined components such as yeast extract. In the past growth in chemically defined media has been successfully achieved for *Clostridia* species, including *C. perfringens* [35], *Clostridium botulinum* [102] and *C. difficile* [43]. Based on the knowledge derived from the metabolic model it should be possible to develop a chemically defined medium for *R. ilealis* CRIB^T as well. Preliminary attempts in this direction have failed, and one possibility for lack of growth could be the lack of amino acids or, more specifically, specific transporters that would allow the organism to take them up from the environment if supplied in a defined

medium. To this end, it is interesting to note that genome analysis revealed the presence of an oligopeptide transporter as well as a range of different endopeptidases, suggesting that some amino acids might only be taken up as oligopeptides.

We are just starting to elucidate the composition and the function of the microbial communities in the mammalian small intestine. Recently we have reported the isolation and characterization of *R. ilealis* CRIB^T from the small intestine of a rat (**Chapter 4** [32]). In rats this species was identified to be a dominant member of the ileal microbiota (**Chapter 3** [33]). Here we applied a holistic systems biology approach, involving several fields of wet and dry biology, to study *R. ilealis* CRIB^T, a natural and abundant inhabitant of the small intestinal tract of rats, with close relatives also present at high numbers in the human small intestine (**Chapter 6**). Its ability to use host sugars that are liberated by other microbes, suggest *R. ilealis* CRIB^T is in competition with mucus-degrading microbes, like *B. thetatotaomicron* or *Akkermansia muciniphila*. In conclusion, *R. ilealis* is a species that is able to utilize an array of carbohydrates using different and partially redundant pathways. In contrast, it has only limited ability to *de novo* synthesize amino acids and vitamins, and hence the organism shows an adaption to a nutrient-rich environment in which carbohydrates and exogenous sources of amino acids and vitamins are abundantly available.

Acknowledgements

The authors thank Hans Heilig for his help with the genomic DNA isolations and Jasper Koehorst for his help with the genome annotation. This work was supported by a grant of SenterNovum (Agentschap NL), an agency of the Dutch Ministry of Economic Affairs (FND-07013).

References

1. Alikhan N.F., N.K. Petty, N.L. Ben Zakour and S.A. Beatson (2011). BLAST Ring Image Generator (BRIG): simple prokaryote genome comparisons. *BMC genomics*; 12: 402.
2. Almagro-Moreno S. and E.F. Boyd (2009). Insights into the evolution of sialic acid catabolism among bacteria. *BMC Evol Biol*; 9: 118.
3. Ammons D., J. Rampersad and G.E. Fox (1999). 5S rRNA gene deletions cause an unexpectedly high fitness loss in *Escherichia coli*. *Nucleic Acids Res*; 27 (2): 637-642.
4. Backhed F., R.E. Ley, J.L. Sonnenburg, D.A. Peterson and J.I. Gordon (2005). Host-bacterial mutualism in the human intestine. *Science*; 307 (5717): 1915-1920.
5. Baldoma L. and J. Aguilar (1988). Metabolism of L-fucose and L-rhamnose in *Escherichia coli*: aerobic-anaerobic regulation of L-lactaldehyde dissimilation. *J Bacteriol*; 170 (1): 416-421.
6. Barrangou R., E. Altermann, R. Hutkins, R. Cano and T.R. Klaenhammer (2003). Functional and comparative genomic analyses of an operon involved in fructooligosaccharide utilization by *Lactobacillus acidophilus*. *Proc Natl Acad Sci USA*; 100 (15): 8957-8962.
7. Becker D.J. and J.B. Lowe (2003). Fucose: biosynthesis and biological function in mammals. *Glycobiology*; 13 (7): 41R-53R.
8. Bensaadi-Merchermek N., J.C. Salvado, C. Cagnon, S. Karama and C. Mouches (1995). Characterization of the unlinked 16S rDNA and 23S-5S rRNA operon of *Wolbachia pipientis*, a prokaryotic parasite of insect gonads. *Gene*; 165 (1): 81-86.
9. Bergner H., O. Simon, T. Zebrowska and R. Munchmeyer (1986). Studies on the secretion of amino acids and of urea into the gastrointestinal tract of pigs. 3. Secretion of urea determined by continuous intravenous infusion of 15N-urea. *Archiv fur Tierernahrung*; 36 (6): 479-490.
10. Bland C., T.L. Ramsey, F. Sabree, M. Lowe, K. Brown, N.C. Kyrpides and P. Hugenholtz (2007). CRISPR recognition tool (CRT): a tool for automatic detection of clustered regularly interspaced palindromic repeats. *BMC bioinformatics*; 8: 209.
11. Boetzer M., C.V. Henkel, H.J. Jansen, D. Butler and W. Pirovano (2011). Scaffolding pre-assembled contigs using SSPACE. *Bioinformatics*; 27 (4): 578-579.
12. Boetzer M. and W. Pirovano (2012). Toward almost closed genomes with GapFiller. *Genome Biol*; 13 (6): R56.
13. Boisvert S., F. Laviolette and J. Corbeil (2010). Ray: simultaneous assembly of reads from a mix of high-throughput sequencing technologies. *J Comput Biol*; 17 (11): 1519-1533.
14. Booijink C.C., E.G. Zoetendal, M. Kleerebezem and W.M. de Vos (2007). Microbial communities in the human small intestine: coupling diversity to metagenomics. *Future Microbiol*; 2 (3): 285-295.
15. Burge S.W., J. Daub, R. Eberhardt, J. Tate, L. Barquist, E.P. Nawrocki, S.R. Eddy, *et al.* (2013). Rfam 11.0: 10 years of RNA families. *Nucleic Acids Res*; 41 (Database issue): D226-232.
16. Calusinska M., T. Happe, B. Joris and A. Wilmotte (2010). The surprising diversity of clostridial hydrogenases: a comparative genomic perspective. *Microbiology*; 156 (Pt 6): 1575-1588.
17. Claudel-Renard C., C. Chevalet, T. Faraut and D. Kahn (2003). Enzyme-specific profiles for genome annotation: PRIAM. *Nucleic Acids Res*; 31 (22): 6633-6639.
18. Commichau F.M., F.M. Rothe, C. Herzberg, E. Wagner, D. Hellwig, M. Lehnik-Habrink, E. Hammer, *et al.* (2009). Novel activities of glycolytic enzymes in *Bacillus subtilis*: interactions with essential proteins involved in mRNA processing. *Mol Cell Proteomics : MCP*; 8 (6): 1350-1360.
19. Czyzewski B.K. and D.N. Wang (2012). Identification and characterization of a bacterial hydrosulphide ion channel. *Nature*; 483 (7390): 494-497.
20. Darling A.E., B. Mau and N.T. Perna (2010). progressiveMauve: multiple genome alignment with gene gain, loss and rearrangement. *PLoS One*; 5 (6): e11147.
21. Delcher A.L., A. Phillippy, J. Carlton and S.L. Salzberg (2002). Fast algorithms for large-scale genome alignment and comparison. *Nucleic Acids Res*; 30 (11): 2478-2483.
22. Derrien M., M.W. van Passel, J.H. van de Bovenkamp, R.G. Schipper, W.M. de Vos and J. Dekker (2010). Mucin-bacterial interactions in the human oral cavity and digestive tract. *Gut Microbes*; 1 (4): 254-268.
23. Deutscher J. (2008). The mechanisms of carbon catabolite repression in bacteria. *Curr Opin Microbiol*; 11 (2): 87-93.
24. Dhillon A., S. Goswami, M. Riley, A. Teske and M. Sogin (2005). Domain evolution and functional diversification of sulfite reductases. *Astrobiology*; 5 (1): 18-29.
25. Dupuy B., G. Daube, M.R. Popoff and S.T. Cole (1997). *Clostridium perfringens* urease genes are plasmid borne. *Infect Immun*; 65 (6): 2313-2320.

26. Elia M. and J.H. Cummings (2007). Physiological aspects of energy metabolism and gastrointestinal effects of carbohydrates. *Eur J Clin Nutr*; 61 Suppl 1: S40-74.
27. Fischer J.D., G.L. Holliday and J.M. Thornton (2010). The CoFactor database: organic cofactors in enzyme catalysis. *Bioinformatics*; 26 (19): 2496-2497.
28. Flint H.J., K.P. Scott, S.H. Duncan, P. Louis and E. Forano (2012). Microbial degradation of complex carbohydrates in the gut. *Gut Microbes*; 3 (4): 289-306.
29. Fuller M.F. and P.J. Reeds (1998). Nitrogen cycling in the gut. *Annual review of nutrition*; 18: 385-411.
30. Ganzle M.G. and R. Follador (2012). Metabolism of oligosaccharides and starch in lactobacilli: a review. *Front Microbiol*; 3: 340.
31. Gao S., W.K. Sung and N. Nagarajan (2011). Opera: reconstructing optimal genomic scaffolds with high-throughput paired-end sequences. *J Comput Biol*; 18 (11): 1681-1691.
32. Gerritsen J., S. Fuentes, W. Grievink, L. van Niftrik, B.J. Tindall, H.M. Timmerman, G.T. Rijkers, *et al.* (2014). Characterization of *Romboutsia ilealis* gen. nov., sp. nov., isolated from the gastro-intestinal tract of a rat, and proposal for the reclassification of five closely related members of the genus *Clostridium* into the genera *Romboutsia* gen. nov., *Intestinibacter* gen. nov., *Terrisporobacter* gen. nov. and *Asaccharospora* gen. nov. *Int J Syst Evol Microbiol*; 64 (Pt 5): 1600-1616.
33. Gerritsen J., H.M. Timmerman, S. Fuentes, L.P. van Minnen, H. Panneman, S.R. Konstantinov, F.M. Rombouts, *et al.* (2011). Correlation between protection against sepsis by probiotic therapy and stimulation of a novel bacterial phylotype. *Appl Environ Microbiol*; 77 (21): 7749-7756.
34. Goh Y.J., C. Zhang, A.K. Benson, V. Schlegel, J.H. Lee and R.W. Hutkins (2006). Identification of a putative operon involved in fructooligosaccharide utilization by *Lactobacillus paracasei*. *Appl Environ Microbiol*; 72 (12): 7518-7530.
35. Goldner S.B., M. Solberg and L.S. Post (1985). Development of a minimal medium for *Clostridium perfringens* by using an anaerobic chemostat. *Appl Environ Microbiol*; 50 (2): 202-206.
36. Heinemann M., A. Kummel, R. Ruinatscha and S. Panke (2005). In silico genome-scale reconstruction and validation of the *Staphylococcus aureus* metabolic network. *Biotechnol Bioeng*; 92 (7): 850-864.
37. Heinken A., S. Sahoo, R.M. Fleming and I. Thiele (2013). Systems-level characterization of a host-microbe metabolic symbiosis in the mammalian gut. *Gut Microbes*; 4 (1): 28-40.
38. Hernandez D., P. Francois, L. Farinelli, M. Osteras and J. Schrenzel (2008). *De novo* bacterial genome sequencing: millions of very short reads assembled on a desktop computer. *Genome research*; 18 (5): 802-809.
39. Hooper L.V., J. Xu, P.G. Falk, T. Midtvedt and J.I. Gordon (1999). A molecular sensor that allows a gut commensal to control its nutrient foundation in a competitive ecosystem. *Proc Natl Acad Sci USA*; 96 (17): 9833-9838.
40. Hunter S., P. Jones, A. Mitchell, R. Apweiler, T.K. Attwood, A. Bateman, T. Bernard, *et al.* (2012). InterPro in 2011: new developments in the family and domain prediction database. *Nucleic Acids Res*; 40 (Database issue): D306-312.
41. Hyatt D., G.L. Chen, P.F. Locascio, M.L. Land, F.W. Larimer and L.J. Hauser (2010). Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC bioinformatics*; 11: 119.
42. Ingram-Smith C., A. Gorrell, S.H. Lawrence, P. Iyer, K. Smith and J.G. Ferry (2005). Characterization of the acetate binding pocket in the *Methanosarcina thermophila* acetate kinase. *J Bacteriol*; 187 (7): 2386-2394.
43. Karasawa T., S. Ikoma, K. Yamakawa and S. Nakamura (1995). A defined growth medium for *Clostridium difficile*. *Microbiology*; 141 (Pt 2): 371-375.
44. Kopke M., M. Straub and P. Durre (2013). Clostridium difficile is an autotrophic bacterial pathogen. *PLoS One*; 8 (4): e62157.
45. Kopylova E., L. Noe and H. Touzet (2012). SortMeRNA: fast and accurate filtering of ribosomal RNAs in metatranscriptomic data. *Bioinformatics*; 28 (24): 3211-3217.
46. Koren S., G.P. Harhay, T.P. Smith, J.L. Bono, D.M. Harhay, S.D. McVey, D. Radune, *et al.* (2013). Reducing assembly complexity of microbial genomes with single-molecule sequencing. *Genome Biol*; 14 (9): R101.
47. Kortman G.A., M. Raffatellu, D.W. Swinkels and H. Tjalsma (2014). Nutritional iron turned inside out: intestinal stress from a gut microbial perspective. *FEMS Microbiol Rev*.
48. Lagesen K., P. Hallin, E.A. Rodland, H.H. Staerfeldt, T. Rognes and D.W. Ussery (2007). RNAmmer: consistent and rapid annotation of ribosomal RNA genes. *Nucleic Acids Res*; 35 (9): 3100-3108.
49. Lang J.M., A.E. Darling and J.A. Eisen (2013). Phylogeny of bacterial and archaeal genomes using conserved genes: supertrees and supermatrices. *PLoS One*; 8 (4): e62510.
50. Lange K., F. Hugenholtz, H. Schols, M. Kleerebezem, H. Smidt, M. Müller and G.J.E.J. Hooiveld (2014). Distinct responses of mucosal transcriptional profiles and luminal microbiota composition and SCFA concentrations to dietary fibers in murine colon. *Submitted*.
51. Langmead B. and S.L. Salzberg (2012). Fast gapped-read alignment with Bowtie 2. *Nature Methods*; 9 (4): 357-359.

52. Langmead B., C. Trapnell, M. Pop and S.L. Salzberg (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol*; 10 (3): R25.
53. Latendresse M., M. Krummenacker, M. Trupp and P.D. Karp (2012). Construction and completion of flux balance models from pathway databases. *Bioinformatics*; 28 (3): 388-396.
54. LeBlanc D.J. and R.P. Mortlock (1971). Metabolism of D-arabinose: a new pathway in *Escherichia coli*. *J Bacteriol*; 106 (1): 90-96.
55. Lee Z.M., C. Bussema, 3rd and T.M. Schmidt (2009). rrnDB: documenting the number of rRNA and tRNA genes in bacteria and archaea. *Nucleic Acids Res*; 37 (Database issue): D489-493.
56. Li H., B. Handsaker, A. Wysoker, T. Fennell, J. Ruan, N. Homer, G. Marth, *et al.* (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics*; 25 (16): 2078-2079.
57. Li M., B. Wang, M. Zhang, M. Rantalainen, S. Wang, H. Zhou, Y. Zhang, *et al.* (2008). Symbiotic gut microbes modulate human metabolic phenotypes. *Proc Natl Acad Sci USA*; 105 (6): 2117-2122.
58. Linke C.M., S.A. Woodiga, D.J. Meyers, C.M. Buckwalter, H.E. Salhi and S.J. King (2013). The ABC transporter encoded at the pneumococcal fructooligosaccharide utilization locus determines the ability to utilize long- and short-chain fructooligosaccharides. *J Bacteriol*; 195 (5): 1031-1041.
59. Lowe T.M. and S.R. Eddy (1997). tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res*; 25 (5): 955-964.
60. Makarova K.S., D.H. Haft, R. Barrangou, S.J. Brouns, E. Charpentier, P. Horvath, S. Moineau, *et al.* (2011). Evolution and classification of the CRISPR-Cas systems. *Nat Rev Microbiol*; 9 (6): 467-477.
61. Marshall B.J., L.J. Barrett, C. Prakash, R.W. McCallum and R.L. Guerrant (1990). Urea protects *Helicobacter* (*Campylobacter*) *pylori* from the bactericidal effect of acid. *Gastroenterology*; 99 (3): 697-702.
62. Martin M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *2011*; 17 (1).
63. Miller R.S. and L.C. Hoskins (1981). Mucin degradation in human colon ecosystems. Fecal population densities of mucin-degrading bacteria estimated by a "most probable number" method. *Gastroenterology*; 81 (4): 759-765.
64. Mitsuhashi H., Y. Nojima, T. Tanaka, K. Ueki, A. Maezawa, S. Yano and T. Naruse (1998). Sulfite is released by human neutrophils in response to stimulation with lipopolysaccharide. *J Leukoc Biol*; 64 (5): 595-599.
65. Mobley H.L., M.D. Island and R.P. Hausinger (1995). Molecular biology of microbial ureases. *Microbiol Rev*; 59 (3): 451-480.
66. Ng K.M., J.A. Ferreyra, S.K. Higginbottom, J.B. Lynch, P.C. Kashyap, S. Gopinath, N. Naidu, *et al.* (2013). Microbiota-liberated host sugars facilitate post-antibiotic expansion of enteric pathogens. *Nature*; 502 (7469): 96-99.
67. Oultram J.D., I.D. Burr, M.J. Elmore and N.P. Minton (1993). Cloning and sequence analysis of the genes encoding phosphotransbutyrylase and butyrate kinase from *Clostridium acetobutylicum* NCIMB 8052. *Gene*; 131 (1): 107-112.
68. Ouwerkerk J.P., W.M. de Vos and C. Belzer (2013). Glycobiome: bacteria and mucus at the epithelial interface. *Best Pract Res Clin Gastroenterol*; 27 (1): 25-38.
69. Overbeek R., M. Fonstein, M. D'Souza, G.D. Pusch and N. Maltsev (1999). The use of gene clusters to infer functional coupling. *Proc Natl Acad Sci USA*; 96 (6): 2896-2901.
70. Qin J., R. Li, J. Raes, M. Arumugam, K.S. Burgdorf, C. Manichanh, T. Nielsen, *et al.* (2010). A human gut microbial gene catalogue established by metagenomic sequencing. *Nature*; 464 (7285): 59-65.
71. Quigley E.M. (2013). Gut bacteria in health and disease. *Gastroenterol Hepatol*; 9 (9): 560-569.
72. Quinlan A.R. and I.M. Hall (2010). BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics*; 26 (6): 841-842.
73. Reichardt N., S.H. Duncan, P. Young, A. Belenguer, C. McWilliam Leitch, K.P. Scott, H.J. Flint, *et al.* (2014). Phylogenetic distribution of three pathways for propionate production within the human gut microbiota. *ISME J*; 8 (6): 1323-1335.
74. Ridlon J.M., D.J. Kang and P.B. Hylemon (2006). Bile salt biotransformations by human intestinal bacteria. *J Lipid Res*; 47 (2): 241-259.
75. Robbe C., C. Capon, B. Coddeville and J.C. Michalski (2004). Structural diversity and specific distribution of O-glycans in normal human mucins along the intestinal tract. *Biochem J*; 384 (Pt 2): 307-316.
76. Robbe C., C. Capon, E. Maes, M. Rousset, A. Zweibaum, J.P. Zanetta and J.C. Michalski (2003). Evidence of regio-specific glycosylation in human intestinal mucins: presence of an acidic gradient along the intestinal tract. *J Biol Chem*; 278 (47): 46337-46348.
77. Roggentin P., G. Gutschker-Gdaniec, R. Schauer and R. Hobrecht (1985). Correlative properties for a differentiation of two *Clostridium sordellii* phenotypes and their distinction from *Clostridium bifermentans*. *Zentralbl Bakteriol Mikrobiol Hyg A*; 260 (3): 319-328.
78. Rutherford J.C. (2014). The emerging role of urease as a general microbial virulence factor. *PLoS Path*; 10 (5):

e1004062.
79. Saulnier D.M., D. Molenaar, W.M. de Vos, G.R. Gibson and S. Kolida (2007). Identification of prebiotic fructooligosaccharide metabolism in *Lactobacillus plantarum* WCFS1 through microarrays. *Appl Environ Microbiol*; 73 (6): 1753-1765.
80. Schmieder R. and R. Edwards (2011). Quality control and preprocessing of metagenomic datasets. *Bioinformatics*; 27 (6): 863-864.
81. Schwartz J.J., A. Gazumyan and I. Schwartz (1992). rRNA gene organization in the Lyme disease spirochete, *Borrelia burgdorferi*. *J Bacteriol*; 174 (11): 3757-3765.
82. Scott K.P., S.W. Gratz, P.O. Sheridan, H.J. Flint and S.H. Duncan (2013). The influence of diet on the gut microbiota. *Pharmacol Res*; 69 (1): 52-60.
83. Scott K.P., J.C. Martin, G. Campbell, C.D. Mayer and H.J. Flint (2006). Whole-genome transcription profiling reveals genes up-regulated by growth on fucose in the human gut bacterium "*Roseburia inulinivorans*". *J Bacteriol*; 188 (12): 4340-4349.
84. Sebaihia M., B.W. Wren, P. Mullany, N.F. Fairweather, N. Minton, R. Stabler, N.R. Thomson, *et al.* (2006). The multidrug-resistant human pathogen *Clostridium difficile* has a highly mobile, mosaic genome. *Nature Genet*; 38 (7): 779-786.
85. Sharma V., M.F. Prere, I. Canal, A.E. Firth, J.F. Atkins, P.V. Baranov and O. Fayet (2014). Analysis of tetra- and hepta-nucleotides motifs promoting -1 ribosomal frameshifting in Escherichia coli. *Nucleic Acids Res*; 42 (11): 7210-7225.
86. Stahl M., L.M. Friis, H. Nothaft, X. Liu, J. Li, C.M. Szymanski and A. Stintzi (2011). L-fucose utilization provides *Campylobacter jejuni* with a competitive advantage. *Proc Natl Acad Sci USA*; 108 (17): 7194-7199.
87. Stams A.J., J.B. Van Dijk, C. Dijkema and C.M. Plugge (1993). Growth of syntrophic propionate-oxidizing bacteria with fumarate in the absence of methanogenic bacteria. *Appl Environ Microbiol*; 59 (4): 1114-1119.
88. Suthers P.F., M.S. Dasika, V.S. Kumar, G. Denisov, J.I. Glass and C.D. Maranas (2009). A genome-scale metabolic reconstruction of *Mycoplasma genitalium*, iPS189. *PLoS Comput Biol*; 5 (2): e1000285.
89. Suzek B.E., H. Huang, P. McGarvey, R. Mazumder and C.H. Wu (2007). UniRef: comprehensive and non-redundant UniProt reference clusters. *Bioinformatics*; 23 (10): 1282-1288.
90. Tatusov R.L., E.V. Koonin and D.J. Lipman (1997). A genomic perspective on protein families. *Science*; 278 (5338): 631-637.
91. ter Braak C.F.J. and P. Smilauer (2012). Canoco Reference Manual and User's Guide: Software for Ordination,Version 5.0. Ithaca, NY, USA: Microcomputer Power.
92. Trapnell C., D.G. Hendrickson, M. Sauvageau, L. Goff, J.L. Rinn and L. Pachter (2013). Differential analysis of gene regulation at transcript resolution with RNA-seq. *Nat Biotechnol*; 31 (1): 46-53.
93. UniProt-Consortium (2014). Activities at the universal protein resource (UniProt). *Nucleic Acids Res*; 42 (Database issue): D191-D198.
94. Van den Bogert B., J. Boekhorst, R. Herrmann, E.J. Smid, E.G. Zoetendal and M. Kleerebezem (2013). Comparative genomics analysis of *Streptococcus* isolates from the human small intestine reveals their adaptation to a highly dynamic ecosystem. *PLoS One*; 8 (12): e83418.
95. van den Bogert B., O. Erkus, J. Boekhorst, M. de Goffau, E.J. Smid, E.G. Zoetendal and M. Kleerebezem (2013). Diversity of human small intestinal *Streptococcus* and *Veillonella* populations. *FEMS Microbiol Ecol*; 85 (2): 376-388.
96. van der Oost J., E.R. Westra, R.N. Jackson and B. Wiedenheft (2014). Unravelling the structural and mechanistic basis of CRISPR-Cas systems. *Nat Rev Microbiol*; 12 (7): 479-492.
97. van Hijum S.A., S. Kralj, L.K. Ozimek, L. Dijkhuizen and I.G. van Geel-Schutten (2006). Structure-function relationships of glucansucrase and fructansucrase enzymes from lactic acid bacteria. *Microbiol Mol Biol Rev*; 70 (1): 157-176.
98. Vimr E.R. (2013). Unified theory of bacterial sialometabolism: how and why bacteria metabolize host sialic acids. *ISRN microbiology*; 2013: 816713.
99. Vimr E.R., K.A. Kalivoda, E.L. Deszo and S.M. Steenbergen (2004). Diversity of microbial sialic acid metabolism. *Microbiol Mol Biol Rev*; 68 (1): 132-153.
100. Vital M., A.C. Howe and J.M. Tiedje (2014). Revealing the bacterial butyrate synthesis pathways by analyzing (meta) genomic data. *mBio*; 5 (2): e00889.
101. Vital M., C.R. Penton, Q. Wang, V.B. Young, D.A. Antonopoulos, M.L. Sogin, H.G. Morrison, *et al.* (2013). A gene-targeted approach to investigate the intestinal butyrate-producing bacterial community. *Microbiome*; 1 (1): 8.
102. Whitmer M.E. and E.A. Johnson (1988). Development of improved defined media for *Clostridium botulinum* serotypes A, B, and E. *Appl Environ Microbiol*; 54 (3): 753-759.
103. Wiedenheft B., S.H. Sternberg and J.A. Doudna (2012). RNA-guided genetic silencing systems in bacteria and archaea. *Nature*; 482 (7385): 331-338.
104. Williams D., W.L. Trimble, M. Shilts, F. Meyer and H. Ochman (2013). Rapid quantification of sequence repeats to

resolve the size, structure and contents of bacterial genomes. *BMC genomics*; 14: 537.
105. Xu J., M.K. Bjursell, J. Himrod, S. Deng, L.K. Carmichael, H.C. Chiang, L.V. Hooper, *et al.* (2003). A genomic view of the human-*Bacteroides thetaiotaomicron* symbiosis. *Science*; 299 (5615): 2074-2076.
106. Zhang Z., J. Geng, X. Tang, H. Fan, J. Xu, X. Wen, Z.S. Ma, *et al.* (2014). Spatial heterogeneity and co-occurrence patterns of human mucosal-associated intestinal microbiota. *ISME J*; 8 (4): 881-893.
107. Zoetendal E.G., J. Raes, B. van den Bogert, M. Arumugam, C.C. Booijink, F.J. Troost, P. Bork, *et al.* (2012). The human small intestinal microbiota is driven by rapid uptake and conversion of simple carbohydrates. *ISME J*; 6 (7): 1415-1426.

Supplemental material

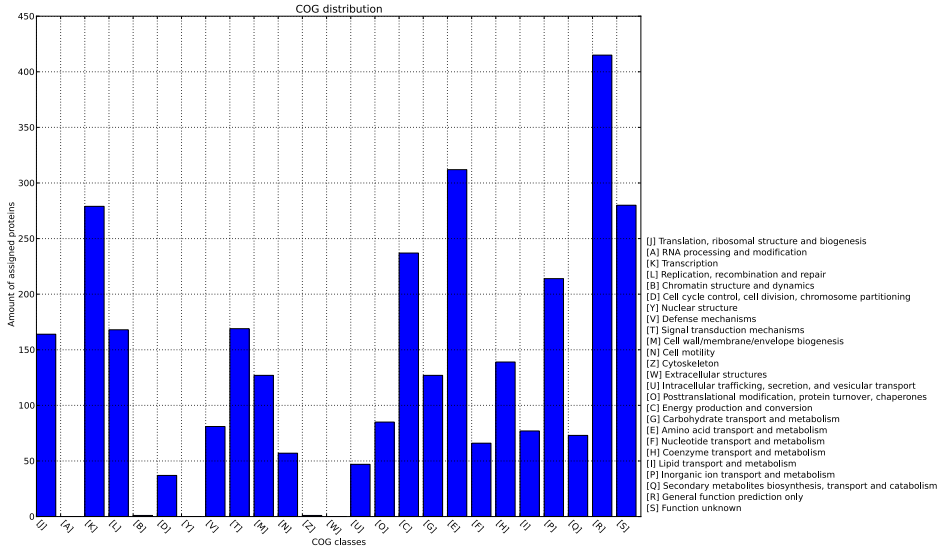

Figure S1. Distribution of COG categories in R. ilealis CRIB^T.

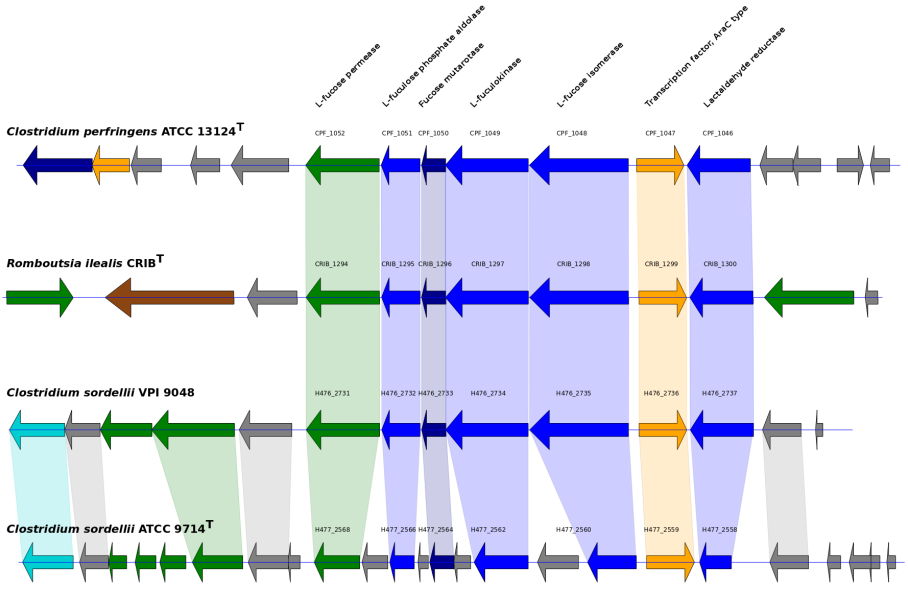

Figure S2. Organization of the L-fucose degradation gene cluster of *R. ilealis* CRIB^T compared to similar gene clusters found in *C. perfringens* (strain ATCC^T) and close relative *C. sordellii* (strains VPI 9048 and ATCC 9714^T). In the current annotation of *C. sordellii* ATCC 9714^T open reading frame prediction seems to be suboptimal, with a considerable amount of potentially wrong stop codons. Genes are color-coded by predicted function: transporter (green), metabolic enzyme with EC number (blue), metabolic enzyme with preliminary EC number/without assigned EC number (dark blue), transcriptional regulator (yellow), hypothetical/unknown protein (grey), protein involved in DNA processing (brown), protein involved in vitamin metabolism (light blue).

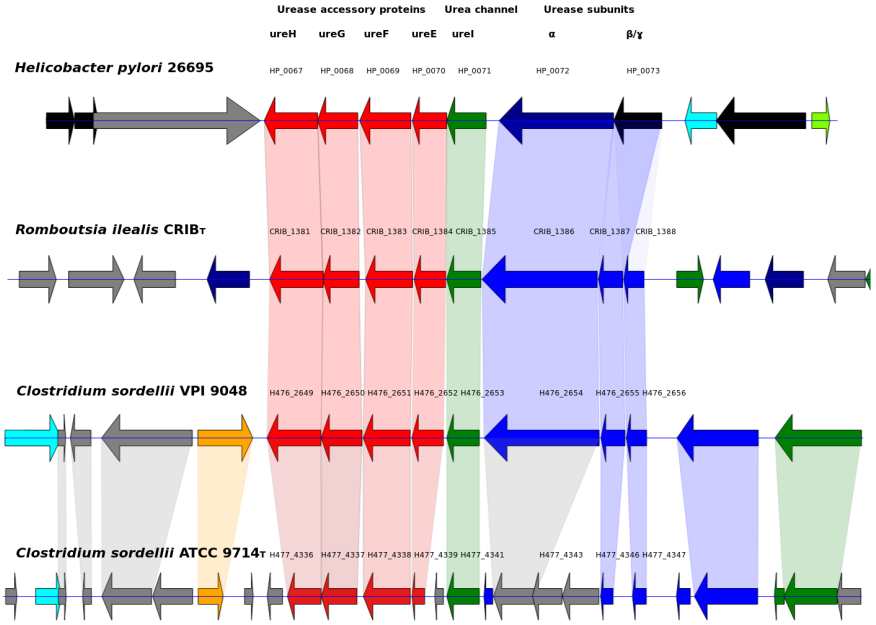

Figure S3. Organization of the urease gene cluster of *R. ilealis* CRIB^T compared to similar gene clusters found in *H. pylori* (strain 26695) and close-relative *C. sordellii* (strains VPI 9048 and ATCC 9714^T). In the current annotation of *C. sordellii* ATCC 9714^T open reading frame prediction seems to be suboptimal, with a considerable amount of potentially wrong stop codons. Genes are color-coded by predicted function: urease accessory proteins (red), transporter (green), metabolic enzyme with EC number (blue), metabolic enzyme with preliminary EC number/without assigned EC number (dark blue), transcriptional regulator (yellow), hypothetical/unknown protein (grey), protein involved in vitamin metabolism (light blue), ribosomal proteins (green).

Table S1. Growth characteristics of the individual cultures used for whole-genome transcriptome analysis of R. ilealis CRIB^T grown in four experimental conditions.

	Mid-exponential phase (~8-10h)		Stationary phase (24h)	
	OD_{600nm}	pH	OD_{600nm}	pH
Glucose 1	0.595	6.99	1.168	6.19
Glucose 2	0.585	6.95	1.072	6.26
Glucose 3	0.560	6.92	0.992	6.11
Average	**0.580**	**6.95**	**1.077**	**6.19**
FOS 1	0.492	6.90	1.149	6.37
FOS 2	0.550	6.89	1.079	6.26
FOS 3	0.526	6.90	1.069	6.33
Average	**0.523**	**6.90**	**1.099**	**6.32**
L-fucose 1	0.547	6.94	0.71	6.31
L-fucose 2	0.520	6.93	0.61	6.37
L-fucose 3	0.563	6.87	0.618	6.36
Average	**0.543**	**6.91**	**0.646**	**6.35**
Control 1 (preconditioned on glucose)	0.315	6.94	0.358	6.88
Control 2 (preconditioned on FOS)	0.403	6.98	0.33	6.97
Control 3 (preconditioned on L-fucose)	0.413	6.97	0.367	6.93
Average	**0.377**	**6.96**	**0.352**	**6.93**

Table S2. Summary of the RNA-seq raw data analysis for whole-genome transcriptome analysis of R. ilealis CRIB^T grown in four experimental conditions.

	Glucose			FOS		
	1	2	3	1	2	3
Total no. of reads	15,574,335	11,030,037	15,657,610	12,652,647	11,948,224	14,122,138
Total no. of rRNA reads	1,724,209	599,335	108,411	933,615	2,444,855	2,337,475
Percentage of rRNA reads	11	5	1	7	20	17
Total no. of reads after rRNA removal	13,850,126	10,430,702	15,549,199	11,719,032	9,503,369	11,784,663
Total no. of bases removed as adapters	52,6331,427	232,474,074	495,884,518	353,295,955	285,235,443	338,892,680
Percentage of bases removed as adapters	25	15	21	20	20	19
Total no. of reads after quality filtering	1,2893,890	9,797,525	14,588,980	10,910,861	8,795,679	11,023,854
Percentage of reads that passed quality filtering	93	94	94	93	93	94
Mean length of reads that passed quality filtering	112	127	118	120	121	121
Percentage of initial bases after all filtering steps	62	75	73	69	59	63
Total no. of reads mapped to the genome	1,266,0782	9,663,639	14,374,304	10,747,772	8,607,651	10,832,643
Percentage of reads mapping to the genome	98	99	99	99	98	98

	L-Fucose			Control		
	1	2	3	1 (glucose)	2 (FOS)	3 (L-fucose)
Total no. of reads	14,215,916	15,527,737	9,086,655	15,522,354	13,691,129	1,022,1853
Total no. of rRNA reads	2,689,251	1,115,485	5,677,746	1,085,502	160,691	382,228
Percentage of rRNA reads	19	7	62	7	1	4
Total no. of reads after rRNA removal	11,526,665	14,412,252	3,408,909	14,436,852	13,530,438	9,839,625
Total no. of bases removed as adapters	360,471,843	559,814,123	124,300,857	461,476,164	443,584,851	223,111,337
Percentage of bases removed as adapters	21	26	24	21	22	15
Total no. of reads after quality filtering	10,769,006	13,474,757	3,124,596	13,550,929	1,253,1571	9,261,668
Percentage of reads that passed quality filtering	93	94	92	94	93	94
Mean length of reads that passed quality filtering	119	111	115	118	118	127
Percentage of initial bases after all filtering steps	60	64	26	69	72	77
Total no. of reads mapped to the genome	10,545,697	13,067,619	2,990,835	13,347,802	12,228,370	9,128,748
Percentage of reads mapping to the genome	98	97	96	99	98	99

Table S3. Overview of the genes that were significantly upregulated during growth of R. ilealis CRIB^T in the presence of glucose in comparison to at least one of the other three conditions [L-fucose, control (absence of additional carbon source) and FOS] as identified by RNA-seq analysis. Differential gene expression values that did not meet the criteria for significance (≥1.5 log2(fold change) and q value ≤0.05) are color-coded in dark-grey.

