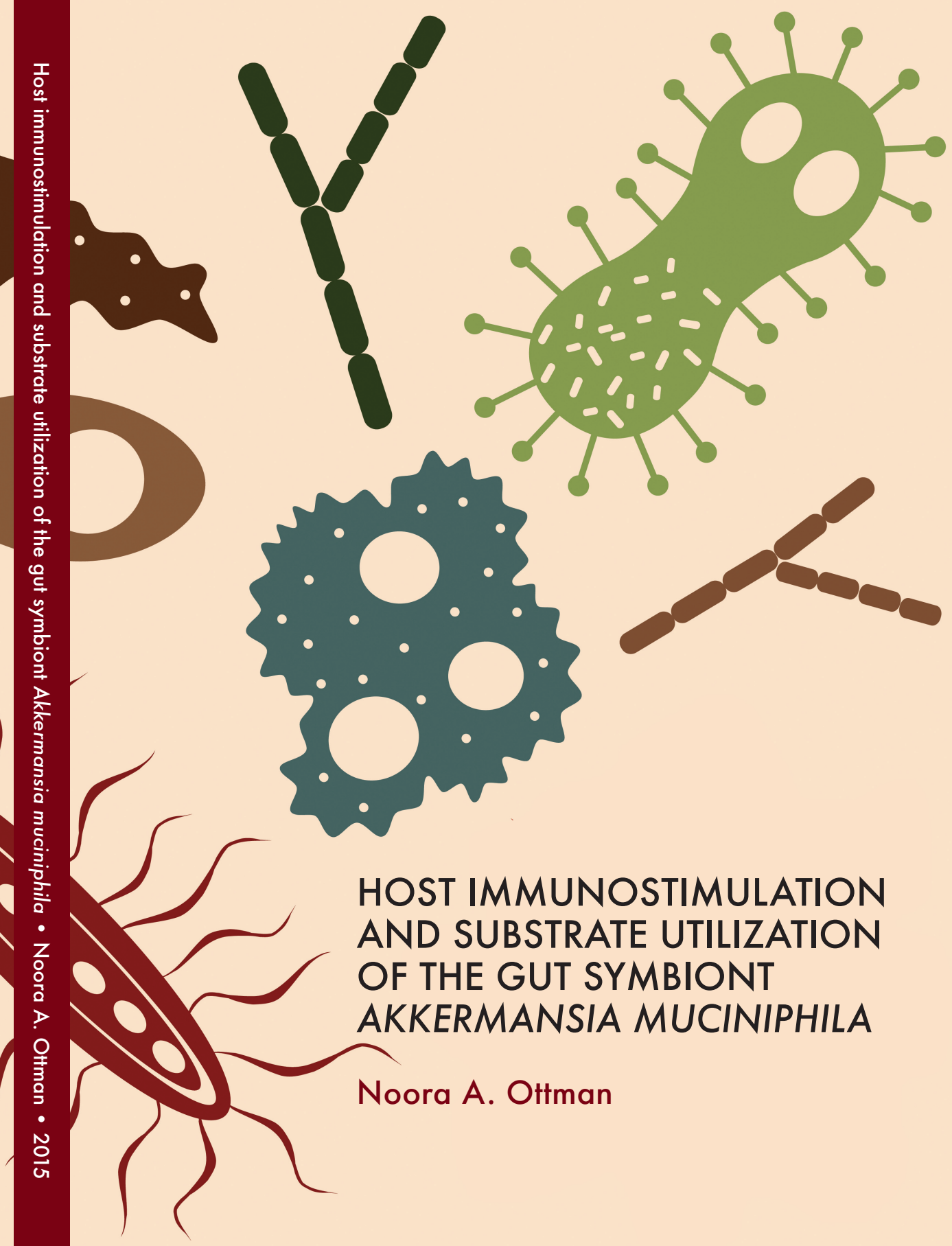


HOST IMMUNOSTIMULATION AND SUBSTRATE UTILIZATION OF THE GUT SYMBIONT *AKKERMANSIA MUCINIPHILA*

Noora A. Ottman

Host immunostimulation and substrate utilization of the gut symbiont *Akkermansia muciniphila* • Noora A. Ottman • 2015

Host immunostimulation and substrate utilization of the gut symbiont *Akkermansia muciniphila*

Noora A. Ottman

Thesis committee

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

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

Co-promotor
Dr C. Belzer
Postdoctoral researcher, Laboratory of Microbiology

Other members
Prof. Dr E.J. Smid, Wageningen University
Prof. Dr H.F.J. Savelkoul, Wageningen University
Prof. Dr T. Van de Wiele, Ghent University, Belgium
Dr A.C. Ouwehand, DuPont Nutrition and Health, Kantvik, Finland

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

Host immunostimulation and substrate utilization of the gut symbiont *Akkermansia muciniphila*

Noora A. Ottman

Thesis

submitted in fulfilment of the requirements for the degree of doctor
at Wageningen University
by the authority of the Rector Magnificus
Prof. Dr A.P.J. Mol,
in the presence of the
Thesis Committee appointed by the Academic Board
to be defended in public
on Wednesday 16 September 2015
at 4 p.m. in the Aula.

Noora A. Ottman
Host immunostimulation and substrate utilization of the gut symbiont *Akkermansia muciniphila*,
208 pages.

PhD thesis, Wageningen University, Wageningen, NL (2015)
With references, with summary in English

ISBN 978-94-6257-456-4

Summary

The human gastrointestinal tract is colonized by a complex community of micro-organisms, the gut microbiota. The majority of these are bacteria, which perform various functions involved in host energy metabolism and immune system stimulation. The field of gut microbiology is continuously expanding as novel species are isolated and high-throughput techniques are developed. The research focus is shifting from DNA-based techniques, looking at microbial community composition, to techniques relying on analysis of RNA and proteins, which reveal more about the activity and functionality of the microbiota.

The mucosa-associated microbiota forms a distinct population in the gut, and is influenced by the close proximity of the epithelial layer and nutrients present in the mucus layer. One of the key players in this community is the mucus degrader *Akkermansia muciniphila*. This Gram-negative, anaerobic bacterium can use mucin, the main component of mucus, as the sole carbon and nitrogen source for growth. *A. muciniphila* belongs to the phylum *Verrucomicrobia* and is present in the majority of humans, starting from early life. Interestingly, the levels of *A. muciniphila* are negatively correlated with several disorders, including inflammatory bowel diseases and diabetes. *A. muciniphila* lives in a symbiosis with its host, harvesting energy from mucin; whether the relationship is mutualistic, and thereby also beneficial to the host, remains to be discovered. In this thesis, the ability of *A. muciniphila* to utilize the host-derived glycans mucin and human milk oligosaccharides was studied in detail. In addition, the host-bacterial interactions were examined by immunological assays, focusing especially on the effect of *A. muciniphila* outer membrane proteins on host immune response.

The genome of *A. muciniphila* encodes numerous enzymes involved in mucin degradation. Transcriptome analysis comparing the gene expression of *A. muciniphila* grown on mucin or the non-mucin sugar glucose confirmed the activity of these genes and revealed most of them to be upregulated in the presence of mucin. This was also confirmed by a proteome analysis, reinforcing the adaptation of *A. muciniphila* to the mucosal environment. A genome-based metabolic model was constructed to test amino acid auxotrophy, vitamin biosynthesis, and sugar-degrading capacities of *A. muciniphila*. The model predicted *A. muciniphila* to be able to synthesize all the essential amino acids, with the exception of threonine, which was added to the mucin-free medium designed to test *A. muciniphila* growth on single sugars. *A. muciniphila* was able to individually metabolize all the main monomeric sugars present in mucin, albeit with limited efficiency in comparison to mucin. As mucin shares structural similarities with human milk oligosaccharides (HMOs), which stimulate the

bacterial community colonizing the gut in early life, growth of *A. muciniphila* on human milk and its components was tested. *A. muciniphila* showed metabolic activity on human milk and one of the HMOs, 2'-fucosyllactose. Comparison of *A. muciniphila* activity during growth on human milk or mucin revealed that the expression of genes involved in mucin degradation was similar for both experimental conditions, suggesting that *A. muciniphila* might be capable of also using the corresponding gene products for utilization of human milk glycans. The capacity to survive in the early life environment by degrading and consuming human milk components would be beneficial for *A. muciniphila* during initial colonization before reaching the mucosal layer in the intestine.

Several mouse studies have indicated that *A. muciniphila* is able to modulate the host immune system, possibly to the benefit of the host, but not much is known about its immunological mechanism of action. The cell envelope structures of bacteria can have a big influence on their immunostimulatory capacities, and therefore the outer membrane (OM) proteome of *A. muciniphila* was characterized. The membrane structure of *A. muciniphila* is also of interest because it belongs to the *Planctomycetes-Verrucomicrobia-Chlamydiae* superphylum, which contains bacteria with features that differentiate them from classical Gram-negative bacteria, including a complex endomembrane system. Mass spectrometry data, coupled with bioinformatics analysis, revealed the presence of highly abundant OM proteins involved in secretion, transport and biogenesis of the Gram-negative membranes, as well as proteins predicted to take part in formation of the fimbriae-like structures observed in *A. muciniphila* by electron microscopy. Live *A. muciniphila* and the identified OM proteins induced production of a wide range of cytokines and activated the intestinal Toll-like receptors 2 and 4. Moreover, a 30 kDa protein that was predicted to form a part of the fimbriae, increased transepithelial resistance, indicating it may be involved in improving gut barrier function.

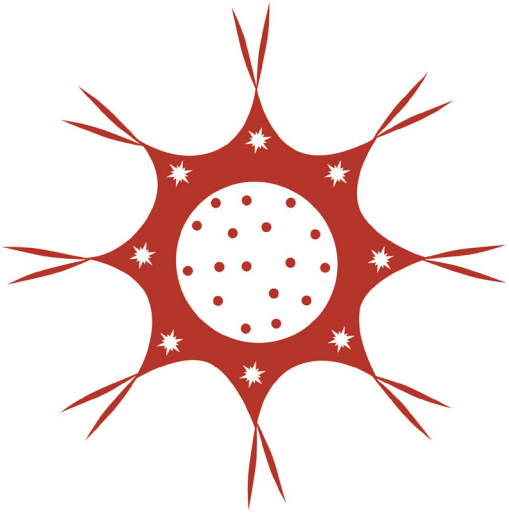
Based on the evidence from *in vitro* and *in vivo* studies, *A. muciniphila* is a promising candidate for a next-generation probiotic. However, further confirmation of causal relationships between disease development and presence of this species in the gut is required. The findings of this thesis provide valuable insight into the bacterial lifestyle and host interactions of the gut symbiont *A. muciniphila*.

TABLE OF CONTENTS

Mummolle

Chapter 1

General introduction and thesis outline



A woman is known by the company she keeps

What does your microbiota tell about you? More than you may think. The microbial communities covering our bodies, inside and out, are associated with age, genetic background, geographic location, diet, health status and use of medication (Ursell *et al.*, 2012; Xu *et al.*, 2015). This is particularly true for the gut microbiota (Lozupone *et al.*, 2012). With over a thousand identified microorganisms, including members of *Eukarya*, *Archaea* and *Bacteria*, the microbial community residing in the gastrointestinal (GI) tract is essential to the well-being of humans (Rajilic-Stojanovic and de Vos, 2014).

The GI tract extends from the mouth to the anal canal, dividing into several separate parts. The oral cavity, and even the acidic stomach, have their distinct microbial communities (Lopetuso *et al.*, 2014; Zaura *et al.*, 2014), but from here on the focus is on the microorganisms present in the small and large intestine. The small intestine can be divided into duodenum, jejunum and ileum, whereas the large intestine is divided into cecum, ascending, transverse, descending and sigmoid colon (Figure 1). Physicochemical differences, including pH, redox potential, nutrient supplies, transit time and host secretions account for the differences in microbial density and composition between these compartments (Gerritsen *et al.*, 2011). The small intestinal microbiota contains $10^2 – 10^7$ cells per gram of content, while the levels go up to $10^{10} – 10^{12}$ cells/g in the large intestine and feces (Dethlefsen *et al.*, 2006; El Aidy *et al.*, 2015).

Each person harbors their own unique microbiota, with around 160 different microbial species per individual (Qin *et al.*, 2010). A part of this community forms a stable core microbiota, which persists throughout life and can be largely shared between individuals (Faith *et al.*, 2013). The majority of the bacteria in the gut microbiota belong to the phyla *Bacteroidetes*, *Firmicutes* and *Actinobacteria*, with members of *Proteobacteria*, *Fusobacteria*, *Lentisphareae*, *Spriochaetes* and *Verrucomicrobia* also present (Zoetendal *et al.*, 2008). The small intestinal microbiota is rich in *Streptococcus* spp. and *Veillonella* spp., and shows large diet-related fluctuations (Booijink *et al.*, 2010b; van den Bogert *et al.*, 2013). The large intestinal microbiota is strongly affected by slow transit time, anaerobic environment and presence of complex carbohydrates, which were not digested in the small intestine. Contrary to the small intestinal microbiota, the large intestinal microbiota is more stable and has a higher species diversity (Eckburg *et al.*, 2005; Salonen and de Vos, 2014). Importantly, along the GI tract there are two separate bacterial niches, the lumen and the mucus layer. Considerable variation between surface-adherent and luminal microbial populations have been observed (Ringel *et al.*, 2015; Zoetendal *et al.*, 2002).

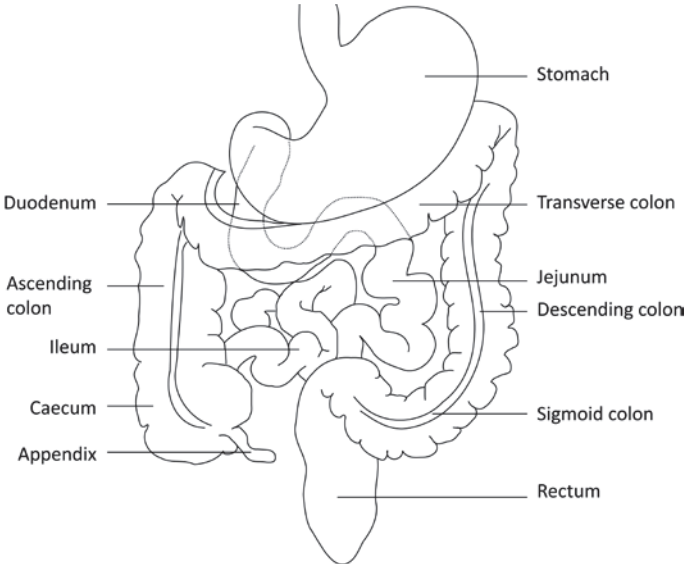


Figure 1. Schematic presentation of the human lower gastrointestinal tract.

The relationship between the host and the intestinal microbes is mutualistic, indicating that both the microorganisms and the human benefit from the activity of each other. Bacteria are beneficial for the host for example by harvesting energy, competitively excluding pathogens and producing essential vitamins and amino acids (Rajilic-Stojanovic, 2013). The host provides a living environment for the bacteria, which use both diet-derived and host-derived nutrients as energy sources. In the intestine bacterial activity leads to production and metabolism of various metabolites, including short-chain fatty acids (SCFA), branched-chain fatty acids (BCFA), vitamins and nitrogenous compounds (Cummings and Macfarlane, 1997). SCFA are produced by microbial fermentation of both diet and host-derived complex polysaccharides and can have an effect on the health of the host through various mechanisms. The main SCFA are acetate, propionate and butyrate. They have been implicated to play a part in metabolic and inflammatory disorders such as obesity, diabetes and inflammatory bowel diseases (IBD), through the activation of specific G protein-coupled receptors and modification of transcription factors of the host (Puertollano *et al.*, 2014). After absorption from the colon, SCFA are metabolized at different sites of the body. Colonic mucosal epithelial cells use butyrate as substrate for energy producing pathways, liver cells use propionate and residual butyrate as substrate for hepatic gluconeogenesis, and muscle cells conserve energy from the oxidation of acetate (Wong *et al.*, 2006). The modulation of the gut microbiota by pro- and prebiotics is often targeted at increasing the levels of beneficial SCFA (De Preter *et al.*, 2011).

Metabolism of host-derived glycans in the gut

In addition to diet-derived glycans, the gut microbiota can feed on host-derived glycans. The main sources of these are the mucus layer, which provides a stable carbon and energy source, and human milk, which is only present in the intestine early in life. Both of these contain complex oligosaccharides, which stimulate the growth of bacteria that are able to utilize them, and thus select for specific bacterial populations.

Mucus layer and mucin-degrading bacteria

The entire gastrointestinal tract is covered by a mucus layer consisting mainly of large glycoproteins (mucins), water (~95 %) and low amounts of salts, lipids and several intestinal proteins (lysozyme, defensins), immunoglobulins, growth factors, trefoil factors and released epithelial cells (Clamp and Ene, 1989; Hamer *et al.*, 2009). The thickness of the mucus layer varies depending on the location. In the small intestine it consists of a single layer because food components need to be taken up. However, in the stomach and colon there are two layers as there is a need for additional protection. The looser outer layer is colonized by mucin-degrading bacteria whereas the dense inner layer is largely devoid of bacteria (Johansson *et al.*, 2008). Mucus is constantly secreted by goblet cells leading to up to 10 liters of mucus production each day in humans (Cone, 2009). In addition to serving as a niche for commensal bacteria, the mucus layer has many other purposes. It can function as a lubricant facilitating passage of food particles, as a selective barrier to allow the passage of nutrients to the epithelial cells, and as a protective system against mechanical damage, pathogens or harmful substances such as toxins (Gouyer *et al.*, 2011).

The main components of the mucus layer, mucins, are complex glycoproteins characterized by a heavily glycosylated protein backbone. Approximately 70 – 80 % of the glycosylation is O-linked. The MUC gene family, with more than 20 genes identified so far, is responsible for production of mucin in humans (Dekker *et al.*, 2002; Gum, 1992). Mucin molecules are grouped into gel-forming mucins (MUC2, MUC5AC, MUC5B, MUC6), which form extremely large polymers, and transmembrane mucins (MUC1, MUC3, MUC4, MUC12, MUC13, MUC16, MUC17), which cover the apical membranes of epithelial cells. MUC5AC, MUC6 and MUC1 are found in the stomach, and MUC2 and MUC6 along with several different transmembrane mucins in the small intestine (Pelaseyed *et al.*, 2014). In the colon the gel-forming mucus is MUC2, and the transmembrane mucins are composed of MUC1, MUC3, MUC4, MUC12, MUC13 and MUC17 (Johansson *et al.*, 2013).

Mucin is rich in serine and threonine, and these together with proline, alanine and glycine make up to 80 % of the total amino acid content (Schrager, 1970). All mucins contain at least one so called PTS domain, which is abundant in proline, threonine and serine, and these PTS domains account for the great variability in mucin length and extent of glycan attachment. The oligosaccharide structures in mucin are highly diverse with fucose, galactose, N-acetylgalactosamine (GalNAc), N-acetylglucosamine (GlcNAc) and N-acetylneuraminic acid (sialic acid) as the main sugar monomers (Johansson *et al.*, 2011). Out of the eight core structures of the mucin *O*-glycan chain, core 1 – 4 glycans are most commonly found in intestinal mucins (Tailford *et al.*, 2015a) (Figure 2A). The core structures can be extended with different sugar residues by attaching one of the multiple termination structures (Figure 2B). The presence of these structures varies along the GI tract and can even be affected by the gut bacteria (Hooper *et al.*, 1999; Robbe *et al.*, 2004).

For a bacterium to be able to use mucin as an energy source, it must possess the enzymes needed for mucin degradation. Some bacteria may benefit from mucin also indirectly by feeding off the sugars released from mucin by other bacteria. Various members of the phyla *Bacteroidetes*, *Firmicutes*, *Actinobacteria* and *Verrucomicrobia* have been shown to be able to consume mucins. One of the well-studied mucin-degrading organisms is *Bacteroides thetaiotaomicron*, whose genome contains specific polysaccharide-utilization loci (PULs), which include genes coding for mucin-degrading enzymes (Martens *et al.*, 2008). PULs are also found in the genome of *Bacteroides fragilis* (Marcobal *et al.*, 2011). In addition, several bifidobacteria, including *Bifidobacterium bifidum* and *Bifidobacterium longum* subsp. *infantis*, are capable of growing on mucin as the sole carbon source (Abe *et al.*, 2010; Hoskins *et al.*, 1985; Ruas-Madiedo *et al.*, 2008). Two members of the *Lachnospiraceae*, *Ruminococcus gnavus* and *Ruminococcus torques*, can also utilize mucin (Hoskins *et al.*, 1985; Png *et al.*, 2010). Recently, it has been shown that the genome of *R. gnavus* encodes a specific trans-sialidase that allows this potential pathogen to degrade sialomucin in a dedicated way (Tailford *et al.*, 2015b). Finally, *Akkermansia muciniphila* was isolated from a human fecal sample due to its ability to thrive on mucin (Derrien *et al.*, 2004). *A. muciniphila* and its mucin-degrading capacities will be discussed in detail later in this chapter. In addition to the abovementioned commensal bacteria, many pathogens have evolved enzymes to degrade mucus (McGuckin *et al.*, 2011).

The essential enzymes needed for mucin degradation are proteases, sulfatases and multiple glycosyl hydrolases, such as sialidases, fucosidases, ß-galactosidases, α-N-acetylglucosaminidases and α-N-acetylgalactosaminidases (Tailford *et al.*, 2015a). Generally bacteria have the ability to produce only some of these enzymes and the

complete degradation of mucin molecules requires the cooperation of several bacterial species.

Human milk oligosaccharides

The other important host-derived glycans in regards to the microbiota are the human milk oligosaccharides (HMOs) present in human milk. While lactose is the most abundant source of carbohydrate in human milk, HMOs are particularly important in stimulating the bacterial community colonizing the gut in early life. HMOs are only partially digested in the babies' small intestine and thus they function as natural prebiotics, selectively stimulating the growth of beneficial bacteria in the colon.

Human milk contains 5 – 15 g/l HMOs, and over 200 different HMO structures exist (Zivkovic *et al.*, 2011). Interestingly, the structures of HMOs resemble those of mucin, suggesting that bacteria capable of degrading mucin may also be able to consume HMOs. The monosaccharides present in HMOs are glucose, galactose, N-acetylgalactosamine, fucose and sialic acid (N-acetylneuraminic acid). Lactose is found at the reducing end of all milk oligosaccharides described so far. Addition of fucose or N-acetylneuraminic acid to lactose by the action of glycosyltransferases leads to formation of small soluble milk glycans, which are resistant to digestion by human enzymes (Engfer *et al.*, 2000) (Figure 2C). More complex HMOs are formed by sequential elongation and decoration of these simple structures. HMOs can be linear, branched, fucosylated and/or sialylated, usually with lacto-*N*-tetraose or lacto-*N*-neotetraose as core structures (Kobata, 2010) (Figure 2C). The genotype of the mother plays a role in which HMOs are present in human milk as the enzymes necessary for the synthesis of these structures are not equally distributed amongst humans. Fucosylation of HMOs is dependent on the status of the mother in regards to two histo-blood group systems: Lewis and Secretor (Thurl *et al.*, 1997). The Secretor gene encodes the α1-2 fucosyltransferase FUT2, and the Lewis gene encodes the α1-3/4 fucosyltransferase FUT3. Based on the production of these fucosyltransferases, four milk groups can be distinguished. These genotypes affect the fucosylation of mucin in a similar manner (Figure 2B). The secretor genotype has been shown to be associated with gut microbiota composition (Rausch *et al.*, 2011; Wacklin *et al.*, 2011; Wacklin *et al.*, 2014).

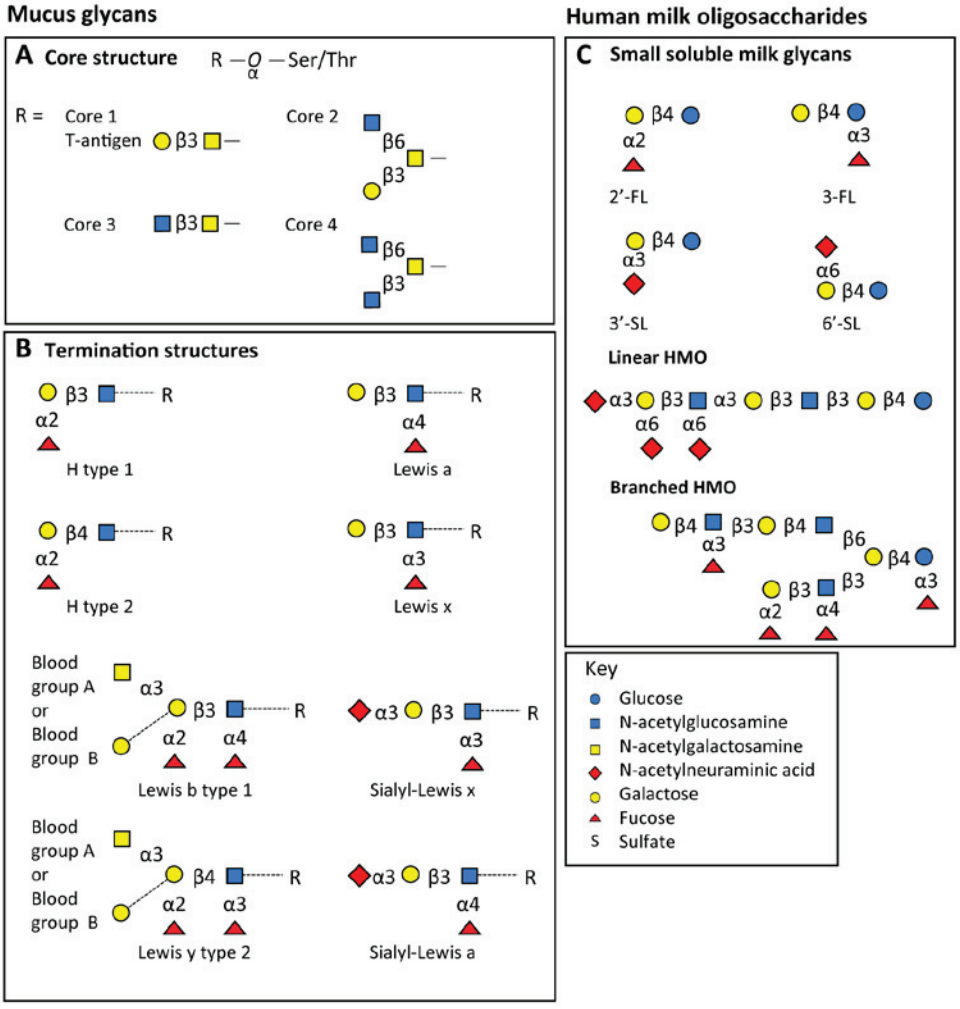


Figure 2. Mucus glycans and human milk oligosaccharides share structural similarities. (A) The four most commonly found mucin core structures in the intestine. (B) Mucin molecule termination structures. (C) Examples of small soluble milk glycans, linear HMO and branched HMO. 2'-FL, 2′-fucosyllactose; 3-FL, 3-fucosyllactose; 3'-SL, 3′-sialyllactose; 6'-SL, 6′-sialyllactose. Adapted from (Castanys-Munoz *et al.*, 2013; Ouwerkerk *et al.*, 2013; Wu *et al.*, 2010).

Host-derived glycans are particularly important in early life when few types of glycans transit the gut, and the carbohydrate composition is an important factor in guiding the establishment of the gut microbiota. Colonization of the gut and composition of the infant microbiota is further discussed in **Chapter 2**.

The mucosal immune system

The human host is continuously exposed to the external environment through food, microorganisms and other substances that enter the GI tract. The mucosal immune system is thus faced with the challenge of discriminating between potential pathogens and commensal microbiota. The gut-associated lymphoid tissue (GALT) contains immune cells responsible for coordinating this antigenic challenge by remaining immunologically hypo-responsive to commensal bacteria while retaining the capacity to respond to pathogens (Artis, 2008).

In the intestine, the epithelial monolayer functions as a barrier between the intestinal lumen and lamina propria where the immune cells reside. As discussed above, the mucus layer forms a protective layer on top of the epithelial cells. The epithelium contains several different cell types including enterocytes, enteroendocrine cells, paneth cells and goblet cells, which secrete mucin as mentioned before. Enteroendocrine cells secrete peptide hormones and regulate food intake, glucose homeostasis and gastric emptying (Mellitzer *et al.*, 2010). Paneth cells produce bactericidal products such as lysozyme, defensins, phospholipases and RegIII proteins (Sato *et al.*, 2011). Enterocytes are the most abundant cell types and their main function is nutrient absorption and metabolism. Additionally, they take up and process luminal antigens and thereby play an immunosensory role in the gut by secreting signaling molecules, such as cytokines, in response to luminal signals (Snoeck *et al.*, 2005).

