Exploring the Human Intestinal Microbiome in Health and Disease

Johan Dicksved
Faculty of Natural Resources and Agricultural Sciences
Department of Microbiology
Uppsala

Doctoral Thesis
Swedish University of Agricultural Sciences
Uppsala 2008

Acta Universitatis agriculturae Sueciae
2008:30

ISSN 1652-6880
ISBN 978-91-85913-63-3

Tryck: SLU Service/Repro, Uppsala 2008

Exploring the Human Intestinal Microbiome in Health and Disease

Abstract

In this thesis, molecular tools were used to study bacteria inhabiting the gastrointestinal tract of humans. One aim was to determine whether certain lifestyle factors, such as an anthroposophic lifestyle, or living on a farm, had an impact on the bacterial composition in fecal samples collected from 90 children. The bacterial composition was determined by terminal restriction fragment length polymorphism (T-RFLP). Each child had a unique bacterial community, but anthroposophic children had a higher bacterial diversity than farm children. A second aim was to study the bacterial communities in fecal samples and biopsies collected from patients with Crohn's disease (CD), compared to healthy individuals. Several molecular approaches were used to characterize the microbiota in identical twins, including healthy twins pairs and twins that were discordant (one has CD and one is healthy) or concordant (both have CD). The bacterial profiles of healthy twins were highly similar whereas in twins discordant for CD they were very different. An imbalance in the microbiota was observed in a subset of individuals with CD and was correlated with the ileal disease phenotype (ICD). A reduced diversity of *Firmicutes*, and in particular, a depletion of *Faecalibacterium prausnitzii* was correlated with ICD. By contrast there was an increased prevalence and abundance of *E. coli* in ICD individuals. These findings suggest that specific members of the gut microbiota play a functional role in ICD and this could have clinical significance. A third aim was to characterize the microbiota in stomach biopsies from individuals with stomach cancer compared to healthy individuals. The gastric cancer microbiota had a high bacterial diversity and a similar composition to that in dyspeptic controls. In both cases, the bacterial composition was dominated by streptococci and other *Firmicutes* and a low abundance of *Helicobacter pylori*, contradicting earlier reports of *H. pylori* dominance in the stomach. In conclusion, the bacterial community in the human gastrointestinal tract is more complex than originally thought and disturbances in this community are indicative of some disease states.

Keywords: Microbiota, terminal-restriction fragment length polymorphism (T-RFLP), gastrointestinal tract, identical twins, inflammatory bowel disease (IBD), Crohn's disease, gastric cancer, *Faecalibacterium prausnitzii*, lifestyle, diversity.

Author's address: Johan Dicksved, Department of Microbiology, SLU
Box 7025, 750 07 Uppsala, Sweden
E-mail: Johan.Dicksved@mikrob.slu.se

Contents

List of Publications

This thesis is based on the work contained in the following papers, referred to by Roman numerals in the text:

I Dicksved, J., Flöistrup, H., Bergström, A., Rosenquist, M., Pershagen, G., Scheynius, A., Roos, S., Alm, J.S., Engstrand, L., Braun-Fahrländer, C., von Mutius, E., and Jansson J.K. (2007). Molecular fingerprinting of the fecal microbiota of children raised according to different lifestyles. *Appl. Environ. Microbiol.* 73(7), 2284-2289.

II Dicksved, J., Halfvarson, J., Rosenquist, M., Järnerot, G., Tysk, C., Apajalahti, J., Engstrand, L., and Jansson, J.K. (2008). Molecular analysis of the gut microbiota of identical twins with Crohn's disease. *ISME J.* in press.

III Willing, B., Halfvarson, J., Dicksved, J., Rosenquist, M., Järnerot, G., Tysk, C., Engstrand, L., and Jansson, J.K. Twin studies shed light on differences in mucosa-associated microbiota according to Crohn's disease phenotype (manuscript).

IV Dicksved, J., Lindberg, M., Rosenquist, M., Enroth, H., Jansson J.K., and Engstrand, L. Molecular characterization of the stomach microbiota in patients with gastric cancer and controls (manuscript).

Abbreviations

CD	Crohn's disease
ICD	Ileal Crohn's disease
CCD	Colon Crohn's disease
T-RFLP	Terminal restriction fragment length polymorphism
TRF	Terminal restriction fragment
SCFA	Short chain fatty acid
AIEC	Adherent invasive Escherichia coli
MAP	*Mycobacterium avium* subsp. *paratuberculosis*

1 Introduction

The world of microbes, although invisible to the human eye, encompasses all types of ecological niches. For a microbiologist it is particularly intriguing to know that one of the most complex microbial ecosystems on earth lies enclosed within our bodies. The number of microbes residing in our gut exceeds the number of our own body cells by a factor of ten and thousands of different microbial species reside in our intestinal tract (Bäckhed *et al.*, 2005). Together, all of these microorganisms (collectively called the microbiota) harbor a repertoire of protein encoding genes that by far exceeds the gene pool found in our human genome. For more than a century researchers have been trying to understand the role of the intestinal microbiota for our well-being. In the past few decades increasing evidence has revealed that the intestinal microbiota harbors a tremendous potential to interact and interfere with the host, in both positive and negative ways. The relationship between the host and the gut microbiota is symbiotic and mutualistic interactions include microbial breakdown of otherwise non-digestible compounds that reach the large bowel. In addition, the products from microbial fermentation processes contribute to up to ten per cent of the host energy uptake (Macfarlane & Macfarlane, 2007), and are thought to be important to maintain gut health (Pryde *et al.*, 2002). Furthermore, the microbiota produces some essential vitamins and can have a functional barrier effect, protecting from pathogen invasion by colonizing mucosal surfaces (Stecher & Hardt, 2008). Studies of germ free animals have provided insights into the importance of the gut microbiota for maturation of the host. For example, it has been demonstrated that the microbiota plays an essential role in shaping our immune system (Hooper *et al.*, 2003; Mazmanian *et al.*, 2005), and intriguingly, it may exert behavioral effects on humans (Parracho *et al.*, 2005). Potential negative impacts in addition to pathogenesis include conversion of drugs or dietary compounds to carcinogenic compounds (Hope *et al.*, 2005). Importantly, the microbiota

has been implicated to play a critical role in several diseases of the gastrointestinal tract. Inflammatory bowel diseases, intestinal cancers, allergies and obesity are examples of such diseases; yet, the microbial role in the etiology of these diseases largely remains an enigma.

Despite a century of research, the work to elucidate the role of the microbial gut inhabitants in health and disease is still in its infancy. However, recent advances in new technologies and a large medical interest in this field of science will hopefully provide more insights to understanding the link between microbial ecology and human diseases. Recently, there has been research interest in sequencing the next human genome; i.e. the human microbiome. This is a global large scale effort with the aim to define a core microbiome for mankind, and to understand which host and environmental factors contribute to the dynamics of the human microbiome (Turnbaugh *et al.*, 2007).

2 The healthy gut microbiota

Starting in the mouth, the human digestive system runs through the body and approximately seven meters later it ends with the anus. This system is the home of an enormous number of organisms including bacteria, yeast, fungi, protozoa and viruses. Due to great physiological differences along the digestive tract, the presence and function of microbial communities varies largely at different sites of our body. Many of the bacteria are normal commensal inhabitants and colonizers of some part of the gut, whereas others are just transient, and follow the body fluids through the system. However, the relative abundances and contributions of indigenous versus transient members of the microbiota is currently not known. Regardless of intestinal site and whether they are residents or transients, many of these microbes play an important role in the gastrointestinal tract.

2.1 Function of the microbiota

The commensal microbiota have important and specific functions for the host. These can be divided into three main groups; trophic, protective and metabolic (Guarner & Malagelada, 2003). Trophic functions include control of proliferation, differentiation of epithelial cells and immune system maturation. The protective role of the microbiota is by occupying intestinal surfaces and creating a stability of the system and a milieu that prevents invasion from exogenous microbes, for example, by production of antimicrobial compounds. Metabolic functions include breakdown of non-digestible compounds by anaerobic fermentation, such as resistant starch and plant polysaccharides, which generates short chain fatty acids (SCFA) that serve as growth signals and fuel for the intestinal epithelium as well as for other microbes. In particular, butyrate has gained large attention for its health promoting properties (D'Argenio & Mazzacca, 1999; Pryde *et al.*, 2002; Wong *et al.*, 2006). Another example of metabolic function provided

by the microbiota is the production of some vitamins, such as, vitamin K and B_{12}, synthesis of amino acids and digestion of proteins. In exchange, the host provides a stable habitat with a constant flow of food and nutrients.

2.2 Establishment of the microbiota

We are born sterile but immediately after birth a microbial community starts to establish (O'Hara & Shanahan, 2006). The first inocula come from the mother, by the birth canal, and/or a fecal-oral route, and from the surrounding area. For example, the same serotype of *Escherichia coli* was found in the mouths of newborn babies as in feces from their mothers (Brook *et al.*, 1979). The microbial composition and colonization pattern varies widely from baby to baby (Palmer *et al.*, 2007). During the first months of life the community is highly variable and dominated with a facultative anaerobic microbiota. When space and nutrients are not limiting, bacteria with high division rates are more prevalent. However, over time, when the system becomes more densely populated and nutrient supply is limited, a more specialized microbiota takes over (Hooper *et al.*, 1999). The microbiota of the children becomes progressively more complex and at the age of two the microbial community is stabilized and an adult-like flora established with a large predominance of anaerobic microbes (Favier *et al.*, 2002). Without outer environmental stresses the dominant microbiota is considered to be remarkably stable over time (Ley *et al.*, 2006b; Vanhoutte *et al.*, 2004; Zoetendal *et al.*, 1998). However, by studying subgroups of the microbiota, temporal variation was observed in the lactobacilli population and *Enterococcus* population, whereas *Bacteroides* spp. and bifidobacteria remained relatively stable (Jernberg *et al.*, 2007; Vanhoutte *et al.*, 2004).

2.3 The microbiota of the upper digestive tract

A complex microbiota of more than 700 different species with representatives from six phyla (Figure 1) has been found in the oral cavity of humans (Aas *et al.*, 2005). The mouth is comprised of many different types of surfaces, all with a distinct microbial profile. However, some genera are common to all sites of the mouth, and these are *Gemella*, *Granulicatella*, *Streptococcus*, and *Veillonella* (Aas *et al.*, 2005). Oral microbial communities are commonly found in biofilms and it is probable that a metabolic inter-dependency is important for shaping these communities (Jenkinson & Lamont, 2005). Bacteria are introduced in the esophagus either by swallowing or by reflux from a colonized stomach (Pei *et al.*, 2004). Previously, it was generally thought that the esophagus was sterile or contained transients from the oral cavity due to the short retention time and

flushing of foods and fluids ingested by the host. However, Pei et al. (2004) showed that a diverse microbiota colonizes the mucosal surfaces of the esophagus. They found 95 phylotypes with representatives from six different phyla (Figure 1). Many of these were novel but culturable. In addition, this microbiota shares many members that usually are found in the oral cavity.

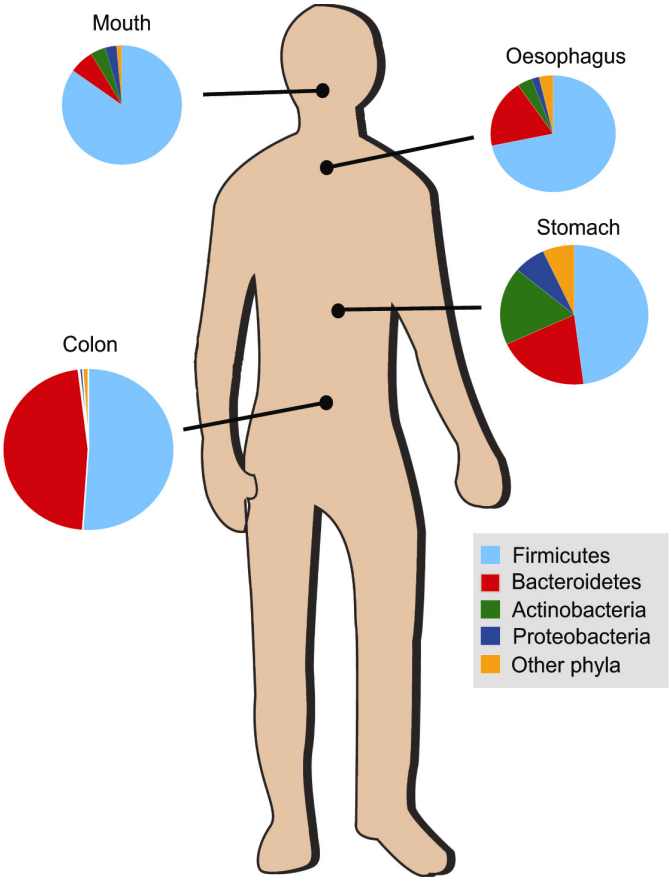

Figure 1. Distributions of bacterial phyla at different anatomical sites of the body. Each area corresponds to the abundance of phylotypes belonging to different phyla. Data derived from Dethlefsen et al. (2007), with permission from the publisher. Between 3 and 11 individuals were studied per habitat.

The stomach is the first line in the body's defense that microbes encounter as they enter the gastrointestinal tract. The acidic environment is thought to be an effective barrier, killing off most of the microbes in their passage through the stomach into the small bowel. In addition, a short transit time

contributes to the relatively low number of bacteria that are found in the stomach (Guarner, 2006). Previously the healthy stomach was considered to be otherwise sterile and not colonized by any microbes (Hentges, 1983). However, the gastric mucosa was hiding a secret that was discovered more than two decades ago (Marshall & Warren, 1984). *Helicobacter pylori* is nowadays a well studied colonizer of the gastric mucosa. Expression of the urease enzyme allows its survival in the acidic environment during its active passage into the mucus layer of the stomach lining where the pH is almost neutral. Additional bacteria have been found in the stomach, such as streptococci and lactobacilli, enterobacteria and other members of the oral cavity (Savage, 1977). There is no evidence that other bacteria than *H. pylori* can colonize the healthy stomach. However, streptococci are frequently isolated and some species may have the ability to survive the acidic environment. For example, *Streptococcus salivarius* exerts urease activity (Sissons & Hancock, 1993) and one can hypothesize that this can enable it to persist in the stomach. Recent advances in molecular biology have shown a more diversified gastric microbiota than previously thought. For example, Bik et al. (2006) found 128 phylotypes from gastric mucosa, with representatives from six bacterial divisions (Figure 1). This set of organisms was distinguished from what has been found in the mouth and esophagus, suggesting that the stomach is home for a distinct microbial ecosystem (Bik *et al.*, 2006).

2.4 Microbiota of the small bowel

The small bowel, comprised of duodenum, jejunum and ileum (Figure 2), is the main site for digestion of foods and absorption of nutrients in the body. Bile, pancreatic secretions and flushing mechanisms restrict the bacterial density in the proximal parts of the small bowel (O'Hara & Shanahan, 2006) whereas a lower transit rate increases bacterial density dramatically in the distal regions (Booijink *et al.*, 2007). Due to technical difficulties in sampling, the residential microbiota is not well characterized. According to current knowledge, the duodenal and jejunal microbiota is predominated by facultative anaerobes, mainly acid tolerant lactobacilli and streptococci (Hayashi *et al.*, 2005; Wang *et al.*, 2005). The obligate anaerobic *Firmicutes* that dominate the microbiota of the large bowel are not commonly found in the upper parts of the small bowel (Hayashi *et al.*, 2005). However, the distal part of the ileum hosts a more diverse microbiota and it resembles that of the colonic microbiota.

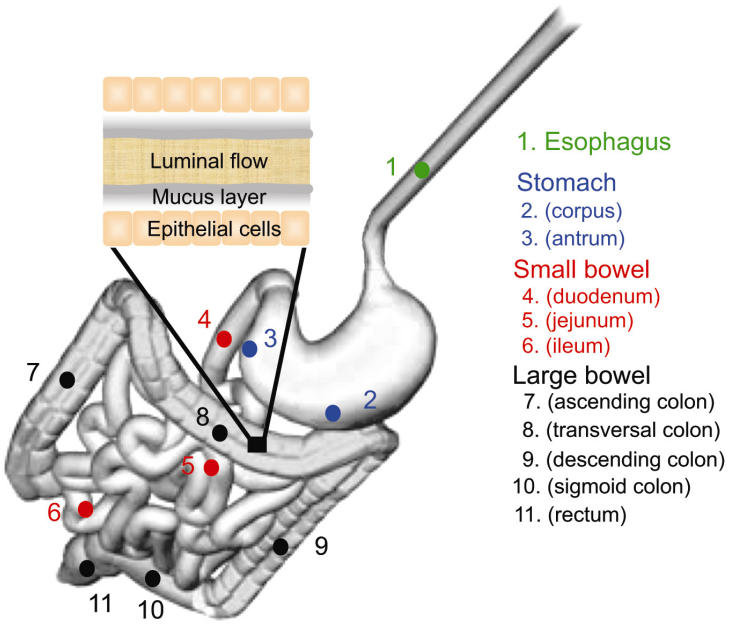

Figure 2. Overview of the human gastrointestinal tract. Within the gastrointestinal tract, mucus, a complex polymeric gel, covers the epithelial cell surface so it is not in contact with the luminal flow. This illustration showing the structure within the bowel is simplified and does not reflect true proportions.

2.5 Microbiota of the large bowel

The large bowel, comprised of ascending, transverse, descending colon and rectum (Figure 2), is the primary site for microbial persistence in the human body. The number of bacteria is estimated to be 10^{11}-10^{12} bacteria/gram colonic content (O'Hara & Shanahan, 2006). The anaerobic condition promotes fermentation and obligate anaerobic bacteria outnumber facultative anaerobic bacteria. Even if the bacterial diversity is very high, the large bowel contains few phyla and is totally dominated by two; i.e. *Bacteroidetes* and *Firmicutes* comprise over 90 % of the bacteria found in the large bowel (Figure 1; Ley *et al.*, 2006b). Despite great physiological differences between the proximal and distal parts of the colon, bacteria are uniformly distributed along the length of the large bowel. There are, however, some differences in bacterial types found in biopsies compared with those found in feces (Eckburg *et al.*, 2005; Lepage *et al.*, 2005; Nielsen *et al.*, 2003; Zoetendal *et al.*, 2002).

The number of bacteria found in the luminal region of the intestinal lining outnumbers the bacteria found within the mucus layer (Figure 2). Approximately 10^5 to 10^6 bacteria have been detected per biopsy from the colonic wall (Zoetendal *et al.*, 2002). In contrast, several studies have failed to detect bacteria on the epithelial surface from healthy individuals (Kleessen *et al.*, 2002; Swidsinski *et al.*, 2007; van der Waaij *et al.*, 2005). It has been challenging to get correct estimates of the number and types of bacteria in the bowel using traditional cultivation-based approaches because the majority of the bacteria found in the large bowel have not yet been cultured. This challenge was exemplified in a large survey of the gut microbiota using molecular approaches; i.e. cloning and sequencing (Eckburg *et al.*, 2005), with the finding that 80% of the species represented by the clones had never previously been cultured.

2.6 Factors influencing the microbial structure

The dominant microbial communities of the large bowel have shown a remarkable stability over time during adulthood, (Ley *et al.*, 2006b; Vanhoutte *et al.*, 2004; Zoetendal *et al.*, 1998) yet the microbial community profiles are unique for each individual (Eckburg *et al.*, 2005; Ley *et al.*, 2006b). However, there are numerous external factors that have potential to influence the microbial composition in the gut. Which factors and to what extent these can contribute are poorly understood. Therefore, it is important to set up well-designed studies to reduce the number of confounding factors as far as possible to better elucidate the role of specific factors on the gut microbial composition.