			Glucose vs L-fucose		Glucose vs control		Glucose vs FOS	
	locus	annotation	log2(fold change)	q value	log2(fold change)	q value	log2(fold change)	q value
1	CRIB_2231	Aldehyde-alcohol dehydrogenase 2	3.90	0.0003	3.17	0.0002	0.76	0.3131
2	CRIB_2018	PTS system glucose-specific EIIA component	3.74	0.0003	2.89	0.0002	1.58	0.0018
3	CRIB_2017	PTS system glucose-specific EIIBC component	3.33	0.0003	2.54	0.0002	1.31	0.0018
4	CRIB_2019	Transcription antiterminator LicT	3.06	0.0003	2.58	0.0002	1.11	0.0018
5	CRIB_877	c-di-GMP-I	2.63	0.0003	0.86	0.0101	-0.10	0.7471
6	CRIB_1635	FeoA domain	2.48	0.0003	1.71	0.0002	1.53	0.0018
7	CRIB_186	Central glycolytic genes regulator	2.48	0.0003	3.76	0.0002	1.03	0.0200
8	CRIB_934	Rhodanese-like protein	2.31	0.0003	2.65	0.0002	1.32	0.0018
9	CRIB_1634	Ferrous iron transport protein B	2.09	0.0003	0.89	0.0152	0.85	0.0338
10	CRIB_187	Glyceraldehyde-3-phosphate dehydrogenase	2.05	0.0003	3.45	0.0002	0.60	0.2333
11	CRIB_2505	Protein translocase subunit SecY	2.05	0.0003	2.20	0.0002	0.53	0.2790
12	CRIB_601	RNA polymerase sigma-70 factor	2.00	0.0003	-0.99	0.0224	-3.58	0.0018
13	CRIB_2506	Adenylate kinase	2.00	0.0003	2.28	0.0002	0.68	0.2158
14	CRIB_2507	Methionine aminopeptidase 1	2.00	0.0003	2.28	0.0002	0.68	0.2158
15	CRIB_439	PF09924 protein	1.99	0.0003	2.17	0.0002	0.81	0.0219
16	CRIB_2510	30S ribosomal protein S13	1.98	0.0003	2.02	0.0002	0.83	0.0682
17	CRIB_2038	Heme oxygenase	1.98	0.0003	0.14	0.6834	0.51	0.1518
18	CRIB_472	YqzL-like protein	1.97	0.0003	1.93	0.0002	0.74	0.0514
19	CRIB_473	DNA repair protein RecO	1.97	0.0003	1.93	0.0002	0.74	0.0514
20	CRIB_474	Domain of unknown function (DUF4342)	1.97	0.0003	1.93	0.0002	0.74	0.0514
21	CRIB_2369	Veg protein	1.91	0.0003	0.80	0.0202	-0.07	0.8691
22	CRIB_2492	50S ribosomal protein L16	1.89	0.0006	2.57	0.0002	1.10	0.0611
23	CRIB_2496	50S ribosomal protein L24	1.88	0.0003	2.48	0.0002	0.83	0.1008
24	CRIB_2486	50S ribosomal protein L4	1.88	0.0094	2.19	0.0006	0.66	0.4126
25	CRIB_2487	50S ribosomal protein L23	1.88	0.0094	2.19	0.0006	0.66	0.4126
26	CRIB_1247	Arsenate reductase	1.88	0.0003	1.28	0.0002	0.64	0.0267
27	CRIB_2511	30S ribosomal protein S11	1.86	0.0014	2.35	0.0002	1.01	0.0830
28	CRIB_1457	Sodium/alanine symporter	1.86	0.0003	1.50	0.0002	0.61	0.0910
29	CRIB_895	Ferric anguibactin-binding protein	1.83	0.0003	0.47	0.2352	0.58	0.2159

			Glucose vs L-fucose		Glucose vs control		Glucose vs FOS	
	locus	annotation	log2(fold change)	q value	log2(fold change)	q value	log2(fold change)	q value
30	CRIB_797	Protein of unknown function (DUF1659)	1.83	0.0003	1.56	0.0002	0.47	0.1254
31	CRIB_197	tmRNA	1.82	0.0061	0.74	0.1741	-0.09	0.9173
32	CRIB_470	GTPase Era	1.82	0.0003	2.07	0.0002	0.74	0.0357
33	CRIB_796	Protein of unknown function (DUF2922)	1.81	0.0003	1.51	0.0002	0.83	0.0081
34	CRIB_2043	T-box	1.81	0.0003	0.59	0.0308	0.36	0.2102
35	CRIB_1596	Transporter	1.80	0.0003	1.60	0.0002	0.27	0.3010
36	CRIB_935	TVP38/TMEM64 inner membrane protein YdjZ	1.80	0.0003	2.62	0.0002	0.79	0.0090
37	CRIB_2509	Ribosomal protein L36p	1.79	0.0003	2.27	0.0002	0.91	0.0056
38	CRIB_2501	50S ribosomal protein L18	1.78	0.0003	2.24	0.0002	0.83	0.0815
39	CRIB_892	Ferric anguibactin transport system permease	1.77	0.0003	0.50	0.2183	0.20	0.6777
40	CRIB_893	Ferric anguibactin transport system permease	1.77	0.0003	0.50	0.2183	0.20	0.6777
41	CRIB_894	Iron(3+)-hydroxamate import ATP-binding protein	1.77	0.0003	0.50	0.2183	0.20	0.6777
42	CRIB_634	Uncharacterized BCR, YitT family COG1284	1.76	0.0003	1.34	0.0002	0.29	0.3278
43	CRIB_1470	Hypothetical protein	1.76	0.0003	1.87	0.0002	1.25	0.0018
44	CRIB_2102	Preprotein translocase, subunit YajC	1.73	0.0003	1.51	0.0002	0.89	0.0090
45	CRIB_1862	Branched-chain amino acid transport system carrier	1.72	0.0003	0.72	0.0437	0.42	0.2765
46	CRIB_2079	Hypothetical protein	1.72	0.0003	0.93	0.0002	0.77	0.0081
47	CRIB_2493	Ribosomal protein L29p/L35e	1.70	0.0003	1.98	0.0002	0.45	0.1474
48	CRIB_255	Iron compound ABC transporter, iron	1.70	0.0010	0.59	0.1873	0.91	0.0964
49	CRIB_1864	50S ribosomal protein L19	1.69	0.0003	1.66	0.0002	0.68	0.1518
50	CRIB_2485	50S ribosomal protein L3	1.68	0.0010	2.15	0.0002	0.48	0.3819
51	CRIB_2503	Ribosomal protein L30p/L7e	1.67	0.0003	1.88	0.0002	0.66	0.0289
52	CRIB_2192	Ribosomal protein L35	1.67	0.0003	1.65	0.0002	0.85	0.0244
53	CRIB_2490	50S ribosomal protein L22	1.66	0.0006	2.36	0.0002	0.67	0.1723
54	CRIB_2508	Translation initiation factor IF-1	1.65	0.0003	1.39	0.0002	0.13	0.6697
55	CRIB_2139	DEAD-box ATP-dependent RNA helicase CshA	1.65	0.0003	2.11	0.0002	0.99	0.0154
56	CRIB_2372	SEC-C domain protein	1.64	0.0003	1.85	0.0002	1.11	0.0018
57	CRIB_2441	tRNA-specific adenosine deaminase	1.62	0.0003	1.49	0.0002	1.14	0.0018
58	CRIB_2513	DNA-directed RNA polymerase subunit alpha	1.61	0.0012	1.56	0.0002	0.50	0.3046
59	CRIB_1253	L-cystine uptake protein TcyP	1.61	0.0003	1.51	0.0002	0.28	0.3861
60	CRIB_2491	30S ribosomal protein S3	1.61	0.0034	2.25	0.0002	0.87	0.1366
61	CRIB_437	UBA/THIF-type NAD/FAD binding protein	1.61	0.0003	1.70	0.0002	0.43	0.1474

			Glucose vs L-fucose		Glucose vs control		Glucose vs FOS	
	locus	annotation	log2(fold change)	q value	log2(fold change)	q value	log2(fold change)	q value
62	CRIB_2494	Ribosomal protein S17p/S11e	1.60	0.0006	1.91	0.0002	0.48	0.2686
63	CRIB_2489	30S ribosomal protein S19	1.59	0.0003	1.76	0.0002	0.20	0.6029
64	CRIB_2499	30S ribosomal protein S8	1.59	0.0016	2.28	0.0002	0.62	0.2219
65	CRIB_890	Protein of unknown function (DUF3793)	1.58	0.0003	0.63	0.0134	0.49	0.0834
66	CRIB_2497	50S ribosomal protein L5	1.58	0.0016	2.08	0.0002	0.61	0.2193
67	CRIB_182	Small acid-soluble spore protein	1.57	0.0003	2.67	0.0002	1.25	0.0018
68	CRIB_183	Methyltransferase domain protein	1.57	0.0003	2.67	0.0002	1.25	0.0018
69	CRIB_1812	Ribosome maturation factor RimP	1.57	0.0003	1.68	0.0002	0.69	0.0313
70	CRIB_1595	Tubulin/FtsZ, C-terminal	1.57	0.0003	1.52	0.0002	0.50	0.0928
71	CRIB_1352	Radical SAM domain protein	1.57	0.0003	0.85	0.0153	0.93	0.0056
72	CRIB_2495	50S ribosomal protein L14	1.56	0.0028	2.49	0.0002	1.03	0.0566
73	CRIB_2069	Sugar transferase	1.55	0.0003	1.26	0.0002	0.76	0.0270
74	CRIB_2498	Ribosomal protein S14	1.55	0.0003	2.10	0.0002	0.77	0.0275
75	CRIB_1633	Virus attachment protein p12 family	1.53	0.0059	-0.07	0.8720	-0.37	0.3511
76	CRIB_471	Magnesium transporter MgtE	1.52	0.0003	1.58	0.0002	0.64	0.0663
77	CRIB_2425	Ribosomal protein L34p	1.51	0.0003	0.87	0.0144	0.60	0.1137
78	CRIB_2407	Single-stranded DNA-binding protein	1.51	0.0016	1.63	0.0002	0.85	0.0882
79	CRIB_2450	Protein of unknown function (DUF2508)	1.50	0.0006	1.59	0.0002	0.26	0.4135
80	CRIB_436	3D domain protein	1.50	0.0003	1.66	0.0002	0.37	0.2266
81	CRIB_1735	Hypothetical protein	1.50	0.0003	1.66	0.0002	0.83	0.0018
82	CRIB_2408	30S ribosomal protein S6	1.49	0.0003	1.69	0.0002	0.85	0.0256
83	CRIB_188	Phosphoglycerate kinase	1.49	0.0014	2.88	0.0002	0.65	0.1607
84	CRIB_2488	50S ribosomal protein L2	1.48	0.0030	2.19	0.0002	0.60	0.2694
85	CRIB_2520	30S ribosomal protein S9	1.48	0.0016	1.72	0.0002	1.04	0.0308
86	CRIB_2347	glmS	1.46	0.0003	1.86	0.0002	0.90	0.0018
87	CRIB_2500	50S ribosomal protein L6	1.44	0.0033	2.06	0.0002	0.62	0.2260
88	CRIB_2476	50S ribosomal protein L10	1.43	0.0006	1.56	0.0002	0.20	0.6297
89	CRIB_190	2,3-bisphosphoglycerate-independent	1.43	0.0029	2.84	0.0002	0.63	0.1970
90	CRIB_2512	30S ribosomal protein S4	1.43	0.0012	1.75	0.0002	0.51	0.2730
91	CRIB_191	Enolase	1.42	0.0029	2.44	0.0002	0.69	0.2212
92	CRIB_2502	30S ribosomal protein S5	1.41	0.0034	2.28	0.0002	0.84	0.0867
93	CRIB_1674	30S ribosomal protein S2	1.40	0.0019	1.63	0.0002	0.78	0.1000

			Glucose vs L-fucose		Glucose vs control		Glucose vs FOS	
	locus	annotation	log2(fold change)	q value	log2(fold change)	q value	log2(fold change)	q value
94	CRIB_2337	S-adenosylmethionine synthase	1.40	0.0042	1.94	0.0002	0.79	0.1037
95	CRIB_2504	50S ribosomal protein L15	1.40	0.0029	2.14	0.0002	0.89	0.0636
96	CRIB_2480	30S ribosomal protein S12	1.39	0.0030	1.55	0.0002	0.62	0.1874
97	CRIB_2514	50S ribosomal protein L17	1.39	0.0040	1.78	0.0002	0.78	0.0982
98	CRIB_1403	Hypothetical protein	1.38	0.0006	1.60	0.0002	0.72	0.0358
99	CRIB_136	SEC-C motif-containing protein	1.36	0.0003	1.56	0.0002	0.80	0.0154
100	CRIB_424	T-box	1.34	0.0003	1.73	0.0002	1.18	0.0018
101	CRIB_1834	Metallo beta-lactamase superfamily protein	1.34	0.0032	1.56	0.0002	0.89	0.0466
102	CRIB_365	Catabolite control protein A	1.32	0.0003	1.84	0.0002	0.64	0.0414
103	CRIB_2484	30S ribosomal protein S10	1.32	0.0016	2.13	0.0002	0.68	0.0956
104	CRIB_1932	RNA methyltransferase, RsmD	1.30	0.0003	1.69	0.0002	0.83	0.0056
105	CRIB_2519	50S ribosomal protein L13	1.30	0.0045	1.93	0.0002	1.03	0.0335
106	CRIB_2428	S4 domain-containing protein	1.29	0.0003	1.57	0.0002	0.66	0.0290
107	CRIB_957	Transcriptional regulator, ArsR	1.28	0.0003	1.54	0.0002	0.08	0.7861
108	CRIB_2348	Phosphoglucosamine mutase	1.28	0.0040	1.84	0.0002	1.08	0.0224
109	CRIB_1621	GTP-binding protein TypA/BipA homolog	1.27	0.0045	1.79	0.0002	0.54	0.2239
110	CRIB_1764	CBS domain protein	1.27	0.0003	1.58	0.0002	0.55	0.0558
111	CRIB_1702	Small-conductance mechanosensitive channel	1.27	0.0003	1.87	0.0002	0.67	0.0178
112	CRIB_1842	Alanine--tRNA ligase	1.27	0.0006	1.68	0.0002	0.64	0.0830
113	CRIB_1912	Exodeoxyribonuclease 7 large subunit	1.24	0.0019	1.66	0.0002	0.59	0.1461
114	CRIB_1913	O-sialoglycoprotein endopeptidase, peptidase M22	1.24	0.0019	1.66	0.0002	0.59	0.1461
115	CRIB_1914	Transcription termination protein, NusB	1.24	0.0019	1.66	0.0002	0.59	0.1461
116	CRIB_2090	D-tyrosyl-tRNA(Tyr) deacylase	1.22	0.0003	1.54	0.0002	0.42	0.1564
117	CRIB_2235	Conserved domain protein	1.22	0.0003	1.55	0.0002	0.45	0.1518
118	CRIB_2390	Protein PhnA	1.22	0.0003	1.63	0.0002	0.58	0.0623
119	CRIB_308	UPF0371 protein CD630 08980	1.21	0.0003	1.66	0.0002	0.41	0.2239
120	CRIB_2344	UDP-N-acetylglucosamine 1-carboxyvinyltransferase	1.19	0.0021	1.50	0.0002	0.47	0.2142
121	CRIB_189	Triosephosphate isomerase	1.17	0.0026	2.44	0.0002	0.37	0.3088
122	CRIB_2400	Adenylosuccinate synthetase	1.16	0.0029	1.59	0.0002	0.74	0.0572
123	CRIB_2479	DNA-directed RNA polymerase subunit beta	1.16	0.0184	1.59	0.0002	0.50	0.3437
124	CRIB_1966	Isoleucine--tRNA ligase	1.14	0.0049	1.62	0.0002	0.55	0.1730
125	CRIB_2475	L10_leader	1.14	0.0042	1.76	0.0002	0.40	0.2260

			Glucose vs L-fucose		Glucose vs control		Glucose vs FOS	
	locus	annotation	log2(fold change)	q value	log2(fold change)	q value	log2(fold change)	q value
126	CRIB_1664	Capsular polysaccharide biosynthesis)	1.14	0.0026	1.60	0.0002	0.56	0.0844
127	CRIB_2338	SAM	1.13	0.0329	2.00	0.0006	0.91	0.0405
128	CRIB_12	T-box	1.13	0.0014	1.56	0.0002	0.53	0.0818
129	CRIB_1271	ABC transporter F member 2	1.11	0.0003	1.53	0.0002	0.44	0.1342
130	CRIB_1978	PDZ domain protein	1.09	0.0006	1.67	0.0002	0.42	0.1858
131	CRIB_1673	Elongation factor Ts	1.08	0.0135	1.55	0.0002	0.76	0.0928
132	CRIB_2472	Transcription antitermination protein NusG	1.07	0.0085	1.58	0.0002	0.63	0.1223
133	CRIB_967	3-hydroxyacyl-[acyl-carrier-protein] dehydratase	1.07	0.0033	1.68	0.0002	0.73	0.0484
134	CRIB_961	Acyl carrier protein	1.05	0.0003	1.62	0.0002	0.60	0.0372
135	CRIB_1954	UPF0701 protein YloC	1.03	0.0010	1.61	0.0002	0.57	0.0601
136	CRIB_2346	Glutamine--fructose-6-phosphate aminotransferase	1.02	0.0191	1.65	0.0002	0.74	0.1046
137	CRIB_1902	Thiamin pyrophosphokinase, catalytic region	1.01	0.0014	1.51	0.0002	0.62	0.0485
138	CRIB_2478	DNA-directed RNA polymerase subunit beta	1.00	0.0429	1.66	0.0002	0.50	0.3371
139	CRIB_969	Acetyl-CoA carboxylase, biotin carboxylase, subunit	0.99	0.0298	1.56	0.0002	0.72	0.1138
140	CRIB_970	Acetyl-CoA carboxylase, biotin carboxylase, subunit	0.99	0.0298	1.56	0.0002	0.72	0.1138
141	CRIB_1594	ATP-binding cassette sub- F member 2	0.98	0.0053	1.55	0.0002	0.25	0.4715
142	CRIB_1688	Sodium/amino acid symporter	0.97	0.0196	1.50	0.0002	-0.36	0.3828
143	CRIB_1952	DNA-directed RNA polymerase, omega subunit	0.93	0.0003	1.64	0.0002	0.83	0.0018
144	CRIB_2104	Hypothetical protein	0.92	0.0010	1.55	0.0002	0.80	0.0065
145	CRIB_1953	Guanylate kinase	0.90	0.0066	1.81	0.0002	0.75	0.0268
146	CRIB_1677	DEAD-box ATP-dependent RNA helicase CshA	0.86	0.0021	1.50	0.0002	0.70	0.0159
147	CRIB_2470	Ribosomal protein L33p	0.85	0.0045	1.94	0.0002	0.47	0.1099
148	CRIB_2471	Preprotein translocase, SecE subunit	0.82	0.0122	1.63	0.0002	0.43	0.1698
149	CRIB_1730	KWG Leptospira	0.82	0.0150	1.53	0.0002	0.45	0.1614
150	CRIB_2473	Ribosomal protein L11p/L12e	0.80	0.0451	1.56	0.0002	0.54	0.1647
151	CRIB_180	Glyoxalase protein	0.79	0.0070	1.78	0.0002	0.75	0.0192
152	CRIB_184	33 kDa chaperonin	0.74	0.0316	2.53	0.0002	0.52	0.1242
153	CRIB_1876	Elongation factor P	0.72	0.0641	1.56	0.0002	1.16	0.0056
154	CRIB_1672	Uridylate kinase	0.70	0.0477	1.61	0.0002	0.87	0.0200
155	CRIB_962	Enoyl-(acyl-carrier-protein) reductase	0.67	0.1648	1.60	0.0002	0.76	0.1351
156	CRIB_963	Malonyl CoA-acyl carrier protein transacylase	0.67	0.1648	1.60	0.0002	0.76	0.1351
157	CRIB_17	Ribosomal RNA small subunit methyltransferase A	0.65	0.0151	1.55	0.0002	0.57	0.0469

Table S4. Overview of the genes that were significantly upregulated during growth of R. ilealis CRIB^T in the presence of FOS in comparison to at least one of the other three conditions [glucose, L-fucose and control (absence of additional carbon source)] as identified by RNA-seq analysis. Differential gene expression values that did not meet the criteria for significance (≥1.5 log2(fold change) and q value ≤0.05) are color-coded in dark-grey.

			FOS vs glucose		FOS vs L-fucose		FOS vs control	
	locus tag	annotation	log2(fold change)	q value	log2(fold change)	q value	log2(fold change)	q value
1	CRIB_603	Nucleic-acid-binding protein containing Zn-ribbon	4.12	0.0018	4.51	0.0007	2.59	0.0005
2	CRIB_602	Hypothetical protein	3.88	0.0018	5.34	0.0007	2.81	0.0005
3	CRIB_601	RNA polymerase sigma-70 factor	3.58	0.0018	5.60	0.0007	2.61	0.0005
4	CRIB_150	ABC transporter, substrate-binding protein	2.99	0.0018	2.01	0.0007	2.12	0.0005
5	CRIB_151	*ß-fructofuranosidase*	2.79	0.0018	1.99	0.0007	2.21	0.0005
6	CRIB_1461	PTS system sucrose-specific EIIBC component	2.74	0.0018	1.74	0.0007	1.69	0.0005
7	CRIB_149	Protein LplC	2.69	0.0018	1.73	0.0007	1.68	0.0005
8	CRIB_152	Fructokinase-2	2.64	0.0018	1.81	0.0007	1.88	0.0005
9	CRIB_148	Protein LplB	2.53	0.0018	1.56	0.0007	1.69	0.0005
10	CRIB_1458	Fructokinase-2	2.32	0.0018	1.61	0.0007	1.47	0.0005
11	CRIB_1509	Transcriptional regulator, GntR family	1.86	0.0018	2.21	0.0007	2.31	0.0005
12	CRIB_1507	ABC transporter, permease protein	1.65	0.0018	1.09	0.0007	1.12	0.0005
13	CRIB_1508	ABC transporter related protein	1.65	0.0018	1.09	0.0007	1.12	0.0005
14	CRIB_1459	RDD family protein	1.55	0.0018	0.70	0.0148	0.72	0.0065
15	CRIB_1460	*ß-fructofuranosidase*	1.55	0.0018	0.70	0.0148	0.72	0.0065
16	CRIB_1506	P-loop containing nucleoside triphosphate	1.52	0.0018	0.74	0.0035	0.88	0.0005
17	CRIB_2230	Hypothetical protein	0.56	0.0876	1.58	0.0007	1.40	0.0005
18	CRIB_1633	Virus attachment protein p12 family	0.37	0.3511	1.92	0.0007	0.29	0.4432
19	CRIB_1688	Sodium/amino acid symporter	0.36	0.3828	1.35	0.0007	1.87	0.0005
20	CRIB_1729	LrgA family protein	0.11	0.7094	0.98	0.0007	1.57	0.0005
21	CRIB_877	c-di-GMP-I	0.10	0.7471	2.73	0.0007	0.92	0.0005
22	CRIB_612	FAD-dependent pyridine nucleotide-disulphide	0.10	0.8315	1.60	0.0007	0.18	0.6911
23	CRIB_197	tmRNA	0.09	0.9173	1.94	0.0007	0.85	0.0826
24	CRIB_2369	Veg protein	0.07	0.8691	1.99	0.0007	0.84	0.0165
25	CRIB_2508	Translation initiation factor IF-1	-0.13	0.6697	1.55	0.0007	1.26	0.0005
26	CRIB_892	Ferric anguibactin transport system permease	-0.20	0.6777	1.58	0.0007	0.28	0.5672
27	CRIB_893	Ferric anguibactin transport system permease	-0.20	0.6777	1.58	0.0007	0.28	0.5672
28	CRIB_894	Iron(3+)-hydroxamate import ATP-binding protein	-0.20	0.6777	1.58	0.0007	0.28	0.5672
29	CRIB_2489	30S ribosomal protein S19	-0.20	0.6029	1.41	0.0007	1.56	0.0005

			FOS vs glucose		FOS vs L-fucose		FOS vs control	
	locus tag	annotation	log2(fold change)	q value	log2(fold change)	q value	log2(fold change)	q value
30	CRIB_1596	Transporter	-0.27	0.3010	1.55	0.0007	1.33	0.0005
31	CRIB_189	Triosephosphate isomerase	-0.37	0.3088	0.81	0.0426	2.07	0.0005
32	CRIB_2493	Ribosomal protein L29p/L35e	-0.45	0.1474	1.26	0.0007	1.53	0.0005
33	CRIB_2485	50S ribosomal protein L3	-0.48	0.3819	1.22	0.0219	1.67	0.0005
34	CRIB_184	33 kDa chaperonin	-0.52	0.1242	0.24	0.5266	2.00	0.0005
35	CRIB_2505	Protein translocase subunit SecY	-0.53	0.2790	1.55	0.0007	1.67	0.0005
36	CRIB_187	Glyceraldehyde-3-phosphate dehydrogenase	-0.60	0.2333	1.47	0.0025	2.85	0.0005
37	CRIB_2488	50S ribosomal protein L2	-0.60	0.2694	0.90	0.1002	1.58	0.0005
38	CRIB_2499	30S ribosomal protein S8	-0.62	0.2219	0.99	0.0599	1.66	0.0005
39	CRIB_190	2,3-bisphosphoglycerate-independent	-0.63	0.1970	0.82	0.1019	2.21	0.0005
40	CRIB_188	Phosphoglycerate kinase	-0.65	0.1607	0.87	0.0547	2.24	0.0005
41	CRIB_2487	50S ribosomal protein L23	-0.66	0.4126	1.24	0.1073	1.53	0.0106
42	CRIB_2486	50S ribosomal protein L4	-0.66	0.4126	1.24	0.1073	1.53	0.0106
43	CRIB_2490	50S ribosomal protein L22	-0.67	0.1723	1.02	0.0271	1.70	0.0005
44	CRIB_2506	Adenylate kinase	-0.68	0.2158	1.35	0.0064	1.61	0.0005
45	CRIB_2507	Methionine aminopeptidase 1	-0.68	0.2158	1.35	0.0064	1.61	0.0005
46	CRIB_191	Enolase	-0.69	0.2212	0.75	0.1675	1.75	0.0005
47	CRIB_2231	Aldehyde-alcohol dehydrogenase 2	-0.76	0.3131	3.17	0.0007	2.41	0.0005
48	CRIB_935	TVP38/TMEM64 inner membrane protein YdjZ	-0.79	0.0090	1.02	0.0007	1.82	0.0005
49	CRIB_2496	50S ribosomal protein L24	-0.83	0.1008	1.08	0.0186	1.65	0.0005
50	CRIB_186	Central glycolytic genes regulator	-1.03	0.0200	1.47	0.0007	2.73	0.0005
51	CRIB_2019	Transcription antiterminator LicT	-1.11	0.0018	1.97	0.0007	1.47	0.0005
52	CRIB_2017	PTS system glucose-specific EIIBC component	-1.31	0.0018	2.05	0.0007	1.24	0.0005
53	CRIB_2018	PTS system glucose-specific EIIA component	-1.58	0.0018	2.19	0.0007	1.31	0.0005

Table S5. Overview of the genes that were significantly upregulated during growth of R. ilealis CRIB^T in the presence of L-fucose in comparison to at least one of the other three conditions [glucose, FOS and control (absence of additional carbon source)] as identified by RNA-seq analysis. Differential gene expression values that did not meet the criteria for significance (≥1.5 log2(fold change) and q value ≤0.05) are color-coded in dark-grey.

			L-fucose vs glucose		L-fucose vs FOS		L-fucose vs control	
	locus	annotation	log2(fold change)	q value	log2(fold change)	q value	log2(fold change)	q value
1	CRIB_1298	L-fucose isomerase	6.71	0.0003	5.79	0.0008	4.64	0.0023
2	CRIB_1296	Fucose mutarotase	6.53	0.0003	5.70	0.0008	4.85	0.0023
3	CRIB_1297	Carbohydrate kinase, FFGY family	6.53	0.0003	5.70	0.0008	4.85	0.0023
4	CRIB_1295	L-fuculose phosphate aldolase	6.47	0.0003	5.60	0.0008	4.81	0.0023
5	CRIB_1300	Lactaldehyde reductase	6.37	0.0003	5.50	0.0008	4.48	0.0023
6	CRIB_1294	L-fucose permease	5.66	0.0003	5.05	0.0008	4.25	0.0023
7	CRIB_2242	Beta-N-acetylglucosaminidase/beta-glucosidase	3.50	0.0003	2.75	0.0008	-0.13	0.9987
8	CRIB_614	N-acetylneuraminate lyase	3.44	0.0003	2.75	0.0008	0.71	0.1206
9	CRIB_613	N-acetylmannosamine-6-phosphate 2-epimerase,	2.87	0.0003	2.24	0.0008	0.20	0.9465
10	CRIB_1299	Transcriptional regulator, AraC	2.74	0.0003	2.17	0.0008	0.58	0.0993
11	CRIB_426	L-threonine dehydratase catabolic TdcB	2.72	0.0003	2.05	0.0008	0.31	0.9509
12	CRIB_615	Sodium:solute symporter protein	2.69	0.0003	2.16	0.0008	0.62	0.2396
13	CRIB_616	YhcH/YjgK/YiaL protein	2.58	0.0003	1.95	0.0008	0.42	0.3907
14	CRIB_2011	Carbon starvation protein A	2.45	0.0003	1.72	0.0008	0.08	0.9987
15	CRIB_630	MgsA/globin-like domain-containing protein	2.33	0.0003	1.72	0.0008	-0.42	0.4253
16	CRIB_650	Altronate oxidoreductase	2.31	0.0003	1.65	0.0008	-0.11	0.9987
17	CRIB_651	Altronate hydrolase	2.31	0.0003	1.65	0.0008	-0.11	0.9987
18	CRIB_652	Uronate isomerase	2.29	0.0003	1.61	0.0008	-0.21	0.9804
19	CRIB_649	2-dehydro-3-deoxygluconokinase	2.13	0.0003	1.22	0.0008	-0.54	0.2557
20	CRIB_1439	Hypothetical protein	2.05	0.0490	0.19	0.8672	-0.02	0.9987
21	CRIB_1429	Hypothetical protein	2.05	0.0434	0.33	0.6722	-0.20	0.9987
22	CRIB_756	Hypothetical protein	2.02	0.0073	1.10	0.0801	-0.45	0.8569
23	CRIB_1067	CRISPR-associated nuclease/helicase Cas3	1.96	0.0003	1.85	0.0008	0.40	0.7762
24	CRIB_1302	Hypothetical protein	1.93	0.0151	1.08	0.0865	0.38	0.9420
25	CRIB_1776	Hypothetical protein	1.89	0.0295	0.29	0.6677	-0.13	0.9987
26	CRIB_1066	CRISPR-associated protein cas5	1.88	0.0003	1.89	0.0008	0.19	0.9332
27	CRIB_1696	Perfringolysin O regulator protein	1.88	0.0003	1.29	0.0008	-1.48	0.0023
28	CRIB_1070	CRISPR-associated endoribonuclease Cas2	1.88	0.0003	1.75	0.0008	0.26	0.7962
29	CRIB_1345	Protein of unknown function (DUF2922)	1.81	0.0120	0.66	0.2731	0.09	0.9987

			L-fucose vs glucose		L-fucose vs FOS		L-fucose vs control	
	locus	annotation	log2(fold change)	q value	log2(fold change)	q value	log2(fold change)	q value
30	CRIB_1064	CRISPR-associated cxxc cxxc protein Cst1	1.79	0.0003	1.84	0.0008	0.33	0.7732
31	CRIB_2244	Uronate isomerase	1.77	0.0003	0.46	0.2489	-1.07	0.0510
32	CRIB_2246	2-keto-3-deoxygluconate kinase	1.75	0.0006	0.58	0.2627	-0.63	0.6516
33	CRIB_2247	4-Hydroxy-2-oxoglutarate aldolase /	1.75	0.0006	0.58	0.2627	-0.63	0.6516
34	CRIB_2248	Short-chain dehydrogenase/reductase SDR	1.75	0.0003	0.65	0.1444	-0.42	0.8356
35	CRIB_629	PTS system, lactose/cellobiose family IIC	1.70	0.0003	1.04	0.0008	-0.63	0.1052
36	CRIB_1398	Melibiose carrier protein	1.68	0.0003	0.99	0.0008	-0.15	0.9740
37	CRIB_2245	Mannonate dehydratase	1.68	0.0003	0.46	0.3356	-0.67	0.5157
38	CRIB_1068	CRISPR-associated protein cas4	1.68	0.0003	1.46	0.0008	0.19	0.9678
39	CRIB_1069	CRISPR-associated endonuclease Cas1	1.68	0.0003	1.46	0.0008	0.19	0.9678
40	CRIB_1740	Hypothetical protein	1.65	0.0003	0.78	0.0611	0.07	0.9987
41	CRIB_810	Glycine/sarcosine/betaine reductase selenoprotein B	1.64	0.0003	0.95	0.0014	0.14	0.9928
42	CRIB_1065	Devr CRISPR-associated regulatory protein	1.63	0.0003	1.61	0.0008	0.23	0.9465
43	CRIB_1601	Beta-galactosidase	1.62	0.0003	0.87	0.0060	-0.21	0.9740
44	CRIB_1014	Hypothetical protein	1.53	0.0003	0.62	0.0563	0.36	0.7221
45	CRIB_1013	Antirepressor, phage associated	1.53	0.0003	1.23	0.0008	0.61	0.0753
46	CRIB_1063	CRISPR-associated protein cas6	1.52	0.0003	1.53	0.0008	0.10	0.9987
47	CRIB_1607	Glycoside-Pentoside-Hexuronide (GPH):cation	1.51	0.0003	0.59	0.0124	-0.56	0.1506
48	CRIB_809	PTS system, lactose/cellobiose-specific, IIC	1.50	0.0003	0.74	0.0008	0.07	0.9987
49	CRIB_807	PTS system, lactose/cellobiose-specific, IIA	1.50	0.0016	0.68	0.0975	0.34	0.8569
50	CRIB_937	Domain of Unknown Function (DUF1540)	1.50	0.0006	0.64	0.0624	-0.01	0.9987
51	CRIB_172	Hypothetical protein	1.12	0.0852	2.00	0.0339	0.69	0.7356
52	CRIB_2073	S-layer homoloy domain-containing protein	0.90	0.0427	2.37	0.0008	-0.02	0.9987
53	CRIB_1670	Hypothetical protein	0.56	0.0460	1.08	0.0008	1.77	0.0023
54	CRIB_184	33 kDa chaperonin	-0.74	0.0316	-0.24	0.5265	1.76	0.0023

Table S6. Overview of the genes that were significantly upregulated during growth of R. ilealis CRIB^T in basal medium in absence of an additional carbon source (control condition) in comparison to at least one of the other three conditions (glucose, FOS and L-fucose) as identified by RNA-seq analysis. Differential gene expression values that did not meet the criteria for significance (≥1.5 log2(fold change) and q value ≤0.05) are color-coded in dark-grey.

			Control vs glucose		Control vs FOS		Control vs L-Fucose	
	locus	annotation	log2(fold change)	q value	log2(fold change)	q value	log2(fold change)	q value
1	CRIB_2242	Beta-N-acetylglucosaminidase/beta-glucosidase	3.59	0.0002	2.87	0.0005	0.13	0.9987
2	CRIB_1696	Perfringolysin O regulator protein	3.33	0.0002	2.76	0.0005	1.48	0.0023
3	CRIB_2244	Uronate isomerase	2.81	0.0002	1.52	0.0005	1.07	0.0510
4	CRIB_630	MgsA/globin-like domain-containing protein	2.72	0.0002	2.13	0.0005	0.42	0.4253
5	CRIB_614	N-acetylneuraminate lyase	2.71	0.0002	2.04	0.0005	-0.71	0.1206
6	CRIB_649	2-dehydro-3-deoxygluconokinase	2.65	0.0002	1.76	0.0005	0.54	0.2557
7	CRIB_613	N-acetylmannosamine-6-phosphate 2-epimerase,	2.65	0.0002	2.04	0.0005	-0.20	0.9465
8	CRIB_2334	Ribosomal subunit interface protein	2.55	0.0002	2.62	0.0005	2.02	0.0023
9	CRIB_652	Uronate isomerase	2.46	0.0002	1.81	0.0005	0.21	0.9804
10	CRIB_756	Hypothetical protein	2.42	0.0007	1.53	0.0013	0.45	0.8569
11	CRIB_650	Altronate oxidoreductase	2.38	0.0002	1.75	0.0005	0.11	0.9987
12	CRIB_651	Altronate hydrolase	2.38	0.0002	1.75	0.0005	0.11	0.9987
13	CRIB_426	L-threonine dehydratase catabolic TdcB	2.37	0.0002	1.74	0.0005	-0.31	0.9509
14	CRIB_2246	2-keto-3-deoxygluconate kinase	2.35	0.0002	1.21	0.0029	0.63	0.6516
15	CRIB_2247	4-Hydroxy-2-oxoglutarate aldolase	2.35	0.0002	1.21	0.0029	0.63	0.6516
16	CRIB_2011	Carbon starvation protein A	2.35	0.0002	1.63	0.0005	-0.08	0.9987
17	CRIB_2245	Mannonate dehydratase	2.31	0.0002	1.12	0.0038	0.67	0.5157
18	CRIB_629	PTS system, lactose/cellobiose family IIC	2.31	0.0002	1.66	0.0005	0.63	0.1052
19	CRIB_633	PTS system, lactose/cellobiose family IIA	2.29	0.0002	2.51	0.0005	1.26	0.0023
20	CRIB_632	PTS system, lactose/cellobiose family IIB	2.28	0.0002	2.43	0.0005	1.05	0.0023
21	CRIB_1429	Hypothetical protein	2.23	0.0182	0.52	0.3587	0.20	0.9987
22	CRIB_616	YhcH/YjgK/YiaL protein	2.14	0.0002	1.53	0.0005	-0.42	0.3907
23	CRIB_2248	Short-chain dehydrogenase/reductase SDR	2.13	0.0002	1.06	0.0062	0.42	0.8356
24	CRIB_1299	Transcriptional regulator, AraC	2.13	0.0002	1.58	0.0005	-0.58	0.0993
25	CRIB_1298	L-fucose isomerase	2.06	0.0002	1.14	0.0005	-4.64	0.0023
26	CRIB_615	Sodium:solute symporter protein	2.05	0.0002	1.54	0.0005	-0.62	0.2396
27	CRIB_1607	Glycoside-Pentoside-Hexuronide (GPH):cation	2.04	0.0002	1.15	0.0005	0.56	0.1506
28	CRIB_1312	Glucuronide permease	1.98	0.0002	1.31	0.0005	0.54	0.5611
29	CRIB_1776	Hypothetical protein	1.98	0.0110	0.41	0.4176	0.13	0.9987

			Control vs glucose		Control vs FOS		Control vs L-Fucose	
	locus	annotation	log2(fold change)	q value	log2(fold change)	q value	log2(fold change)	q value
30	CRIB_1704	Maltodextrin import ATP-binding protein MsmX	1.94	0.0002	1.26	0.0005	0.75	0.1076
31	CRIB_1451	Kae1-associated kinase Bud32	1.93	0.0002	1.78	0.0005	1.11	0.0023
32	CRIB_2281	Chaperone protein htpG	1.90	0.0002	0.68	0.0875	1.37	0.0141
33	CRIB_655	TRAP-type C4-dicarboxylate transport system, large	1.88	0.0002	1.09	0.0005	0.52	0.5738
34	CRIB_1300	Lactaldehyde reductase	1.87	0.0002	1.01	0.0005	-4.48	0.0023
35	CRIB_1345	Protein of unknown function (DUF2922)	1.87	0.0023	0.74	0.1094	0.09	0.9987
36	CRIB_2241	Conserved domain protein	1.84	0.0002	0.86	0.0046	0.67	0.2155
37	CRIB_2240	Beta-N-acetylglucosaminidase/beta-glucosidase	1.83	0.0002	0.81	0.0368	0.41	0.8569
38	CRIB_1510	RNA polymerase sigma factor RpoD	1.82	0.0002	1.48	0.0005	0.92	0.0150
39	CRIB_2249	Glucuronide carrier protein homolog	1.81	0.0002	0.41	0.1967	0.63	0.2866
40	CRIB_1601	Beta-galactosidase	1.80	0.0002	1.08	0.0005	0.21	0.9740
41	CRIB_654	TRAP-type C4-dicarboxylate transport system, small	1.80	0.0002	1.00	0.0005	0.56	0.2303
42	CRIB_1398	Melibiose carrier protein	1.79	0.0002	1.13	0.0005	0.15	0.9740
43	CRIB_708	Cystathionine beta-lyase PatB	1.77	0.0002	1.51	0.0005	0.86	0.0192
44	CRIB_2138	spore_II_AA: anti-sigma F factor antagonist	1.75	0.0002	1.47	0.0005	0.82	0.0115
45	CRIB_978	Hypothetical protein	1.74	0.0015	0.67	0.0989	0.78	0.4437
46	CRIB_1275	Glucosamine-6-phosphate deaminase	1.73	0.0002	0.83	0.0013	0.93	0.0041
47	CRIB_171	Small acid-soluble spore protein	1.73	0.0002	1.10	0.0023	0.38	0.7973
48	CRIB_1384	Urease accessory protein UreE	1.67	0.0002	0.68	0.0134	0.45	0.5655
49	CRIB_1386	Urease subunit alpha	1.66	0.0002	0.76	0.0023	0.29	0.7662
50	CRIB_1296	Fucose mutarotase	1.66	0.0002	0.84	0.0029	-4.85	0.0023
51	CRIB_1297	Carbohydrate kinase, FFGY family	1.66	0.0002	0.84	0.0029	-4.85	0.0023
52	CRIB_1066	CRISPR-associated protein cas5	1.65	0.0002	1.69	0.0005	-0.19	0.9332
53	CRIB_1295	L-fuculose phosphate aldolase	1.65	0.0002	0.78	0.0005	-4.81	0.0023
54	CRIB_2314	Orotate phosphoribosyltransferase	1.62	0.0002	1.68	0.0005	1.21	0.0023
55	CRIB_554	Hypothetical protein	1.60	0.0002	1.48	0.0005	1.23	0.0041
56	CRIB_1321	Domain of unknown function (DUF3794)	1.58	0.0002	0.56	0.0378	0.23	0.9301
57	CRIB_1070	CRISPR-associated endoribonuclease Cas2	1.58	0.0002	1.49	0.0005	-0.26	0.7962
58	CRIB_1521	Protein of unknown function (DUF1659)	1.55	0.0067	0.73	0.1146	0.22	0.9987
59	CRIB_2413	Transcriptional regulator	1.55	0.0002	1.08	0.0005	0.47	0.4253
60	CRIB_1740	Hypothetical protein	1.55	0.0004	0.69	0.0355	-0.07	0.9987
61	CRIB_603	Nucleic-acid-binding protein containing Zn-ribbon	1.54	0.0002	-2.59	0.0005	1.92	0.0023

			Control vs glucose		Control vs FOS		Control vs L-Fucose	
	locus	annotation	log2(fold change)	q value	log2(fold change)	q value	log2(fold change)	q value
62	CRIB_1385	AmiS/UreI transporter	1.54	0.0002	1.01	0.0005	0.41	0.6212
63	CRIB_591	Extracellular solute-binding protein 1	1.53	0.0002	0.91	0.0005	0.47	0.4394
64	CRIB_1606	Beta-galactosidase BgaA	1.52	0.0002	0.87	0.0005	0.35	0.5924
65	CRIB_1067	CRISPR-associated nuclease/helicase Cas3	1.52	0.0002	1.44	0.0005	-0.40	0.7762
66	CRIB_1216	Cold shock protein	1.51	0.0035	0.73	0.0637	0.26	0.9890
67	CRIB_1381	Urease accessory protein UreH	1.51	0.0002	0.74	0.0023	0.40	0.4800
68	CRIB_1382	Urease accessory protein UreG	1.51	0.0002	0.74	0.0023	0.40	0.4800
69	CRIB_1302	Hypothetical protein	1.51	0.0208	0.69	0.1623	-0.38	0.9420
70	CRIB_1064	CRISPR-associated cxxc cxxc protein Cst1	1.41	0.0002	1.50	0.0005	-0.33	0.7732
71	CRIB_23	Stage V sporulation protein G, SpoVG	1.30	0.0002	1.77	0.0005	1.94	0.0023
72	CRIB_178	UPF0597 protein CD630 32320	1.26	0.0002	1.08	0.0038	1.61	0.0023
73	CRIB_602	Hypothetical protein	1.09	0.0280	-2.81	0.0005	2.54	0.0023
74	CRIB_373	6S	1.04	0.0007	0.57	0.0140	1.90	0.0023
75	CRIB_601	RNA polymerase sigma-70 factor	0.99	0.0218	-2.61	0.0005	3.00	0.0023
76	CRIB_2073	S-layer homoloy domain-containing protein	0.88	0.0344	2.38	0.0005	0.02	0.9987
77	CRIB_1690	Thioredoxin, possibly sporulation protein	0.24	0.4995	0.47	0.1949	1.65	0.0023
78	CRIB_1633	Virus attachment protein p12 family	0.07	0.8755	-0.29	0.4432	1.64	0.0192
79	CRIB_2038	Heme oxygenase	-0.14	0.6875	0.38	0.2749	1.88	0.0023
80	CRIB_877	c-di-GMP-I	-0.86	0.0084	-0.92	0.0005	1.82	0.0023

6 CHAPTER

Comparative genomics and functional analysis of the genus *Romboutsia* provides insight into adaptation to an intestinal lifestyle

Jacoline Gerritsen*, Bastian Hornung*, Ivelina Staneva, Jarmo Ritari, Lars Paulin, Ger T. Rijkers, Peter Schaap, Willem M. de Vos and Hauke Smidt

* These authors contributed equally to this work.

Abstract

Cultivation-independent surveys have shown that the recently described genus *Romboutsia* within the family *Peptostreptococcaceae* is more diverse than previously acknowledged. The majority of *Romboutsia*-associated 16S rRNA gene sequences have an intestinal origin. However, the specific roles that *Romboutsia* species play in the intestinal tract are largely unknown. To gain more insight in the metabolic and functional capabilities of the genus *Romboutsia*, strain FRIFI was isolated from human ileostoma effluent, and the complete genome of strain FRIFI was sequenced together with that of *Romboutsia lituseburensis* A25K^T. On the basis of phenotypic, phylogenetic and genomic analyses it was concluded that strain FRIFI constitutes a novel species within the genus *Romboutsia* for which the name *Romboutsia hominis* sp. nov. is proposed. An evaluation of the common traits of the recently defined genus *Romboutsia* was done based on comparative genome analysis of '*R. hominis*' FRIFI and *R. lituseburensis* A25K^T together with the recently elucidated genome of the type species *Romboutsia ilealis* CRIB^T and the draft genome of '*Clostridium dakarense*' FF1^T, which was found to be related to the *Romboutsia* species. These analyses have shown that the genus *Romboutsia* contains a versatile array of metabolic capabilities with respect to carbohydrate utilization, fermentation of single amino acids, anaerobic respiration and metabolic end products. *R. ilealis* CRIB^T and '*R. hominis*' FRIFI, both of small intestinal origin, have a significantly smaller genome compared to the other two *Romboutsia* species. Main differences were found in their ability to utilize specific carbohydrates, to synthesize vitamins and other cofactors, and to assimilate nitrogen. In addition, differences were found in the presence of bile metabolism and motility-related gene clusters.

Introduction

Recently the genus *Romboutsia* within the family *Peptostreptococcaceae* was described to allocate *Romboutsia ilealis*, a novel species isolated from the small intestinal tract of rats, and *Romboutsia lituseburensis*, which was previously known as *Clostridium lituseburensis* (**Chapter 4** [39]). Members of the *Peptostreptococcaceae* are commonly found in diverse environments, but sequence-based studies have reported the presence of *Romboutsia*-like 16S rRNA gene sequences predominantly in samples of intestinal origin. Clone-library based molecular surveys have shown the occurrence of *Romboutsia*-like 16S rRNA sequences in intestinal content samples (duodenum, jejunum, ileum and colon) from dogs [108] and cows [98], and faecal samples from rats [1], polar bears [43] and porpoises [80]. In addition, more recent pyrosequencing-based studies have identified similar phylotypes in faecal samples from humans and other mammals [69], human ileal biopsies [70], mouse faecal samples [131], ileal biopsies from pigs [77], and ileal contents from rats [128]. There have been recent reports on the isolation of related strains, such as strain TC1 that was isolated from the hide of a cow, but was probably derived from faecal contamination of the hide [81], and '*Clostridium dakarense*' that was isolated from the stool of a 4-month-old Senegalese child [71]. Altogether these studies suggest that the genus *Romboutsia* is probably more diverse than previously appreciated and many of its members may have the intestinal tract as a natural habitat. Nevertheless, the specific roles that members of the *Romboutsia* species play especially in the small intestinal tract remain largely unknown, because of the still limited availability of cultured representatives and their genomes, and the restricted accessibility of their natural habitat, the small intestinal tract.