Peyer's patches are specialized lymphoid structures, which are located in the mucosa of the small intestine. They form areas where luminal antigens are sampled via microfold (M) cells, which lack surface microvilli and the normal mucus layer (Mowat, 2003). In addition to M cells, sampling of antigens is done by specialized dendritic cells, which can penetrate between the epithelial cells and sample directly from the lumen (Mann *et al.*, 2013). Antigen sampling is crucial for downstream immunological responses. M cells pass on the antigens to professional antigen presenting cells (APCs) such as macrophages, B cells and dendritic cells (Artis, 2008). The antigens are presented to naive T cells which in turn differentiate into defense-directed T helper (T_H) cells and tolerance-associated regulatory T (T_{reg}) cells (Hooper and Macpherson, 2010). Differentiated T cells secrete cytokines, which can be either proinflammatory or anti-inflammatory and a balanced ratio between these controls immune homeostasis. As the microbiota acts as a major source of antigens for the GI immunosensory cells, it critically affects cellular host responses and is capable of regulating and sustaining immune homeostasis. The differentiation of B cells leads to cells that produce immunoglobulin A (IgA), which is transported to the epithelium and released as

secretory IgA (sIgA). Commensal bacteria and soluble antigens are coated by sIgA to inhibit their binding to the epithelium, and by doing this sIgA promotes intestinal barrier function and helps to maintain host-bacterial mutualism (Fagarasan *et al.*, 2010).

The ability of the immune cells described above to recognize, capture and process antigens is based on their pattern recognition receptors (PRRs), which recognize microbe-associated molecular patterns (MAMPs) (Wells *et al.*, 2010). These patterns are shared by a variety of microbial species and include for example lipopolysaccharide (LPS), lipoteichoic acid (LTA), flagellin and bacterial DNA. Most PRRs belong to one of the four families: Toll-like receptors (TLRs), NOD-like receptors (NLRs), C-type lectin receptors (CLRs) and retinoic acid-inducible gene I (RIG-I)-like receptors (RLRs) (Kamada *et al.*, 2013). The receptors are expressed either at the cell surface or in intracellular membrane vesicles. Following antigen binding to a PRR, a complex signaling cascade is triggered that leads to expression of genes associated with defense mechanisms. In the case of TLRs, the specificity of the response is governed by differential recruitment of the adaptor proteins MyD88, MAL, TRIF, SARM and TRAM, which in turn leads to activation of key regulatory pathways such as nuclear factor kappa B (NF-κB) or mitogen-activated protein (MAP) kinases (O'Neill and Bowie, 2007). The molecular structures recognized by the known human TLRs and NLRs, and their cellular locations are shown in Table 1.

Animal studies have shown that the host intestinal immune system requires the presence of commensal microbes for its development and proper functioning (Hooper *et al.*, 2012). Germ-free mice have a severely impaired immune system, and they lack for example sIgA and isolated lymphoid follicles in the small intestine. Perturbation of the balance between the host and the microbiota caused by microbial aberrations can lead to immune dysregulation and susceptibility to diseases. Changes in microbiota composition are associated with many diseases, including the autoimmune diseases type 1 diabetes, coeliac disease and rheumatoid arthritis (McLean *et al.*, 2015). However, the question of whether the microbial deviations are a cause or an effect still remains, and further research is needed to examine the interplay between the host and its microbes. Restoration of a healthy gut microbiota by fecal microbiota transplantation (FMT) has the potential to be an efficient therapeutic option in the treatment of several diseases, and also an effective way to study the involvement of the gut microbiota in these diseases (Cammarota *et al.*, 2014). FMT has shown promising results in the treatment of recurrent *Clostridium difficile* infection, metabolic syndrome and IBD (Angelberger *et al.*, 2013; van Nood *et al.*, 2013; Vrieze *et al.*, 2012).

Table 1. Localization, ligands and their origin of Toll-like receptors (TLRs) and NOD-like receptors (NLRs). Abbreviations used in this table: G-, Gram-negative bacteria; G+, Gram-positive bacteria.

Receptor	Localization	Ligand	Origin of ligand
TLR1	Cell surface	Triacyl lipoproteins	G- bacteria
TLR2	Cell surface	Lipoteichoic acid	Bacteria
		Lipoproteins	
		Peptidoglycan	
		Lipomannan	Mycobacteria
		Zymosan	Yeast
TLR3	Endolysosomal membrane	Double-stranded RNA	Viruses
TLR4	Cell surface	Lipopolysaccharide	G- bacteria
	Endolysosomal membrane		
TLR5	Cell surface	Flagellin	Bacteria
TLR6	Cell surface	Diacyl lipoproteins	G+ bacteria
		Lipoteichoic acid	Mycoplasma
		Zymosan	Yeast
TLR7	Endolysosomal membrane	Single-stranded RNA	RNA viruses
			Bacteria
TLR8	Endolysosomal membrane	Single-stranded RNA	Viruses
			Bacteria
TLR9	Endolysosomal membrane	CpG-containing DNA	Bacteria
		RNA:DNA hybrids	Viruses
TLR10	Cell surface	Unknown	Unknown
NOD1	Cell cytoplasm	Meso-diaminopimelic acid	bacteria
NOD2	Cell cytoplasm	Muramyl dipeptide	Bacteria

Akkermansia muciniphila – a mucin specialist like no other

The Gram-negative, non-motile, anaerobic bacterium *Akkermansia muciniphila* was isolated from a human fecal sample in a quest to discover new mucin-degrading bacteria (Derrien *et al.*, 2004). *A. muciniphila* cells are oval-shaped and most often observed singly or in pairs (Figure 3). The complete genome of *A. muciniphila* MucT (ATTC BAA-835) is composed of one circular chromosome of 2,664,102 bp and has a total of 2,176 predicted protein-coding sequences (van Passel *et al.*, 2011).

A. muciniphila is the only intestinal member of the *Verrucomicrobia* phylum described so far, and it is present in the majority of the human population. The levels of *A. muciniphila* reach 1 to 4 % of the total fecal microbiota, and it has also been detected abundantly in biopsies from the mucosal layer in the colon (Collado *et al.*, 2007; Derrien *et al.*, 2008; Lyra *et al.*, 2012; Png *et al.*, 2010). *A. muciniphila* colonizes the infant gut relatively early, as fecal samples from 1-month-old infants already have low levels of this bacterium (Collado *et al.*, 2007). After this, the amount of *A. muciniphila* increases, reaching a level close to that observed in adults within a year. In addition to Finnish and Danish infants, *A. muciniphila* has been observed to be present in Malawian infants, albeit in low levels (Grzeskowiak *et al.*, 2012). Besides humans, *Akkermansia*-like organisms are found in many mammals, such as mice (Sonoyama *et al.*, 2010), hamsters (Sonoyama *et al.*, 2009), squirrels (Carey *et al.*, 2013), guinea pigs (Hildebrand *et al.*, 2012), rabbits (Zeng *et al.*, 2015), donkeys (Liu *et al.*, 2014) and also non-mammals such as pythons (Costello *et al.*, 2010), chickens (Belzer and de Vos, 2012) and zebrafish (Roeselers *et al.*, 2011).

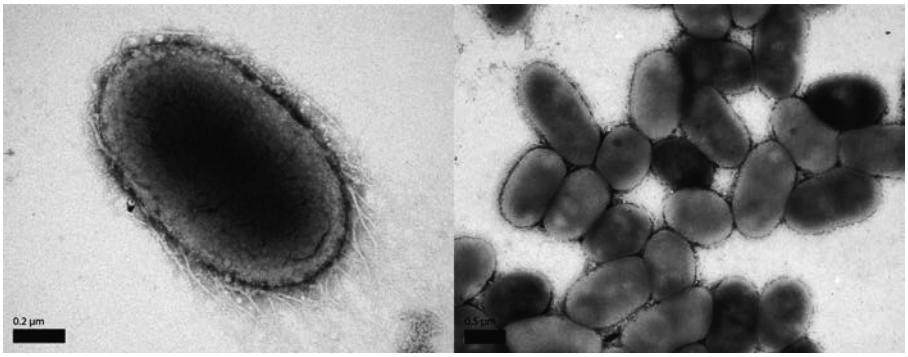


Figure 3. Electron microscope images of *A. muciniphila*. Images by Justus Reunanen, University of Helsinki.

A. muciniphila is capable of using mucin as the sole carbon and nitrogen source and has so far been shown to grow on human MUC2 and pig gastric mucin (Tailford *et al.*, 2015a). The genome sequence revealed 61 proteins (2.8 % of all proteins) predicted to be involved in the degradation of mucin (van Passel *et al.*, 2011). These proteins include numerous enzymes such as glycosyl hydrolases, proteases, sulfatases and sialidases. No mucus-binding domains were found in the genome, and it was later shown that the bacteria do not adhere to human colonic mucus but do bind to intestinal cells (Reunanen *et al.*, 2015). A metaproteomic study of two human fecal samples identified close to 200 proteins deriving from *A. muciniphila*-like bacteria (Rooijers *et al.*, 2011). Apart from the housekeeping proteins, the

majority of proteins were predicted to be involved in carbohydrate transport and metabolism as well as amino acid transport and metabolism. Various mucin-degrading enzymes were among the identified proteins, reflecting the intestinal lifestyle of *A. muciniphila*. The high mucin-degrading capacity has also been shown *in vivo* in mice with a complex gut ecosystem. In this setting *A. muciniphila* and *Bacteroides acidifaciens* were the most important host-protein foragers as determined by stable isotope labeling of threonine (Berry *et al.*, 2013). During nutrient deprivation *A. muciniphila* has a competitive advantage over bacteria not capable of degrading mucin. This is demonstrated in studies where the relative abundance of *A. muciniphila* increases during fasting in humans and animals (Carey *et al.*, 2013; Dill-McFarland *et al.*, 2014; Remely *et al.*, 2015a; Sonoyama *et al.*, 2009). As a result of mucin fermentation *A. muciniphila* produces acetate and propionate. In addition to mucin, *A. muciniphila* can grow on glucose, GlcNAc, and GalNAc, but with a lower growth rate and final density than on mucin and only in the presence of a large amount of tryptone, peptone, casitone and yeast extract (Derrien *et al.*, 2004).

Interestingly, soon after its discovery, the levels of *A. muciniphila* in the gut have been negatively correlated with many diseases. Some examples of these include IBD (James *et al.*, 2014; Png *et al.*, 2010; Rajilic-Stojanovic *et al.*, 2013; Vigsnaes *et al.*, 2012), appendicitis (Swidsinski *et al.*, 2011), obesity (Karlsson *et al.*, 2012; Remely *et al.*, 2015b), autism (Wang *et al.*, 2011) and atopy (Candela *et al.*, 2012; Drell *et al.*, 2015). The most convincing evidence of the beneficial effect of *A. muciniphila* comes from studies linking the bacteria to metabolic disorders, such as diabetes. The administration of the prebiotic oligofructose to genetically obese mice increased the abundance of *A. muciniphila* by ~100-fold (Everard *et al.*, 2011). Thereafter, a decrease of *A. muciniphila* in obese and type 2 diabetic mice was restored with prebiotic feeding, and increase of the bacteria was correlated with an improved metabolic profile (Everard *et al.*, 2013). In the same study treatment of mice with live but not dead *A. muciniphila* reversed the high-fat diet-induced metabolic disorders and increased the intestinal levels of endocannabinoids that control inflammation, the gut barrier, and gut peptide secretion. These results were soon confirmed by another study, where feeding *A. muciniphila* to mice on a high-fat diet enhanced their glucose tolerance and increased the number of goblet cells and adipose tissue-resident CD4 Foxp3 Tregs (Shin *et al.*, 2014). In the same study, treatment of the obese mice with metformin, a medicine used to treat type 2 diabetes, was associated with increased levels of *A. muciniphila*, along with improvement of the glycaemic profile. The enrichment effect of metformin on the growth of *A. muciniphila* was later confirmed *in vitro* and *in vivo* (Lee and Ko, 2014). Possible protective effects of *A. muciniphila* on both type 1 and type 2

diabetes have been observed also in several other studies (Anhe *et al.*, 2014; Ellekilde *et al.*, 2014; Hansen *et al.*, 2012; Marietta *et al.*, 2013; Zhang *et al.*, 2013).

Based on these findings, *A. muciniphila* has been proposed as a candidate for a next-generation probiotic (Neef and Sanz, 2013). The results point to an anti-inflammatory role of *A. muciniphila* via interactions with intestinal epithelial cells, but the exact mechanisms of the host-microbial interaction behind these beneficial effects remain unknown. Furthermore, it is not clear if *A. muciniphila* only exerts positive effects on the host. Increased levels of *A. muciniphila*, or genes deriving from it, have been positively correlated with type 2 diabetes (T2D) (Qin *et al.*, 2012), dextran sodium sulfate (DSS)-induced colitis (Berry *et al.*, 2012; Hakansson *et al.*, 2014), and colorectal cancer (CRC) (Baxter *et al.*, 2014; Weir *et al.*, 2013; Zackular *et al.*, 2013). However, some caution is needed for the interpretation of these studies. The association to T2D could not be reproduced in a well-controlled cohort in absence of the use of metformin, suggesting that the initial association of *A. muciniphila* with T2D may have been due to metformin treatment (Karlsson *et al.*, 2013). Moreover, *A. muciniphila* extracellular vesicles ameliorated DSS colitis (Kang *et al.*, 2013), and the increase of *A. muciniphila* in the DSS models may be just a reflection of increased mucus production. Finally, in mice or humans with CRC food intake is reduced, which is known to increase the levels of *A. muciniphila* as mentioned above.

Going beyond genomics

For most of its existence, the field of gut microbiology has been dependent on the culture and isolation of microbes in the laboratory. However, recent advances in culture-independent, sequence-based methods have revolutionized the manner in which bacterial communities in the gut can be studied. Metagenomics and other 'meta-omics' applications provide powerful tools for studying gut ecosystem dynamics. However, as the hype of these research methods is settling down, scientists are now starting to recognize that culture-dependent methods still have crucially important roles in the interpretation of metagenomic data, progressing understanding of the microbiota, and developing novel therapeutics aimed at improving gut health (Allen-Vercoe, 2013; Rajilic-Stojanovic and de Vos, 2014; Walker *et al.*, 2014).

One of the benefits of the availability of cultured microbes is the ability to obtain completely sequenced genomes from these species. Still, the genome only tells about the genetic potential of the organism, and while predictions of the functionality can be made based on this, other experiments are needed to verify the predictions. Gene expression profiles can be studied by transcriptome analysis, which has so far

mainly been done by microarray analysis. As sequencing techniques have evolved and become more affordable, sequencing of RNA sequence libraries (RNA-seq) has shown to be an effective method for gene expression analysis (Creecy and Conway, 2015; Forde and O'Toole, 2013). While transcriptome analysis gives a view on what the organism is aiming to accomplish, a proteomic approach will describe what was accomplished by revealing which proteins were actually produced. Proteomics techniques have also advanced vastly in the recent years, and they are increasingly being used for microbial research. Sample preparation techniques have evolved from the classical 2DE-based proteomics to include other gel-based methods as well as gel-free approaches based on 'in solution' digestion (Otto *et al.*, 2014). Modern mass spectrometry-based quantitative techniques usually rely on either metabolic or chemical labeling of the sample, but label-free quantification methods are also becoming exceedingly popular. A variety of mass analyzers can be used depending on the throughput requirements of the research. Some examples of these are quadrupole instruments, ion trap instruments and time of flight (TOF) instruments, as well as the orbitrap instrument, which is one of the most commonly used instruments in the current proteomics applications due to its high mass resolution and high mass accuracy (Graham *et al.*, 2011). In addition to studying the proteome of single organisms, the improved proteomics techniques and instruments have stimulated the field of metaproteomics, in which the collective proteome of microbial communities, in this case the gut microbiota, is studied (Kolmeder and de Vos, 2014).

In this thesis, the information derived from the genome of *A. muciniphila* is used in combination with transcriptomic and proteomic techniques to gain insight into the metabolism, membrane structure and immunomodulatory capacities of this prominent member of the gut microbiota.

Research aim and thesis outline

The aim of the research described in this thesis was to unravel the mechanisms of immune stimulation and substrate utilization of the human commensal bacterium *Akkermansia muciniphila*. Transcriptome and proteome analyses in combination with *in vitro* growth experiments and immune assays were used to reach this goal.

Chapter 2 provides an overview of the functionality of human intestinal microbiota based on available metagenome, metatranscriptome, and metaproteome data. It highlights the importance of early life nutrition, dietary habits in adults and health status on the activity of the gut microbiota. As such, it complements the information summarized above on the effect of the microbiota on the host.

In **Chapter 3** the substrate utilization capacities of *A. muciniphila* are studied and new metabolic pathways are described. The focus is on the environmental response by using high throughput transcriptomic and proteomic techniques that indicate differences between bacteria feeding on mucin or the non-mucin sugar glucose. In addition, a genome-based metabolic model is generated and validated with *in vitro* growth experiments in which the carbohydrate degradation abilities, amino acid requirements and final fermentation profiles of the bacteria are examined.

The metabolic capacity of *A. muciniphila* is further investigated in **Chapter 4**, where its ability to degrade and consume human milk glycans is observed. This suggests that these bacteria may benefit from the breast milk environment during colonization in early life, before settling into their niche, the mucus layer.

After establishment in the gut, *A. muciniphila* is in close contact with the host and other members of the microbiota. Bacterial outer membrane proteins play an important role in this cross-talk and thus they are characterized in **Chapter 5**. This has led to the confirmation of the presence of a Gram-negative outer membrane and identification of candidate proteins for immune stimulation.

The role of outer membrane proteins with respect to the immunostimulatory properties of *A. muciniphila* is elucidated in **Chapter 6**. A specific gene cluster potentially involved in the formation of fimbriae is identified as a key component in activating TLR2 and cytokine response as well as increasing transepithelial resistance. These results indicate that the outer membrane protein composition of *A. muciniphila* prompts immunological homeostasis in the gut mucosa by keeping the immune system alert and by improving gut barrier function.

Chapter 7 summarizes the findings of this thesis, discusses the outcome in a broader setting and provides future perspectives for research.

Chapter 2

The function of our microbiota: who is out there and what do they do?

Noora Ottman, Hauke Smidt, Willem M. de Vos, and Clara Belzer
Frontiers in Cellular and Infection Microbiology, Aug 9;2:104 (2012)

Abstract

Current metagenomics developments provide a portal into the functional potential and activity of the intestinal microbiota. The comparative and functional metagenomics approaches have made it possible to get a molecular snap shot of microbial function at a certain time and place. To this end, metagenomics is a DNA-based approach, metatranscriptomics studies the total transcribed RNA, metaproteomics focuses on protein levels and metabolomics describes metabolic profiles. Notably, the metagenomic toolbox is rapidly expanding and has been instrumental in the generation of draft genome sequences of over 1000 human associated microorganisms as well as an astonishing 3.3 million unique microbial genes derived from the intestinal tract of over 100 European adults. Remarkably, it appeared that there are at least 3 clusters of co-occurring microbial species, termed enterotypes, that characterize the intestinal microbiota throughout various continents. The human intestinal microbial metagenome further revealed unique functions carried out in the intestinal environment and provided the basis for newly discovered mechanisms for signaling, vitamin production and glycan, amino-acid and xenobiotic metabolism. Individual's age, diet and health status are all factors that contribute to the composition and activity of the microbiota in several ways. Utilization of high throughput techniques is gradually increasing the knowledge available on the causalities of these phenomena. In this review we will focus on our current understanding of the functionality of the intestinal microbiota.

Introduction

The human intestinal microbiota is known to play a key role in several metabolic, nutritional, physiological and immunological processes and recent years have seen a rapid development in the techniques for studying this previously overlooked organ (O'Hara and Shanahan, 2006). The human microbiota is established after birth and starts out as a dynamic ecosystem, dominated by bifidobacteria, that stabilizes during the first 2-3 years (Koenig *et al.*, 2011; Scholtens *et al.*). During life the microbial composition increases in both diversity and richness (Scholtens *et al.*) (Figure 1) and reaches highest complexity in the human adult, with several hundred species-level phylotypes dominated by the phyla *Bacteroidetes* and *Firmicutes* (Rajilic-Stojanovic *et al.*, 2009). Each human individual reaches a homeostatic climax composition, which likely remains relatively stable during most of a healthy adult's life. Although the individual microbial composition has an 'individual core' that varies at the bacterial phylotype level and depends on the depth of the analysis (Jalanka-Tuovinen *et al.*, 2011; Zoetendal *et al.*, 2008), the overall phylogenetic profile can be categorized into a limited number of well-balanced host-microbial symbiotic states, the so-called enterotypes (Arumugam *et al.*, 2011). At the late stages of life the microbiota composition becomes again less diverse and more dynamic, characterized by a higher *Bacteroides* to *Firmicutes* ratio, increase in *Proteobacteria* and decrease in *Bifidobacterium* (Biagi *et al.*, 2010).

The establishment of the bacterial ecosystem in early life is suggested to play a role in the microbial composition and disease susceptibility throughout life (Scholtens *et al.*, 2012). A different microbiota composition is associated with chronic intestinal disorders and the severity of perturbation during disease and after antibiotic use (Sekirov *et al.*, 2010). Diet is another important factor in microbiota composition development. Early in life there is already an impact of the diet on the microbiome: the microbiota of breastfed and formula-fed infants was found to differ significantly in both composition and diversity. Breastfed babies contain a microbiota that is more heterogeneous than that of formula-fed babies and contain a higher taxonomic diversity (Schwartz *et al.*, 2012). In addition, food habits can influence microbiota composition, and malnutrition results in lower abundance of *Bacteroidetes* that are shown to be specialized in breaking down the carbohydrates in energy rich western diet foods. Diet-related diseases such as allergies and obesity are also characterized by microbiota changes. Obesity is characterized by a typical *Firmicutes* to *Bacteroides* ratio. Energy harvest potential and short chain fatty acids (SCFA) are determined by the microbiota composition and have a direct effect on the host epithelial cell energy availability. A microbiota stimulated with

probiotic microbes can even decrease the incidence of infant diarrhea and atopic eczema due to host immune stimulation (Niers *et al.*, 2009; Sjogren *et al.*, 2009).

Numerous meta-omics approaches have vastly increased the knowledge available on the genome, activity and functionality of the complex ecosystem residing in the human gut. By far the most commonly applied technique is metagenomics, which is based on direct isolation and, in most cases, sequencing of the complete genetic material obtained from an environmental sample, such as the intestine. However, one of the biggest drawbacks of this technique is its inability to display the actual metabolic activity due to the fact that it detects both expressed and non-expressed genes. In addition, it may generate information from dead cells as it is known that more than half of the cells in fecal samples are non-viable or heavily damaged (Ben-Amor *et al.*, 2005). Instead of focusing on microbiota composition the purpose of this review is to combine the available knowledge on microbial genomics with reports on the functional metagenomics i.e. transcriptomics and proteomics approaches. This combination is expected to provide a refined understanding of the role of the microbiota and its capabilities in regulating human health.

Role of the microbiota in early and late life

Early life

During natural birth, a newborn is exposed to the environmental, mainly maternal, microbiota which commences the acquisition of what we assume is a normal microbiota. The mode of delivery strongly affects the composition of the microbiota. In the case of caesarean delivery (C-section), other environmental bacteria form the basis for the microbiota instead of vaginal and fecal bacteria from the mother, reportedly resulting in a substantial reduction of bifidobacteria (Biasucci *et al.*, 2008). In a comparison of the microbiota of babies delivered either vaginally or via C-section, it was shown that the newborns harbored undifferentiated bacterial communities across skin, oral, nasopharyngeal and gut habitats regardless of delivery mode, and that the microbiota of C-section babies was similar to the skin communities of the mothers whereas vaginally delivered infants acquired bacterial communities resembling the vaginal microbiota of their mothers (Dominguez-Bello *et al.*, 2010). Other factors influencing the microbiota are the type of infant feeding, gestational age, infant hospitalization and antibiotic use by the infant. The microbiota of breastfed infants is dominated by bifidobacteria whereas the counts of *Escherichia coli*, *Clostridium difficile*, *Bacteroides fragilis* and lactobacilli are higher in exclusively formula-fed infants (Penders *et al.*, 2006).

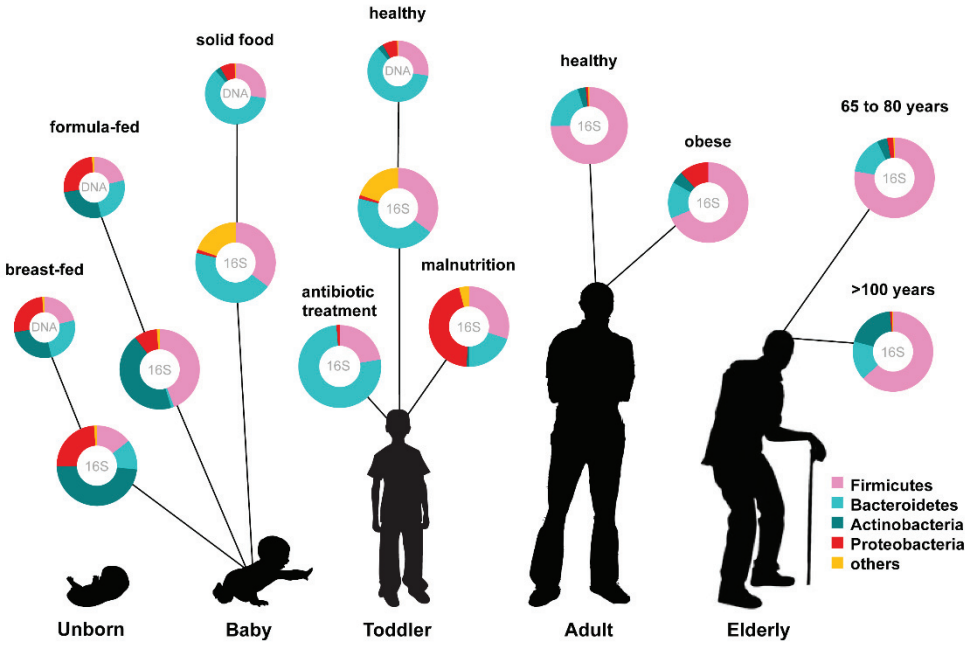


Figure 1. Human microbiota: onset and shaping through life stages and perturbations. The graph provides a global overview of the relative abundance of key phyla of the human microbiota composition in different stages of life. Measured by either 16S RNA or metagenomic approaches (DNA). Data arriving from: Babies breast and formula fed (Schwartz *et al.*, 2012), baby solid food (Koenig *et al.*, 2011), toddler antibiotic treatment (Koenig *et al.*, 2011), toddler healthy or malnourished (Monira *et al.*, 2011), adult, elderly and centenarian healthy (Biagi *et al.*, 2010), adult obese (Zhang *et al.*, 2009).

The composition of the intestinal microbiota plays an important role in immune system development, and it is possible that childhood allergies are related to differences in the microbiota (Sjogren *et al.*, 2009). The intestinal defense of the preterm infant is rather immature and exaggerates inflammatory responses that can be evoked by both commensal and pathogenic bacteria (Nanthakumar *et al.*, 2000). Thus, the first microbes colonizing the intestinal tract hold a pivotal role. Once the core microbiota has developed, it stabilizes and is expected to become less sensitive to modification. The question is at what age does the microbiota become adult-like and recent data with large cohorts of babies in various parts of the world indicate that this is at ages after at least 3 years (Yatsunenko *et al.*, 2012).

The succession of the microbial ecosystem in the intestinal tract of newborns is a complicated process, which is not yet fully understood. The increasing diversity of the microbiota is believed to have an effect on the functional gene content over time. Several studies have provided insight in the infant gut community structure and its perturbations during early development and highlighted the impact of weaning (Favier *et al.*, 2003; Favier *et al.*, 2002). Moreover, a recent 2.5-year case study was reported, where sixty fecal samples were collected from a healthy infant (Koenig *et al.*, 2011) (Figure 1). The results of this study showed a gradual increase in the phylogenetic diversity of the microbiome over time. Life events such as changes in diet, illnesses and antibiotic treatments were associated with large shifts in the abundances of major groups in this single infant. Assignment of gene functions to the metagenomic data from this study revealed an enrichment of carbohydrate-metabolizing genes involved in lactate utilization from the very beginning of life. Interestingly, during an exclusive breast-milk diet genes facilitating the breakdown of plant-derived polysaccharides were already present, suggesting that the microbiota is metabolically prepared to receive simple plant-derived foods. This is consistent with other observations that showed high similarity in the proportions of Clusters of Orthologous Groups (COG) encoding proteins specialized for the transport and metabolism of plant polysaccharides or COGs encoding proteins transporting and metabolizing human milk oligosaccharides (HMOs) between infant and maternal microbiota samples (Vaishampayan *et al.*, 2010) (Table 1). Baby gut microbiomes are also enriched in functions involved in using glycans represented in breast milk and the intestinal mucosa, even more so in the microbiomes of Amerindian and Malawian babies compared with US babies, possibly reflecting differences in the glycan content of breast milk (Yatsunenko *et al.*, 2012).