2.6.1 Host genetics

The structure of the intestinal microbiota has also been shown to be dependent on host genetics. Zoetendal et al. (2001) studied the fecal microbiota of individuals with different degrees of genetic relatedness, ranging from monozygotic twins to unrelated individuals. Monozygotic twins that lived apart for years showed high similarities in their microbial profiles whereas marital partners that lived in the same environment and with comparable feeding habits, did not. Van de Merwe et al. (1983), compared the culturable fraction of the fecal microbiota between mono and dizygotic twins and found higher similarities within monozygotic pairs. This difference was confirmed in young children, living in the same household (Stewart *et al.*, 2005). Animal studies have also supported the importance of genetic influence on the microbial composition. Ley et al. (2005), showed in a mouse model that three generations of offspring shared some similarities in their cecal microbiota, and these similarities were not seen when different

families of mice were compared. Taken together, these findings strongly argue that genetic factors provide a strong force in shaping the gut microbiota. However, colonization history and inoculum at birth from mother to child cannot be excluded to also be of importance for the similar microbial communities found among twins.

2.6.2 Birth delivery mode

The birth delivery mode has also been shown to influence the composition of the microbiota, with slower diversification and lack of anaerobic species, such as clostridia, in infants delivered by cesarean section (Grönlund *et al.*, 1999). In another study, a lower prevalence of clostridial species in the microbiota of a cesarean section delivered child, was prolonged over a period of several years (Salminen *et al.*, 2004).

2.6.3 Geographical impacts

There have also been reports that geographical region can have an impact on the gut microbial composition. For example, the gut microbiota has been shown to differ between westernized developed countries, Asian and developing countries. *H. pylori* has been found to be more prevalent in developing countries than in developed countries (Genta *et al.*, 1995). In another study, the microbiota of Estonian infants was shown to differ from the microbiota of Swedish infants (Björksten *et al.*, 1999). However, it has not yet been properly investigated whether these region- based discordances are based on the local environment, climate, genetic, dietary, or other factors connected to different lifestyles. In one study, local variations in the microbiota of the western European population were correlated to age, i.e. adult compared with elderly (Mueller *et al.*, 2006). For example, country and age interactions were observed for German and Italians with inverse interactions for the predominating *Clostridium coccoides* and *Bacteroides-Prevotella* groups. Contrary to this observation, no significant differences in the microbial composition were observed of individuals from five European countries (Lay *et al.*, 2005).

2.6.4 Influence of ageing

Due to an altered physiology in elderly, with a decreased intestinal motility, reduced secretion of gastric acid, and a change of dietary habits and lifestyle, the microbiota in elderly persons differs from younger adults, although with great individual variations. In elderly persons, it is common to have a reduction in numbers and diversity of *Bacteroides* and bifidobacteria, and a reduced production of SCFA and amylolytic activity. In addition, increased numbers of facultative anaerobes, fusobacteria, clostridia and eubacteria but decreased numbers of bifidobacteria have been reported in elderly people

(Woodmansey, 2007). Although there is large disagreement in the findings between studies, probably due to individual differences and use of different methods, the abundance and diversity of bifidobacteria is consistently reported to be decreased in elderly individuals (Mueller *et al.*, 2006; Woodmansey, 2007).

2.6.5 Influence of diet

Diet is often of interest for research studies due to its potential for modulation of the intestinal microbiota of the host, either in a beneficial or detrimental way. Dietary habits are considered to impact the microbial composition in the gut, in particular during the first years of life. For example, the fecal microbiota differs between breast-fed and formula fed infants, with more lactic acid bacteria and bifidobacteria in the breast-fed infants (Balmer & Wharton, 1991; Harmsen *et al.*, 2000). However broadly defined diets such as a "Western diet", comprised of a high intake of fat and animal proteins but low fiber, and a "Japanese diet", with less fat and more vegetables, have only revealed moderate differences in the gut microbiota involving a few bacterial genera (Finegold & Sutter, 1978). Similar observations were made by Aries et al. (1971), that only observed small differences in the bacterial composition when persons on a normal mixed diet were compared with strict vegetarians (Aries *et al.*, 1971). In another study a shift in the bacterial SCFA profile was observed between individuals that had an uncooked extreme vegan diet for one month compared to individuals that had a westernized diet. However, this shift was not reflected in the bacterial community composition (Peltonen *et al.*, 1992). Diet is likely to have larger effects on microbial communities in the small intestinal due to a more diversified substrate supply in that location compared to the colon. The dietary effects of the small bowel microbiota could explain the large temporal variations of the microbiota in small intestinal digesta (Booijink *et al.*, 2007) but not in feces (Zoetendal *et al.*, 1998). However, recently it was shown that diet could promote a division wide change of the microbiota. Obese people that were randomly assigned a fat restricted or carbohydrate restricted low calorie diet over a year, showed a pronounced increase in the abundance of *Bacteroidetes* with a corresponding decrease of *Firmicutes* (Ley *et al.*, 2006b).

2.6.6 Impact of antibiotics

It is well known that the microbiota can be strongly affected by antibiotic treatment. Recently, it was shown that the gut microbial composition could be disrupted for a long period of time, up to 2 years in some cases (Jernberg *et al.*, 2007; Löfmark *et al.*, 2006), depending on the antibiotics used. Jernberg et al. (2007) showed that the *Bacteroides* community never re-

established to its original composition within two years after a 7-day clindamycin treatment. Even if the original pre-antibiotic microbial composition is disrupted, it is generally assumed that the function of the entire microbial community is re-established (Norin, 1997). Still, more studies are necessary to determine what effect a population shift of the microbiota will have over a long period of time on host physiology.

2.6.7 Pre- and probiotics

The beneficial effects of fermented foods on human health were known long before the existence of bacteria was discovered. It was later found that fermented foods contain bacteria, mainly lactic acid bacteria (LAB) and bifidobacteria and these have since been extensively studied for their beneficial effects on human health. Probiotics are living bacterial food ingredients with beneficial effects on human health (Fuller, 1986). Bifidobacteria and LAB, especially lactobacilli, are commonly used as probiotics agents. In addition, specific dietary compounds have been shown to alter the metabolism or stimulate growth of beneficial groups of bacteria, such as inulin or fructooligosaccarides that stimulate the growth of bifidobacteria and lactobacilli in the gut (Van Loo, 2004). These compounds are collectively called prebiotics (Gibson & Roberfroid, 1995). Both pre- and probiotics and have been implicated to be protective against several clinical complications, such as allergy development (Abrahamsson *et al.*, 2007), inflammatory bowel diseases (IBD; Sartor, 2004), irritable bowel syndrome (IBS; Kajander *et al.*, 2008) and acute diarrhea (Canani *et al.*, 2007).

3 The microbiota and gut related disorders

A balance between the host immune system and the commensal gut microbiota is crucial for maintaining health. When this balance is disturbed, the host–microbe relationship can progress towards a disease state. Although some members of the commensal gut microbiota protect our body from foreign infections, other members are partly responsible for several of the disorders within the gastrointestinal tract. For example, some microbes in the mouth are responsible for two of the most common bacterial infections in humans, namely caries and periodontitis. Moreover, the discovery of *H. pylori* and its role in the development of gastritis, duodenal ulcers and gastric cancer gained large attention. In addition, the gut microbiota has been implicated to play an important role in several other disorders of the gastrointestinal tract, such as inflammatory bowel diseases (IBD), including Crohn's disease (CD) and ulcerative colitis (UC), irritable bowel syndrome (IBS), as well as colorectal cancer (CRC) and allergies. How, when and to which extent the microbiota is involved in the mechanisms causing these diseases remains an enigma. The mechanisms are thought to be multifactorial for most, if not all disorders of the gastrointestinal tract, including host susceptibility, microbial factors and environmental triggers.

In recent decades incidence rates of several of the immune dysfunctional disorders in the human gastrointestinal tract have increased dramatically (Guarner *et al.*, 2006). This increase in autoimmune disease incidence is, however, only observed in developed countries. A high level of hygiene has been identified as one risk factor for developing allergies, IBD, type I diabetes, and multiple sclerosis (Guarner *et al.*, 2006). The correlation between the modern hygienic lifestyle and the increase in allergies has given rise to the hygiene hypothesis, suggesting a link between decreased pathogen exposure and developing inflammatory disorders (Strachan, 1989). For example, there is supportive data, that carriage of helminths, a multicellular

worm parasite that is eliminated in western societies, might protect the host from developing immunological disorders (Guarner *et al.*, 2006). However, an alternative interpretation suggests that disturbances of the intestinal microbiota occur due to an increased use of antibiotics and "Westernized" dietary habits that disrupt the mechanisms involved in the development of immune tolerance (Noverr & Huffnagle, 2004). There are many hypotheses regarding which microbial mechanisms are underlying these diseases. One is the "pathogen hypothesis", where researchers have tried to correlate gastrointestinal disorders with infection of recognized pathogens. One example of a microbe that has been investigated as a potential invasive pathogen in gastrointestinal disorders is *Mycobacterium avium* subsp. *paratuberculosis* (MAP). For example, this bacterium has been implicated in CD development (McFadden *et al.*, 1987; Yoshimura *et al.*, 1987). Another hypothesis, suggests that breakdown of the balance between beneficial and harmful bacteria (dysbiosis) initiates an overly aggressive immune response (Tamboli *et al.*, 2004), or leads to proliferation of toxin-producing bacteria, such as in the situation for *Clostridium difficile* mediated diarrhea (Pothoulakis, 1996).

3.1 Allergy and asthma

There has been a constant increase of asthma and allergic diseases in western countries over the last few decades. Within any given country, asthma and allergies are more prevalent in urban than in rural areas (Guarner *et al.*, 2006). The assumptions underlying the hygiene hypothesis have resulted in a large interest in the study of the microbiota in individuals with allergy or asthma and such studies have revealed a high correlation between allergies and a disrupted microbial composition (Alm *et al.*, 1999; Björksten *et al.*, 1999; Kalliomäki *et al.*, 2001; Wang *et al.*, 2008; Wickens *et al.*, 1999). The hygiene hypothesis emphasizes the importance of the microbial composition that establishes in newborn infants (Strachan, 2000). By comparing the microbiota of infants that were only three weeks old, the microbial community composition was found to differ between atopic and non-atopic infants, mainly with higher proportions of bifidobacteria in the non-atopic infants. However, this difference in the microbial composition was not reflected when the children were three months old (Kalliomäki *et al.*, 2001). Furthermore, the microbial diversity was significantly lower in infants with atopic disease compared with non-atopic disease (Wang *et al.*, 2008). Differences in the gut microbiota have also been observed between Estonian (low prevalence of allergies) and Swedish infants (high prevalence of allergies), mainly with higher counts of lactobacilli and eubacteria in Estonian infants and instead higher counts of clostridia in Swedish infants

(Sepp *et al.*, 1997). Interestingly, the Estonian infant microbiota reflected the microbiota of infants in Western European countries several decades earlier, when the prevalence of allergies was low in Western Europe (Sepp *et al.*, 1997). In addition, the non-allergic infants within these respective countries had higher counts of lactobacilli and bifidobacteria whereas higher counts of coliforms and *Staphylococcus aureus* were reported among the allergic infants (Björksten *et al.*, 1999). This finding supported the hypothesis that beneficial properties of lactobacilli and bifidobacteria, commonly used as probiotics, were protective against allergy development. Probiotic bacteria, in particular lactobacilli, have shown encouraging effects in prevention of allergy development (Isolauri *et al.*, 2000; Rosenfeldt *et al.*, 2003; Viljanen *et al.*, 2005). An increased prevalence of lactobacilli and lactic acid bacteria was shown in the infant microbiota of anthroposophic children (Alm *et al.*, 2002). Children raised according to the anthroposophic lifestyle, for example Steiner school children, are a low risk group for allergy development.

3.2 Inflammatory bowel disease

Inflammatory bowel disease refers to chronic relapsing disorders of the intestinal tract with yet unknown etiologies. It is the collective term for a group of heterogeneous tissue reactions, with similar clinical behavior. The most familiar subtypes of IBD are CD and UC. CD can affect the whole intestinal tract and is characterized by discontinuous involvement of various portions of the gut, whereas UC is restricted to the colon and rectum and is generally characterized by a continuous inflammation of the large bowel. Heterogeneities in response to medication and clinical behavior within CD and UC have further divided it into distinct disease phenotypes. For example, CD can be classified into a few main phenotypes, with respect to the disease location. The classical phenotype is located in the terminal ileum, i.e. ileal CD (ICD). Inflammation located in the colon is referred to as colonic CD (CCD) and inflammation involving both ileum and colon is addressed as ileocolonic disease (ICCD). Less common but still present is inflammation of the upper intestinal tract in some individuals (Satsangi *et al.*, 2006).

The highest incidence of IBD is in Northern Europe, the UK and North America (Loftus, 2004). An increasing rate of disease incidence during the past few decades has now started to stabilize. Instead, the prevalence in low incidence areas, such as Southern European countries, has started to increase (Baumgart & Carding, 2007). Genetics plays an important role for the disease prevalence with higher concordance rates (both are sick) in

monozygotic twins than in dizygotic twins with CD (approximately 50% versus <10%). For UC the genetic factors are not as strong as for CD, shown with lower concordant rates in both mono- and dizygotic twins (approximately 15% and <5% respectively) (Orholm *et al.*, 2000; Thompson *et al.*, 1996; Tysk *et al.*, 1988). Still, a large number of monozygotic twins are discordant for the disease (one is sick and one healthy), which emphasizes that other factors in addition to genetic predisposition play an important role in disease prevalence. To date, a number of environmental factors have been identified to be important for relapse of the disease. The most evident of these is smoking cigarettes, and this is directly associated with an aggravation of the disease, and acceleration of the need for surgery in patients in remission (Baumgart & Sandborn, 2007). Interestingly, smoking has the opposite effect in UC patients. Other factors, such as diet and childhood infections have been associated with IBD but data are conflicting (Loftus, 2004).

There are several lines of evidence that argue for a microbial involvement in the disease etiology: 1) inflammation occurs in the regions with the highest bacterial densities, 2) diversion of the fecal stream induces clinical remission of active CD, however disease relapses shortly after the bowel continuity is restored (Harper *et al.*, 1983), 3) germ free animals do not develop colitis (Sartor, 2003), 4) a subset of CD patients have polymorphisms in genes that encode receptors that recognize bacterial antigens (Ahmad *et al.*, 2002), 5) antibiotics and probiotics have shown therapeutic efficacy in IBD patients (Sartor, 2004), 6) the epithelial barrier is leaky with defects in mucus production in patients with IBD (Swidsinski *et al.*, 2007), 7) decreased levels of antimicrobial peptides (alpha defensins) have been demonstrated in ileal tissues from a subset of CD patients (Wehkamp *et al.*, 2005) and 8) a large number of studies have reported an altered bacterial composition and higher bacterial loads in IBD patients compared with healthy individuals (Baumgart *et al.*, 2007; Frank *et al.*, 2007; Kleessen *et al.*, 2002; Seksik *et al.*, 2003; Swidsinski *et al.*, 2002; Swidsinski *et al.*, 2005). In addition, the severity of the disease correlates with the bacterial density found in the intestinal mucosa (Swidsinski *et al.*, 2002).

There are two main hypotheses of how microbes can contribute in the development and progression of IBD. The search for a pathogen, responsible for the disease was originally focused on MAP. This bacterium, or traces of it, has been identified in blood and tissues of IBD patients (McFadden *et al.*, 1987; Yoshimura *et al.*, 1987) and furthermore it has been shown to be responsible for Johne's disease, a disease with clinically similar symptoms in cattle (Greenstein, 2003). However, attempts to detect MAP in

CD tissues have resulted in conflicting and inconsistent findings (Chiodini, 1989). Therefore, it was postulated that MAP infection could be responsible for CD in a subset of patients that were susceptible or that this organism may selectively but accidentally colonize the ulcerated mucosa of CD patients but not initiate inflammation. Currently, most recent publications do not specifically point out MAP as a causal agent in CD, however its role in CD is neither excluded (Sartor, 2005).

Several studies have identified a higher prevalence of enterobacteria and in particular *E. coli* in mucosal tissues from IBD patients (Baumgart *et al.*, 2007; Darfeuille-Michaud *et al.*, 1998; Kotlowski *et al.*, 2007; Swidsinski *et al.*, 2002). Most attention as been focused on adherent invasive *E. coli* (AIEC), that was isolated from and within ileal tissues of CD patients (Darfeuille-Michaud *et al.*, 1998). This bacterium has been found to persist and multiply within macrophages (Baumgart *et al.*, 2007). Furthermore, it has been reported almost exclusively localized to ileal tissues from CD patients with the ileal disease phenotype. Moreover, a recent study by Mpofu et al. (2007), showed that supernatants from MAP cultures impaired the ability of adherent monocytes and monocyte-derived macrophages to kill phagocyted *E. coli* (Mpofu *et al.*, 2007) and that could explain the increased prevalence of *E. coli* attached to and within the epithelia.

There have also been reports of an imbalance (dysbiosis) between beneficial and detrimental bacteria in the intestinal microbiota that is correlated to CD (Baumgart *et al.*, 2007; Manichanh *et al.*, 2006; Scanlan *et al.*, 2006; Seksik *et al.*, 2003). The most prevalent observation is a reduction in the diversity of *Firmicutes* in CD patients, with reduced numbers of specific members of the microbiota, mostly within the *Clostridium leptum* (*Clostridium* cluster IV) group (Baumgart *et al.*, 2007; Frank *et al.*, 2007; Manichanh *et al.*, 2006; Scanlan *et al.*, 2006). This group is thought to be functionally important as one of the major producers of SCFA, particularly butyrate, which has been shown to act as a fuel and inflammation repressor for the colonic epithelial cells (Pryde *et al.*, 2002). This observation has been supported by the observation of decreased fecal levels of acetate and butyrate in CD patients (Marchesi *et al.*, 2007; Takaishi *et al.*, 2007).

3.3 Gastric cancer

Gastric cancer is a worldwide problem and one of the leading causes of deaths due to cancer, although its incidence in western societies has been decreasing in recent years (Lochhead & El-Omar, 2007). The major known

risk factor for developing gastric cancer is *H. pylori*. This bacterium by itself is classified as a carcinogen by the Agency for Research into Cancer. However, only a small percentage of the individuals that carry this bacterium will develop cancer, suggesting that other factors, such as host susceptibility, dietary, lifestyle factors or other bacteria residing in the stomach may contribute towards cancer development. Persistent colonization of *H. pylori* can lead to divergent responses. Most carriers will develop mild pangastritis, a condition that does not alter the gastric physiology and without disease symptoms (Smith *et al.*, 2006). However, in some individuals, persistent colonization will progress into a disease state. Antral predominant gastritis is associated with increased production of gastric acid, and an increased risk in developing duodenal ulcer disease, but instead a low risk in developing gastric cancer (Hansson *et al.*, 1996). By contrast, corpus predominant gastritis, leads to a reduced production of gastric acid and increased risk in developing gastric cancer (Uemura *et al.*, 2001). An altered physiology with a reduced capacity for acid secretion allows survival and proliferation of bacteria that are normally not thought to survive the acidic conditions of the stomach. The role of these invaders is not known but it is hypothesized that they have the capacity to convert nitrite to carcinogenic N-nitroso compounds and trigger the release of reactive oxygen species (Blaser & Atherton, 2004). Furthermore, bacterial overgrowth of the stomach and the small bowel can contribute to malabsorption, diarrhea, and possibly to bacterial translocation of the epithelial layer (Husebye, 2005). The bacterial composition of the stomach has been studied in patients using acid-reducing drugs (Adamsson *et al.*, 1999; Mowat *et al.*, 2000; Sanduleanu *et al.*, 2001), but rarely in gastric cancer patients (Sjöstedt *et al.*, 1988), and those studies have been restricted to detection of the culturable fraction of the stomach microbiota. The current knowledge of the gastric microbiota in patients using acid reducing drugs has revealed a predominance of the species normally found in the upper respiratory tract.