A genomic and functional analysis was done on *R. ilealis* CRIB^T, the type strain of the genus *Romboutsia* (**Chapter 5**). This analysis showed *R. ilealis* CRIB^T to be a flexible anaerobe that is adapted to a nutrient-rich environment in which carbohydrates and exogenous sources of amino acids and vitamins are abundantly available. In addition, genome mining revealed potential mechanisms by which *R. ilealis* is adapted to the small intestinal environment.

The type strain of the second validly described species within the genus *Romboutsia*, *R. lituseburensis* A25K^T, is not of intestinal origin, but was originally isolated from soil and humus from Côte d'Ivoire [94]. Based on 16S rRNA gene identity, very similar strains have been isolated in recent years: strain H17 that was isolated from the main anaerobic digester of a biogas plant (EU887828; Nayak *et al.*), strain VKM B-2279 was isolated from a p-toluene sulfonate degrading community [105], strain 2ER371.1 that was isolated from waste of biogas plants [23], and strain E2 that was isolated from a cellulose degrading community enriched from mangrove soil [37]. In contrast to the intestinal origin of the overwhelming majority of *Romboutsia*-associated 16S rRNA gene sequences retrieved by cultivation-independent surveys, all these strains have been isolated from non-intestinal environments, thus suggesting an adaption of *R. lituseburensis* to life outside a host.

To gain more insight in metabolic and functional capabilities of the genus *Romboutsia* we set out to isolate a *Romboutsia* strain of human small intestinal origin, resulting in the isolation of strain FRIFI from ileostoma effluent of a human adult and the subsequent proposal of *Romboutsia hominis* sp. nov. Here we present the complete genomes of '*R. hominis*' FRIFI and *R. lituseburensis* A25K^T, together with an evaluation of the common traits of this recently defined genus based on comparative genome analysis, including the recently elucidated genome of the type species *R. ilealis* CRIB^T (**Chapter 5**) and the draft genome of *C. dakarense* FF1^T.

Materials and methods

Sample collection

Ileostomy effluent samples were obtained from seven human volunteers with an ileostoma, but who were otherwise healthy. The study was approved by the local medical ethical committees of University Medical Centre Utrecht (The Netherlands) and University Medical Centre St. Radboud (Nijmegen, The Netherlands). A waiver for informed consent was given due to the non-invasive nature of the study. The samples were collected by taking material directly from the ileostomy bag and preservation either as glycerol suspension (25% v/v) for subsequent cultivation or storage at –80°C for DNA extraction. In addition, ileum content samples were collected from Sprague-Dawley rats as part of a previous project (**Chapter 3** [40]).

DNA extraction and quantitative PCR

Total DNA was extracted from ileostoma effluent samples (1 mL), enrichment culture samples (1 mL), and pieces of rat ileum (0.25 g), using the repeated bead-beating method previously described [103]. For the rat ileum samples an additional digestion step was done before DNA extraction by incubation in 1 mL lysisbuffer containing 10 µL proteinase K (20 mg/mL) for 1 h at 55°C [34]. DNA was quantified using a NanoDrop ND-1000 spectrophotometer (NanoDrop Technologies, Wilmington, DE). Extracted DNA was stored at –20°C for downstream applications. Relative abundance of *R. ilealis* was determined using quantitative PCR (qPCR) and the primer pair CRIB-61F/CRIB-235R as described previously (**Chapter 3** [40]).

Isolation of *R. hominis* FRIFI

Isolation of strain FRIFI was performed by serial dilution of an ileostoma effluent sample in liquid anoxic bicarbonate-buffered CRIB medium (pH 7.0), prepared as previously described (**Chapter 4** [39]). Enrichment was achieved by repeated serial dilutions. Cell pellets were collected of 1 mL samples taken from the enrichment cultures and total DNA was extracted from using the FastDNA Spin kit for soil (MP Biomedicals, Solon, OH) according to the manufacturer's instructions. The increase in relative abundance of the

target phylotype was followed in the subsequent serial dilutions using qPCR and the *R. ilealis*-specific primer pair CRIB-61F/CRIB-235R as described previously (**Chapter 3** [40]).

After primary isolation, accomplished by repeated rounds of fast transfers (8 hours of incubation), the strain was purified by repeated plating of single colonies on solid DSMZ 104b medium containing 0.8% gelrite (Carl Roth). The purified strain was stored as glycerol suspension (25% v/v) at –80°C. *R. hominis* FRIFI^T was deposited at the culture collection of the Leibniz-Institut DSMZ (Braunschweig, Germany) under accession number DSM 28814.

PCR amplification of 16S rRNA genes and bar-coded pyrosequencing

A subset of samples was selected for 16S rRNA gene-based microbial composition profiling, which was done as previously described [119]. Briefly, barcoded amplicons from the V1-V2 region of 16S rRNA genes were generated by PCR using the 27F-DegS primer [118] that was appended with the titanium sequencing adaptor A and an 8-nt sample-specific barcode [46] at the 5′-end, and an equimolar mix of two reverse primers (338R I and II [45] that are based on three previously published probes EUB 338 I, II, and III [24]) that carried the titanium adaptor B at the 5′-end. PCRs were performed using a thermocycler GS0001 (Gene Technologies, Braintree, UK) in a total volume of 100 µL containing 1× HF buffer (Finnzymes, Vantaa, Finland), 2 µL PCR Grade Nucleotide Mix (Roche, Diagnostics GmbH, Mannheim, Germany), 2 U of Phusion Hot Start II High-Fidelity DNA polymerase, 500 nM of a forward and the reverse primer mix (Biolegio BV, Nijmegen, The Netherlands), and 0.2–0.4 ng/µL of template DNA. The amplification program consisted of an initial denaturation at 98°C for 30 s; 30 cycles of denaturation at 98°C for 10 s, annealing at 56°C for 20 s, and elongation at 72°C for 20 s; and a final extension at 72°C for 10 min The size of the PCR products (~375 bp) was confirmed by gel electrophoresis using 5 µL of the amplification reaction mixture on a 1% (w/v) agarose gel containing 1× SYBR Safe (Invitrogen, Grand Island, NY). PCR products were purified with the High Pure Cleanup Micro Kit (Roche) using 10 µL nuclease-free water for elution and quantified using the Qubit fluorometer and the Quant-iTdsDNA high-sensitivity assay kit according to manufacturer's instructions (Invitrogen). Purified PCR products were pooled in approximately equimolar amounts and subjected to agarose gel size selection by electrophoresis, followed by excision and purification using a DNA gel extraction kit (Millipore, Billerica, MA). Purified amplicon pools were subjected to pyrosequencing using a Genome Sequencer FLX in combination with titanium chemistry (GATC-Biotech, Konstanz, Germany).

Sequence processing, taxonomic assignment, and analysis of diversity and phylogeny

Pyrosequencing data were analysed using the QIIME pipeline v1.8.0 [14]. During demultiplexing a quality filtering step was performed using default settings, discarding reads shorter than 200 bp and reads with an average quality score of <25. Sequencing

error correction was performed using Acacia v1.53 [9] and subsequent chimera removal was performed using USEARCH v6.1 by both reference-based and *de novo* methods [29]. Clustering of the sequences in operational taxonomic units (OTUs) was done using UCLUST v1.2.22q [29] based on open reference OTU picking at 97% sequence identity using the SILVA SSURef dataset release 111 [93]. Taxonomy was assigned to the representative sequences for each OTU with USEARCH 6.1 against the SILVA SSURef dataset release 111. OTU tables at different taxonomic levels were generated with QIIME.

16S rRNA gene-based phylogenetic analysis

Near full-length 16S rRNA gene fragments from *R. hominis* FRIFI were amplified and sequenced as previously described (**Chapter 4** [39]). The obtained sequence was aligned with reference sequences using the online SINA aligner [92]. Phylogenetic analysis was carried out using ARB [75]. A neighbour-joining tree was constructed using the "The all-species living tree" project dataset release LTPs115 [130]. Bootstrap values were calculated based on 1000 replications. Tree topology was confirmed using maximum-likelihood methods.

Growth conditions and strain characterization

R. hominis FRIFI, *R. lituseburensis* A25K^T (DSM 797) and *R. ilealis* CRIB^T (DSM 25109) were routinely grown in liquid CRIB-medium (pH 7.0). For induction of sporulation the strains were grown in Duncan-Strong medium [28] and cooked meat medium (226730; BD).

Substrate utilization of *R. hominis* FRIFI was determined using the API 50 CH and API 20 A systems (bioMérieux) according to the manufacturer's instructions except that liquid peptone-yeast extract (PY) medium was used for inoculation and the strips were incubated anoxically as previously described (**Chapter 4** [39]).

For analysis of metabolites produced during growth, *R. lituseburensis* A25K^T and *R. hominis* FRIFI were grown in liquid PY medium to which either D-glucose, D-fructose or glycerol were added to a final concentration of 0.5% (w/v) under anoxic conditions. Growth was determined spectrophotometrically by measuring optical density at 600 nm. Metabolites (including short chain fatty acids) were analysed by HPLC using a Metacarb 67H column (Varian, Middelburg, The Netherlands). Acetate, butyrate, ethanol, formate, lactate, propionate, iso-butyrate, iso-valerate, valerate, succinate, 1,2-propanediol and 1,3-propanediol were identified and quantified with standard curves prepared for each individual compound.

For DNA-DNA hybridization (DDH) analyses, *R. lituseburensis A25K^T*, *R. hominis* FRIFI and *R. ilealis* CRIB^T were grown overnight at 37°C, and cells were sent to the Leibniz-Institut DSMZ. Cells were disrupted by using a French pressure cell (Thermo Spectronic), and the DNA in the crude cell lysate was purified by chromatography on hydroxyapatite as described by Cashion *et al.* [16]. DDH was carried out as previously described [26] under

consideration of the modifications described by Huss *et al.* [51]. DDH values reported are based on mean percentage reassociation ± standard deviation (n=2).

For genomic DNA extraction, *R. lituseburensis* A25K^T and *R. hominis* FRIFI were grown overnight at 37°C. DNA extraction was performed as described previously [117]. DNA quality and concentrations were determined by spectrophotometric analysis using NanoDrop (Thermo Scientific) and by electrophoresis on a 1.0% (w/v) agarose gel. DNA was stored at 4°C until subsequent sequencing.

Genome sequencing and assembly

Genome sequencing of *R. hominis* FRIFI was carried out at the Institute of Biotechnology, University of Helsinki (Finland) on a PacBio RS II, resulting in 134,366 PacBio reads and a total amount of 464,930,600 bases. Assembly was performed with PacBio SMRT analysis pipeline v2.2 and the HGAP protocol [18]. Default settings were used except for: minimum sub-read length 500, minimum polymerase read length quality 0.80, minimum seed read length 7,000, split target into chunks 1, alignment candidate per chunk 24, genome size 3,000,000, target coverage 30, overlapper error rate 0.06, overlapper mini length 40, and overlapper k-mer 14.

Genome sequencing of *R. lituseburensis* A25K^T was performed at GATC Biotech (Konstanz, Germany). One MiSeq library was generated on an Illumina MiSeq Personal Sequencer with 250 bp paired-end reads and an insert size of 500 bp, which resulted in 772,051 paired-end reads. Additionally one PacBio library was generated on a PacBio RS II, which resulted in 441,151 subreads and in total 998,181,178 bases. A hybrid assembly was carried out with MiSeq paired-end and PacBio circular consensus sequencing (CCS) reads. For the MiSeq paired-end reads first all rRNA reads were removed with SortMeRNA v1.9 [65] using default settings except for an increased error value of 20% for the adapters, and using the reverse complement of the adapters as well. Quality trimming was performed with PRINSEQ Lite v0.20.0 [104] with a minimum sequence length of 40 and a minimum quality of 30 on both ends and as mean quality of the read. The assembly was done in parallel using two different assemblers. Ray v2.3 was used for the MiSeq paired-end dataset and the PacBio CCs dataset, using default settings except for a k-mer value of 75. The PacBio SMRT analysis pipeline v2.2 was run on the SMRT-cell subreads with the protocol RS_HGAP_Assembly_2, using default settings except that the number of seed read chunks was set to 1, the minimum seed read length was set to 7000, the alignment candidate per chunk was set to 24 and the estimated genome size was reduced to 4 MB. Both assemblies were merged, and duplications were identified based on BLASTN hits. Duplicate contigs were discarded if they had a hit with at least 99% sequence identity within a bigger contig, which spanned at least 98% of contig query length. Furthermore, contigs with a length of less than 500 bp were discarded. The remaining contigs were merged with CAP3 v.12/21/07 [49] with an overlap length cut-off of 5,000 bp and a minimum

identity of 90%. A circular element was detected within this assembly, based on BLASTP results of the predicted proteins (e-value 0.0001), and this was excluded from the further assembly process, but added to the final assembly result. Scaffolding of the contigs was done with SSPACE-LongRead [8] and the PacBio CCS reads using default settings. Further scaffolding was done with Contiguator v2.7.4 [35] using the genome of *R. hominis* FRIFI as reference genome and applying default settings.

Genome annotation

Annotation was carried out with an in-house pipeline. Prodigal v2.5 was used for prediction of protein coding DNA sequences (CDS) [52], InterProScan 5RC7 for protein annotation [50], tRNAscan-SE v1.3.1 for prediction of tRNAs [73] and RNAmmer v1.2 for the prediction of rRNAs [66]. Additional protein function predictions were derived via BLAST identifications against the UniRef50 [109] and Swissprot [116] databases (download August 2013). Subsequently, the annotation was further enhanced by adding EC numbers via PRIAM version March 06, 2013 [19]. Non-coding RNAs were identified using rfam_scan.pl v1.04, on release 11.0 of the RFAM database [11]. CRISPRs were annotated using CRISPR Recognition Tool v1.1 [7]. A further step of automatic curation was performed by weighting the annotation of the different associated domains, penalizing uninformative functions (e.g. "Domain of unknown function"), and prioritizing functions of interest (e.g. domains containing "virus", "phage", "integrase" for phage related elements; similar procedure for different other functions).

Whole genome analysis

Copy numbers of the 16S rRNA gene from published genomes were derived from the rrnDB v4.0.0 [68]. Average nucleotide identity (ANI) values were calculated with JSpecies v1.2.1 [100] by pairwise comparisons of available genomes within the family *Peptostreptococcaceae*.

Homology between the CDS of the *Romboutsia* strains was determined via best bidirectional BLAST hit [85] at the amino acid level with an e-value cut-off of 0.0001. To evaluate the core and pan metabolism of the *Romboutsia* strains, the four annotated genomes were supplied to Pathway tools v18 [56], and a limited amount of manual curation was performed to remove obvious false positives. Next, the pathway databases were exported via the built-in lisp interface and the exported data was merged. A reaction was considered to be in the core metabolism if it was present in all four databases, else it was considered to be in the pan metabolism. Both parts were then separately or jointly reimported into Pathway tools for further analyses.

Assignment of sporulation-related genes

Genes were matched to the list of essential and non-essential sporulation-related genes compiled by Galperin *et al.* [36] via different methods. Firstly, the protein-coding genes *of*

Bacillus subtilis subsp. *subtilis* 168 were annotated via InterProScan and the respective *B. subtilis* sporulation-related proteins were matched to the proteins of the four *Romboutsia* genomes, if they contained at least 50% of the same domains. In case multiple matches were possible, the match with the highest number of similar domains was picked. The matches were manually curated and arbitrary proteins and/or false hits were excluded. For every protein that did not have any match via the domains, the best bidirectional BLAST hit (e-value cut-off of 0.0001) was used instead. Secondly, the genome of *R. ilealis* CRIB^T was manually curated with respect to putative sporulation-related genes. In case the genomes of the other *Romboutsia* strains did not have any corresponding match for one of the proteins whereas a manually curated hit was present in *R. ilealis* CRIB^T, the best bidirectional hit was assigned. Genomes were manually checked for further missing essential sporulation- and germination-related genes as defined by Galperin *et al.* [36]. Function curation was performed with assistance of the *B. subtilis* wiki (http://subtiwiki.uni-goettingen.de/).

Nucleotide sequence accession number

The genomes of *R. hominis* FRIFI^T and *R. lituseburensis* A25K^T (raw data and annotated assembly) have been deposited at the European Nucleotide Archive under project numbers PRJEB7106 and PRJEB7306, respectively.

Results and discussion

Isolation of *R. hominis* FRIFI

To gain more insight in metabolic and functional capabilities of the genus *Romboutsia* in relation to a small intestinal habitat, the aim was to obtain a *Romboutsia* isolate of human small intestinal origin. Ileostoma effluent samples were collected and based on a pre-screening for the presence of 16S rRNA gene sequences similar to the rat-derived strain of *R. ilealis* using qPCR and the primer pair CRIB-61F/CRIB-235R previously designed to specifically detect *R. ilealis* (**Chapter 3** [40]). Based on these results, four samples were chosen to start enrichment cultures. In the end, strain FRIFI was isolated from an ileostoma effluent sample (IS) obtained from a human adult after repeated serial dilution of the sample in anoxic bicarbonate-buffered medium (pH 7.0) supplemented with rich-medium components such as peptone, yeast extract and beef extract, different carbon sources (glucose, cellobiose, maltose, soluble starch) and a low amount of bile salts (**Chapter 4** [39]). Subsequent analysis of the near full-length 16S rRNA gene sequence that was obtained from the isolate demonstrated that strain FRIFI was, as expected from the match with the *R. ilealis*-specific primer pair, related to *R. ilealis* CRIB^T (97.7% 16S rRNA gene sequence identity).

Based on the pre-screening using the *R. ilealis*-specific primer pair, a subset of samples was selected for 16S rRNA gene-based microbial composition profiling. Firstly, these

were the four ileostoma effluent samples that were selected as inocula for enrichment cultivation based on positive signals in the qPCR, including the sample from which strain FRIFI was isolated (IS5). Secondly, three enrichment culture samples were selected, including the enrichments that eventually had led to the isolation of strain FRIFI (E2 and E3). Furthermore, seven rat ileum samples that showed a high relative abundance of *R. ilealis* as determined by qPCR were selected from a previous study leading to the identification and subsequent isolation of R. ilealis CRIB^T (**Chapter 3** [40] and **Chapter 4** [39]).

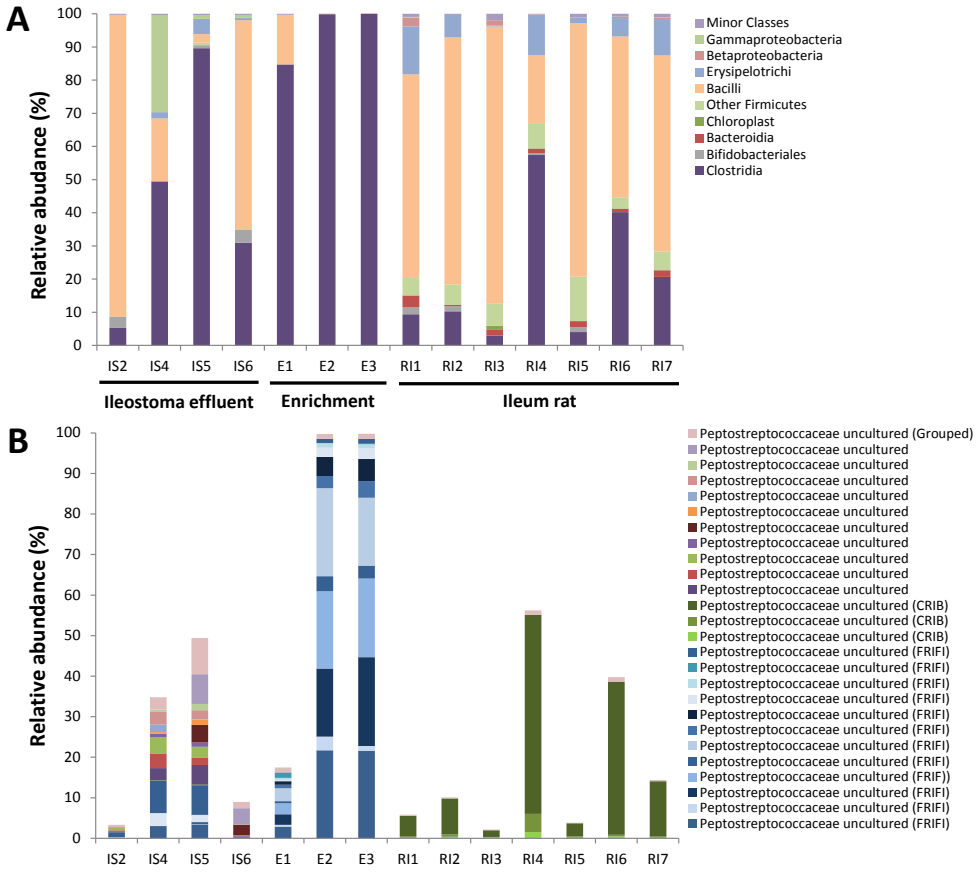

Figure 1.(**A**) Relative abundance of the dominant microbial class level groups in selected intestinal samples and enrichment culture samples. Sample codes indicate ileostoma effluent samples from human volunteers (IS2, IS4, IS5 and IS6), enrichment cultures from ileostoma effluent (E1-E3), and samples from the ileum of rats (RI1-RI7). Class level groups with a relative abundance lower than 1% in all samples were grouped in 'Minor classes'. (**B**) Relative abundance of individual OTUs classified as family *Peptostreptococcaceae*.

Microbial composition profiling was done using pyrosequencing technology, confirming that the enrichment samples contained a highly enriched phylotype that was present in the original ileostoma effluent sample, and that corresponded to the 16S rRNA gene of the newly isolated *R. hominis* FRIFI (Figure 1). At the OTU level (97% sequence identity) this species was also present in the other three ileostoma effluent samples, but in none of the rat ileum samples. In contrast, the rat ileum samples contained phylotypes that were highly similar to the 16S rRNA gene sequence of *R. ilealis* CRIB^T, but that were absent or in low abundance in the ileostoma effluent samples. Whereas the richness of the *R. ilealis*-like phylotypes was low in the rat ileum samples, the ileostoma effluent samples contained multiple similar but distinct phylotypes. This suggests that multiple populations of *Romboutsia*-like phylotypes can co-occur in the human intestinal tract of which one representative was enriched and isolated, *R. hominis* FRIFI.

General features of the *Romboutsia* genomes

To gain more insight in metabolic and functional capabilities of the genus *Romboutsia*, the genomes of *R. hominis* FRIFI and *R. lituseburensis* strain A25K^T were sequenced and compared to the genome *R. ilealis* CRIB^T that was previously sequenced (**Chapter 5**). In addition, the recently partially sequenced genome of *Clostridium dakarense* strain FF1^T was included for comparative genome analysis as well [71], since based on 16S rRNA gene sequence analysis this strain was found to be closely-related to the *Romboutsia* species *R. ilealis* and *R. lituseburensis* (97.3% and 96.0% 16S rRNA gene sequence identity, respectively) and should therefore most likely be assigned to the genus *Romboutsia* as well.

Table 1. General features of the *Romboutsia* genomes

	R. hominis FRIFI	R. lituseburensis A25K^T	R. ilealis CRIB^T	C. dakarense FF1^T
Status current assembly of the chromosome	1 scaffold, 1 gap	2 scaffolds, 11 gaps	1 scaffold, 3 gaps	257 scaffolds
Genome size (Mb)	3.09	3.88	2.58	3.74
Chromosome size (Mb)	3.09	3.78	2.52	3.74
Plasmid size (Mb)	-	0.98	0.06	-
G+C content (%)	27.8	28.2	27.9	28.0
Total no. of CDS	2,852	3,662	2,359	3,843
No. of rRNA genes				
16S rRNA genes	16	_*	14	_*
23S rRNA genes	17	_*	14	_*
5S rRNA genes	15	_*	14	_*
No. of tRNAs	107	118	109	65
No. of ncRNAs	82	116	28	67
CRISPR repeats	-	-	1*71	1*41

* Number of rRNA genes cannot accurately be estimated since a) some of rRNA genes are situated next to an assembly gap of unknown size and may therefore be duplicates (*R. lituseburensis* A25K^T) and b) the quality of the assembly was too low (*C. dakarense* FF1^T).

R. hominis FRIFI^T contains a single, circular chromosome of 3,090,335 bp with an overall G+C content of 27.8% (Table 1). The chromosome encodes 2,852 predicted coding sequences (CDS) of which 83% have a function assigned. *R. lituseburensis* strain A25K^T contains one circular chromosome of 3,776,615 bp and one circular plasmid of 97,957 bp, with an overall G+C content of 28.2% (Table 1). The chromosome encodes 3,535 CDS and the plasmid 123 CDS, of which 82% have a function assigned. In addition, one segment of 4,101 bp, containing one 5S rRNA and four CDS (one two-component system and two subunits of an ABC transporter), could not be placed. The plasmid of *R. lituseburensis* A25K^T encodes two plasmid replication proteins, transporters, transcription factors, hydrolases and acyltransferases.

Impact of a high number of ribosomal operons on sequence assembly efforts

Gaps in whole genome assemblies are usually located in repetitive regions that include ribosomal operons that can appear multiple times in the genome. Also for the *Romboutsia* genomes, the presence of a high number of rRNA operons has been problematic for genome assembly. The assembly of *R. hominis* FRIFI contains one gap situated in a long stretch of ribosomal operons. The assembly of the chromosome of *R. lituseburensis* A25K^T contains eleven gaps, of which six are generated due to the scaffolding with the reference. Nine of the eleven gaps are situated within or neighbouring rRNA operons or tRNA clusters. For *C. dakarense* FF1^T the number of ribosomal operons could not be determined due to the lower quality of the current assembly that consists of 257 scaffolds.

A total of 16 copies of the 16S rRNA gene were identified in the genome of *R. hominis* FRIFI. This is the highest copy number reported for the 16S rRNA gene in prokaryotic species to date (see Table S1 in the supplemental material). High numbers of rRNA operons have been proposed to be indicative for fast growth and to allow microbes to respond quickly to changes in available resources [60]. In addition, it has been shown for *B. subtilis* that multiple rRNA operons are essential for efficient cell growth and sporulation, but maybe more importantly, also for outgrowth after germination of spores [129]. This suggests that a high copy number of the rRNA operon is essential for successful sporulation and germination. Most of the species currently containing the highest numbers of reported rRNA operons do indeed belong to the spore-forming families *Bacillaceae* and *Clostridiaceae* (Table S1). Pairwise sequence identity of the 16S rRNA sequences showed that within the genome of strain FRIFI there was an average sequence identity of 99.3% and the lowest identity between individual copies was 98.4%. Sequence divergence in the 16S rRNA gene is not uncommon within individual prokaryotic genomes [88, 121]. However, it should be noted that for *R. hominis* FRIFI the divergence is located in the V1-V2 region of the 16S rRNA gene (see Figure S1 in the supplemental material), which is commonly used for sequence-based bacterial community analyses [95]. In this region the average sequence identity was only 96.5% and the lowest identity was only 92.3%. Consequence of this

divergence is that during identity clustering into operational taxonomic units (OTUs) the different copies of the 16S rRNA gene of *R. hominis* FRIFI end up in different OTUs, even at the level of 97% identity, resulting in an overestimation of the diversity in *Romboutsia* phylotypes. In addition, the *R. ilealis*-specific primer pair used to evaluate the enrichment process of strain FRIFI is only targeting half of the copies (Figure S1), since the variation is positioned in the regions targeted by the primers CRIB-61F and CRIB-235R. In comparison, *R. ilealis* CRIB^T contains little variation in the 16S rRNA gene sequence regarding both the V1-V2 region and the whole 16S rRNA gene (both >99% sequence identity), despite that also in this genome 14 copies of the 16S rRNA gene were identified.

Phylogeny of the genus *Romboutsia*

A 16S rRNA gene based neighbour joining-tree was constructed with a representative copy of the 16S rRNA gene of the four species *R. ilealis, R. lituseburensis, R. hominis* FRIFI and *C. dakarense* and type strains of close-related species (Figure 2). Based on their 16S rRNA gene sequence these four species form a monophyletic group and both *R. hominis* FRIFI and *C. dakarense* can be assigned to the genus *Romboutsia*.

DNA-DNA hybridization (DDH) is still widely used in bacterial taxonomy in cases where strains share more than 97% 16S rRNA gene sequence identity [111]. DDH was done for *R. ilealis* CRIB^T with *R. lituseburensis* A25K^T (15.5% ± 0.8%) and *R. hominis* FRIFI (13.9% ± 3.3%). The reassociation values were well below the cut-off point of 70% that was recommended by Wayne et al. [123] for species delineation, thereby indicating that just as R. lituseburense A25K^T, *R. hominis* FRIFI does not belong to the species *R. ilealis*.

Having the complete genome sequences, it is possible to obtain *in silico* analogues of DDH values based on computational comparisons between two genomes [62, 100]. Average nucleotide identity (ANI) has been proposed to be the next 'gold standard' for species delineation. ANI values were calculated by pairwise comparisons of the complete genomes of the four proposed *Romboutsia* species and closely-related species (see Table S2 in the supplemental material). In all cases the ANI-values for pairwise comparisons of the different *Romboutsia* strains amounted to 76.3-79.1%, which is well below the cut-off point of 95-96% that is generally accepted to be the ANI threshold for species delineation [44, 57, 100], affirming the novel species assignment of *R. hominis* FRIFI.

Based on 16S rRNA gene sequence and whole genome-based comparisons we conclude that *R. ilealis, R. lituseburensis, R. hominis* FRIFI and *C. dakarense* form a genus level monophyletic clade that justifies including *C. dakarense* and *R. hominis* FRIFI as novel species within the genus *Romboutsia*.

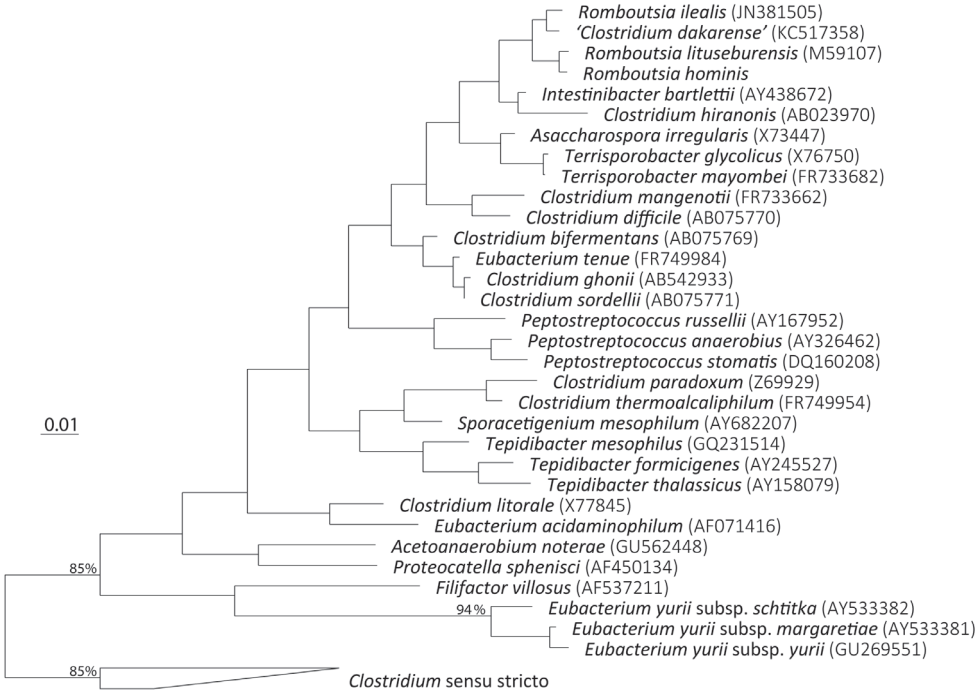

Figure 2. Neighbour-joining tree of Romboutsia species and closely-related species within the family Peptostreptococcaceae based on 16S rRNA gene sequence data. The 16S rRNA gene sequences were aligned using the SINA aligner [92]. The phylogenetic tree was reconstructed using ARB [75] using the "The all-species living tree" project dataset release LTPs115 [130]. Only bootstrap values >70% are shown at branch nodes. Bar, 1% sequence divergence. GenBank accession numbers are given in parentheses. The 16S rRNA gene sequences of several Clostridium sensu stricto species were used as an outgroup.

Comparative genomic analysis

The number of protein coding genes per genome within the various strains was quite variable, ranging from 2,360 to 3,843 (Table 1). The number of putative homologous proteins among the four *Romboutsia* genomes was determined via amino acid level best bidirectional hits (Figure 3). In total 1,421 genes were shared between all four strains, the core genome, accounting for 37% to 60% of the total gene count in the individual genomes, providing a first insight in the genomic heterogeneity within the genus. The bigger the genome, the more unique genes were present, ranging from 14% to 28% of the total gene count.

The comparative genome analysis showed a general conservation of the genomic structure of the genus *Romboutsia* around the replication start site, while synteny appeared to be lost towards the replication end point. For most pairwise comparisons synteny was lost at a quarter of the genome in both up- and downstream directions, making roughly half of the genomes syntenic. Breaks of synteny appeared to be related to specific recombination events. For example, compared to the other genomes synteny is

absent in *R. ilealis* CRIB^T due to the insertion of a prophage, whereas the regions both up- and downstream are syntenic. At another spot in the genome of *R. ilealis* CRIB^T synteny is lost due to phage-related genes found around the tmRNA gene, which has been reported to be a common insertion site for phages [125]. The position of the tmRNA itself is roughly equal in all three genomes, but no synteny could be observed in its vicinity. Strain/species specific gene clusters, like the CRISPR-Cas system or the fucose degradation pathway present in *R. ilealis* CRIB^T, appeared to be situated more towards the less conserved replication end point. One point of conservation in the less conserved area is an inversion of one part of the butyrate fermentation pathway, which is absent in *R. ilealis* CRIB^T, but inverted between *R. hominis* FRIFI and *R. lituseburensis* A25K^T. Some significant deletion events appeared to have occurred, since they can be observed in the conserved areas of the genomes. For example, the pili encoding gene cluster, which was found in the genomes of *R. lituseburensis* A25K^T and *C. dakarense* FF1^T close to the replication start site, is absent in the genomes of *R. ilealis* CRIB^T and *R. hominis* FRIFI except for a twitching motility protein encoding gene. Another example is the biosynthesis cluster for Vitamin B12, which is also located close to the replication start site. While this cluster is complete in *R. lituseburensis* A25K^T, *C. dakarense* FF1^T and *R. hominis* FRIFI, the remnants of the cluster are still visible in *R. ilealis* CRIB^T, having a deletion in the middle of nine proteins, which prevents the biosynthesis of cob(II)yrinate a,c-diamide. This cluster is also situated next to an rRNA operon, of which only the one in *R. ilealis* CRIB^T has an integrase inserted.

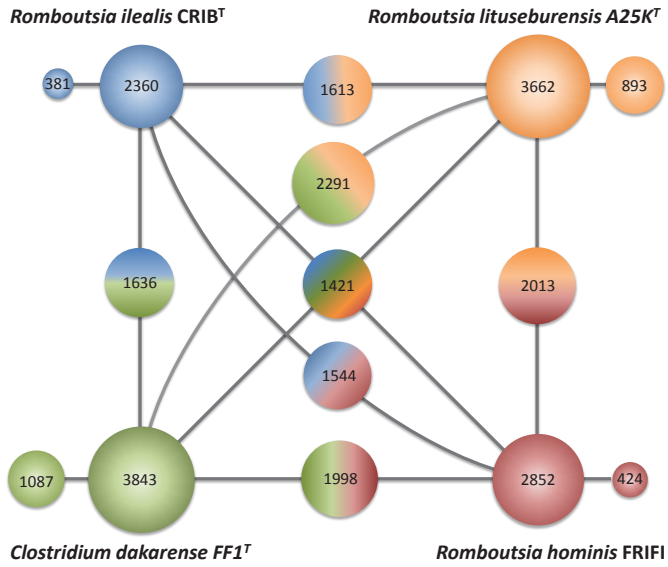

Figure 3. Overview of the number of homologous genes shared between the four *Romboutsia* genomes available to date. Also the total number of genes and the number of unique genes are indicated for each genome. The circles are colour-coded by the *Romboutsia* strains they represent: blue, *R. ilealis* CRIB^T; yellow; *R. lituseburense* A25K^T; green, *C. dakarense* FF1^T; red, *R. hominis* FRIFI. Area of the circle is representative for the number of genes.

Core and pan metabolism of genus *Romboutsia*

An overview of the core metabolism of the *Romboutsia* strains and strain-specific metabolic features is provided in Figure 4. All four *Romboutsia* strains can ferment carbohydrates via the glycolysis, and possess the non-oxidative pentose phosphate pathway. Moreover, from the genomes it was predicted that all strains have the capability to synthesize (and degrade) all nucleotides, cell wall components, fatty acids and siroheme. In addition, it was predicted that the *Romboutsia* strains can only produce a limited set of amino acids. In turn they are, however, also able to ferment numerous amino acids. Furthermore various pathways for assimilation of nitrogen were predicted, as well as a pathway for production of the quorum sensing associated autoinducer AI-2. Some of the metabolic highlights will be discussed in the following paragraphs.

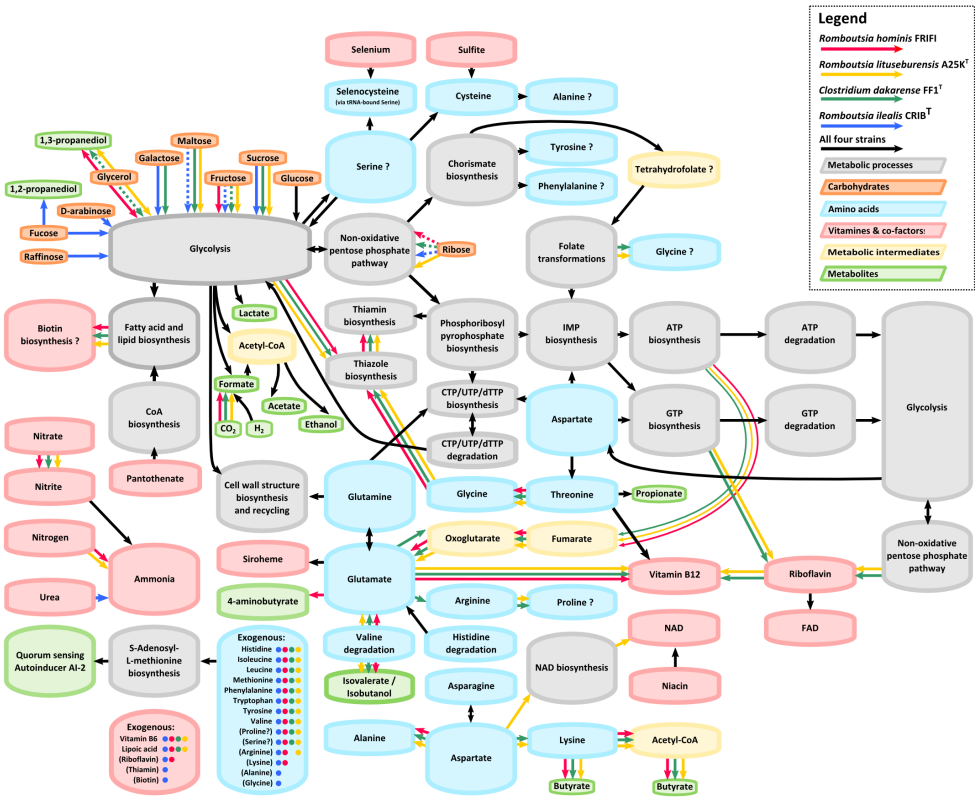

Figure 4. Overview of the core metabolism and strain-specific metabolic features predicted from the genomes of the four *Romboutsia* strains. Dotted lines indicate instances where reported experimental observations did not match genome-based predictions.

Metabolic features: fermentation and anaerobic respiration

Similar fermentation end products were observed for *R. ilealis* CRIB^T, *R. hominis* FRIFI and *R. lituseburensis A25K^T* during growth on glucose, namely formate, acetate and a small amount of lactate (**Chapter 4**, data not shown for *R. hominis* FRIFI). The fermentation end products for *C. dakarense* FF1^T have not been reported, however, based on gene presence/absence it is possible to make predictions (Figure 4). The pathways leading to formate, acetate and lactate production previously described for *R. ilealis* CRIB^T (**Chapter 5**) were also found in the other *Romboutsia* strains suggesting that all are able to produce formate, acetate and lactate. This was experimentally confirmed for *R. hominis* FRIFI and *R. lituseburensis* A25K^T.

In the genomes of *R. hominis* FRIFI, *R. lituseburensis* A25K^T and *C. dakarense* FF1^T a reductive pathway for the metabolism of glycerol was predicted, comprising glycerol dehydratase and 1,3-propanediol dehydrogenase [22]. This suggests that these strains are able to ferment glycerol and produce 1,3-propanediol as one of the fermentation end products. For both *R. hominis* FRIFI and *R. lituseburensis* A25K^T this was confirmed experimentally (data not shown). Fermentation of glycerol was, however, not reported for *C. dakarense* FF1^T [71]. For *R. lituseburensis* the ability to produce 1,3-propanediol was previously reported [23]. In the genomes of *R. hominis* FRIFI, *R. lituseburensis* A25K^T and *C. dakarense* FF1^T also the oxidative pathway for glycerol metabolism, comprising glycerol dehydrogenase and dihydroxyacetone kinase, could be identified suggesting that both strains should be able to use glycerol as sole carbon and energy source.

With respect to the potential to produce butyrate, the acetyl-CoA and lysine pathways, both leading to butyrate synthesis, could be predicted from the genomes of *R. hominis* FRIFI, *R. lituseburensis* A25K^T and *C. dakarense* FF1^T [122]. The pathways are co-located in the genome, suggestion that the acetoacetyl-CoA formed during lysine fermentation can be directly used as substrate in the acetyl-CoA pathway for additional energy conservation [122]. A lysine-specific permease was predicted in the genomes of *R. hominis* FRIFI, *R. lituseburensis* A25K^T and *C. dakarense* FF1^T, suggesting that exogenous lysine can serve as energy source for these strains. Since an acetyl-CoA acetyltransferase was also found in the gene cluster, a fully functioning carbohydrate-driven acetyl-CoA pathway is expected. For the final step in butyrate production, a phosphate butyryl transferase/ butyrate kinase (*buk*) gene cluster was identified in the genomes of the three (potential) butyrate-producing *Romboutsia* strains, with two paralogues being predicted in the genome of R. *lituseburensis* A25K^T. Butyrate (and iso-valerate) production was observed for *R. lituseburensis* A25K^T during *in vitro* growth on undefined medium components such as beef extract, peptone and casitone (but not on yeast extract). The addition of a carbohydrate (e.g. glucose) resulted in a redirection of the fermentation pathways towards other end products such as formate (data not shown).