Table 1. Percentages of COG categories expressed in the gut microbiota

Population	Sample size (n)	COG categories (percentage of all genes)							Reference
		C	E	G	L	M	J	O	
mother 1m	1	5	6	11	10	8,5	3	3	Vaishampayan et al., 2010
mother 11m	1	6	10	12	5.5	6	5	2.5	"
infant 1m	1	7	7,5	11,5	6	6	4,5	4	"
infant 11m	1	4	11	12	7.5	6	5	3	"
female twin pair[1]	2	14	n/a	16	n/a	n/a	19	12	Verberkmoes et al., 2009
healthy volunteers	10	6.0 - 13.0	3.5 - 7.0	9.5 - 22.0	3.5 - 11.0	1.5 - 8.0	9.0 - 15.0	2.5 - 14.0	Gosalbes et al., 2011
female cotwin (TS28)[2]	1	8	7	8	6	6	9	6	Turnbaugh et al., 2010
female cotwin (TS29)[2]	1	8	6	9	5	5	10	7	"
female cotwin (TS28)[3]	1	5	6	8	9	6	7	5	"
female cotwin (TS29)[3]	1	5	7	9	9	6	6	4	"

COG descriptions: C Energy production and conversion; E Amino acid transport and metabolism; G Carbohydrate transport and metabolism; L DNA replication, recombination and repair; M Cell envelope biogenesis, outer membrane; J Translation, ribosomal structure and biogenesis; O Posttranslational modification, protein turnover, chaperones

[1] out of the core proteome

[2] out of genes with high relative expression

[3] out of genes with low relative expression

Recently it has been shown that the use of specific human milk derived glycans such as HMO utilization is not exclusive to certain well-known infant colonizers, such as *Bifidobacterium* species, since members of the genus *Bacteroides* can also use milk glycans as a sole carbon and energy source (Marcobal *et al.*, 2010). Moreover, it has been shown that *Bacteroides thetaiotaomicron* responds to common structural motifs found in oligosaccharides from mother's milk and intestinal mucin glycans, suggesting that HMOs may mimic mucus glycans to attract mucin-adapted resident mutualists to an infant microbiota (Marcobal *et al.*, 2011). However, specific HMO components select for HMO-adapted species such as *Bifidobacterium longum* subsp. *infantis*, and provide a selective advantage to this species *in vivo* when biassociated with *B. thetaiotaomicron* in the gnotobiotic mouse gut. The complex oligosaccharide mixture within HMOs thus attracts both mutualistic mucus-adapted species and HMO-adapted bifidobacteria to the infant intestine that likely facilitate both milk and future solid food digestion.

Little is known of the effect of diet on the composition and in particular the activity of the developing gut microbiota. Comparison of host epithelial cell gene expression and microbiota profile between breast- and formula-fed infants demonstrated that differences in the diet of infants can have an influence on the host gene expression via the gut microbiota (Schwartz et al., 2012). Virulence characteristics of the microbiota were the only functional properties that were found to differ among these two groups. Further analysis of the host transcriptome revealed a subset of eleven immunity and mucosal defense-related genes exhibiting evidence of a multivariate relationship with microbiome virulence characteristics. This provides additional proof for the capability of human milk to promote the mutualistic interactions between the mucosal immune system and the microbiome in maintaining intestinal homeostasis. Gene content analysis of the gut microbiome of 110 individuals including both adults and babies from Venezuela, Malawi and the United States revealed age-related changes in the metabolism of vitamins B12 (cobalamin) and folate (Yatsunenko *et al.*, 2012). Genes involved in *de novo* biosynthesis of folate decreased with age whereas genes encoding most enzymes associated with cobalamin biosynthesis increased, correlating with previous data of blood levels of these vitamins in different age groups (Monsen *et al.*, 2003).

The key players in the neonate gut are the bifidobacteria, which dominate the microbial community of human milk-fed infants. A number of studies using metagenomic approaches have also demonstrated the importance of this genus in the developing gut (Turroni *et al.*, 2012; Yatsunenko *et al.*, 2012), while at the same time other studies have reported low abundance or even absence of bifidobacteria (Koenig *et al.*, 2011; Palmer *et al.*, 2007), most likely due to technical biases related

to DNA extraction protocols or the selected PCR primers. Genome analysis of *Bifidobacterium longum* subsp. *infantis* revealed a nutrient-utilization strategy targeting milk-derived molecules which are not of nutritional value to the infant (Sela *et al.*, 2008). The proteomic profile of the organism grown on HMOs confirmed the activity of these genes (Sela *et al.*, 2008). This suggests *B. longum* subsp. *infantis* coevolved with its infant host and under the presence of human milk compounds.

Furthermore, the type of milk, either mother's milk or formula, determines the colonization with different types of bifidobacteria. Breastfed infants contain a high abundance of *Bifidobacterium breve*. In contrast, fecal samples from standard formula-fed infants lacked detectable amounts of this *B. breve* but contained *B. longum*. Remarkably, infants that received breast milk and later a prebiotic formula consisting of a standard formula milk containing a mixture of specific galacto- and fructo-oligosaccharides, continued to harbor a *B. breve*-dominant fecal population (Boesten *et al.*, 2011).

Transcriptional analysis of the response of *B. longum* to human milk and formula milk indicated upregulation of genes involved in carbohydrate metabolism in breast milk, endorsing the concept that the bifidogenic effect of breast milk is primarily based on its oligosaccharides (Gonzalez *et al.*, 2008). Moreover, the same study found upregulation of putative genes for cell surface type 2 glycoprotein-binding fimbriae associated with attachment and colonization in the intestine in both breast milk and formula milk when compared to semisynthetic medium with glucose. Transcriptome analysis of *B. breve* in a mouse model also showed differential expression of genes encoding for the production of type IVb tight adherence pili, which are essential for efficient *in vivo* murine gut colonization (O'Connell Motherway *et al.*, 2011).

Another study comparing the bifidobacterial transcriptome of breast-fed infants and prebiotic-containing formula-fed infants showed that in the beginning of the intervention breast-fed infants had higher counts of bifidobacteria compared to the formula-fed infants (Klaassens *et al.*, 2009). However, during the intervention the bacterial numbers and species diversity of *Bifidobacterium* increased significantly in the formula-fed infants, possibly on account of the galacto- and fructo-oligosaccharides in the formula. These prebiotics have also previously been shown to shift the bifidobacterial quantities towards those of breast-fed infants (Knol *et al.*, 2005). The metatranscriptome analysis in babies revealed that the most prominent functions of the transcripts were related to carbohydrate metabolism, with higher expression of genes encoding these functions in breast-fed infants

compared to formula-fed infants (Klaassens *et al.*, 2009). This included significant expression of genes involved in HMO degradation. Moreover, the expression of genes involved in folate production was observed in all babies indicating that intestinal bifidobacteria produced this important vitamin involved in neural development. In the same study, a gene for bifidobacterial transaldolase, which is a key enzyme of the non-oxidative phase of the pentose phosphate pathway, was expressed in samples from all infants. Bifidobacterial transaldolase was also found in the only metaproteome study thus far to look at the infant gut microbiota (Klaassens *et al.*, 2007). Production of the protein spot on a 2D-gel corresponding to this protein was increased over time suggesting an increase in the numbers and activity of bifidobacteria in the infant's gut. Understanding the factors relating to the existence and host interactions of bifidobacteria and linking the functionality of this early intestinal colonizer to specific diets and groups of healthy or diseased individuals may eventually lead to the possibility of guiding the development of the microbiota. This can be achieved with pro- and prebiotic supplemented infant formulas that are aimed at increasing the bacterial diversity and a more optimal bifidobacterial community composition.

Late life

In addition to the beginning of life, the microbiota also undergoes significant changes towards the other extremity of life, old age. These alterations, however, are not clear-cut partially due to the various physiological changes that the elderly go through. These include factors such as modifications in lifestyle, nutritional behavior, increase in infection rates and inflammatory diseases, and therefore the need for more medication. All of these issues will certainly also affect the composition and activity of the microbiota, but the course and mechanisms behind these changes are not yet completely understood.

The process of ageing has been demonstrated to have a negative effect on the diversity of the microbiota, but different studies have reported conflicting results on the age-related changes with regard to the two major phylogenetic groups. Assessment of the gut microbiota of the elderly with quantitative PCR revealed high levels of *Escherichia coli* and *Bacteroidetes* as well as a significant difference in the *Firmicutes* to *Bacteroidetes* ratio for adults (10.9) and elderly individuals (0.6) (Mariat *et al.*, 2009). In this study the total bacterial counts for adults and seniors were comparable whereas another study, employing cytosine (%G + C) profiling and 16S rRNA gene sequencing, described a significant reduction in overall numbers of microbes in elderly subjects compared to young adults (Makivuokko *et al.*, 2010). They also observed lower numbers of *Firmicutes* and an

increase in *Bacteroidetes*, with lowered amounts of known butyrate producers belonging to *Clostridium* cluster XIVa.

Another study, which included young (20 - 40 years old), elderly (60 – 80 years old) and an additional group of centenarian citizens (~100 years old), clearly demonstrated that the process of ageing coincides with decreasing microbiota diversity (Biagi *et al.*, 2010) (Figure 1). By using the Human Intestinal Tract Chip (HITChip) and qPCR, they observed that the composition of microbiota was quite similar between the young and the elderly groups represented by dominant portions of *Firmicutes* and *Bacteroidetes* (95% of total bacteria). The centenarian group also showed a dominant portion of *Firmicutes* and *Bacteroidetes* (93% of total bacteria). The *Firmicutes/Bacteroidetes* ratios obtained for the centenarians, elderly and young adults were 3.6, 5.1 and 3.9, respectively. However, there was a significant decrease in the *Firmicutes* subgroup *Clostridium* cluster XIV and an increase in *Bacilli* in the centenarian group. Furthermore, there was a significant increase in several facultative anaerobes, members of the *Proteobacteria* phylum, many of which constitute opportunistic pathogens. This rearrangement of the microbiota does not seem to be in favor of the aging subjects that showed an increased level of circulating inflammatory cytokines. These were inversely associated with bacteria belonging to *Clostridium* cluster XIV and *Clostridium* cluster IV that include the main butyrate-producers in the gut. Butyrate has been associated with a range of health effects from anti-inflammatory properties to enhancement of intestinal barrier function (Macfarlane and Macfarlane, 2011).

Recently, pyrosequencing of tagged PCR-amplified 16S rRNA genes was applied to characterize the fecal microbiota of 161 seniors aged 65 years and older in the ELDERMET consortium (Claesson *et al.*, 2011). In this extensive study the elderly microbiota was observed to be dominated by the phylum *Bacteroidetes* (57%) compared with *Firmicutes* (40%). However, the proportions of the major phyla showed extraordinary variation between individuals, with the proportion of *Bacteroidetes* ranging from 3% to 92% and *Firmicutes* from 7% to 94%. In addition to the general composition, also the core microbiota of the elderly differed substantially with that of young adults, characterized by a shift to a more *Clostridium* cluster IV-dominated community in the elderly. The microbiota of the elderly showed temporal stability for the majority of subjects as revealed by analysis of 3-month follow-up samples.

These studies indicate that there undoubtedly are fluctuations in the elderly microbiota, but both the threshold for an “aged” microbiota and the trends for these changes seem to be highly variable. Some of these differences may be explained by

country-specific dietary habits, as the most recent studies used separate cohorts from two different European countries, Italy (Biagi *et al.*, 2010) and Ireland (Claesson *et al.*, 2011). The living environments of elderly people are highly dependent on their health status, with healthier seniors living independently and subjects with medical issues often living in nursing homes. These factors can also influence the aging gut microbiota. Follow-up studies assessing the function of the elderly gut microbiota by functional metagenomic techniques already applied for the infant and adult microbiota will shed more light on these issues and reveal prospects for possible dietary interventions aimed at improving the health of the elderly.

Microbiota activity in response to diet

Host dietary habits appear to affect gut microbiota composition, but the actual association between different diets and the microbial community composition as well as the underlying causes for this are still unclear. Although there was no clear environmental or genetic explanation found for the initial clustering of the enterotypes (Arumugam *et al.*, 2011), these were found to be strongly associated with long-term diets, with protein and animal fat correlating with the enterotype characterized by high levels of *Bacteroidetes*, and carbohydrates with the *Prevotella* enterotype (Wu *et al.*, 2011). Differences in microbiota composition as a result of diverging dietary habits was also shown in a comparison of the microbiota of European children, who consumed a diet high in animal protein, sugar, starch and fat and low in fiber, and children from Burkina Faso, where the predominantly vegetarian diet consists mainly of carbohydrates, fiber and non-animal protein (De Filippo *et al.*, 2010). The European microbiome was enriched with *Firmicutes* and *Proteobacteria*, whereas *Actinobacteria* and *Bacteroidetes* were more represented in the African children. Interestingly, *Xylanibacter* and *Prevotella* were only present in the children from Burkina Faso, leading the authors to hypothesize that members of these genera could improve the ability to extract calories from indigestible polysaccharides commonly consumed in rural Africa indicating a coevolution of the microbial community with the polysaccharide-rich diet. Malnourished children from poor socio-economic status families in Bangladesh were found to have lower diversity of gut microbiota compared to healthy children from moderate to high income families in the same region, characterized by lower relative abundance of *Bacteroidetes* and a dominance of *Proteobacteria* (Monira *et al.*, 2011). The authors suggest that the low presence of *Bacteroidetes*, which are known to digest complex dietary material and thus improve energy extraction from various foods, and the higher presence of potentially pathogenic *Proteobacteria* might contribute to explaining the poor health of the malnourished children.

In a metagenome study, short-term dietary intervention (high-fat/low-fiber or low-fat/high-fiber diets) lead to rapid changes in the microbiome composition but was not sufficient to shift individuals between the two enterotypes described in the same study (Wu *et al.*, 2011). Few functional gene categories, including bacterial secretion system, protein export and lipoic acid metabolism, differentiated between the two test diets suggesting a shift in selected bacterial functions in response to the dietary changes. Microbiome analysis of subjects on a diet rich in protein, typically consumed in the US, showed enrichment of multiple Enzyme Commission (EC) groups when compared with Malawian and Amerindian subjects consuming a diet high in carbohydrates (Yatsunenko *et al.*, 2012). These included degradation of glutamine and other amino acids, catabolism of simple sugars, vitamin biosynthesis and bile salt metabolism. Degradation of glutamine has earlier been found to be overrepresented in carnivorous mammalian microbiomes, while glutamate synthase, which was enriched in Malawian/Amerindian microbiomes, was present in higher proportions in herbivorous mammalian microbiomes (Muegge *et al.*, 2011).

Several metatranscriptome and metaproteome studies describing the human intestinal microbiota have confirmed the importance of bacterial functions related to carbohydrate metabolism in the colon. Enrichment of these genes has earlier been shown in metagenomic studies of the human gut (Gill *et al.*, 2006; Kurokawa *et al.*, 2007; Turnbaugh *et al.*, 2009a). Metatranscriptome analysis of fecal samples from two healthy volunteers found that most expressed genes (26 % of all sequenced and annotated transcripts) were involved in the metabolism of carbohydrate (Booijink *et al.*, 2010a). Recently the majority of bifidobacterial transcripts within the fecal community of adults were also reported to be involved in metabolism of carbohydrates of plant origin (Klaassens *et al.*, 2011).

Similar results were seen in a transcriptional analysis of fecal samples from a monozygotic, obese twin pair (Turnbaugh *et al.*, 2010) (Table 1), and metatranscriptomics analysis of fecal samples from ten healthy volunteers (Gosalbes *et al.*, 2011) (Table 1). Metatranscriptomic data from the less studied small intestinal microbiota showed enrichment in sugar phosphotransferase (PTS) and other carbohydrate transport systems, as well as energy- and central metabolic, and amino acid conversion pathways as compared with the metagenome (Zoetendal *et al.*, 2012b). This suggests rapid uptake and fermentation of available simple sugars by the small intestinal microbiota, compared to the degradation of more complex carbohydrates by the bacteria in the colon. The importance of carbohydrate metabolism is also evident from the enormous amount of carbohydrate-active enzymes (CAZymes) present in the gut microbiome. By

applying a multi-step functional screening procedure of a metagenomic library from the feces of a volunteer following a fiber-rich diet, 73 CAZymes from 35 different families were recently discovered (Tasse *et al.*, 2010).

Shotgun metaproteomics approach used to identify microbial proteins in fecal samples from a female twin pair identified several COG categories more highly represented in the microbial metaproteome compared to the average metagenome (Verberkmoes *et al.*, 2009) (Table 1). A high proportion of the proteins that were equally abundant in both samples were from common gut bacteria, such as *Bacteroides*, *Bifidobacterium* and *Clostridium*. These included proteins involved in translation, carbohydrate metabolism and energy production. In another study, two human fecal samples were analyzed and the functions of the identified proteins were predicted (Rooijers *et al.*, 2011). The most abundantly present COGs were involved in translation, energy production and conversion as well as carbohydrate transport and metabolism, which supports the findings of studies linking the microbiota with carbohydrate metabolism (Kovatcheva-Datchary *et al.*, 2009). The study also pointed out the abundance of *Akkermansia muciniphila*, the only intestinal member of the *Verrucomicrobia*, within the microbiota and showed that most of the proteins produced by these bacteria are involved in carbohydrate transport and metabolism as well as amino acid transport and metabolism. This is in line with observation that *A. muciniphila* can use mucin as the sole carbon and nitrogen source (Derrien *et al.*, 2008). The fecal samples were also subject to metagenome sequencing and the phylogenetic diversity was determined with two approaches, 16S rRNA sequence analysis of the metagenomic data sets and an abundance analysis of the metagenomic sequences using a synthetic metagenome as reference set. The results showed that *Bacteroidetes*, *Firmicutes*, *Actinobacteria*, *Verrucomicrobia* and *Proteobacteria* were the dominant groups in the microbiota of the study subjects.

These results were further confirmed by analyzing the gut metaproteome of three healthy subjects over a period of six to twelve months (Kolmeder *et al.*, 2012). In this study, proteins involved in carbohydrate transport and metabolism accounted for over 10 % of the detected proteins, forming a part of the core metaproteome found in all the test subjects. The glycolysis pathway, in particular, was noticeable with several related enzymes identified. After assigning the spectral hits for each COG functional category per phylum, it was apparent that *Firmicutes* and *Actinobacteria* were responsible for the active carbohydrate metabolism, while *Bacteroidetes* showed more mixed functions. Both *Firmicutes* and *Bacteroidetes* were found to have an active carbohydrate metabolism on a transcriptional level in an earlier report (Gosalbes *et al.*, 2011). Furthermore, Kolmeder and co-authors

(2012) observed that the majority of the identified actinobacterial peptides were predicted to be involved in sugar metabolism. The importance of carbohydrate metabolism has been shown also previously for the core genome of bifidobacteria (Bottacini *et al.*, 2010). Temporal analysis showed that the metaproteome is stable over time, as is the microbial composition of the gut, suggesting that homeostasis in function and composition of the intestinal microbiota are tightly linked (Kolmeder *et al.*, 2012).

Recently, a metatranscriptomics approach with RNA-seq has been applied to investigate the effect of a fermented milk product (FMP) containing several probiotics on the gut microbiome of gnotobiotic mice colonized with a model human gut microbiota and monozygotic twins (McNulty *et al.*, 2011). There were no or minimal changes observed in the bacterial species composition in mice and humans after consumption of FMP. Still, transcriptional analysis revealed significant changes in numerous metabolic pathways, especially in carbohydrate metabolism, in both mice and human subjects. The question, however, is whether this reflects a functional difference in the colon or is a result of technical or biological effects such as variations in the transit time of the fecal material used for this analysis.

Metagenomic approaches combined with studies using gnotobiotic animals colonized with only a few known microorganisms or even the entire human fecal microbiota, provide a powerful tool for examining the relationship between the host and the functionality of the microbial community under controlled conditions. A study of humanized gnotobiotic mice transplanted with either fresh or frozen adult human fecal microbial communities into germ-free C57BL/6J mice revealed a stable and heritable colonization which enabled a diet intervention, where the mice were switched from a low-fat, plant polysaccharide to a high-fat, high-sugar diet (Turnbaugh *et al.*, 2009b). This diet change induced a structural shift in the microbiota within one day and presented an enrichment for various KEGG pathways involved in nutrient processing compared to the control diet. Metatranscriptome analysis of rRNA-depleted RNA isolated from the ceca of the humanized mice demonstrated a clear difference in the gene expression of the mice on the Western diet compared to the control group, with upregulation of clusters containing *Clostridium innocuum* strain SB23 genes encoding Western diet-associated transcripts (pyruvate formate-lyase, phosphotransferase system, phosphoglycerate kinase) and *Firmicutes* gene clusters encoding ABC-type sugar transport systems.

A shift in the microbial community was also seen after switching both wild-type and RELMß-deficient mice to a high-fat diet, indicating that the diet itself was responsible for the detected changes independent of obesity (Hildebrandt *et al.*, 2009). RELMß is a colonic goblet cell-specific gene, whose expression is dependent on the presence of the gut microbiome. After the dietary switch the amounts of *Proteobacteria*, *Firmicutes* and *Actinobacteria* increased whereas *Bacteroidetes* decreased, as measured from fecal samples. Analysis of gene functions revealed a decrease in the number of metabolic genes under the high-fat condition, possibly as a result of nutrient deficiency. However, as also noted by Turnbaugh and colleagues (2009b), a group of genes for ABC-transporters increased in abundance, indicating adaptation to the high-fat diet by enhancing nutrient intake in an environment with limited substrate availability.

Mice colonized with 10 sequenced human gut bacteria, and fed with a series of refined diets showed that casein concentration was highly correlated with the yield of total DNA per fecal pellet in all 17 test diets (Faith *et al.*, 2011). The abundance of all of the ten species was significantly associated with casein, with seven of them positively correlated with casein concentration and three negatively correlated. None of the diets caused significant changes in the gene expression of the bacterial species, analyzed by RNA-seq, but high expression of genes predicted to be involved in pathways using amino acids as substrates for nitrogen, as energy and/or carbon sources were found for the species positively correlated with casein.

In conclusion, the studies to date endorse the concept that the intestinal microbiota thrives on using polysaccharides and peptides, which are indigestible to human (Guarner and Malagelada, 2003). The metagenomic data is confirmed on a functional level by the metatranscriptomics and metaproteomics data. The composition of the microbiota in the colon is dominated by *Firmicutes* that appear to be active in carbohydrate metabolism whereas *Bacteroidetes* show activity in a number of functions like energy production and conversion as well as amino acid transport and metabolism, in addition to carbohydrate metabolism. The complex polysaccharides are degraded by a specialized microbial community and the released oligosaccharides can in turn be used by other commensal bacteria. In this manner, diet is has a crucial influence on the intestinal microbial activity.

Microbial imbalances and disease

Inflammatory bowel diseases

The gut microbiota has been connected to several diseases, with obesity and inflammatory bowel diseases (IBD) representing the most studied disorders to date. Most research about potential differences of microbiota related to different disease states has so far focused on describing the composition and diversity of the microbiome in patients compared to healthy subjects, and consequently revealing interesting associations between them. In order to get a better understanding of the underlying mechanisms of the relationship between the microbial communities and specific disorders, functional microbiomic approaches need to be employed.

Despite exhaustive research efforts, the etiology and pathogenesis of IBD, including Crohn's disease (CD) and ulcerative colitis (UC) have stayed unclear. The causes of these intestinal diseases are most likely linked with both human gene- and microbiome-associated factors (Pflughoeft and Versalovic, 2012). Crohn's disease and ulcerative colitis patients seem to harbor separate microbial communities both from each other and healthy subjects, and also have lower bacterial diversity compared to healthy people (Dicksved *et al.*, 2008; Manichanh *et al.*, 2006; Qin *et al.*, 2010). Several bacterial groups have been implied to be either increased or decreased in association with IBD. However, it is not clear whether this dysbiosis is the reason for the inflammation in IBD, or simply something caused by the disturbed environment in the GI tract.

Metagenomic studies and microarray analyses have demonstrated a reduction of *Firmicutes*, such as *Faecalibacterium prausnitzii*, in Crohn's disease (Kang *et al.*, 2010; Manichanh *et al.*, 2006). A 16S rRNA gene pyrosequencing study of twin pairs who were concordant or discordant for Crohn's disease or ulcerative colitis showed a clear division in the microbial composition between CD and healthy individuals but not between UC and healthy individuals (Willing *et al.*, 2010). There were more *Firmicutes* detected for colonic involvement CD and less for ileum localized CD (ICD) compared to healthy subjects. In addition to *F. prausnitzii*, also other core members of the microbiota, such as *Roseburia*, were less abundant in ICD. Interestingly, a separate study analyzing the same samples showed clear shifts in metabolic profiles corresponding to the same bacterial groups (Jansson *et al.*, 2009). Pathways with differentiating metabolites included those involved in the metabolism and or synthesis of amino acids, fatty acids, bile acids and arachidonic acid.

A recent analysis of the fecal microbiota of UC patients in relapse and remission further confirmed the reduction of bacterial diversity in these patients and showed that this mainly affects members of the *Clostridium* cluster IV within the phylum *Firmicutes* (Rajilic-Stojanovic *et al.*, 2013). The authors also speculated on the role of SCFA in UC as they reported reduced numbers of butyrate-producing bacteria, along with other studies (Frank *et al.*, 2007), and a disturbed abundance of typical propionate producers. A depletion of one propionate producer, *A. muciniphila*, was observed in the fecal samples while another one, *Megamonas* sp. was increased. *A. muciniphila* was previously found to be decreased in biopsies of patients with UC with an associated increase in *Ruminococcus* sp. (Png *et al.*, 2010). The role of butyrate and propionate, both of which have anti-inflammatory properties (Tedelind *et al.*, 2007), in UC is still under debate (Chapman *et al.*, 1994; Roediger *et al.*, 1997). These, and forthcoming studies will eventually help in screening and diagnosing IBD patients.

Obesity & metabolic syndrome

Obesity and obesity-associated metabolic disorders, such as metabolic syndrome and type 2 diabetes have been suggested to be associated with the composition and function of the intestinal microbiota. Initial research showed an increase in the relative abundance of *Firmicutes* and decrease in *Bacteroidetes* in both obese mice (Ley *et al.*, 2005) and humans (Ley *et al.*, 2006), but later studies have failed to endorse these findings and showed inconsistent results with respect to the changes in the microbiota of obese people (Nadal *et al.*, 2009; Santacruz *et al.*, 2010; Santacruz *et al.*, 2009; Schwiertz *et al.*, 2010; Zhang *et al.*, 2009) (Figure 1). In addition, the transfer of the gut microbiota of obese (*ob/ob*) mice to germ-free wild-type mice causes an increase in fat mass in the recipients, indicating that the obese microbiota has an increased capacity to harvest energy from the diet (Turnbaugh *et al.*, 2006).

Departing from these findings, scientists are now trying to unravel the mechanisms behind the observations. One study found that loss of Toll-like receptor (TLR) 5, which is a transmembrane protein recognizing bacterial flagellin, in a mouse model results in a phenotype resembling human metabolic syndrome (Vijay-Kumar *et al.*, 2010). The authors speculated that the loss of this receptor alters the microbiota inducing low-grade inflammatory signaling, which eventually leads to hyperphagia and metabolic syndrome. In another study, TLR2-deficient mice, which are protected from diet-induced insulin resistance under germ-free settings, developed a condition reminiscent of metabolic syndrome after colonization (Caricilli *et al.*, 2011). The microbiota of the mice showed notable increase in *Firmicutes* and slight

increase in *Bacteroidetes* compared to controls. The authors suggested that the mechanisms by which the TLR2-deficient mice became insulin resistant and, later, obese could be related to increased capacity for energy harvesting from the diet or alternatively to increased level of lipopolysaccharides (LPS) caused by increased gut permeability and LPS absorption. Recently, it was shown that antibiotic treatment with vancomycin for diet-induced obese mice significantly reduced the proportions of *Firmicutes* and *Bacteroidetes*, and increased *Proteobacteria* (Murphy *et al.*, 2012). These changes were associated with improvement in the metabolic abnormalities associated with obesity, by reducing body weight gain and improving inflammatory and metabolic health of the host. Based on these studies, it seems plausible that the ability of the gut microbiota to regulate inflammatory responses play an important role in the complex mechanisms behind obesity and metabolic syndrome. Still, more long-term studies in animal models and humans are required to acquire a clearer picture of the relationship between the intestinal microbiota and different diseases.

Concluding remarks

The complexity of the microbiota–host interactions has been the prime obstacle in defining microbial functionality at a postgenomic level. The recent technical advances in analyzing genomes, transcriptomes and proteomes of complex bacterial consortia and intra- and interspecies metabolic networks help to tackle this problem and will enable systems-level analyses of the crosstalk between the microbiota and the host.