3.4 Colorectal cancer

A western diet, comprising a high intake of fat and animal protein but low fiber content is associated with a higher risk in developing colon cancer, compared to a Japanese diet and vegan diet which are associated with a low risk of developing colon cancer (Hayashi 2002). However, Japanese who adopt a westernized diet, develop colon cancer more frequently (Finegold 1978). Case control studies have shown that dietary fat and high consumption of red meat, in particular processed meat, are associated with an increased risk of developing colorectal cancer (CRC; Guarner 2006). In

contrast, high intake of fruits and vegetables and whole grain cereals, fish and calcium is associated with a decreased risk. Meat is a rich source of sulphur, which promotes growth of sulphate reducing bacteria (SRB). This group of bacteria produces hydrogen sulphide, a toxic compound that can suppress butyrate utilization, inhibit the synthesis of mucus and methylation of DNA (O'Keefe, 2008). Gut bacteria exhibit a number of characteristics that can contribute to development of CRC. For example, they are able to produce genotoxic compounds that can activate pro-carcinogens to DNA reactive agents. Animal models have shown that colon cancer not was developed in germ free animals whereas 70% of the animals in the same model with a gut microbiota developed colonic tumors (Kado *et al.*, 2001). A number of bacterial species have been associated with CRC but without any consistency. For example, increased numbers of *Bacteroides* have been associated with an increased risk for CRC (Hill *et al.*, 1971; Moore & Moore, 1995). Also, increased numbers of both *Clostridium* and *Lactobacillus* have been found in individuals with CRC (Kanazawa *et al.*, 1996). By contrast, *Lactobacillus* spp have been associated with a reduced risk (Moore & Moore, 1995). *Streptococcus bovis* have been implicated in colonic neoplasia, and antigens extracted from its cell wall have provoked proinflammatory responses in animal models (Biarc *et al.*, 2004; Ellmerich *et al.*, 2000).

4 Methods to study the gut microbiota

4.1 Cultivation versus Molecular approaches

Traditionally, microbes in the gastrointestinal tract have been studied using cultivation-based techniques. The great advantage with cultivation is that isolates can be recovered and further studied for their ability to utilize different substrates and other physiological parameters. However, cultivation suffers from several drawbacks. It is very labor intensive and does not give a representative picture of the intestinal microbiota. Only a small proportion (approximately 20 %) of the microbes residing within our intestinal system can be cultivated with the approaches currently available (Eckburg *et al.*, 2005). This number is even lower for other ecological niches, such as soil, where it is estimated that only 0.1-10% of the bacteria can be cultivated (Torsvik *et al.*, 1998). In recent decades development of molecular approaches has enabled us to bypass the necessity for cultivation, thereby opening a new window to explore the world of microbes, however, mostly viewed behind the screen of a computer. The most common methods rely upon analysis of the prokaryotic 16S ribosomal RNA gene that encodes part of the small subunit of the ribosome. This gene, approximately 1.5 kb long, is present in all prokaryotes in various copy numbers and has enough sequence conservation for accurate alignments and at the same time enough sequence variability to detect differences between different species. This has therefore been used as a tool to determine the phylogenetic relationships between different microorganisms. The widespread use of the 16S rRNA gene in microbial ecology has provided an extensive public DNA database, including close to 500,000 16S rRNA gene sequences. The sequence database is growing with a tremendous speed.

4.2 Microbial community fingerprinting

Microbial community fingerprinting is the collective term for a number of molecular methods that have recently been developed to study the composition of microbial communities in different environments. Unifying for all molecular fingerprinting approaches is that they rely on multi-template PCR reactions for assessment of microbial community structures, or "fingerprints". Depending on the approach used, they have potential for high throughput, are relatively cheap and not too laborious. However some methods lack resolution and are limited in their identification abilities. Nevertheless, these approaches are very useful for studying the dynamics of microbial communities due to different treatment or temporal effects or for comparisons between groups of individuals. These fingerprinting methods have been used extensively in the field of environmental microbial ecology and more recently to provide essential knowledge of factors contributing to the dynamics and composition of the gut microbiota.

4.2.1 DGGE/TGGE

One of the most common fingerprinting approaches to assess the structure of the intestinal microbiota is denaturing gradient gel electrophoresis (DGGE), or temperature gradient gel electrophoresis (TGGE). These methods separate multitemplate PCR products as bands on gels according to GC content, dependent on melting behaviors of the amplicons as they migrate through the gels (Figure 3). The amplicons become denatured according to their GC content during gel migration. In DGGE, the denaturing agent is commonly a mixture of urea and formamide, while TGGE uses a temperature gradient. The resulting banding patterns on gels can be compared visually on the gel or analyzed by cluster analysis. One advantage of this approach is that bands of interest can be excised and sequenced to obtain information about the species they represent. Muyzer et al. (1993), was the first to use DGGE to obtain fingerprints of complex microbial communities in soil samples (Muyzer *et al.*, 1993). More recently this method has been used to monitor the temporal stability of the microbiota in health and disease (Scanlan *et al.*, 2006; Zoetendal *et al.*, 1998), including impacts of diet, antibiotics and probiotics (McCracken *et al.*, 2001; Vanhoutte *et al.*, 2006). Furthermore, it has been used to assess the microbial composition at some other sites of the body, such as the stomach (Monstein *et al.*, 2000) and mouth (Zijnge *et al.*, 2006).

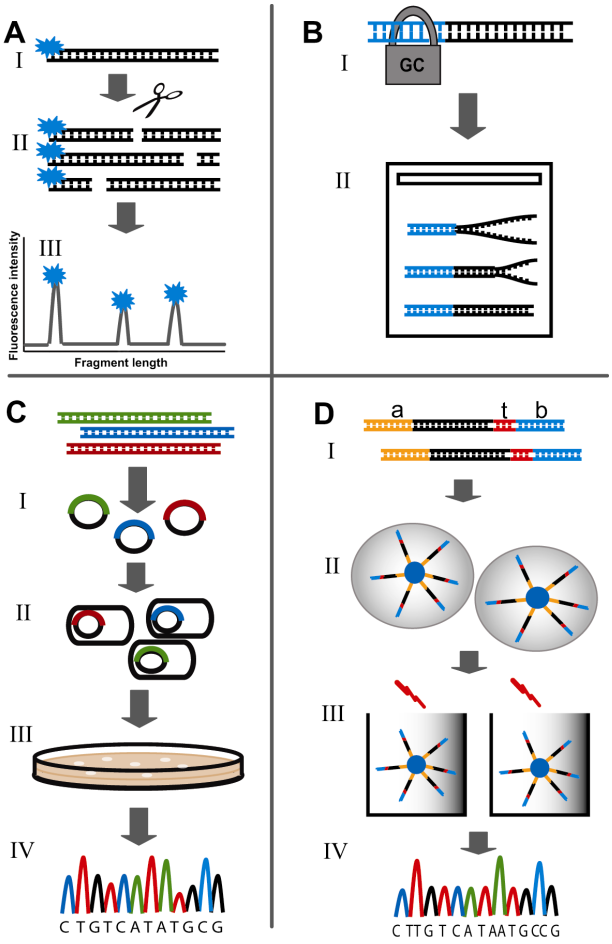

Figure 3. Flowchart of four PCR-based methods for microbial community fingerprinting, **A**) T-RFLP, **B**) DGGE/TGGE, **C**) Cloning and sequencing, **D**) 454-FLX pyrosequencing. (**A**) In T-RFLP, PCR amplification with a fluorescently labeled primer (I) and restriction digestion of the amplicons provides a restriction pattern (II) where only the terminal labeled fragments will be detected during electrophoresis and provides a community fingerprint (III). (**B**) In DGGE/TGGE, a primer tagged with a GC clamp (I) prevents complete denaturation of amplicons in gel electrophoresis (II) where amplicons are separated according to GC content during migration against a thermal (TGGE) or chemical gradient (DGGE). (**C**) In cloning and sequencing, PCR amplicons are cloned into plasmid vectors (I) and transformed into bacteria (II) that are multiplied on agar plates (III). Plasmids are recovered from the clones and inserted fragments are sequenced (IV). (**D**) In 454-FLX pyrosequencing, the amplicon is tagged with adaptors, (a and b). In addition, a sample specific barcode (t) allows tracking of each read back to its original sample (I). The fragments are bound, though the adaptor sequences, to oligo-nucleotide coated beads, favoring one fragment per bead. Fragments are emulsified in a micro reactor (II) and the amplified DNA is sequenced using pyrosequencing (III), which is performed by flowing nucleotides over wells and measuring light emission (III) that gives rise to a sequence (IV).

4.2.2 T-RFLP

Terminal-restriction fragment length polymorphism (T-RFLP) is another powerful fingerprinting approach that has been widely used in microbial ecology (Jernberg *et al.*, 2007; Kitts, 2001; Liu *et al.*, 1997; Wang *et al.*, 2008)(**Papers I-IV**). In this method the PCR product is end labeled with a fluorescent dye. Restriction digestion of the amplicons provides a restriction pattern, but only the terminal labeled fragment will be detected during electrophoresis based on its fluorescence that is detected by an automated sequencer. Sequence polymorphisms between species provide heterogeneous terminal fragment lengths (TRFs) and their corresponding lengths (base pairs) are determined based on a standard (Figure 3). The intensity of the fluorescence will also be registered and reflects a quantitative measure of abundances of the individual TRFs. However, this quantitative measure should only be considered as relative quantitation, i.e. relative abundance. The output raw data from T-RFLP analysis is visualized in an electropherogram. The electropherograms can be compared visually between samples, however with increased risk for misinterpretations of the data due to variations in running conditions, amplicon amounts, etc. (Hayashi *et al.*, 2002; Sakamoto *et al.*, 2003). It is more appropriate to normalize the data by applying size and abundance threshold values. Relative abundances of individual TRFs can be calculated as percent areas by summing the total TRF peak area and dividing each individual TRF by the sum of TRFs. In this way the comparisons between samples are not restricted by the amount loaded on the gel or into the capillary. In addition, by constructing a consensus profile of technical replicates, false positive peaks can be excluded and biases introduced by PCR decreased. T-RFLP can be used to visualize diversity (of TRFs or of dominant populations; see below), structure and dynamics of the microbiota in a statistically adequate manner. The method has been shown to be highly reproducible (Lueders & Friedrich, 2003; Osborn *et al.*, 2000). T-RFLP has been successfully used, for example, to monitor changes in the microbiota due to antibiotic administration (Jernberg *et al.*, 2007), to assess the diversity of fecal microbial communities in allergic infants (Wang *et al.*, 2008) and the community structure of the oral microbiota (Sakamoto *et al.*, 2005). In addition, it has constituted a methodological backbone in this thesis, used in all **Papers** (**I-IV**). One weakness of T-RFLP is that the data does not rely on identification alone since the direct link between TRFs and sequence data is lacking. Another disadvantage is that one TRF can potentially originate from several phylogenetically unrelated species. This method is more suitable for rapid comparisons of microbial community compositions in different samples. If the goal is to still try to identify the species corresponding to a specific TRF there are usually two options: 1) to use

information from multiple enzyme digestions in parallel and to compare the terminal restriction patterns from the different enzymes to a database (Edlund *et al.*, 2006; Marsh *et al.*, 2000). This approach introduces some uncertainties due to structural differences between the fluorophors attached to the internal standard and the sample DNA fragments that may introduce a difference in the true and observed fragment lengths (Kaplan & Kitts, 2003). 2) Another option is to create a clone library to find sequences corresponding to specific TRF lengths when digested *in silico*. In addition, T-RFLP analysis of the clones in the library can first be used to find out which clones should be sequenced to match a specific TRF of interest in a T-RFLP profile (**Papers II** and **III**). However representatives of TRFs with low abundances are usually hard to track in clone libraries, due to the requirements of screening many clones to detect rare sequences.

4.2.3 Cloning and sequencing

Cloning and sequencing of 16S rRNA genes, (Figure 3), yields a lot of information, depending on the number of clones sequenced, but is time consuming and associated with high costs. The usual dilemma in cloning surveys is to select between high resolution allowing in depth analysis from a few samples, or lower resolution allowing more samples to be included in the analysis. A few large scale-cloning surveys have provided valuable information of the microbial diversity at different anatomical sites of the body (Aas *et al.*, 2005; Bik *et al.*, 2006; Eckburg *et al.*, 2005; Pei *et al.*, 2004).

4.2.4 454 pyrosequencing

Recent methodological breakthroughs have enabled large scale cloning surveys that permit in depth analysis of each sample as well as a large set of samples. Recently, 454 pyrosequencing has been proposed as a new powerful technique for assessment of complex microbial communities (Margulies 2005). In this technique, DNA fragments from e.g. a multitemplate PCR are bound to beads under conditions that favor one fragment per bead. A PCR amplification step occurs in emulsified micro reactors surrounding the beads (Figure 3). The amplified DNA can then be sequenced using the light emission technique pyrosequencing (Ronaghi *et al.*, 1998). This approach provides a tremendous number of DNA sequences, usually several hundred thousand sequences per run. The read length using this type of technology was originally very short (approximately 50-100 bp), but recent improvements (i.e. 454-FLX) have increased the read length and it is now possible to obtain approximately 200-500 base pair long sequence reads. By using a sample specific barcode on one of the primers, each read can be tracked back to its original sample. Using this

approach several samples can be run in parallel on a 454 picotiter plate. This technique is expected to be more widely applied for many types of environmental samples, and we were the first to our knowledge to use the 454-FLX method for the intestinal microbiota (manuscript in preparation). Although this approach is not subjected to cloning bias it still has the biases introduced by PCR. In addition, this is currently a very expensive approach due to the associated running costs and it is challenging because the large amounts of data require extensive bioinformatics skills and processing. Despite these disadvantages, many microbial ecologists feel that this sort of technique will eventually surplace many of the other fingerprinting methods currently in use, in particular when costs eventually decrease.

4.3 FISH

Direct quantification of the number of bacteria in a sample can be performed by Fluorescent In Situ Hybridization (FISH). This method relies on detection of bacteria using fluorescently labeled oligonucleotide probes. The fluorescently labeled bacteria are usually detected by epiflourescent microscopy, confocal laser microscopy or flow cytometry (Zoetendal *et al.*, 2004). Depending on the selection of probes, FISH can be used to detect bacteria at different phylogenetic levels. It is not based on PCR and thereby free from biases introduced by PCR or DNA recovery. It is however dependent on reliable sequence data, and laborious if identification at the species level is required. In addition, only a few probes can be used per analysis. This approach has been used in several studies to quantify specific groups of bacteria in both fecal samples as well as in biopsies (Mueller *et al.*, 2006; Swidsinski *et al.*, 2002).

4.4 Quantitative PCR

Another way for direct quantification of the number of bacteria in a sample is to use quantitative PCR (qPCR). This method uses a chemiluminscent or fluorescent reaction to determine the kinetics of product accumulation during PCR amplification with specific primers for a specific group or species of bacteria. It is then possible to use the product accumulation rate curves to back calculate to the number of original target molecules in a sample. This method is limited by the specificity of the primers used during PCR and has mainly been used to detect and quantify bacteria at the genus or phylum level. The 16S rRNA gene is commonly used as a target molecule however, due to heterogeneous gene copy numbers in some species, quantification based on 16S rRNA gene copies can be misleading. One way to get around this problem is to amplify other genes that are

present in single copies per genome, such as some "housekeeping genes" or to amplify a functional gene specific for a given group or species to increase specificity (**Paper III**).

4.5 G+C fractionation

Percent guanine + cytosine (%G+C) profiling is another approach for assessment of complex microbial community structures, such as that found in the gastrointestinal tract (**Paper II**). This method relies on separation of bacterial chromosomes in a density gradient based on differences in their %G+C contents, imposed by an AT-dependent DNA-binding dye (Apajalahti *et al.*, 1998). With a differential ultracentrifugation step, the chromosomal DNA can be fractionated and the %G+C content, represented by each gradient fraction, determined based on DNA standards with known %G+C composition. The advantage with this method is that it is independent of prior sequence information (e.g. primer and probe sequences) and thus free of biases such as those related to PCR and cloning. Furthermore, the %G+C fractionation offers the opportunity to further analyze bacterial genomes with a defined interval of GC content (Kassinen *et al.*, 2007). The disadvantage is a poor resolution and no possibility to identify community members, without additional PCR or cloning steps.

4.6 Method considerations

As in all analyses there are a number of limitations and biases along the road that are important to keep in mind when it comes to interpretation of data. A good study design is of great importance to be able to draw powerful conclusions. For example, the sample type, number of samples, methodological approach, environmental data, etc. are factors that could be highly important to take into consideration before starting a study. For studies of the human gut, it is valuable to have knowledge about external or host related factors that might affect the gut microbial composition and to use this information when designing studies to reduce the number of confounding factors when possible.

4.6.1 Sampling

Sample origins and types should be carefully considered depending on the scientific question to be addressed. For example, some studies have shown differences in the luminal and the mucosa-associated microbiota in the gut (Eckburg *et al.*, 2005; Lepage *et al.*, 2005; Sanduleanu *et al.*, 2001; Zoetendal *et al.*, 2002). As mentioned above, in some studies the fecal microbiota has been shown to differ from the colonic microbiota found on mucosal

biopsies. When studying intestinal diseases, it is often stated that the mucosal microbiota are more relevant to study due to their closer interaction and ability to crosstalk with the host epithelial cell layer. However, mucosal samples are harder to obtain than fecal samples and sampling is invasive to the human host. By contrast, fecal samples are non- invasive, easier to collect, and thereby more feasible for use in dynamic studies of the microbiota. Another consideration to take into account is that biopsy collection is usually preceded by a washout of the bowel. However, this procedure might introduce a biased view the mucosa-associated community when comparing with a fecal sample or an unwashed biopsy. In one study, it was shown that the fecal microbiota differed from its original composition after colonoscopy in three out of five patients. This effect was observed 6-8 weeks after colonoscopy and it is not known how long this disruption of the fecal microbiota lasts (Mai *et al.*, 2006).

4.6.2 DNA extraction

One of the bottlenecks that is of critical importance in molecular studies of the human gut is the extraction and purification of DNA from the fecal and biopsy samples. There are numerous extraction protocols available to purify DNA from different types of matrixes. Several different extraction methods have been compared (McOrist *et al.*, 2002; Scupham *et al.*, 2007), but without a consensus result. Use of bead beating is generally thought to be most suitable for efficient lysis of Gram-positive microbes (Zoetendal *et al.*, 2001b). However, the bead beating time needs to be optimized so that it is long enough to disrupt the cell wall of hard to lyse cells, but not too long so as to shear DNA. In addition, it is of critical importance to have a good DNA quality for downstream analyses, without PCR inhibiting agents, which are common in fecal samples. Finally, sample replication is important to ensure reproducibility of the extraction method. Therefore, sample extractions are commonly performed in two or more replicates.