All four *Romboutsia* strains contain genes encoding for the respective enzymes of the

Wood-Ljungdahl pathway. A formate dehydrogenase was predicted for all strains except *R. ilealis* CRIB^T. Similar to *Clostridium difficile* [64], *C. dakarense* FF1^T contains multiple copies of formate dehydrogenase encoding genes. The presence of formate dehydrogenase together with a complete Wood-Ljungdahl pathway categorizes them as potential acetogens, microbes that can grow autotrophically using CO_2 and H_2 as carbon and energy source. This provides them with metabolic flexibility in addition to heterotrophic growth on organic compounds. The role of acetogens in the intestinal tract is not well studied. They have been proposed to play an important role in hydrogen disposal, in addition to methanogens and sulfate reducers [42, 67].

The genes encoding for a complete membrane-bound electron transport system were identified in the genomes of *R. hominis* FRIFI, *R. lituseburensis* A25K^T and *C. dakarense* FF1^T, similar to the Rnf system identified in microbes such as *Clostridium tetani, Clostridium ljungdahlii* and *C. difficile*. In these species the system is suggested to be used to generate a proton gradient for energy conservation in microbes without cytochromes. In *C. tetani* the system is proposed to play a role in the electron flow from reduced ferredoxin, via NADH to the NADH-consuming dehydrogenase of the butyrate synthesis pathway [10]. In addition, the Rnf system is proposed to be used by *C. ljungdahlii* during autotrophic growth [63]. In the genome of *R. ilealis* CRIB^T remnants of an Rnf electron transport system could be found, which may be a result of genome reduction since also no butyrate synthesis pathway or acetogenic pathway were found (**Chapter 5**).

Genomes of all four *Romboutsia* strains contain genes predicted to encode a sulfite reductase of the AsrABC type. Inducible sulfite reductases are directly linked to the regeneration of NAD^+, which plays a role in energy conservation and growth, as well as to detoxification of sulfite [27]. *R. hominis* FRIFI, however, appears to lack the formate/nitrite transporter family protein that was found in the vicinity of the predicted sulfite reductase in the other strains similarly to *C. difficile* where it was characterized as a hydrosulfide ion channel that exports the toxic metabolites out of the cell [21].

Carbohydrate metabolism

The carbohydrate utilization of the different *Romboutsia* strains is shown in Table 2. *R. hominis* FRIFI is limited in the carbohydrates it is able to grow on, since only glucose, fructose and glycerol were able to support growth.

Absence/presence of specific gene clusters could explain most of the differences found in carbohydrate utilization. In some cases discrepancies between genome-based predictions and experimental observations could be explained by the absence of specific transporters. The only carbohydrate all strains were able to use was glucose. Although a ribose degradation cluster was found in all four *Romboutsia* strains, only *R. lituseburensis* A25K^T was found to be able to grow on ribose *in vitro*. It could be confirmed that the L-fucose degradation pathway was only present in *R. ilealis* CRIB^T, coinciding with its

ability to use D-arabinose. In the genome of *R. hominis* FRIFI a fructose-specific PTS system was identified together with a 1-phosphofructokinase, similar to *R. lituseburensis* A25K^T. A similar gene cluster could also be identified in *C. dakarense* FF1^T suggesting that growth on D-fructose should be possible, however, it was reported that *C. dakarense* FF1^T is not able to do so [71]. The gene cluster could not be identified in *R. ilealis* CRIB^T, confirming previous experimental observations that this organism is not able to grow on exogenous D-fructose. No transport system for sucrose was found in *R. hominis* FRIFI, which corresponds with the observed lack of growth on sucrose. The PTS system identified in *R. ilealis* CRIB^T to be involved in fructo-oligosaccharide (FOS) degradation (**Chapter 5**) was also found in *R. lituseburensis* A25K^T and *C. dakarense* FF1^T. It is interesting to note that for *R. lituseburensis* A25K^T a similar organization of the gene cluster was observed, including the presence of a ß-fructofuranosidase overlapping with an RDD family protein, the gene organization that for *R. ilealis* CRIB^T was hypothesized to encode for an external membrane-bound enzymatically active protein. The other two sucrose degradation gene clusters identified in *R. ilealis* CRIB^T, including the ABC-type transporter containing gene cluster that was upregulated during growth on FOS (**Chapter 5**), could not be identified in the other *Romboutsia* genomes. The gene clusters for degradation of galactose (encoding a galactokinase, galactose 1-phosphate uridylyltransferase and UDP-glucose 4-epimerase) and raffinose utilization were only found in *R. ilealis* CRIB^T. A maltose degradation gene cluster was identified in *R. ilealis* CRIB^T, *R. lituseburensis* A25K^T and *C. dakarense* FF1^T, in the latter together with a glucose-specific PTS system. *R. lituseburensis* A25K^T and *R. hominis* FRIFI can both metabolise glycerol, and a gene cluster for glycerol utilization could also be identified in both genomes together with a predicted glycerol transporter. However, as mentioned before, a similar gene cluster could also be identified in *C. dakarense* FF1^T suggesting that *C. dakarense* FF1^T should be able to metabolise glycerol as well.

Table 2. Carbohydrate utilization by the *Romboutsia* strains determined with the API50 system.

	R. hominis FRIFI	R. lituseburensis A25K^T	R. ilealis CRIB^T	C. dakarense FF1^T
D-arabinose	-	-	+	-
L-Fucose	-	-	+	-
Fructose	+	+	-*	-†
Glucose	+	+	+	+
Galactose	-	-	+	+
Glycerol	+	+	-	-†
Maltose	-	+	-*	+
Raffinose	-	-	+	-
Ribose	-†	+	-†	-†
Sucrose	-	+	+	+

*Only very weak growth on these substrates was observed in liquid mono-cultures.

† Reported experimental observations do not match genome-based predictions.

In addition to the experimentally examined carbohydrates, parts of the sialic acid utilization gene cluster previously predicted in the genome of *R. ilealis* CRIB^T (**Chapter 5**) were also present in *R. hominis* FRIFI and *C. dakarense* FF1^T. In addition, *C. dakarense* FF1^T contains a gene predicted to encode a sialidase that contains a signal peptide, indicating that this is an extracellular sialidase, similar to the extracellular sialidase found in the related species *C. perfringens* [17].

Several complete phosphoenolpyruvate (PEP):carbohydrate phosphotransferase system (PTS) transporters, consisting of several Enzyme II (EII) subunits, were identified in the four *Romboutsia* genomes. To this end, *R. hominis* FRIFI, *R. lituseburensis* A25K^T and *C. dakarense* FF1^T have a PTS transporter belonging to the mannose/fructose/sorbose family that is unique among PTS permease families in that they contain a fourth subunit (EIID), while most PTS systems consist of only three subunits (EIIA, EIIB and EIIC) [91]. Of related species for which genomes are available, only *Clostridium sordellii* and *Clostridium bifermentans* contain a similar transporter (determined by BLAST search). The specificity of this particular PTS transporter in the *Romboutsia* strains remains to be determined. With respect to the other PTS transporters identified, it was observed that in some cases an EIIA subunit is likely used to complete multiple PTS systems as it has been previously described [91]. For example, the glucose-specific EIIA subunit is probably also used in the sucrose/FOS-specific PTS systems of *R. ilealis* CRIB^T, *C. dakarense* FF1^T and *R. lituseburensis* A25K^T , and the predicted N-acetylglucosamine-specific PTS systems of *R. hominis* FRIFI, *C. dakarense* FF1^T and *R. lituseburensis* A25K^T.

A gene cluster for ethanolamine utilization was predicted in *C. dakarense* FF1^T and *C. lituseburensis* A25K^T. Ethanolamine (and glycerol) is produced by phosphodiesterases from phosphatidylethanolamine, which is an abundant phospholipid in both mammalian and bacterial cell membranes and is therefore abundantly present in the intestinal tract. The enzymes for ethanolamine utilization are active in a microcompartment in the cell, of which the structural proteins are also encoded in the ethanolamine gene cluster [38]. By the breakdown to acetaldehyde and ammonia, ethanolamine can be used as a carbon and/or nitrogen source for the microbes able to utilize phosphatidylethanolamine. It is interesting to note that phosphatidylethanolamine was not found in the cell membranes of *R. ilealis* CRIB^T and *R. lituseburensis* A25K^T, something that clearly distinguishes members of the *Peptostreptococcaceae* from *Clostridium* sensu stricto species (**Chapter 4** [39]). Similar gene clusters could be identified in the related species *C. difficile* [89] and by homology also in C. *bifermentans* and *C. mangenottii.*

Fermentation of individual amino acids

Species belonging to the class *Clostridia* are known for their capabilities to ferment amino acids. Of the four *Romboutsia* strains, *R. lituseburensis* A25K^T and *C. dakarense* FF1^T appear to be the most resourceful. All four *Romboutsia* strains are predicted to be

able to ferment L-histidine via glutamate using a histidine ammonia lyase. In addition, fermentation of L-threonine was predicted using a L-threonine dehydratase resulting in propionate production, which has been described for *R. lituseburensis* [94]. Fermentation of L-serine into pyruvate using an L-serine dehydratase was predicted for all four strains as well. As mentioned before, *R. hominis* FRIFI, *R. lituseburensis* A25K^T and *C. dakarense* FF1^T are predicted to be able to ferment L-lysine. In addition, *R. lituseburensis* A25K^T and *C. dakarense* FF1^T are predicted to ferment glycine using the glycine reductase pathway found in other related species including *C. difficile* [2, 54]. A corresponding complex has also been identified in *R. hominis* FRIFI, but it is likely to be non-functional, due to a loss of two of the three subunits. Furthermore, the ability to ferment L-arginine (using an arginine deiminase) and L-glutamate (using a Na^+-dependent glutaconyl-CoA decarboxylase) was predicted for *C. lituseburensis* A25K^T and *C. dakarense* FF1T as well. A glutamate decarboxylase was predicted for *R. hominis* FRIFI, suggesting the ability to decarboxylate glutamate to 4-aminobutyrate (GABA) for this strain only.

Table 3. Overview of genome-based predictions for amino acid requirements of the four *Romboutsia* strains. In case only one or two enzymes are missing in either salvage or *de novo* pathway leading to the production of an amino acid, this is indicated in parentheses.

	R. hominis FRIFI	R. lituseburensis A25K^T	R. ilealis CRIB^T	C. dakarense FF1^T
Alanine	+	+	-	+
Arginine	-	-	-	+
Asparagine	+	+	+	+
Aspartic acid	+	+	+	+
Cysteine	+	+	+	+
Glutamic acid	+	+	+	+
Glutamine	+	+	+	+
Glycine	+	+	- (-1)	+
Histidine	-	-	-	-
Isoleucine	-	-	-	-
Leucine	-	-	-	-
Lysine	- (-1)	+	- (-1)	+
Methionine	-	-	-	-
Phenylalanine	- (-1)	- (-1)	- (-1)	-
Proline	-	- (-1)	-	- (-1)
Serine	- (-2)	- (-1)	- (-1)	- (-1)
Threonine	- (-1)	- (-1)	- (-2)	- (-2)
Tryptophan	-	-	-	-
Tyrosine	- (-1)	- (-1)	- (-1)	-
Valine	-	-	-	-

Amino acid and vitamin requirements

Pathways for (*de novo*) synthesis of amino acids were identified in the four *Romboutsia* strains (Table 3). All four strains show similar dependencies on exogenous amino acid sources. Based on genome predictions, *R. lituseburensis* A25K^T and *C. dakarense* FF1^T are able to synthesize lysine from aspartate, whereas the last enzyme in this pathway is missing in the genomes of *R. hominis* FRIFI and *R. ilealis* CRIB^T. In addition, *R. hominis* FRIFI, *R. lituseburensis* A25K^T and *C. dakarense* FF1^T are predicted to synthesise alanine from aspartate and glycine from threonine. Common to all organisms is that the prephenate dehydratase for the biosynthesis of phenylalanine and tyrosine is missing, although all other enzymes for the biosynthesis of chorismate and for the further conversion to both amino acids are present.

The urease gene cluster, previously identified in *R. ilealis* CRIB^T (**Chapter 5**), could not be identified in the three other *Romboutsia* strains. A nitrogenase encoding gene cluster was identified in the genomes of *R. hominis* FRIFI and *R. lituseburensis* A25K^T, suggesting that these two strains are able to fix N_2.

All four strains encode one or several oligopeptide transporters of the OPT family [74] In addition, two oligopeptide transport systems (*Opp* and *App*) [31, 61] were predicted in *R. hominis* FRIFI, *R. lituseburensis A25K^T* and *C. dakarense* FF1^T (*R. ilealis* FRIFI is missing the OppA), whereas they were absent in *R. ilealis* CRIB^T. Based on the predicted amino acid dependencies, it can be concluded that these *Romboutsia* strains are adapted to an environment rich in amino acids and peptides.

The metabolic capabilities of the four *Romboutsia* species are comparable regarding the ability to produce certain vitamins and other cofactors (shown in Figure 4). None of them are predicted to be able to synthesize vitamin B6, lipoic acid or pantothenate, however, they are all able to produce siroheme from glutamate and CoA from pantothenate. As previously described for *R. ilealis* CRIB^T (**Chapter 5**), the pathway for *de novo* folate biosynthesis is present via the pABA branch, however, a gene encoding dihydrofolate reductase, could not be identified in any of the four *Romboutsia* strains. However, since this enzyme is essential in both *de novo* and salvage pathways of tetrahydrofolate (THF), it is highly likely that a functional equivalent is encoded in the genomes. The biosynthetic capabilities of *R. lituseburensis* A25K^T, *C. dakarense* FF1^T and *R. hominis* FRIFI are larger than that of *R. ilealis* CRIB^T, as they are all three predicted to produce biotin, thiamin and vitamin B12. The gene clusters for biotin and thiamine biosynthesis are located in the more variable regions of the genomes as discussed above, and the vitamin B12 biosynthesis pathway is incomplete in *R. ilealis* CRIB^T due to a deletion, as mentioned earlier. Only *R. lituseburensis* A25K^T and *C. dakarense* FF1^T are predicted to have the capacity to produce riboflavin *de novo*. Furthermore, *R. lituseburensis* A25K^T is, as the only non-host derived organism in this comparison, the only strain that can synthesize NAD *de novo*.

Bile resistance

One of the challenges for microbes living in the intestinal tract is to deal with the host secreted bile acids. The bile acid pool size and composition has been shown to modulate the size and composition of the intestinal microbiota and vice versa [25, 102]. Bile acids can undergo a variety of bacterial transformations including deconjugation, dehydroxylation and epimerization. In both *R. ilealis* CRIB^T and *R. hominis* FRIFI a choloylglycine hydrolase encoding gene was identified. Bile salt hydrolases (BSHs), also known as conjugated bile acid hydrolases (CBAHs), from the choloylglycine hydrolase family are widespread among Gram-positive and Gram-negative intestinal microbes [101]. They are involved in the hydrolysis of the amide linkage in conjugated bile acids, releasing primary bile acids. There is a large heterogeneity among BSHs, including with respect to their substrate specificity (e.g. specificity towards either taurine or glycine conjugated bile salts) [25]. The choloylglycine hydrolase of *R. ilealis* CRIB^T and *R. hominis* FRIFI differ significantly from each other (32% identity at amino acid level), suggesting a different origin. The BSH of *R. ilealis* CRIB^T and *R. hominis* FRIFI show at the amino acid level 52% and 33% identity, respectively, with the choloylglycine hydrolase CBAH-1 from *Clostridium perfringens* [20].

In addition to the possible BSH, a bile acid 7α-dehydratase encoding gene could be identified in *R. hominis* FRIFI. This enzyme is part of the multi-step 7α/ß-dehydroxylation pathway that is involved in the transformation of primary bile acids into secondary bile acids. So far, this pathway has been found exclusively in a small number of anaerobic intestinal bacteria all belonging to the *Firmicutes* [101]. The presence of this pathway allows microbes to use primary bile acids as an electron acceptor, resulting in for increased ATP formation and growth. High levels of secondary bile acids, however, are associated with diseases of the host such as cholesterol gallstone disease and cancers of the GI tract [6, 79], although it should be noted that the evidence that bacteria capable of 7α-dehydroxylation are directly involved in the pathogenesis of these diseases is still limited. The pathway has been extensively studied in the human isolate *Clostridium scindens* VPI 12708 (formerly known as *Eubacterium* sp. strain VPI 12708 [58]). In addition, 7α-dehydroxylation activity was also reported for *Clostridium hiranonis* [124] and *C. sordellii* [47], close relatives of the *Romboutsia* species. Extensive characterization of the 7α/ß-dehydroxylation pathway in *C. scindens* VPI 12708 has demonstrated that the genes involved are encoded by a large bile acid inducible (*bai*) operon [101]. For *R. hominis* FRIFI several other genes were identified in the vicinity of the bile acid 7α-dehydratase gene that showed homology to the genes in the *bai* operon, however, some other (essential) genes seem to be missing. From gene presence/absence it was therefore not possible to predict whether *R. hominis* FRIFI has 7α/ß-dehydroxylation activity.

Toxins and virulence-related genes

The class *Clostridia* contains some well-known pathogens, including *C. difficile* and

C. sordellii, for which several toxins have been characterized in depth [90]. No homologues of the genes coding for the toxins of *C. difficile* (toxin A, toxin B, binary toxin) or *Clostridium botulinum* could be found in the genomes of the four *Romboutsia* strains. The genomes of *R. ilealis* CRIB^T and *C. dakarense* FF1^T both encode a predicted protein that was annotated as a putative septicolysin (CRIB_2392 and DAK_3451) since it shares 56% and 49% identity to a protein that has been characterized as an oxygen-labile hemolysin in *Clostridium septicum* [82]. The exact role of septicolysin in potential pathogenesis is, however, not known. Homologues are not found in other related species. A homologue for the alpha toxin of *Clostridium perfringens* [90, 113], a phospholipase C protein involved in the aetiology of gas gangrene caused by *C. perfringens* [112], was found by BLAST search in the genomes of *R. hominis* FRIFI, *R. lituseburensis* A25K^T and *C. dakarense* FF1^T (49.4-54.3% identity at the amino acid level). In addition, *R. lituseburensis* A25K^T and *C. dakarense* FF1^T were both predicted to contain a protein very homologous to the perfringolysin O (theta toxin) of *C. perfringens*, which is a thiol-activated cytolysin that forms large homo-oligomeric pore complexes in cholesterol-containing membranes, which is also involved in gas gangrene aetiology. By BLAST search similar proteins could also be found in the genomes of *C. sordellii* and *C. bifermentans*, which are close relatives of the *Romboutsia* strains. There are, however, many homologous enzymes produced by other bacteria that do not have similar toxigenic properties as the *C. perfringens* proteins [112]. For example, the phospholipase C proteins produced by *C. bifermentans* and *C. sordellii*, were found to have significantly less haemolytic activity than the homologues protein of *C. perfringens* (51% and 53.4% indentity on amino acid level, respectively) [55, 115]. The predictions regarding the presence of potential toxin-encoding genes in the *Romboutsia* strains was done based on homology; the enzymatic activity of the gene products will have to be determined in the future to see whether some of the *Romboutsia* strains have toxigenic properties.

6

Motility

Motility was observed for *R. hominis* FRIFI, *R. lituseburensis* A25K^T and *C. dakarense* FF1^T, but not for *R. ilealis* CRIB^T, as previously reported. In general, different appendages can be found on bacterial surfaces that provide bacteria with the ability to swim in liquids or move on surfaces via gliding or twitching motility [4, 120]. In the genomes of *R. hominis* FRIFI and *R. lituseburensis* A25K^T a large gene cluster for the synthesis of flagella could be identified. In addition, homologues could be identified in the genome of *C. dakarense* FF1^T as well, however, due to the quality of the current assembly the genes were spread across several scaffolds. The organization of the flagella gene cluster is very similar to that in the genome of *C. difficile*. The formation of flagella involves a whole array of different components, including the core protein flagellin. Post-translational modification of flagellin by glycosylation is an important process both for the flagellar assembly and biological function, and genes involved in these modifications are often found in the

vicinity of the structural flagellin genes [72]. This is also the case for *R. hominis* FRIFI and *R. lituseburensis* A25K^T, and these genes are found in an intra-flagellar synthesis locus similar to the situation in *C. difficile* 630 [106]. In *R. ilealis* CRIB^T no genes encoding flagellar proteins or genes involved in chemotaxis could be identified in line with the lack of motility (**Chapter 4** [39]). Flagella are dominant innate immune activators in the intestinal tract as flagellin molecules can be recognized by host cell-surface and cytoplasmatic pattern recognition receptors [15, 41, 48]. The role for flagella in virulence and pathogenicity of *C. difficile* is a topic of interest, however, their exact contribution is still unknown [3]. The flagellin proteins of some of the most abundant motile commensal microbes that are found in the human intestinal tract, *Eubacterium* and *Roseburia* species, have recently also shown to possess pro-inflammatory properties [83].

One type of gliding motility involves the extension, attachment and retraction of type IV pili (TFP), which pull the bacterium towards the site of attachment [53]. In the genomes of *C. dakarense* FF1^T and *R. lituseburensis* A25K^T a complete set of genes for the assembly of Type IV pili could be identified. In contrast, in the genomes of *R. ilealis* CRIB^T and *R. hominis* FRIFI only remnants could be identified.

Cell surface proteins

A fibronectin-binding protein encoding gene (*fbpA*) has been identified in the genome of all four *Romboutsia* strains. Similar genes have been identified in the genomes of closely related species, including *C. difficile*, *C. glycolicum* and *C. bartlettii*. A mutant of *fbpA* was constructed in *C. difficile* strain 630 and characterized *in vitro* and *in vivo*, indicating that fibronectin-binding protein plays a minor role in intestinal colonization [5].

In *R. ilealis* CRIB^T a gene cluster for capsule biosynthesis (CapABCD) was predicted (CRIB_210-215). Homologues are found in *I. bartlettii* and *T. glycolicus*, but are absent in the other three *Romboutsia* strains. The gene cluster is found in a genomic area that is situated around tmRNA and scattered phage elements mentioned above, showing overall no synteny to the other *Romboutsia* genomes. These genes are involved in the formation of a capsule composed of poly-gamma-glutamate (PGA) found in a limited number of Gram-positive bacteria [12]. The impact of the transposase, which is inserted between the *capA* and *capD* gene, on capsule formation remains to be determined.

Sporulation

Sporulation is a trait found only in certain low G+C Gram-positive bacteria belonging to the *Firmicutes* [84]. The formation of metabolically dormant endospores is an important strategy used by bacteria to survive environmental challenges such as nutrient limitation. These endospores are resistant to extreme exposures (e.g. high temperatures, freezing, radiation and agents such as antibiotics and most detergents) that would kill vegetative cells. The ability to form endospores was also studied for the four *Romboutsia* strains. *R.*

lituseburensis A25K^T readily forms mature spores, especially during growth in Duncan-Strong medium and Cooked meat medium, that both contain large quantities of proteose peptone, and spore formation was observed in almost every cell (data not shown). Also for *R. hominis* FRIFI^T sporulation was observed (data not shown) and for *C. dakarense* FF1^T sporulation has been reported as well [71]. Previously, the endospore forming capabilities of *R. ilealis* CRIB^T have been studied (**Chapter 4** [39]). Using different media and incubation conditions it was observed that the process of sporulation appears to be initiated, however, no free mature spores could be observed.

The whole process of sporulation and subsequent spore germination involves the expression of hundreds of genes in a highly regulated manner. At a molecular level the process is best understood in the model organism *Bacillus subtilis* [110]. For species belonging to the class *Clostridia*, the process of sporulation is mainly studied in microbes in which sporulation has been shown to play a big role in other processes such as virulence (*C. perfringens*, *C. difficile*, *C. botulinum* and *C. tetani*) or solvent production (*Clostridium acetobutylicum*). Studying these microbes has made it clear that there are significant differences in the sporulation and germination process in species belonging to the class *Clostridia* compared to *Bacilli* species [36, 87, 127]. The *B. subtilis* proteins involved in the early stages of sporulation [onset (stage I), commitment and asymmetric cell division (stage II), and engulfment (stage III)] are largely conserved in *Clostridia* species, however, many of the proteins that play a role in later stages (cortex formation (stage IV), spore coat maturation (stage V), mother cell lysis and spore release) appear to be less conserved. For example, l*imited spore outer layer conservation was observed in C. difficile compared to B. subtilis [86]*. Comparative genomic based-studies have tried to define the minimal set of genes essential for sporulation in clostridial species, however, this has appeared to be challenging [36, 87, 126]. In all spore-formers, initiation of sporulation is controlled by the transcription factor Spo0A, a highly conserved master regulator of sporulation. Phosphorylation of Spo0A leads to the activation of a tightly regulated cascade involving several sigma factors that regulate the further expression of a multitude of genes involved in sporulation. There are, however, significant differences in the regulation of the sporulation pathway between different *Clostridium* species [33] of which we do not yet completely understand the impact on sporulation itself, highlighting that there is still a big gap in our knowledge on the complex process of sporulation.

The genomes of the four *Romboutsia* genomes were mined for homologues of sporulation specific genes according to Galperin *et al*. [36]. All four *Romboutsia* strains have similar sets of sporulation-related genes, with *R. ilealis* CRIB^T having the least number of genes (147 genes) and *R. lituseburensis* A25K^T having the most (183 genes) (see Table S3 in the supplemental material). The only protein that is deemed essential for sporulation, but which was only found in the genome of *R. lituseburensis* A25K^T, is the Stage V sporulation protein S that has been implicated to increase sporulation [97]. For *R. lituseburensis* A25K^T,

the sporulation regulator Spo0E was predicted to be absent, due to a point mutation in the start codon of the corresponding gene, changing it to an alternative start codon. This regulator is suggested to be involved in the prevention of sporulation under certain circumstances [96]; impact of the point mutation on transcription of the gene and subsquent regulation of sporulation in *R. lituseburensis* A25K^T will have to be determined. Interestingly, the stage V sporulation proteins AA and AB, encoded by *spoVAA* and s*poVAB*, that are essential for sporulation in *Bacilli* since mutants lead to the production of immature spores [114], are absent in sporulating *Clostridium* species, but are present in all four *Romboutsia* strains. Furthermore, of the four strains, *R. lituseburensis* A25K^T is the only one that contains the *sps* operon that has been shown to be involved in spore surface adhesion [13]. Absence of this operon in *B. subtilis* resulted in defective germination, and more hydrophobic and adhesive spores, however, given that these proteins are also absent in nearly all clostridial species, their role in sporulation and germination in the *Romboutsia* strains still has to be determined. As also noted by Galperin *et al.* [36], there are other species that have been demonstrated to be spore-forming but which also lack some of the genes that are deemed to be essential, e.g. *spoIIB*, *spoIIM*, and other proteins from the second sporulation stage in *Lysinibacillus sphaericus* C3-41v [36]. In comparison, it is interesting to note that the genome of *C. hiranonis*, a close relative of the genus *Romboutsia* (and *C. difficile)*, appears to contain only 21 of the essential sporulation genes, missing for example most of the proteins related to the second and third stage of sporulation, while *C. hiranonis* is known to form spores ([59], and own observations). Altogether, based on gene presence/absence it is not possible to predict whether these *Romboutsia* strains are indeed able to successfully complete the process of sporulation and release endospores. An asporogenous phenotype could be credited to the absence or mutation of a single gene.

Initiation of sporulation is still a topic of interest. Accessory gene regulatory (*agr)*-dependent quorum sensing, and thus most likely cell density, has been proposed to play an important role in efficient sporulation [107]. For *C. difficile*, however, quorum-sensing has been shown not to play a role in initiation of sporulation and recently a more direct link between nutrient availability and sporulation was suggested [30]. The uptake of peptides by the Opp and App oligopeptide transport systems appears to prevent initiation of sporulation in nutrient rich environments [31]. Both transport systems are absent in *R. ilealis* CRIB^T, but are present in the three other *Romboutsia* strains.

During sporulation, a number of species produce inclusion bodies and granules that are visible by phase contrast and electron microscope. This is also true for *R. lituseburensis* A25K^T in which electron translucent bodies are visible in TEM pictures (data not shown), similar to the carbohydrate or polyhydroxybutyrate inclusions observed in for example *Clostridium pasteurianum* [76], *C. acetobutylicum* [99] and *C. botulinum* [32]. *The development of these inclusion bodies appears* to coincide with the initiation of sporulation. Based on

this observation, it can be speculated that by intracellular accumulation of a carbon and energy source these microbes ensure they can complete the sporulation process with only limited dependence on external carbon and energy sources.

Concluding remarks

Based on the comparative genome analysis presented here we can conclude that the genus *Romboutsia* contains a versatile array of metabolic capabilities with respect to carbohydrate utilization, fermentation of single amino acids, anaerobic respiration and metabolic end products. The presence of bile converting enzymes and pathways related to host-derived carbohydrates, point towards adaption to a life in the (small) intestine of mammalian hosts. For each *Romboutsia* strain unique properties were found. However, since currently only one genome was available for each species, it is impossible to unequivocally predict which properties may apply to the whole species and which properties are strain-specific. Isolation and genome sequencing of additional strains from diverse environments is needed to provide a more in-depth view of the metabolic capabilities at the species- as well as the genus level and to reveal specific properties that relate to adaptation to an intestinal lifestyle.

Lastly, we propose that strain FF1^T is assigned to the genus *Romboutsia* and propose the name *Romboutsia dakarensis*. However, since the proposal of *Clostridium dakarense* as a novel species is not yet validly published, a formal proposition for name change is not possible. In addition, we conclude that on the basis of phenotypic, phylogenetic and genomic analyses, strain FRIFI constitutes a novel species within the genus *Romboutsia* for which we propose the name *Romboutsia hominis* sp. nov.

Acknowledgements

We appreciate the help of Ineke Spee (UMC Utrecht) and Heleen Vromen, Ellen Geerards and Kim van der Maas-Tromp (all from UMC St. Radboud) for assistance with collecting ileostoma effluent samples. Furthermore, we are grateful for the participants who provided ileostoma effluent samples. The authors would like to thank Wilma Akkermans-van Vliet for help with DNA extractions and Tom van den Bogert, Javier Garcia Ramiro and Lennart Kleinjans for help and advice on the pyrosequencing analyses. Harri Kangas, Juhana Kammonen and Fitsum Tamene are acknowledged for PacBio sequencing and assembly. We would also like to thank Jasper Koehorst for help with the annotation, Bart Nijsse for assistance with diverse software solutions, and Jesse van Dam for help with the Pathway Tools lisp interface. In addition, we would like to thank William Trimble from the Argonne National Laboratory for helpful discussions concerning genome assembly with MiSeq data sets. Research was partially funded by a grant of SenterNovem (FND-07013) and ERC advanced grant Microbes Inside (250172).

References

1. Abnous K., S.P. Brooks, J. Kwan, F. Matias, J. Green-Johnson, L.B. Selinger, M. Thomas, *et al.* (2009). Diets enriched in oat bran or wheat bran temporally and differentially alter the composition of the fecal community of rats. *J Nutr*; 139 (11): 2024-2031.
2. Andreesen J.R. (2004). Glycine reductase mechanism. *Curr Opin Chem Biol*; 8 (5): 454-461.
3. Baban S.T., S.A. Kuehne, A. Barketi-Klai, S.T. Cartman, M.L. Kelly, K.R. Hardie, I. Kansau, *et al.* (2013). The role of flagella in *Clostridium difficile* pathogenesis: comparison between a non-epidemic and an epidemic strain. *PLoS One*; 8 (9): e73026.
4. Bardy S.L., S.Y. Ng and K.F. Jarrell (2003). Prokaryotic motility structures. *Microbiology*; 149 (Pt 2): 295-304.
5. Barketi-Klai A., S. Hoys, S. Lambert-Bordes, A. Collignon and I. Kansau (2011). Role of fibronectin-binding protein A in *Clostridium difficile* intestinal colonization. *J Med Microbiolog*; 60 (Pt 8): 1155-1161.
6. Berr F., G.A. Kullak-Ublick, G. Paumgartner, W. Munzing and P.B. Hylemon (1996). 7 alpha-dehydroxylating bacteria enhance deoxycholic acid input and cholesterol saturation of bile in patients with gallstones. *Gastroenterology*; 111 (6): 1611-1620.
7. Bland C., T.L. Ramsey, F. Sabree, M. Lowe, K. Brown, N.C. Kyrpides and P. Hugenholtz (2007). CRISPR recognition tool (CRT): a tool for automatic detection of clustered regularly interspaced palindromic repeats. *BMC bioinformatics*; 8: 209.
8. Boetzer M. and W. Pirovano (2014). SSPACE-LongRead: scaffolding bacterial draft genomes using long read sequence information. *BMC bioinformatics*; 15: 211.
9. Bragg L., G. Stone, M. Imelfort, P. Hugenholtz and G.W. Tyson (2012). Fast, accurate error-correction of amplicon pyrosequences using Acacia. *Nat Methods*; 9 (5): 425-426.
10. Bruggemann H., S. Baumer, W.F. Fricke, A. Wiezer, H. Liesegang, I. Decker, C. Herzberg, *et al.* (2003). The genome sequence of *Clostridium tetani*, the causative agent of tetanus disease. *Proc Natl Acad Sci USA*; 100 (3): 1316-1321.
11. Burge S.W., J. Daub, R. Eberhardt, J. Tate, L. Barquist, E.P. Nawrocki, S.R. Eddy, *et al.* (2013). Rfam 11.0: 10 years of RNA families. *Nucleic Acids Res*; 41 (Database issue): D226-232.
12. Candela T. and A. Fouet (2006). Poly-gamma-glutamate in bacteria. *Mol Microbiol*; 60 (5): 1091-1098.
13. Cangiano G., T. Sirec, C. Panarella, R. Isticato, L. Baccigalupi, M. De Felice and E. Ricca (2014). The sps gene products affect germination, hydrophobicity and protein adsorption of *Bacillus subtilis* spores. *Appl Environ Microbiol*.
14. Caporaso J.G., J. Kuczynski, J. Stombaugh, K. Bittinger, F.D. Bushman, E.K. Costello, N. Fierer, *et al.* (2010). QIIME allows analysis of high-throughput community sequencing data. *Nat Methods*; 7 (5): 335-336.
15. Carvalho F.A., I. Nalbantoglu, J.D. Aitken, R. Uchiyama, Y. Su, G.H. Doho, M. Vijay-Kumar, *et al.* (2012). Cytosolic flagellin receptor NLRC4 protects mice against mucosal and systemic challenges. *Mucosal Immunol*; 5 (3): 288-298.
16. Cashion P., M.A. Holder-Franklin, J. McCully and M. Franklin (1977). A rapid method for the base ratio determination of bacterial DNA. *Anal Biochem*; 81 (2): 461-466.
17. Chiarezza M., D. Lyras, S.J. Pidot, M. Flores-Diaz, M.M. Awad, C.L. Kennedy, L.M. Cordner, *et al.* (2009). The NanI and NanJ sialidases of *Clostridium perfringens* are not essential for virulence. *Infect Immun*; 77 (10): 4421-4428.
18. Chin C.S., D.H. Alexander, P. Marks, A.A. Klammer, J. Drake, C. Heiner, A. Clum, *et al.* (2013). Nonhybrid, finished microbial genome assemblies from long-read SMRT sequencing data. *Nature methods*; 10 (6): 563-569.
19. Claudel-Renard C., C. Chevalet, T. Faraut and D. Kahn (2003). Enzyme-specific profiles for genome annotation: PRIAM. *Nucleic Acids Res*; 31 (22): 6633-6639.
20. Coleman J.P. and L.L. Hudson (1995). Cloning and characterization of a conjugated bile acid hydrolase gene from *Clostridium perfringens*. *Appl Environ Microbiol*; 61 (7): 2514-2520.
21. Czyzewski B.K. and D.N. Wang (2012). Identification and characterization of a bacterial hydrosulphide ion channel. *Nature*; 483 (7390): 494-497.
22. da Silva G.P., M. Mack and J. Contiero (2009). Glycerol: a promising and abundant carbon source for industrial microbiology. *Biotechnol Adv*; 27 (1): 30-39.
23. Dabrowski S., E. Zablotna, D. Pietrewicz-Kubicz and A. Dlugolecka (2012). Screening of environmental samples for bacteria producing 1,3-propanediol from glycerol. *Acta Biochim Pol*; 59 (3): 353-356.
24. Daims H., A. Bruhl, R. Amann, K.H. Schleifer and M. Wagner (1999). The domain-specific probe EUB338 is insufficient for the detection of all Bacteria: development and evaluation of a more comprehensive probe set. *Syst Appl Microbiol*; 22 (3): 434-444.
25. Dawson P.A. and S.J. Karpen (2014). Intestinal transport and metabolism of bile acids. *J Lipid Res*.
26. De Ley J., H. Cattoir and A. Reynaerts (1970). The quantitative measurement of DNA hybridization from renaturation rates. *Eur J Biochem*; 12 (1): 133-142.

27. Dhillon A., S. Goswami, M. Riley, A. Teske and M. Sogin (2005). Domain evolution and functional diversification of sulfite reductases. *Astrobiology*; 5 (1): 18-29.
28. Duncan C.L. and D.H. Strong (1968). Improved medium for sporulation of *Clostridium perfringens*. *Appl Microbiol*; 16 (1): 82-89.
29. Edgar R.C. (2010). Search and clustering orders of magnitude faster than BLAST. *Bioinformatics*; 26 (19): 2460-2461.
30. Edwards A.N. and S.M. McBride (2014). Initiation of sporulation in *Clostridium difficile*: a twist on the classic model. *FEMS Microbiol Lett*; 358 (2): 110-118.
31. Edwards A.N., K.L. Nawrocki and S.M. McBride (2014). Conserved Oligopeptide Permeases Modulate Sporulation Initiation in *Clostridium difficile*. *Infect Immun*; 82 (10): 4276-4291.
32. Emeruwa A.C. and R.Z. Hawirko (1973). Poly-beta-hydroxybutyrate metabolism during growth and sporulation of *Clostridium botulinum*. *J Bacteriol*; 116 (2): 989-993.
33. Fimlaid K.A., J.P. Bond, K.C. Schutz, E.E. Putnam, J.M. Leung, T.D. Lawley and A. Shen (2013). Global analysis of the sporulation pathway of *Clostridium difficile*. *PLoS Genet*; 9 (8): e1003660.
34. Fuentes S., M. Egert, M. Jimenez-Valera, M. Monteoliva-Sanchez, A. Ruiz-Bravo and H. Smidt (2008). A strain of Lactobacillus plantarum affects segmented filamentous bacteria in the intestine of immunosuppressed mice. *FEMS Microbiol Ecol*; 63 (1): 65-72.
35. Galardini M., E.G. Biondi, M. Bazzicalupo and A. Mengoni (2011). CONTIGuator: a bacterial genomes finishing tool for structural insights on draft genomes. *Source Code Biol Med*; 6: 11.
36. Galperin M.Y., S.L. Mekhedov, P. Puigbo, S. Smirnov, Y.I. Wolf and D.J. Rigden (2012). Genomic determinants of sporulation in Bacilli and Clostridia: towards the minimal set of sporulation-specific genes. *Environ Microbiol*; 14 (11): 2870-2890.
37. Gao Z.M., X. Xu and L.W. Ruan (2014). Enrichment and characterization of an anaerobic cellulolytic microbial consortium SQD-1.1 from mangrove soil. *Appl Microbiol Biotechnol*; 98 (1): 465-474.
38. Garsin D.A. (2010). Ethanolamine utilization in bacterial pathogens: roles and regulation. *Nat Rev Microbiol*; 8 (4): 290-295.
39. Gerritsen J., S. Fuentes, W. Grievink, L. van Niftrik, B.J. Tindall, H.M. Timmerman, G.T. Rijkers, *et al.* (2014). Characterization of *Romboutsia ilealis* gen. nov., sp. nov., isolated from the gastro-intestinal tract of a rat, and proposal for the reclassification of five closely related members of the genus *Clostridium* into the genera *Romboutsia* gen. nov., *Intestinibacter* gen. nov., *Terrisporobacter* gen. nov. and *Asaccharospora* gen. nov. *Int J Syst Evol Microbiol*; 64 (Pt 5): 1600-1616.
40. Gerritsen J., H.M. Timmerman, S. Fuentes, L.P. van Minnen, H. Panneman, S.R. Konstantinov, F.M. Rombouts, *et al.* (2011). Correlation between protection against sepsis by probiotic therapy and stimulation of a novel bacterial phylotype. *Appl Environ Microbiol*; 77 (21): 7749-7756.
41. Gewirtz A.T., T.A. Navas, S. Lyons, P.J. Godowski and J.L. Madara (2001). Cutting edge: bacterial flagellin activates basolaterally expressed TLR5 to induce epithelial proinflammatory gene expression. *J Immunol*; 167 (4): 1882-1885.
42. Gibson G.R., J.H. Cummings, G.T. Macfarlane, C. Allison, I. Segal, H.H. Vorster and A.R. Walker (1990). Alternative pathways for hydrogen disposal during fermentation in the human colon. *Gut*; 31 (6): 679-683.
43. Glad T., P. Bernhardsen, K.M. Nielsen, L. Brusetti, M. Andersen, J. Aars and M.A. Sundset (2010). Bacterial diversity in faeces from polar bear (*Ursus maritimus*) in Arctic Svalbard. *BMC Microbiol*; 10: 10.
44. Goris J., K.T. Konstantinidis, J.A. Klappenbach, T. Coenye, P. Vandamme and J.M. Tiedje (2007). DNA-DNA hybridization values and their relationship to whole-genome sequence similarities. *Int J Syst Evol Microbiol*; 57 (Pt 1): 81-91.
45. Guss A.M., G. Roeselers, I.L. Newton, C.R. Young, V. Klepac-Ceraj, S. Lory and C.M. Cavanaugh (2011). Phylogenetic and metabolic diversity of bacteria associated with cystic fibrosis. *ISME J*; 5 (1): 20-29.
46. Hamady M., J.J. Walker, J.K. Harris, N.J. Gold and R. Knight (2008). Error-correcting barcoded primers for pyrosequencing hundreds of samples in multiplex. *Nat Methods*; 5 (3): 235-237.
47. Hayakawa S. and T. Hattori (1970). 7alpha-dehydroxylation of cholic acid by *Clostridium bifermentans* strain ATCC 9714 and *Clostridium sordellii* strain NCIB 6929. *FEBS Lett*; 6 (2): 131-133.
48. Hayashi F., K.D. Smith, A. Ozinsky, T.R. Hawn, E.C. Yi, D.R. Goodlett, J.K. Eng, *et al.* (2001). The innate immune response to bacterial flagellin is mediated by Toll-like receptor 5. *Nature*; 410 (6832): 1099-1103.
49. Huang X. and A. Madan (1999). CAP3: A DNA sequence assembly program. *Genome Res*; 9 (9): 868-877.
50. Hunter S., P. Jones, A. Mitchell, R. Apweiler, T.K. Attwood, A. Bateman, T. Bernard, *et al.* (2012). InterPro in 2011: new developments in the family and domain prediction database. *Nucleic Acids Res*; 40 (Database issue): D306-312.
51. Huss V.A.R., H. Festl and K. Schleifer (1983). Studies on the spectrophotometric determination of DNA hybridization from renaturation rates. *Syst Appl Microbiol*; 4: 184-192.