There are multiple reports providing circumstantial evidence to support the concept that microbiota composition and activity influence host metabolism and disease development. These examples include the differences in microbiota composition and microbiota expressed proteins of breast feeding as compared to formula fed babies (Schwartz *et al.*, 2012), differences between microbiota composition and activity between healthy and malnourished infants (Monira *et al.*, 2011), differences in the microbiota composition of elderly and centenarians as compared to youngsters (Biagi *et al.*, 2010) and differences in microbiota composition and activity between humans that are either lean or obese (Ley *et al.*, 2005; Zhang *et al.*, 2009) and healthy or suffering of IBD (Willing *et al.*, 2010). The data suggest that the activity and composition of the microbiota is affected by food intake and genetic background of the host. Most findings are supported by animal studies but there is also data on human subjects. The field of functional microbiomics is still rapidly advancing with continuously emerging new techniques and results. Nevertheless, a lot of times the high throughput techniques fail to correlate bacterial species and genome content

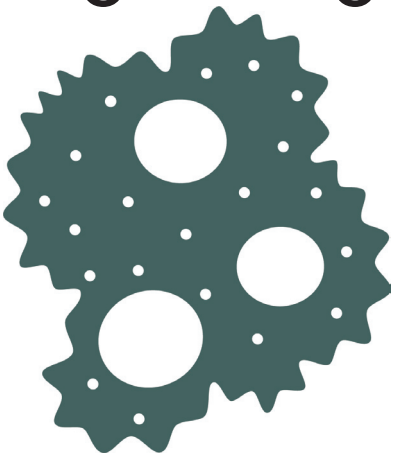
to function due to the lack of characterized isolates and genes. It is important to identify the regulating parameters of the functioning intestinal ecosystem to gain insight into the influence of the microbiota on human development, aging and disease.

Acknowledgements

We are grateful to our colleagues and collaborators for stimulating discussions. This work was partly funded by grant EC FP7 Marie Curie Cross-talk project (PITN-GA-2008-215553) and 250172-MicrobesInside of the European Research Council.

Chapter 3

Metabolic capacities of *Akkermansia muciniphila* reflect a mucin-degrading lifestyle



Noora Ottman, Mark Davids, Sjef Boeren, Peter J. Schaap, Vitor A. P. Martins dos Santos, Hauke Smidt, Clara Belzer, and Willem M. de Vos

Abstract

The composition and activity of the microbiota residing in the human gastrointestinal tract is primarily shaped by the nutrients derived either from food or the host. Bacteria colonizing the mucus layer have evolved to use mucin, the main component of mucus, as a source of energy. One of the members of the mucosa-associated microbiota is *Akkermansia muciniphila*, which is capable of producing an extensive repertoire of mucin-degrading enzymes according to genome analysis and experimental data. To further study the substrate utilization abilities of *A. muciniphila*, we constructed a genome-based metabolic model to test sugar degrading capacities. The findings of the model were validated by *in vitro* growth experiments, which showed *A. muciniphila* to be able to grow on the mucin-derived monosaccharides N-acetylglucosamine, N-acetylgalactosamine, fucose, and galactose. The uptake of these sugars, and also the non-mucin sugar glucose, was enhanced in the presence of mucin, indicating additional mucin-derived components are needed for optimal growth. Transcriptome analysis comparing the gene expression of *A. muciniphila* grown on mucin or glucose confirmed the activity of the genes involved in mucin degradation, and revealed most of them to be upregulated in the presence of mucin. *A. muciniphila* grown on glucose showed a considerable stress response and upregulation of specific genes, such as Amuc_1094, which was identified as a glucokinase-encoding gene induced by glucose. The transcriptional response was confirmed by proteome analysis, reinforcing the adaptation of *A. muciniphila* to the mucosal environment. These new findings provide molecular insights into the lifestyle of *A. muciniphila*, and further confirm its role as a mucin specialist in the gut.

Introduction

Microbial community composition and activity in the human gastrointestinal (GI) tract are largely dependent on the nutrition sources accessible to the gut microbiota (Lacroix *et al.*, 2015; Nicholson *et al.*, 2012; Zoetendal *et al.*, 2012a). Availability of diet-derived nutrients changes according to the diet of the host, whereas host-derived nutrients, such as mucin, are constantly available. Therefore, the capacity to use host-derived glycans can confer a competitive advantage to the bacteria residing in the gut (Koropatkin *et al.*, 2012; Martens *et al.*, 2008).

The intestinal mucus layer, which is one of the bacterial niches in the gut, serves several purposes: it acts as a lubricant for food passing over membranes, as a selective barrier to allow the passage of nutrients to the epithelial cells, and as a protective system against mechanical damage or harmful substances such as pathogens and toxins (Cone, 2009). Regarding the commensal microbiota, the mucus layer offers the bacteria substrates for growth, adhesion and protection (Johansson *et al.*, 2013). Oppositely, microbiota composition has been shown to influence mucus barrier properties, with bacterial species commonly related to inflammatory conditions increasing the permeability of the mucus layer and thus decreasing the barrier function (Jakobsson *et al.*, 2015).

The mucus layer covers the entire GI tract and can be divided into inner and outer layer in the stomach and the colon, whereas the small intestine only has one layer. It is believed that in the colon commensal bacteria only exist on the outer layer, whereas the inner layer is meant to keep the bacteria at a distance from the epithelial cells and to enforce immune tolerance in the gut by delivering antimicrobial proteins and IgA (Johansson *et al.*, 2011; Johansson *et al.*, 2008). Mucins, the main components of the mucus layer, are large glycoproteins, secreted by goblet cells in the epithelium and consisting of a protein backbone decorated with a variety of carbohydrate chains. The main sugar monomers present in mucin are fucose, galactose, N-acetylgalactosamine (GalNAc), N-acetylglucosamine (GlcNAc) that in some cases are modified with sialic acid (N-acetylneuraminic acid) and sulfate. These glycans are attached to the protein backbone via O-glycosidic linkages to serine and threonine residues (Johansson *et al.*, 2011).

One of the key players in the mucus-associated microbiota is *Akkermansia muciniphila*, which colonizes the GI tract of a considerable part of the human population (Derrien *et al.*, 2008). *A. muciniphila* is the only isolated intestinal member of the *Verrucomicrobia* phylum. It has been associated with a healthy intestine and numbers of *A. muciniphila* in the gut are inversely correlated with inflammatory bowel diseases (Png *et al.*, 2010; Rajilic-Stojanovic *et al.*, 2013),

appendicitis (Swidsinski *et al.*, 2011), obesity (Karlsson *et al.*, 2012) and diabetes (Zhang *et al.*, 2013). *A. muciniphila* seems to be highly adapted to its living environment as it has a great capacity of breaking down mucin. Analysis of its 2.7-Mb genome predicted 61 proteins (2.8 % of all proteins) involved in the degradation of mucin (Derrien *et al.*, 2010; van Passel *et al.*, 2011). These enzymes include various proteases, sulfatases and glycosyl hydrolases, including sialidases.

A. muciniphila has previously been shown to produce acetate and propionate during its growth on mucin. *A. muciniphila* is also able to grow on glucose and on the mucin-derived amino sugars GlcNAc and GalNAc, albeit much less efficiently than on mucin, and only in the presence of large amounts of peptone, tryptone, casitone and yeast extract (Derrien *et al.*, 2004). The specific substrate requirements, enzyme activities and metabolic pathways behind these observations, however, are not well understood. Moreover, the physiological growth parameters during growth on non-mucus substrates have been poorly described. To assess the metabolic potential of *A. muciniphila*, we constructed a genome-scale, constraint-based model of its metabolism and transport. We validated it experimentally with *in vitro* growth experiments measuring short-chain fatty acid (SCFA) production, and the global transcriptional and proteomic response as a read-out, based on RNA-seq and advanced mass spectrometry, respectively. In addition, we compared the metabolic responses of *A. muciniphila* grown on mucin to growth on mucin-derived monosugars (GlcNAc, GalNAc, fucose, galactose, and the decorating sialic acid), and the non-mucin sugar glucose. Moreover, the results were used to identify mucin-specific pathways and provided indications for the hierarchal use of sugar sources, reflecting the adaption of *A. muciniphila* to a mucus–degrading lifestyle.

Materials & methods

Bacterial growth conditions. *Akkermansia muciniphila* MucT (ATTC BAA-835) was grown in a basal medium as described previously except without the addition of rumen fluid (Derrien *et al.*, 2004) and supplemented with 1 mM L-threonine and 8 g/l tryptone. This medium was termed Basal Tryptone Threonine Medium (BTTM). BTTM was supplemented with carbon and nitrogen sources (Table 1) and used to test the growth of *A. muciniphila* on single sugars. Hog gastric mucin (Type III; Sigma-Aldrich, St. Louis, MO, USA) was purified by ethanol precipitation as described previously (Miller and Hoskins, 1981). All medium components were purchased from Sigma-Aldrich, except for tryptone (Oxoid Ltd, Basingstoke, Hampshire, England). Incubations were performed in serum bottles sealed with

butyl-rubber stoppers at 37ºC under anaerobic conditions provided by a gas phase of 182 kPa (1.5 atm) N_2/CO_2. Culture volume was 20 ml. The cultures were inoculated with 2×10^7 bacterial cells from bacterial cultures that were three times transferred on the respective carbohydrate before start of the experiment, except for the cultures with galactose or a mixture of more than one sugar, which were inoculated from a culture grown on mucin. Growth was followed spectrophotometrically by determining the optical density at 600 nm (OD600).

For RNA-seq analysis, the bacteria were grown on mucin (0.5 %), D-glucose (20 mM) or a mixture of D-glucose (10 mM) and mucin (0.25 %). Cultures with only glucose were supplemented with 16 g/l tryptone and 10 mM threonine. The culture volume for cultures containing mucin was 50 ml and for cultures with only glucose 100 ml. Cells from cultures containing mucin were harvested for RNA-extraction at exponential phase 10 - 12 h of incubation and cells from cultures containing only glucose at 32 – 33 h (Figure 1, Figure S1, Table S1, S2, S3). For each condition triplicate cultures were used. Cells were collected by centrifugation at 4800 g for 15 min at 4°C. Cell pellets were suspended to Trizol® Reagent (Ambion, Life Technologies, Carlsbad, CA, USA) and stored at −80°C until RNA was purified.

High-performance liquid chromatography. For fermentation product analysis, samples were taken at different time points during the incubation. One ml of bacterial culture was centrifuged and the supernatant was stored at −20°C until further analysis. Substrate conversion and product formation were measured with a Thermo Scientific Spectrasystem high-performance liquid chromatography (HPLC) system equipped with a Varian Metacarb 67H 300 × 6.5 mm column kept at 45°C and run with 0.005 mM sulfuric acid as eluent. The eluent had a flow of 0.8 ml/min and metabolites were detected by determining the refractive index and identified by using standards of pure compounds as described previously (van Gelder *et al.*, 2012).

RNA-purification. Total RNA was isolated by a method combining the Trizol® Reagent and the RNeasy Mini kit (QIAGEN GmbH, Hilden, Germany) essentially as describe previously (Chomczynski, 1993; Zoetendal *et al.*, 2006). Briefly, 0.2 ml of chloroform was added per 1 ml of Trizol® Reagent containing the cell pellet. The tube was shaken for 15 s, incubated for 2-3 minutes at room temperature and centrifuged at 12 000 g for 15 min at 4°C. Phase Lock Gel heavy tubes (5 Prime GmbH, Hamburg) were used during phase separation. The upper aqueous phase was removed into a new tube. Subsequently, the RNA isolation was continued with the RNA Cleanup according to the manufacturer's instructions for the RNeasy Mini kit. Genomic DNA was removed by an on-column DNase digestion step during RNA purification (DNase I recombinant, RNase-free, Roche Diagnostics GmbH,

Mannheim, Germany). For the RNA purified from cultures containing only glucose as the growth substrate TURBO™ DNase Treatment was performed according to manufacturer's instructions (Ambion, Life Technologies). Yield and RNA quality was assessed using the Experion™ RNA StdSens Analysis Kit in combination with the Experion™ System (Bio-Rad Laboratories, Inc., Hercules, CA, USA). Depletion of rRNA was performed using the Ribo-Zero™ Kit for bacteria (Epicentre, Madison, WI, USA) according to manufacturer's instructions. The success of the rRNA depletion step was checked using the Experion™ RNA StdSens Analysis Kit in combination with the Experion™ System. Library construction for whole-transcriptome sequencing (RNA-seq) was done using the ScriptSeq™ v2 RNA-Seq Library Preparation Kit in combination with ScriptSeq™ Index PCR primers (Epicentre) according to the manufacturer's instructions.

Transcriptome analysis. The barcoded cDNA libraries were sent to BaseClear (Leiden, The Netherlands) where they were pooled, and 50 bp sequencing (single end reads) was performed on two lanes using the Illumina HiSeq2500 platform in combination with the TruSeq Rapid SBS and TruSeq Rapid SR Cluster Kits (Illumina).

Reads were mapped to the genome with Bowtie2 v2.2.1 (Langmead and Salzberg, 2012) using default settings, and BAM files were converted with SAMtools v0.1.19 (Li *et al.*, 2009). BEDTools v2.17.0 was used to determine the read count for each protein coding region (Quinlan and Hall, 2010). Only reads with a minimum 30% length overlap and mapped on the correct strand were counted. Details on the RNA-seq raw data analysis can be found in Table S4 in the supplemental material. Differential gene expression was assessed using edgeR (Robinson *et al.*, 2010) with default trimmed mean of M-values (TMM) settings. Raw sequence files can be found at the ENA repository under sample group ERG004891.

Proteome analysis. The whole proteome of *A. muciniphila* was obtained from cultures grown with mucin or glucose as the carbon source. Bacterial cells from an overnight 2 ml culture were spun down, washed with PBS and suspended to 500 µl of PBS. Cells were lysed by sonication, using a Branson sonifier equipped with a 3 mm tip (four pulses of 20 s with 30 s rest on ice in-between each pulse, strength of the pulse was 4). The samples were stored in 2 ml low binding tubes (Eppendorf, Hamburg, Germany) at −20 °C. Qubit® Protein Assay Kit (Life technologies, Oregon, USA) was used according to the manufacturer's instructions to determine the protein content of cell extracts. Samples were loaded on a 10 % acrylamide separation gel (25201, Precise™ Protein Gels, Thermo Scientific, Rockford, IL, USA) using the mini-PROTEAN 3 cell (Bio-Rad Laboratories, Hercules, CA, USA). The electrophoresis procedure was according to the manufacturer's instructions.

Gels were stained using CBB R-250 as indicated in the protocol of the mini-PROTEAN 3 system.

In-gel digestion of proteins and purification of peptides were done following a modified version of the protocol described earlier (Rupakula *et al.*, 2013). Disulfide bridges in proteins were reduced by covering whole gels with reducing solution (10 mM dithiothreitol, pH 7.6, in 50 mM NH_4HCO_3), and the gels were incubated at 60°C for 1 h. Alkylation was performed for 1 h by adding 25 ml of iodoacetamide solution (10 mM iodoacetamide in 100 mM Tris-HCl, pH 8.0). Gels were thoroughly rinsed with double distilled H_2O water in between steps. Each lane of SDS-PAGE gels was cut into five slices, and the slices were cut into approximately 1 mm^3 cubes and transferred to separate 0.5 ml protein LoBind tubes (Eppendorf, Hamburg, Germany). Enzymatic digestion was done by adding 50 µl of trypsin solution (5 ng/µl trypsin in 50 mM NH_4HCO_3) to each tube, and by incubating at room temperature overnight with gentle shaking. Extraction of peptides was performed with manual sonication in an ultrasonic water bath for 1 s before the supernatant was transferred to a clean protein LoBind tube. Trifluoroacetic acid (10 %) was added to the supernatant to reach a pH between 2 and 4. The supernatant was used for liquid chromatography-tandem mass spectrometry (LC-MS/MS) analysis. Samples were measured by nLC–MS/MS with a Proxeon EASY nLC and a LTQ-Orbitrap XL mass spectrometer as previously described (Lu *et al.*, 2011).

LC – MS data analysis (false discovery rates were set to 0.01 on peptide and protein level) and additional result filtering (minimally 2 peptides necessary for protein identification of which at least one is unique and at least one is unmodified) was performed as described previously (Rupakula *et al.*, 2013). To analyze the abundance of proteins, their label-free quantification (LFQ) intensities were compared (Cox *et al.*, 2014). Non-existing LFQ intensity values due to not enough quantified peptides were substituted with a value lower than the LFQ intensity value for the least abundant, detected peptide. Differential protein abundance was assessed using edgeR similarly to the transcriptome analysis.

Construction of a metabolic model. A genome-scale constrained-based model of metabolism and transport in *A. muciniphila* was constructed. A draft model was generated using pathway-tools GPR mapping and gap filling for improvement of genomic annotation (Karp *et al.*, 2010). The model was manually curated in the MATlab COBRA environment to get a functional model (Schellenberger *et al.*, 2011). By considering maximization of growth yield as objective function, the model was subsequently used to qualitatively test amino acid auxotrophy, vitamin and cofactor biosynthesis, and sugar degrading capacities.

Results & discussion

The metabolic model predicted *A. muciniphila* to be able to synthesize all the essential amino acids, with the exception of threonine. This is a first adaptation to the mucosal environment as threonine is abundantly present in mucin (Schrager, 1970). Furthermore, according to the metabolic model, *A. muciniphila* is able to individually metabolize all the main sugars present in mucin (fucose, galactose, GlcNAc, N-acetylneuraminic acid), except for GalNAc. Hence, a mucin-free medium, BTTM, containing threonine and the amino acid source tryptone was designed to test *A. muciniphila* growth on single sugars. The model-based predicted fermentation profiles for each of the monomeric sugars are shown in Table 1. The Gibbs energy of the reactions was calculated using eQuilibrator (Flamholz *et al.*, 2012). The current version of the metabolic model contains 746 reactions and 736 metabolites to which a total of 337 unique genes are associated.

Table 1. Summary of the predicted degradation of mucin-derived monomeric sugars by *A. muciniphila*. ΔrG'° is the standard reaction Gibbs energy (1M), ΔrG'm is the reaction Gibbs energy with 1 mM standard which is more suitable for comparing reactions in biological settings. The pathways for degradation of sugars were derived from the constructed metabolic model.

Substrate	Degradation reaction	ΔrG'm	ΔrG'°
Glucose	Glucose(aq) ⇌ 4/3 Propionate(aq) + 2/3 Acetate(aq) + 2/3 CO_2(aq) + 2/3 H_2O(l)	-340,1	-311,6
Galactose	Galactose(aq) ⇌ 4/3 Propionate(aq) + 2/3 Acetate(aq) + 2/3 CO_2(aq) + 2/3 H_2O(l)	-347,1	-318,5
Fucose	Fucose(aq) + 1/3 H_2O(l) ⇌ 1,2-Propanediol(aq) + 2/3 Acetate(aq) + 1/3 Propionate(aq) + 2/3 CO_2(aq)	-187,3	-158,8
GlcNAc	N-Acetyl-D-glucosamine(aq) + 4/3 H_2O(l) ⇌ 5/3 Acetate(aq) + Ammonia(aq) + 4/3 Propionate(aq) + 2/3 CO_2(aq)	-375,6	-312,9
GalNAc	N-Acetyl-D-galactosamine(aq) + 4/3 H_2O(l) ⇌ 5/3 Acetate(aq) + 3/3 Ammonia(aq) + 4/3 Propionate(aq) + 2/3 CO_2(aq)	-376,1	-313,3
NeuAc	N-Acetylneuraminate(aq) + 5/3 H_2O(l) ⇌ 5/3 Propionate(aq) + 7/3 Acetate(aq) + 4/3 CO_2(aq) + Ammonia(aq)	-477,2	-385,9

GlcNAc, N-acetylglucosamine; GalNAc, N-acetylgalactosamine; NeuAc, N-acetylneuraminic acid

Qualitative testing of the model

It was confirmed that *A. muciniphila* could grow, but with limited efficiency as compared to mucin, on fucose, GlcNAc and GalNAc as the main energy source on BTTM (Figure 1B, C, D). Measuring growth rates was in some cases challenging because a clear exponential phase was absent, suggesting that some unknown

limitation was occurring. However, it was obvious that growth rates on single sugars were markedly lower than on mucin (Table 2). Moreover, while GlcNAc, GalNAc and to some extent fucose were supporting growth, N-acetylneuraminic acid did not (**Chapter 4**). Utilization of fucose was confirmed by the presence of 1,2-propanediol, an end product of fucose fermentation, in the culture supernatant. In addition to N-acetylneuraminic acid, also galactose did not support growth in our *in vitro* experiments even though the genome predicts the ability to ferment galactose through the Leloir pathway. We proceeded to test the utilization of galactose in combination with mucin, and observed that galactose was not only taken up but it was depleted very rapidly, between 15 and 24 h (Figure 2G). This suggests that *A. muciniphila* indeed has the system for efficient galactose metabolism, but additional, mucin-derived components are needed for induction of the galactose utilization pathway. A transporter for galactose has not been identified in the genome of *A. muciniphila,* and as fermentation of galactose requires several steps in the metabolic process, further research is needed to recognize the inducible genes.

Table 2. Growth rates of *A. muciniphila* on different carbon and nitrogen sources

Growth substrate	μ (h^{-1})
Mucin 0.5 %	0.41 ± 0.05
Mucin 0.25 %	0.38 ± 0.04
D-glucose 10 mM + mucin 0.25 %	0.37 ± 0.03
L-Fucose 10 mM + mucin 0.25 %	0.40 ± 0.04
N-acetylglucosamine 10 mM + mucin 0.25 %	0.40 ± 0.03
N-acetylgalactosamine 10 mM + mucin 0.25 %	0.42 ± 0.04
D-galactose 10 mM + mucin 0.25 %	0.39 ± 0.02
D-glucose 20 mM[1]	0.16 ± 0.01
L-fucose 20 mM[1]	0.08 ± 0.01
N-acetylglucosamine 20 mM[1]	0.17 ± 0.02
N-acetylgalactosamine 20 mM[1]	0.18 ± 0.03
D-galactose 20 mM[1]	0.09 ± 0.03
D-glucose 2.5 mM + L-fucose 2.5 mM + GlcNAc 2.5 mM + GalNAc 2.5 mM[1]	0.20 ± 0.01
D-glucose 10 mM + L-fucose 10 mM[1]	0.15 ± 0.02

[1]Medium was supplemented with L-threonine (10 mM) and tryptone (8 g/l).
GlcNAc; N-acetylglucosamine, GalNAc; N-acetylgalactosamine

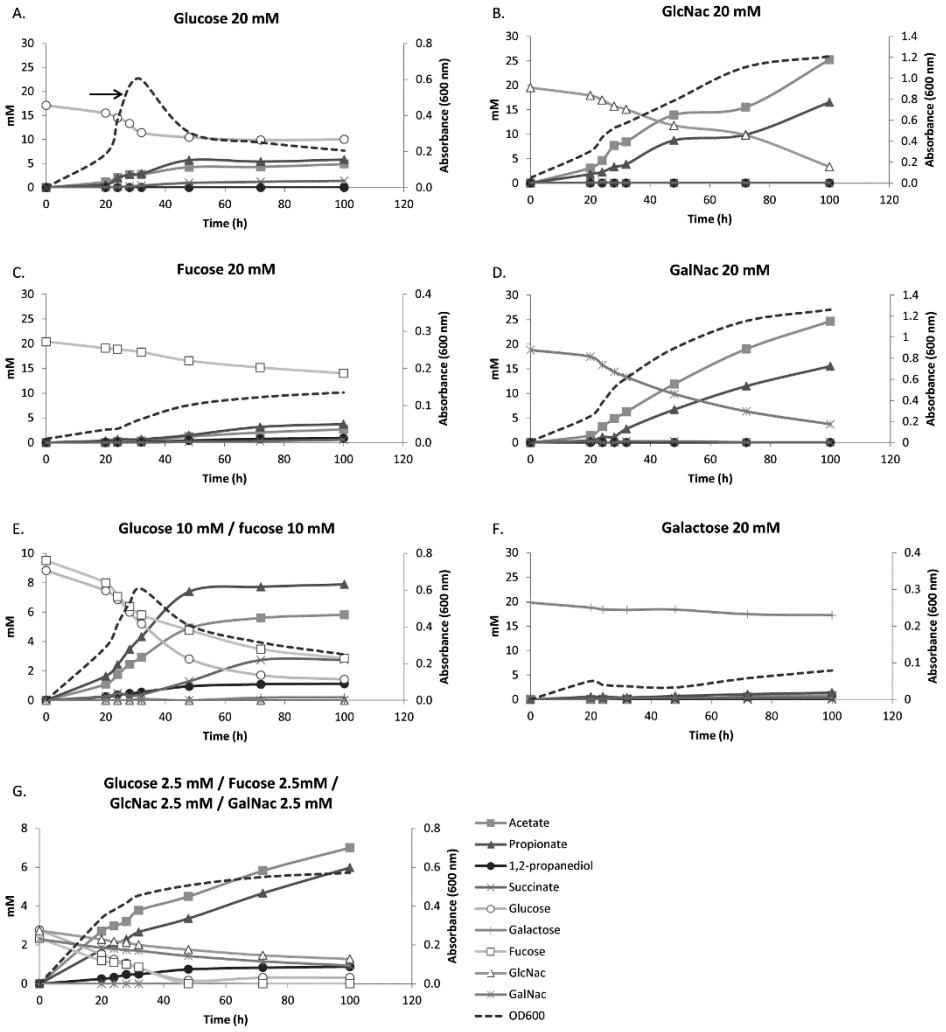


Figure 1. *A. muciniphila* utilization of glucose and mucin-derived sugars. *A. muciniphila* is able to efficiently grow on glucose (A), N-acetylglucosamine (GlcNAc) (B) and N-acetylgalactosamine (GalNAc) (D). Highest yields are reached when grown on GlcNAc or GalNAc. Galactose (F) and fucose (C) degradation by *A. muciniphila* is weak. However, in a mixture (G) glucose and fucose are the preferred substrates. Values represent mean of three replicate experiments. Arrow indicates representative sampling point for RNA-extraction for the transcriptome analysis of *A. muciniphila* grown on glucose.

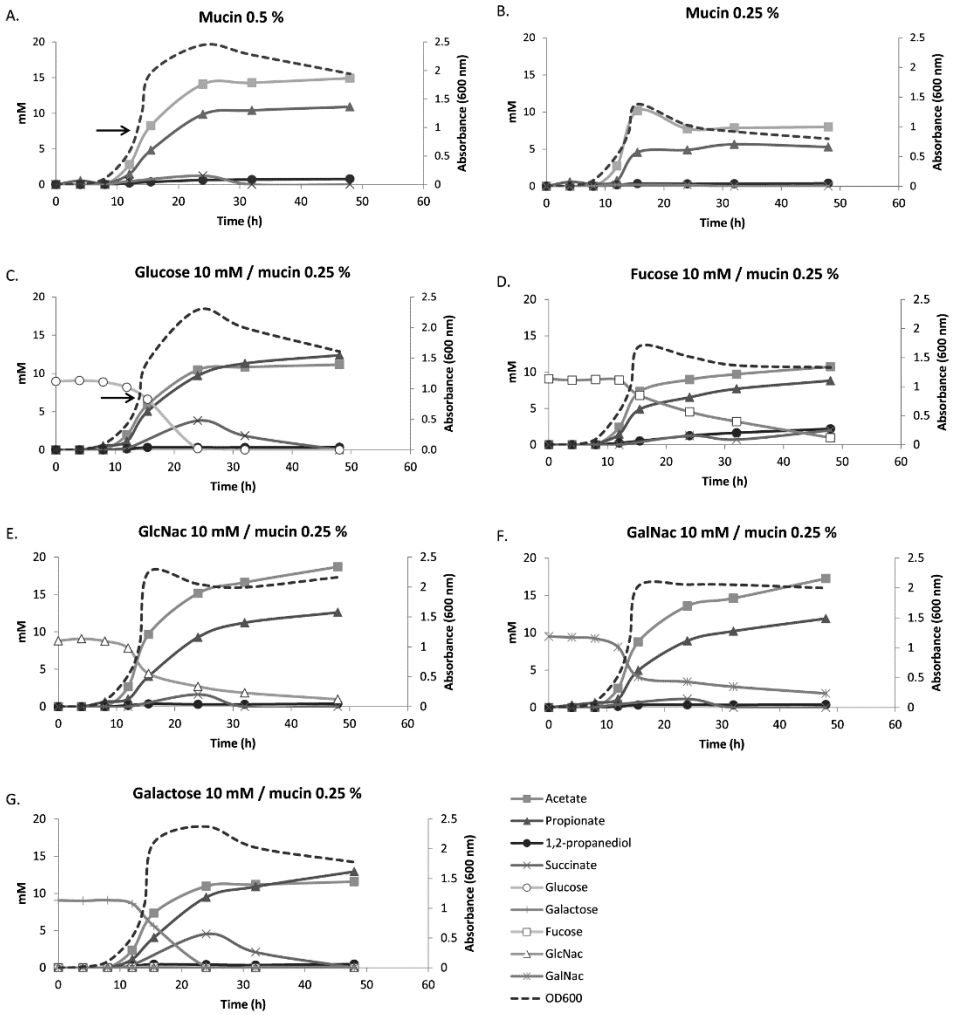


Figure 2. *A. muciniphila* degrades single sugars more efficiently in the presence of mucin. Decreasing the concentration of mucin from 0.5 % (A) to 0.25 % (B) decreases yield, but it can be compensated with the addition of single sugars. Added sugars, including fucose (D) and galactose (G), are depleted within 48 h of growth. Values represent mean of three replicate experiments. Arrows indicate sampling point for RNA-extraction for the transcriptome analysis of *A. muciniphila*.