4.6.3 PCR amplification

There are a number of biases and limitations associated with multitemplate PCR reactions. The first critical step is selection of PCR primers. The constantly expanding databases of 16S rRNA genes have revealed that primers considered as "universal" exclude many sequences. The specificity of the primers also deserves careful attention. For example, when mucosal biopsies are studied the bacterial universal primers should not target or bind to the human genome. Furthermore, the numbers of PCR cycles has been shown to reflect the bacterial diversity in clone libraries, with a higher diversity correlating with a lower cycle number (Bonnet *et al.*, 2002). In

addition, the template to product ratio can be affected by the number of PCR cycles (Polz & Cavanaugh, 1998). PCR based approaches will not reveal any information about whether the bacteria that are detected are alive or dead. In one study it was shown that only two thirds of the cells recovered from feces are alive and the rest are dead (Ben-Amor *et al.*, 2005).

One must keep in mind that despite all the technical biases and limitations these approaches might introduce, as long as the same reagents and conditions are used for the sample sets, that molecular fingerprinting approaches are very powerful, reproducible and rapid tools for determination of structural changes in microbial communities. For example, T-RFLP was found to be highly reproducible and in direct proportion to the abundance of the sample template (Clement & Kitts, 2000; Dunbar *et al.*, 2000).

4.7 Statistical approaches in community analyses

4.7.1 Multivariate data analysis

The use of molecular methods for analyzing bacterial communities from different environments has increased extensively during the past decade. These methods provide large sets of data that are commonly interpreted with multivariate statistical methods. There is no single standard method for examining differences in microbial community structure. In general, data from fingerprinting approaches are analyzed by ordination approaches, such as principal component analysis (PCA; **Papers I** and **II**; Jernberg *et al.*, 2005; Kaplan & Kitts, 2004; Wang *et al.*, 2004), correspondence analysis (CA; Edlund *et al.*, 2006), and non-metric multidimensional scaling (nMDS; Rees *et al.*, 2004). The main use of these multivariate methods is to reduce the complexity in the dataset and to identify patterns that unify samples. For example, two samples with high agreements in their species compositions will group tightly together whereas two samples that have a dissimilar species composition will be separated in ordination space. Statistical significance of differences in species composition between clusters from ordination is usually assessed with permutation tests (**Papers II** and **IV**). In regular linear models, the reference distribution for hypothesis testing can be derived mathematically from the assumptions of the test. In a permutation test, the reference distribution is instead determined from the obtained data by a repeated shuffling of the samples (Legendre & Legendre, 1998).

Cluster analysis is also commonly used for data generated from molecular fingerprints (Blackwood *et al.*, 2003; Kitts, 2001; Sakata *et al.*, 2005). In

cluster analysis, pair-wise comparisons of samples are calculated as distances and sorted according to the calculated value of the distance. There are several different ways to describe the similarity of a pair of samples, and thereby a large number of distance coefficients that can be used for these calculations (Legendre & Legendre, 1998). Jaccard, Dice or Sorensen coefficients are appropriate for presence/absence analysis, but ignore species abundance (**Papers I** and **II**). Euclidean distance, the distance coefficient in PCA, is not considered to be appropriate for datasets with a large number of zeroes (Legendre & Legendre, 1998), which is commonly encountered in T-RFLP analysis. Instead, the Bray Curtis index is more suitable and is thought to give a good representative picture of the structural differences of the data (**Papers III** and **IV**).

The use of multivariate tools for interpretation of large sets of data has become invaluable to identify gradients or subgroups within a sample cohort. However, the complex nature of biological data and differences in sampling designs should be taken into account prior to selection of the analysis method. There is a danger in only relying on the observation from a multivariate statistical approach without further confirmation. Different methods sometimes give inconsistent results because they are used incorrectly.

4.7.2 Diversity indices

It is generally thought that a high diversity, including microbial diversity, is beneficial for an ecosystem providing a higher resilience to ecological disturbances (Bäckhed *et al.*, 2005). Diversity is typically described in terms of species richness and evenness. Richness reflects the number of species in a system but does not take their relative proportions into consideration. Evenness reveals the distributional composition of species in a community by including their relative abundances. There are a number of different indices to describe diversity. Simpson's index of diversity (D) is commonly used, and incorporates both species evenness and richness (Legendre & Legendre, 1998). Another index to assess diversity is Shannon's index, (Legendre & Legendre, 1998) which is divided into a richness index (H) and equitability index (E). Microbial diversity assessment based on T-RFLP or DGGE/TGGE analysis is not a measure of "true" diversity, because only the most dominant members of the community are assessed. This must therefore be taken into account when these tools are used for estimations of microbial diversity. However, these methods can provide a relative diversity measure for comparisons between samples and serve as a diversity indicator.

5 Results obtained during the project

5.1 Lifestyle influences of the microbiota

Some lifestyles, such as an anthroposophic lifestyle and living on a farm are associated with a low risk in developing asthma and allergies (Alm *et al.*, 1999; Braun-Fahrländer *et al.*, 2002). Certain characteristics of the anthroposophic lifestyle, such as a restrictive use of antibiotics, antipyretics and vaccinations, have been epidemiologically associated with a reduced risk for allergies (Flöistrup *et al.*, 2006). This lifestyle also includes certain dietary habits, such as a high intake of organic and biodynamic foods that often include naturally fermented vegetables. It was previously shown that infants in anthroposophic families had a higher prevalence of lactic acid bacteria than other infants (Alm *et al.*, 2002). Therefore, one aim of this study was to investigate if an anthroposophic lifestyle impacted the intestinal microbiota in children. Another interesting group to study is children raised on farms because these children are consistently exposed to elevated levels of endotoxins and mould components (Schram-Bijkerk *et al.*, 2005). In addition, the farming lifestyle includes a number of features that may influence the microbial composition in the gut, such as a diet often including unpasteurized dairy products and exposure to animals and their stools. In **Paper I**, the fecal microbiota was studied in 90 children from anthroposophic families, selected from their attendance in Steiner schools that subscribe to this lifestyle philosophy, or children growing up on farms. In addition, control groups for the respective lifestyles were included, i.e., living in a similar environment but not exhibiting these lifestyles. The samples derived from an epidemiological multicenter study, PARSIFAL (Alfvén *et al.*, 2006), with the aims to identify factors in the anthroposophic and farming lifestyle that contribute to the protection against allergy

development. Fecal microbial profiles were assessed using T-RFLP for investigation of the microbial compositions in these groups.

5.1.1 Microbial diversity associated with lifestyle factors

Multivariate statistics revealed no significant differences in the gut microbial communities between the groups of children. Also, it was not possible to correlate the TRF profiles to any of the explanatory variables, such as lifestyle characteristics, age, gender or geographical origin, provided from the questionnaire data. However, when the bacterial diversity was calculated, based on the richness and evenness of TRFs in the T-RFLP data, significant differences were obtained in comparisons between groups. Most prominent was a higher bacterial diversity in children from anthroposophic families (Steiner school children) and a lower diversity in farm children ($P = 0.0001$; Figure 4). Several characteristics associated with the respective lifestyles correlated with the observed differences in diversity (Figure 4). However, it was not possible to identify a specific factor responsible for a particular increase or decrease in diversity.

Some general correlations to particular factors can however be made. For example, antibiotics are known to have a large potential to disrupt the microbial composition in the gut, and we found that the children that never used antibiotics at all had a higher microbial diversity than the ones that had taken antibiotics. In addition, the dietary regimes in general differ largely between these groups, as mentioned above. These dietary discrepancies may also contribute to the measured differences in gut bacterial diversity between these lifestyles. Interestingly, the factors that correlated with the highest and lowest bacterial diversity have epidemiologically been identified to be protective against allergy and asthma development (Alfvén *et al.*, 2006; Flöistrup *et al.*, 2006).

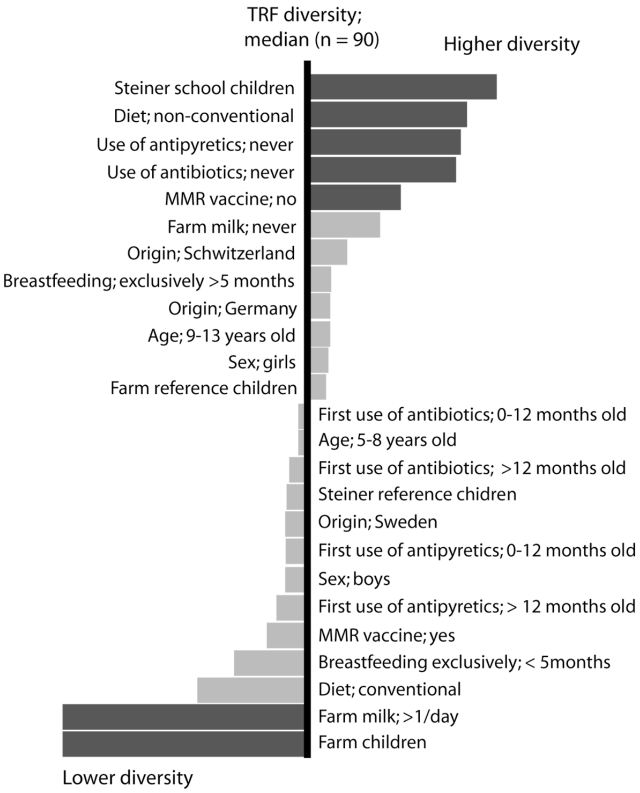

Figure 4. Bacterial diversity based on T-RFLP data in fecal samples collected from 90 children raised according to different lifestyles. The line in the center represents median diversity for all sampled individuals. The size of each bar represents the deviation of the median value for each factor in comparison with the median for the whole sample set. Factors associated with specific lifestyles (i.e. anthroposophic, respective farming lifestyle) are highlighted in dark grey.

5.1.2 Impacts of microbial diversity

In general, it is considered that a high diversity is beneficial for an ecosystem, enabling the community with a greater ability to respond to disturbances (Bäckhed *et al.*, 2005). One can hypothesize that a higher microbial diversity in infants would result in a broader range of immune triggering compounds that would be beneficial in shaping the immune system. In support of this hypothesis, a reduced diversity was observed in the fecal microbiota of one-week-old infants with atopic eczema compared with non-atopic infants (Wang *et al.*, 2008). Interestingly, in **Paper I**, the lowest measured diversity correlated with the farming lifestyle and associated factors, which is thought to be protective against allergy development. This

result suggests that the microbial diversity is not crucial, as long as the microbial consortia present in the intestinal tract are able to retain ecosystem function. Another possibility is that the microbial diversity may play a more crucial role during the first weeks of life. Kalliomäki et al. (2001), reported differences in the microbiota of three weeks old atopic and non-atopic infants. However, these differences could not be reproduced when the infants were three months old and in **Paper I** the children were older (5-13).

Despite inter-individual differences in the microbiota between individuals analyzed in **Paper I**, some TRFs were common for all subjects. For example TRFs 223 and 272 were present in 89/90 and 90/90 subjects respectively. In **Paper I**, we postulated that the bacteria represented by these TRFs could represent core residents of the intestinal microbiota. In addition, these TRFs had higher abundances in the farm children. The attempts to identify the bacterial species that corresponded to these TRF sizes relied only on the existing sequences deposited in public databases and were not conclusive. However, in **Paper III**, correlation between TRF data and sequence data from mucosal biopsies indicated that these TRFs most likely corresponded to the *Firmicutes* phylum, and in particular TRF 223 was specifically correlated to sequences related to *Faecalibacterium prausnitzii*. This bacterium is a producer of butyrate and could represent a core bacterial resident of the intestinal tract. This finding supports the idea that keystone species may obviate the need for a high microbial diversity (Ley *et al.*, 2006a).

5.1.3 Lactic acid bacteria–like microbiota

The dietary habits associated with the anthroposophic lifestyle include naturally fermented foods preserved with lactobacilli. Studies of probiotics including lactobacilli have shown promising effects of these bacteria in prevention of allergies (Abrahamsson *et al.*, 2007; Kukkonen *et al.*, 2007). Therefore, in **Paper I** primers targeting lactic acid bacteria were used in the T-RFLP approach. In eight out of the 90 samples, it was not possible to obtain a PCR product, presumably as the LAB were under the detection limit in those samples. Interestingly, none of these negative samples derived from the Steiner school children. The obtained profiles for the LAB-like community divided the individuals into two groups, dependent on different dominant TRFs. In 30% of the individuals a TRF corresponding to *Eubacterium biforme* or *E. cylindroides* dominated the profile. By contrast, 70% of individuals were dominated by a TRF that correlated with a previously uncultured bacterial clone derived from the intestinal tract. Interestingly, individuals with both TRFs were negligible. The divergent prevalence of

these TRFs was not possible to correlate to specific lifestyle factors, yet, with one exception, the presence of the TRFs corresponding to the eubacterium clones were more prevalent among the farm children and consumers of farm milk. The biological significance of the prevalence of different species, targeted using this approach is however not known. Although lactobacilli were detected the TRFs that correlated to *Lactobacillus* spp. were not among the most dominating ones, suggesting that the lactobacilli were present in lower abundances in the general community.

5.2 The microbiota of twins with Crohn's disease

Genetic factors are important for influencing susceptibility for development of CD. Several susceptibility loci have been identified in the human genome (Weersma *et al.*, 2007). Importantly, studies of identical twins have revealed that although host genetics is important, other factors are also involved in CD development, such as stimuli from gut microbes. Because identical twins have an identical genetic makeup, and often comparable dietary and lifestyle habits, they are an model system for dissection of other factors that can influence CD. In particular, study of twins that are discordant for the disease (one is healthy and one has CD), provides a means to untangle the respective contributions of genetics from other factors, since the twin pairs provide each other's genetically matched control. Numerous studies have investigated the composition and dynamics of the intestinal microbiota in CD, but until now, this has not been studied in identical twins.

In **Papers II** and **III**, the fecal and mucosal microbiota (5 biopsy locations, ranging from terminal ileum to rectum) of ten twin pairs were explored with different molecular approaches to identify microbial components of importance in CD. Six of these pairs were discordant, and four concordant for CD. In addition, fecal samples from eight healthy twin pairs (5 monozygotic and 3 dizygotic) were included to study similarities in the microbiota within healthy pairs (**Paper II**). The main approach used to assess the microbial population structure was T-RFLP using both general and group specific primers (**Papers II** and **III**). In addition, the results were verified with quantitative PCR (**Paper III**) and in a subset of samples, with %G+C profiling (**Paper II**) and conventional cloning (**Papers II** and **III**). The new and promising technique, 454-FLX pyrosequencing, was also used to assess the fecal microbiota from all discordant and concordant twin pairs with CD and the mucosal microbiota from two biopsy locations, from the majority of the twins (ileum; samples from 16 individuals, and descending colon; samples from 14 individuals). In total, 136,848 16S rRNA gene

sequences were recovered and classified against the databases, (RDP II classifier), with a discrepancy at 95% identity (manuscript in preparation).

5.2.1 Differences in the microbiota within twin pairs

In **Paper II**, it was shown that healthy twins have high similarities in their fecal microbial profiles (Figure 5) and this finding is consistent with the published literature (Stewart *et al.*, 2005; Van de Merwe *et al.*, 1983; Zoetendal *et al.*, 2001a). By contrast, comparison of the fecal microbiota of the discordant twin pairs revealed large differences between the healthy and the sick individuals within any given pair (Figure 5). In addition, the concordant CD twin pairs, did not have similar microbial profiles (Figure 5) and the observed differences were in the same range as for the discordant pairs. Furthermore, analysis of the mucosal microbiota in the biopsies supported the observations from feces and showed a large disagreement in the microbiota within discordant and concordant twin pairs. (Figure 5, **Paper III**). As previously reported (Lepage *et al.*, 2005; Zoetendal *et al.*, 2002) there was a high consistency of the microbial profiles in biopsies collected at different locations within an individual, and in **Paper III**, we found that this was true regardless of the disease state (Figure 5).

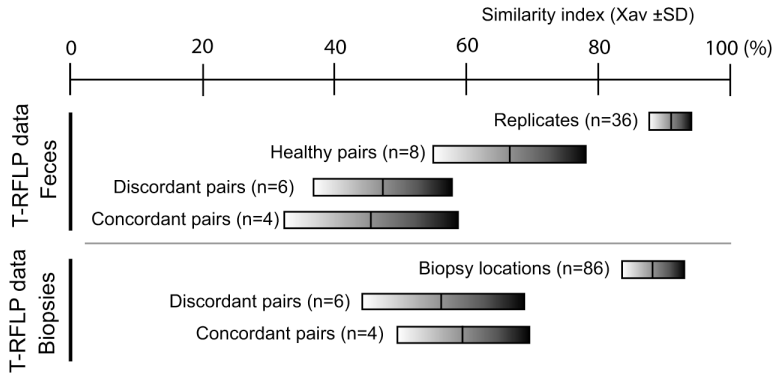

Figure 5. Similarity scores for pair wise comparisons of microbial community profiles derived from T-RFLP data (feces and biopsies). Reproducibility of the T-RFLP method (from duplicated fecal samples), is indicated by "replicates" and averages with standard deviations are represented in the upper box. Similarities of T-RFLP profiles derived from different locations in the terminal ileum and large bowel are indicated "biopsy location". In addition, similarity scores were calculated for healthy, discordant and concordant twins. The data originates from pair wise comparisons of Manhattan distances. The boxes represent the mean and standard deviations for all comparisons made within a group. n; represents the number of pair wise comparisons performed. ICD; ileal Crohn's disease, CCD; colonic Crohn's disease.

5.2.2 Clustering according to the disease phenotype

Multivariate data analysis revealed that there was a subset of the CD patients that grouped separately from the others. This clustering was correlated to the disease phenotype rather than the genetic relatedness. The CD patients with the ileal disease phenotype clustered together clearly separated from the healthy individuals, whereas the group with the disease located in the colon were not significantly different from the healthy individuals (**Papers II** and **III**). This result was consistent along the length of the large bowel as well as in feces. In addition, this result was consistent for the different analytical approaches used (Figure 6). Both T-RFLP and 454-FLX pyrosequencing differentiate the bacterial communities in patients with ICD. By contrast, the CCD patients cluster with the healthy individuals suggesting that the microbiota is impacted more in the ileal disease phenotype than in the colonic phenotype. Several other reports have only been able to demonstrate an imbalance in a subset of CD patients (Baumgart *et al.*, 2007; Frank *et al.*, 2007). Due to the heterogeneity of this disease, the current findings lend support to the hypothesis that the composition of the gut microbiota may play a more important role in the ileal disease phenotype. There are a number of other findings that support this idea. Mutations in the NOD2/CARD15 region, encoding receptors for bacterial recognition, are correlated to the ileal disease phenotype (Ahmad *et al.*, 2002). Furthermore, deficiency in the production of alpha defensins produced by Paneth cells in the ileum correlates with the ileal disease phenotype, and mainly with patients that have mutations in NOD2/CARD15 region (Wehkamp *et al.*, 2005). In addition, the ileal disease phenotype has been associated with higher levels of ileum associated *E. coli* with invasive properties (Baumgart *et al.*, 2007; Darfeuille-Michaud *et al.*, 1998).