52. Hyatt D., G.L. Chen, P.F. Locascio, M.L. Land, F.W. Larimer and L.J. Hauser (2010). Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics*; 11: 119.
53. Imam S., Z. Chen, D.S. Roos and M. Pohlschroder (2011). Identification of surprisingly diverse type IV pili, across a broad range of gram-positive bacteria. *PLoS One*; 6 (12): e28919.
54. Jackson S., M. Calos, A. Myers and W.T. Self (2006). Analysis of proline reduction in the nosocomial pathogen *Clostridium difficile*. *J Bacteriol*; 188 (24): 8487-8495.
55. Karasawa T., X. Wang, T. Maegawa, Y. Michiwa, H. Kita, K. Miwa and S. Nakamura (2003). Clostridium sordellii phospholipase C: gene cloning and comparison of enzymatic and biological activities with those of Clostridium perfringens and Clostridium bifermentans phospholipase C. *Infect Immun*; 71 (2): 641-646.
56. Karp P.D., S. Paley and P. Romero (2002). The Pathway tools software. *Bioinformatics*; 18.
57. Kim M., H.S. Oh, S.C. Park and J. Chun (2014). Towards a taxonomic coherence between average nucleotide identity and 16S rRNA gene sequence similarity for species demarcation of prokaryotes. *Int J Syst Evol Microbiol*; 64 (Pt 2): 346-351.
58. Kitahara M., F. Takamine, T. Imamura and Y. Benno (2000). Assignment of *Eubacterium* sp. VPI 12708 and related strains with high bile acid 7alpha-dehydroxylating activity to *Clostridium scindens* and proposal of *Clostridium hylemonae* sp. nov., isolated from human faeces. *Int J Syst Evol Microbiol*; 50 Pt 3: 971-978.
59. Kitahara M., F. Takamine, T. Imamura and Y. Benno (2001). Clostridium *hiranonis* sp. nov., a human intestinal bacterium with bile acid 7alpha-dehydroxylating activity. *Int J Syst Evol Microbiol*; 51 (Pt 1): 39-44.
60. Klappenbach J.A., J.M. Dunbar and T.M. Schmidt (2000). rRNA operon copy number reflects ecological strategies of bacteria. *Appl Environ Microbiol*; 66 (4): 1328-1333.
61. Koide A. and J.A. Hoch (1994). Identification of a second oligopeptide transport system in *Bacillus subtilis* and determination of its role in sporulation. *Mol Microbiol*; 13 (3): 417-426.
62. Konstantinidis K.T. and J.M. Tiedje (2005). Genomic insights that advance the species definition for prokaryotes. *Proc Natl Acad Sci USA*; 102 (7): 2567-2572.
63. Kopke M., C. Held, S. Hujer, H. Liesegang, A. Wiezer, A. Wollherr, A. Ehrenreich, *et al.* (2010). *Clostridium ljungdahlii* represents a microbial production platform based on syngas. *Proc Natl Acad Sci USA*; 107 (29): 13087-13092.
64. Kopke M., M. Straub and P. Durre (2013). *Clostridium difficile* is an autotrophic bacterial pathogen. *PLoS One*; 8 (4): e62157.
65. Kopylova E., L. Noe and H. Touzet (2012). SortMeRNA: fast and accurate filtering of ribosomal RNAs in metatranscriptomic data. *Bioinformatics*; 28 (24): 3211-3217.
66. Lagesen K., P. Hallin, E.A. Rodland, H.H. Staerfeldt, T. Rognes and D.W. Ussery (2007). RNAmmer: consistent and rapid annotation of ribosomal RNA genes. *Nucleic Acids Res*; 35 (9): 3100-3108.
67. Lajoie S.F., S. Bank, T.L. Miller and M.J. Wolin (1988). Acetate production from hydrogen and [13C]carbon dioxide by the microflora of human feces. *Appl Environ Microbiol*; 54 (11): 2723-2727.
68. Lee Z.M., C. Bussema, 3rd and T.M. Schmidt (2009). rrnDB: documenting the number of rRNA and tRNA genes in bacteria and archaea. *Nucleic Acids Res*; 37 (Database issue): D489-493.
69. Ley R.E., M. Hamady, C. Lozupone, P.J. Turnbaugh, R.R. Ramey, J.S. Bircher, M.L. Schlegel, *et al.* (2008). Evolution of mammals and their gut microbes. *Science*; 320 (5883): 1647-1651.
70. Li E., C.M. Hamm, A.S. Gulati, R.B. Sartor, H. Chen, X. Wu, T. Zhang, *et al.* (2012). Inflammatory bowel diseases phenotype, *C. difficile* and NOD2 genotype are associated with shifts in human ileum associated microbial composition. *PLoS One*; 7 (6): e26284.
71. Lo C.I., A.K. Mishra, R. Padhmanabhan, B. Samb, A.G. Sow, C. Robert, C. Couderc, *et al.* (2013). Non-contiguous finished genome sequence and description of *Clostridium* dakarense sp. nov. *Stand Genomic Sci*; 9 (1): 14-27.
72. Logan S.M. (2006). Flagellar glycosylation - a new component of the motility repertoire? *Microbiology*; 152 (Pt 5): 1249-1262.
73. Lowe T.M. and S.R. Eddy (1997). tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res*; 25 (5): 955-964.
74. Lubkowitz M.A., D. Barnes, M. Breslav, A. Burchfield, F. Naider and J.M. Becker (1998). Schizosaccharomyces pombe isp4 encodes a transporter representing a novel family of oligopeptide transporters. *Mol Microbiol*; 28 (4): 729-741.
75. Ludwig W., O. Strunk, R. Westram, L. Richter, H. Meier, Yadhukumar, A. Buchner, *et al.* (2004). ARB: a software environment for sequence data. *Nucleic Acids Res*; 32 (4): 1363-1371.
76. Mackey B.M. and J.G. Morris (1971). Ultrastructural changes during sporulation in *Clostridium pasteurianum*. *J Gen Microbiol*; 66 (1): 13.
77. Mann E., S. Schmitz-Esser, Q. Zebeli, M. Wagner, M. Ritzmann and B.U. Metzler-Zebeli (2014). Mucosa-associated bacterial microbiome of the gastrointestinal tract of weaned pigs and dynamics linked to dietary calcium-

phosphorus. *PLoS One*; 9 (1): e86950.
78. Martin M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *2011*; 17 (1).
79. McGarr S.E., J.M. Ridlon and P.B. Hylemon (2005). Diet, anaerobic bacterial metabolism, and colon cancer: a review of the literature. *J Clin Gastroenterol*; 39 (2): 98-109.
80. McLaughlin R.W., M. Chen, J. Zheng, Q. Zhao and D. Wang (2012). Analysis of the bacterial diversity in the fecal material of the endangered Yangtze finless porpoise, *Neophocaena phocaenoides asiaeorientalis. Mol Biol Rep*; 39 (5): 5669-5676.
81. Moschonas G. and D.J. Bolton (2013). Characterization of a potentially novel 'blown pack' spoilage bacterium isolated from bovine hide. *J Appl Microbiol*; 114 (3): 771-777.
82. Moussa R.S. (1958). Complexity of toxins from *Clostridium septicum* and *Clostridium chauvoei*. *J Bacteriol*; 76 (5): 538-545.
83. Neville B.A., P.O. Sheridan, H.M. Harris, S. Coughlan, H.J. Flint, S.H. Duncan, I.B. Jeffery, *et al.* (2013). Pro-inflammatory flagellin proteins of prevalent motile commensal bacteria are variably abundant in the intestinal microbiome of elderly humans. *PLoS One*; 8 (7): e68919.
84. Onyenwoke R.U., J.A. Brill, K. Farahi and J. Wiegel (2004). Sporulation genes in members of the low G+C Gram-type-positive phylogenetic branch (*Firmicutes*). *Arch Microbiol*; 182 (2-3): 182-192.
85. Overbeek R., M. Fonstein, M. D'Souza, G.D. Pusch and N. Maltsev (1999). The use of gene clusters to infer functional coupling. *Proc Natl Acad Sci USA*; 96 (6): 2896-2901.
86. Paredes-Sabja D., A. Shen and J.A. Sorg (2014). *Clostridium difficile* spore biology: sporulation, germination, and spore structural proteins. *Trends Microbiol*; 22 (7): 406-416.
87. Paredes C.J., K.V. Alsaker and E.T. Papoutsakis (2005). A comparative genomic view of clostridial sporulation and physiology. *Nat Rev Microbiol*; 3 (12): 969-978.
88. Pei A.Y., W.E. Oberdorf, C.W. Nossa, A. Agarwal, P. Chokshi, E.A. Gerz, Z. Jin, *et al.* (2010). Diversity of 16S rRNA genes within individual prokaryotic genomes. *Appl Environ Microbiol*; 76 (12): 3886-3897.
89. Pitts A.C., L.R. Tuck, A. Faulds-Pain, R.J. Lewis and J. Marles-Wright (2012). Structural insight into the *Clostridium difficile* ethanolamine utilisation microcompartment. *PLoS One*; 7 (10): e48360.
90. Popoff M.R. and P. Bouvet (2013). Genetic characteristics of toxigenic *Clostridia* and toxin gene evolution. *Toxicon*; 75: 63-89.
91. Postma P.W., J.W. Lengeler and G.R. Jacobson (1993). Phosphoenolpyruvate:carbohydrate phosphotransferase systems of bacteria. *Microbiol Rev*; 57 (3): 543-594.
92. Pruesse E., J. Peplies and F.O. Glockner (2012). SINA: accurate high-throughput multiple sequence alignment of ribosomal RNA genes. *Bioinformatics*; 28 (14): 1823-1829.
93. Quast C., E. Pruesse, P. Yilmaz, J. Gerken, T. Schweer, P. Yarza, J. Peplies, *et al.* (2013). The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res*; 41 (Database issue): D590-596.
94. Rainey F.A., B.J. Hollen and A. Small (2009). Genus 1. *Clostridium*. In: *Bergey's Manual of Systematic Bacteriology*. 2nd Edn, Vol. 3: The *Firmicutes*, pp 738-828. Edited by P. De Vos, G.M. Garrity, D. Jones, N.R. Krieg, *et al.* New York: Springer.
95. Rajilic-Stojanovic M., H.G. Heilig, D. Molenaar, K. Kajander, A. Surakka, H. Smidt and W.M. de Vos (2009). Development and application of the human intestinal tract chip, a phylogenetic microarray: analysis of universally conserved phylotypes in the abundant microbiota of young and elderly adults. *Environ Microbiol*; 11 (7): 1736-1751.
96. Reder A., U. Gerth and M. Hecker (2012). Integration of sigmaB activity into the decision-making process of sporulation initiation in *Bacillus subtilis*. *J Bacteriol*; 194 (5): 1065-1074.
97. Resnekov O., A. Driks and R. Losick (1995). Identification and characterization of sporulation gene spoVS from *Bacillus subtilis*. *J Bacteriol*; 177 (19): 5628-5635.
98. Reti K.L., M.C. Thomas, L.J. Yanke, L.B. Selinger and G.D. Inglis (2013). Effect of antimicrobial growth promoter administration on the intestinal microbiota of beef cattle. *Gut pathogens*; 5: 8.
99. Reysenbach A.L., N. Ravenscroft, S. Long, D.T. Jones and D.R. Woods (1986). Characterization, biosynthesis, and regulation of granulose in *Clostridium acetobutylicum*. *Appl Environ Microbiol*; 52 (1): 185-190.
100. Richter M. and R. Rosselló-Móra (2009). Shifting the genomic gold standard for the prokaryotic species definition. *Proc Natl Acad Sci USA*; 106 (45): 19126-19131.
101. Ridlon J.M., D.J. Kang and P.B. Hylemon (2006). Bile salt biotransformations by human intestinal bacteria. *J Lipid Res*; 47 (2): 241-259.
102. Ridlon J.M., D.J. Kang, P.B. Hylemon and J.S. Bajaj (2014). Bile acids and the gut microbiome. *Cur Opin Gastroenterol*; 30 (3): 332-338.
103. Salonen A., J. Nikkila, J. Jalanka-Tuovinen, O. Immonen, M. Rajilic-Stojanovic, R.A. Kekkonen, A. Palva, *et al.* (2010).

Comparative analysis of fecal DNA extraction methods with phylogenetic microarray: effective recovery of bacterial and archaeal DNA using mechanical cell lysis. *J Microbiol Methods*; 81 (2): 127-134.
104. Schmieder R. and R. Edwards (2011). Quality control and preprocessing of metagenomic datasets. *Bioinformatics*; 27 (6): 863-864.
105. Shcherbakova V.A., N.A. Chuvil'skaia, N.P. Golovchenko, N.E. Suzina, A.M. Lysenko, K.S. Laurinavichus and V.K. Akimenko (2003). [Analysis of the anaerobic microbial community capable of degrading p-toluene sulphonate]. *Mikrobiologiia*; 72 (6): 752-758.
106. Stabler R.A., M. He, L. Dawson, M. Martin, E. Valiente, C. Corton, T.D. Lawley, *et al.* (2009). Comparative genome and phenotypic analysis of *Clostridium difficile* 027 strains provides insight into the evolution of a hypervirulent bacterium. *Genome Biol*; 10 (9): R102.
107. Steiner E., J. Scott, N.P. Minton and K. Winzer (2012). An agr quorum sensing system that regulates granulose formation and sporulation in *Clostridium acetobutylicum*. *Appl Environ Microbiol*; 78 (4): 1113-1122.
108. Suchodolski J.S., J. Camacho and J.M. Steiner (2008). Analysis of bacterial diversity in the canine duodenum, jejunum, ileum, and colon by comparative 16S rRNA gene analysis. *FEMS Microbiol Ecol*; 66 (3): 567-578.
109. Suzek B.E., H. Huang, P. McGarvey, R. Mazumder and C.H. Wu (2007). UniRef: comprehensive and non-redundant UniProt reference clusters. *Bioinformatics*; 23 (10): 1282-1288.
110. Tan I.S. and K.S. Ramamurthi (2014). Spore formation in *Bacillus subtilis*. *Environ Microbiol Rep*; 6 (3): 212-225.
111. Tindall B.J., R. Rosselló-Móra, H.J. Busse, W. Ludwig and P. Kämpfer (2010). Notes on the characterization of prokaryote strains for taxonomic purposes. *Int J Syst Evol Microbiol*; 60 (Pt 1): 249-266.
112. Titball R.W. (2005). Gas gangrene: an open and closed case. *Microbiology*; 151 (Pt 9): 2821-2828.
113. Titball W.R., S.E.C. Hunter, K.L. Martin, B.C. Morris, A.D. Shuttleworth, T. Rubidge, D.W. Anderson, *et al.* (1989). Molecular Cloning and Nucleotide Sequence of the Alpha-Toxin (Phospholipase C) of *Clostridium perfringens*. *Infect Immun*; 57 (2): 367-376.
114. Tovar-Rojo F., M. Chander, B. Setlow and P. Setlow (2002). The products of the spoVA operon are involved in dipicolinic acid uptake into developing spores of Bacillus subtilis. *J Bacteriol*; 184 (2): 584-587.
115. Tso J.Y. and C. Siebel (1989). Cloning and expression of the phospholipase C gene from Clostridium perfringens and Clostridium bifermentans. *Infect Immun*; 57 (2): 468-476.
116. UniProt-Consortium (2014). Activities at the universal protein resource (UniProt). *Nucleic Acids Res*; 42 (Database issue): D191-D198.
117. Van den Bogert B., J. Boekhorst, R. Herrmann, E.J. Smid, E.G. Zoetendal and M. Kleerebezem (2013). Comparative genomics analysis of *Streptococcus* isolates from the human small intestine reveals their adaptation to a highly dynamic ecosystem. *PLoS One*; 8 (12): e83418.
118. van den Bogert B., W.M. de Vos, E.G. Zoetendal and M. Kleerebezem (2011). Microarray analysis and barcoded pyrosequencing provide consistent microbial profiles depending on the source of human intestinal samples. *Appl Environ Microbiol*; 77 (6): 2071-2080.
119. van den Bogert B., O. Erkus, J. Boekhorst, M. de Goffau, E.J. Smid, E.G. Zoetendal and M. Kleerebezem (2013). Diversity of human small intestinal *Streptococcus* and *Veillonella* populations. *FEMS Microbiol Ecol*; 85 (2): 376-388.
120. Van Gerven N., G. Waksman and H. Remaut (2011). Pili and flagella biology, structure, and biotechnological applications. *Prog Mol Biol Transl Sci*; 103: 21-72.
121. Vetrovsky T. and P. Baldrian (2013). The variability of the 16S rRNA gene in bacterial genomes and its consequences for bacterial community analyses. *PLoS One*; 8 (2): e57923.
122. Vital M., A.C. Howe and J.M. Tiedje (2014). Revealing the bacterial butyrate synthesis pathways by analyzing (meta) genomic data. *mBio*; 5 (2): e00889.
123. Wayne L.G., D.J. Brenner, R.R. Colwell, P.A.D. Grimont, O. Kandler, M.I. Krichevsky, L.H. Moore, *et al.* (1987). International Committee on Systematic Bacteriology. Report of the ad hoc committee on reconciliation of approaches to bacterial systematics *Int J Syst Bacteriol*; 37 (4): 463-464.
124. Wells J.E. and P.B. Hylemon (2000). Identification and characterization of a bile acid 7alpha-dehydroxylation operon in *Clostridium* sp. strain TO-931, a highly active 7alpha-dehydroxylating strain isolated from human feces. *Appl Environ Microbiol*; 66 (3): 1107-1113.
125. Williams K.P. (2002). Integration sites for genetic elements in prokaryotic tRNA and tmRNA genes: sublocation preference of integrase subfamilies. *Nucleic Acids Res*; 30 (4): 866-875.
126. Xiao Y. (2014).Clostridium perfringens sporulation, germination and outgrowth in food: a functional genomics approach. PhD thesis, Wageningen University.
127. Xiao Y., C. Francke, T. Abee and M.H. Wells-Bennik (2011). Clostridial spore germination versus bacilli: genome mining and current insights. *Food Microbiol*; 28 (2): 266-274.

128. Xu D., J. Gao, M. Gillilland, 3rd, X. Wu, I. Song, J.Y. Kao and C. Owyang (2014). Rifaximin alters intestinal bacteria and prevents stress-induced gut inflammation and visceral hyperalgesia in rats. *Gastroenterology*; 146 (2): 484-496.
129. Yano K., T. Wada, S. Suzuki, K. Tagami, T. Matsumoto, Y. Shiwa, T. Ishige, *et al.* (2013). Multiple rRNA operons are essential for efficient cell growth and sporulation as well as outgrowth in *Bacillus subtilis*. *Microbiology*; 159 (Pt 11): 2225-2236.
130. Yarza P., M. Richter, J. Peplies, J. Euzeby, R. Amann, K.H. Schleifer, W. Ludwig, *et al.* (2008). The All-Species Living Tree project: a 16S rRNA-based phylogenetic tree of all sequenced type strains. *Syst Appl Microbiol*; 31 (4): 241-250.
131. Yoshimoto S., T.M. Loo, K. Atarashi, H. Kanda, S. Sato, S. Oyadomari, Y. Iwakura, *et al.* (2013). Obesity-induced gut microbial metabolite promotes liver cancer through senescence secretome. *Nature;* 499 (7456): 97-101.

Supplemental material

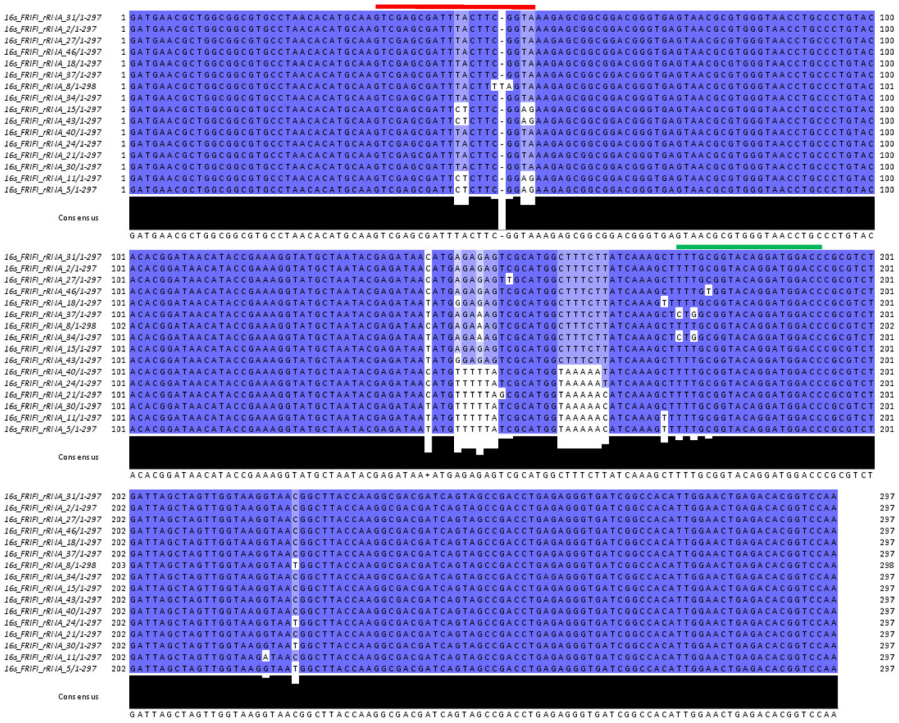

Figure S1. Alignment of the V1-V2 regions of 16 copies of the 16S rRNA gene that are present in the genome of *Romboutsia hominis* FRIFI. Nucleotides are coloured by percentage identity. The red and green line indicate the positions of the *R.* ilealis-specific primers 61F and 235R, respectively.

Table S1. Overview of copy numbers of the 16S rRNA gene reported for the bacterial species containing more than ten copies. Numbers are derived from the rrnDB v4.0.0 [68]. In cases with more than one genome available for a particular species, the copy number is provided for each genome.

Species name	No. of genomes	List of 16S rRNA gene counts
***Romboutsia hominis* FRIFI**	**1**	**16**
Brevibacillus brevis	1	15
Clostridium paradoxum	1	15
Photobacterium profundum	1	15
Bacillus thuringiensis	14	15, 14, 14, 14, 14, 14, 13, 13, 13, 13, 13, 13, 12, 12
Bacillus weihenstephanensis	6	14, 14, 14, 14, 14, 14
Clostridium beijerinckii	1	14
Romboutsia ilealis	**1**	**14**
Shewanella violacea	1	14
Paenibacillus mucilaginosus	2	14, 13
Paenibacillus polymyxa	4	14, 14, 13, 12
Bacillus cereus	13	14, 14, 14, 14, 13, 13, 13, 13, 13, 13, 12, 12, 11
Bacillus cytotoxicus	1	13
Vibrio natriegens	1	13
Aliivibrio fischeri	1	12
Aliivibrio salmonicida	1	12
Bacillus toyonensis	1	12
Bacillus sp. 1NLA3E	1	12
Clostridium saccharobutylicum	1	12
Paenibacillus sp. JDR-2	1	12
Shewanella sediminis	1	12
Solibacillus silvestris	1	12
Bacillus megaterium	3	12, 11, 11
Vibrio parahaemolyticus	4	12, 11, 11, 11
Clostridium acetobutylicum	3	11, 11, 11
Clostridium saccharoperbutylacetonicum	1	11
Shewanella pealeana	1	11
Vibrio alginolyticus	1	11
Vibrio sp. Ex25	1	11
Bacillus anthracis	6	11, 11, 11, 11, 11, 10
Clostridium difficile	2	11, 10
Clostridium botulinum	12	11, 11, 10, 9, 9, 9, 9, 9, 9, 9, 8, 8

Table S2. Overview of the ANI values calculated by pairwise comparisons of available genomes within the family *Peptostreptococcaceae*. In cases with more than one genome available for a particular species, the ANI values are above the cut-off point of 95-96% ANI used for species delineation (indicated in light-grey).

	Species name	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1	*Bacillus subtilis* subsp. *subtilis* 168	100	61.0	64.7	62.0	64.3	64.4	61.7	61.4	64.8	63.8	64.4	65.2	64.7	65.1	64.2	65.4	64.9	64.0
2	*Intestinibacter bartlettii* DORA_8_9	61.3	100	**98.5**	71.5	71.7	71.8	71.6	73.7	71.0	71.7	71.7	66.9	66.8	66.7	72.8	73.3	72.6	71.7
3	*Intestinibacter bartlettii* DSM 16795^T	63.7	**98.2**	100	72.0	72.3	73.2	72.2	73.7	71.7	72.1	72.0	68.1	67.8	67.8	73.1	73.7	73.2	72.1
4	*Clostridium bifermentans* ATCC 19299	62.7	71.0	71.5	100	**96.0**	73.4	73.1	72.3	70.8	80.5	80.4	67.0	67.0	66.9	75.0	74.8	74.8	73.8
5	*Clostridium bifermentans* ATCC 638^T	64.9	71.1	72.4	**96.0**	100	73.7	72.9	72.3	71.4	80.5	80.5	68.5	68.2	68.4	75.3	75.3	75.6	74.0
6	*Clostridium* difficile 630	64.8	71.8	73.4	73.3	73.9	100	**99.4**	72.5	72.6	74.0	73.6	71.5	70.7	70.5	75.7	75.9	75.3	73.1
7	*Clostridium difficile* CD196	61.1	71.4	71.9	72.9	73.0	**99.8**	100	71.9	71.3	72.9	73.0	67.1	67.1	66.6	75.0	74.8	74.4	72.7
8	*Terrisporobacter glycolicus* ATCC 14880^T	61.0	73.0	73.1	71.9	72.0	71.8	71.5	100	70.2	72.1	72.0	66.6	66.5	66.2	73.2	73.3	73.0	71.6
9	*Clostridium hiranonis* DSM 13275^T	61.8	70.9	71.5	71.1	71.2	72.0	71.4	70.8	100	71.1	70.9	68.8	68.7	68.0	71.6	72.1	71.6	70.7
10	*Clostridium sordellii* ATTC 9714^T	62.2	71.3	71.7	80.5	80.5	73.4	72.9	72.6	71.0	100	**98.0**	67.5	67.5	67.4	75.2	74.9	75.0	74.7
11	*Clostridium sordellii* VPI 9048	62.4	71.3	71.6	80.3	80.3	73.3	73.0	72.3	71.0	**98.0**	100	67.3	67.2	66.9	75.2	74.8	74.9	74.5
12	*Peptostreptococcus anaerobius* 653-L	61.6	66.9	67.7	67.1	67.4	70.5	66.9	67.5	68.9	67.5	67.4	100	**98.6**	74.1	67.7	68.2	67.6	67.3
13	*Peptostreptococcus anaerobius* VPI 4330^T	61.8	66.7	67.4	67.1	67.4	69.5	67.0	66.9	68.9	67.2	67.0	**98.8**	100	73.7	67.4	67.7	67.5	67.1
14	*Peptostreptococcus stomatis* DSM 17678^T	61.8	66.9	67.3	66.7	67.1	69.5	66.8	66.6	67.9	67.1	67.1	74.1	73.9	100	67.2	67.6	67.1	67.2
15	**Clostridium dakarense FF1^T**	61.9	72.3	72.6	74.8	74.9	75.5	75.1	73.5	71.2	75.1	75.0	67.1	66.9	66.9	100	77.8	78.5	76.3
16	**Romboutsia ilealis CRIB^T**	66.8	73.2	74.4	75.1	75.8	76.4	75.4	74.0	73.0	75.8	75.5	70.1	69.9	69.9	78.8	100	78.7	77.1
17	*Romboutsia lituseburensis* A25K^T	67.4	72.1	73.9	75.1	75.8	75.7	74.7	73.5	72.7	75.5	75.4	69.8	69.3	69.8	79.1	78.6	100	77.1
18	*Romboutsia hominis* FRIFI	66.5	72.8	74.1	75.2	75.8	75.8	74.9	73.6	72.8	76.4	76.2	69.8	69.5	69.8	78.7	78.4	79.1	100

Table S3. Overview of sporulation-related genes in the four *Romboutsia* genomes. Genes from *Bacillus subtilis* subsp. *subtilis* 168, to which no homologues could be identified, are omitted. In case multiple candidate loci were detected, all are mentioned. Loci that are assigned to more than one gene from *B. subtilis* subsp. *subtilis* 168 are marked with an asterisk.

Gene name	Synonym	Locus in *B.subtilis* subsp. *subtilis* 168	Locus in *R. ilealis* CRIB^T	Locus in *R. hominis* FRIFI	Locus in *R. lituseburensis* A25K^T	Locus in *C. dakarense* FF1^T
Identification of sporulation genes that are conserved in all Bacilli and Clostridia						
spo0A		BSU24220	CRIB_1904	FRIFI_2159	RLITU_2811	DAK_0332
spo0J		BSU40960	CRIB_2415	FRIFI_2795	RLITU_3485	DAK_2705
sigE	spoIIGB	BSU15320	CRIB_1992	FRIFI_2269	RLITU_2916	DAK_1342
sigF	spoIIAC	BSU23450	CRIB_2136	FRIFI_2447	RLITU_3121	DAK_3848
sigG	spoIIIG	BSU15330	CRIB_1991	FRIFI_2268	RLITU_2915	DAK_1343
sigH	spo0H	BSU00980	CRIB_2468	FRIFI_2836	RLITU_3537	DAK_1219
sigK	spoIIIC spoIVCB	BSUW23 12795	CRIB_1886	FRIFI_2141	RLITU_2793	DAK_0352
spmA		BSU23180	CRIB_842	FRIFI_0010	RLITU_0011	DAK_2856
spmB		BSU23170	CRIB_843	FRIFI_0011	RLITU_0012	DAK_2855
obgE		BSU27920	CRIB_416	FRIFI_0507	RLITU_0517	DAK_1460
spoIIAA		BSU23470	CRIB_2138	FRIFI_2449	RLITU_3123	DAK_3850
spoIIAB		BSU23460	CRIB_2137	FRIFI_2448	RLITU_3122	DAK_3849
spoIID		BSU36750	CRIB_2343	FRIFI_2687	RLITU_3377	DAK_1825
spoIIE	spoIIH	BSU00640	CRIB_42	FRIFI_0041	RLITU_0053	DAK_3064
spoIIGA		BSU15310	CRIB_1993	FRIFI_2270	RLITU_2918	DAK_1341
spoIIM		BSU23530	CRIB_1896	FRIFI_2151	RLITU_2803	DAK_0342
spoIIP		BSU25530	CRIB_445	FRIFI_0542	RLITU_0558	DAK_0743
spoIIR		BSU36970	CRIB_2365	FRIFI_2717	RLITU_3400	DAK_2541
spoIIIAA		BSU24430	CRIB_1923	FRIFI_2178	RLITU_2830	DAK_0311
spoIIIAB		BSU24420	CRIB_1922	FRIFI_2177	RLITU_2829	DAK_0312
spoIIIAC		BSU24410	CRIB_1921	FRIFI_2176	RLITU_2828	DAK_0313
spoIIIAD		BSU24400	CRIB_1920	FRIFI_2175	RLITU_2827	DAK_0314
spoIIIAE		BSU24390	CRIB_1919	FRIFI_2174	RLITU_2826	DAK_0315
spoIIIAF		BSU24380	CRIB_1918	FRIFI_2173	RLITU_2825	DAK_0316
spoIIIAG		BSU24370	CRIB_1917	FRIFI_2172	RLITU_2824	DAK_0317
spoIIIAH		BSU24360	CRIB_1916	FRIFI_2171	RLITU_2823	DAK_0318
spoIIID		BSU36420	CRIB_2341	FRIFI_2685	RLITU_3375	DAK_1827
spoIIIE		BSU16800	CRIB_1795	FRIFI_2005	RLITU_2672	DAK_1125
spoIIIJ		BSU41040	CRIB_2422	FRIFI_2802	RLITU_3492	DAK_2712
spoIVA		BSU22800	CRIB_1974	FRIFI_2251	RLITU_2898	DAK_0262
spoIVB		BSU24230	CRIB_1905	FRIFI_2160	RLITU_2812	DAK_2787
spoVAC		BSU23420	CRIB_2133	FRIFI_2444	RLITU_3118	DAK_0671 / DAK_3845
spoVAD		BSU23410	CRIB_2132	FRIFI_2443	RLITU_3117	DAK_3844 / DAK_0670
spoVAEB		BSU23402	CRIB_2131	FRIFI_2442	RLITU_3116	DAK_0669 / DAK_3843
spoVC		BSU00530	CRIB_30	FRIFI_0029	RLITU_0041	DAK_3078
spoVD	ftsI	BSU15170	CRIB_2004	FRIFI_2279	RLITU_2931	DAK_1329
spoVG		BSU00490	CRIB_23	FRIFI_0023	RLITU_0026	DAK_3096
spoVK		BSU17420	CRIB_2194	FRIFI_1768	RLITU_3196	DAK_0980
spoVS		BSU16980			RLITU_1653	

Gene name	Synonym	Locus in *B.subtilis* subsp. *subtilis* 168	Locus in *R. ilealis* CRIB^T	Locus in *R. hominis* FRIFI	Locus in *R. lituseburensis* A25K^T	Locus in *C. dakarense* FF1^T
spoVT		BSU00560	CRIB_33	FRIFI_0032	RLITU_0044	DAK_3073
dapA		BSU16770	CRIB_82	FRIFI_0168	RLITU_0200	DAK_3602
dapB		BSU22490	CRIB_83	FRIFI_0167	RLITU_0199	DAK_3601
cwlD		BSU01530	CRIB_2521	FRIFI_2890	RLITU_3590	DAK_0610
stoA	spoIVH ykvV	BSU13840	CRIB_1690	FRIFI_1892	RLITU_2518	DAK_1979
ylbJ		BSU15030	CRIB_1929	FRIFI_2198	RLITU_2849	DAK_0293
yabP		BSU00600	CRIB_38	FRIFI_0037	RLITU_0049	DAK_3068
yabQ		BSU00610	CRIB_39	FRIFI_0038	RLITU_0050	DAK_3067
yqfC		BSU25360	CRIB_465	FRIFI_0626	RLITU_0592	DAK_2383
yqfD		BSU25350	CRIB_466	FRIFI_0627	RLITU_0593	DAK_2382
alr	yncD	BSU17640	CRIB_841	FRIFI_0068 / FRIFI_0433	RLITU_2625 / RLITU_0081	DAK_2857 / DAK_3035
dacB		BSU23190	CRIB_1675	FRIFI_1852	RLITU_2462	DAK_0725
lgt		BSU34990	CRIB_2007	FRIFI_2282	RLITU_2934	DAK_1322
gpr		BSU25540	CRIB_444	FRIFI_0541	RLITU_0557	DAK_0744
jag		BSU41030	CRIB_2421	FRIFI_2801	RLITU_3491	DAK_2711
ytvI		BSU29160	CRIB_1578	FRIFI_0439	RLITU_0447	DAK_1857
yyaC		BSU40950	CRIB_2339	FRIFI_2682	RLITU_3372 / RLITU_1704	DAK_1830
spoVB		BSU27670	CRIB_34 * / CRIB_377 / CRIB_589	FRIFI_0033 * / FRIFI_0785 / FRIFI_0466	RLITU_1122 / RLITU_0474 / RLITU_0045 *	DAK_2610 / DAK_3072 *
ykvU		BSU13830	CRIB_34 *	FRIFI_0033 *	RLITU_0045 *	DAK_3072 *
Identification of sporulation genes that are conserved in all Bacilli and most Clostridia						
spo0E		BSU13640	CRIB_144	FRIFI_0130		DAK_3014 / DAK_2020 / DAK_1809
spoIVFB		BSU27970	CRIB_182	FRIFI_0176	RLITU_0203	DAK_3605 *
spoVAF		BSU23390			RLITU_2061	DAK_3306
spoVE	ftsW	BSU15210	CRIB_2000	FRIFI_2275	RLITU_2927	DAK_1333
cotJC		BSU06910	CRIB_522 / CRIB_483	FRIFI_0648 / FRIFI_1024	RLITU_0853 / RLITU_0614	DAK_2361 / DAK_0995 / DAK_3575
yabG		BSU00430	CRIB_2370	FRIFI_2722	RLITU_3405	DAK_2831
ydhD		BSU05710	CRIB_1736	FRIFI_1942	RLITU_2593	DAK_1104
yhaX		BSU09830	CRIB_1499	FRIFI_1432		
yhbA		BSU08910	CRIB_2206	FRIFI_2522	RLITU_3210	DAK_0110
yhbB		BSU08920	CRIB_2057 *		RLITU_2990 *	DAK_0890 *
yhjR		BSU10610			RLITU_3652	
yyaA		BSU40990	CRIB_2417	FRIFI_2797	RLITU_3487	DAK_2707
yjcM		BSU11910	CRIB_2057 *	FRIFI_2339	RLITU_2116 / RLITU_2990 *	DAK_3857 * / DAK_3858 / DAK_0890 * / DAK_1748

Gene name	Synonym	Locus in *B.subtilis* subsp. *subtilis* 168	Locus in *R. ilealis* CRIB^T	Locus in *R. hominis* FRIFI	Locus in *R. lituseburensis* A25K^T	Locus in *C. dakarense* FF1^T
yqfU		BSU25110	CRIB_711 / CRIB_1569 / CRIB_1727	FRIFI_1932 / FRIFI_1334	RLITU_2582	DAK_1095 / DAK_3972 / DAK_0248 / DAK_0412 / DAK_1197
yunB		BSU32350	CRIB_2129	FRIFI_2440	RLITU_3114	DAK_2788
yhcQ		BSU09180		FRIFI_2077 * / FRIFI_1553 *		DAK_3172 * / DAK_3153 *
yraD		BSU26990		FRIFI_2077 * / FRIFI_1553 *		DAK_3172 * / DAK_3153 *
yraF		BSU26960		FRIFI_2077 * / FRIFI_1553 *		DAK_3172 * / DAK_3153 *
yusN		BSU32860	CRIB_305	FRIFI_2077 * / FRIFI_1553 *		DAK_3172 * / DAK_3153 *
rapD		BSU36380				DAK_1669
sspA		BSU29570	CRIB_171 *	FRIFI_0154 *	RLITU_0187 *	DAK_3384 *
sspB		BSU09750	CRIB_171 *	FRIFI_0154 *	RLITU_0187 *	DAK_3384 *
sspC		BSU19950	CRIB_171 *	FRIFI_0154 *	RLITU_0187 *	DAK_3384 *
sspD		BSU13470	CRIB_171 *	FRIFI_0154 *	RLITU_0187 *	DAK_3384 *
Identification of sporulation genes that are conserved in most Bacilli and some Clostridia						
spo0M		BSU08760			RLITU_1449	
spoVAA		BSU23440	CRIB_2135	FRIFI_2446	RLITU_3120	DAK_3847
spoVAB		BSU23430	CRIB_2134	FRIFI_2445	RLITU_3119	DAK_3846
cotH		BSU36060		FRIFI_1803	RLITU_1113	DAK_0520 / DAK_0521 / DAK_1039
cotJA		BSU06890	CRIB_524	FRIFI_0650	RLITU_0616	DAK_3573
cotJB		BSU06900	CRIB_523	FRIFI_0649	RLITU_0615	DAK_3574
yisY		BSU10900	CRIB_1479	FRIFI_0995 *	RLITU_1637	DAK_2521 *
yqfS		BSU25130	CRIB_2209	FRIFI_2525	RLITU_3213	DAK_0113
ykuD		BSU14040				DAK_1840 *
ypqA		BSU22240	CRIB_1858	FRIFI_2112	RLITU_2758	DAK_1621
yfhM		BSU08590	CRIB_1518	FRIFI_0995 *	RLITU_1604	DAK_2521 *
ytgP		BSU30050	CRIB_286	FRIFI_0353	RLITU_0342	DAK_0252
ytxC		BSU28960	CRIB_2198	FRIFI_2513	RLITU_3200	DAK_0984
Other sporulation genes						
spoVID		BSU28110				DAK_1840 *
safA	yrbA exsA	BSU27840				DAK_1840 *
sspF		BSU00450				DAK_3605 *
cwlH		BSU25710				DAK_1434
gerM		BSU28380				DAK_2170
ahpF		BSU40100				DAK_1939
dppA		BSU12920				DAK_2858 *
dciAA	DppA	BSU12920				DAK_2858 *
ydfS		BSU05540				DAK_3729
yqkA		BSU23670				DAK_2990
yckK	tcyA	BSU03610			RLITU_1129 / RLITU_2054	DAK_3120

Gene name	Synonym	Locus in *B.subtilis* subsp. *subtilis* 168	Locus in *R. ilealis* CRIB^T	Locus in *R. hominis* FRIFI	Locus in *R. lituseburensis* A25K^T	Locus in *C. dakarense* FF1^T
oppA	spo0KA	BSU11430		FRIFI_1568	RLITU_102 / RLITU_0350	DAK_2261 / DAK_0009
sleL	yaaH	BSU00160			RLITU_0806	DAK_1841
yqhH		BSU24580			RLITU_1041	
yusA	metQ	BSU32730			RLITU_1045	DAK_3344
pbpE		BSU34440			RLITU_1191	
yhcO		BSU09165			RLITU_1212	
gerE		BSU28410			RLITU_1453	DAK_2480 / DAK_2479
sinR		BSU24610			RLITU_1707	DAK_1296
cgeB		BSU19790			RLITU_2271 *	
ykvP		BSU13780			RLITU_2271 *	DAK_1840 *
spsG		BSU37850			RLITU_2345	
spsE		BSU37870			RLITU_2348	
spsF		BSU37860			RLITU_2349	
spsC		BSU37890			RLITU_2350	
kinE		BSU13530			RLITU_2363	DAK_0757 / DAK_3420
ykuL		BSU14130			RLITU_2632	
spsK		BSU37820			RLITU_3038	
yjmD	gutB0	BSU12330			RLITU_3044	
spsJ		BSU37830			RLITU_3049	
spsL		BSU37810			RLITU_3050	DAK_0066
spsI		BSU37840			RLITU_3051	
splB	spI	BSU13930			RLITU_3297	
oppC	spo0KC	BSU11450		FRIFI_0084 / FRIFI_1566	RLITU_0099 / RLITU_0349_ Oligopeptide	DAK_2258 / DAK_0007
oppB	spo0KB	BSU11440		FRIFI_0085 / FRIFI_1567	RLITU_0100 / RLITU_0348	DAK_2259 / DAK_0008
oppD	spo0KD	BSU11460		FRIFI_0083 / FRIFI_1565	RLITU_0098 / RLITU_0351	DAK_2257 / DAK_0004
oppF	spo0KE	BSU11470		FRIFI_0082 / FRIFI_1564	RLITU_0097 / RLITU_0352	DAK_2256 / DAK_0003
ytrE		BSU30420		FRIFI_0205 *	RLITU_1066 *	
slrR		BSU34380		FRIFI_0260		
yuzA		BSU31380		FRIFI_0413	RLITU_0001 / RLITU_1439	DAK_2883
spsB		BSU37900		FRIFI_0707		
tlp		BSU18030		FRIFI_0871		
arsB		BSU25790		FRIFI_0921	RLITU_3646	DAK_0478
paiA		BSU32150		FRIFI_1027		DAK_1443
yhfW		BSU10390		FRIFI_1457	RLITU_0951	DAK_3668
cotQ	yvdP	BSU34520		FRIFI_1556		DAK_1558
yybI		BSU40630		FRIFI_1593	RLITU_1015	
ytrB		BSU30450		FRIFI_1679		
yocH		BSU19210		FRIFI_1853	RLITU_2464	DAK_1840 *
dltC		BSU38520		FRIFI_2045	RLITU_0270	DAK_2746