A. muciniphila was also able to grow on BTTM when the non-mucin sugar glucose was added as a carbon source (Figure 1A). Interestingly, growth on glucose was exponential, but only ~6 mM glucose was utilized, after which the OD600 dropped and glucose utilization slowed down considerably. A similar trend was observed when the bacteria were grown on a mixture of glucose and fucose (Figure 1E). This indicates a depletion of an unknown factor, possibly amino acids or proteins, which prevents further utilization of glucose and growth. Conversely, when grown on GlcNAc or GalNAc, the exponential phase was very short after which the OD600 increased linearly until the end of measurements (100 h of incubation), and substrate utilization was stable (Figure 1B, D).

The main difference between these sugars is the presence of an N-acetyl group in GlcNAc and GalNAc, as compared to glucose and fucose. One benefit of metabolizing amino sugars is the production of ammonia, which can neutralize the environment, which is acidified as a result of carbohydrate catabolism to organic acids, such as acetate and propionate. Furthermore, the production of ammonia inside the cells has the dual advantage of neutralizing the cytoplasm and providing a nitrogen source. This eliminates the need to devote energy to importing ammonium ions or other nitrogen sources from the environment.

Growing *A. muciniphila* on BTTM with a mixture of the single sugars (fucose, glucose, GlcNAc and GalNAc; 2.5 mM each) resulted in co-metabolism of all the sugars, but with differences in consumption rates (Figure 1G). The consumption of glucose and fucose was three times faster than the consumption of GlcNAc and GalNAc at 20 h of incubation, and GlcNAc and GalNAc still remained in the medium after 100 h. After the depletion of glucose and fucose, the OD600 increased linearly and no dip was observed, unlike in the case of glucose alone or in combination with fucose. This faster and longer growth indicates that the culture grown on a mixture of sugars shows the combined characteristics of *A. muciniphila* growth on both non-amino sugars (glucose and fucose) as well as hexosamines (GlcNAc and GalNAc). A simple explanation could be that the amino sugars serve as a nitrogen source, and that the non-amino sugars glucose and fucose are rapidly transported and used for energy generation.

The co-metabolism of sugars could be a consequence of the adaptation of *A. muciniphila* to the mucin environment, where a mixture of sugars is constantly present and the bacteria produce a wide range of enzymes simultaneously to degrade them. To our surprise, the non-amino sugars were preferred over the thermodynamically more favorable hexosamines (Table 1). This could be some form of catabolite repression of monosaccharides even though the canonical catabolite repression protein was not found to be encoded by the genome of *A. muciniphila*.

The release and subsequent consumption of fucose, galactose and sialic acid from mucin by *A. muciniphila* was observed using high-performance anion-exchange chromatography (Derrien, 2007). However, the exact degradation pattern of mucin by *A. muciniphila* is not known, and the faster consumption of fucose in comparison to GlcNAc and GalNAc may reflect the manner of degradation taking place *in vivo*. The faster uptake of glucose in the presence of mucin-derived sugars may also reflect the situation in the gut, where glucose might be present together with mucin.

Interestingly, the genome predicted two amylase-encoding genes (Amuc_1637, Amuc_1812), the expression of which was found to be upregulated on glucose in comparison to mucin (see below), indicating that *A. muciniphila* prefers to use glucose. We also tested the growth of *A. muciniphila* on the di- and trisaccharides maltose, melibiose, trehalose and raffinose, but observed no growth or metabolic activity (data not shown).

Eventually all the metabolized sugars ended up going through glycolysis, but the routes leading up to it and the final fermentation profiles were found to be different. Differences in growth kinetics of *A. muciniphila* grown on glucose as opposed to mucin-derived sugars may be due transcriptional regulation of transporters or kinases in response to environmental conditions. Overall, these data suggest that the high-energy nitrogen-carbon compounds are important for *A. muciniphila*. This reflects the adaptation of *A. muciniphila* to the intestinal environment, as GlcNAc and GalNAc are abundant in mucin.

Quantitative testing of the model

Our results indicate that the BTTM medium with addition of single sugars as carbon sources does not support optimal growth, as the growth rates are higher on mucin medium. When grown on 0.5 % mucin, *A. muciniphila* grows to an OD600 of ~2.5 and produces acetate and propionate, and low amounts of 1,2-propanediol and succinate (Figure 2A). As the exact decomposition route of mucin by *A. muciniphila* is unknown, the expected ratios of fermentation products cannot be predicted. Assuming that mucin contains signaling molecules essential for optimal growth, we set out to lower the mucin concentration and compensate carbon availability by the addition of single sugars. Reducing the concentration of mucin from 0.5 % to 0.25 % resulted in lower biomass yields (OD600 of ~1.4) and also halved the amount of fermentation products, indicating substrate availability is the limiting factor (Figure 2B). Addition of 10 mM of the tested individual sugars together with 0.25 % mucin partially compensated for the decrease in OD600 and amount of fermentation products (Figure 2C to 2G). This in turn implies that the energy

source limits the growth. More importantly, all the sugars, including fucose and galactose were used by *A. muciniphila*, and were depleted 48 h after incubation.

The metabolic model predicted a 1:2 ratio of acetate to propionate upon growth on glucose as the sole carbon source, but we detected equal amounts of these fermentation products. For growth on GlcNAc and GalNAc, a 5:4 ratio of acetate to propionate was expected, but we observed 1.5 to 2 times more acetate than propionate. The carbon recovery for glucose, GlcNAc and GalNAc cultures was 80 – 90 %. These discrepancies between the metabolic model and experimental data suggest the presence of an alternative electron sink. Succinate was produced as an intermediate compound in the propionate pathway. Succinate can function as an alternative electron sink (Mattam and Yazdani, 2013), and may partially explain the higher than expected acetate to propionate ratios. Succinate production co-occurred with high growth rate and long exponential phase as cultures with glucose or mucin as the carbon source showed higher succinate levels compared to other cultures. In these conditions, conversion from succinate to propionate seems to be a metabolic bottleneck, and as excess succinate is accumulated it is excreted outside the cell. In *Bacteroides* the balance between succinate and propionate is controlled by both CO_2 and vitamin B12 availability (Fischbach and Sonnenburg, 2011). Higher succinate levels can be a means to reduce CO_2 levels because conversion of succinate to propionate releases CO_2. It may also be that during fast growth, the rate of B12 uptake limits conversion of succinate to propionate, as B12 is the cofactor for methylmalonyl-CoA mutase, which is required in the process. B12 is not essential for growth of *A. muciniphila*, but a part of the B12 biosynthesis pathway was identified from its genome. The uptake of succinate at the stationary phase suggests that there is an energetic gain to further metabolize it into propionate. This is supported by the model where the final step is coupled to sodium transport, which in turn fuels proton translocation and ATP generation.

Interestingly, acetate is not produced from acetyl-CoA via the commonly used acetate kinase. Instead the CoA group is transferred to succinate to generate succinyl-CoA, an intermediate of the propionate cycle, and conversely ATP is generated via the succinyl-CoA ligase complex. This interlink between the acetate and propionate pathway may serve to enhance thermodynamic properties of the primary ATP generating pathways.

As predicted by the metabolic model, 1,2-propanediol was produced when *A. muciniphila* was grown on fucose as the sole carbon source. However, the ratio of 1,2-propanediol produced in comparison to acetate and propionate was lower than predicted (3:2:1, respectively). The carbon recovery for *A. muciniphila* cultures grown on fucose was 87 %. The amount of 1,2-propanediol should be equal to the

amount of fucose consumed, but we observed less than 1 mM of 1,2-propanediol after consumption of 6 mM fucose. Furthermore, the levels of propionate production were higher than expected. A pathway in which 1,2-propanediol is converted to propionate exists in some members of the gut microbiota, but the genes required for the conversion are not present in *A. muciniphila* (Reichardt *et al.*, 2014).

It has previously been shown that growth rates and bacterial cell yields are higher on oligosaccharides than on their monomeric moieties (Palframan *et al.*, 2003; Tsujikawa *et al.*, 2013; Van der Meulen *et al.*, 2004). It can be speculated that *A. muciniphila* substrate transporter systems are designed for oligosaccharide transport and metabolic features have been optimized for growth on complex carbohydrates.

Transcriptomic and proteomic response to *A. muciniphila* growth on mucin and glucose

To answer some of the questions that arose from the growth experiments, a whole-genome transcriptome analysis of *A. muciniphila* grown on mucin, glucose and a mixture of the two was performed. In addition, we analyzed the proteome of *A. muciniphila* grown on mucin or glucose with mass spectrometry. The proteome analysis showed highly similar results compared to the transcriptional response corroborating the findings of both methods (Figure 3). As the similarity is so high (Spearman rank coefficient: 0.581), we focused here mainly on the transcriptome data, as these were deeper and more global as compared to the proteome results.

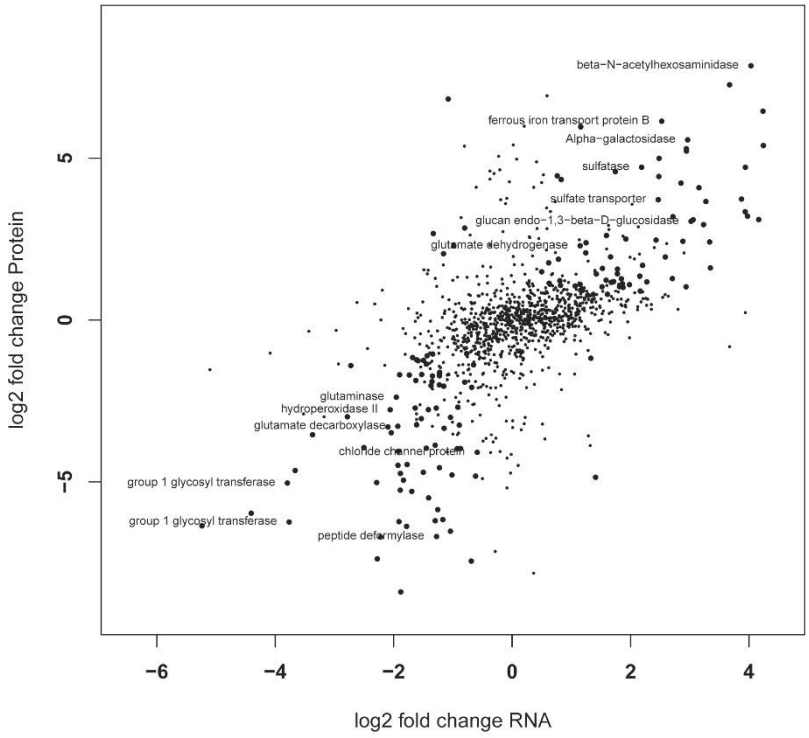


Figure 3. Correlation between fold change in relative abundance on transcript and protein level. Genes indicated with a bigger circle represent genes, which were significantly differentially abundant on both transcriptome and proteome level. Positive values correspond to the mucin condition while negative values reflect the glucose condition. The Spearman rank coefficient was calculated to be 0.581.

Results showed the largest difference in transcriptome profiles between the mucin and glucose cultures, while the transcriptomes of cultures grown on the mixture were highly similar to the mucin cultures. Differential gene expression (DGE) showed only 57 genes to be significantly (p-value < 0.05) differentially expressed between the mucin and the mixture (mucin+glucose) cultures. Downregulated in the mixture cultures were mainly genes encoding hydrolases, indicating a potential lower requirement for carbohydrate sources normally obtained by release from the mucin substrate (Figure 4B). Amongst the upregulated genes were those encoding glycosyl transferases and a glucose kinase (Amuc_1094). Comparing the glucose culture with the mucin culture showed many more genes to be significantly differentially expressed (1074 in total and 657 more than 2-fold). In addition to the processes already seen in the comparison between the mixture and the mucin culture, genes involved in protein biosynthesis and energy metabolism were upregulated in the mucin condition (Figure 4A). The mass spectrometry analysis of *A. muciniphila* proteome identified peptides from 1129 proteins (52 % of all predicted protein-coding sequences). The abundance of 207 proteins differed significantly (fold change ≥ 2, p-value < 0.05) between mucin and glucose-grown *A. muciniphila*.

Under mucin condition genes encoding ribosomal proteins, tRNA-charging proteins and most amino acid biosynthesis pathways were upregulated. Biosynthesis pathways for cysteine, proline, glutamate and glutamine were downregulated. While cysteine and proline are both highly abundant in the mucin peptide backbone this likely reflects the uptake of these amino acids. Interestingly the incorporation of ammonia via glutamine synthase (Amuc_1252) was downregulated under mucin conditions, indicating a reduced requirement for ammonia assimilation.

Glutamate decarboxylase (Amuc_0372) was significantly upregulated in glucose cultures on both transcriptome and proteome level. The glutamate decarboxylase (GAD) system is known to be activated in response to acid stress and multiple other stresses in many bacteria (Feehily and Karatzas, 2013). Oxidative stress response (Amuc_1466, Amuc_2070) and nucleotide excision repair genes (Amuc_1452, Amuc_1555, Amuc_1237, Amuc_0683, Amuc_1938, Amuc_0011) were also upregulated in the glucose cultures, indicating an organism dealing with a non-optimal, stressful environment. Other processes upregulated in the glucose cultures were polysaccharide biosynthesis (Amuc_2096, Amuc_2077, Amuc_2078), inorganic ion transport and motility (Amuc_0166, Amuc_1101, Amuc_1584).

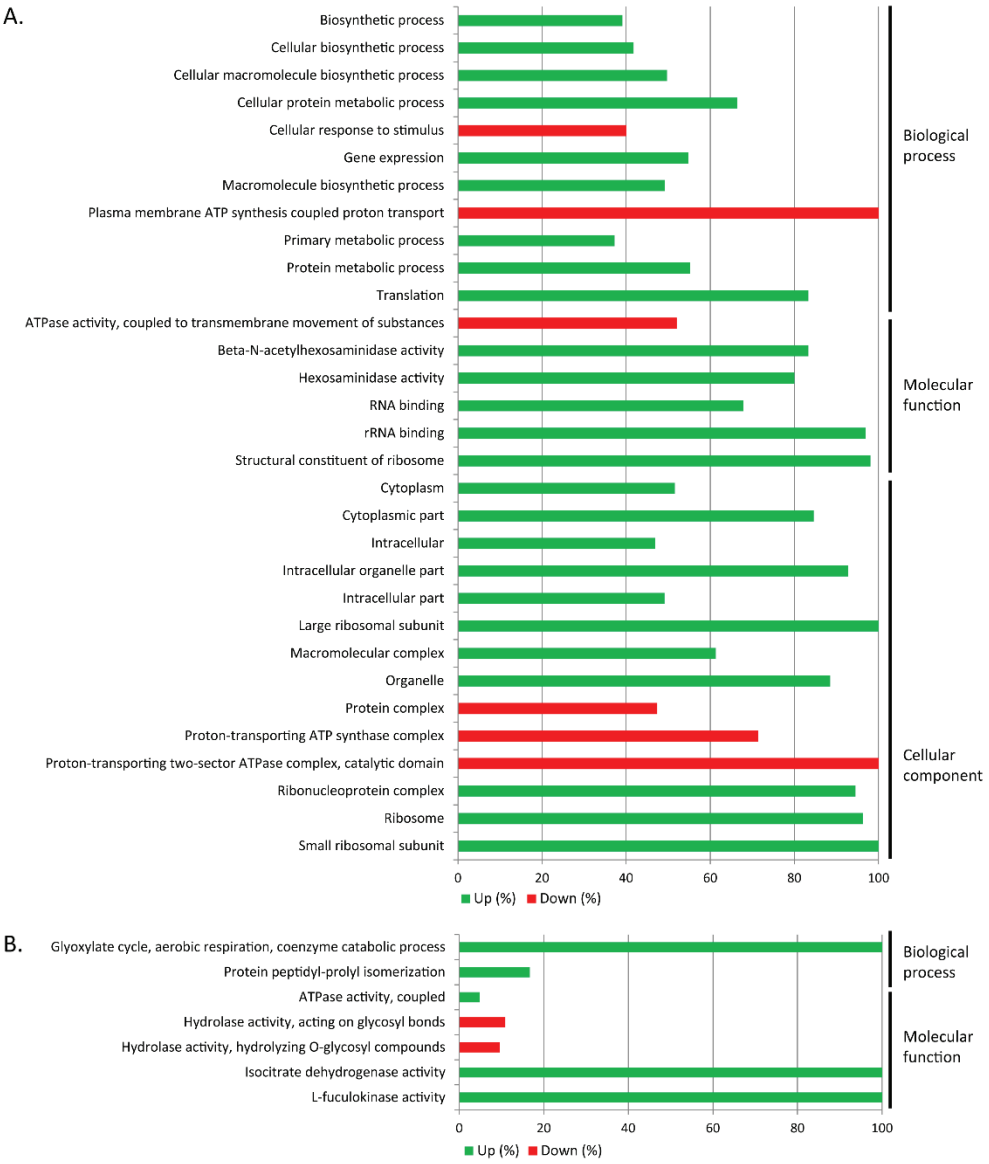


Figure 4. Gene ontology grouping of differentially regulated genes in *A. muciniphila*. Each bar represents the ratio of the number of differentially expressed genes belonging to the GO terms shown in the y-axis. (A) *A. muciniphila* grown on mucin vs. glucose. Green, upregulated in mucin; red, downregulated in mucin. (B) *A. muciniphila* grown on mucin vs. mixture (mucin+glucose). Green, upregulated in mixture; red, downregulated in mixture.

A total of 30 hydrolases implicated in mucin degradation were significantly upregulated in the mucin condition compared to glucose. 17 of these were also significantly more abundant in the proteome analysis. Conversely, in the glucose condition 15 glycosyl transferases were significantly upregulated, including the entire gene cluster Amuc_2079 to Amuc_2098, which contains nine glycosyl transferase encoding genes. It is likely that *A. muciniphila* can produce exopolysaccharides (EPS) as a protection to the sometimes harsh conditions present in the intestinal environment. Moreover, the production of EPS appears to be induced by glucose and this is likely to be the case when *A. muciniphila* is not anymore located in the protective mucus layer but has a luminal location where other intestinal bacteria are competing for food-derived substrates, such as glucose.

Out of the 12 sulfatases predicted to be encoded in the genome of *A. muciniphila*, six were significantly upregulated in the mucin condition, along with the sulfate transporter (Amuc_0840). On proteome level one of the sulfatases (Amuc_0491) and the sulfate transporter (Amuc_0840) were also significantly more abundant in the mucin condition in comparison to glucose condition. Mucin contains substantial levels of sulfate covalently bound to the oligosaccharide chains (Wright *et al.*, 2000), and the sulfatases expressed by *A. muciniphila* most likely desulfate this to increase the susceptibility of the mucin to degradation by other mucin-degrading enzymes (Derrien, 2007).

Out of all the significantly differentially expressed hypothetical proteins, almost 70 % were upregulated in the glucose cultures on both transcriptome and proteome level. More insight on the adaption of *A. muciniphila* to glucose could be gained by further research on the functions of these proteins.

In conclusion, the upregulation of genes and proteins involved in mucin degradation in *A. muciniphila* grown on mucin is consistent with the adaptation of *A. muciniphila* to this environment. Besides the identified glucokinase (Amuc_1094), it was more challenging to pin down genes specific to the glucose environment. As the growth rate and yield of *A. muciniphila* are higher during growth on mucin than on glucose, many of the changes in gene expression may be involved in a global response to fast growth, rather than the specific carbon source.

Supplemental data

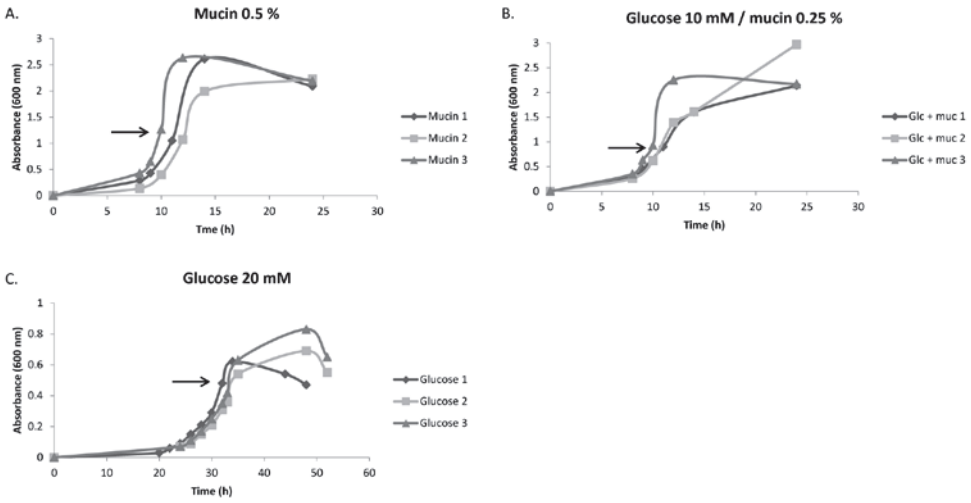


Figure S1. *A. muciniphila* cultures were harvested in mid-exponential phase for RNA-extraction. *A. muciniphila* was grown on mucin (A), a mixture of glucose and mucin (B), or glucose (C). Three replicates of each condition were sampled. Arrows indicate time points for sampling.

Table S1. Fermentation end products of *A. muciniphila* during growth on different substrates. Samples were obtained during mid-exponential phase (~10-12 h incubation, used for transcriptome analysis) and in stationary phase (24 h incubation). Values represent means of triplicate cultures with standard deviations. N.D., not detected.

	Acetate		Propionate		Succinate		1,2-propanediol	
	10-12 h	24 h	10-12 h	24 h	10-12 h	24 h	10-12 h	24 h
Mucin 0.5 %	6.1 ± 2.4	17.1 ± 3.6	4.1 ± 2.3	12.4 ± 3.6	1.8 ± 1.0	0.1 ± 0.1	0.0 ± 0.0	0.7 ± 0.1
Glucose 10 mM / mucin 0.25 %	6.0 ± 3.7	12.2 ± 3.0	3.1 ± 4.0	12.8 ± 5.0	2.0 ± 0.6	1.1 ± 0.5	0.1 ± 0.1	0.1 ± 0.2

Table S2. Degradation of sugars in *A. muciniphila* cultures during growth on different substrates. Values represent means of triplicate cultures with standard deviations. N.D., not detected.

	Glucose		Fucose		Galactose	
	10-12 h	24 h	10-12 h	24 h	10-12 h	24 h
Mucin 0.5 %	N.D.	N.D.	0.1 ± 0.1	N.D.	0.1 ± 0.1	N.D.
Glucose 10 mM / mucin 0.25 %	-1.7 ± 0.6	-8.3 ± 0.0	N.D.	N.D.	0.3 ± 0.2	0.0 ± 0.1

Table S3. Glucose degradation and fermentation end products of *A. muciniphila* during growth on glucose. Samples were obtained during mid-exponential phase (~32-33 h incubation, used for transcriptome analysis) and in stationary phase (48 h incubation). Values represent means of triplicate cultures with standard deviations.

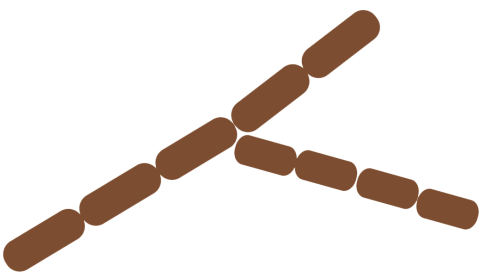
	Glucose		Acetate		Propionate		Succinate	
	32-33 h	48 h	32-33 h	48 h	32-33 h	48 h	32-33 h	48 h
Glucose 20 mM	-4.9 ± 0.9	-11.5 ± 2.8	1.7 ± 0.7	5.5 ± 0.6	1.6 ± 0.1	8.0 ± 0.9	0.6 ± 0.1	1.8 ± 0.2

Table S4. Summary of the RNA-seq raw data analysis for transcriptome analysis of *A. muciniphila* grown on glucose, mucin and a mixture of mucin and glucose.

	Glucose			Mucin			Mucin + glucose		
	1	2	3	1	2	3	1	2	3
Total no. of reads	11047878	10545771	18497197	11645389	17180153	11743614	16724960	14660237	17949331
Total no. of reads mapped to the genome	10302029	10206943	17953823	11459340	16915727	11596495	16215886	14271680	17217612
Percentage of reads mapped to the genome	93.2	96.8	97.1	98.4	98.5	98.7	97.0	97.3	95.9
Total no. of reads mapped to protein coding regions	7063596	6904150	12096349	8399083	11766433	8157964	12130637	9767040	11811538

Chapter 4

Mucolytic activity of *Akkermansia muciniphila* enables human milk degradation and growth



Noora Ottman, Mark Davids, Rocio Martin, Jan Knol, Hauke Smidt, Willem M. de Vos, and Clara Belzer

Abstract

Microbial colonization of the infant gut is affected by several environmental factors, and the acquired gut microbiota composition may be associated with health in later life. One of the elements shaping the microbiota is the glycan composition of human milk. Human milk oligosaccharides (HMOs) function as natural prebiotics stimulating the growth of beneficial bacteria. The mucus-resident human commensal *Akkermansia muciniphila* is present in the infant intestine from early life and associated with a healthy intestine in adults. In this study, we evaluated the ability of *A. muciniphila* to persist in the human milk environment. Growth experiments demonstrated metabolic activity of *A. muciniphila* grown on human milk, lactose and one of the HMOs, 2'-fucosyllactose. Transcriptome analysis by RNA-seq showed upregulation of genes involved in lactose degradation, carbohydrate transport and translational activity in *A. muciniphila* grown on human milk in comparison to mucin. However, the expression of genes involved in mucin degradation was not affected by the experimental conditions, suggesting that *A. muciniphila* is capable of using these genes also for utilization of human milk glycans due to the similar structures of HMOs and mucin. The capacity to survive in the early life environment by degrading and consuming human milk components might be beneficial for *A. muciniphila* during initial colonization before reaching the mucosal layer in the intestine.

Introduction

Early life microbial colonization can have an influence on health in later life, including the risk of developing diseases such as obesity, inflammatory bowel diseases and asthma (Arrieta *et al.*, 2014; Nylund *et al.*, 2014). As bacterial colonization coincides with development of the infant's immune system, these two events are expected to be closely linked to each other (Martin *et al.*, 2010; Wopereis *et al.*, 2014). Apart from the host immune system, environmental conditions such as the duration of breast feeding, use of formula's, and antibiotic use also affect acquisition of intestinal homeostasis (Mueller *et al.*, 2015). While fecal material is the most likely source of the new-born's inoculum, it has recently been suggested that microbes present in placenta, vagina, and breast milk also play a role in gut microbiota colonization (Aagaard *et al.*, 2014; Fernandez *et al.*, 2013).

Dietary glycans have a major impact on microbiota composition, and in the infant gut the main source of glycans are human milk oligosaccharides (HMOs). Human milk contains 5 – 15 g/l HMOs, and over 200 different HMO structures exist (Zivkovic *et al.*, 2011). The major monosaccharides present in HMOs are glucose, galactose, N-acetylglucosamine, N-acetylgalactosamine, fucose, and *N*-acetylneuraminic acid (sialic acid). These sugars form complex glycans by numerous different linkages between them, creating both linear and branched structures (Wu *et al.*, 2010). Host genetics also play a role in defining which types of HMOs are present. Human milk can be assigned to four groups based on host expression of α1-2-fucosyltransferase FUT2 and α1-3/4-fucosyltransferase FUT3 genes, which lead to different degrees of fucosylation (Bode, 2012). Interestingly, it was recently shown that relative abundances of some bacterial genera, e.g. *Bifidobacterium*, in infant feces can be predicted by the type of HMOs present in their mother's milk (Wang *et al.*, 2015). Even in adults there appears to be a correlation between the FUT2 secretor type and the colonic microbiota, with *Bifidobacterium* spp. and to some extent *Akkermansia* being associated with the presence of fucose (Wacklin *et al.*, 2014).

The composition of the infant gut microbiota is less diverse and temporally stable than that of the adult microbiota. HMOs are only partially digested in the small intestine, and thus they function as natural prebiotics, selectively stimulating the growth of beneficial bacteria, such as bifidobacteria (Zivkovic *et al.*, 2011) and lactobacilli (Bidart *et al.*, 2014). Several bacteria have been found to be able to grow on human milk, or HMOs, as their sole energy source. These include for example *Bifidobacterium longum* (Gonzalez *et al.*, 2008), *Bacteroides thetaiotaomicron* and *Bacteroides fragilis* (Marcobal *et al.*, 2011). Most of these bacteria are also able to grow on mucin (Derrien *et al.*, 2010).

The structure of HMOs shows chemical similarities to mucin. As few other glycans transit the infant gut, microorganisms capable of utilizing mucin and/or human milk have a selective advantage for colonization over other organisms in early life. It has been reported that in *B. thetaiotaomicron* and *B. fragilis* the same genes that are used to harvest mucin glycans are induced during consumption of HMOs (Marcobal *et al.*, 2011).