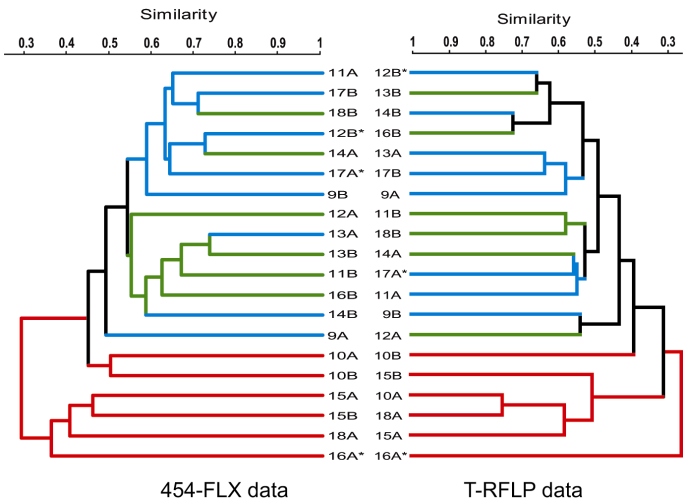

Figure 6. Clustering of microbial profiles according to the species composition in fecal samples originated from discordant and concordant twins. Bray Curtis index was used as a distance matrix and the dendrogram was based on UPGMA clustering representing data from T-RFLP analysis and 454-FLX. The numbers represent the twin pairs and the letters (A or B) represent the individuals within a pair. Coloring of branches: red, ICD; blue, CCD; and green, healthy individuals. *Represent individuals that originally were classified as ileocolonic disease phenotype (**Paper II**), but were reclassified as predominantly having ICD or CCD (**Paper III**).

5.2.3 Microbial diversity in Crohn's disease

The TRF diversity was higher among the healthy individuals compared to CD patients. In addition, among the discordant twin pairs, the healthy twin consistently had the higher diversity index (**Papers II** and **III**). Diversity measures based on T-RFLP data are however restricted to the dominant community members. The same diversity index was used to calculate the microbial diversity based on the 454-FLX data. This method allows a more in depth analysis of the microbial diversity. Interestingly, the results obtained with the T-RFLP data were in good agreement with the data obtained with 454-FLX. From the 454-FLX data, it was shown that the diversity among the *Firmicutes* was reduced in the ICD patients but not in CCD patients compared to healthy individuals (Figure 7). A previous study showed that the fecal microbiota from six CD patients revealed a clear reduction in diversity within the *Firmicutes* phylum when a metagenomic library was compared between healthy individuals and CD patients (Manichanh *et al.*, 2006). Most prominent was a depletion of members of the *Clostridium leptum*

group (*Clostridium* cluster IV) (Manichanh *et al.*, 2006)(Manichanh 2005). Strikingly, all of these patients had CD with ileal involvement.

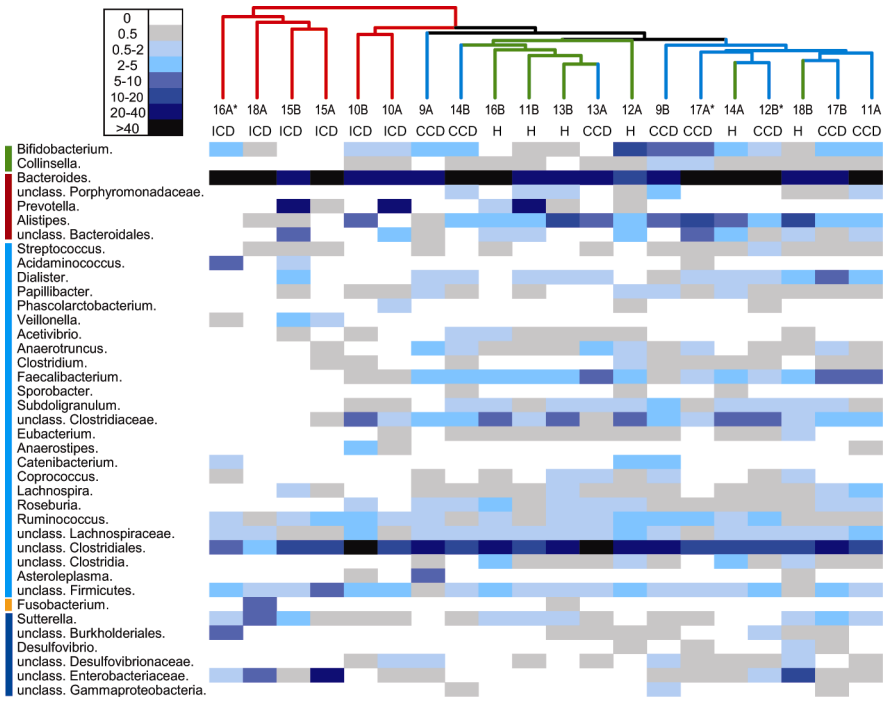

Figure 7. Relative abundances of sequences obtained using the 454-FLX approach that were classified to the genus level. Explanation of the abundance intervals is shown in a progressive blue scale with higher abundances correlating with darker colors. The abundance profile for each sample represents a column and the abundance for each genus as a row. Color scale to the left represents which phylum the genus belongs to: green, *Actinobacteria*; red, *Bacteroidetes*; light blue, *Firmicutes*; orange, *Fusobacteria*; Dark blue, *Proteobacteria*. To reduce space of the figure, genera that did not have any representatives with abundances over 0.5 % in any of the samples were removed. The samples were sorted according to their clustering with Bray Curtis index as a distance matrix and UPGMA as a dendrogram. Branch color represents: red, ICD; blue, CCD; and green, healthy individuals. *Represents individuals that originally were classified as ileocolonic disease phenotype (**Paper II**), but were reclassified as predominantly having ICD or CCD (**Paper III**).

5.2.4 Targeting subgroups of the microbiota

5.2.4.1 Bacteroides spp.

There were some indications when the fecal T-RFLP data were analyzed, that there was a difference in abundance for some of the TRFs that putatively could be identified as *Bacteroides* species (**Paper II**). This genus is

one of the dominant members of the microbiota, (confirmed with the 454-FLX data, manuscript in preparation), and some species are capable of inducing colitis in an animal model (Sartor, 2003). Therefore, the *Bacteroides* community was studied in more detail using primers that target the *B. fragilis* subgroup. Obtained TRF profiles showed very high similarities for some of the twin pairs but this similarity was not specifically correlated with the disease status (**Paper II**). However, in agreement with the results for the dominant microbiota, the disease phenotype discriminated the clustering of the *Bacteroides* data was segregating according to disease phenotype. In particular, the TRF that correlated with *B. uniformis* (TRF 262, Figure 8) had a significantly lower abundance within the ICD patients (P =0.005). Instead, there was a trend towards increased abundances of the TRFs correlated with *B. ovatus* (TRF 264, P =0.08) and *B. vulgatus* (TRFs 83 and 142, P = 0.12) in the subset with ICD compared with the healthy individuals and the CCD patients (Figure 8).

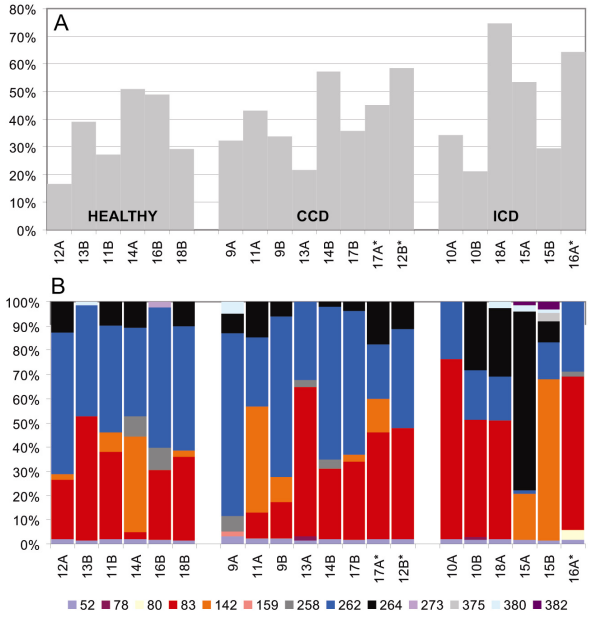

Figure 8. Relative abundances (A) and diversities (B) of *Bacteroides* community members found within the fecal microbiota of discordant and concordant twins. Panel A, shows the relative abundances of sequences classified as *Bacteroides* spp. (454-FLX data) for each individual, grouped according to disease phenotype. Panel B, shows the relative abundances of the detected TRFs, representing members from the *B. fragilis* subgroup. Each bar represents the T-RFLP profile obtained from an individual. The colors representing the different TRFs, are shown in the bottom of the figure. *Represent individuals that originally were classified as ileocolonic disease phenotype (**Paper II**), but were reclassified as predominantly having ICD or CCD (**Paper III**).

The 454-FLX data showed that the abundance of the *Bacteroides* genus was not significantly higher in the ICD patients, compared with CCD patients or healthy individuals. Furthermore, *Bacteroides* abundances were higher in fecal samples than in biopsies for most of the individuals. Other studies have assessed the abundance and diversity of *Bacteroides* spp. in the gut, however with conflicting results (Frank *et al.*, 2007; Gophna *et al.*, 2006; Ott *et al.*, 2004; Scanlan *et al.*, 2006; Swidsinski *et al.*, 2002). A predominance of *Bacteroides,* and foremost *B. vulgatus*, was previously reported from CD patients (Gophna *et al.*, 2006). In addition, higher levels of *Bacteroides* were found on the mucosa of CD patients compared to healthy individuals (Swidsinski *et al.*, 2002). In contrast, Scanlan et al. (2006) failed to amplify *Bacteroides* in 39% for the CD patients and argued that this was not a methodological failure but depleted levels of this genus. In support of that finding, Takaishi et al. (2007) reported reduced abundances of *B. uniformis* as well as *B. ovatus* and *B. vulgatus* in CD patients. The inconsistent results found in abundance of *Bacteroides* can reflect differences in patient populations and methodologies, but it can also argue for CD disease as a group of related disorders rather than a distinct disease.

5.2.4.2 E. coli

Increasing interest has been directed towards *E. coli* and its prevalence in CD patients. Elevated levels of epithelium adherent and invasive strains of *E. coli* (AIEC) have been reported in the ileal mucosa (Darfeuille-Michaud *et al.*, 1998) an in particular in CD patients with the ileal disease phenotype (Baumgart *et al.*, 2007). In **Paper III** the relative abundance of *E. coli* was investigated with qPCR. In agreement with the literature, the prevalence and abundance of *E. coli* was correlated with ICD and was rarely present in the healthy individuals. However, in contrast to previous publications, the results in **Paper III** showed that *E. coli* was not restricted to the ileal mucosa (Figure 9). The higher incidence and relative abundance of *E. coli* was constant along the large bowel, which is consistent with data from an older study (Keighley *et al.*, 1978).

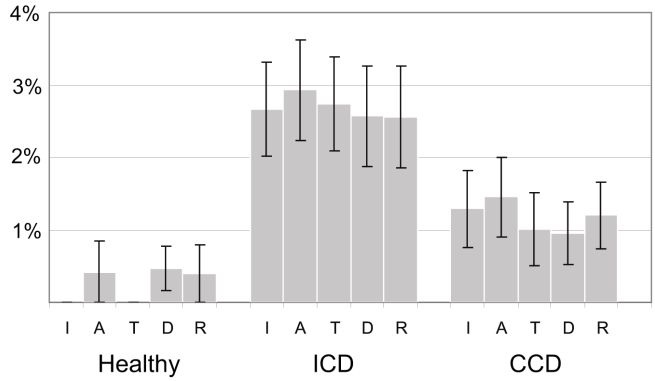

Figure 9. Relative abundance of *E. coli* along the digestive tract of the twins, grouped according to their disease state. Each bar shows the average abundance of *E. coli* for a specific biopsy location and error bars represent standard error of the mean. The sample location is abbreviated below each bar, I; ileum, A; ascending colon, T; transversal colon, D; descending colon, R; rectum. Abbreviations for disease phenotypes, ICD; ileal Crohn's disease, CCD; colonic Crohn's disease.

The increased levels of *E. coli*, observed in **Paper III** as well as in other studies could be of clinical significance because this bacterium has a number of features that may be of importance in CD. For example, animal models have shown that *E. coli* increases the permeability of the epithelium, although without physical damage (Garcia-Lafuente *et al.*, 2001). An altered intestinal permeability has been observed in CD patients as well as among their healthy relatives (Hilsden *et al.*, 1996).

The ileal disease phenotype is associated with a deficiency in release of antimicrobial peptides, such as alpha-defensins (Wehkamp *et al.*, 2005). This could explain the higher prevalence of *E. coli* in tissues from ICD patients. However, alpha defensins are produced by paneth cells restricted to the ileum, and are not produced in colonic tissues. A deficiency in these antimicrobial peptides would presumably increase a localized response with increase of bacteria associated to ileal tissues. The consistent increase of *E. coli* along the digestive tract, in ICD patients suggest that other mechanisms are involved.

5.2.4.3 Faecalibacterium

The analysis of the TRF data in **Papers II** and **III** identified one TRF that was discriminated according to the disease phenotype. TRF 223 was present in all biopsies and fecal samples from the healthy group and the CCD patients, with average abundances of 11% and 15% respectively. However, in patients with ICD, this TRF was absent or present in scarce abundances

(Figure 10). Cloning and sequencing revealed that this TRF correlated with sequences similar to *Faecalibacterium prausnitzii* (Figure 10). The depleted abundances of *F. prausnitzii* were verified with the 454-FLX data and with qPCR data (Figure 10). By comparing the twins, is was shown that both the concordant pairs with ICD and the diseased twin in the pairs with ICD had no or moderate abundance of *Faecalibacterium* whereas the healthy individuals in the ICD pairs as well as all pairs with CCD had this bacterium in substantial levels.

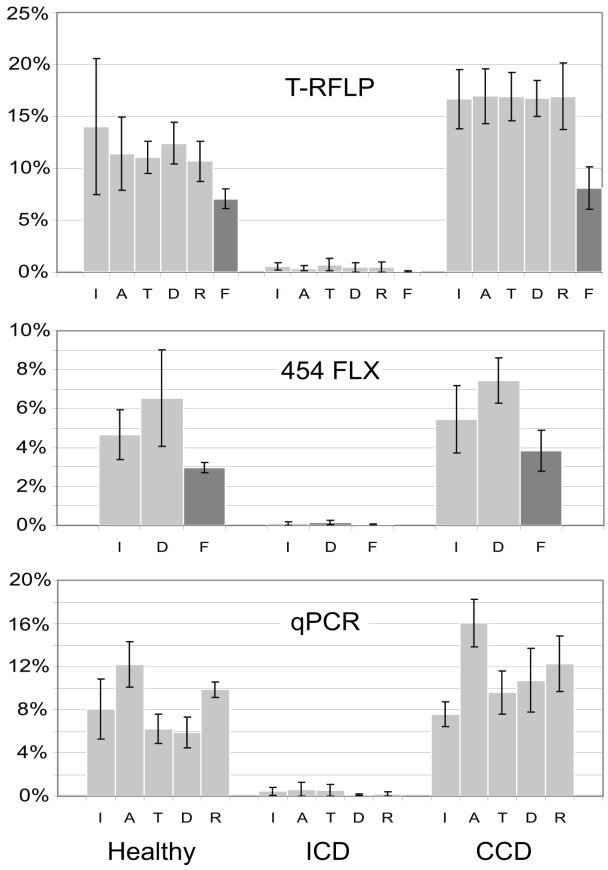

Figure 10. Abundance of *Faecalibacterium* in healthy individuals, ICD and CCD patients. The upper panel shows the relative abundance of TRF 223, the mid panel, relative abundance of *Faecalibacterium* from the 454 data, and the lower panel shows the abundance of *F. prausnitzii* assessed with qPCR. Sample location is indicated below each panel, with the abbreviations, I; Ileum, A; ascending colon, T; transversal colon, D; descending colon, R; rectum, and F; feces.

Studies of the intestinal microbiota have revealed large inter-individual differences. Despite great differences between individuals the prevalence of *F. prausnitzii* seems to be consistent in almost all individuals. This has been shown in several studies (Lay *et al.*, 2005; Mueller *et al.*, 2006; Suau *et al.*, 2001) and in **Paper I**, where TRF 223, representative of *F. prausnitzii* was present in 89/90 children. However, several studies have reported a depletion of members of the *C. leptum* group (which includes *F. prausnitzii* as the dominant member) in CD patients (Baumgart *et al.*, 2007; Frank *et al.*, 2007; Manichanh *et al.*, 2006; Martinez-Medina *et al.*, 2006; Scanlan *et al.*, 2006; Swidsinski *et al.*, 2008; Vasquez *et al.*, 2007). In the literature, this depletion is usually restricted to a subset of CD patients. The data from **Papers II**, **III** and the 454-FLX data (manuscript in preparation), suggest that this depletion is correlated to the disease phenotype. Interestingly, a recent publication showed results that are in agreement with our findings, with scarce abundances or absence of *Faecalibacterium* in patients with the ileal disease phenotype but similar abundances in healthy individuals and those with the colonic disease phenotype (Baumgart *et al.*, 2007). These authors suggest that localized dysbiosis of the mucosa associated bacteria is casually related to ileal inflammation. However, in that study, only the ileal microbiota was analyzed. The results in **Papers II**, **III** and the 454-FLX data, do not support the hypothesis that dysbiosis is localized to a specific region, although higher numbers were found in the biopsies at all locations than in feces. The depletion of *F. prausnitzii*, was consistent along the digestive tract independent of the analytical approach used. In addition, previous investigations comparing inflamed and non-inflamed tissues have failed to detect differences suggesting that dysbiosis is consistent along the length of the colon and not localized in CD patients.

Interestingly, it has been shown that depleted levels of *F. prausnitzii* can be restored with high doses of immune suppressive drugs, suggesting that the absence of *F. prausnitzii* is due to an activated immune response, which specifically eradicates selective groups of bacteria (Swidsinski *et al.*, 2008). Removal of medication resulted in a gradual reduction of *F. prausnitzii* until it was completely absent (Swidsinski *et al.*, 2008). *F. prausnitzii* is known as one of the major producers of butyrate in the bowel. Production of short chain fatty acids, and in particular butyrate is thought to be important for maintaining health (Pryde *et al.*, 2002). Interestingly, reduced levels of butyrate have been reported from CD patients previously (Marchesi *et al.*, 2007; Takaishi *et al.*, 2007). Besides its properties as fuel for the epithelium, butyrate is thought suppress expression of proinflammatory cytokines

(Segain *et al.*, 2000) and may be of critical importance for maintaining health.

5.3 The microbiota in patients with gastric cancer

Persistent infection of *H. pylori* is associated with an increased risk for developing atrophic gastritis and gastric cancer. It is generally accepted that *H. pylori* causes gastritis but its role in the progression into gastric cancer is not evident. In the published literature, it has been suggested that the increased pH of the atrophic stomach allows survival of more genotoxic bacteria that may be of toxicological significance (Blaser & Atherton, 2004; Heavey & Rowland, 2004). However, few studies have characterized the non-*Helicobacter* bacterial flora in patients with gastric cancer and the studies performed to date have mainly relied on cultivation techniques, which has been shown to discriminate against a large fraction of the gut microbiota residing within the gastrointestinal tract (Bik *et al.*, 2006; Eckburg *et al.*, 2005).

In **Paper IV**, the stomach microbiota of ten patients with gastric cancer was characterized and compared with five dyspeptic controls that visited the hospital for abdominal pain, but without any signs of morphological changes on the gastric mucosa. The bacterial community structure was assessed with T-RFLP and by cloning and sequencing of 16S rRNA genes that were PCR amplified from biopsies collected from a subset of the cancer patients. The cancer patients as well as the controls had a diverse assemblage of TRFs and each individual had a unique bacterial community profile. There was no significant difference in the TRF profiles between cancer patients and the control group ($P=0.12$). It was not possible to conclude if the stomach of the gastric cancer patients were subject for bacterial overgrowth since the bacterial load was not quantified and furthermore, no data of the gastric pH was collected in any of the individuals. Despite these shortcomings some interesting and novel information was obtained about the composition of the bacterial communities in individuals with stomach cancer.