Gene name	Synonym	Locus in *B.subtilis* subsp. *subtilis* 168	Locus in *R. ilealis* CRIB^T	Locus in *R. hominis* FRIFI	Locus in *R. lituseburensis* A25K^T	Locus in *C. dakarense* FF1^T
dltA		BSU38500		FRIFI_2047	RLITU_0268	DAK_2744
dltD		BSU38530		FRIFI_2048	RLITU_0267	DAK_2743 / DAK_3220
ytfI		BSU29510		FRIFI_2071	RLITU_2720	
tuaG		BSU35550		FRIFI_2347 *	RLITU_3004 *	DAK_0883 *
sleB		BSU22930		FRIFI_2716		DAK_2540
adhB		BSU26970		FRIFI_2766		
yxbB		BSU39890		FRIFI_2769		
yrkC		BSU26560		FRIFI_0756 / FRIFI_1138 / FRIFI_2763	RLITU_1750 / RLITU_0303 / RLITU_1610	DAK_3191 / DAK_3450 / DAK_1559 / DAK_2331 / DAK_1709
yttP		BSU29630		FRIFI_1502 / FRIFI_0787	RLITU_0850	DAK_1547
yybM		BSU40590		FRIFI_1692 / FRIFI_2603 / FRIFI_0214 / FRIFI_2605	RLITU_1142 / RLITU_3660 / RLITU_1068 / RLITU_1144 / RLITU_1067 / RLITU_1246 / RLITU_3661 / RLITU_3596 / RLITU_0650 / RLITU_3674	DAK_2096 / DAK_0188 / DAK_0631 / DAK_0633 / DAK_3625 / DAK_2037
yfhP		BSU08620		FRIFI_2089 / FRIFI_1054	RLITU_2734	DAK_3234
ytkD		BSU30630	CRIB_101	FRIFI_0100	RLITU_0123	DAK_2274
exoA		BSU22010	CRIB_1042		RLITU_1577	DAK_1495
ytkL		BSU29410	CRIB_1305			DAK_1586
yngK		BSU18280	CRIB_1306	FRIFI_2342	RLITU_2993	DAK_0887
ytlA		BSU30595	CRIB_1424	FRIFI_1522	RLITU_1856	DAK_0065
ytlC		BSU30610	CRIB_1447	FRIFI_0205 *	RLITU_1066 *	DAK_0063
ytlD		BSU30620	CRIB_1448	FRIFI_1462	RLITU_1831	DAK_0064
asnO		BSU10790	CRIB_1455		RLITU_1730	DAK_1839
hmp		BSU13040	CRIB_1533			
yocI		BSU19220	CRIB_1602	FRIFI_1888	RLITU_2512	DAK_1739
cwlC		BSU17410	CRIB_1642	FRIFI_1631	RLITU_0927	
cotSA		BSU30910	CRIB_1659			
sodA		BSU25020	CRIB_1699	FRIFI_1902 / FRIFI_1046	RLITU_2534	DAK_0454 / DAK_0870
kgsA		BSU00420	CRIB_17	FRIFI_0017	RLITU_0019	DAK_3104
ylbC		BSU14960	CRIB_1724	FRIFI_1930	RLITU_2580	DAK_2049
pdaA	yfjs	BSU07980	CRIB_1739	FRIFI_1948	RLITU_2600	DAK_1518
kapD		BSU31470	CRIB_1778	FRIFI_1987	RLITU_2654	DAK_1142
ywkD		BSU37020	CRIB_180	FRIFI_0163	RLITU_1184	DAK_3300

Gene name	Synonym	Locus in *B.subtilis* subsp. *subtilis* 168	Locus in *R. ilealis* CRIB^T	Locus in *R. hominis* FRIFI	Locus in *R. lituseburensis* A25K^T	Locus in *C. dakarense* FF1^T
ytfJ		BSU29500	CRIB_1823	FRIFI_2070	RLITU_2719	DAK_3253
dacF		BSU23480	CRIB_1829	FRIFI_2084	RLITU_2724	DAK_3241
cggR		BSU33950	CRIB_186	FRIFI_0180	RLITU_0207	DAK_0919
lspA		BSU15450	CRIB_1960	FRIFI_2230	RLITU_2881	DAK_1408
yteA		BSU30840	CRIB_1962	FRIFI_2232	RLITU_2883	DAK_1406
sbp		BSU15270	CRIB_1995		RLITU_2920	DAK_1339
cotR		BSU34530	CRIB_2050	FRIFI_2329	RLITU_2983	DAK_0899
dltB		BSU38510	CRIB_2054	FRIFI_2046	RLITU_0269	DAK_2745
yrzE		BSU27690	CRIB_2101	FRIFI_2412	RLITU_3087	DAK_2240
yrbC		BSU27820	CRIB_2112	FRIFI_2423	RLITU_3099	DAK_2203
yvaB		BSU33540	CRIB_2148	FRIFI_1787	RLITU_1679	
uxuA		BSU12340	CRIB_2245	FRIFI_1546	RLITU_1331	DAK_3180
uxuB	yjmF	BSU12350	CRIB_2248	FRIFI_1549	RLITU_1328	DAK_3177
spoIVCA		BSU25770	CRIB_231		RLITU_3655	DAK_1689
ydcI		BSU04780	CRIB_2331	FRIFI_2674	RLITU_3364	DAK_1013
ykvT		BSU13820	CRIB_2364		RLITU_3399	
yxjC		BSU39000	CRIB_2393	FRIFI_2772		
soj		BSU40970	CRIB_2416	FRIFI_2796	RLITU_3640	DAK_2706
dnaA		BSU00010	CRIB_2426	FRIFI_2806	RLITU_3496	DAK_2716
ytxH		BSU29770	CRIB_2432	FRIFI_2812	RLITU_3502	DAK_2722
rsbW		BSU04720	CRIB_2435	FRIFI_2815	RLITU_3505	DAK_2725
ykvI		BSU13710	CRIB_301	FRIFI_0378	RLITU_0368	DAK_0227
yqhQ		BSU24490	CRIB_50	FRIFI_0048	RLITU_0061	DAK_3055
yloB		BSU15650	CRIB_560	FRIFI_1447	RLITU_0930	DAK_2024
dppE		BSU12960	CRIB_578	FRIFI_0087	RLITU_0103	
dppD		BSU12950	CRIB_581	FRIFI_0205 *	RLITU_1066 *	
yabC		BSU00360	CRIB_6	FRIFI_0006	RLITU_0007	DAK_2877
tpx		BSU29490	CRIB_607	FRIFI_0745	RLITU_1367	DAK_1007
uxaB		BSU12380	CRIB_650			
uxaA		BSU12390	CRIB_651			
sbcC	yiry	BSU10650	CRIB_754	FRIFI_0436	RLITU_0444 / RLITU_0443	DAK_2841 / DAK_0328 / DAK_2842
yoaR		BSU18720	CRIB_759	FRIFI_1714	RLITU_3207	DAK_0991
ytrF		BSU30410	CRIB_765	FRIFI_1571		
ykfA		BSU12970	CRIB_844	FRIFI_0434	RLITU_2624	DAK_2854
ctpB		BSU35240	CRIB_86	FRIFI_0076	RLITU_0089	DAK_3027
yesJ		BSU06920	CRIB_929	FRIFI_0944	RLITU_1714	
yqgT		BSU24830	CRIB_993			
spoIIQ		BSU36550	CRIB_1775 * / CRIB_531 * / CRIB_2342 *	FRIFI_1985 * / FRIFI_0657 * / FRIFI_2686 *	RLITU_3376 * / RLITU_2648 *	DAK_1826 * / DAK_3563 * / DAK_1149 *
spoIVFA		BSU27980	CRIB_1775 * / CRIB_531 * / CRIB_2342 *	FRIFI_1985 * / FRIFI_0657 * / FRIFI_2686 *	RLITU_3376 * / RLITU_2648 *	DAK_1826 * / DAK_3563 * / DAK_1149 *
lytH	yunA	BSU32340	CRIB_1775 * / CRIB_531 * / CRIB_2342 *	FRIFI_1985 * / FRIFI_0657 * / FRIFI_2686 *	RLITU_3376 * / RLITU_2648 *	DAK_1826 * / DAK_3563 * / DAK_1149 *

Gene name	Synonym	Locus in *B.subtilis* subsp. *subtilis* 168	Locus in *R. ilealis* CRIB^T	Locus in *R. hominis* FRIFI	Locus in *R. lituseburensis* A25K^T	Locus in *C. dakarense* FF1^T
ymxH	ylmC family	BSU16720	CRIB_1797 / CRIB_1986	FRIFI_2008 / FRIFI_2263	RLITU_2676 / RLITU_2910	DAK_1122 / DAK_1348
seaA		BSU22850	CRIB_1980 / CRIB_1981	FRIFI_2258 / FRIFI_2257	RLITU_2904 / RLITU_2905	DAK_0255 / DAK_0256
cgeD		BSU19760	CRIB_2065 / CRIB_1900	FRIFI_2347 * / FRIFI_1833 / FRIFI_2155 / FRIFI_2348 / FRIFI_2350	RLITU_0793 / RLITU_3042 / RLITU_3004 * / RLITU_2807 / RLITU_2995 / RLITU_3052 / RLITU_0794 / RLITU_3014 / RLITU_3013 / RLITU_2247 / RLITU_3040 / RLITU_2274	DAK_0338 / DAK_0882 * / DAK_1670 / DAK_0883 * / DAK_0879 / DAK_0881
dnaD		BSU22350	CRIB_2398 / CRIB_1526	FRIFI_1169 / FRIFI_0970 / FRIFI_2776	RLITU_3465 / RLITU_1810 / RLITU_1189	DAK_2687
yviA		BSU35480	CRIB_956 / CRIB_1409 / CRIB_1378 / CRIB_1802	FRIFI_1655 / FRIFI_2518 / FRIFI_0788 / FRIFI_1887 / FRIFI_0133	RLITU_1996 / RLITU_0165 / RLITU_2510 / RLITU_3205 / RLITU_2680 / RLITU_1991 / RLITU_1132	DAK_3017 / DAK_1738 / DAK_1967 / DAK_0279 / DAK_0989

7 CHAPTER

General discussion

Next-generation probiotics: novel probiotic strains of intestinal origin

As discussed in **Chapter 2**, most of the strains on the probiotic market belong to the bacterial genera *Bifidobacterium* and *Lactobacillus,* and few belong to other genera like *Enterococcus*, *Escherichia*, *Bacillus* or *Streptococcus*. However, it is obvious that the limited number of bacterial genera that contain currently acknowledged probiotic strains does not cover the complete diversity of the commensal intestinal microbiota. Therefore it is possible, maybe even likely, that there are other genera that include strains with higher probiotic effectiveness than 'traditional' probiotic genera. Nevertheless, after decades of research on bacterial strains for their potential use as probiotics, not many new players have made it into the arena.

Most of the bacterial strains currently available on the probiotic market are selected based on specific properties such as resistance to acids and bile salts, high viability or the simple fact that they are on the list of microbes safe for human consumption [25]. Moreover, some strains are selected for their functional properties, such as stimulation of immune cells, adherence to epithelial cells, or their ability to inhibit growth of potential pathogens [23]. Large collections of strains can be screened for (functional) properties using *in vitro* assays. Based on the results obtained with these screening assays, a small set of strains can be selected that merits further investigation. Since novel bacterial strains are isolated on a daily basis, and increasing numbers of (high-throughput) *in vitro* screening approaches are developed, an ever growing collection of microbes can be screened for their *in vitro* activities. This approach was followed for the strain selection for the probiotic mixture used in the study that was described in **Chapter 3** [93].

The insights from culture-independent studies have led to another strategy for the identification and selection of novel probiotic strains. As discussed in **Chapter 2,** culture-independent studies have delivered and will continue to deliver a wealth of information on microbial composition and diversity, and more recently also on functionality, in health and disease [104, 111]. This information can be used to identify specific microbial species or strains that are underrepresented in specific diseases and may potentially be supplemented in order to improve health (next-generation probiotics). For example, commensal butyrate-producing bacteria have been suggested to have probiotic potential for patients suffering from IBD [59]. Among the proposed candidates is the butyrate-producing microbe *Faecalibacterium prausnitzii*, an abundant species in the human intestinal tract [109]. Although this species is found in the top 57 of most common species found in the human intestinal tract in ≥90% of the studied population as reported by Qin *et al.*, the range in relative abundance can be over a thousand-fold [74]. Sokol *et al.* were the first to report about an association between low abundance of *F. prausnitzii* in the intestinal tract and remission time of Crohn's disease [89]. Since then, numerous studies have shown an association between the relative abundance of *F. prausnitzii* and a number of other diseases. Based on these findings, it has been suggested that the relative abundance

of *F. prausnitzii* can serve as a biomarker for intestinal health [63]. Moreover, it has been demonstrated in animal model studies that *F. prausnitzii* has probiotic properties as well. In a mouse model the supplementation of *F. prausnitzii* (and its spent medium) resulted in a marked reduction of colitis and a tendency to counterbalance the colitis-associated aberrations that were termed dysbiosis [88, 107]. In this example, the characterization of compositional dysbiosis has led to the identification of possible commensal bacterial species that could be used to restore dysbiosis and subsequently improve health. Similarly, Machiels *et al.* have observed decreased numbers of *Roseburia hominis* (and *F. prausnitzii*) in ulcerative colitis patients [57]. Based on these findings, a probiotic potential for *R. hominis* as butyrate-producing commensal has been suggested as well [92]. A different butyrate-producer with probiotic potential is *Butyricicoccus pullicaecorum*. A decreased abundance of the genus *Butyricicoccus* has been found in patients suffering from IBD [24]. In addition, supplementation of B. *pullicaecorum* resulted in an attenuation of disease in a rat colitis model [24].

Another intestinal commensal that has been suggested to be a biomarker for intestinal health based on culture-independent studies is *Akkermansia muciniphila*, a mucus-degrading bacterium that belongs to the phylum *Verrucomicrobia* [4]. Recently, the administration of *A. muciniphila* to high-fat diet-induced obese mice was shown to restore gut barrier function and increase the thickness of the mucus layer [26]. Based on these observations a probiotic potential for *A. muciniphila* has been proposed for the prevention or treatment of obesity and associated metabolic disorders.

Although none of the strains/species described in the aforementioned studies have been reportedly tested in humans or are available on the probiotic market, they do show the potential of using microbiomic insights to pinpoint specific commensal microbes that may have a beneficial effect on the health of the host. A similar approach was followed for *Romboutsia ilealis* in the work described in this thesis. As described in **Chapter 3**, our study using a rat model for acute pancreatitis revealed that the relative abundance of a specific phylotype, which was named CRIB (an abbreviation of commensal rat ileum bacterium) at the time, was correlated with severity of pancreatitis and associated sepsis. **Chapter 4** described how a representative of this phylotype was isolated, characterized, and given the name *Romboutsia ilealis* CRIB^T. Based on the findings described in **Chapter 3**, an association between a high relative abundance of *R. ilealis* and reduced duodenal bacterial overgrowth, bacterial translocation to remote organs, pancreas pathology, and proinflammatory cytokine levels in plasma was observed. Based on these findings, a probiotic potential was proposed for *R. ilealis*. To demonstrate a direct role for *R. ilealis* in attenuation of disease severity in acute pancreatic, we set out to repeat the animal experiment. This time, an additional group of animals was included that received *R. ilealis* CRIB^T instead of either placebo or the multispecies probiotic mixture that was used in the previous study. However, we were not able to finish the experiment since

the mortality rates were higher than expected due to technical problems with disease induction. This particular animal model is highly sensitive; fine-tuning of the disease inducing compounds (bile salt and cerulein) in relation to weight of the animal appeared to be an essential factor. However, preliminary results did indicate that the rats that were supplemented with either the multispecies probiotic mixture Ecologic® 641 or *R. ilealis* CRIB^T showed lower weight loss, higher survival rates, and had less bacterial translocation compared to placebo-treated animals. While potential effectiveness in acute pancreatitis was shown, direct application of *R. ilealis* CRIB^T in this population of critically ill patients is not feasible [5]. We therefore decided not to pursue this line of research any further, but rather focus on the potential probiotic properties of *R. ilealis* CRIB^T instead. As part of this investigation we performed a genomic and transcriptomic analysis of *R. ilealis* CRIB to get an overall insight in its genetic and functional potential (**Chapter 5)**. Based on these insights we were able to pinpoint specific properties that demonstrate an adaptation of *R. ilealis* CRIB^T to the small intestinal tract, such as fast uptake of simple sugars, specific genes involved in urea and bile metabolism, and potential mechanisms for competition with mucus-degrading microbes. We were able to demonstrate that *R. ilealis* CRIB^T is able to use L-fucose, a predominantly host-derived carbon source. Recently it was demonstrated that the host is able to regulate the fucosylation of its intestinal epithelial cells in response to pathogen-induced stress. This suggests that the ability of microbes to use fucose as an energy source may contribute to the protection of the host against infections by endogenous opportunistic pathogens [73].

The fact that *R. ilealis* CRIB^T is a small intestinal microbe mainly found in the terminal ileum, made it difficult to find an association between relative abundance of *R. ilealis* CRIB^T and other diseases. This is due to the sampling difficulties that are the result of the inaccessibility of this part of the intestinal tract as described in **Chapter 2** and other studies performed in our laboratory [9, 102, 103]. Using the *R. ilealis*-specific primer set that was developed as described in **Chapter 3**, we examined the abundance of *R. ilealis* in samples obtained from regions close to the small intestine, such as ileoanal pouch samples (both luminal and biopsy samples from healthy and inflamed pouches) and appendix samples (both healthy and inflamed). However, since the sample sets were small and the abundance was low, we were not able to find significant differences between patients and controls. Furthermore, none of these sample sets represented what we think would be typical small intestinal microbiota. The results from the ileostoma sample screening that was described in **Chapter 6**, provided insights in the *Romboutsia* species diversity in the human intestinal tract, and ultimately led to the isolation of a novel *Romboutsia* species, *R. hominis*, from the human small intestine.

The path from novel isolate to novel probiotic strain

It has been proposed that the ideal probiotic strain must have the following properties to

be effective and commercially attractive [30, 70]:

- Be able to survive the passage through the gastrointestinal tract (tolerant to acids, bile and digestive enzymes);
- Have a demonstrated beneficial effect on the host;
- Be safe for (human) consumption (non-pathogenic, non-toxic and free of significant adverse side effects);
- Be compatible with product matrix, processing and storage conditions to maintain the desired properties (amenable to cultivation on an industrial scale, high viability, stable on storage and oxygen tolerant).

Some of the work on *R. ilealis* CRIB^T that has not been discussed in this thesis so far, and that addressed the necessary steps needed to elucidate its (potential) probiotic properties, will be discussed in the following paragraphs. Altogether the works forms a case study for the steps to make and hurdles to overcome in the characterization of a novel isolate and marketing it as a novel probiotic strain.

Isolation source

Bacterial strains currently on the probiotic market have been isolated from a broad variety of sources: dairy and dairy-related products, fermented foods, the intestinal tract (not limited to the human intestinal tract), and even soil [30]. In the early probiotic literature you can find "be of human origin" as one of the properties of effective probiotic strains [70, 91]. Although often suggested, it is debatable whether the most effective human probiotic strains should originate from the human intestine as it only refers to from what source the particular strain was isolated, and it does not necessarily imply adaptation to this environment. Although some microbes appear to be true inhabitants of the intestinal tract, others act more like tourists who are introduced to the ecosystem via food, water, or any other substance that enters the intestinal tract and they are transient members of the intestinal microbiota [110]. Therefore, a strain isolated from human faeces could, for example, very well be derived from the fermented foods that were consumed. A recent comparative genomic and functional analysis of 100 *Lactobacillus rhamnosus* strains in comparison with the well-known probiotic strain LGG [21], showed that true intestinal inhabitants could be identified based on traits that were indicative of adaptation to an intestinal niche, in contrast to isolates that were derived from other niches [21].

As probiotic microbes have to withstand the same conditions as members of the intestinal microbiota, an intestinal origin is thought to be a preferred property of probiotic microbes. However, the impact of host specificity (i.e. adaptation to a certain host species) on probiotic effectivity is still largely unknown. Nonetheless, it has been observed that some microbes of intestinal origin do appear to show host adaptation as result of niche-specialization [32]. The different niches in the intestinal tract are determined by anatomical, immunological, and physiological characteristics of the host species [110].

Therefore, different strains may express habitat preferences that may differ between host species. For example, cell surface-associated compounds have shown to be (partially) responsible for strain specificity [11]. Therefore, it is possible that cell surface-associated ligands have evolved in response to differences in host receptors, resulting in host specificity of different microbial strains. Since probiotic effectivity of a given strain depends on the mechanism by which it exerts its probiotic activities, host specificity may not be an issue for effectiveness of each probiotic strain. For example, inhibition of pathogens by the excretion of certain antimicrobial compounds is most likely not dependent on the host species. However, specific immune modulation by ligand binding to host pattern recognition receptors is probably much dependent on the host species [13]. This issue is still largely ignored since the effectivity of probiotic strains for human application is commonly tested in animal models.

For *R. ilealis* we have attempted to isolate and characterize strains of different host species (i.e. rat, mice and human). Although we have not been able to isolate new strains belonging to the same species, we did manage to isolate a strain, named *R. hominis* FRIFI (abbreviation for first *Romboutsia* isolate from ileostoma), from the human small intestine. This isolate appeared to be a close relative, but clearly distinct novel species compared to *R. ilealis* (**Chapter 6)**. We observed the presence of multiple *Romboutsia*-like populations in rat ileum and human ileostoma effluent samples of which the microbial composition was determined by high resolution 16S rRNA gene pyrosequencing (see also **Chapter 6**). Although we focussed on obtaining only a single isolate from these samples, for future research it may be interesting to obtain multiple isolates from the same sample in order to study the microdiversity at strain-level similar to what was done for small intestinal streptococci [103]. It has been shown that examining the microdiversity of microbes can provide insight in the niche-specific adaptations [72, 102].

Gastrointestinal survival

Since probiotic microbes have to be alive according to the definition of probiotics, survival during gastrointestinal tract passage is an important criterion used in the selection of potential probiotic microbes. Especially the conditions in the stomach and the small intestine can be harsh for microbes, due to low pH in these regions and the presence of bile and digestive enzymes (as discussed in **Chapter 1**). *In vitro* models can be used to simulate these conditions and to investigate the potential survival rate of a specific microbe during gastrointestinal passage. However, it must be noted that although probiotic microbes have to be alive according to the latest definition of a true probiotic, the ability to survive gastrointestinal passage as such is insufficient to be classified as a probiotic.

For *R. ilealis* CRIB^T, two independent experiments were performed to study the gastrointestinal survival. These experiments were hampered by the absence of an efficient

enumeration system. Due to the fact that *R. ilealis* CRIB^T did not grow on culture plates after exposure to stressful conditions, there was no efficient plating system available for this anaerobe. However, in these conditions we were able to quantify the number of viable cells by growth in liquid medium. Therefore, the most-probably number (MPN) method was used as an alternative for viability counting on plates. Due to the difficulties with determining viability using culture-based approaches, we examined the possibility of using molecular methods. For this a qPCR approach was applied using the *R. ilealis*-specific primer set that was developed as described in **Chapter 3** in combination with a DNA-intercalating dye (propidium monoazide, PMA) that can be used to distinguish between live and dead cells based on integrity of the cell membranes [69]. This approach was optimized for *R. ilealis* CRIB^T, which involved determining 1) optimal time for light exposure, 2) optimal PMA incubation time, 3) optimal PMA concentration and 4) cytotoxicity of PMA for *R. ilealis* CRIB^T cells. Using this approach we were able to distinguish between live and dead cells of *R. ilealis* CRIB^T in mixtures of both cell types, confirming that in the future this approach can be used in experiments where viability has to be determined. Also for other (probiotic) strains, quantification of live and dead cells by live/dead qPCR may be a suitable alternative to the viable plate count method, although strain-specific optimization of the protocol is necessary [50]. An additional advantage of this approach is that it can be used to specifically detect viable cells of target populations in environmental samples, although there are still some issues to consider, for example with respect to dye specificity [27, 28].

As mentioned before, two independent experiments were performed to study the gastrointestinal survival of *R. ilealis* CRIB^T. Firstly, the tolerance of *R. ilealis* CRIB^T to gastric and small intestine conditions was examined under fed conditions (i.e. in the presence of a simulated food medium) similar to a set-up recently reported for *B. pullicaecorum* [35]. Using an anaerobic batch culture set-up, the viability and cultivability of *R. ilealis* CRIB^T was examined before and after exposure to low pH (pH 2 – pH 6), and after exposure to bile and pancreatic juices under fed conditions. The results of these experiments showed that there was at maximum one log decrease in cultivability of *R. ilealis* CRIB^T, suggesting that *R. ilealis* CRIB^T has a high tolerance to the simulated gastrointestinal conditions. Secondly, the tolerance of *R. ilealis* CRIB^T to gastric and small intestine conditions was examined in the absence of a simulated food medium (fasted conditions) [93]. The results showed that already at a pH below 5, survival of *R. ilealis* CRIB^T was less than 1%, suggesting that only in fed conditions *R. ilealis* CRIB^T is able to withstand low pH conditions. However, it is very difficult to compare the results from these two independent experiments since different experimental conditions were used, the sample size was small, and different methods were used to determine viability and cultivability (live/dead analysis with flow cytometry and plate counts in the first set-up, live/dead analysis with qPCR in the second set-up). It is possible that the urease identified in the genome of CRIB^T (**Chapter 5**) influences

resistance of *R. ilealis* CRIB^T to low pH conditions by a similar mechanism that protects *Helicobacter pylori* against the low pH in the stomach [58]. Expression of urease-encoding genes is only induced in specific conditions [65], which in the case of *R. ilealis* CRIB^T may be the presence of a simulated food medium. To examine this possibility, the gastrointestinal survival experiment should be repeated using both fed and fasted conditions in the same experimental set-up, and in addition to viability and cultivability, gene expression could be examined in relation to the ability of *R. ilealis* CRIB^T to tolerate gastrointestinal conditions.

Functional characterization: potential mechanisms of probiotic action

In general, probiotics are thought to act in a number of ways: 1) effects within the intestinal tract by interaction with other microbes or direct metabolic activities; 2) local effects by interactions with the mucus layer, intestinal epithelial cells, mucosal immune system, and the enteric immune system; 3) systemic effects through host interactions that go beyond the intestinal tract by signalling to other organs (e.g. the brain) or the systemic immune system [76]. The task of unravelling the underlying mechanisms of action for specific probiotic strains is daunting. Currently, we have the most profound knowledge on probiotic mechanisms of action and specific effector molecules regarding *Lactobacillus* species [12, 53, 82, 84]. However, most of the probiotic mechanisms are not yet understood at a molecular level and therefore knowledge on the gene repertoire involved is still in its infancy [12, 94]. An exception is the production of short chain fatty acids (SCFA) such as acetate, propionate, and butyrate. For these metabolic end products various mechanistic pathways have been worked out that include direct effects on the host [40, 80, 87]. Given the complexity of the intestinal microbiota and the numerous microbe-microbe and host-microbe interactions, we are most probably just beginning to grasp the potential mechanisms of probiotic action.

Regarding microbe-microbe interactions, numerous mechanisms have been described such as competition for nutrients, competition for adhesion sites, decrease of luminal pH, and production of anti-microbial compounds, as discussed in **Chapter 1**. For *R. ilealis* CRIB^T we have examined some of these mechanisms at the genomic level (**Chapter 5**). Metabolic properties of *R. ilealis* CRIB^T include the production of metabolic end products (i.e. formate, acetate, and lactate) that either directly interact with the host or can be further metabolized by other microbes. In addition, the production of these acidic metabolites results in a decrease of the pH that might may be inhibitory for other (pathogenic) microbes. Furthermore, the ability of *R. ilealis* CRIB^T to use host-derived carbohydrates may result in competition for nutrients with mucus-degrading microbes.
In addition to microbe-microbe interactions, host-microbe interactions are plentiful in the intestinal tract. For example, microbes, and thus also probiotic microbes, can influence intestinal barrier function. Possible mechanisms behind this effect are the modulation

of specific signalling pathways involved in mucus production, tight junction function, and/or apoptosis of epithelial cells [12, 101]. From the rat study described in **Chapter 3** there are indications that *R. ilealis* may have protective effects on intestinal barrier function, because increased abundance of *R. ilealis* was correlated with significantly less extra-intestinal bacterial translocation, which indicates an increased intestinal barrier function. However, so far we do not have any direct evidence that supports impact of *R. ilealis* CRIB^T on barrier function.

Host-microbe interactions also occur at the level of immune-modulation. Especially the identification of the microbial effector molecules responsible for functional modulation of host immune cells is a subject of current research. However, as mentioned before, the host-microbe interactions at a molecular level are still largely unknown [12, 53]. The observation in the rat study that relative abundance of *R. ilealis* was correlated with altered plasma cytokine levels (**Chapter 3**), suggests that this organism modulates the immune system. Based on this observation, pilot experiments were done in which T-cell differentiation was examined by co-cultures of *R. ilealis* CRIB^T with peripheral blood mononuclear cells (PBMCs) (Figure 1) [20]. Based on the T cell subset-specific transcription factor expression during co-culturing it was concluded that *R. ilealis* CRIB^T induces the differentiation of naïve T-cell into T helper 17 (T_H17) cells, and to a lesser extent T_H1 cells. Our pilot data indicate that *R. ilealis* CRIB^T is able to induce differentiation of (peripheral blood derived) CD4$^+$ precursor cells into T_H17 cells *in vitro*. The effect of *in vivo* intestinal exposure to *R. ilealis* CRIB^T on the development of mucosal T_H17 cells, as well as their functionality, remains to be established.

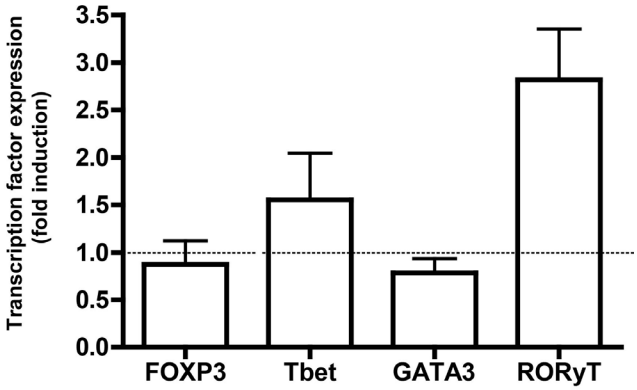

Figure 1. Modulation of T-cell differentiation by *R. ilealis* CRIB^T. PBMCs were isolated from healthy human volunteers and grown in the presence of *R. ilealis* CRIB^T (1:10 ratio). Messenger RNA (mRNA) was isolated from lysed CD4$^+$ T-cells, and the expression of transcription factors specific for CD4$^+$ T cell subsets was investigated using reverse transcriptase PCR. FOXP3, Tbet, GATA3 and RORγT are predominantly expressed by Treg, T_H1, T_H2 and T_H17 cells, respectively [6]. The relative expression of transcription factor mRNA of PMBC co-cultured with *R. ilealis* CRIB^T was compared to cells cultured in medium alone (fold induction is shown) [20].

T_H17 cells play a crucial role in host defence by control and elimination of invasive infections [60]. On the other hand, there is a proinflammatory role for T_H17 cells in inflammation and autoimmune diseases [86]. Since the initial discovery in 2005, considerable progress has been made in our understanding of T_H17 cells [3, 60, 86]. Although aberrant T_H17 responses are found to be involved in diseases such as rheumatoid arthritis, asthma, and Type-1 diabetes [86], the role of T_H17 cells in normal physiology has received little attention. Recent studies revealed that distinct microbial species can induce the accumulation of T_H17 cells, which is most likely part of the mucosal response against pathogens [7, 60]. For example, segmented filamentous bacteria have shown to be responsible for a strong induction of T_H17 cells. It has been found that T_H17 cells are specifically enriched in the small intestinal lamina propria and that an intestinal microbiota is necessary for the development of T_H17 cells [43]. Future studies will provide more insight in the role of specific microbes in maintaining a high number of T_H17 cells in the small intestinal lamina propria, which is presumably required to maintain mucosal protection against pathogens [54].

Safety

For every microbe, an unambiguous safety status is a prerequisite to be eligible as a probiotic strain. Most of the current approved probiotic strains are generally recognized as safe for human consumption, because they have been found in fermented foods that we have been consuming for years, or even centuries. The 'traditional' lactic acid bacteria used in foods and dietary supplements have a qualified presumed safety (QPS) status, which is used by authorities such as the European Food Safety Authority (EFSA) to assess safety of probiotic strains [25]. However, in the proposed use of 'non-traditional' species for probiotic applications, safety is an issue. In addition, the more recent use of intestinal isolates that are suggested to be consumed in high cell numbers by individuals with potentially compromised health, has raised the question of safety [105]. How to approach the safety evaluation of probiotic strains is still debated [78]. However, as stated in recent papers that review the safety of probiotics "there is no such thing as zero risk" [85]. From the probiotic studies that have been performed so far, it is now generally acknowledged that probiotics are not beneficial in all circumstances. In a recent review, a safety assessment was made of probiotics for human use, concluding that, considering the widespread use of probiotics, there are only a low number of documented correlations between adverse events and the consumption of probiotics [78].

For a next-generation probiotic strain it is still difficult to get guidance on what safety studies have to be performed to ensure safe application in humans, and to obtain QPS status. Considerations that have to be taken into account related to safe consumption of live microbes include their potential to produce toxins and to cause infections. At a genomic level the presence of transferrable antibiotic-resistance genes and virulence-related genes will immediately raise concerns about using such a strain as probiotic.

However, especially the presence of possible virulence factors is difficult to assess, because the properties that are virulence factors for some pathogens, are just colonization factors for other commensal microbes. For *R. ilealis* CRIB^T, we examined the presence of well-known toxin-encoding genes in the genome (**Chapter 6**). Furthermore, we examined resistance of *R. ilealis* CRIB^T to a few frequently used antibiotics (**Chapter 4**). However, a full safety assessment of *R. ilealis* CRIB^T for its use as probiotic for human consumption will comprise additional *in vitro* and *in vivo* studies.

Although there is lack of guidance, attempts are made to bring next-generation probiotic strains to the market. Recently an extensive series of experiments have been performed to determine the safety status of a new isolate, *Bacteroides xylanisolvens* DSM 23964, which included *in vitro* bacterial reverse mutation assay, *in vitro* chromosomal aberration assay, and 90 day subchronic repeated oral toxicity studies in mice, and tolerance studies in healthy humans [95-97]. Another example is *Clostridium butyricum* FERM BP-2789 that also has an extensive safety dossier, which is currently under review by EFSA for its use as a probiotic for human purposes. A product containing a similar *C. butyricum* strain, has already been available on the Japanese probiotic market for several years [83].

Product development

Initially promising strains might not make it as successful probiotic strains since they do not pass the product development process. One important aspect is the viability of the strain throughout shelf-life. One way to preserve viable bacterial cells for prolonged time is by drying the cells [15, 34]. The dried cells can be added to the food chain. Drying of microbial cells is usually done by freeze-drying or spray-drying. To this end, we have examined whether freeze-drying can potentially be used to preserve viable *R. ilealis* CRIB^T cells (Figure 2). In general, the whole freeze-drying process encompasses several steps that all have significant influence on the survival rate after freeze-drying [81]. In the first step, which includes growing the cells, the fermentation conditions during growth were shown to have a significant impact on the survival rate. This includes the addition of specific carbohydrate components to the medium [19]. In addition, the growth stage in which the cells are freeze-dried can have a major impact as well. To facilitate cell survival during drying, the cells are often formulated beforehand. The formulation forms a matrix that embeds the cells and protects them during freezing and drying from various detrimental stresses imposed on the cells. For this purpose, sugars are by far the most widely used compounds [112]. Trehalose, a well-studied cryopreservative, was used for the freeze-drying of *R. ilealis* CRIB^T. A special issue that relates to the freeze-drying of anaerobic microbes is the fact that reducing agents may have to be included to protect the cells against the oxygen they will be exposed to during the whole formulation process. However, the addition of such agents may intervene with the successful formation of an amorphous glassy matrix during the drying process, which is in general essential for reaching high

survival rates. The third and fourth steps in the process are the freezing and drying of the cells. Viability after drying can only be determined after the last step, which is rehydration. The way cells are rehydrated can have a major influence on the final survival rate.

As mentioned before, determining the survival rates for *R. ilealis* CRIB^T has been difficult since the cells did not grow on culture plates after exposure to the stressful conditions involved in the freeze-drying process. Therefore, the MPN method was used as an alternative for viability counting. Good survival rates of *R. ilealis* CRIB^T were obtained after freeze-drying (Figure 2), as only a one-log decrease was observed in the MPN counts after freeze-drying using a 10% trehalose solution supplemented with sodium sulfide combined with a rapid deep-freezing using liquid nitrogen. However, still many questions remained regarding 1) the most optimal growth phase, 2) the influence of cell density on survival, 3) the most optimal trehalose concentration, and 4) the influence of rehydration on survival. In addition, shelf-life during storage has not yet been assessed.

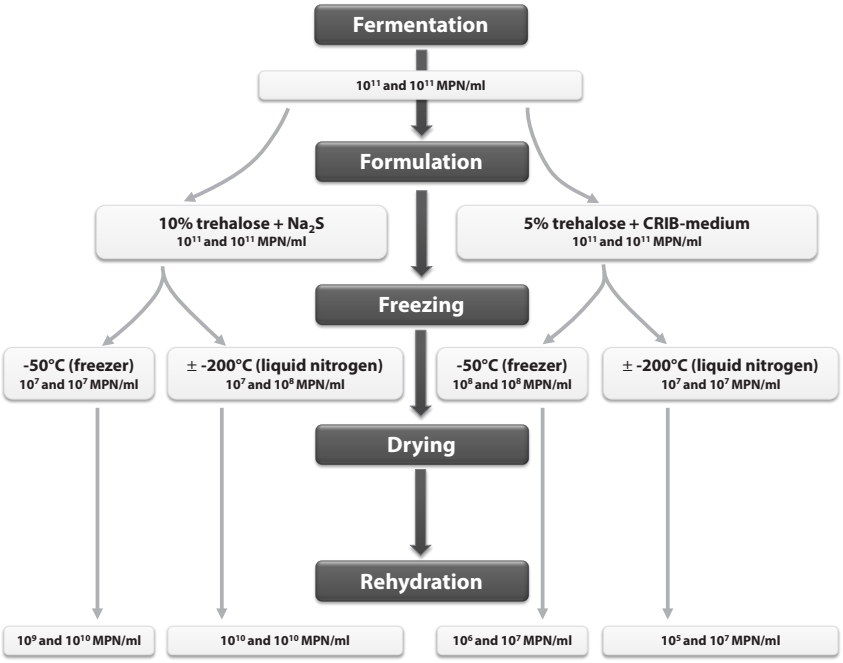

Figure 2. Overview of the set-up used for the freeze-drying of *R. ilealis* CRIB^T. Most-probably number (MPN) method was used for viability counting and the individual results of duplicate measurements are indicated. The effect of several protective agents (trehalose in combination with or without CRIB-medium, sucrose and ficoll) on the survival rate after formulation was determined, without or in combination with reducing agents (L-cysteine, sodium sulfide or titanium citrate). Of these combinations, trehalose with sodium sulfide resulted in the highest survival rates. Based on MPN/ml, we did not observe a significant difference between survival rates of cells frozen by rapid deep-freezing using liquid nitrogen (± -200°C) or cells gradually frozen by placing them in a -50°C freezer. A pilot-plant freeze-dryer (Lyostar II, FTS kinetics, NY, USA) was used for freeze-drying with a final volume of 1.0 ml. The fact that MPN counts where higher after rehydration compared to after freezing, even though the same cell concentrations were used, indicated that cells were injured during the drying process and therefore unable to grow, which was restored during the subsequent rehydration process.

Freeze-drying and subsequent storage of anaerobic bacteria thus can be challenging, as oxygen often cannot be completely excluded. For practical reasons, most of the current probiotic strains are aerobic or facultative anaerobic or can at least withstand some oxygen like, for example, many of the probiotic *Bifidobacterium* species. It is difficult to get strictly anaerobic microbes into the food-chain [90]. However, from an ecological point of view, organisms present in the intestinal tract are generally anaerobic or microaerophilic. The fact that recently *F. prausnitzii*, which is considered to be a strict anaerobe, was observed to be able to withstand oxic conditions in the presence of riboflavin or cysteine, suggests there are ways to protect anaerobic organisms from the detrimental effects of exposure to oxic conditions [45]. Micro-encapsulation of live microbial cells, by which they are immobilized in a carrier material, has been proposed as an alternative to freeze-drying [42]. Sophisticated coating techniques can protect the microbial cells against the potential detrimental effects of processing and storage. Micro-encapsulation may therefore be a solution to the current difficulties in the preparation of probiotic products containing anaerobic microbes in industrial environments [34]. An added advantage of micro-encapsulation is that a coating may ensure that viable microbes are delivered to specific regions of the intestinal tract. This provides an opportunity for potential probiotic microbes that have a low gastrointestinal tract survival by nature. Although micro-encapsulation has much potential, successful encapsulation of viable microbial cells and effective targeted delivery to the intestinal tract is still very challenging [34].

Some strains may be delivered in the form of spores, which are more thermostable by nature. The ability to form spores has been regarded advantageous for probiotic application as a way to deliver strictly anaerobic microbes in a viable condition into the gastrointestinal tract [1, 34, 79]. However, there are some issues regarding the microencapsulation of spores. For example, spores have to germinate in order to become metabolically active. It is known that only specific substances trigger germination. The right circumstances have to be present for the spores to become active at the right spot in the intestinal tract [34]. Since *R. ilealis* CRIB^T belongs to a group of spore-forming bacteria, we attempted to optimize the sporulation conditions for this organism as spores can be an alternative to freeze-drying of vegetative cells for processing and long-term storage of the organisms. Despite numerous efforts to induce sporulation, as shortly described in **Chapter 4**, efficient sporulation of *R. ilealis* CRIB^T was never achieved. Although sporulation seems to be initiated, free spores have not been observed. From the genome characterization of *R. ilealis* CRIB^T (**Chapter 5** and **Chapter 6**) we were also not able to pinpoint obvious defects in the sporulation machinery. However, sporulation is a complex process encompassing hundreds of genes that have not all been identified and characterized yet [33]. Hence unrecognized defects in the sporulation of *R. ilealis* CRIB^T may be present.

Alternative: stimulation of endogenous populations

Rather than directly administering a probiotic strain, stimulation of the endogenous population could be an alternative. There are various ways to stimulate endogenous microbes. One approach would be to administer a probiotic strain that is able to specifically stimulate the outgrowth of the target species. The mechanism behind this approach could be the production of specific metabolites that in turn can be used by others, similar to a co-metabolism mechanism recently described for *Ruminococcus bromii* [48, 49, 114]. The results from the rat study described in **Chapter 3** suggest that such an approach could be feasible, since we observed a significant increase in relative abundance of *R. ilealis* in the rats that had received a multispecies probiotic mixture. However, we have not been able to unravel the exact mechanism behind a possible stimulation of *R. ilealis* by one or more of the strains in the multispecies mixture.