Bacteria that are able to degrade mucin efficiently may benefit from the early life environment, where mucin is already present but the microbial complexity is still low. One of the most prominent mucin specialists in the adult mucosa is *Akkermansia muciniphila*. It colonizes the infant gut relatively early, as fecal samples from 1-month-old infants already have low levels of this bacterium (Collado *et al.*, 2007). Subsequently, the amount of *A. muciniphila* increases, reaching a level close to that observed in adults within a year. *Verrucomicrobia*, most likely represented by *A. muciniphila*, have also been identified in the microbiota of human breast tissue (Urbaniak *et al.*, 2014). Notably, two studies have reported lower levels of *A. muciniphila* in breast-fed infants in comparison to formula-feeding (Azad *et al.*, 2013; Bergstrom *et al.*, 2014). However, the study by Azad et al. had a relatively low sample size (24 infants) and lacked longitudinal data as only one sample from each 4-month-old infant was analyzed. Bergstrom et al. studied the microbiota composition of the infants at ages 9, 18 and 36 months, and *A. muciniphila* showed decreased levels in breast-fed in comparison to formula-fed infants only at 9 months of age with relatively low statistical power ($p=0.045$). In a more recent study made with 98 Swedish infants, the abundance and prevalence of *A. muciniphila* increased between 4 and 12 months of age, and showed no significant change depending on delivery mode or type of feeding (Backhed *et al.*, 2015).

A. muciniphila colonization is established in early life and thus, we hypothesized that its mucin degrading ability is favored within the milk environment. This, in turn, might be beneficial for bacterial survival during passage through the gastrointestinal tract, aiding the bacteria to reach its niche, the mucus layer. To study this, we tested the growth of *A. muciniphila* on human milk and its compounds, including HMOs. Additionally, we performed a whole-genome transcriptome analysis to discover whether *A. muciniphila* utilizes its mucin-degrading genes for degradation of human milk compounds.

Materials & methods

Bacterial growth conditions. *A. muciniphila* Muc^T (ATTC BAA-835) was grown in a basal medium as described previously, except without the addition of rumen fluid (Derrien *et al.*, 2004). The medium was supplemented with either hog gastric mucin (0.5 %, Type III; Sigma-Aldrich, St. Louis, MO, USA) purified by ethanol precipitation as described previously (Miller and Hoskins, 1981), human milk (10 %), lactose (10 mM, Sigma-Aldrich), glucose (10 mM, Sigma-Aldrich), galactose (10 mM, Sigma-Aldrich), *N*-acetylneuraminic acid (10 mM, Sigma-Aldrich), 2'-fucosyllactose (10 mM, Danone, Utrecht, The Netherlands), or 3'-sialyllactose (10 mM, Danone). Human breast milk was provided by a healthy donor after approximately 3-6 months of lactation collected under sterile conditions, and stored at −20 ºC until use. Except for the incubations with mucin and human milk, the medium was also supplemented with Bacto™ casitone (BD, Sparks, MD, USA), BBL™ yeast extract (BD), tryptone (Oxoid Ltd, Basingstoke, Hampshire, England) and peptone (Oxoid ltd) (2 g/l each) and L-threonine (Sigma-Aldrich) (2 mM). Incubations were performed in serum bottles sealed with butyl-rubber stoppers at 37ºC under anaerobic conditions provided by a gas phase of 182 kPa (1.5 atm) N_2/CO_2. Growth was measured by spectrophotometer as optical density at 600 nm (OD600). Due to turbidity, measuring the OD was not possible for incubations with human milk. Therefore, cultures were plated on agar plates, and colony forming units (CFUs) were counted. Plates were made by preparing BBL™ Brain Heart Infusion (BD) medium according to the manufacturer's instructions and supplementing it with 1 % bacteriological agar (Oxoid), 0.5 % (w/v) hog gastric mucin and 0.05 % (w/v) cysteine (Sigma-Aldrich). Plates were incubated anaerobically at 37ºC.

For RNA-seq analysis, the bacteria were grown either on mucin or human milk as described above. Cells were harvested at ~17 h (OD600 of ~1.8) of incubation for mucin cultures and ~48 h of incubation for human milk cultures. For each condition, triplicate cultures were set up. Cells were collected by centrifugation at 4800 × g for 15 min at 4°C. Cell pellets were suspended to Trizol® Reagent (Ambion, Life Technologies, Carlsbad, CA, USA) and stored at −80°C until RNA-purification.

RNA-purification. Total RNA was isolated by a method combining the Trizol® Reagent and the RNeasy Mini kit (QIAGEN GmbH, Hilden, Germany) essentially as describe previously (Chomczynski, 1993; Zoetendal *et al.*, 2006). Briefly, 0.2 ml of chloroform was added per 1 ml of Trizol® Reagent containing the cell pellet. The tube was shaken for 15 s, incubated for 2-3 minutes at room temperature and centrifuged at 12 000 g for 15 min at 4°C. Phase Lock Gel heavy tubes (5 Prime

GmbH, Hamburg) were used during phase separation. The upper aqueous phase was removed into a new tube. After this, the RNA isolation was continued with the RNA Cleanup kit according to the manufacturer's instructions for RNeasy Mini kit. Genomic DNA was removed by on-column DNase digestion step during RNA purification (DNase I recombinant, RNase-free, Roche Diagnostics GmbH, Mannheim, Germany). Yield and RNA quality was assessed using the Experion™ RNA StdSens Analysis Kit in combination with the Experion™ System (Bio-Rad Laboratories, Inc., Hercules, CA, USA). Depletion of rRNA was performed using the Ribo-Zero™ Kit for bacteria (Epicentre, Madison, WI, USA) according to manufacturer's instructions. Success of the rRNA depletion step was checked using the Experion™ RNA StdSens Analysis Kit in combination with the Experion™ System. Library construction for whole-transcriptome sequencing (RNA-seq) was done by ScriptSeq™ v2 RNA-Seq Library Preparation Kit in combination with ScriptSeq™Index PCR primers (Epicentre) according to the manufacturer's instructions.

Transcriptome analysis. The barcoded cDNA libraries were sent to GATC (Konstanz, Germany) where they were pooled, and paired 100 bp sequencing was performed on two lanes using the Illumina HiSeq2500 platform in combination with the TruSeq Rapid SBS (200 cycles) and TruSeq Rapid SR Cluster Kits (Illumina). One of the milk samples was split between two lanes, and marked as 3a and 3b.

Quality trimming was performed with PRINSEQ Lite v0.20.3 (Schmieder and Edwards, 2011) with a minimum sequence length of 30 bp and minimum quality of 30 on both ends of the read and as mean quality. Reads were mapped to the genome with Bowtie2 v2.2.1 (Langmead and Salzberg, 2012) using default settings and BAM files were converted with SAMtools v0.1.19 (Li *et al.*, 2009). BEDTools v2.17.0 was used to determine the read count for each protein coding region (Quinlan and Hall, 2010). Only reads with a minimum 30 % length overlap and mapped on the correct strand were counted. Details on the RNA-seq raw data analysis can be found in Table S1 in the supplemental material. The reads not mapping to the genome of *A. muciniphila* were confirmed to not be due to human or bacterial contamination by aligning them against the NCBI's nucleotide sequence database using MEGAblast. Differential gene expression was assessed using edgeR (Robinson *et al.*, 2010) with default trimmed mean of M-values TMM settings. Due to a potential background signal in the milk samples in a second analysis read counts were quantile normalized before running regular edgeR. Only genes that were significantly differentially expressed using both methods were considered for further analysis.

High-performance liquid chromatography. For fermentation product analysis, samples were obtained at different time points of the incubations. One ml of bacterial culture was centrifuged and the supernatant was stored at −20°C until the HPLC analysis. Substrate conversion and product formation were measured with a Thermo Scientific Spectrasystem high-performance liquid chromatography (HPLC) system equipped with a Varian Metacarb 67H 300 × 6.5 mm column kept at 45°C and running with 0.005 mM sulfuric acid as eluent. The eluent had a flow of 0.8 ml/min and metabolites were detected by determining the refractive index and identified by using standards of pure compounds as described previously (van Gelder *et al.*, 2012).

Results

Metabolic activity of *A. muciniphila* on human milk

Growth of *A. muciniphila* on human milk resulted in release of glucose, galactose, fucose and N-acetylgalactosamine (GalNAc) (Figure S1). These monomeric sugars were most likely cleaved from lactose and HMOs present in the human milk. In total, 3.8 mM of lactose was utilized, and 1.8 mM glucose, 1.9 mM galactose, 0.3 mM fucose and 0.5 mM of GalNAc were released (Figure 1C, Table 1).

Growth of *A. muciniphila* in human milk was determined in several ways. First, the *A. muciniphila* CFU counts increased 5-fold during the first 8 h of incubation. *A. muciniphila* grown on mucin, its natural substrate, was used as a control (Figure 1A). Moreover, HPLC analysis showed that a low amount of propionate, one of the main fermentation products of *A. muciniphila* grown on mucin, was already detectable after 8 h of incubation on human milk (Figure 1B, Table 2). The highest cell number, 6.4×10^7 cells/ml, was reached after 24 h. At this point both acetate and propionate were detected in the spent medium of human milk grown cells (Table 2). After 48 h incubation 45 % of bacteria were still viable in comparison to the situation after incubation for 24 h. Metabolic activity had continued, as amounts of acetate and propionate had increased, and 1,2-propanediol, an end product of fucose fermentation, could now be detected. Remarkably, after 8 days of growth on human milk viable cells were still detectable (Figure 1A).

Next, the growth kinetics of *A. muciniphila* on various milk-derived sugars (glucose, lactose, galactose, *N*-acetylneuraminic acid, 2'-fucosyllactose or 3'-sialyllactose) was tested. *N*-acetylneuraminic acid did not support growth. In all other cases weak growth with varying kinetics was observed, apart from glucose, which supported rapid growth under these conditions in the presence of high levels of protein hydrolysates, as described previously (Derrien *et al.*, 2004) (Figure 2).

Metabolic analyses indicated lactose and fucose release as a result of 2'-fucosyllactose degradation (Figure S2). In addition, very small chromatographic peaks indicated that *A. muciniphila* was able to degrade lactose into glucose and galactose after six days of incubation (Figure S3).

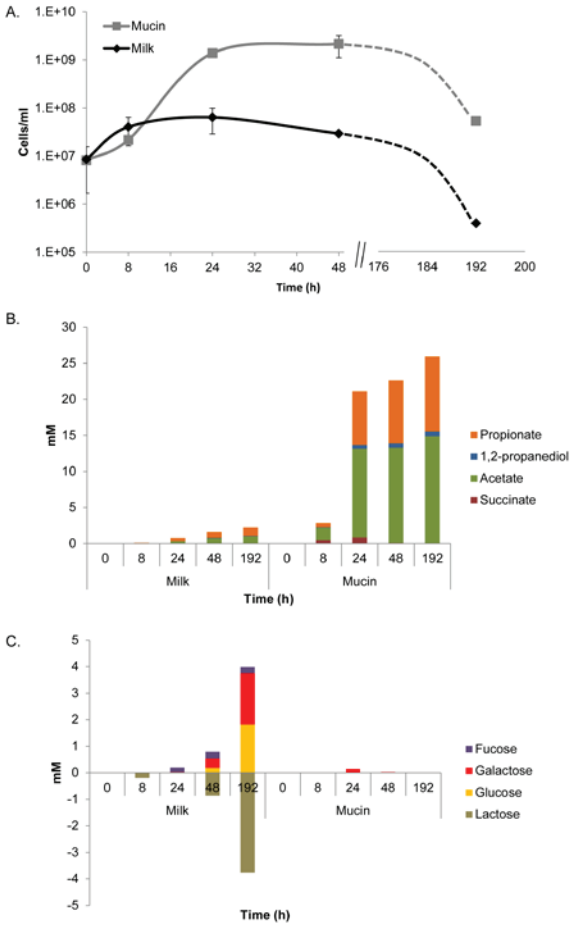


Figure 1. *A. muciniphila* growth on human milk shows a mucin-like metabolic profile. (A) *A. muciniphila* was incubated with either human milk or mucin as the sole carbon and nitrogen source. Serial dilutions were plated and CFU counts determined after 8 h, 24 h, 48 h and 192 h of incubation. Mean and standard deviations from three experiments are shown. (B) *A. muciniphila* produces acetate and propionate when grown on human milk. Fermentation products of *A. muciniphila* grown on mucin are shown as comparison. The exact amounts of fermentation products are shown in Table 2. (C) *A. muciniphila* releases galactose, glucose and fucose from human milk.

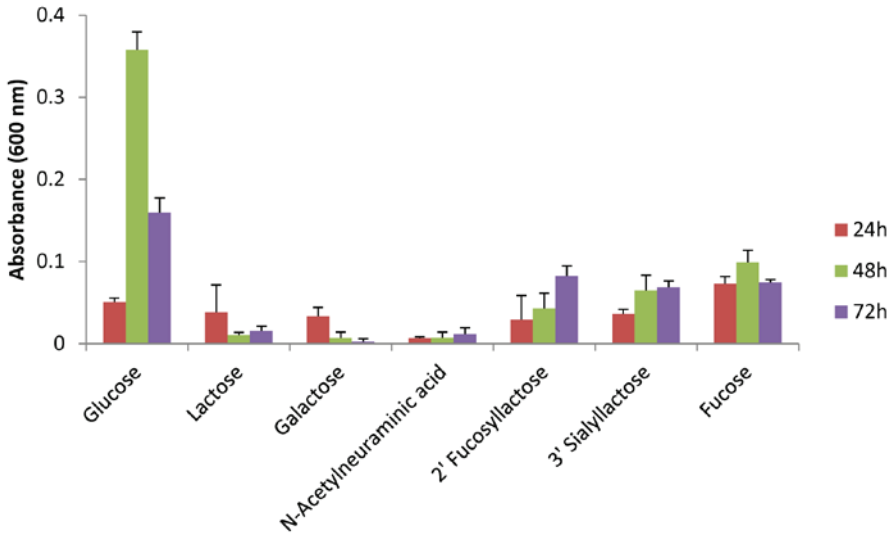


Figure 2. *A. muciniphila* shows weak growth on human milk-derived components as the sole carbon source. *A. muciniphila* was incubated on glucose, lactose, galactose, *N*-acetylneuraminic acid, 2'-fucosyllactose, 3'-sialyllactose or fucose for 72 h, and growth was monitored every 24 h by spectrophotometer as optical density at 600 nm. Mean and standard deviations from three experiments are shown.

Table 1. Sugar utilization and release of *A. muciniphila* grown on human milk and mucin

		Lactose	Glucose	Galactose	Fucose	GalNAc
Milk	0h	15.9 ± 0.3	N.D.	N.D.	N.D.	N.D.
	8h	15.7 ± 0.5	N.D.	N.D.	N.D.	0.3 ± 0.2
	24h	15.9 ± 1.0	N.D.	N.D.	0.2 ± 0.0	0.3 ± 0.2
	48h	15.0 ± 1.8	0.2 ± 0.1	0.4 ± 0.1	0.3 ± 0.1	0.3 ± 0.1
	192h	12.1 ± 0.9	1.8 ± 0.4	1.9 ± 0.4	0.2 ± 0.1	0.5 ± 0.2
Mucin	0h	N.D.	N.D.	N.D.	N.D.	N.D.
	8h	N.D.	N.D.	N.D.	N.D.	N.D.
	24h	N.D.	N.D.	0.1 ± 0.1	N.D.	N.D.
	48h	N.D.	N.D.	N.D.	N.D.	N.D.
	192h	N.D.	N.D.	N.D.	N.D.	N.D.

Mean and standard deviations from three experiments are shown. N.D., not detected.

Table 2. Fermentation products of *A. muciniphila* grown on human milk and mucin

		Acetate	Propionate	Succinate	1,2-propanediol
Milk	0h	N.D.	N.D.	N.D.	N.D.
	8h	N.D.	0.1 ± 0.2	N.D.	N.D.
	24h	0.3 ± 0.3	0.4 ± 0.1	N.D.	N.D.
	48h	0.7 ± 0.2	0.8 ± 0.2	N.D.	0.1 ± 0.1
	192h	1.0 ± 0.1	1.2 ± 0.4	N.D.	0.1 ± 0.1
Mucin	0h	N.D.	N.D.	N.D.	N.D.
	8h	1.7 ± 0.4	0.6 ± 0.2	0.5 ± 0.1	0.1 ± 0.1
	24h	12.3 ± 1.5	7.5 ± 1.8	0.8 ± 0.3	0.5 ± 0.1
	4 h	13.2 ± 0.8	8.8 ± 1.6	0.1 ± 0.1	0.6 ± 0.0
	192h	14.9 ± 0.5	10.4 ± 0.8	N.D.	0.7 ± 0.0

Mean and standard deviations from three experiments are shown. N.D., not detected.

Transcriptome analysis during growth on human milk and mucin

After demonstrating that *A. muciniphila* is capable of metabolizing human milk and the milk components lactose and 2'-fucosyllactose, we set out to discover the differences in gene expression between growth on human milk and mucin. This was done by a transcriptome analysis based on RNA-seq. Sampling of *A. muciniphila* grown on human milk was done at 48 h incubation, because according to the fermentation profile, the cells were most metabolically active at this time point, with 0.7 mM acetate and 0.8 mM propionate produced on average. For mucin cultures, sampling was done after ~17 h incubation, when the cells were in late-exponential phase. For the whole-genome transcriptome analysis, triplicate cultures of the two different conditions (i.e. 6 cultures), a total of 49,288,456 reads (with size of 100 bp) were generated by RNA-seq (overview in table S1). After quality filtering, significantly differentially regulated genes were identified by pairwise comparisons using a cut off of fold change ≥ 2 and p-value ≤ 0.05. During growth on human milk 445 genes were upregulated compared to growth on mucin, whereas 350 genes were down-regulated (Figure 3). Out of all the differentially expressed genes, 291 encoded for hypothetical proteins.

Gene ontology (GO) analysis of the gene expression profiles indicated an enrichment of six molecular functions, 16 biological processes and 11 cellular component genes during growth of *A. muciniphila* in milk as compared to mucin

(Figure 4). All but two of these (response to stimulus and proton-transporting ATP synthase complex) were upregulated in human milk cultures. RNA binding was a highly enriched molecular function during growth in human milk, whereas protein and tRNA metabolic processes, gene expression, and translation were among the enriched biological processes. The cellular process category, which describes the location in the cell where the gene product executes its function, was enriched for cytoplasm and ribosomes. On individual gene level the differential regulation of translation was evident due to the 46 ribosomal proteins and 18 tRNA metabolism genes that were upregulated in the human milk condition in comparison to mucin. The ribosomal proteins were expressed at up to 35-fold higher levels in the human milk cultures in comparison to mucin cultures.

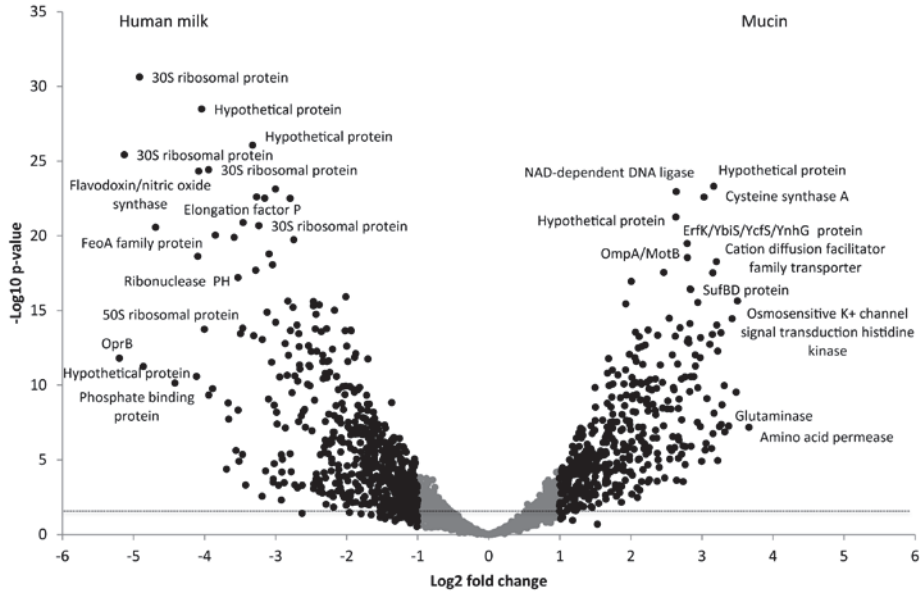


Figure 3. Volcano plot showing p-values correlated to fold changes in gene expression observed in cultures grown in the presence of human milk vs. mucin. Negative fold change indicates upregulation in human milk condition. The dashed black line shows where $p = 0.05$ with points above the line having $p < 0.05$ and points below the line having $p > 0.05$. Genes having a fold-change less than 2 are shown in gray.

The expression of 23 out of 109 (20 %) genes predicted to be involved in carbohydrate metabolism was significantly influenced by the experimental conditions. This indicates that the majority of genes (80%) for carbohydrate-degrading enzymes were expressed in similar amounts in *A. muciniphila* grown on human milk and mucin. Out of the 79 genes predicted to be involved in mucin degradation, eight were upregulated and 17 were downregulated on human milk (Table 3). The upregulated genes were predicted to encode three proteins involved in glycolysis (Amuc_0210, Amuc_0562, Amuc_0721), a beta-N-acetylhexosaminidase (Amuc_2109) and two beta-glucanase precursors (Amuc_0697, Amuc_0698). The highest fold changes were seen for genes encoding 6-phosphofructokinase (Amuc_0210, 6.4-fold higher in milk) and fructose-bisphosphate aldolase (Amuc_0721, 3.7-fold higher in milk), which are both essential enzymes in glycolysis. Among the downregulated genes were 14 that encoded different glycosyl hydrolases (Table 3). For the majority of mucin-degradation genes there were no significant differences in expression levels between the two conditions. These genes were predicted to encode nine beta-N-acetylhexosaminidases, two alpha-N-acetylglucosaminidases, two exo-alpha-sialidases and three enzymes involved in fucose transport and degradation.

The genome of *A. muciniphila* encodes for eight putative ß-galactosidases, which could be involved in the degradation of lactose. The expression of four of these ß-galactosidase genes (Amuc_0539, Amuc_0824, Amuc_1667, Amuc_1686) was not affected by the environmental conditions. However, Amuc_1811, which encodes for a probable ß-galactosidase according to KEGG orthology, was expressed 5.5-fold more during growth on human milk condition than on mucin. The remaining three ß-galactosidase genes (Amuc_0290, Amuc_0771, Amuc_1666) were expressed 2 to 3-fold higher during growth on mucin than milk.

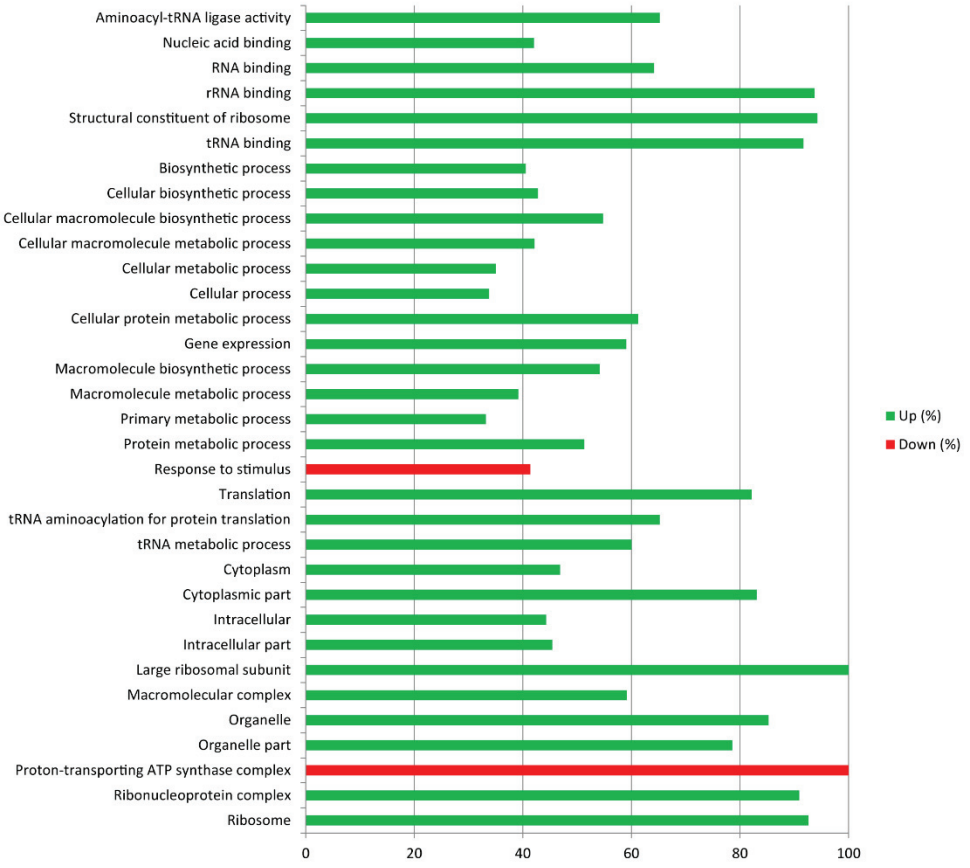


Figure 4. Gene ontology grouping of differentially regulated genes in *A. muciniphila* grown on human milk vs. mucin. Each bar represents the ratio of the number of differentially expressed genes (green, upregulated in human milk vs. mucin; red, downregulated in human milk vs. mucin) belonging to the GO terms shown in the y-axis.

Table 3. Expression of genes predicted to be involved in mucin degradation in *A. muciniphila* grown on human milk in comparison to mucin. Negative values indicate upregulation in human milk condition and positive values indicate downregulation in comparison to mucin condition. Significantly different gene expression values are color-coded in grey.

Locus tag	Annotation	log2(fold change)	p-value
Amuc_0210	6-phosphofructokinase	-2.7	0.00
Amuc_0721	fructose-bisphosphate aldolase	-1.9	0.00
Amuc_0562	Triose-phosphate isomerase	-1.6	0.00
Amuc_2109	beta-N-acetylhexosaminidase	-1.5	0.00
Amuc_1699	aminoglycoside phosphotransferase	-1.3	0.20
Amuc_1829	class II aldolase/adducin family protein	-1.2	0.00
Amuc_0697	beta-glucanase precursor	-1.2	0.00
Amuc_0698	beta-glucanase precursor	-1.0	0.01
Amuc_1822	glucosamine-6-phosphate isomerase	-1.0	0.00
Amuc_0623	glycosyl hydrolase family protein	-0.9	0.05
Amuc_1815	beta-N-acetylhexosaminidase	-0.8	0.04
Amuc_1309	aldose 1-epimerase	-0.8	0.10
Amuc_1463	hypothetical protein	-0.7	0.01
Amuc_2018	beta-N-acetylhexosaminidase	-0.5	0.23
Amuc_1830	L-fuculokinase	-0.5	0.13
Amuc_1686	beta-galactosidase	-0.4	0.14
Amuc_0029	NAD-dependent epimerase/dehydratase	-0.4	0.25
Amuc_0846	coagulation factor 5/8 type domain-containing protein	-0.3	0.57
Amuc_1245	Mannose-6-phosphate isomerase	-0.3	0.22
Amuc_1667	glycoside hydrolase family 2	-0.3	0.60
Amuc_0397	beta-N-acetylhexosaminidase	-0.2	0.68
Amuc_1220	alpha-N-acetylglucosaminidase	-0.2	0.52
Amuc_1669	beta-N-acetylhexosaminidase	-0.2	0.64
Amuc_0517	raffinose synthase	-0.2	0.73
Amuc_1008	glycoside hydrolase family protein	-0.1	0.70
Amuc_1637	alpha amylase	-0.1	0.71
Amuc_1547	hypothetical protein	-0.0	1.00
Amuc_0075	PfkB domain-containing protein	0.0	0.89
Amuc_1975	glucose-6-phosphate isomerase	0.1	0.82
Amuc_2148	beta-N-acetylhexosaminidase	0.1	0.74
Amuc_0060	alpha-N-acetylglucosaminidase	0.1	0.75
Amuc_1261	aldose 1-epimerase	0.6	0.66
Amuc_0868	beta-N-acetylhexosaminidase	0.2	0.69
Amuc_0017	oxidoreductase domain-containing protein	0.2	0.61
Amuc_0052	Hyalurononglucosaminidase	0.2	0.59
Amuc_0369	beta-N-acetylhexosaminidase	0.3	0.58
Amuc_1187	Alpha-galactosidase	0.3	0.35
Amuc_0097	ROK family protein	0.3	0.31

Locus tag	Annotation	log2(fold change)	p-value
Amuc_1671	6-phosphofructokinase	0.3	0.16
Amuc_2164	glycoside hydrolase family protein	0.5	0.08
Amuc_0729	aminoglycoside phosphotransferase	0.5	0.04
Amuc_1835	Exo-alpha-sialidase	0.6	0.03
Amuc_1757	phosphotransferase system, phosphocarrier protein HPr	0.6	0.08
Amuc_1621	4-alpha-glucanotransferase	0.7	0.03
Amuc_2136	Glycoside hydrolase, family 20, catalytic core	0.7	0.02
Amuc_1260	glycoside hydrolase family protein	0.8	0.09
Amuc_1924	beta-N-acetylhexosaminidase	0.8	0.01
Amuc_1216	oxidoreductase domain-containing protein	0.8	0.01
Amuc_2019	beta-N-acetylhexosaminidase	0.8	0.00
Amuc_0010	alpha-L-fucosidase	0.8	0.02
Amuc_1870	Alpha-glucosidase	0.8	0.00
Amuc_1481	6-phosphofructokinase	0.9	0.03
Amuc_0724	glucan endo-1,3-beta-D-glucosidase	0.9	0.00
Amuc_1125	UDP-glucose 4-epimerase	0.9	0.00
Amuc_0875	glycoside hydrolase family protein	1.1	0.18
Amuc_1120	hypothetical protein	1.1	0.00
Amuc_0771	beta-galactosidase	1.1	0.00
Amuc_0824	glycoside hydrolase family 2	1.2	0.00
Amuc_2108	glycoside hydrolase family protein	1.2	0.00
Amuc_1032	beta-N-acetylhexosaminidase	1.2	0.00
Amuc_0216	hypothetical protein	1.2	0.00
Amuc_0920	oxidoreductase domain-containing protein	1.2	0.00
Amuc_1832	L-fucose isomerase	1.2	0.42
Amuc_0290	glycoside hydrolase family 2	1.3	0.00
Amuc_0948	N-acetylglucosamine-6-phosphate deacetylase	1.3	0.00
Amuc_0392	coagulation factor 5/8 type domain-containing protein	1.3	0.62
Amuc_0539	glycoside hydrolase family 2	1.4	0.49
Amuc_0863	glycosyl hydrolase family protein	1.5	0.84
Amuc_1666	glycoside hydrolase family 2	1.5	0.01
Amuc_1833	L-fucose transporter	1.6	0.06
Amuc_1868	glycoside hydrolase family protein	1.7	0.00
Amuc_0855	glycoside hydrolase clan GH-D	1.7	0.00
Amuc_1420	hypothetical protein	1.7	0.00
Amuc_0186	glycoside hydrolase family protein	1.7	0.00
Amuc_0969	galactokinase	2.0	0.00
Amuc_0146	alpha-L-fucosidase	2.0	0.00
Amuc_2100	aminoglycoside phosphotransferase	2.4	0.00
Amuc_0625	Exo-alpha-sialidase	2.4	0.59
Amuc_1812	alpha-amylase	3.1	0.37

The highest fold change in the human milk samples in comparison to mucin was observed for the gene encoding carbohydrate-selective porin OprB (Amuc_1085), which was expressed 36-fold higher on milk than on mucin. In the genus *Pseudomonas* OprB has been described as an outer membrane channel for monosaccharide uptake, with a preference for glucose (van den Berg, 2012). The expression of *Pseudomona putida* OprB is transcriptionally induced by glucose, and repressed when the bacteria are grown on other substrates (Shrivastava *et al.*, 2011). It should be noted that during growth on milk approximately 1.8 mM of glucose was released from lactose due to the ß-galactosidase activity of *A. muciniphila*, and that glucose is not present in our mucin condition.