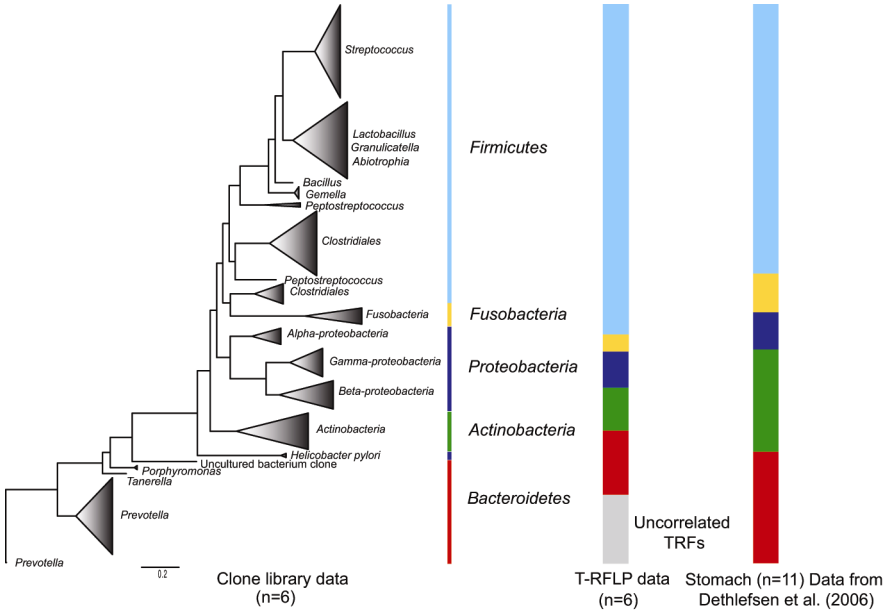

Figure 11. Phylogenetic distribution of clones recovered from gastric cancer patients. Terminal restriction fragment abundances resulting from the T-RFLP analyses of the gastric cancer patients that were correlated to 16S rRNA gene sequences from the same individuals are represented by the bar graph in the center and compared with the phylogenetic distribution of 16S rRNA gene sequences isolated from the stomach of healthy subjects to the right (Dethlefsen *et al.*, 2007). Colors represent different phyla: light blue, *Firmicutes*; orange, *Fusobacteria*; Dark blue, *Proteobacteria*; green, *Actinobacteria*; red, *Bacteroidetes*; and grey, TRFs that could not be correlated to 16S rRNA gene sequences.

To be able to correlate the data from the TRF analysis, cloning and sequencing of 16S rRNA genes was performed in a subset of the gastric cancer patients. In total the gastric cancer microbiota revealed 102 phylotypes, predominantly belonging to *Firmicutes* (Figure 11). Five phyla were represented among the sequenced clones (*Firmicutes*, *Bacteroidetes*, *Proteobacteria*, *Actinobacteria* and *Fusobacteria*) and many of these are normally found in the oral cavity (Aas *et al.*, 2005). The proportions of the phyla found in the clone libraries agree with a previous study of the gastric microbiota of healthy individuals that was also obtained using the same approach (Figure 11; Dethlefsen *et al.*, 2007). *H. pylori* was rarely found among the sequenced clones and was correlated to a TRF with low relative abundance. A scarce presence of *H. pylori* in stomachs with an altered gastric physiology has been observed previously (Blaser & Atherton, 2004). The non-Helicobacter microbiota may play an as yet undiscovered role for progression into gastric cancer. **Paper IV** at least presents the first step towards understanding the microbial community composition in patients

with gastric cancer. Further studies are required to establish whether any of the bacterial populations detected play a role in gastric cancer development or progression.

6 Conclusions

The microbial profile within an individual is unique for that individual. However, in healthy identical twins, genetic factors or a shared microbial colonization history from infancy and early childhood, partly diffuse this individuality. This was confirmed by a very similar composition and diversity in the microbiota of identical twins. However, in twins discordant for CD, these similarities were no longer apparent suggesting that disease is one factor that can cause an imbalance in the composition of the commensal gut microbiota.

Factors associated with the anthroposophic lifestyle were shown to correlate with an increased microbial diversity, which is thought to be generally beneficial for resilience against disturbances. However, presence of keystone members of the microbiota may compensate for a high microbial diversity, supported by the finding that farm children had a low microbial diversity, although, this lifestyle is also associated with a low risk in developing allergies and asthma.

The hypothesis that keystone members are important for health is supported by our finding that depleted prevalence of potential key-members of the gut microbiota was correlated to the ileal disease phenotype in twins with CD. Specific disturbances in the microbiota of patients with ICD were correlated to a lower prevalence of *Firmicutes*, and most striking was the absence or scarce presence of *F. prausnitzii* along the length of the large bowel as well as in fecal samples from the ICD patients. Furthermore, the prevalence of *E. coli* was associated with ICD and was not, as previously reported, restricted to ileal tissues, but consistent to all sampling locations in the colonic mucosa.

Characterization of the microbial composition in the stomach from gastric cancer patients revealed a higher bacterial diversity than previously estimated using cultivation-based approaches. Bacterial populations from five different phyla were present. However, the diversity and community structure of the individuals with gastric cancer were not significantly different from those found in dyspeptic controls.

The different methods used in this thesis have shown a high agreement that strengthens the validity of the results. For example, T-RFLP has a limitation in that it can only detect the dominant groups in complex microbial communities. By contrast, the 454-FLX approach enables deeper exploration of microbial communities due to the large number of sequences that can be generated. Therefore, one might expect a different species distribution to be obtained when comparing these two approaches. However, our results showed very high agreement between these two methods. Since T-RFLP is currently more rapid and inexpensive than 454-FLX sequencing, our results encourage the initial screening of large numbers of samples using the T-RFLP approach prior to selection of those for deeper assessment using 454-FLX.

6.1 Future perspectives

The key question of importance for inflammatory diseases of the gut remains unanswered. Is the inflammation itself causing a shift in the composition of the gut microbiota or is it the shift in the microbiota that causes inflammation? To be able to answer that question the first critical issue is to understand the relative contributions of the host and environmental factors to the dynamics of the microbiota. The current improvement in technologies, such as 454-FLX pyrosequencing, enables a more in depth analysis of the structural composition of the microbial community than was previously possible. Future advances in this technology are in the pipeline with predicted longer length reads. In addition, new sequencing approaches are being developed that might lower the sequencing costs and make this approach more feasible for screening large numbers of samples. This will provide important insights about microbial diversity and dynamics in large sets of individuals. In parallel, the functional repertoire of genes represented in the human microbiome and the expression of these genes as well as the metabolic products of active microbial populations in the gut (in a healthy compared to a diseased state) should be explored. This information would be useful to correlate structural shifts in the microbiota to an impaired metabolic function and will provide important insights into the beneficial or detrimental roles of specific microorganisms as an initiator/maintainer, or

guardian against intestinal diseases. These studies require good model systems to detect significant differences in human populations, and one excellent example is to continue to study identical twins, as in **Papers II** and **III**.

7 References

Aas, J.A., Paster, B.J., Stokes, L.N., Olsen, I. & Dewhirst, F.E. (2005). Defining the normal bacterial flora of the oral cavity. *Journal of Clinical Microbiology* 43(11), 5721-5732.

Abrahamsson, T.R., Jakobsson, T., Böttcher, M.F., Fredrikson, M., Jenmalm, M.C., Björksten, B. & Oldaeus, G. (2007). Probiotics in prevention of IgE-associated eczema: a double-blind, randomized, placebo-controlled trial. *Journal of Allergy and Clinical Immunology* 119(5), 1174-1180.

Adamsson, I., Nord, C.E., Lundquist, P., Sjöstedt, S. & Edlund, C. (1999). Comparative effects of omeprazole, amoxycillin plus metronidazole versus omeprazole, clarithromycin plus metronidazole on the oral, gastric and intestinal microflora in *Helicobacter pylori*-infected patients. *Journal of Antimicrobial Chemotherapy* 44(5), 629-640.

Ahmad, T., Armuzzi, A., Bunce, M., Mulcahy-Hawes, K., Marshall, S.E., Orchard, T.R., Crawshaw, J., Large, O., de Silva, A., Cook, J.T., Barnardo, M., Cullen, S., Welsh, K.I. & Jewell, D.P. (2002). The molecular classification of the clinical manifestations of Crohn's disease. *Gastroenterology* 122(4), 854-866.

Alfvén, T., Braun-Fahrländer, C., Brunekreef, B., von Mutius, E., Riedler, J., Scheynius, A., van Hage, M., Wickman, M., Benz, M.R., Schram, D., Ublagger, E., Waser, M., Pershagen, G. & Group., T.P.S. (2006). Allergic diseases and atopic sensitization in children related to farming and anthroposophic lifestyle - the PARSIFAL study. *Allergy* 61(4), 414-421.

Alm, J.S., Swartz, J., Björksten, B., Engstrand, L., Engström, J., Kuhn, I., Lilja, G., Möllby, R., Norin, E., Pershagen, G., Reinders, C., Wreiber, K. & Scheynius, A. (2002). An anthroposophic lifestyle and intestinal microflora in infancy. *Pediatric Allergy and Immunology* 13(6), 402-411.

Alm, J.S., Swartz, J., Lilja, G., Scheynius, A. & Pershagen, G. (1999). Atopy in children of families with an anthroposophic lifestyle. *Lancet* 353(9163), 1485-1488.

Apajalahti, J.H., Sarkilahti, L.K., Maki, B.R., Heikkinen, J.P., Nurminen, P.H. & Holben, W.E. (1998). Effective recovery of bacterial DNA and percent-guanine-plus-cytosine-based analysis of community structure in the gastrointestinal tract of broiler chickens. *Applied and Environmental Microbiology* 64(10), 4084-4088.

Aries, V.C., Crowther, J.S., Drasar, B.S., Hill, M.J. & Ellis, F.R. (1971). The effect of a strict vegetarian diet on the faecal flora and faecal steroid concentration. *Journal of Pathology* 103(1), 54-56.

Balmer, S.E. & Wharton, B.A. (1991). Diet and faecal flora in the newborn: iron. *Archives of Disease in Childhood* 66(12), 1390-1394.

Baumgart, D.C. & Carding, S.R. (2007). Inflammatory bowel disease: cause and immunobiology. *Lancet* 369(9573), 1627-1640.

Baumgart, D.C. & Sandborn, W.J. (2007). Inflammatory bowel disease: clinical aspects and established and evolving therapies. *Lancet* 369(9573), 1641-1657.

Baumgart, M., Dogan, B., Rishniw, M., Weitzman, G., Bosworth, B., Yantiss, R., Orsi, R.H., Wiedmann, M., McDonough, P., Kim, S.G., Berg, D., Schukken, Y., Scherl, E. & Simpson, K.W. (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 Journal* 1(5), 403-418.

Ben-Amor, K., Heilig, H., Smidt, H., Vaughan, E.E., Abee, T. & de Vos, W.M. (2005). Genetic diversity of viable, injured, and dead fecal bacteria assessed by fluorescence-activated cell sorting and 16S rRNA gene analysis. *Applied and Environmental Microbiology* 71(8), 4679-4689.

Biarc, J., Nguyen, I.S., Pini, A., Gosse, F., Richert, S., Thierse, D., Van Dorsselaer, A., Leize-Wagner, E., Raul, F., Klein, J.P. & Schöller-Guinard, M. (2004). Carcinogenic properties of proteins with pro-inflammatory activity from *Streptococcus infantarius* (formerly *S.bovis*). *Carcinogenesis* 25(8), 1477-1484.

Bik, E.M., Eckburg, P.B., Gill, S.R., Nelson, K.E., Purdom, E.A., Francois, F., Perez-Perez, G., Blaser, M.J. & Relman, D.A. (2006). Molecular analysis of the bacterial microbiota in the human stomach. *Proceedings of the National Academy of Science U S A* 103(3), 732-737.

Björksten, B., Naaber, P., Sepp, E. & Mikelsaar, M. (1999). The intestinal microflora in allergic Estonian and Swedish 2-year-old children. *Clinical and Experimental Allergy* 29(3), 342-346.

Blackwood, C.B., Marsh, T., Kim, S.H. & Paul, E.A. (2003). Terminal restriction fragment length polymorphism data analysis for quantitative comparison of microbial communities. *Applied and Environmental Microbiology* 69(2), 926-932.

Blaser, M.J. & Atherton, J.C. (2004). *Helicobacter pylori* persistence: biology and disease. *Journal of Clinical Investigation* 113(3), 321-333.

Bonnet, R., Suau, A., Dore, J., Gibson, G.R. & Collins, M.D. (2002). Differences in rDNA libraries of faecal bacteria derived from 10- and 25-cycle PCRs. *International Journal of Systematic and Evolutionary Microbiology* 52(Pt 3), 757-763.

Booijink, C.C., Zoetendal, E.G., Kleerebezem, M. & de Vos, W.M. (2007). Microbial communities in the human small intestine: coupling diversity to metagenomics. *Future Microbiology* 2, 285-295.

Braun-Fahrländer, C., Riedler, J., Herz, U., Eder, W., Waser, M., Grize, L., Maisch, S., Carr, D., Gerlach, F., Bufe, A., Lauener, R.P., Schierl, R., Renz, H., Nowak, D. & von Mutius, E. (2002). Environmental exposure to endotoxin and its relation to asthma in school-age children. *New England Journal of Medicine* 347(12), 869-877.

Brook, I., Barrett, C.T., Brinkman, C.R., 3rd, Martin, W.J. & Finegold, S.M. (1979). Aerobic and anaerobic bacterial flora of the maternal cervix and newborn gastric fluid and conjunctiva: a prospective study. *Pediatrics* 63(3), 451-455.

Bäckhed, F., Ley, R.E., Sonnenburg, J.L., Peterson, D.A. & Gordon, J.I. (2005). Host-bacterial mutualism in the human intestine. *Science* 307(5717), 1915-20.

Canani, R.B., Cirillo, P., Terrin, G., Cesarano, L., Spagnuolo, M.I., De Vincenzo, A., Albano, F., Passariello, A., De Marco, G., Manguso, F. & Guarino, A. (2007). Probiotics for treatment of acute diarrhoea in children: randomised clinical trial of five different preparations. *British Medical Journal* 335(7615), 340.

Chiodini, R.J. (1989). Crohn's disease and the mycobacterioses: a review and comparison of two disease entities. *Clinical Microbiology Reviews* 2(1), 90-117.

Clement, B.G. & Kitts, C.L. (2000). Isolating PCR-quality DNA from human feces with a soil DNA kit. *Biotechniques* 28(4), 640-642, 644, 646.

D'Argenio, G. & Mazzacca, G. (1999). Short-chain fatty acid in the human colon. Relation to inflammatory bowel diseases and colon cancer. *Advances in Experimental Medicine and Biology* 472, 149-158.

Darfeuille-Michaud, A., Neut, C., Barnich, N., Lederman, E., Di Martino, P., Desreumaux, P., Gambiez, L., Joly, B., Cortot, A. & Colombel, J.F. (1998). Presence of adherent *Escherichia coli* strains in ileal mucosa of patients with Crohn's disease. *Gastroenterology* 115(6), 1405-1413.

Dethlefsen, L., McFall-Ngai, M. & Relman, D.A. (2007). An ecological and evolutionary perspective on human-microbe mutualism and disease. *Nature* 449(7164), 811-818.

Dunbar, J., Ticknor, L.O. & Kuske, C.R. (2000). Assessment of microbial diversity in four southwestern United States soils by 16S rRNA

gene terminal restriction fragment analysis. *Applied and Environmental Microbiology* 66(7), 2943-2950.

Eckburg, P.B., Bik, E.M., Bernstein, C.N., Purdom, E., Dethlefsen, L., Sargent, M., Gill, S.R., Nelson, K.E. & Relman, D.A. (2005). Diversity of the human intestinal microbial flora. *Science* 308(5728), 1635-1638.

Edlund, A., Soule, T., Sjöling, S. & Jansson, J.K. (2006). Microbial community structure in polluted Baltic Sea sediments. *Environmental Microbiology* 8(2), 223-232.

Ellmerich, S., Schöller, M., Duranton, B., Gosse, F., Galluser, M., Klein, J.P. & Raul, F. (2000). Promotion of intestinal carcinogenesis by *Streptococcus bovis*. *Carcinogenesis* 21(4), 753-756.

Favier, C.F., Vaughan, E.E., De Vos, W.M. & Akkermans, A.D. (2002). Molecular monitoring of succession of bacterial communities in human neonates. *Applied and Environmental Microbiology* 68(1), 219-226.

Finegold, S.M. & Sutter, V.L. (1978). Fecal flora in different populations, with special reference to diet. *American Journal of Clinical Nutrition* 31(10 Suppl), S116-S122.

Flöistrup, H., Swartz, J., Bergström, A., Alm, J.S., Scheynius, A., van Hage, M., Waser, M., Braun-Fahrländer, C., Schram-Bijkerk, D., Huber, M., Zutavern, A., von Mutius, E., Ublagger, E., Riedler, J., Michaels, K.B., Pershagen, G. & The Parsifal Study Group (2006). Allergic disease and sensitization in Steiner school children. *Journal of Allergy and Clinical Immunology* 117(1), 59-66.

Frank, D.N., St Amand, A.L., Feldman, R.A., Boedeker, E.C., Harpaz, N. & Pace, N.R. (2007). Molecular-phylogenetic characterization of microbial community imbalances in human inflammatory bowel diseases. *Proceedings of the National Academy of Science U S A* 104(34), 13780-13785.

Fuller, R. (1986). Probiotics. *Society for Applied Bacteriology Symposium Series* 15, 1S-7S.

Garcia-Lafuente, A., Antolin, M., Guarner, F., Crespo, E. & Malagelada, J.R. (2001). Modulation of colonic barrier function by the composition of the commensal flora in the rat. *Gut* 48(4), 503-507.

Genta, R.M., Gurer, I.E. & Graham, D.Y. (1995). Geographical pathology of *Helicobacter pylori* infection: is there more than one gastritis? *Annals of Medicine* 27(5), 595-599.

Gibson, G.R. & Roberfroid, M.B. (1995). Dietary modulation of the human colonic microbiota: introducing the concept of prebiotics. *Journal of Nutrition* 125(6), 1401-1412.

Gophna, U., Sommerfeld, K., Gophna, S., Doolittle, W.F. & Veldhuyzen van Zanten, S.J. (2006). Differences between tissue-associated intestinal microfloras of patients with Crohn's disease and ulcerative colitis. *Journal of Clinical Microbiology* 44(11), 4136-4141.

Greenstein, R.J. (2003). Is Crohn's disease caused by a mycobacterium? Comparisons with leprosy, tuberculosis, and Johne's disease. *Lancet Infectious Diseases* 3(8), 507-514.

Grönlund, M.M., Lehtonen, O.P., Eerola, E. & Kero, P. (1999). Fecal microflora in healthy infants born by different methods of delivery: permanent changes in intestinal flora after cesarean delivery. *Journal of Pediatric Gastroenterology and Nutrition* 28(1), 19-25.

Guarner, F. (2006). Enteric flora in health and disease. *Digestion* 73 Suppl 1, 5-12.

Guarner, F., Bourdet-Sicard, R., Brandtzaeg, P., Gill, H.S., McGuirk, P., van Eden, W., Versalovic, J., Weinstock, J.V. & Rook, G.A. (2006). Mechanisms of disease: the hygiene hypothesis revisited. *Nature Clinical Practice Gastroenterologt & Hepatology* 3(5), 275-284.