A second approach to achieve stimulation of specific endogenous microbes is to administer a prebiotic. A prebiotic is defined as "a selectively fermented ingredient that results in specific changes in the composition and/or activity of the gastrointestinal microbiota, thus conferring benefits upon host health, although also the definition of this term is debated [36, 77]. This prebiotic approach is more straight-forward, since the ability of specific microbes to grow on certain prebiotics substrates can be easily examined using *in vitro* screenings. The work described in **Chapter 5** showed that *in vitro* the growth of *R. ilealis* CRIB^T can be stimulated by fructo-oligosaccharides (FOS and inulin), substances that already showed to have prebiotic effects by significant stimulation of bifidobacteria [77]. The specificity of stimulation when there is competition from other microbes in a complex ecosystem can, however, only be tested using more complex *in vitro* systems or *in vivo*. Multi-compartment *in vitro* models that simulate the dynamic physiological and physicochemical characteristics of specific regions of the intestinal tract are often used to study gut microbial activity and dynamics [56, 98, 108]. However, before it was possible to study the stimulation of *R. ilealis* CRIB^T as member of a complex community, an *in vitro* microbial community had to be established with *R. ilealis* CRIB^T as a stable member. Therefore, the ability of *R. ilealis* CRIB^T to colonize and retain itself in the presence of a complex community was studied using two independent set-ups. Firstly, the Simulator of the Human Intestinal Microbial Ecosystem (SHIME) was used. The SHIME-model consists of five connected vessels, simulating the stomach, small intestine, and the ascending, transverse, and descending colon [66]. Three times per day, SHIME feed and pancreatic juice was added to the stomach and small intestine compartments, respectively. The development of a stable microbial community in the SHIME model is achieved by introduction of faeces-derived microbes to the three colon-representing vessels [99]. Recently, the SHIME model was expanded with mucin-covered microcosms, M-SHIME, that create a mucosal environment within the *in vitro* model [100]. To study the ability of *R. ilealis* CRIB^T to establish itself in an already stable microbial intestinal

community, the M-SHIME set-up was chosen. More specifically, the first (proximal) colon vessel was studied, as the microbial communities in this vessel were thought to be most representative of a colonized ileum-like community, because the establishment of a stable ileum-like community within the vessel representing the small intestine has not yet been achieved. After faecal inoculation of the SHIME, the luminal and mucosal microbial communities were allowed to stabilize during a ten-day period. Next, stabilization was monitored for a five-day period, followed by a five-day period in which the first colon vessel was inoculated with *R. ilealis* CRIB^T on a daily basis, followed by another five-day wash-out period. The abundance of *R. ilealis* CRIB^T was followed over time by qPCR using the *R. ilealis*-specific primer set that was developed as described in **Chapter 3**. However, *R. ilealis* CRIB^T could not be detected during the inoculation nor during the wash-out period, suggesting that the organism was not able to retain itself or multiply in the proximal colon vessel. Secondly, in addition to the SHIME model, we used a colonized *in vitro* ileum model that was recently generated in analogy to the TIM-2 model of the large intestine, previously developed by TNO (Stolaki *et al.*, unpublished) [62]. This system was also initially inoculated with faecal material after which the microbial community is allowed to stabilize over a ten-day period. After this period the system was inoculated with *R. ilealis* CRIB^T and samples that were taken on a daily basis were analysed by qPCR to follow the abundance of *R. ilealis* CRIB^T in the system. However, also in this set-up, *R. ilealis* CRIB^T was not able to retain itself in the system; within hours the abundance of *R. ilealis* CRIB^T was below the detection limit of the qPCR. Even subsequent inoculations of the model with *R. ilealis* CRIB^T together with an additional carbohydrate source, on which *R. ilealis* CRIB^T is able to grow (inulin, FOS or raffinose; **Chapter 4**), did not enable *R. ilealis* CRIB^T to stay within the system. From the results of these two independent experiments it can be concluded that it is difficult, even when using advanced *in vitro* systems, to study *R. ilealis* CRIB^T in the presence of other microbes in a complex ecosystem. However, since we did not study the colonization of these systems in the absence of a complex microbial community it is not possible to say whether *R. ilealis* CRIB^T is not able to withstand the conditions in these particular systems or that it was not able to compete with the other microbes present.

As we were not able to study stimulation of R. ilealis CRIBT by prebiotic and/or probiotics using in vitro systems, we chose to perform an in vivo study with healthy human volunteers. The main difficulty of this approach was the intestinal region of interest, the small intestine. Human intervention studies targeting the small intestine and especially the last part of the small intestine, the ileum are difficult to perform since it is very challenging to obtain samples from this region [8, 103]. For this study intraluminal sampling of small intestinal fluid was done using a naso-ileal catheter as previously described [115]. The presence of several sampling ports allowed independent sampling at three locations in the small intestine (duodenum, jejunum, and ileum). Based on the

knowledge from previous studies we decided to use a synbiotic approach (probiotics in combination with a prebiotic) in an attempt to stimulate endogenous Romboutsia populations. The synbiotic used in this study consisted of a combination of a multispecies probiotic (Ecologic® 825; 1.5 × 1010 CFU/day; Winclove Probiotics, The Netherlands) and a prebiotic (FOS P06; 10 g/day; Winclove Probiotics, The Netherlands), identical to the FOS that was able to support growth of R. ilealis CRIBT in the study described in **Chapter 4**.

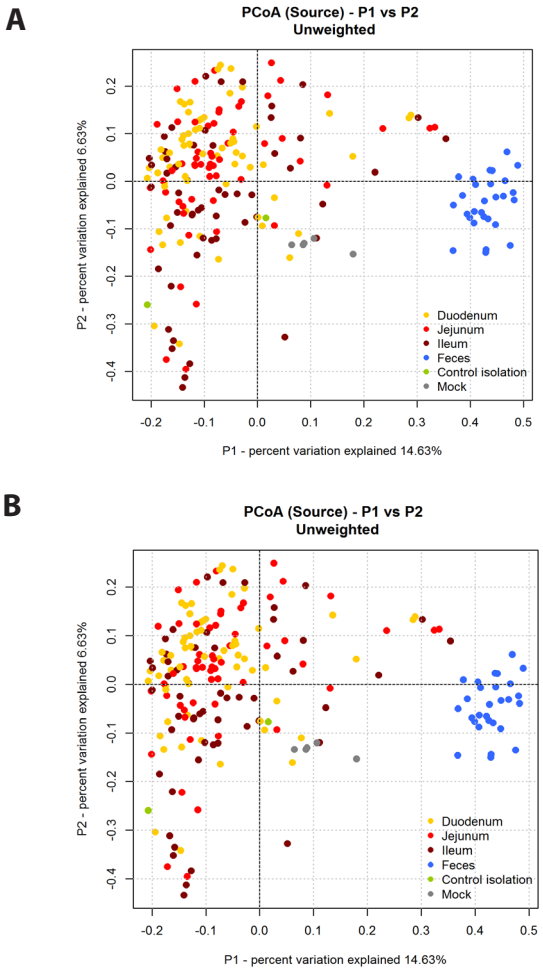

Figure 3. First two principal components (P1 and P2) from unweighted UniFrac-based PCoA analysis of the microbial community in duodenal, jejunal, ileal, and faecal samples obtained from 20 healthy subjects with colouring of the samples by (**A**) source and (**B**) intervention. MiSeq sequencing data was analysed using an adapted workflow described by Navas-Molina et al. [68]. Mock communities represent bacterial communities that were analysed alongside the intestinal samples in each amplicon pool that was sequenced in different libraries. Grouping of the mock communities underpins the reproducibility of the sequencing strategy employed here.

A randomized double-blind placebo-controlled parallel design was used to study the effect of the synbiotic on microbiota composition and activity, in addition to its effects on intestinal barrier function and immune modulation in twenty healthy human volunteers. Faecal and small intestinal fluid samples were collected before and after two weeks of synbiotic intervention. Microbiota composition was studied by 16S rRNA gene barcoded sequencing of the V1-V2 region using the Illumina MiSeq platform. Preliminary principal coordinates analysis (PCoA) of the community composition data revealed a clear and distinct grouping for small intestinal and faecal samples (Figure 3A), reflecting the differences in the overall composition along the intestinal tract as described previously [8, 115]. However, a more refined analysis did not show distinct groupings for samples obtained from different sampling regions (Figure 3A). This contrasts earlier findings that duodenal and jejunal microbial populations are quite distinct from ileal populations [115]. This could be attributable to the fact that it is not easy to determine the precise sampling site in the small intestine with the use of a catheter. Furthermore, there was no apparent grouping of the samples by treatment, i.e. placebo *vs.* synbiotic (Figure 3B). However, more in-depth analyses are necessary to examine the effects on specific microbial groups, as well as individual subject-specific trends in spatio-temporal microbiota dynamics. When considering specifically the effect on the family Peptostreptococcaceae, we observed a trend towards stimulation of this microbial group by the synbiotic in the ileum and faecal samples (Figure 4). However, more in-depth analyses are needed to evaluate whether we can identify specific species within this family that were stimulated.

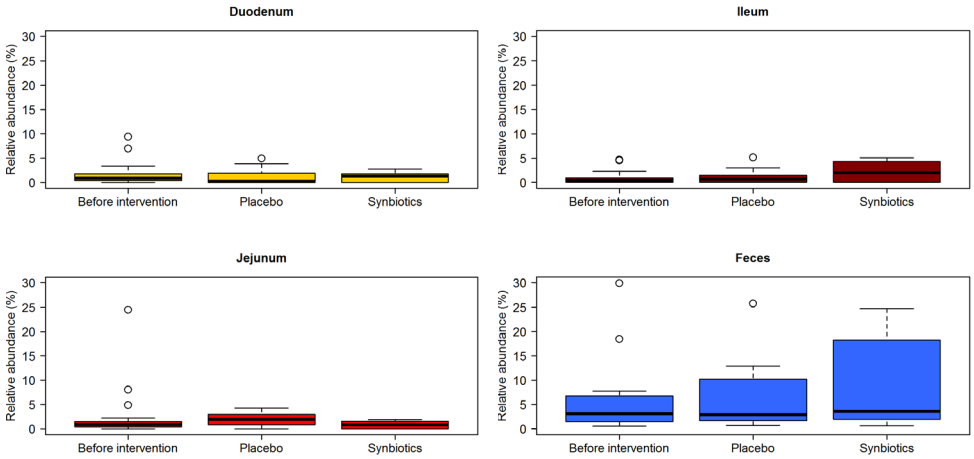

Figure 4. Relative contribution of *Peptostreptococcaceae* in the duodenum, jejunum, ileum and faecal samples obtained from 20 healthy human subjects before and after a two week supplementation period with either placebo or synbiotic, as determined by 16S rRNA gene sequencing on the Illumina MiSeq platform.

Classification of the *Peptostreptococcaceae*

The work presented in this thesis resulted in isolation and characterization of two novel species belonging to the novel genus *Romboutsia* within the family *Peptostreptococcaceae*. In **Chapter 4** the isolation and characterization of *R. ilealis* CRIB^T was described and in **Chapter 6** this was done for *R. hominis* FRIFI. Furthermore, the work described in **Chapter 4** has led to the reclassification of some species previously classified as *Clostridium*, to novel genera within the family *Peptostreptococcaceae*. This work has been a first step in resolving some of the taxonomical issues within the family *Peptostreptococcaceae*, but more reassignments will have to follow in the future. While taxonomic revisions are widely applauded by the microbial taxonomists and molecular microbiologists among us, it might be more difficult to accept in other fields of life science, since changes in names frequently lead to confusion. Therefore species and genus-level reassignments, that involve subsequent name changes should only been done after careful consideration to limit the number of name changes for a particular organism. Recently, a major reclassification was proposed for 78 *Clostridium* species to 6 new genera based on whole genome comparisons [113]. For example, *Clostridium difficile* and 16 close relatives (all *Clostridium* cluster XI species), were proposed to be renamed as *Peptoclostridium*, a novel genus within the family *Peptostreptococcaceae*. Although the attempt was admirable, one can argue that the grouping proposed by the authors does not justify genus-level groups, and therefore these name changes may not be valid. However, these names changes are already appearing in the sequence databases and therefore cause confusion, especially because the taxonomist community has not formally accepted them.

In **Chapter 6** the strength of a genomic approach in characterizing a novel species has been demonstrated. However, a genomic approach also has its limitations, and bacterial classification based on genomic information should be done with care. Recently, many proposals for novel species based on whole genome sequence descriptions are published, but most of the descriptions lack a solid taxonomic evaluation and by wrongfully assigning names to novel species they add to the taxonomic confusion. Recent examples are the incorrectly proposed species '*Clostridium dakarense*' and '*Clostridium jeddahense*' [51, 55, 64] that have only been given the name *Clostridium* since their most closely related species were given that name. For '*Clostridium dakarense*' we have shown that assignment to the genus *Clostridium* was incorrect (**Chapter 6**). The fact that genomic and taxonomic characterization of a novel species can go hand in hand is nicely illustrated by the proposal of the novel species *Clostridium bornimense* that was characterized from both taxonomic and genomic points of view [37, 38].

Species concept in a genomic era

Sequencing technologies have considerably advanced over the past few years and it is nowadays more than affordable to sequence complete microbial genomes. The first

reported bacterial genome was that of *Haemophilus influenza* Rd, which was sequenced in 1995 [29]. Now almost twenty years later there are thousands of bacterial genomes available in public databases. One might think that the simplest method for classification of bacteria would be to compare their genome sequences. Whole genome-based trees are thought to have greater phylogenetic resolution than single-gene topologies such as 16S rRNA gene sequence-based trees [18]. However, the difficulties of classifying bacteria based on whole genome sequences become apparent when one realizes that bacterial genomes are much more dynamic and diverse than previously assumed. Mutations are inevitable and therefore genotypic differences between bacterial cells potentially arise at every cell division [2, 22]. As a result of selection, some of these mutations can persist in a population. This phenomenon will also occur in eukaryotic species, but because (most) bacteria can multiply faster than eukaryotic cells and have haploid genomes, evolution takes place at a much faster speed [2, 14]. A bigger challenge in modern-day taxonomy, however, relates to the recent discovery of extensive horizontal gene transfer between organisms [31, 52]. Horizontal gene transfer, also known as lateral gene transfer, is the exchange of genetic material between individual cells of different species. With the increasing number of available bacterial genomes, the number of examples of horizontal gene transfer between individuals of different species is growing rapidly. Especially in the intestinal tract, an environment under constant external pressure and with high density of microbial cells, horizontal gene transfer occurs frequently. Antibiotic resistance genes, for example, are spread among various species via transport on plasmids, small DNA molecules that are physically separated from the chromosomal DNA within a cell [10, 41, 106]. These processes have shown to be important phenomena in the evolution of many microbes. However, they pose a significant dilemma for taxonomists since they have to deal with the fact that different genes within the same organism can have different phylogenetic origins. As discussed in **Chapter 1**, taxonomy is all about identifying and characterizing species, not specimens (strains). The current difficulty is to know the species variability and to separate what is intra- and inter-species variability. Some species are highly clonal, but for others it has been described that strains of the same species can differ in up to 30% of the genes in their genomes [47]. A well-known example is *E. coli* H0157, responsible for a foodborne illness due to contaminated meat called Hamburger disease. Its genome was determined to be 859 Kb larger than the intestinal isolate *E. coli* K12 [39]. Do these strains still belong to the same species, or should they be separated in multiple species? And what to do with co-existing phylogenetically closely related populations that are physiologically quite distinct [67]? This discussion constitutes a conceptual dilemma, but it actually starts to matter when describing ecosystem composition and talking about the number of species present, their diversity and ecosystem stability.

The current bacterial classification is shaped by historic reasoning. However, it is generally acknowledged that a classification should rest on the highest quality of data

available at the time, and a classification should be adjusted accordingly when these data change. Therefore, old taxonomic assignments need to be adjusted in order to get to a system that is (more) predictive of the genetic relatedness of the grouped taxa. In addition, as long as we are studying life, novel organisms will be discovered, which will necessitate refinement of classifications. The enormous amount of information that is produced on a daily basis regarding the chemical and genetic composition of bacterial cells, allows for a more sophisticated approach of bacterial taxonomy and phylogeny. How genomics can be integrated in the taxonomy and systematics of prokaryotes is still up for discussion [16, 17, 44, 46, 61, 75]. Acknowledgement of the extensive impact of horizontal gene transfer on bacterial evolution may force us to give up on a hierarchical classification system for microbial life [71]. The only way the field of microbiology will truly advance is when we accept that taxonomy will always be on the move.

Concluding remarks and future perspectives

The work presented in this thesis describes the various stages of a journey that started with the discovery of a new bacterial species in the ileum of a rat and has led to an in-depth characterization of a novel genus, *Romboutsia*, which is present within the intestinal tract of humans and other mammals. By the isolation and characterization of *Romboutsia* species of intestinal origin, we have gained an initial insight into the genetic diversity within the genus *Romboutsia* and the adaptation of *Romboutsia* species to a life in the intestinal tract. Based on sequence-based community analyses, it has become clear that the diversity within the genus is not limited to the species isolated and/or characterized in this thesis. Related strains have already been isolated by others, but they have never been (validly) described and deposited, partially due to the taxonomic issues associated with this group of microbes. Characterization of the genus *Romboutsia* has been an endurance run, since it involved taxonomic reclassification of a number of closely related species, a labour-intensive and time-consuming task. However, now that the genus has been described and defined, it will enable others to add new species. Future isolation and characterization of additional *Romboutsia* species and strains will provide more detailed information regarding the metabolic and functional properties of these microbes. Furthermore, the isolation of additional strains of different origins will help to unravel the mechanisms by which these species have adapted to a life in a specific niche within the host, or provide insight in host specialization.

Based on the work that was done on *R. ilealis* CRIB^T specifically, it is to date not possible to come to a definitive conclusion on whether the differences that were observed in the rat study described in **Chapter 2** were the direct effect of health-beneficial properties of *R. ilealis*. It may still be possible that the relative abundance of *R. ilealis* was a biomarker for health, but the possibility that *R. ilealis* was just a bystander in the observed health-promoting effects by the multispecies probiotic mixture, should also not be ruled out. As

we know that different strains may exert different effects based on specific capabilities and enzymatic activities, even within one species, it might very well be possible that the collection of additional and/or other *R. ilealis* isolates during our journey would have led us elsewhere. This thesis shows the potential for newly identified strains to exert health-promoting effects with the ultimate possibility for them to be classified as probiotics. Although the probiotic market is still largely dominated by *Lactobacillus* and *Bifidobacterium* species, the number of commensal microbes to which potential beneficial effects have been attributed is growing. Further integration of culture-dependent and culture-dependent studies will increase our knowledge on the microbial contribution to host's health and disease. As discussed, these insights can be used to identify specific microbial species or even strains that may be beneficial to certain target populations. Although current approaches to identify next-generation probiotic strains are largely driven by the search for compositional differences, an even more effective approach would be to focus on functional differences. Microbiomic approaches can be used to identify functional genes that regulate health and disease, which in turn can be used to select potential probiotic strains based on specific functional properties. Currently, probiotic strains are screened for probiotic properties using *in vitro* models, but it is still difficult to translate *in vitro* findings into significant clinical effectiveness as the environment controls the expression of certain functionalities. Elucidation of the molecular mechanisms behind probiotic mechanisms will enable future selection of new probiotic strains based on a combination of specific properties. In addition, in the search for new probiotic strains one should not forget that the intestinal tract consist of various regions each containing its own microbial community adapted to that niche. Depending on the mechanism(s) of action, the search for more effective probiotic strains should be extended to currently largely unexplored regions of the intestinal tract, such as the small intestine.

References

1. Bader J., A. Albin and U. Stahl (2012). Spore-forming bacteria and their utilisation as probiotics. *Benef Microbes*; 3 (1): 67-75.
2. Barrick J.E. and R.E. Lenski (2013). Genome dynamics during experimental evolution. *Nat Rev Genet*; 14 (12): 827-839.
3. Basso A.S., H. Cheroutre and D. Mucida (2009). More stories on Th17 cells. *Cell Res*; 19 (4): 399-411.
4. Belzer C. and W.M. de Vos (2012). Microbes inside - from diversity to function: the case of *Akkermansia*. *ISME J*; 6 (8): 1449-1458.
5. Besselink M.G., H.C. van Santvoort, E. Buskens, M.A. Boermeester, H. van Goor, H.M. Timmerman, V.B. Nieuwenhuijs, *et al.* (2008). Probiotic prophylaxis in predicted severe acute pancreatitis: a randomised, double-blind, placebo-controlled trial. *Lancet*; 371 (9613): 651-659.
6. Bettelli E., M. Oukka and V.K. Kuchroo (2007). T(H)-17 cells in the circle of immunity and autoimmunity. *Nat Immunol*; 8 (4): 345-350.
7. Blaschitz C. and M. Raffatellu (2010). Th17 cytokines and the gut mucosal barrier. *J Clin Immunol*; 30 (2): 196-203.
8. Booijink C.C., S. El-Aidy, M. Rajilic-Stojanovic, H.G. Heilig, F.J. Troost, H. Smidt, M. Kleerebezem, *et al.* (2010). High temporal and inter-individual variation detected in the human ileal microbiota. *Environ Microbiol*; 12 (12): 3213-3227.
9. Booijink C.C., E.G. Zoetendal, M. Kleerebezem and W.M. de Vos (2007). Microbial communities in the human small intestine: coupling diversity to metagenomics. *Future Microbiol*; 2 (3): 285-295.
10. Broaders E., C.G. Gahan and J.R. Marchesi (2013). Mobile genetic elements of the human gastrointestinal tract: potential for spread of antibiotic resistance genes. *Gut Microbes*; 4 (4): 271-280.
11. Bron P.A., S. Tomita, A. Mercenier and M. Kleerebezem (2013). Cell surface-associated compounds of probiotic lactobacilli sustain the strain-specificity dogma. *Curr Opin Microbiol*; 16 (3): 262-269.
12. Bron P.A., P. van Baarlen and M. Kleerebezem (2012). Emerging molecular insights into the interaction between probiotics and the host intestinal mucosa. *Nat Rev Microbiol*; 10 (1): 66-78.
13. Bryant C.E. and T.P. Monie (2012). Mice, men and the relatives: cross-species studies underpin innate immunity. *Open Biol*; 2 (4): 120015.
14. Burke M.K. (2012). How does adaptation sweep through the genome? Insights from long-term selection experiments. *Proc Biol Sci*; 279 (1749): 5029-5038.
15. Chávez B.E. and A.M. Ledeboer (2007). Drying of Probiotics: Optimization of Formulation and Process to Enhance Storage Survival. *Drying Technology*; 25 (7-8): 1193-1201.
16. Chun J. and F.A. Rainey (2014). Integrating genomics into the taxonomy and systematics of the *Bacteria* and *Archaea*. *Int J Syst Evol Microbiol*; 64 (Pt 2): 316-324.
17. Coenye T., D. Gevers, Y. Van de Peer, P. Vandamme and J. Swings (2005). Towards a prokaryotic genomic taxonomy. *FEMS Microbiol Rev*; 29 (2): 147-167.
18. Coenye T. and P. Vandamme (2003). Extracting phylogenetic information from whole-genome sequencing projects: the lactic acid bacteria as a test case. *Microbiology*; 149 (Pt 12): 3507-3517.
19. Corcoran B.M., R.P. Ross, G.F. Fitzgerald and C. Stanton (2004). Comparative survival of probiotic lactobacilli spray-dried in the presence of prebiotic substances. *J Appl Microbiol*; 96 (5): 1024-1039.
20. de Roock S., M. van Elk, M.O. Hoekstra, B.J. Prakken, G.T. Rijkers and I.M. de Kleer (2011). Gut derived lactic acid bacteria induce strain specific CD4(+) T cell responses in human PBMC. *Clin Nutr*; 30 (6): 845-851.
21. Douillard F.P., A. Ribbera, R. Kant, T.E. Pietila, H.M. Jarvinen, M. Messing, C.L. Randazzo, *et al.* (2013). Comparative genomic and functional analysis of 100 Lactobacillus rhamnosus strains and their comparison with strain GG. *PLoS Genet*; 9 (8): e1003683.
22. Drake J.W. (1991). A constant rate of spontaneous mutation in DNA-based microbes. *Proc Natl Acad Sci USA*; 88 (16): 7160-7164.
23. Dunne C., L. O'Mahony, L. Murphy, G. Thornton, D. Morrissey, S. O'Halloran, M. Feeney, *et al.* (2001). *In vitro* selection criteria for probiotic bacteria of human origin: correlation with *in vivo* findings. *Am J Clin Nutr*; 73 (2 Suppl): 386S-392S.
24. Eeckhaut V., K. Machiels, C. Perrier, C. Romero, S. Maes, B. Flahou, M. Steppe, *et al.* (2013). *Butyricicoccus pullicaecorum* in inflammatory bowel disease. *Gut*; 62 (12): 1745-1752.
25. EFSA (2013). Scientific Opinion on the maintenance of the list of QPS biological agents intentionally added to food and feed (2013 update). *The EFSA Journal*; 11 (11): 3349.
26. Everard A., C. Belzer, L. Geurts, J.P. Ouwerkerk, C. Druart, L.B. Bindels, Y. Guiot, *et al.* (2013). Cross-talk between

Akkermansia muciniphila and intestinal epithelium controls diet-induced obesity. *Proc Natl Acad Sci USA*; 110 (22): 9066-9071.

27. Fittipaldi M., F. Codony, B. Adrados, A.K. Camper and J. Morato (2011). Viable real-time PCR in environmental samples: can all data be interpreted directly? *Microb Ecol*; 61 (1): 7-12.
28. Fittipaldi M., A. Nocker and F. Codony (2012). Progress in understanding preferential detection of live cells using viability dyes in combination with DNA amplification. *J Microbiol Methods*; 91 (2): 276-289.
29. Fleischmann R.D., M.D. Adams, O. White, R.A. Clayton, E.F. Kirkness, A.R. Kerlavage, C.J. Bult, *et al.* (1995). Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd. *Science*; 269 (5223): 496-512.
30. Fontana L., M. Bermudez-Brito, J. Plaza-Diaz, S. Munoz-Quezada and A. Gil (2013). Sources, isolation, characterisation and evaluation of probiotics. *Br J Nutr*; 109 Suppl 2: S35-50.
31. Fraser C., W.P. Hanage and B.G. Spratt (2007). Recombination and the nature of bacterial speciation. *Science*; 315 (5811): 476-480.
32. Frese S.A., A.K. Benson, G.W. Tannock, D.M. Loach, J. Kim, M. Zhang, P.L. Oh, *et al.* (2011). The evolution of host specialization in the vertebrate gut symbiont *Lactobacillus reuteri*. *PLoS Genet*; 7 (2): e1001314.
33. Galperin M.Y., S.L. Mekhedov, P. Puigbo, S. Smirnov, Y.I. Wolf and D.J. Rigden (2012). Genomic determinants of sporulation in Bacilli and Clostridia: towards the minimal set of sporulation-specific genes. *Environ Microbiol*; 14 (11): 2870-2890.
34. Gawkowski D. and M.L. Chikindas (2013). Non-dairy probiotic beverages: the next step into human health. *Benef Microbes*; 4 (2): 127-142.
35. Geirnaert A., A. Steyaert, V. Eeckhaut, B. Debruyne, J.B. Arends, F. Van Immerseel, N. Boon, *et al.* (2014). *Butyricicoccus pullicaecorum*, a butyrate producer with probiotic potential, is intrinsically tolerant to stomach and small intestine conditions. *Anaerobe*; 30C: 70-74.
36. Gibson G.R., K.P. Scott and R.A. Rastall (2010). Dietary prebiotics: current status and new definition. *Food Science and Technology Bulletin: Functional Foods*; 7: 1-19.
37. Hahnke S., J. Striesow, M. Elvert, X.P. Mollar and M. Klocke (2014). *Clostridium bornimense* sp. nov., isolated from a mesophilic, two-phase, laboratory-scale biogas reactor. *Int J Syst Evol Microbiol*; 64 (Pt 8): 2792-2797.
38. Hahnke S., D. Wibberg, G. Tomazetto, A. Puhler, M. Klocke and A. Schluter (2014). Whole genome sequence of *Clostridium bornimense* strain M2/40 isolated from a lab-scale mesophilic two-phase biogas reactor digesting maize silage and wheat straw. *J Biotechnol*; 184: 199-200.
39. Hayashi T., K. Makino, M. Ohnishi, K. Kurokawa, K. Ishii, K. Yokoyama, C.G. Han, *et al.* (2001). Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genomic comparison with a laboratory strain K-12. *DNA Res*; 8 (1): 11-22.
40. Hong Y.H., Y. Nishimura, D. Hishikawa, H. Tsuzuki, H. Miyahara, C. Gotoh, K.C. Choi, *et al.* (2005). Acetate and propionate short chain fatty acids stimulate adipogenesis via GPCR43. *Endocrinology*; 146 (12): 5092-5099.
41. Huddleston J.R. (2014). Horizontal gene transfer in the human gastrointestinal tract: potential spread of antibiotic resistance genes. *Infect Drug Resist*; 7: 167-176.
42. Islam M.A., C.H. Yun, Y.J. Choi and C.S. Cho (2010). Microencapsulation of live probiotic bacteria. *J Microbiol Biotechnol*; 20 (10): 1367-1377.
43. Ivanov, II, L. Frutos Rde, N. Manel, K. Yoshinaga, D.B. Rifkin, R.B. Sartor, B.B. Finlay, *et al.* (2008). Specific microbiota direct the differentiation of IL-17-producing T-helper cells in the mucosa of the small intestine. *Cell Host Microbe*; 4 (4): 337-349.
44. Kämpfer P. and S.P. Glaeser (2012). Prokaryotic taxonomy in the sequencing era - the polyphasic approach revisited. *Environ Microbiol*; 14 (2): 291-317.
45. Khan M.T., J.M. van Dijl and H.J. Harmsen (2014). Antioxidants keep the potentially probiotic but highly oxygen-sensitive human gut bacterium Faecalibacterium prausnitzii alive at ambient air. *PLoS One*; 9 (5): e96097.
46. Konstantinidis K. and E. Stackebrandt (2013). Defining Taxonomic Ranks. In: *The Prokaryotes*. pp 229-254. Edited by E. Rosenberg, E. DeLong, S. Lory, E. Stackebrandt, *et al.*: Springer Berlin Heidelberg.
47. Konstantinidis K.T. and J.M. Tiedje (2005). Genomic insights that advance the species definition for prokaryotes. *Proc Natl Acad Sci USA*; 102 (7): 2567-2572.
48. Kovatcheva-Datchary P. (2010).Analyzing the functionality of the human intestinal microbiota by stable isotope probing. PhD thesis. Wageningen University.
49. Kovatcheva-Datchary P., M. Egert, A. Maathuis, M. Rajilic-Stojanovic, A.A. de Graaf, H. Smidt, W.M. de Vos, *et al.* (2009). Linking phylogenetic identities of bacteria to starch fermentation in an in vitro model of the large intestine by RNA-based stable isotope probing. *Environ Microbiol*; 11 (4): 914-926.
50. Kramer M., N. Obermajer, B. Bogovic Matijasic, I. Rogelj and V. Kmetec (2009). Quantification of live and dead

probiotic bacteria in lyophilised product by real-time PCR and by flow cytometry. *Appl Microbiol Biotechnol*; 84 (6): 1137-1147.

51. Lagier J.C., F. Bibi, D. Ramasamy, E.I. Azhar, C. Robert, M. Yasir, A.A. Jiman-Fatani*, et al.* (2014). Non contiguous-finished genome sequence and description of *Clostridium jeddahense* sp. nov. *Stand Genomic Sci*; 9 (3): 1003-1019.
52. Lawrence J.G. and A.C. Retchless (2009). The interplay of homologous recombination and horizontal gene transfer in bacterial speciation. *Methods Mol Biol*; 532: 29-53.
53. Lee I.C., S. Tomita, M. Kleerebezem and P.A. Bron (2013). The quest for probiotic effector molecules - unraveling strain specificity at the molecular level. *Pharmacol Res*; 69 (1): 61-74.
54. Lee Y.K. and S.K. Mazmanian (2014). Microbial learning lessons: SFB educate the immune system. *Immunity*; 40 (4): 457-459.
55. Lo C.I., A.K. Mishra, R. Padhmanabhan, B. Samb, A.G. Sow, C. Robert, C. Couderc*, et al.* (2013). Non-contiguous finished genome sequence and description of *Clostridium dakarense* sp. nov. *Stand Genomic Sci*; 9 (1): 14-27.
56. Macfarlane G.T. and S. Macfarlane (2007). Models for intestinal fermentation: association between food components, delivery systems, bioavailability and functional interactions in the gut. *Curr Opin Biotechnol*; 18 (2): 156-162.
57. Machiels K., M. Joossens, J. Sabino, V. De Preter, I. Arijs, V. Eeckhaut, V. Ballet*, et al.* (2014). A decrease of the butyrate-producing species *Roseburia hominis* and *Faecalibacterium prausnitzii* defines dysbiosis in patients with ulcerative colitis. *Gut*; 63 (8): 1275-1283.
58. Marshall B.J., L.J. Barrett, C. Prakash, R.W. McCallum and R.L. Guerrant (1990). Urea protects *Helicobacter* (*Campylobacter*) *pylori* from the bactericidal effect of acid. *Gastroenterology*; 99 (3): 697-702.
59. Marteau P. (2013). Butyrate-producing bacteria as pharmabiotics for inflammatory bowel disease. *Gut*; 62 (12): 1673.
60. McGeachy M.J. and S.J. McSorley (2012). Microbial-induced Th17: superhero or supervillain? *J Immunol*; 189 (7): 3285-3291.
61. Meier-Kolthoff J.P., H.P. Klenk and M. Goker (2014). Taxonomic use of DNA G+C content and DNA-DNA hybridization in the genomic age. *Int J Syst Evol Microbiol*; 64 (2): 352-356.
62. Minekus M., M. Smeets-Peeters, A. Bernalier, S. Marol-Bonnin, R. Havenaar, P. Marteau, M. Alric*, et al.* (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 (1): 108-114.
63. Miquel S., R. Martin, O. Rossi, L.G. Bermudez-Humaran, J.M. Chatel, H. Sokol, M. Thomas*, et al.* (2013). *Faecalibacterium prausnitzii* and human intestinal health. *Curr Opin Microbiol*; 16 (3): 255-261.
64. Mishra A.K., J.C. Lagier, C. Robert, D. Raoult and P.E. Fournier (2012). Non-contiguous finished genome sequence and description of *Clostridium senegalense* sp. nov. *Stand Genomic Sci*; 6 (3): 386-395.
65. Mobley H.L., M.D. Island and R.P. Hausinger (1995). Molecular biology of microbial ureases. *Microbiol Rev*; 59 (3): 451-480.
66. Molly K., M. Vande Woestyne and W. Verstraete (1993). Development of a 5-step multi-chamber reactor as a simulation of the human intestinal microbial ecosystem. *Appl Microbiol Biotechnol*; 39 (2): 254-258.
67. Moore L.R., G. Rocap and S.W. Chisholm (1998). Physiology and molecular phylogeny of coexisting *Prochlorococcus* ecotypes. *Nature*; 393 (6684): 464-467.
68. Navas-Molina J.A., J.M. Peralta-Sánchez, A. González, P.J. McMurdie, Y. Vázquez-Baeza, Z. Xu, L.K. Ursell*, et al.* (2013). Chapter Nineteen - Advancing Our Understanding of the Human Microbiome Using QIIME. In: *Methods in Enzymology*. Volume 531, pp 371-444. Edited by: Academic Press.
69. Nocker A., P. Sossa-Fernandez, M.D. Burr and A.K. Camper (2007). Use of propidium monoazide for live/dead distinction in microbial ecology. *Appl Environ Microbiol*; 73 (16): 5111-5117.
70. Ouwehand A.C., S. Salminen and E. Isolauri (2002). Probiotics: an overview of beneficial effects. *Antonie Van Leeuwenhoek*; 82 (1-4): 279-289.
71. Pace N.R., J. Sapp and N. Goldenfeld (2012). Phylogeny and beyond: scientific, historical, and conceptual significance of the first tree of life. *Proc Natl Acad Sci USA*; 109 (4): 1011-1018.
72. Pena A., H. Teeling, J. Huerta-Cepas, F. Santos, P. Yarza, J. Brito-Echeverria, M. Lucio*, et al.* (2010). Fine-scale evolution: genomic, phenotypic and ecological differentiation in two coexisting *Salinibacter* ruber strains. *ISME J*; 4 (7): 882-895.
73. Pickard J.M., C.F. Maurice, M.A. Kinnebrew, M.C. Abt, D. Schenten, T.V. Golovkina, S.R. Bogatyrev*, et al.* (2014). Rapid fucosylation of intestinal epithelium sustains host-commensal symbiosis in sickness. *Nature*.
74. Qin J., R. Li, J. Raes, M. Arumugam, K.S. Burgdorf, C. Manichanh, T. Nielsen*, et al.* (2010). A human gut microbial gene catalogue established by metagenomic sequencing. *Nature*; 464 (7285): 59-65.
75. Ramasamy D., A.K. Mishra, J.C. Lagier, R. Padhmanabhan, M. Rossi, E. Sentausa, D. Raoult*, et al.* (2014). A polyphasic

strategy incorporating genomic data for the taxonomic description of novel bacterial species. *Int J Syst Evol Microbiol*; 64 (Pt 2): 384-391.

76. Rijkers G.T., S. Bengmark, P. Enck, D. Haller, U. Herz, M. Kalliomaki, S. Kudo, *et al.* (2010). Guidance for substantiating the evidence for beneficial effects of probiotics: current status and recommendations for future research. *J Nutr*; 140 (3): 671S-676S.
77. Roberfroid M., G.R. Gibson, L. Hoyles, A.L. McCartney, R. Rastall, I. Rowland, D. Wolvers, *et al.* (2010). Prebiotic effects: metabolic and health benefits. *Br J Nutr*; 104 Suppl 2: S1-63.
78. Sanders M.E., L.M. Akkermans, D. Haller, C. Hammerman, J. Heimbach, G. Hormannsperger, G. Huys, *et al.* (2010). Safety assessment of probiotics for human use. *Gut Microbes*; 1 (3): 164-185.
79. Sanders M.E., L. Morelli and T.A. Tompkins (2003). Sporeformers as human probiotics: *Bacillus*, *Sporolactobacillus*, and *Brevibacillus*. *Compr Rev Food Sci F*; 2 (3): 101-110.
80. Scheppach W. (1994). Effects of short chain fatty acids on gut morphology and function. *Gut*; 35 (1 Suppl): S35-38.
81. Schoug Å. (2009).A dry phase of life: Freeze-drying and storage stability of *Lactobacillus coryniformis* Si3 in sucrose-based formulations. PhD thesis. Uppsala: Swedish University of Agricultural Sciences.
82. Segers M.E. and S. Lebeer (2014). Towards a better understanding of Lactobacillus rhamnosus GG - host interactions. *Microb Cell Fact*; 13 Suppl 1: S7.
83. Seki H., M. Shiohara, T. Matsumura, N. Miyagawa, M. Tanaka, A. Komiyama and S. Kurata (2003). Prevention of antibiotic-associated diarrhea in children by *Clostridium butyricum* MIYAIRI. *Pediatr Int*; 45 (1): 86-90.
84. Sengupta R., E. Altermann, R.C. Anderson, W.C. McNabb, P.J. Moughan and N.C. Roy (2013). The role of cell surface architecture of lactobacilli in host-microbe interactions in the gastrointestinal tract. *Mediators Inflamm*; 2013: 237921.
85. Shanahan F. (2012). A commentary on the safety of probiotics. *Gastroenterol Clinics North Am*; 41 (4): 869-876.
86. Singh R.P., S. Hasan, S. Sharma, S. Nagra, D.T. Yamaguchi, D. Wong, H. Bh, *et al.* (2014). Th17 cells in inflammation and autoimmunity. *Autoimmun Rev*; 13 (12): 1174-1181.
87. Smith P.M., M.R. Howitt, N. Panikov, M. Michaud, C.A. Gallini, Y.M. Bohlooly, J.N. Glickman, *et al.* (2013). The microbial metabolites, short-chain fatty acids, regulate colonic Treg cell homeostasis. *Science*; 341 (6145): 569-573.
88. Sokol H., B. Pigneur, L. Watterlot, O. Lakhdari, L.G. Bermudez-Humaran, J.J. Gratadoux, S. Blugeon, *et al.* (2008). *Faecalibacterium prausnitzii* is an anti-inflammatory commensal bacterium identified by gut microbiota analysis of Crohn disease patients. *Proc Natl Acad Sci USA*; 105 (43): 16731-16736.
89. Sokol H., P. Seksik, J.P. Furet, O. Firmesse, I. Nion-Larmurier, L. Beaugerie, J. Cosnes, *et al.* (2009). Low counts of *Faecalibacterium prausnitzii* in colitis microbiota. *Inflamm Bowel Dis*; 15 (8): 1183-1189.
90. Talwalkar A. and K. Kailasapathy (2004). The role of oxygen in the viability of probiotic bacteria with reference to *L. acidophilus* and *Bifidobacterium* spp. *Curr Issues Intest Microbiol*; 5 (1): 1-8.
91. Teitelbaum J.E. and W.A. Walker (2002). Nutritional impact of pre- and probiotics as protective gastrointestinal organisms. *Ann Rev Nutr*; 22: 107-138.
92. Tilg H. and S. Danese (2014). *Roseburia hominis*: a novel guilty player in ulcerative colitis pathogenesis? *Gut*; 63 (8): 1204-1205.
93. Timmerman H.M., L.E. Niers, B.U. Ridwan, C.J. Koning, L. Mulder, L.M. Akkermans, F.M. Rombouts, *et al.* (2007). Design of a multispecies probiotic mixture to prevent infectious complications in critically ill patients. *Clin Nutr*; 26 (4): 450-459.
94. Turroni F., M. Ventura, L.F. Butto, S. Duranti, P.W. O'Toole, M.O. Motherway and D. van Sinderen (2014). Molecular dialogue between the human gut microbiota and the host: a *Lactobacillus* and *Bifidobacterium* perspective. *Cell Mol Life Sci: CMLS*; 71 (2): 183-203.
95. Ulsemer P., K. Toutounian, G. Kressel, J. Schmidt, U. Karsten, A. Hahn and S. Goletz (2012). Safety and tolerance of *Bacteroides xylanisolvens* DSM 23964 in healthy adults. *Benef Microbes*; 3 (2): 99-111.
96. Ulsemer P., K. Toutounian, J. Schmidt, U. Karsten and S. Goletz (2012). Preliminary safety evaluation of a new *Bacteroides xylanisolvens* isolate. *Appl Environ Microbiol*; 78 (2): 528-535.
97. Ulsemer P., K. Toutounian, J. Schmidt, J. Leuschner, U. Karsten and S. Goletz (2012). Safety assessment of the commensal strain *Bacteroides xylanisolvens* DSM 23964. *Regul Toxicol Pharmacol: RTP*; 62 (2): 336-346.
98. Van de Wiele T.R., A.G. Oomen, J. Wragg, M. Cave, M. Minekus, A. Hack, C. Cornelis, *et al.* (2007). Comparison of five *in vitro* digestion models to *in vivo* experimental results: lead bioaccessibility in the human gastrointestinal tract. *J Environ Sci Health A Tox Hazard Subst Environ Eng*; 42 (9): 1203-1211.
99. Van den Abbeele P., C. Grootaert, M. Marzorati, S. Possemiers, W. Verstraete, P. Gerard, S. Rabot, *et al.* (2010). Microbial community development in a dynamic gut model is reproducible, colon region specific, and selective for *Bacteroidetes* and *Clostridium cluster* IX. *Appl Environ Microbiol*; 76 (15): 5237-5246.