On the mucin grown *A. muciniphila* the highest fold change compared to milk grown cells was found for a gene encoding an amino acid permease (Amuc_0037, 13-fold upregulation). This indicates increased transport of amino acids into the cell. In addition to this a glutaminase-encoding gene (Amuc_0038) was upregulated 10-fold and a gene that codes for glutamate decarboxylase (Amuc_0372) was upregulated 6-fold in the presence of mucin in comparison to human milk.

Other genes with high fold changes in the milk condition in comparison to mucin were for example those encoding flavodoxin/nitric oxide synthase (Amuc_1899, 17-fold), ribonuclease PH (Amuc_0395, 11-fold) and elongation factor P (Amuc_0874, 10-fold) (Figure 3). Flavodoxins are electron-transfer proteins that function in various electron transport systems, and nitric oxide synthases produce nitric oxide from L-arginine and NADPH (Wang *et al.*, 2007). Ribonuclease PH and elongation factor P are involved in translation and tRNA processing (Doerfel and Rodnina, 2013; Wen *et al.*, 2005).

A gene cluster involved in phosphate uptake and transport was identified in the genome of *A. muciniphila* (Amuc_1302 to Amuc_1308). The cluster contains the genes coding for the four-component Pst transporter, which is used to sense phosphate, as well as two genes encoding for PhoU, which is essential for the repression of the Pho regulon at high phosphate conditions (Bergwitz and Juppner, 2011). Five genes in this cluster were significantly upregulated during growth on human milk, and the remaining two genes were also upregulated, albeit not significantly (Figure 5A). In bacteria the phosphate system is active by default and only excess phosphate turns the system off. Also a gene cluster involved in iron transport was upregulated in milk grown bacteria (Figure 5B). These genes are a part of the Feo system, which is the major bacterial route for acquiring the ferrous (Fe^{2+}) form of iron from the environment, and has an import role in iron uptake under the anaerobic-microaerophilic conditions of the gastrointestinal tract (Cartron *et al.*, 2006). The Feo system is generally composed of three major units:

FeoA, FeoB and FeoC. However, FeoC is not found in the genome of *A. muciniphila*, and two genes (Amuc_1089, Amuc_1090) are annotated as FeoA.

Interestingly, the genes encoding two cytochromes (cytochrome d ubiquinol oxidase, subunit II; Amuc_1694 and cytochrome bd ubiquinol oxidase subunit I; Amuc_1695) were significantly upregulated in the mucin grown *A. muciniphila* in comparison to human milk grown bacteria. This could indicate increased respiration in the mucin environment, possibly reflecting the presence of oxygen in the mucus layer close to the epithelial cells (Ouwerkerk *et al.*, 2013).

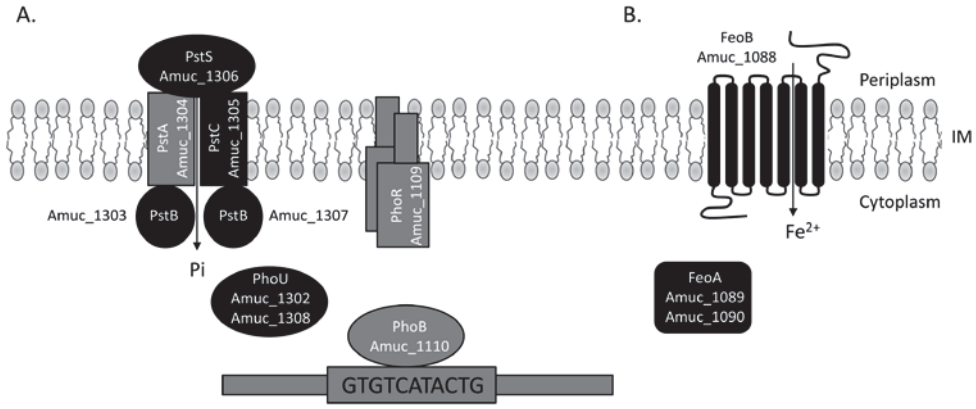


Figure 5. Phosphate and iron uptake are upregulated in *A. muciniphila* grown on human milk in comparison to mucin. Genes encoding for the proteins marked in black were upregulated in *A. muciniphila* grown on human milk in comparison to mucin. (A) Phosphate (Pi) is sensed by binding to the four-component Pst transporter (PstS, PstA, PstB, PstC), which activates a two component signaling system composed of the sensory histidine kinase PhoR and the transcription factor, PhoB that is known to bind DNA and modulate gene expression. The conformational states of PstSCAB are sensed by PhoR through PhoU as a consequence of Pi transport, leading to modulation of the kinase/phosphatase equilibrium toward the appropriate response (Gardner *et al.*, 2014). Under high phosphate conditions the sensing system is inhibited. (B) The Feo system is involved in transport of ferrous iron (Fe^{2+}) from the environment. FeoA is a small, soluble SH3-domain located in the cytosol, and FeoB is a large protein with a C-terminal integral inner membrane domain which likely functions as the Fe^{2+} permease (Cartron *et al.*, 2006). IM; inner membrane.

Genes involved in cobalamin (vitamin B12) modification were upregulated when *A. muciniphila* was grown on human milk. This was evident based on the upregulation of four genes from the cobalamin biosynthesis pathway (cobyrinic acid ac-diamide synthase; Amuc_0079 and Amuc_2037, cobalbumin biosynthesis protein; Amuc_1678, cobalamin 5'-phosphate synthase; Amuc_1680). In addition to this, a gene involved in cobalt transport (cobalt ABC transporter ATPase; Amuc_1198) was also upregulated during growth on milk. The growth medium for both mucin and human milk cultures was supplemented with the same amount of cyanocobalamin. Human milk contains vitamin B12 tightly bound to haptocorrin, and the concentration varies depending on the diet of the mother, especially intake of animal source foods (Allen, 2012). Increased succinate to propionate ratio has been observed in *A. muciniphila* in the absence of B12 in comparison to when B12 was provided (Belzer et al., unpublished results). No succinate was detected during growth on human milk, and propionate was produced, suggesting B12 was not limiting for *A. muciniphila*.

Discussion

We demonstrated that *A. muciniphila* was able to grow in human breast milk based on an increase in cell numbers, metabolic end products and changes in global gene expression. The magnitude of growth was still limited in comparison to growth on mucin, which further confirmed the role of *A. muciniphila* as a mucin specialist. The expression of the majority of genes involved in mucin degradation was not affected by the experimental condition. However, the global transcriptome analysis showed upregulation of ß-galactosidase as well as carbohydrate-selective porin OprB and proteins involved in translation and protein synthesis during growth on human milk in comparison to mucin.

A. muciniphila was not able to grow on lactose as the sole carbon source, but a significant amount of lactose (3.8 mM) was utilized during growth on human milk. Only 1.8 mM glucose and 1.9 mM galactose were detected in the supernatant, and hence it is probable that *A. muciniphila* used intact lactose, explaining the upregulation of the intracellular ß-galactosidase and glycolytic genes. The inability of *A. muciniphila* to grow on lactose alone may be due to the absence of growth-stimulating factors in the synthetic medium supplemented with lactose, or presence of stimulating factors in the breast milk. We identified eight putative ß-galactosidase genes in the genome of *A. muciniphila*. From these, the expression of Amuc_1811 was affected most by the change in environmental conditions, with five times higher levels detected during growth on human milk in comparison to mucin. This predicted enzyme is the only one of the eight ß-galactosidases in the genome of *A. muciniphila* that does not have a signal peptide sequence. This suggests that

either lactose is transported into the cell or that the enzyme is released to the environment from lysed cells. The other ß-galactosidases might have other, secretory functions, possibly involved in mucin utilization, notably as their specificity cannot be accurately predicted from the protein sequence.

During growth on 2'-fucosyllactose fucose and lactose were released into the spent medium, indicating that *A. muciniphila* breaks down the α2 bond between galactose and fucose. Besides 2'-fucosyllactose, the same bond exists in many of the termination structures of the mucin molecule (**Chapter 1**). Therefore, it is expected that *A. muciniphila* has the enzymatic capacity to cleave fucose from these structures, and use it as an energy source. Utilization of fucose during growth on human milk was confirmed by the presence of 1,2-propanediol in our analysis. Further assays with *A. muciniphila* α-L-fucosidases are needed to demonstrate the enzymatic activity in both mucin and HMOs. The genome of *A. muciniphila* encodes for five different α-L-fucosidases, with predicted end locations in both the periplasm and extracellular space, suggesting a variety of possible functions for these enzymes.

The single gene most affected by the change in environmental conditions was that coding for the carbohydrate-selective porin OprB (Amuc_1085). It was upregulated in the milk grown cells in comparison to the mucin medium, which does not contain glucose. This gene has been shown to be induced by glucose in other bacteria (Shrivastava *et al.*, 2011; van den Berg, 2012). We cannot conclude that this is the case in *A. muciniphila* as comparison of gene expression between *A. muciniphila* grown on mucin and glucose did not show upregulation of OprB in the glucose condition (**Chapter 3**). Another possibility could be induction by galactose or lactose. More galactose is expected to be available in human milk as a result of lactose degradation, than in mucin. Porins allow the diffusion of relatively large molecules through the outer membrane of Gram-negative bacteria, suggesting that small soluble milk glycans, such as 2'-fucosyllactose could also be transported through the porins to the periplasm, where fucosidases can cleave the terminal fucose. The genome of *A. muciniphila* also encodes for another OprB (Amuc_2127), which was expressed 6-fold higher in the mucin condition in comparison to human milk. Further research is needed to verify the role of these porins in *A. muciniphila*.

In *A. muciniphila* grown on mucin, amino acid transport and glutamine/glutamate metabolism were upregulated in comparison to the milk condition. According to KEGG orthology the upregulated amino acid permease (Amuc_0037) is a probable glutamate/γ-aminobutyric acid (GABA) antiporter. The upregulation of these genes may be related to acid resistance. The well described bacterial acid resistance

system 2 (AR2) comprises an amino acid antiporter, which exchanges extracellular L-glutamate with intracellular GABA, and two L-glutamate decarboxylases, which convert L-glutamate to GABA (Richard and Foster, 2004). However, recently another acid resistance system, relying on L-glutamine, was described in *Escherichia coli* (Lu *et al.*, 2013). In this system, L-glutamine is transported into the cell and converted to L-glutamate by a glutaminase. This leads to release of gaseous ammonia, which neutralizes a proton, resulting in elevated intracellular pH in acidic environments. This is more likely to be the system *A. muciniphila* utilizes as in addition to the amino acid permease and glutamate decarboxylase (Amuc_0372), also a glutaminase (Amuc_0038) was upregulated. Furthermore, the glutamate/GABA antiporter is also able to transport L-glutamine (Ma *et al.*, 2012), which is present in mucin (Khan *et al.*, 1999). During growth on mucin, *A. muciniphila* produces high amounts of acetate and propionate, which lowers the extracellular pH and most probably leads to the activation of the described acid resistance system.

The significant upregulation of ribosomal proteins and genes involved in tRNA metabolism during *A. muciniphila* growth on human milk was in stark contrast with the higher growth rate and biomass yield of the bacteria grown on mucin. The bacterial culture grown on mucin was sampled at 17 h of incubation, at which point *A. muciniphila* is approaching the stationary phase. This could lead to repression of the translation machinery, including ribosomal proteins. Ribosomal proteins and aminoacyl-tRNA synthesis have been shown to be upregulated during lag-phase in comparison to late-exponential and stationary-phase bacteria (Rolfe *et al.*, 2012). Upregulation of proteins involved in translation was also observed in lag-phase *Lactococcus lactis* grown on milk in comparison to exponential and stationary phase bacteria (Larsen *et al.*, 2006). Furthermore, in a recent report, *L. lactis* showed very little regulation of gene expression when shifting from low growth rate to high growth rate, which was exemplified by an excess of ribosomal proteins at low growth rates (Goel *et al.*, 2015). The authors suggested that the organism preserves a large overcapacity of ribosomes to be ready to quickly respond and grow when conditions improve. This hypothesis would also fit *A. muciniphila* as during colonization the bacteria would simply have to stay viable on human milk before reaching the intestinal mucus layer where optimal nutrients are abundantly available.

Previously, lag-phase conditions have shown to be marked by higher phosphate and iron uptake in comparison to other growth stages in *Salmonella* spp., presumably preparing these bacteria for different environmental conditions (Rolfe *et al.*, 2012). We observed upregulation of phosphate and iron uptake and transport in *A. muciniphila* during growth on human milk, further suggesting the bacteria were in

the lag-phase. However, this could also be an indication of iron and phosphate limitation in milk as these compounds are tightly linked to casein and/or complexed with citrate, which would lead to enhanced expression of uptake systems by *A. muciniphila*.

During the first weeks and months of life, when the infant's diet is almost exclusively milk, the gut microbiota is dominated by bifidobacteria, with members of *Enterococcaceae*, *Streptococcaceae*, *Lactobacillaceae*, *Clostridiaceae* also present (Matamoros *et al.*, 2013). In this bacterial community, other members may cleave off sugars from HMOs for *A. muciniphila* to use. It is also well possible that some of the wide range of mucin-degrading enzymes that are produced by *A. muciniphila* can be used for degradation of HMOs due to the similar structures of these glycans. In either case, being able to feed off human milk would give *A. muciniphila* a selective advantage during early life colonization. We also demonstrated *A. muciniphila* viability on human milk after an extended incubation period. It is plausible that human milk can act as a carrier for bacteria from the environment, protecting them from the harsh conditions of the upper GI tract before reaching the small and large intestine. For *A. muciniphila* this might be a mechanism for transferring from the mother to the infant. We also found evidence for an acid resistance system, which may further assist *A. muciniphila* in surviving passage through the GI tract.

It has been hypothesized that the reason HMO structures resemble mucin, is to attract mucin-adapted resident mutualists to an infant microbiota. In *B. thetaiotaomicron* the same polysaccharide utilization loci (PULs) associated with mucin degradation were also induced when the bacteria were grown on HMOs (Marcobal *et al.*, 2011). However, the *B. thetaiotaomicron* mutants created in four of the HMO-specific loci showed no defect in growth on HMOs, suggesting more evidence is needed to prove the association between mucin and HMO utilization genes. Despite the related structures of HMOs and mucin, some bacterial species are capable of using only one of these host glycans, indicating host selection that favors the colonization of specific species over others. For example *Bifidobacterium infantis* is adapted to use the simple structures within HMOs, and is not able to use the complex glycan structures found in mucin (Marcobal *et al.*, 2011). Succession of the mucosa-associated microbiota may be involved in the development of the immune system due to its close proximity to the epithelial layer (Brugman *et al.*, 2015; Duerr and Hornef, 2012). Consequently, colonization of the mucosal microbiota with beneficial microbes could be associated with health in later life. Colonization by *A. muciniphila* in early life would be favorable for the host as this bacterium is associated with a healthy gut.

Several studies have investigated the capacity of bifidobacteria, lactobacilli and members of the *Bacteroidetes* to utilize human milk, but this is the first report on *A. muciniphila*, a member of the *Verrucomicrobia*. We observed metabolic activity and upregulation of gene expression involved in lactose degradation, carbohydrate transport and translational activity in *A. muciniphila* grown on human milk in comparison to mucin. The milk environment also activated expression of the same genes that are needed for mucin degradation. These results suggest that in the early life intestine *A. muciniphila* is active and could benefit from this during initial colonization before settling in the intestinal mucus layer.

Supplemental material

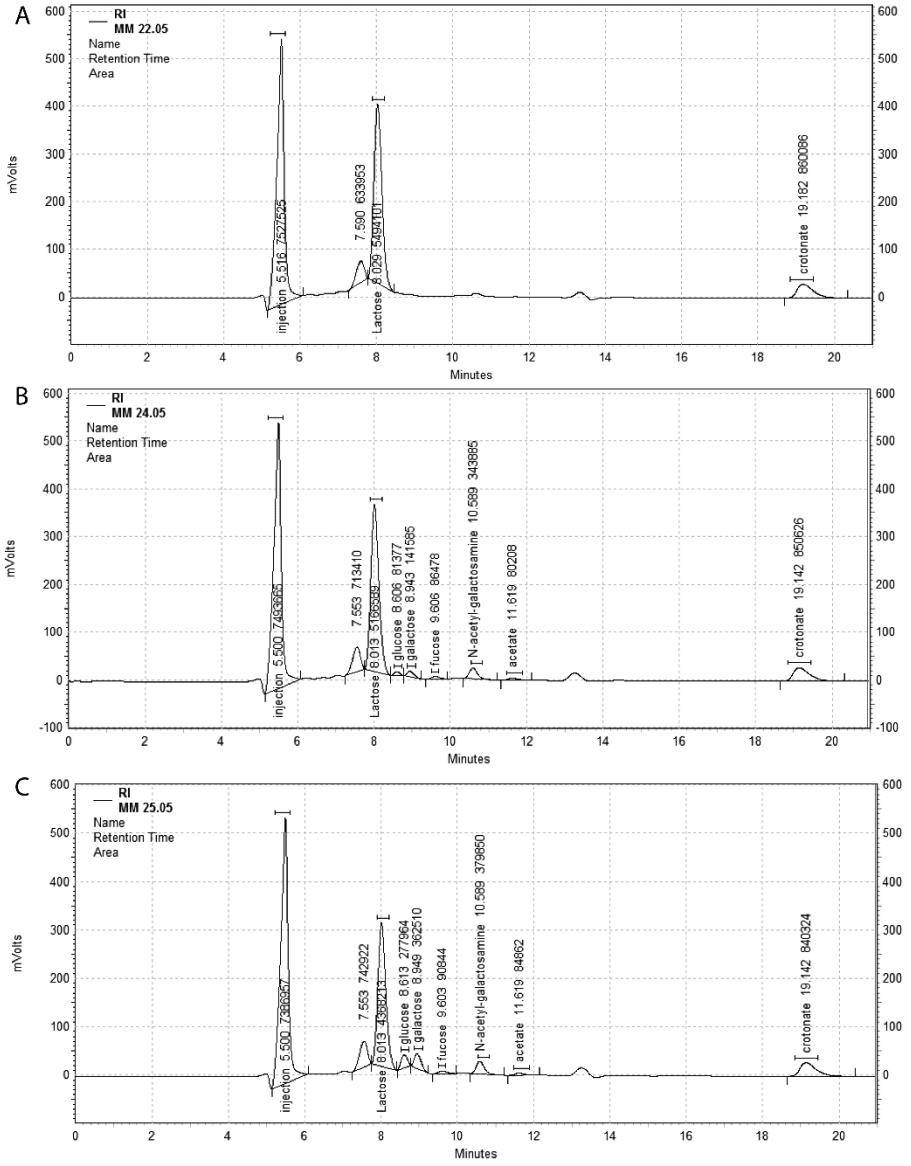


Figure S1. HPLC profile for *A. muciniphila* grown on human milk. (A) At time point 0 h lactose is detected in the supernatant. (B) At time point 48 h the size of the lactose peak has decreased and glucose, galactose, fucose and N-acetylgalactosamine have been released into the supernatant. A small peak for acetate is detected. (C) At time point 72 h lactose degradation has continued and increasing amounts of monomeric sugar have been released. Crotonate was added to the samples as an internal standard.

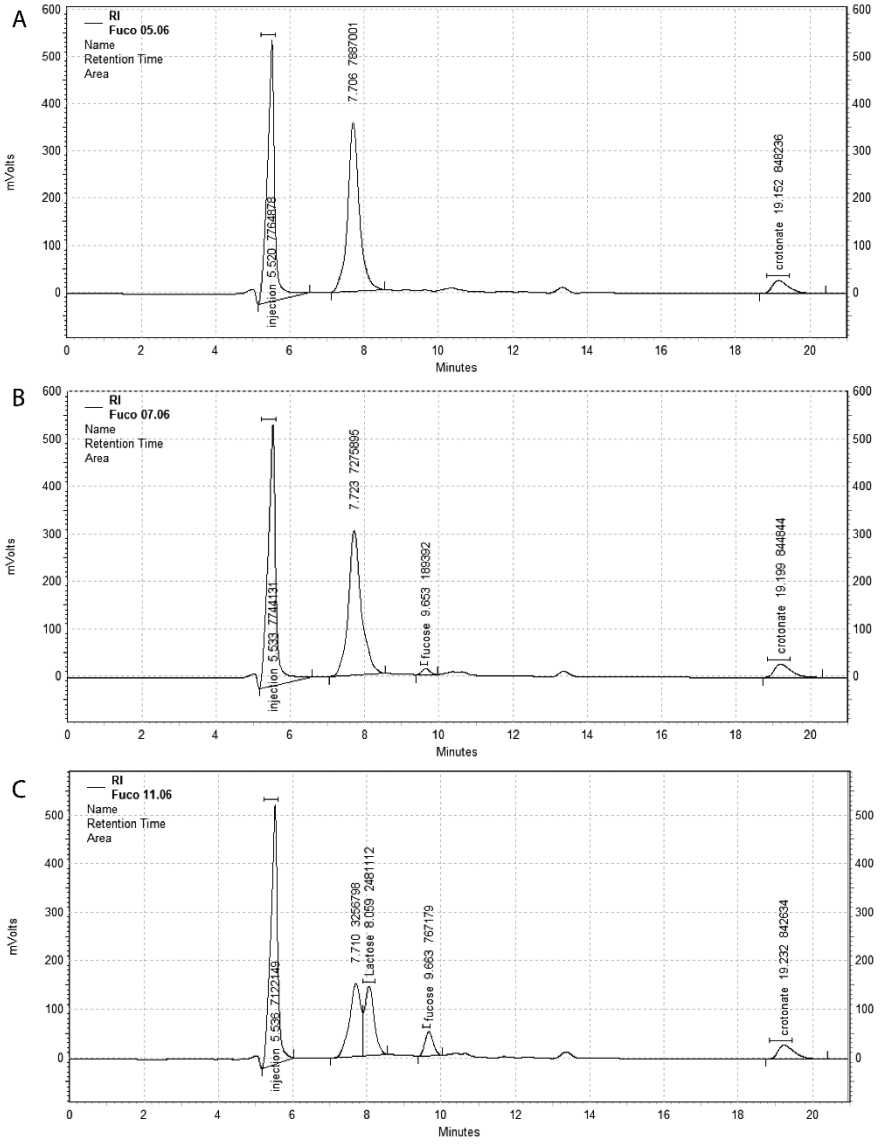


Figure S2. HPLC profile for *A. muciniphila* grown on 2'-fucosyllactose. (A) At time point 0 h an unknown peak at 7.7 min was detected. This is most likely 2'-fucosyllactose. (B) At time point 48 h fucose was released into the supernatant. (C) At time point 144 h (6 days) the size of the peak for 2'-fucosyllactose has decreased and lactose and fucose have been released into the supernatant. Crotonate was added to the samples as an internal standard.

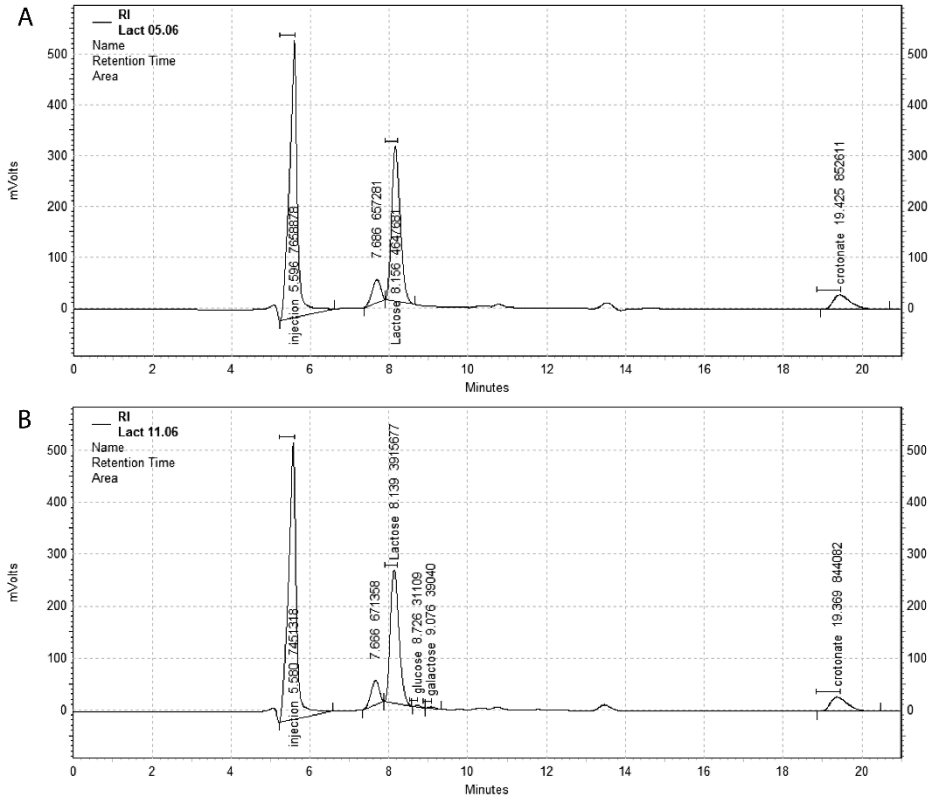
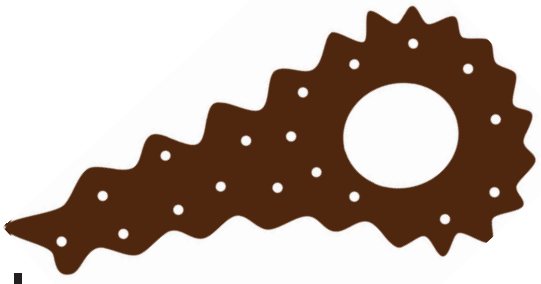


Figure S3. HPLC profile for *A. muciniphila* grown on lactose. (A) At time point 0 h lactose is detected in the supernatant. (B) At time point 144 h (6 days) the size of the lactose peak has decreased slightly and small peaks for glucose and galactose are detected. Crotonate was added to the samples as an internal standard.