Guarner, F. & Malagelada, J.R. (2003). Gut flora in health and disease. *Lancet* 361(9356), 512-519.

Hansson, L.E., Nyren, O., Hsing, A.W., Bergstrom, R., Josefsson, S., Chow, W.H., Fraumeni, J.F., Jr. & Adami, H.O. (1996). The risk of stomach cancer in patients with gastric or duodenal ulcer disease. *New England Journal of Medicine* 335(4), 242-249.

Harmsen, H.J., Wildeboer-Veloo, A.C., Raangs, G.C., Wagendorp, A.A., Klijn, N., Bindels, J.G. & Welling, G.W. (2000). Analysis of intestinal flora development in breast-fed and formula-fed infants by using molecular identification and detection methods. *Journal of Pediatric Gastroenterology and Nutrition* 30(1), 61-67.

Harper, P.H., Truelove, S.C., Lee, E.C., Kettlewell, M.G. & Jewell, D.P. (1983). Split ileostomy and ileocolostomy for Crohn's disease of the colon and ulcerative colitis: a 20 year survey. *Gut* 24(2), 106-113.

Hayashi, H., Sakamoto, M. & Benno, Y. (2002). Fecal microbial diversity in a strict vegetarian as determined by molecular analysis and cultivation. *Microbiology and Immunology* 46(12), 819-831.

Hayashi, H., Takahashi, R., Nishi, T., Sakamoto, M. & Benno, Y. (2005). Molecular analysis of jejunal, ileal, caecal and recto-sigmoidal human colonic microbiota using 16S rRNA gene libraries and terminal restriction fragment length polymorphism. *Journal of Medical Microbiology* 54(Pt 11), 1093-1101.

Heavey, P.M. & Rowland, I.R. (2004). Microbial-gut interactions in health and disease. Gastrointestinal cancer. *Best Practice & Research in Clinical Gastroenterology* 18(2), 323-336.

Hentges, D.J. (1983). *Human intestinal microflora in health and disease* New York: Academic Press.

Hill, M.J., Drasar, B.S., Hawksworth, G., Aries, V., Crowther, J.S. & Williams, R.E. (1971). Bacteria and aetiology of cancer of large bowel. *Lancet* 1(7690), 95-100.

Hilsden, R.J., Meddings, J.B. & Sutherland, L.R. (1996). Intestinal permeability changes in response to acetylsalicylic acid in relatives of patients with Crohn's disease. *Gastroenterology* 110(5), 1395-1403.

Hooper, L.V., Stappenbeck, T.S., Hong, C.V. & Gordon, J.I. (2003). Angiogenins: a new class of microbicidal proteins involved in innate immunity. *Nature Immunology* 4(3), 269-273.

Hooper, L.V., Xu, J., Falk, P.G., Midtvedt, T. & Gordon, J.I. (1999). A molecular sensor that allows a gut commensal to control its nutrient foundation in a competitive ecosystem. *Proceedings of the National Academy of Science U S A* 96(17), 9833-9838.

Hope, M.E., Hold, G.L., Kain, R. & El-Omar, E.M. (2005). Sporadic colorectal cancer-role of the commensal microbiota. *FEMS Microbiology Letters* 244(1), 1-7.

Husebye, E. (2005). The pathogenesis of gastrointestinal bacterial overgrowth. *Chemotherapy* 51 Suppl 1, 1-22.

Isolauri, E., Arvola, T., Sutas, Y., Moilanen, E. & Salminen, S. (2000). Probiotics in the management of atopic eczema. *Clinical and Experimental Allergy* 30(11), 1604-1610.

Jenkinson, H.F. & Lamont, R.J. (2005). Oral microbial communities in sickness and in health. *Trends in Microbiology* 13(12), 589-595.

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

Jernberg, C., Sullivan, Å., Edlund, C. & Jansson, J.K. (2005). Monitoring of antibiotic-induced alterations in the human intestinal microflora and detection of probiotic strains by use of terminal restriction fragment length polymorphism. *Applied and Environmental Microbiology* 71(1), 501-506.

Kado, S., Uchida, K., Funabashi, H., Iwata, S., Nagata, Y., Ando, M., Onoue, M., Matsuoka, Y., Ohwaki, M. & Morotomi, M. (2001). Intestinal microflora are necessary for development of spontaneous adenocarcinoma of the large intestine in T-cell receptor beta chain and p53 double-knockout mice. *Cancer Research* 61(6), 2395-2398.

Kajander, K., Myllyluoma, E., Rajilic-Stojanovic, M., Kyrönpalo, S., Rasmussen, M., Järvenpää, S., Zoetendal, E.G., de Vos, W.M., Vapaatalo, H. & Korpela, R. (2008). Clinical trial: multispecies probiotic supplementation alleviates the symptoms of irritable bowel syndrome and stabilizes intestinal microbiota. *Alimentary Pharmacology &Therapeutics* 27(1), 48-57.

Kalliomäki, M., Kirjavainen, P., Eerola, E., Kero, P., Salminen, S. & Isolauri, E. (2001). Distinct patterns of neonatal gut microflora in infants in whom atopy was and was not developing. *Journal of Allergy and Clinical Immunology* 107(1), 129-34.

Kanazawa, K., Konishi, F., Mitsuoka, T., Terada, A., Itoh, K., Narushima, S., Kumemura, M. & Kimura, H. (1996). Factors influencing the

development of sigmoid colon cancer. Bacteriologic and biochemical studies. *Cancer* 77(8 Suppl), 1701-1706.

Kaplan, C.W. & Kitts, C.L. (2003). Variation between observed and true Terminal Restriction Fragment length is dependent on true TRF length and purine content. *Journal of Microbiological Methods* 54(1), 121-125.

Kaplan, C.W. & Kitts, C.L. (2004). Bacterial succession in a petroleum land treatment unit. *Applied and Environmental Microbiology* 70(3), 1777-1786.

Kassinen, A., Krogius-Kurikka, L., Mäkivuokko, H., Rinttilä, T., Paulin, L., Corander, J., Malinen, E., Apajalahti, J. & Palva, A. (2007). The fecal microbiota of irritable bowel syndrome patients differs significantly from that of healthy subjects. *Gastroenterology* 133(1), 24-33.

Keighley, M.R., Arabi, Y., Dimock, F., Burdon, D.W., Allan, R.N. & Alexander-Williams, J. (1978). Influence of inflammatory bowel disease on intestinal microflora. *Gut* 19(12), 1099-1104.

Kitts, C.L. (2001). Terminal restriction fragment patterns: a tool for comparing microbial communities and assessing community dynamics. *Current Issues in Intestinal Microbiology* 2(1), 17-25.

Kleessen, B., Kroesen, A.J., Buhr, H.J. & Blaut, M. (2002). Mucosal and invading bacteria in patients with inflammatory bowel disease compared with controls. *Scandinavian Journal of Gastroenterology* 37(9), 1034-1041.

Kotlowski, R., Bernstein, C.N., Sepehri, S. & Krause, D.O. (2007). High prevalence of *Escherichia coli* belonging to the B2+D phylogenetic group in inflammatory bowel disease. *Gut* 56(5), 669-75.

Kukkonen, K., Savilahti, E., Haahtela, T., Juntunen-Backman, K., Korpela, R., Poussa, T., Tuure, T. & Kuitunen, M. (2007). Probiotics and prebiotic galacto-oligosaccharides in the prevention of allergic diseases: a randomized, double-blind, placebo-controlled trial. *Journal of Allergy and Clinical Immunology* 119(1), 192-198.

Lay, C., Rigottier-Gois, L., Holmstrom, K., Rajilic, M., Vaughan, E.E., de Vos, W.M., Collins, M.D., Thiel, R., Namsolleck, P., Blaut, M. & Dore, J. (2005). Colonic microbiota signatures across five northern European countries. *Applied and Environmental Microbiology* 71(7), 4153-4155.

Legendre, P. & Legendre, L. (1998). *Numerical ecology*. 2nd English edition. Amsterdam: Elsevier. (Developments in environmental modelling ; 20

Lepage, P., Seksik, P., Sutren, M., de la Cochetiere, M.F., Jian, R., Marteau, P. & Dore, J. (2005). Biodiversity of the mucosa-associated microbiota is stable along the distal digestive tract in healthy individuals and patients with IBD. *Inflammatory Bowel Diseases* 11(5), 473-480.

Ley, R.E., Bäckhed, F., Turnbaugh, P., Lozupone, C.A., Knight, R.D. & Gordon, J.I. (2005). Obesity alters gut microbial ecology. *Proceedings of the National Academy of Science U S A* 102(31), 11070-11075.

Ley, R.E., Peterson, D.A. & Gordon, J.I. (2006a). Ecological and evolutionary forces shaping microbial diversity in the human intestine. *Cell* 124(4), 837-848.

Ley, R.E., Turnbaugh, P.J., Klein, S. & Gordon, J.I. (2006b). Microbial ecology: human gut microbes associated with obesity. *Nature* 444(7122), 1022-1023.

Liu, W.T., Marsh, T.L., Cheng, H. & Forney, L.J. (1997). Characterization of microbial diversity by determining terminal restriction fragment length polymorphisms of genes encoding 16S rRNA. *Applied and Environmental Microbiology* 63(11), 4516-4522.

Lochhead, P. & El-Omar, E.M. (2007). *Helicobacter pylori* infection and gastric cancer. *Best Practice & Research in Clinical Gastroenterology* 21(2), 281-97.

Loftus, E.V., Jr. (2004). Clinical epidemiology of inflammatory bowel disease: Incidence, prevalence, and environmental influences. *Gastroenterology* 126(6), 1504-1517.

Lueders, T. & Friedrich, M.W. (2003). Evaluation of PCR amplification bias by terminal restriction fragment length polymorphism analysis of small-subunit rRNA and mcrA genes by using defined template mixtures of methanogenic pure cultures and soil DNA extracts. *Applied and Environmental Microbiology* 69(1), 320-326.

Löfmark, S., Jernberg, C., Jansson, J.K. & Edlund, C. (2006). Clindamycin-induced enrichment and long-term persistence of resistant *Bacteroides* spp. and resistance genes. *Journal of Antimicrobial Chemotheraphy* 58(6), 1160-1167.

Macfarlane, G.T. & Macfarlane, S. (1997). Human colonic microbiota: ecology, physiology and metabolic potential of intestinal bacteria. *Scand J Gastroenterol Suppl* 222, 3-9.

Macfarlane, G.T. & Macfarlane, S. (2007). Models for intestinal fermentation: association between food components, delivery systems, bioavailability and functional interactions in the gut. *Current Opinion in Biotechnology* 18(2), 156-162.

Mai, V., Greenwald, B., Morris, J.G., Jr., Raufman, J.P. & Stine, O.C. (2006). Effect of bowel preparation and colonoscopy on post-procedure intestinal microbiota composition. *Gut* 55(12), 1822-1823.

Manichanh, C., Rigottier-Gois, L., Bonnaud, E., Gloux, K., Pelletier, E., Frangeul, L., Nalin, R., Jarrin, C., Chardon, P., Marteau, P., Roca, J. & Dore, J. (2006). Reduced diversity of faecal microbiota in Crohn's disease revealed by a metagenomic approach. *Gut* 55(2), 205-211.

Marchesi, J.R., Holmes, E., Khan, F., Kochhar, S., Scanlan, P., Shanahan, F., Wilson, I.D. & Wang, Y. (2007). Rapid and noninvasive metabonomic characterization of inflammatory bowel disease. *Journal of Proteome Research* 6(2), 546-551.

Marsh, T.L., Saxman, P., Cole, J. & Tiedje, J. (2000). Terminal restriction fragment length polymorphism analysis program, a web-based research tool for microbial community analysis. *Applied and Environmental Microbiology* 66(8), 3616-3620.

Marshall, B.J. & Warren, J.R. (1984). Unidentified curved bacilli in the stomach of patients with gastritis and peptic ulceration. *Lancet* 1(8390), 1311-1315.

Martinez-Medina, M., Aldeguer, X., Gonzalez-Huix, F., Acero, D. & Garcia-Gil, L.J. (2006). Abnormal microbiota composition in the ileocolonic mucosa of Crohn's disease patients as revealed by polymerase chain reaction-denaturing gradient gel electrophoresis. *Inflammatory Bowel Diseases* 12(12), 1136-1145.

Mazmanian, S.K., Liu, C.H., Tzianabos, A.O. & Kasper, D.L. (2005). An immunomodulatory molecule of symbiotic bacteria directs maturation of the host immune system. *Cell* 122(1), 107-118.

McCracken, V.J., Simpson, J.M., Mackie, R.I. & Gaskins, H.R. (2001). Molecular ecological analysis of dietary and antibiotic-induced alterations of the mouse intestinal microbiota. *Journal of Nutrition* 131(6), 1862-1870.

McFadden, J.J., Butcher, P.D., Chiodini, R. & Hermon-Taylor, J. (1987). Crohn's disease-isolated mycobacteria are identical to *Mycobacterium paratuberculosis*, as determined by DNA probes that distinguish between mycobacterial species. *Journal of Clinical Microbiology* 25(5), 796-801.

McOrist, A.L., Jackson, M. & Bird, A.R. (2002). A comparison of five methods for extraction of bacterial DNA from human faecal samples. *Journal of Microbiological Methods* 50(2), 131-139.

Monstein, H.J., Tiveljung, A., Kraft, C.H., Borch, K. & Jonasson, J. (2000). Profiling of bacterial flora in gastric biopsies from patients with *Helicobacter pylori*-associated gastritis and histologically normal control individuals by temperature gradient gel electrophoresis and 16S rDNA sequence analysis. *Journal of Medical Microbiology* 49(9), 817-822.

Moore, W.E. & Moore, L.H. (1995). Intestinal floras of populations that have a high risk of colon cancer. *Applied and Environmental Microbiology* 61(9), 3202-3207.

Mowat, C., Williams, C., Gillen, D., Hossack, M., Gilmour, D., Carswell, A., Wirz, A., Preston, T. & McColl, K.E. (2000). Omeprazole, *Helicobacter pylori* status, and alterations in the intragastric milieu facilitating bacterial N-nitrosation. *Gastroenterology* 119(2), 339-347.

Mpofu, C.M., Campbell, B.J., Subramanian, S., Marshall-Clarke, S., Hart, C.A., Cross, A., Roberts, C.L., McGoldrick, A., Edwards, S.W. & Rhodes, J.M. (2007). Microbial mannan inhibits bacterial killing by macrophages: a possible pathogenic mechanism for Crohn's disease. *Gastroenterology* 133(5), 1487-1498.

Mueller, S., Saunier, K., Hanisch, C., Norin, E., Alm, L., Midtvedt, T., Cresci, A., Silvi, S., Orpianesi, C., Verdenelli, M.C., Clavel, T., Koebnick, C., Zunft, H.J., Dore, J. & Blaut, M. (2006). Differences in fecal microbiota in different European study populations in relation to age, gender, and country: a cross-sectional study. *Applied and Environmental Microbiology* 72(2), 1027-1033.

Muyzer, G., Dewaal, E.C. & Uitterlinden, A.G. (1993). Profiling Of Complex Microbial-Populations By Denaturing Gradient Gel-Electrophoresis Analysis Of Polymerase Chain Reaction-Amplified Genes-Coding For 16s Ribosomal-Rna. *Applied and Environmental Microbiology* 59(3), 695-700.

Nielsen, D.S., Möller, P.L., Rosenfeldt, V., Paerregaard, A., Michaelsen, K.F. & Jakobsen, M. (2003). Case study of the distribution of mucosa-associated *Bifidobacterium* species, *Lactobacillus* species, and other lactic acid bacteria in the human colon. *Applied and Environmental Microbiology* 69(12), 7545-7548.

Norin, K.E. (1997). Influence of antibiotics on some intestinal microflora associated characteristics. *Anaerobe* 3(2-3), 145-148.

Noverr, M.C. & Huffnagle, G.B. (2004). Does the microbiota regulate immune responses outside the gut? *Trends in Microbiology* 12(12), 562-568.

O'Hara, A.M. & Shanahan, F. (2006). The gut flora as a forgotten organ. *EMBO Reports* 7(7), 688-693.

O'Keefe, S.J. (2008). Nutrition and colonic health: the critical role of the microbiota. *Current Opinion in Gastroenterology* 24(1), 51-58.

Orholm, M., Binder, V., Sorensen, T.I., Rasmussen, L.P. & Kyvik, K.O. (2000). Concordance of inflammatory bowel disease among Danish twins. Results of a nationwide study. *Scandinavian Journal of Gastroenterology* 35(10), 1075-1081.

Osborn, A.M., Moore, E.R. & Timmis, K.N. (2000). An evaluation of terminal-restriction fragment length polymorphism (T-RFLP) analysis for the study of microbial community structure and dynamics. *Environmental Microbiology* 2(1), 39-50.

Ott, S.J., Musfeldt, M., Wenderoth, D.F., Hampe, J., Brant, O., Fölsch, U.R., Timmis, K.N. & Schreiber, S. (2004). Reduction in diversity of the colonic mucosa associated bacterial microflora in patients with active inflammatory bowel disease. *Gut* 53(5), 685-693.

Palmer, C., Bik, E.M., Digiulio, D.B., Relman, D.A. & Brown, P.O. (2007). Development of the Human Infant Intestinal Microbiota. *Public Library of Science Biolology* 5(7), 1556-1573.

Parracho, H.M., Bingham, M.O., Gibson, G.R. & McCartney, A.L. (2005). Differences between the gut microflora of children with autistic spectrum disorders and that of healthy children. *Journal of Medical Microbiology* 54(Pt 10), 987-991.

Pei, Z., Bini, E.J., Yang, L., Zhou, M., Francois, F. & Blaser, M.J. (2004). Bacterial biota in the human distal esophagus. *Proceedings of the National Academy of Science U S A* 101(12), 4250-4255.

Peltonen, R., Ling, W.H., Hanninen, O. & Eerola, E. (1992). An uncooked vegan diet shifts the profile of human fecal microflora: computerized analysis of direct stool sample gas-liquid chromatography profiles of bacterial cellular fatty acids. *Applied and Environmental Microbiology* 58(11), 3660-3666.

Polz, M.F. & Cavanaugh, C.M. (1998). Bias in template-to-product ratios in multitemplate PCR. *Applied and Environmental Microbiology* 64(10), 3724-3730.

Pothoulakis, C. (1996). Pathogenesis of *Clostridium difficile*-associated diarrhoea. *European Journal of Gastroenterology & Hepatology* 8(11), 1041-1047.

Pryde, S.E., Duncan, S.H., Hold, G.L., Stewart, C.S. & Flint, H.J. (2002). The microbiology of butyrate formation in the human colon. *FEMS Microbiology Letters* 217(2), 133-9.

Rees, G.N., Baldwin, D.S., Watson, G.O., Perryman, S. & Nielsen, D.L. (2004). Ordination and significance testing of microbial community composition derived from terminal restriction fragment length polymorphisms: application of multivariate statistics. *Antonie Van Leeuwenhoek* 86(4), 339-347.

Ronaghi, M., Uhlen, M. & Nyren, P. (1998). A sequencing method based on real-time pyrophosphate. *Science* 281(5375), 363, 365.

Rosenfeldt, V., Benfeldt, E., Nielsen, S.D., Michaelsen, K.F., Jeppesen, D.L., Valerius, N.H. & Paerregaard, A. (2003). Effect of probiotic *Lactobacillus* strains in children with atopic dermatitis. *Journal of Allergy and Clinical Immunology* 111(2), 389-395.