100. Van den Abbeele P., S. Roos, V. Eeckhaut, D.A. MacKenzie, M. Derde, W. Verstraete, M. Marzorati, *et al.* (2012). Incorporating a mucosal environment in a dynamic gut model results in a more representative colonization by lactobacilli. *Microb Biotechnol*; 5 (1): 106-115.
101. Van den Abbeele P., T. Van de Wiele, W. Verstraete and S. Possemiers (2011). The host selects mucosal and luminal associations of coevolved gut microorganisms: a novel concept. *FEMS Microbiol Rev*; 35 (4): 681-704.
102. Van den Bogert B., J. Boekhorst, R. Herrmann, E.J. Smid, E.G. Zoetendal and M. Kleerebezem (2013). Comparative genomics analysis of *Streptococcus* isolates from the human small intestine reveals their adaptation to a highly dynamic ecosystem. *PLoS One*; 8 (12): e83418.
103. van den Bogert B., O. Erkus, J. Boekhorst, M. de Goffau, E.J. Smid, E.G. Zoetendal and M. Kleerebezem (2013). Diversity of human small intestinal *Streptococcus* and *Veillonella* populations. *FEMS Microbiol Ecol*; 85 (2): 376-388.
104. Van den Bogert B., M.M. Leimena, W.M. De Vos, E.G. Zoetendal and M. Kleerebezem (2011). Functional intestinal metagenomics. In: *Handbook of molecular microbial ecology II: metagenomics in different habitats*. pp 170-190. Edited by F.J. De Bruijn. Willey-Blackwell.
105. van den Nieuwboer M., E. Claassen, L. Morelli, F. Guarner and R.J. Brummer (2014). Probiotic and synbiotic safety in infants under two years of age. *Benef Microbes*; 5 (1): 45-60.
106. van der Meer J.R., W.M. de Vos, S. Harayama and A.J. Zehnder (1992). Molecular mechanisms of genetic adaptation to xenobiotic compounds. *Microbiol Rev*; 56 (4): 677-694.
107. Van Immerseel F., R. Ducatelle, M. De Vos, N. Boon, T. Van De Wiele, K. Verbeke, P. Rutgeerts, *et al.* (2010). Butyric acid-producing anaerobic bacteria as a novel probiotic treatment approach for inflammatory bowel disease. *J Med Microbiol*; 59 (Pt 2): 141-143.
108. Venema K. and P. van den Abbeele (2013). Experimental models of the gut microbiome. *Best Pract Res Clin Gastroenterol*; 27 (1): 115-126.
109. Walker A.W., J. Ince, S.H. Duncan, L.M. Webster, G. Holtrop, X. Ze, D. Brown, *et al.* (2011). Dominant and diet-responsive groups of bacteria within the human colonic microbiota. *ISME J*; 5 (2): 220-230.
110. Walter J. (2008). Ecological role of lactobacilli in the gastrointestinal tract: implications for fundamental and biomedical research. *Appl Environ Microbiol*; 74 (16): 4985-4996.
111. Weinstock G.M. (2012). Genomic approaches to studying the human microbiota. *Nature*; 489 (7415): 250-256.
112. Wessman P., D. Mahlin, S. Akhtar, S. Rubino, K. Leifer, V. Kessler and S. Hakansson (2011). Impact of matrix properties on the survival of freeze-dried bacteria. *J Sci Food Agric*; 91 (14): 2518-2528.
113. Yutin N. and M.Y. Galperin (2013). A genomic update on clostridial phylogeny: Gram-negative spore formers and other misplaced clostridia. *Environ Microbiol*; 15 (10): 2631-2641.
114. Ze X., S.H. Duncan, P. Louis and H.J. Flint (2012). *Ruminococcus bromii* is a keystone species for the degradation of resistant starch in the human colon. *ISME J*; 6 (8): 1535-1543.
115. Zoetendal E.G., J. Raes, B. van den Bogert, M. Arumugam, C.C. Booijink, F.J. Troost, P. Bork, *et al.* (2012). The human small intestinal microbiota is driven by rapid uptake and conversion of simple carbohydrates. *ISME J*; 6 (7): 1415-1426.

7

8 CHAPTER

Appendices

Summary

Humans, like other mammals, are not single-species organisms, but they constitute in fact very complex ecosystems. The extensive network of host-microbe and microbe-microbe interactions is tremendously important for our health, and we are just starting to unravel the mechanisms by which microbes contribute to host health and disease.

Especially the intestinal tract of both humans and mammals contains an enormous diversity of microbial species of which many still remain to be cultured and characterized. There are numerous diseases for which aberrations in composition and diversity of the intestinal microbiota have been reported. Probiotic microorganism defined as "live microorganisms that, when administered in adequate amounts, confer a health benefit on the host" have the potential to modulate the intestinal microbiota and thereby contribute to health and well-being. To this end, the relative abundance of a specific bacterial phylotype, named CRIB, was found to be associated with probiotic-induced changes in gut microbiota and decreased severity of pancreatitis and associated sepsis in an experimental rat model for acute pancreatitis studies. Later, a representative of this phylotype (strain CRIB) was isolated, and characterized using a polyphasic taxonomic approach. The taxonomy of several closely related members of the family *Peptostreptococcaceae* was revised in order to provide a valid systematic name to the isolate, for which *Romboutsia ilealis* was chosen. It was found that the majority of *Romboutsia*-associated 16S rRNA gene sequences have an intestinal origin, however, the specific roles that *Romboutsia* species play in the intestinal tract are largely unknown. To gain more insight in metabolic and functional capabilities of members of the genus *Romboutsia*, efforts towards the isolation of additional representatives were undertaken. This ultimately led to the isolation of a human small intestine-derived representative (strain FRIFI) of another novel *Romboutsia* species which was given the name *R. hominis*. Characterization of both novel species of intestinal origin, i.e. *R. ilealis* and *R. hominis*, belonging to the genus *Romboutsia* at the genomic and functional level provided first insights into the genetic diversity within the genus *Romboutsia* and their adaptation to a life in the (upper) intestinal tract. To this end, *Romboutsia* species are flexible anaerobes that are adapted to a nutrient-rich environment in which carbohydrates and exogenous sources of amino acids and vitamins are abundantly available.

Microbiomic approaches such as those employed in this study can be used to pinpoint specific commensal microbes that might have a beneficial effect on the health of the host. In addition, the combination of genomic and functional analyses with single organisms and complex communities can be used to identify microbial functionalities that are related to health and disease, which in turn can be used to select potential probiotic strains based on specific functional properties. Ultimately, these approaches

will lead to the characterization of (new) beneficial commensal microbes that exert health-promoting effects, with the ultimate possibility for them to be exploited as next-generation probiotics.

Nederlandse samenvatting

Op en in ons lichaam bevinden zich een grote hoeveelheid micro-organismen. Waarschijnlijk meer dan je jezelf voor kunt stellen. Er wordt zelfs geschat dat het menselijke lichaam 10 keer meer microbiële cellen bevat dan menselijke cellen. De meeste van deze micro-organismen zijn te vinden in het maag-darmkanaal. De aanwezigheid van deze grote hoeveelheid micro-organismen is in de meeste gevallen zeer gewenst en ze spelen een belangrijke rol in onze gezondheid. Zo leveren deze micro-organismen bijvoorbeeld een essentiële bijdrage aan onze stofwisseling, onder andere door de productie van energie en vitamines. Er zijn steeds meer aanwijzingen die de hypothese ondersteunen dat als er iets mis is met de samenstelling van de gemeenschap van micro-organismes in onze darm (ook wel darmmicrobiota genoemd) dat direct dan wel indirect effect heeft op onze gezondheid. Maar ondanks het feit dat er met name de laatste jaren veel onderzoek wordt gedaan naar de micro-organismen die zich bevinden in de darm, zijn er nog steeds veel vragen en onduidelijkheden. Zo weten we bijvoorbeeld wel dat de samenstelling van de microbiota in en op ons lichaam verschilt per persoon en per anatomische locatie en dat het maag-darmkanaal van zowel mensen als dieren een enorme verscheidenheid aan soorten micro-organismen bevat, maar zo weten we nog niet precies welke micro-organismen er precies onder welke omstandigheden een positieve bijdrage leveren aan onze gezondheid. Het vergaren van meer kennis over de samenstelling van darmmicrobiota zal bijdragen aan het krijgen van meer inzicht in de rol van micro-organismen in gezondheid en ziekte. Dit alles staat beschreven in **Hoofdstuk 1**. In dit hoofdstuk staat ook beschreven welke technieken we tot onze beschikking hebben om individuele micro-organismen dan wel complexe gemeenschappen van meerdere soorten micro-organismen in detail te bestuderen. Een groot probleem hierbij is dat we van een groot deel van de micro-organismen tot op de dag van vandaag nog niet staat zijn ze te laten groeien onder experimentele condities. Dit maakt het erg lastig om onderzoek naar ze te doen. De enorme ontwikkelingen die er de laatste jaren zijn gemaakt op het gebied van met name moleculaire technieken hebben gezorgd voor een heel nieuw palet aan mogelijkheden om de samenstelling en de functionaliteit van de micro-organismen in ons maag-darmkanaal te bestuderen. Een groot aantal van deze technieken zijn toegepast in de experimenten beschreven in dit proefschrift.

In **Hoofdstuk 2** wordt er eerst nog een literatuuroverzicht gegeven van wat er op het moment van schrijven bekend was over de samenstelling van de darmmicrobiota, en welke technieken er op dat moment gebruikt werden om deze complexe microbiële gemeenschappen in de darm te besturen. Daarbij wordt er ook een overzicht gegeven van de ziektes/aandoeningen waarbij er een verandering in darmmicrobiota samenstelling gevonden is. Het gaat hierbij om aandoeningen als prikkelbare darm syndroom en inflammatoire darmziektes, maar ook om aandoeningen die op het eerste oog niks te

maken lijken te hebben met het maag-darmkanaal zoals autisme, eczeem en reumatoïde artritis. Verder wordt er in dit hoofdstuk ook ingegaan op de mogelijkheden die er zijn om de samenstelling van de darmmicrobiota te veranderen. Hierbij is er specifiek aandacht gegeven aan het gebruik van producten die probiotische micro-organismen bevatten, probiotica genaamd. Probiotische micro-organismen worden gedefinieerd als "Levende micro-organismen, die wanneer in voldoende hoeveelheden toegediend, een gezondheidsbevorderend effect hebben op de gastheer". Probiotica zijn potentieel in staat om de samenstelling van de microbiële gemeenschappen in de darm te beïnvloeden en ze kunnen daarbij bijdragen aan gezondheid en welzijn. Maar er zal nog veel onderzoek gedaan moeten worden voordat we precies begrijpen hoe specifieke soorten van micro-organismen bijdragen aan onze gezondheid.

De daarna volgende hoofdstukken in dit proefschrift beschrijven de verschillende fasen van de reis die is afgelegd om de (gezondheidsbevorderende) eigenschappen van één specifieke bacteriesoort te bestuderen. Deze reis is begonnen met de ontdekking van een nieuwe bacteriesoort in de dunne darm van ratten en is uiteindelijk geëindigd met de diepgaande karakterisatie van een hele nieuwe bacteriële genus, *Romboutsia* genaamd.

In **Hoofdstuk 3** staat de studie beschreven die als doel had om meer inzicht te krijgen in de invloed van de darmmicrobiota op het proces van bacteriële translocatie gedurende acute pancreatitis. Hiervoor werd gebruik van een experimenteel rattenmodel voor acute pancreatitis en werd er een gedetailleerde analyse gemaakt van de microbiële gemeenschap in de verschillende delen van het maagdarmkanaal van deze ratten. Hierbij zagen we dat de aanwezigheid van een tot dan toe onbekende bacterie positief en significant gecorreleerd was met een verbeterde gezondheid van de ratten. Deze nieuwe bacteriesoort hebben we de werktitel 'CRIB' gegeven, wat staat voor 'commensal rat ileum bacterium'. Op basis van de bevindingen beschreven in dit hoofdstuk hebben we de hypothese opgeworpen dat 'CRIB' mogelijk gezondheidsbevorderende en dus probiotische eigenschappen heeft.

Tot dan toe was de aanwezigheid van 'CRIB' alleen op basis van DNA sequenties aangetoond, maar hadden we nog geen bacterie in handen die we in laboratorium konden bestuderen. In **Hoofdstuk 4** staat beschreven hoe we vervolgens de 'CRIB' bacterie geïsoleerd hebben uit een monster verkregen uit de dunne darm van een rat. Na isolatie is deze bacterie vergeleken met andere, reeds beschreven bacteriën. Nadat we hadden aangetoond dat het daadwerkelijk om een nieuwe bacteriesoort ging, hebben we deze bacterie de naam *Romboutsia ilealis* gegeven (wat hierna afgekort wordt als *R. ilealis*). Het isolaat zelf heeft de stamnaam CRIB behouden, waardoor die nu door het leven gaat als *Romboutsia ilealis* CRIB. In dit hoofdstuk staat ook beschreven waarom we

ook vijf nauwgerelateerde bacteriesoorten een andere naam hebben gegeven, oftewel taxonomisch hernoemd waardoor ze niet langer de naam *Clostridium* hebben. Dit hebben we gedaan om te zorgen dat er geen misverstanden meer kunnen zijn over het feit dat deze bacteriesoorten echt anders zijn dan *Clostridium* soorten die reeds beschreven zijn.

Vervolgens staat in **Hoofdstuk 5** beschreven hoe we nog dieper in gegaan zijn op de karakteristieken van *R. ilealis* CRIB. Ditmaal door de genoomsequentie volledig te ontrafelen. In een genoomsequentie staan alle potentiele eigenschappen van een organisme gecodeerd. Met deze informatie kun je voorspellingen doen over wat een organisme allemaal zou moeten kunnen doen. Zo konden we in het genoom van *R. ilealis* CRIB bijvoorbeeld vinden welke suikers deze bacterie zou moeten kunnen omzetten in andere stoffen. Ook vonden we eigenschappen die er op duiden dat *R. ilealis* CRIB is aangepast aan een leven in de dunne darm. Eigenschappen zoals het omzetten van galzouten en ureum, stofjes uitgescheiden door het (menselijk) lichaam. Het vinden van bepaalde eigenschappen in het genoom van een bacterie betekend alleen nog niet dat een bacterie deze eigenschappen ook daadwerkelijk in de praktijk kan laten zien (tot expressie brengen). Om in het lab aan te tonen dat *R. ilealis* CRIB ook daadwerkelijk op bepaalde suikers kan groeien hebben we genexpressie experimenten uitgevoerd, hetgeen ook beschreven wordt in hoofdstuk 4. Hiermee hebben we aangetoond welke genen er tot expressie worden gebracht wanneer *R. ilealis* CRIB groeit op bepaalde suikers (glucose, L-fucose en fructo-oligosaccharides) en welke afvalproducten (metabolieten) er worden gemaakt. Deze metabolieten kunnen vervolgens weer worden opgenomen en gebruikt worden door andere bacteriën of door de gastheer. Al met al hebben we door het onderzoek beschreven in dit hoofdstuk een veel beter inzicht gekregen in wat *R. ilealis* CRIB allemaal kan en welke rol de soort *R. ilealis* mogelijk zou kunnen spelen in het microbiële ecosysteem in de dunne darm.

Een volgende stap was om te kijken of *R. ilealis* ook voorkomt in andere ecosystemen, waarbij we met name geïnteresseerd waren in de aanwezigheid van *R. ilealis* in de darmen van de mens. Op basis van de data die tot dan toe beschikbaar was, waren er aanwijzingen dat er nog meer soortgelijke bacteriën te vinden zouden moeten zijn. Om deze reden hebben we een groot aantal monsters gescreend op de aanwezigheid van *R. ilealis*-achtige sequenties. Hieruit bleek dat *R. ilealis* en soortgelijke bacteriën inderdaad ook in mensen te vinden zijn. Dit staat beschreven in **Hoofstuk 6**. Hier staat ook beschreven hoe we geprobeerd hebben een *R. ilealis* stam te isoleren met menselijk materiaal als uitgangsmateriaal. Het is alleen erg lastig om monsters te nemen uit de dunne darm van mensen, deze zit namelijk goed weggestopt. Als alternatief hebben we gebruik gemaakt van de ontlastingmonsters van mensen met een ileostoma. Dit zijn mensen waarbij de dikke darm operatief is verwijderd en waarbij het einde van de dunne darm is bevestigd

aan de buikwand, eindigend in een stoma. Het materiaal wat in het stoma-zakje terecht komt (ileostoma effluent), wat afkomstig is uit de dunne darm, hebben we gebruikt voor isolatie-experimenten. Uiteindelijk is het ons gelukt een nieuw isolaat te verkrijgen, welke we FRIFI hebben genoemd wat staat voor 'first *Romboutsia* isolaat from ileostoma'. Gelijk aan de experimenten beschreven in hoofdstuk 4 en 5 hebben we ook voor dit isolaat een veelvoud aan karakterisaties uitgevoerd. Uiteindelijk bleek dat het isolaat wel degelijk anders was dan *R. ilealis* CRIB, en dat het ging om een nieuwe bacteriesoort, die we *Romboutsia hominis* hebben genoemd. Al met al heeft de karakterisatie van *R. ilealis* CRIB en *R. hominis* FRIFI, zowel op genomisch als functioneel niveau, ons de eerste inzichten gegeven in de genetische diversiteit die aanwezig is binnen het genus *Romboutsia* en hoe ze zich aangepast hebben aan een leven in de (dunne) darm. Beide *Romboutsia* soorten worden, bijvoorbeeld, gekarakteriseerd door eigenschappen op basis waarvan we conluderen dat ze zich hebben gepast aan een nutriënt-rijke leefomgeving waarin koolhydraten, aminozuren en vitamines rijkelijk aanwezig zijn.

In **Hoofdstuk 7** wordt een overzicht gegeven van de experimenten die er in de afgelopen jaren allemaal zijn uitgevoerd om de mogelijk probiotische eigenschappen van *R. ilealis* CRIB te bekijken. Denk hierbij aan maag-darm-overlevingsexperimenten en experimenten waarbij we naar de stimulatie van immuuncellen hebben gekeken. Verder staan in dit hoofdstuk ook de resultaten beschreven van de experimenten die gedaan zijn met betrekking tot vermarkting van *R. ilealis* CRIB als probiotische stam. In dit kader hebben we specifiek gekeken naar het vriesdrogen van levende *R. ilealis* CRIB cellen. Vriesdrogen is een proces dat wordt toegepast om te zorgen dat levende cellen gedurende lange tijd (jaren) levensvatbaar blijven. Tenslotte staan in dit hoofdstuk ook de eerste resultaten beschreven van de studie beschreven waarbij we met een meterslange katheter bij gezonde vrijwilligers monsters hebben afgenomen op verschillende locaties in de dunne darm. Het ultieme experiment om te zien of *Romboutsia* soorten ook daadwerkelijk in de dunne darm van mensen aanwezig zijn. Tevens hebben de vrijwilligers ook gedurende 2 weken een synbioticum (een probiotisch product gecombineerd met een prebiotische vezel) geconsumeerd met als doel eventueel aanwezige *Romboutsia* soorten te stimuleren. Toekomstige analyses moeten nog uitwijzen wat de precieze resultaten zijn.

Al met al hebben we (nog) niet aankunnen tonen dat de bacteriesoort *R. ilealis* daadwerkelijke gezondheidsbevoderende eigenschappen heeft. Wel hebben we veel inzichten gekregen in de *Romboutsia* soorten die onderdeel uitmaken van de 'normale' darmmicrobiota. Alles samengenomen laat het werk beschreven in dit proefschrift de kracht zien van het gebruik van moleculaire technieken om specieke bacteriën te vinden die mogelijk een positief effect hebben op de gezondheid van de gastheer en welke mogelijk gebruikt kunnen worden als probiotische micro-organismen.

List of abbreviations

AAD	Antibiotic-associated diarrhoea
ABC	ATP-binding cassette
ANI	Average nucleotide identity
BLAST	Basic local alignment search tool
BSH	Bile salt hydrolase
CAS	CRISPR-associated
CCR	Carbon catabolite repression
CCS	Circular consensus sequencing
CD	Crohn's disease
CDS	Coding DNA sequences
CE	Capillary electrophoresis
CFU	Colony forming unit
COG	Clusters of orthologous groups
CRIB	Commensal rat ileum bacterium
CRISPR	Clustered regularly interspaces short palindromic repeats
DDH	DNA-DNA hybridization
DGGE	Denaturing gradient gel electrophoresis
DMA	Dimethyl acetal
EC	Enzyme commission
ECL	Equivalent chain length
EFSA	European food safety authority
FAO	Food and agricultural organization of the United Nations
FAP	Familial anastomosis polyposis
FBA	Flux balance analysis
FDR	False discovery rate
FISH	Fluorescent *in situ* hybridisation
FOS	Fructo-oligosaccharide
FRIFI	First *Romboutsia* isolate from ileostoma
FU	Fluorescence units
G+C	Guanine+Cytosine
GC	Gas chromatography
GI	Gastrointestinal
GL	Glycolipid
HPLC	High-performance liquid chromatography
IBD	Inflammatory bowel disease
IBS	Irritable bowel syndrome
IC	Ileocaecal
ICC	Intraclass correlation coefficient

IFN	Interferon
Ig	Immunoglobulin
IL	Interleukin
IS	Insertion element
KEGG	Kyoto encyclopaedia of genes and genomes
L	Lipid
LH-PCR	Length heterogeneity PCR
LPSN	List of Prokaryotic names with standing in nomenclature
MIC	Minimum inhibitory concentration
MIS	Microbial identification system
MLN	Mesenteric lymph node
MMC	Migrating motor complex
MODS	Multiple organ dysfunction syndrome
MPN	Most-probable number
mRNA	Messenger RNA
MS	Mass spectrometry
ncRNA	Non-coding RNA
NEC	Necrotizing enterocolitis
NMR	Nuclear magnetic resonance
OD	Optical density
OTU	Operational taxonomic unit
PBMC	Peripheral blood mononuclear cells
PCA	Principle component analysis
PCoA	Principle coordinates analysis
PCR	Polymerase chain reaction
PL	Phospholipid
PMA	Propidium monoazide
PTS	Phosphotransferase system
PVX	Polyvitex
PY	Peptone-yeast extract
PYG	Peptone-yeast extract-glucose
qPCR	Quantitative PCR
QPS	Qualified presumed safety
RISA	Ribosomal intergenic spacer analysis
rRNA	Ribosomal RNA
SCFA	Short chain fatty acids
SEM	Standard error of the mean
SFB	Segmented filamentous bacteria
SHIME	Simulator of the human intestinal microbial ecosystem

SIBO	Small intestinal bacterial overgrowth
SIM	Sulfide indole motility
SIRS	Systemic inflammatory response syndrome
SSCP	Single strain conformation polymorphism
SSU	Small subunit
TGGE	Temperature gradient gel electrophoresis
T_H	T helper
TLC	Thin-layer chromatography
TLR	Toll-like receptor
TNF	Tumor necrosis factor
TRAC	Transcript analysis with the aid of affinity capture
T-RF(LP)	Terminal-restriction fragment (length polymorphism)
tRNA	Transfer RNA
UC	Ulcerative colitis
WHO	World health organization

Co-author affiliations

Louis M.A. Akkermans[1]
Susana Fuentes[2]
Hein G. Gooszen[1,3]
Wieke Grievink[2]
Bastian Hornung[2,4]
Sacha A.F.T. van Huijm[5,6]
Sergey R. Konstantinov[2]
L. Paul van Minnen[1]
Laura van Niftrik[7]
Henk Panneman[8]
Lars Paulin[9]
Bernadet Renckens[5]
Ger T. Rijkers[1,10,11,12]
Jarmo Ritari[13]
Frans M. Rombouts[14]
Peter Schaap[4]
Hauke Smidt[2]
Ivelina Staneva[2]
Harro M. Timmerman[1,15]
Brian J. Tindall[16]
Willem M. de Vos[2,13,17]

1 Department of Surgery, University Medical Center, Utrecht, The Netherlands
2 Laboratory of Microbiology, Wageningen University, The Netherlands
3 Department of Operating Rooms, University Medical Center St. Radboud, Nijmegen, The Netherlands
4 Laboratory of Systems and Synthetic Biology, Wageningen University, Wageningen, The Netherlands
5 Nijmegen Centre for Molecular Life Sciences, CMBI, Radboud UMC, Nijmegen, The Netherlands
6 NIZO Food Research, Ede, the Netherlands
7 Department of Microbiology, Institute for Water and Wetland Research, Radboud University Nijmegen, Nijmegen, The Netherlands
8 Dr. Van Haeringen Laboratorium B.V., Wageningen, The Netherlands
9 Institute of Biotechnology, University of Helsinki, Helsinki, Finland
10 Department of Operating Rooms, University Medical Center St. Radboud, Nijmegen, The Netherlands
11 Department of Medical Microbiology and Immunology, St. Antonius Hospital, Nieuwegein, The Netherlands
12 Department of Science, University College Roosevelt, Middelburg, The Netherlands
13 Department of Veterinary Biosciences, University of Helsinki, Finland
14 Laboratory of Food Microbiology, Wageningen University, The Netherlands
15 Winclove Probiotics, Amsterdam, The Netherlands
16 Leibniz-Institut DSMZ – Deutsche Sammlung vor Mikroorganismen und Zellkulturen GmbH, Braunschweig, Germany
17 Department of Bacteriology and Immunology, Haartman Institute, University of Helsinki, Finland

Acknowledgements / Dankwoord

'*Alles komt goed*'

Een motto waaraan ik mijzelf gedurende de afgelopen jaren tijdens mijn promotie-onderzoek regelmatig heb moeten herinneren. Het waren namelijk velen jaren hard werken met zeker een aantal hoogtepunten (gelukkig maar!), maar ook vele dieptepunten. Maar hier zijn we dan aangekomen bij het dankwoord en dan lijkt het er dan toch op of het allemaal goed gekomen is!

Maar ik heb het zeker niet alleen gedaan. Graag wil ik dan ook van de gelegenheid gebruik maken om *alle* mensen te bedanken die professioneel of persoonlijk hebben bijgedragen aan mijn onderzoek. Ik heb in de afgelopen jaren met een groot aantal mensen mogen samenwerken en dit proefschrift was niet tot stand gekomen zonder de hulp van vele anderen, die, ieder op hun eigen manier, hebben bijgedragen. Tevens gaat mijn dank uit naar alle vrijwilligers die hun medewerking hebben verleent aan de verschillende studies.

Hierbij wil ik dan ook graag mijn (co-)promotoren bedanken. Allereerst gaat mijn dank uit naar mij promotoren, Hauke Smidt en Willem de Vos; bedankt dat ik het voorrecht heb gehad mijn PhD te voltooien binnen jullie onderzoeksgroep. Hauke, heel erg bedankt voor alle steun door de jaren heen. De rol van promotor én dagelijkse begeleider was niet altijd gemakkelijk, maar ik ben je dankbaar voor je hulp, adviezen en ondersteuning. Ik ben enorm gegroeid door de ruimte en het vertrouwen dat je me hebt gegegeven om binnen mijn onderzoek m'n eigen keuzes (én fouten) te kunnen maken. Willem, jouw kritische blik en bijdrage gedurende het schrijven van de verschillende manuscripten is enorm waardevol voor mij geweest. Bedankt voor je inspirerende optimisme. Verder wil ik mijn co-promotor Ger Rijkers heel erg bedanken. Ger, bedankt dat je altijd betrokken bent gebleven! Jij zorgde ervoor dat ik niet helemaal op ben gegaan in de 'beestjes'. Jouw vermogen om alles, maar dan ook echt alles, simpel en beeldend uit te kunnen leggen is bewonderenswaardig.

Members of the thesis committee, thank your for your time and critical assessment of this thesis.

'*The end of one journey, is the beginning of the next*', dit was ook zeker het geval voor mijn promotieonderzoek. Ik heb het voorrecht gehad om door te kunnen gaan met de resultaten waarmee de promotieonderzoeken van Harro Timmerman en Paul van Minnen eindigden. Harro en Paul, het is niet gemakkelijk je eigen werk zomaar uit handen te geven, maar ik ben dankbaar voor het feit dat ik op jullie werk heb mogen voortbouwen.

Zonder de hulp van co-auteurs zouden er geen manuscripten geweest zijn, dus bij deze wil dan ook alle co-auteurs bedanken voor hun bijdrage aan de manuscripten. Allereerst, professor Louis Akkermans en professor Hein Gooszen, bedankt voor het feit dat ik onderdeel heb mogen uitmaken van jullie onderzoeksgroep en voor jullie bijdrage aan de dierstudies. Ook wil ik graag professor Frans Rombouts bedanken voor zijn bijdrage aan het 'CRIB' onderzoek. Beste Frans, ik heb het een eer gevonden om het genus *Romboutsia* naar u te mogen vernoemen. Henk Panneman, bedankt voor je bijdrage m.b.t. de T-RFLP data. Verder wil Laura van Nifrik bedanken voor de prachtige plaatjes gemaakt met de elektronenmicroscoop. Als je met zulke kleine organismen werkt als bacteriën, is erg leuk om ze eens van dichterbij te kunnen bekijken! Laura, bedankt voor de uurtjes die je met mij achter de microscoop hebt doorgebracht.

In addition, I would like to thank Susana Fuentes, Wieke Grievink and Sergey Kostantinov. Your joined efforts in the identification and isolation of CRIB have been essential for the start of my research. Furthermore, I would like to thank Brian Tindall for his enormous input on the characterization and reclassification paper. Brian, thank you very much for sharing your knowledge on bacterial taxonomy with me. Sacha van Huijm, Bernadet Renckens, Peter Schaap, Jarmo Ritari and Lars Paulin, thank you very much for the help on the genome manuscripts. And a special thanks to Bastian Hornung for the collaboration on both genome manuscripts. Bastian, I really appreciate it that you never seemed to get tired of my never-ending nitpicking.

Ook wil ik iedereen bedanken die zich ook hebben ingezet bij de verschillende studies beschreven in dit proefschrift.

Speciaal wil ik hier Marco van Es en Isolde Besseling-van der Vaart van Winclove Probiotics bedanken voor hun inzet als projectleider/manager. Marco, ik heb bewondering voor je enorme 'drive' om samenwerkingen tussen wetenschappelijke instellingen en het bedrijfsleven te stimuleren. Isolde, dankzij jouw inzet als projectmanager van het 'CRIB' project heeft dit geleidt tot een succesvolle afsluiting van het project. Ik heb respect voor hoe het je altijd gelukt is om vergaderingen te beleggen, voortgangsverslagen op tijd de deur uit te krijgen, en de diverse contracten te regelen. Dankzij jou hebben we meerdere keren een projectverlenging gekregen, waarvoor mijn dank groot is.

Verder wil ik mijn project-partners bedanken van de afdeling Maag-darm-leverziekten van de Universiteit Maastricht. Professor Ad MasClee, Freddy Troost, Harm-Jan Pieters en Ellen Wilms, bedankt voor jullie bijdrage gedurende alle jaren dat het 'CRIB' project heeft gelopen. Ellen, met jouw komst kwam er toch nog een tweede AIO op het 'CRIB' project. Dankzij jouw enorme inzet, samen met je student Alvaro Garcia Fuentes, is het dan toch gelukt om uiteindelijk nog de vrijwilligersstudie uit te voeren.

A special thanks for the colleagues at the Department of Microbiology of the Swedish University of Agricultural Sciences in Uppsala (Sweden) that welcomed me and my colleague Ye Tian in their lab for a two-month period and allowed me to make use of their freeze-drying facilities. A special thanks to Sebastian Håkansson, Per Wessman and Joakim Bjerketorp, thanks to your true friendliness. I have great memories of my time in Sweden.

Also a special thanks to professor Johan Söderholm and Mats Persborn from the Department of Clinical and Experimental Medicine of the Linköping University (Sweden) for the collaboration on the pouch biopsy samples.

Verder wil ook de collega's van het Laboratorium voor Microbiële Ecologie en Technologie van de Universiteit van Gent (Belgium), Tom van de Wiele and Annelies Geirnaert in het bijzonder, bedanken voor hun samenwerking. Heel erg bedankt dat we de SHIME en maag-darm-overlevingsexperimenten in samenwerking met jullie hebben kunnen doen.

De collega's van de afdeling Pediatrische Immunologie van het Wilhelmina Kinderziekenhuis (Universitair Medisch Centrum Utrecht), Sytze de Roock in het bijzonder, wil ik graag bedanken voor het uitvoeren van de T-cell differentiatie assays.

Gedurende mijn promotieonderzoek heb ik het voorrecht gehad deel uit te mogen maken van meerdere teams.

Allereerst wil ik alle collega's bedanken van de 'Acute Pancreatitis - Mucosale barrière' werkgroep van de afdeling Heelkunde van het Universitair Medisch Centrum Utrecht. Dank voor jullie hulp bij het opzetten en uitvoeren van de dierstudies. Tevens wil ook de stoma-verpleegkundigen bedanken van het Universitair Medisch Centrum Utrecht en het Universitair Medisch Centrum St. Radboud voor het benaderen van de stoma-patiënten, waardoor het voor mij mogelijk was om ileostoma effluent samples te verkrijgen voor mijn onderzoek.

In het bijzonder wil ik hier ook André Verheem bedanken. André, het lijkt alweer zo lang geleden, maar maanden lang was het jij, ik en de ratjes. Bedankt voor je hulp, je enorme inzet (ook buiten werktijd) en voor de gezelligheid. Met plezier denk ik terug aan deze periode. Hierbij wil ik ook gebruik maken van de gelegenheid om alle medewerkers van het GDL te bedanken. Veel dank voor jullie ondersteuning tijdens de dierstudies.

Thanks to all past and present colleagues at the Laboratory of Microbiology and the Laboratory of Systems and Synthetic Biology (Wageningen University). Despite that I was not a 'doorsnee' PhD student, you have made me feel part of the team! It was great working

with all of you and thanks for the friendly atmosphere. Thanks also to the wonderful technical and administrative staff, your help and support has been vital. And thanks to the colourful pallet of colleagues with whom I have had the privilege of sharing an office: Alicia, Farai, Kyle, Lennart, Ludmilla, Noora, Rozelin, Sebastian, Thomas, and Tijs. You had to endure my regular sighing while writing manuscript or emails, and my complaining about yet another failed experiment, and 'superb' singing while wearing head phones. However, it was great to have people around to share the ups & downs of doing a PhD.

I would like to express my gratitude to all the students that chose to do their Bsc thesis, Msc thesis or internship under my supervision: Ivelina Staneva, Shi Wang, Ilias Arefaine, Marlene Escobar, Nguyen Thi Kha Tu, Masja Wever and Joep van Bortel. It was nice to have you all as company in the lab and thanks for all the enthusiasm and hard work.

Alle collega's van Winclove Probiotics (en Spencer Food) wil ik heel erg bedanken voor hun collegialiteit de afgelopen jaren. Ook al was ik de eerste jaren maar zelden 'in huis', toch was ik er altijd thuis. Gelukkig is dat niet veranderd nu ik de laatste tijd meer op kantoor te vinden ben. Ook heel erg bedankt voor de gezelligheid de afgelopen jaren, in het bijzonder tijdens de jaarafsluitingen en Winclove-weekendjes.

En dan nu een speciaal bedankje voor m'n paranimfen, Tom van den Bogert en Maria Stolaki. In de afgelopen jaren heb ik met jullie beide met veel plezier samengewerkt. Jullie bereidheid om iedereen te helpen siert jullie! Tommie, ik ben je dankbaar voor de talloze keren dat ik bij je terecht kon voor hulp, suggesties of gewoon even een praatje. En Maria-tje, altijd een luisterend oor. De vrolijkheid die je altijd met je meebrengt is bewonderenswaardig! Bedankt voor jullie vriendschap en de gezelligheid!

Tot slot, mijn 'thuisfront'. Zonder de steun van familie, vrienden en geliefden zou het me niet gelukt zijn. Bedankt voor alle belangstelling de afgelopen jaren; niet alleen mooie resultaten kon ik met jullie delen, maar zeker ook m'n frustraties. Een promotieonderzoek kost veel tijd en energie, en daarom is prettig om af en toe even stoom af te blazen. Bedankt voor al jullie liefde en ondersteuning door de jaren heen.

Coline

About the author

Jacoline (Coline) Gerritsen was born on the 21th of November 1984 in Amersfoort, The Netherlands. In 2003 she obtained her 'Atheneum' diploma at the Scholengemeenschap Marianum in Groenlo. In that same year, she continued with her study 'Biomedical Sciences' at the University of Utrecht. During her bachelor studies, she developed a special interest in microbiology, immulogy and infectious diseases. Therefore, after obtaining her BSc degree in 2006, she continued with her master studies following the specialization 'Immunity and Infection'. She performed her first Msc internship at the Department of Infectious Diseases and Immunology (Utrecht University) investigating the hemagglutinin-esterases of corona- and toroviruses. This was done under supervision of Dr Raoul de Groot. In her second MSc project, at Sanquin Research in Amsterdam, she worked on immunoglobin-G mediated phagocytosis by recombinant immunoglobulins under supervision of Dr Gestur Vidarsson. Her Msc thesis was focussed on the role of an AraC-like regulator on the adaptation of *Enterococcus faecium* to its environment, which was performed under supervision of Dr Janette Top at the Department of Medical Microbiology of the University Medical Centre in Utrecht. In 2008, she obtained her MSc degree *cum laude*. After her graduation, she started her PhD research on the collaborative 'CRIB' project that involved working at the Laboratory of Microbiology of Wageningen University, at the Department of Surgery of the University Medical Centre Utrecht, and at Winclove Probiotics. The work was done under the supervision of Prof. Dr Hauke Smidt, Prof. Dr Willem M. de Vos and Prof. Dr Ger T. Rijkers and the results of her PhD research project are presented in this thesis. Currently she works as a scientist at Winclove Probiotics.

List of publications

Stapleton N.M., J.T. Andersen, A.M. Stemerding, S.P. Bjarnarson, R.C. Verheul, **J. Gerritsen**, Y. Zhao, M. Kleijer, I. Sandlie, M. de Haas, I. Jonsdottir, C.E. van der Schoot and G. Vidarsson (2011). Competition for FcRn-mediated transport gives rise to short half-life of human IgG3 and offers therapeutic potential. *Nat Commun*; 2: 599.

Gerritsen J., H.M. Timmerman, S. Fuentes, L.P. van Minnen, H. Panneman, S.R. Konstantinov, F.M. Rombouts, H.G. Gooszen, L.M. Akkermans, H. Smidt and G.T. Rijkers (2011). Correlation between protection against sepsis by probiotic therapy and stimulation of a novel bacterial phylotype. *Appl Environ Microbiol*; 77 (21): 7749-7756.

Gerritsen J., H. Smidt, G.T. Rijkers and W.M. de Vos (2011). Intestinal microbiota in human health and disease: the impact of probiotics. *Genes Nutr*; 6 (3): 209-240.

Persborn M., **J. Gerritsen**, C. Wallon, A. Carlsson, L.M. Akkermans and J.D. Soderholm (2013). The effects of probiotics on barrier function and mucosal pouch microbiota during maintenance treatment for severe pouchitis in patients with ulcerative colitis. *Aliment Pharmacol Ther*; 38 (7): 772-783.

Gerritsen J., S. Fuentes, W. Grievink, L. van Niftrik, B.J. Tindall, H.M. Timmerman, G.T. Rijkers, H. Smidt (2014). Characterization of *Romboutsia ilealis* gen. nov., sp. nov., isolated from the gastro-intestinal tract of a rat, and proposal for the reclassification of five closely related members of the genus *Clostridium* into the genera *Romboutsia* gen. nov., *Intestinibacter* gen. nov., *Terrisporobacter* gen. nov. and *Asaccharospora* gen. nov. *Int J Syst Evol Microbiol*; 64 (Pt 5): 1600-1616.

Overview of completed training activities

Discipline specific activities

Meetings

- 10th Gut Day Symposium
  (2008, Utrecht)
- 2nd TNO conference 'Beneficial Microbes'
  (2010, Noordwijkerhout)
- 7th joint Rowett/INRA symposium 'Gut Microbiology: New Insights into Gut Microbial Ecosystems'
  (2010, Aberdeen, Scotland)
- 12th Gut Day Symposium
  (2010, Gent, Belgium)
- FASEB Summer Research Conference 'Probiotics, Intestinal Microbiota and the Host: Physiological and Clinical Implications'
  (2011, Carefree, USA)
- 6th Probiotics, Prebiotics & New Foods conference
  (2011, Rome, Italy)
- 13th Gut Day Symposium
  (2011, Wageningen)
- SFA/ISAPP conference
  (2011, Berkeley, USA)
- 3rd TNO conference 'Beneficial Microbes'
  (2012, Noordwijkerhout)
- ISME14 'The Power of the Small'
  (2012, Copenhagen, Denmark)
- 3rd International workshop 'Gut Microbiota in Health and Disease'
  (2012, Maastricht)
- SFA/ISAPP conference
  (2012, Cork, Ireland)
- ISAPP conference combined with the symposium 'Probiotics, Prebiotics and the Host Microbiome: the Science of Translation'
  2013, New York, USA)
- ISAPP conference combined with the 9th joined Rowett/INRA symposium 'Gut Microbiology: from sequence to funcion'
  (2014, Aberdeen, Scotland)

Courses

- Light in the intestinal tract
  (2009, Helsinki, Finland)
- Introduction course Immunology
  (2010, Utrecht)
- Introductory biostatistics for researchers
  (2010, Utrecht)
- Systems Biology: statistical analysis of ~omics data
  (2010, Wageningen)

General Courses

- VLAG PhD week
  (2009, Maastricht)
- PhD competence assessment
  (2010, Wageningen)
- Techniques for writing and presenting a scientific paper
  (2011, Wageningen)
- Voice matters
  (2012, Wageningen)
- Mobilizing your scientific network
  (2012, Wageningen)
- Career perspectives
  (2012, Wageningen)

Optionals

- Preparation PhD research proposal
- Microbiology PhD study trip
  (2009, northeast coast USA)
- Molecular ecology group meetings
  (weekly)
- PhD/Post-Doc meetings
  (biweekly)

Colophon

The research described in this thesis was financially supported by RVO, an agency of the Dutch Ministry of Economic Affairs (grant number FND 07-013).

The printing of this thesis was financially supported by GATC Biotech, Institut Allergosan, Nutrimmun GmbH, Wageningen University and Winclove Probiotics.

Cover design: Henri Aanstoot & Jacoline Gerritsen
Layout design: Jacoline Gerritsen
Printed by: Gildeprint Drukkerijen - The Netherlands