Sakamoto, M., Hayashi, H. & Benno, Y. (2003). Terminal restriction fragment length polymorphism analysis for human fecal microbiota and its application for analysis of complex bifidobacterial communities. *Microbiol Immunol* 47(2), 133-42.

Sakamoto, M., Umeda, M. & Benno, Y. (2005). Molecular analysis of human oral microbiota. *Journal of Periodontal Research* 40(3), 277-285.

Sakata, S., Tonooka, T., Ishizeki, S., Takada, M., Sakamoto, M., Fukuyama, M. & Benno, Y. (2005). Culture-independent analysis of fecal microbiota in infants, with special reference to Bifidobacterium species. *FEMS Microbiology Letters* 243(2), 417-423.

Salminen, S., Gibson, G.R., McCartney, A.L. & Isolauri, E. (2004). Influence of mode of delivery on gut microbiota composition in seven year old children. *Gut* 53(9), 1388-1389.

Sanduleanu, S., Jonkers, D., De Bruine, A., Hameeteman, W. & Stockbrugger, R.W. (2001). Non-*Helicobacter pylori* bacterial flora during acid-suppressive therapy: differential findings in gastric juice and gastric mucosa. *Alimentary Pharmacology &Therapeutics* 15(3), 379-388.

Sartor, R.B. (2003). Targeting enteric bacteria in treatment of inflammatory bowel diseases: why, how, and when. *Current Opinion in Gastroenterology* 19(4), 358-365.

Sartor, R.B. (2004). Therapeutic manipulation of the enteric microflora in inflammatory bowel diseases: antibiotics, probiotics, and prebiotics. *Gastroenterology* 126(6), 1620-1633.

Sartor, R.B. (2005). Does *Mycobacterium avium* subspecies *paratuberculosis* cause Crohn's disease? *Gut* 54(7), 896-898.

Satsangi, J., Silverberg, M.S., Vermeire, S. & Colombel, J.F. (2006). The Montreal classification of inflammatory bowel disease: controversies, consensus, and implications. *Gut* 55(6), 749-753.

Savage, D.C. (1977). Microbial ecology of the gastrointestinal tract. *Annual Review of Microbiology* 31, 107-133.

Scanlan, P.D., Shanahan, F., O'Mahony, C. & Marchesi, J.R. (2006). Culture-independent analyses of temporal variation of the dominant fecal microbiota and targeted bacterial subgroups in Crohn's disease. *Journal of Clinical Microbiology* 44(11), 3980-3988.

Schram-Bijkerk, D., Doekes, G., Douwes, J., Boeve, M., Riedler, J., Ublagger, E., von Mutius, E., Benz, M.R., Pershagen, G., van Hage, M., Scheynius, A., Braun-Fahrländer, C., Waser, M. & Brunekreef, B. (2005). Bacterial and fungal agents in house dust and wheeze in children: the PARSIFAL study. *Clinical and Experimental Allergy* 35(10), 1272-1278.

Scupham, A.J., Jones, J.A. & Wesley, I.V. (2007). Comparison of DNA extraction methods for analysis of turkey cecal microbiota. *Journal of Applied Microbiology* 102(2), 401-409.

Segain, J.P., Raingeard de la Bletiere, D., Bourreille, A., Leray, V., Gervois, N., Rosales, C., Ferrier, L., Bonnet, C., Blottiere, H.M. & Galmiche, J.P. (2000). Butyrate inhibits inflammatory responses through NFkappaB inhibition: implications for Crohn's disease. *Gut* 47(3), 397-403.

Seksik, P., Rigottier-Gois, L., Gramet, G., Sutren, M., Pochart, P., Marteau, P., Jian, R. & Dore, J. (2003). Alterations of the dominant faecal bacterial groups in patients with Crohn's disease of the colon. *Gut* 52(2), 237-242.

Sepp, E., Julge, K., Vasar, M., Naaber, P., Björksten, B. & Mikelsaar, M. (1997). Intestinal microflora of Estonian and Swedish infants. *Acta Paediatrica* 86(9), 956-961.

Sissons, C.H. & Hancock, E.M. (1993). Urease activity in *Streptococcus salivarius* at low pH. *Archives of Oral Biology* 38(6), 507-516.

Sjöstedt, S., Kager, L., Heimdahl, A. & Nord, C.E. (1988). Microbial colonization of tumors in relation to the upper gastrointestinal tract in patients with gastric carcinoma. *Annals of Surgery* 207(3), 341-346.

Smith, M.G., Hold, G.L., Tahara, E. & El-Omar, E.M. (2006). Cellular and molecular aspects of gastric cancer. *World Journal of Gastroenterology* 12(19), 2979-2990.

Stecher, B. & Hardt, W.D. (2008). The role of microbiota in infectious disease. *Trends in Microbiology* 16(3), 107-114.

Stewart, J.A., Chadwick, V.S. & Murray, A. (2005). Investigations into the influence of host genetics on the predominant eubacteria in the faecal microflora of children. *Journal of Medical Microbiology* 54(Pt 12), 1239-1242.

Strachan, D.P. (1989). Hay fever, hygiene, and household size. *British Medical Journal* 299(6710), 1259-1260.

Strachan, D.P. (2000). Family size, infection and atopy: the first decade of the "hygiene hypothesis". *Thorax* 55 Suppl 1, S2-10.

Suau, A., Rochet, V., Sghir, A., Gramet, G., Brewaeys, S., Sutren, M., Rigottier-Gois, L. & Dore, J. (2001). *Fusobacterium prausnitzii* and related species represent a dominant group within the human fecal flora. *Systematic and Applied Microbiology* 24(1), 139-145.

Swidsinski, A., Ladhoff, A., Pernthaler, A., Swidsinski, S., Loening-Baucke, V., Ortner, M., Weber, J., Hoffmann, U., Schreiber, S., Dietel, M. & Lochs, H. (2002). Mucosal flora in inflammatory bowel disease. *Gastroenterology* 122(1), 44-54.

Swidsinski, A., Loening-Baucke, V., Theissig, F., Engelhardt, H., Bengmark, S., Koch, S., Lochs, H. & Dorffel, Y. (2007). Comparative study of the intestinal mucus barrier in normal and inflamed colon. *Gut* 56(3), 343-350.

Swidsinski, A., Loening-Baucke, V., Vaneechoutte, M. & Doerffel, Y. (2008). Active Crohn's disease and ulcerative colitis can be specifically diagnosed and monitored based on the biostructure of the fecal flora. *Inflammatory Bowel Diseases* 14(2), 147-161.

Swidsinski, A., Weber, J., Loening-Baucke, V., Hale, L.P. & Lochs, H. (2005). Spatial organization and composition of the mucosal flora in patients with inflammatory bowel disease. *Journal of Clinical Microbiology* 43(7), 3380-3389.

Takaishi, H., Matsuki, T., Nakazawa, A., Takada, T., Kado, S., Asahara, T., Kamada, N., Sakuraba, A., Yajima, T., Higuchi, H., Inoue, N., Ogata, H., Iwao, Y., Nomoto, K., Tanaka, R. & Hibi, T. (2007).

Imbalance in intestinal microflora constitution could be involved in the pathogenesis of inflammatory bowel disease. *International Journal of Medical Microbiology*

Tamboli, C.P., Neut, C., Desreumaux, P. & Colombel, J.F. (2004). Dysbiosis in inflammatory bowel disease. *Gut* 53(1), 1-4.

Thompson, N.P., Driscoll, R., Pounder, R.E. & Wakefield, A.J. (1996). Genetics versus environment in inflammatory bowel disease: results of a British twin study. *British Medical Journal* 312(7023), 95-96.

Torsvik, V., Daae, F.L., Sandaa, R.A. & Ovreas, L. (1998). Novel techniques for analysing microbial diversity in natural and perturbed environments. *Journal of Biotechnology* 64(1), 53-62.

Turnbaugh, P.J., Ley, R.E., Hamady, M., Fraser-Liggett, C.M., Knight, R. & Gordon, J.I. (2007). The human microbiome project. *Nature* 449(7164), 804-810.

Tysk, C., Lindberg, E., Järnerot, G. & Floderus-Myrhed, B. (1988). Ulcerative colitis and Crohn's disease in an unselected population of monozygotic and dizygotic twins. A study of heritability and the influence of smoking. *Gut* 29(7), 990-996.

Uemura, N., Okamoto, S., Yamamoto, S., Matsumura, N., Yamaguchi, S., Yamakido, M., Taniyama, K., Sasaki, N. & Schlemper, R.J. (2001). *Helicobacter pylori* infection and the development of gastric cancer. *New England Journal of Medicine* 345(11), 784-789.

Van de Merwe, J.P., Stegeman, J.H. & Hazenberg, M.P. (1983). The Resident Faecal Flora is Determined by Genetic Characteristics of the Host. Implications for Crohn's Disease? *Antonie Van Leeuwenhoek* 49, 119-124.

van der Waaij, L.A., Harmsen, H.J., Madjipour, M., Kroese, F.G., Zwiers, M., van Dullemen, H.M., de Boer, N.K., Welling, G.W. & Jansen, P.L. (2005). Bacterial population analysis of human colon and terminal ileum biopsies with 16S rRNA-based fluorescent probes: commensal bacteria live in suspension and have no direct contact with epithelial cells. *Inflammatory Bowel Diseases* 11(10), 865-871.

Van Loo, J.A. (2004). Prebiotics promote good health: the basis, the potential, and the emerging evidence. *Journal of Clinical Gastroenterology* 38(6 Suppl), S70-75.

Wang, M., Ahrne, S., Antonsson, M. & Molin, G. (2004). T-RFLP combined with principal component analysis and 16S rRNA gene sequencing: an effective strategy for comparison of fecal microbiota in infants of different ages. *Journal of Microbiological Methods* 59(1), 53-69.

Wang, M., Ahrne, S., Jeppsson, B. & Molin, G. (2005). Comparison of bacterial diversity along the human intestinal tract by direct cloning and sequencing of 16S rRNA genes. *FEMS Microbiology Ecology* 54(2), 219-231.

Wang, M., Karlsson, C., Olsson, C., Adlerberth, I., Wold, A.E., Strachan, D.P., Martricardi, P.M., Aberg, N., Perkin, M.R., Tripodi, S., Coates, A.R., Hesselmar, B., Saalman, R., Molin, G. & Ahrne, S. (2008). Reduced diversity in the early fecal microbiota of infants with atopic eczema. *Journal of Allergy and Clinical Immunology* 121(1), 129-134.

Vanhoutte, T., De Preter, V., De Brandt, E., Verbeke, K., Swings, J. & Huys, G. (2006). Molecular monitoring of the fecal microbiota of healthy human subjects during administration of lactulose and *Saccharomyces boulardii*. *Applied and Environmental Microbiology* 72(9), 5990-5997.

Vanhoutte, T., Huys, G., De Brandt, E. & Swings, J. (2004). Temporal stability analysis of the microbiota in human feces by denaturing gradient gel electrophoresis using universal and group-specific 16S rRNA gene primers. *FEMS Microbiology Ecology* 48(3), 437-446.

Vasquez, N., Mangin, I., Lepage, P., Seksik, P., Duong, J.P., Blum, S., Schiffrin, E., Suau, A., Allez, M., Vernier, G., Treton, X., Dore, J., Marteau, P. & Pochart, P. (2007). Patchy distribution of mucosal lesions in ileal Crohn's disease is not linked to differences in the dominant mucosa-associated bacteria: a study using fluorescence in situ hybridization and temporal temperature gradient gel electrophoresis. *Inflammatory Bowel Diseases* 13(6), 684-692.

Weersma, R.K., van Dullemen, H.M., van der Steege, G., Nolte, I.M., Kleibeuker, J.H. & Dijkstra, G. (2007). Review article: Inflammatory bowel disease and genetics. *Alimentary Pharmacology &Therapeutics* 26 Suppl 2, 57-65.

Wehkamp, J., Salzman, N.H., Porter, E., Nuding, S., Weichenthal, M., Petras, R.E., Shen, B., Schaeffeler, E., Schwab, M., Linzmeier, R., Feathers, R.W., Chu, H., Lima, H., Jr., Fellermann, K., Ganz, T., Stange, E.F. & Bevins, C.L. (2005). Reduced Paneth cell alpha-defensins in ileal Crohn's disease. *Proceedings of the National Academy of Science U S A* 102(50), 18129-18134.

Wickens, K., Pearce, N., Crane, J. & Beasley, R. (1999). Antibiotic use in early childhood and the development of asthma. *Clinical and Experimental Allergy* 29(6), 766-771.

Viljanen, M., Savilahti, E., Haahtela, T., Juntunen-Backman, K., Korpela, R., Poussa, T., Tuure, T. & Kuitunen, M. (2005). Probiotics in the treatment of atopic eczema/dermatitis syndrome in infants: a double-blind placebo-controlled trial. *Allergy* 60(4), 494-500.

Wong, J.M., de Souza, R., Kendall, C.W., Emam, A. & Jenkins, D.J. (2006). Colonic health: fermentation and short chain fatty acids. *Journal of Clinical Gastroenterology* 40(3), 235-243.

Woodmansey, E.J. (2007). Intestinal bacteria and ageing. *Journal of Applied Microbiology* 102(5), 1178-1186.

Yoshimura, H.H., Graham, D.Y., Estes, M.K. & Merkal, R.S. (1987). Investigation of association of mycobacteria with inflammatory bowel disease by nucleic acid hybridization. *Journal of Clinical Microbiology* 25(1), 45-51.

Zijnge, V., Welling, G.W., Degener, J.E., van Winkelhoff, A.J., Abbas, F. & Harmsen, H.J. (2006). Denaturing gradient gel electrophoresis as a diagnostic tool in periodontal microbiology. *Journal of Clinical Microbiology* 44(10), 3628-3633.

Zoetendal, E.G., Akkermans, A.D. & de Vos, W.M. (1998). Temperature gradient gel electrophoresis analysis of 16S rRNA from human fecal samples reveals stable and host-specific communities of active bacteria. *Applied and Environmental Microbiology* 64(10), 3854-3859.

Zoetendal, E.G., Akkermans, A.D.L., Akkermans-van Vliet, W.M., de Visser, J.A.G.M. & De Vos, W.M. (2001a). The Host Genotype Affects the Bacterial Community in the Human Gastrointestinal Tract. *Microbial Ecology in Health and Disease* 13, 129-134.

Zoetendal, E.G., Ben-Amor, K., Akkermans, A.D., Abee, T. & de Vos, W.M. (2001b). DNA isolation protocols affect the detection limit of PCR approaches of bacteria in samples from the human gastrointestinal tract. *Systematic and Applied Microbiology* 24(3), 405-410.

Zoetendal, E.G., Collier, C.T., Koike, S., Mackie, R.I. & Gaskins, H.R. (2004). Molecular ecological analysis of the gastrointestinal microbiota: a review. *Journal of Nutrition* 134(2), 465-472.

Zoetendal, E.G., von Wright, A., Vilpponen-Salmela, T., Ben-Amor, K., Akkermans, A.D. & de Vos, W.M. (2002). Mucosa-associated bacteria in the human gastrointestinal tract are uniformly distributed along the colon and differ from the community recovered from feces. *Applied and Environmental Microbiology* 68(7), 3401-3407.

8 Acknowledgements

I want to thank my group of supervisors, **Janet Jansson**, **Magnus Rosenquist**, **Hans Jonsson** and **Stefan Roos**, for your work and interest in the field of gut microbiology during these years. You have all contributed to a stimulating scientific environment for me as a pH student.

In particular, I would like to thank my main supervisor **Janet.** Thank you for giving me this opportunity to dig into the human digestive system! It is so interesting, and I know that it is an interest that we share. You are also a role model for me when it comes to writing scientific texts and I know that this will influence me in times when I need to stand on my own. The time has NEVER been too limited for you when I have asked for your help. Your never-ending enthusiasm for science has been a great source of inspiration for me.

My associate supervisor **Magnus**. I am impressed by your quick adaptation into a new field of science and appreciate your unlimited knowledge about different uses and applications of molecular methods. Your guidance to find the "safe" path in the bioinformatic jungle was an important piece in the scientific puzzle. Your ability to be very supportive and at the same time critical is great! I have enjoyed all our discussions about scientific theories as well as childcare and your candy supply has sometimes been a lifesaver!

My associate supervisors **Stefan** and **Hasse**. It is always inspiring to discuss science with you, in particular topics related to intestinal microbiology and physiology. How many scientific myths remain to be turned inside out? Your interest in my results has been a great source of inspiration and our discussions have improved my ability in raising critical scientific questions.

My best wishes goes to the wonderful people, **Jocke**, **Åsa**, **Lotta**, **Mia U**, **Maria H**, **Karin**, **Veronica**, **Mia T**, **Ella**, and **Ben** that shared the lab with me during all (or part of) these years. It has been a pleasure! You have all contributed to a great working climate, especially during December, when all the Christmas decoration brightens up the lab ;-). In particular, I would like to thank, **Maria H** for keeping the lab in such a good shape, **Jocke** for a shared enthusiasm to discuss gut microbes, **Mia U,** and **Lotta** for answering all my questions about thesis formatting, **Karin** for the fruitful discussions about methods and data interpretation, and last but not least, **Ben** for your contribution to my project in the past year, it has been invaluable, and without you I would not have reached this far.

A also want to thank my co-authors outside the department for a fruitful collaboration. In particular, **Lars Engstrand** for your continuous supply of new interesting clinical material, **Mathilda Lindberg,** for your dedicated work in pushing the detection limit for bacteria in the stomach, and **Jonas Halfvarson** for extending the scientific horizon beyond the field of microbiology into gastroenterology.

My forever roommate **Klara**, for your cheerful attitude to life and for all the laughter you bring to our room. I cannot imagine a better roommate!

Cia Jernberg and **Anna Edlund**, for your supportive help, regarding methodological and analytical aspects of the T-RFLP method in the early days as a PhD student.

Chris, for your support and advice regarding statistical questions and computer software, I hope that I will learn R one day!

I would also like to thank all my **present** and **former** colleagues at the **Department of Microbiology** for your contribution to a stimulating atmosphere and nice lunch/coffee breaks. A special thanks to **Anki**, **Barbro**, **Susanne**, **Sture** and **Ingemar**, for all your help with various practical issues.

Jag vill också passa på att tacka mina vänner utanför forskarvärlden. Jag uppskattar verkligen er vänskap och tycker att det är synd att vi inte hinner ses oftare! Särskilt vill jag tacka **Anders.** Du har alltid funnits där som min bästa vän under alla dessa år. Även om du nu bor längre bort vet jag att du alltid finns där och ställer upp om det behövs.

Jag vill tillägna ett stort tack till mina svärföräldrar **Karin** och **Christer** för att ni så öppenhjärtligt välkomnat mig till familjen Fredlund, och för att ni är en så bra mormor och morfar till Martin och Joakim!

Mamma, **pappa** och mina systrar **Sara** och **Anna,** för att ni alltid finns där, all den kärlek ni givit mig och för att ni alltid ställt upp om jag behövt er hjälp. Det är er förtjänst att jag blivit den person jag är idag, och det är jag stolt över!

Joakim och **Martin**, ni har gett mitt liv en helt ny dimension! Att få busa, mysa eller bara vara med er är helt underbart!

Lisa, det finns inte ord att beskriva den uppskattning jag känner att jag får dela mitt liv med dig!! Jag är enormt imponerad av DIN prestation under den senaste hektiska tiden, då du höggravid fått tagit ett stort ansvar för att ta hand om familjen! Du är en fantastisk mamma till våra barn jag är lycklig över att få dela framtiden med dig. Jag älskar dig!