Table S1. Summary of the RNA-seq raw data analysis for transcriptome analysis of *A. muciniphila* grown on human milk and mucin.

	Milk				Mucin		
	1	2	3a	3b	1	2	3
Total no. of reads	2660890	5512805	6794460	5810961	11298914	11587377	5623049
Total no. of reads after quality filtering	2549997	5294294	6505478	5568130	9227135	10620856	5272525
Percentage of reads passing quality filtering	95.8	96.0	95.7	95.8	81.7	91.7	93.8
Total no. of reads mapped to the genome	1183404	1164455	2114105	2017832	7614634	9500529	4697125
Percentage of reads mapping to the genome	46.4	22.0	32.5	36.2	82.5	89.5	89.1
Total no. of reads mapped to protein coding regions	581666	519060	974248	939825	2213705	4692956	2680434

Chapter 5



Biochemical and immunological characterization of the *Akkermansia muciniphila* outer membrane proteome

Noora Ottman, Sjef Boeren, Judith Klievink, Hauke Smidt,
Clara Belzer, and Willem M. de Vos

Abstract

Akkermansia muciniphila is a common member of the human gut microbiota and belongs to the *Planctomycetes-Verrucomicrobia-Chlamydiae* superphylum. Decreased levels of *A. muciniphila* have been associated with many diseases, and thus it is considered to be a beneficial resident of the intestinal mucus layer. Surface-exposed molecules produced by this organism likely play important roles in colonization and communication with other microbes and the host, but the protein composition of the outer membrane (OM) has not been characterized thus far. We used sarkosyl extraction and sucrose density-gradient centrifugation methods to enrich the OM proteome of *A. muciniphila*. Proteins from these fractions were identified by LC-MS and compared to intracellular and whole proteome fractions. These experimental data were combined with bioinformatics analysis for OM protein prediction, and in total we identified 79 putative OM and membrane-associated extracellular proteins. These included highly abundant proteins involved in secretion, transport and biogenesis of Gram-negative membranes, as well as proteins predicted to take part in formation of the fimbriae-like structures observed in *A. muciniphila* by electron microscopy. In addition, composition of the OM was shown to differ between *A. muciniphila* grown on its natural substrate, mucin, and the non-mucus sugar, glucose. The activation of Toll-like receptor 2 was higher for OM proteins from mucin-grown cells in comparison to intracellular proteins, indicating the presence of immunostimulatory compounds in this fraction. In conclusion, the characterization of *A. muciniphila* OM proteome provides valuable information that can be used for further functional and immunological studies.

Introduction

Akkermansia muciniphila is a Gram-negative, anaerobic bacterium, which colonizes the mucus layer of the human gastrointestinal (GI) tract (Derrien *et al.*, 2004). *A. muciniphila* is considered to be an important member of the GI microbiota, because of the inverse correlation between its abundance and several intestinal disorders, including inflammatory bowel diseases and obesity (Karlsson *et al.*, 2012; Png *et al.*, 2010; Rajilic-Stojanovic *et al.*, 2013). Moreover, experiments with germ-free mice mono-associated with *A. muciniphila*, or conventional mice on a high fat diet that are fed *A. muciniphila*, have shown that *A. muciniphila* plays a role in host immune response, restoration of mucus layer thickness and mucus production (Derrien *et al.*, 2011; Everard *et al.*, 2013; Shin *et al.*, 2013). In addition, extracellular vesicles from *A. muciniphila* were shown to have protective effects on the development of dextran sulfate sodium (DSS) induced colitis in mice (Kang *et al.*, 2013). Finally, *A. muciniphila* has also been shown to adhere to intestinal epithelium and improve enterocyte monolayer integrity of Caco-2 cells (Reunanen *et al.*, 2015). These findings suggest important host-bacteria interactions, the mechanisms of which are yet to be discovered.

Bacterial outer membrane (OM) proteins play important roles in communication with other microbes and the host, as well as in colonization and substrate transport (Galdiero *et al.*, 2012; Tseng *et al.*, 2009). Subcellular fractionation techniques combined with mass spectrometry-based proteomic analysis are powerful tools for identifying proteins in different bacterial compartments. These techniques have been successfully used for studying the protein composition of gut bacteria, for example the outer membrane of *Bacteroides fragilis* and *Bacteroides thetaiotaomicron* (Elhenawy *et al.*, 2014; Wilson *et al.*, 2015), surface proteins of *Propionibacterium freudenreichii* (Le Marechal *et al.*, 2015) and outer membrane vesicles of *Escherichia coli* Nissle 1917 (Aguilera *et al.*, 2014).

A. muciniphila is a member of the *Planctomycetes-Verrucomicrobia-Chlamydiae* (PVC) superphylum, which contains bacteria from several groups and various environments with different lifestyles (Gupta *et al.*, 2012; Kamneva *et al.*, 2012). Bacteria from this superphylum were previously suggested to have a compartmentalized cell plan with a cytoplasmic membrane as the outermost membrane, and an intracytoplasmic membrane containing a condensed nucleoid and ribosomes (Lee *et al.*, 2009). However, these observations have been challenged by more recent data suggesting that the PVC cell plan is actually a variation, not an exception, of the Gram-negative cell plan, and that the bacteria have an outer and an inner membrane (IM) with possible invaginations of the IM inside the cytoplasm (Devos, 2014).

There is limited information available on the membrane structure and composition of *A. muciniphila,* and most reports have focused on *in silico* analysis of *Verrucomicrobia* membranes, instead of experimental approaches (Kamneva *et al.*, 2012; Santarella-Mellwig *et al.*, 2010; Speth *et al.*, 2012). Recently, the proteome of a termite hindgut representative of the *Verrucomicrobia*, *Diplosphaera colotermitum* TAV2, was experimentally studied, but this report did not focus on membrane proteins (Isanapong *et al.*, 2013). The presence of OM biomarkers, including genes involved in lipopolysaccharide (LPS) insertion, in the genome of *A. muciniphila* was confirmed computationally (Speth *et al.*, 2012). We have experimentally verified the presence of LPS in *A. muciniphila* by using a protocol to extract LPS form bacterial cells and applying silver staining to visualize it on a gel (**Chapter 6**). No genes coding for membrane coat-like proteins were found in *A. muciniphila*, unlike in some other *Verrucomicrobia* (Santarella-Mellwig *et al.*, 2010).

Herein we set out to identify and characterize *A. muciniphila* proteins using an integrated approach of proteomics and computational analysis. Successful extraction of OM proteins was established, and the proteins were identified with liquid chromatography-tandem mass spectrometry (LC-MS/MS). The abundance of *A. muciniphila* proteins in the OM fraction was compared to the whole proteome of *A. muciniphila* and a fraction enriched for intracellular proteins. Candidates for OM proteins derived from the proteomics analysis were subjected to computational screening to verify their location in the cell. We also explored if the presence of these proteins was dependent on environmental availability of mucus, resulting in the identification of several OM proteins related to mucosal colonization. Finally, the proteome fractions were analyzed for their capacity to induce immune responses in a cell line model. The results indicate that *A. muciniphila* produces OM proteins involved in secretion and transport in high abundance, and these may be involved in its immunostimulatory capacities.

Materials & methods

Bacterial growth conditions. *Akkermansia muciniphila* Muc^T (ATTC BAA-835) was grown in a basal medium as described previously, except without the addition of rumen fluid (Derrien *et al.*, 2004). The medium was supplemented with either hog gastric mucin (0.5 %, Type III; Sigma-Aldrich, St. Louis, MO, USA) purified by ethanol precipitation as described previously (Miller and Hoskins, 1981), or D-glucose (10 mM, Sigma-Aldrich). The medium with glucose was also supplemented with Bacto™ casitone (BD, Sparks, MD, USA), BBL™ yeast extract (BD), tryptone (Oxoid Ltd, Basingstoke, Hampshire, England), peptone (Oxoid ltd) (2 g/l each) and L-threonine (Sigma-Aldrich) (2 mM). Incubations were performed in serum bottles

sealed with butyl-rubber stoppers at 37ºC under anaerobic conditions provided by a gas phase of 182 kPa (1.5 atm) N_2/CO_2. Growth was measured by following optical density at 600 nm (OD600) using a spectrophotometer.

Bacterial fractionation methods. The membranes of *A. muciniphila* were isolated from liquid cultures with two different methods, using either *N*-lauroylsarcosine (sarkosyl) or sucrose density-gradient centrifugation, as described previously, with minor modifications (Hobb *et al.*, 2009).

Briefly, for sarkosyl treatments, 250 ml cultures of *A. muciniphila* were grown on mucin for 16 h (OD_{600} = 1.56) or on glucose for 40 h (OD_{600} = 0.34). Cells were harvested, washed twice with phosphate buffered saline (PBS) and resuspended in 9 ml 10 mM HEPES, pH 7.4, and lysed by passing the culture three times through a French press (Aminco, American Instrument Co., Inc., Maryland, USA) at 1000 psi (40K cell). The lysed cell preparation was centrifuged at 10 000 g for 10 min at 4°C to remove cell debris and unlysed cells. The membranes were collected by ultracentrifugation of the supernatant at 100 000 g for 1 h at 4°C. The supernatant was collected and stored at −20°C. This sample was later analyzed by LC-MS/MS as the intracellular fraction. The pellet was resuspended in 2 ml 10 mM HEPES, pH 7.4, washed in a total volume of 10 ml 10 mM HEPES, pH 7.4, and spun again in the ultracentrifuge (using the conditions described above). The pellet was resuspended in 5 ml 1 % (w/v) N-lauroylsarcosine (sarkosyl) (Sigma-Aldrich) in 10 mM HEPES, pH 7.4, and incubated at 37°C for 30 min with shaking to solubilize cytoplasmic membranes. The sarkosyl-treated membranes were spun at 100 000 g for 1 h at 4°C and the pellet was washed with 7 ml 10 mM HEPES, pH 7.4. Following the final ultracentrifugation, the pellet containing the OM fraction was resuspended in 1 ml 10 mM HEPES, pH 7.4 and stored at −20°C.

For the sucrose-density gradient centrifugation treatments, 250 ml cultures of *A. muciniphila* were grown on mucin for 16 h (OD600 = 1.51) or on glucose for 40 h (OD600 = 0.35). Cells were harvested, washed twice with phosphate buffered saline (PBS) and resuspended in 7 ml 10 mM HEPES, pH 7.4, and lysed by passing the culture three times through a French press at 1000 psi (40K cell). The lysed cells preparation was centrifuged at 10 000 g for 10 min at 4°C. The supernatant was ultracentrifuged at 100 000 g for 60 min at 4°C to pellet the total membranes. The membrane pellet was washed in 10 ml 10 mM HEPES, 0.05 M EDTA pH 7.5 (HE buffer) and ultracentrifuged again. The final membranes were homogenized in 2 ml HE buffer. Continuous sucrose gradients were prepared by layering sucrose solutions (prepared in HE buffer) into 14 × 95 mm polyallomer ultracentrifuge tubes (Seton Scientific, Petaluma, CA, USA) in the following order: 0.4 ml 60% (w/v), 0.9 ml 55 %, 2.2 ml 50 %, 2.2 ml 45 %, 2.2 ml 40 %, 1.3 ml 35% and 0.4 ml 30

%. Total membranes were layered on top of each gradient, with no more than 2.5 ml per gradient. Sucrose gradients were centrifuged in a TST 41.14-41000 RPM swinging-bucket rotor (Kontron) at 250 000 g for 16 h at 4°C. The sucrose gradient tubes were then removed from the rotor buckets and 500 µl fractions (24 fractions for each gradient) were collected from the bottom of each gradient by puncturing the tube with a needle and allowing the sample to drip out by gravity. The samples were stored in 2 ml low binding tubes (Eppendorf, Hamburg, Germany) at −20°C.

Protein identification by mass spectrometry. To determine the protein content of cell extracts, the Qubit® Protein Assay Kit (Life technologies, Oregon, USA) was used according to the manufacturer's instructions. Samples were loaded on a 10 % acrylamide separation gel (25201, Precise™ Protein Gels, Thermo Scientific, Rockford, IL, USA) using the mini-PROTEAN 3 cell (Bio-Rad Laboratories, Hercules, CA, USA). The electrophoresis procedure was according to the manufacturer's instructions. Gels were stained using Coomassie Brilliant Blue (CBB) R-250 as indicated in the protocol of the mini-PROTEAN 3 system.

In-gel digestion of proteins and purification of peptides were done following a modified version of the protocol described by Rupakula et al., 2013. Disulfide bridges in proteins were reduced by covering whole gels with reducing solution (10 mM dithiothreitol, pH 7.6, in 50 mM NH_4HCO_3), and the gels were incubated at 60°C for 1 h. Alkylation was performed for 1 h by adding 25 ml of iodoacetamide solution (10 mM iodoacetamide in 100 mM Tris-HCl, pH 8.0). Gels were thoroughly rinsed with demineralized water in between steps. Each of the used gel lanes was cut into either five slices (sarkosyl-extracted outer membrane and intracellular proteins) or one slice (sucrose density-gradient centrifugation samples), and the slices were cut into approximately 1 mm^3 cubes and transferred to separate 0.5 ml protein LoBind tubes (Eppendorf, Hamburg, Germany). Enzymatic digestion was done by adding 50 µl of trypsin solution (5 ng/µl trypsin in 50 mM NH_4HCO_3) to each tube, and by incubating at room temperature overnight with gentle shaking. Extraction of peptides was performed with manual sonication in an ultrasonic water bath for 1 s before the supernatant was transferred to a clean protein LoBind tube. Trifluoroacetic acid (10 %) was added to the supernatant to reach a pH between 2 and 4. The supernatant was used for LC-MS/MS analysis. Samples were measured by nLC–MS/MS with a Proxeon EASY nLC and a LTQ-Orbitrap XL mass spectrometer as previously described (Lu *et al.*, 2011).

LC–MS data analysis was performed as described previously (Rupakula *et al.*, 2013) with false discovery rates (FDRs) set to 0.01 on peptide and protein level, and additional result filtering (minimally 2 peptides necessary for protein identification of which at least one is unique and at least one is unmodified). To analyze the

abundance of proteins in the fractions, their label-free quantification (LFQ) intensities were compared (Cox *et al.*, 2014). Non-existing LFQ intensity values due to not enough quantified peptides were substituted with a value lower than the LFQ intensity value for the least abundant, detected peptide.

Enzyme activity assay. Enzyme activity was tested in the OM and intracellular fractions by incubating them at 37°C with colorimetric or fluorimetric substrates, as described previously (Rosendale *et al.*, 2012). The following substrate/enzyme combinations were used: 4-nitrophenyl N-acetyl-ß-D-glucosaminide/N-acetyl-ß-glucosaminidase, p-nitrophenyl-alpha-L-fucoside/fucosidase, PNP N-acetyl-glucosamine-sulfate/GlcNAc-sulfatase, and 2′-(4-methylumbelliferyl)-α-D-N-acetylneuraminic acid/sialidase. Substrates were purchased from Carbosynth (Compton, Berkshire, UK). Reaction volume was 20 µl and substrate concentration was 1 mM in each reaction. All reactions were performed in duplicate.

***In vitro* culture and stimulation of human HEK-Blue hTLR2 cell line.** For the inflammatory pathway analysis the HEK-Blue hTLR2 cell line (Invivogen, CA, USA) was used. Stimulation of the receptors with the corresponding ligands activates NF-κB and AP-1, which induces the production of secreted embryonic alkaline phosphatase (SEAP), the levels of which can be measured by spectrophotometer (Spectramax). The cell line was grown and subcultured to 70 to 80% of confluency using Dulbecco's Modified Eagle Medium (DMEM) supplemented with 4.5 g/l D-glucose, 50 U/ml penicillin, 50 µg/ml streptomycin, 100 µg/ml Normocin, 2 mM L-glutamine, and 10% (v/v) of heat-inactivated Fetal Bovine Serum (FBS) as a maintenance medium. An immune response experiment was carried out by splitting HEK-blue cells in flat-bottom 96-well plates and stimulating them by addition of 20 µl protein fractions (1 µg protein/well). The 96-well plates were incubated for 20 to 24 h at 37°C under oxic atmosphere in an incubator supplemented with 5 % CO_2. Receptor ligand Pam3CSK4 (10 ng/ml) was used as positive control whereas maintenance medium without any selective antibiotics was used as negative control. SEAP secretion was detected by measuring the OD600 at 15 min, 1 h, 2 h, and 3 h after addition of 180 µl of QUANTI-Blue (Invivogen, CA, USA) to 20 µl of induced HEK-Blue hTLR2 supernatant. The data shown here are from the 1 h measurement point. Experiments for TLR2 were performed in triplicate.

Bioinformatics analysis. Proteins were categorized based on results of SignalP v.4.0 using default settings for Gram-negative bacteria (Petersen *et al.*, 2011), LipoP v.1.0 (Juncker *et al.*, 2003), SecretomeP v.2.0 (Bendtsen *et al.*, 2005), TMHMM v.2.0 (Krogh *et al.*, 2001) and BOMP (Berven *et al.*, 2004). Subcellular localization was determined using CELLO v.2.5 (set for Gram-negative bacteria)

(Yu *et al.*, 2004) and PSORTb v3.0 (Yu *et al.*, 2010). BLAST searches were run against the non-redundant (nr) database at http://blast.ncbi.nlm.nih.gov/ using default settings. TIGRFAM and Pfam were used for screening proteins with PEP-CTERM domains (Finn *et al.*, 2014; Selengut *et al.*, 2007).

Results

A. muciniphila OM proteins were isolated using two different methods (sarkosyl and sucrose density-gradient centrifugation) and two different growth conditions (mucin and glucose). In the case of the sucrose density-gradient extraction, 24 fractions were collected from each gradient and their protein concentration was measured. For each condition, four fractions along the gradient were chosen for proteomic analysis by mass spectrometry (Figure 1). Additionally, a fraction enriched for intracellular proteins was analyzed for both conditions. The total number of proteins detected for each fraction is shown in Table 1. The amount of proteins detected from *A. muciniphila* whole proteome samples from a previous study (**Chapter 3**) is shown for comparison.

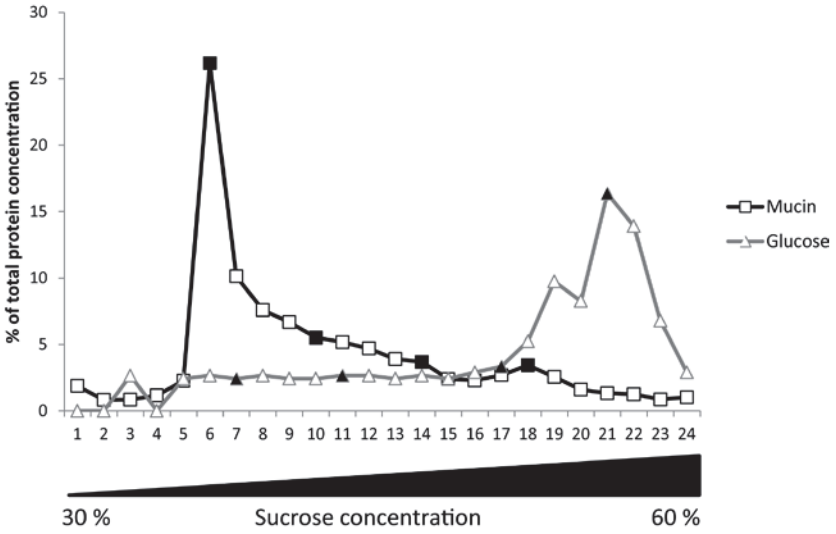


Figure 1. Relative protein concentration of the sucrose-density gradient fractions for mucin and glucose-grown *A. muciniphila*. 24 fractions were collected from each gradient and four fractions from each condition (black markers) were chosen for mass spectrometry analysis.

Table 1. Number of identified proteins from each bacterial fraction analyzed by mass spectrometry.

	Number of proteins	
Sample	Mucin	Glucose
Whole proteome	1054	1087
Sarkosyl method	812	519
Intracellular fraction	965	999
Sucrose density-gradient fraction 1	527	141
Sucrose density-gradient fraction 2	517	325
Sucrose density-gradient fraction 3	507	352
Sucrose density-gradient fraction 4	488	439

Extraction of *A. muciniphila* OM proteins

Successful enrichment of OM proteins was achieved within sarkosyl-extracted OM fractions as verified using the CELLO software, which predicts protein localization (Figure 2). To analyze the abundance of proteins within the fractions from different cellular locations, we compared their LFQ intensities. LFQ intensity is commonly used as a proxy for absolute protein abundance in mass spectrometry data analysis to compare relative amounts between samples (Cox *et al.*, 2014). Total abundance of OM proteins, defined by summed-up LFQ intensities of each protein, was higher in the OM fractions in comparison to the whole proteome and intracellular fractions (Figure 2A, B). The sucrose density-gradient method was used as a complimentary approach to the sarkosyl method. In the sucrose density-gradient samples the abundance of outer membrane proteins was the lowest in the fraction with the lowest sucrose density, and increased along the gradient for both growth conditions (Figure 2C, D).

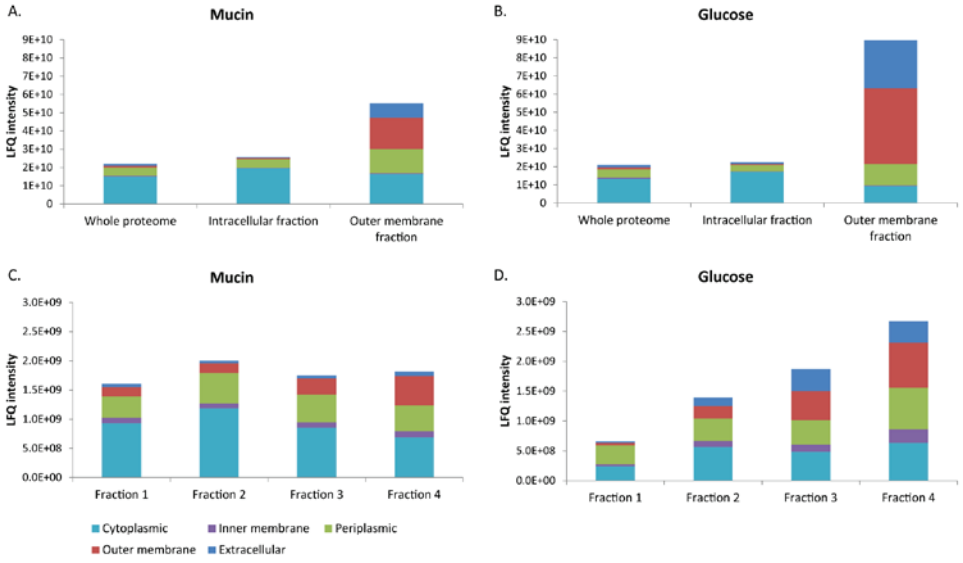


Figure 2. Distribution of proteins from different bacterial compartments in the isolated fractions. Outer membrane and extracellular proteins are enriched in the OM fraction (sarkosyl-method) in comparison to whole proteome and intracellular fraction in both mucin and glucose condition (A, B). The abundance of OM and extracellular proteins increases along the sucrose density-gradient in both conditions (C, D). Label-free quantification (LFQ) intensity was used as a proxy for absolute protein abundance. CELLO software was used to predict protein localization.

The OM-enriched proteins identified as actual OM proteins according to CELLO or PSORTb were chosen for further analyses (Figure 3). Many extracellular proteins were highly abundant in the OM fractions: 48 % of all detected extracellular proteins had a relative LFQ intensity above 8.5 (Figure 4). These proteins may be closely associated with the OM, and thus isolated together with it. Therefore, all extracellular proteins were included in our selection. In addition, 14 cytoplasmic proteins and 12 periplasmic proteins that had an LFQ intensity above 8.5 were added to our preliminary list, to find out whether these could be classified as OM proteins. Our preliminary list contained 145 proteins in total and was further analyzed and refined as described below.

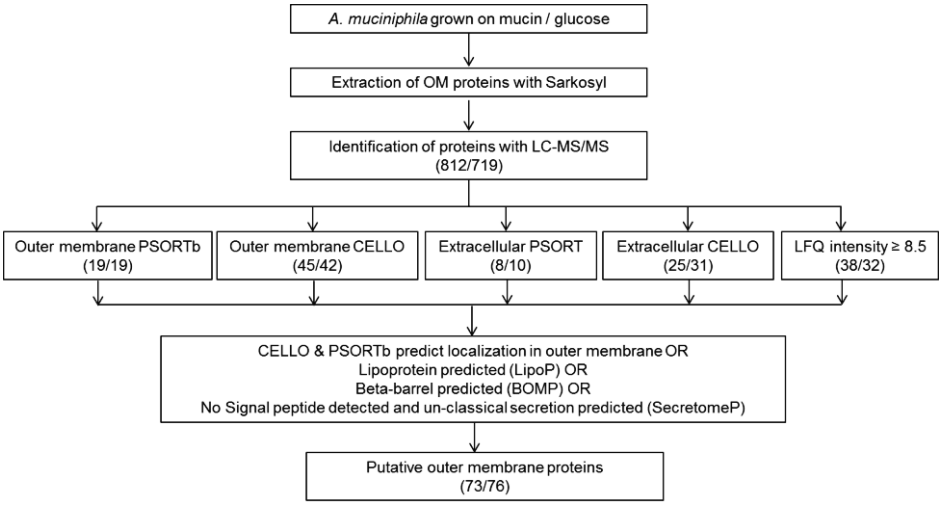


Figure 3. Prediction of outer membrane proteins. All proteins identified from the sarkosyl-extracted OM fraction were subjected to the prediction approach for the identification of outer membrane proteins. The total number of selected proteins for *A. muciniphila* grown on mucin (left) and glucose (right) are shown for each step.

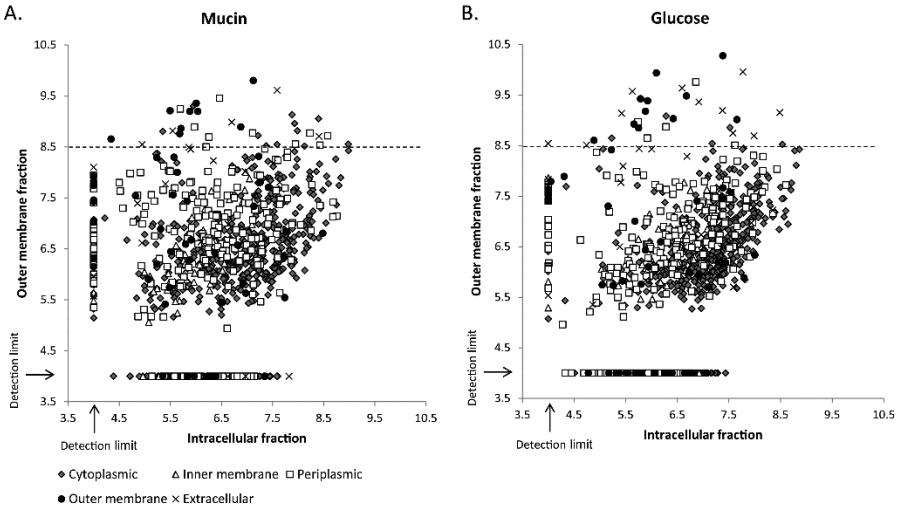


Figure 4. Sarkosyl extraction leads to enrichment of outer membrane and extracellular proteins. Subcellular localization of proteins identified from the OM and intracellular fraction of *A. muciniphila* grown on mucin (A) or glucose (B). CELLO was used to predict protein localization. Relative abundances of the proteins are presented on a log10 scale. Relative abundance of 4.0 represents proteins that were not detected or were under the detection limit. All proteins with LFQ intensity above 8.5 in the sarkosyl fractions (dashed line) were included in the preliminary list of potential OM proteins.

Identification and annotation of new *A. muciniphila* OM proteins

Next, we confirmed the localization of proteins of the preliminary list by applying several criteria: the protein was considered as a potential OM protein if (1) both CELLO and PSORTb predicted the protein to be localized in the OM, (2) LipoP software predicted the protein to be a lipoprotein, (3) BOMP software predicted the presence of beta-barrel(s), or (4) the protein had no signal protein, but SecretomeP predicted unclassical secretion. Proteins that had a signal peptide but did not fill any of the above criteria were assumed to be periplasmic. After applying these criteria, 73 proteins from mucin-grown cells and 76 proteins from glucose-grown cells in the isolated OM fractions were predicted to be OM proteins (Figure 3). Five out of the 14 cytoplasmic and nine out of the 12 periplasmic proteins included in the preliminary list were identified as putative OM proteins. In total, 79 putative OM and membrane-associated extracellular proteins were identified (Table 2). The majority of these were also present in the sucrose density-gradient samples: 62 proteins for the mucin-grown and 58 proteins for the glucose-grown condition.

Additionally, we searched for potential OM proteins, which were not present in either of the sarkosyl-extracted cultures. The *A. muciniphila* genome was found to encode additional 17 proteins, which were predicted as OM proteins according to the applied criteria (see above), and 12 (71 %) of these were uncharacterized proteins (Table 3). Two of the proteins (Amuc_0356 and Amuc_0480) were found to be present in the sucrose density-gradient fractions. Three of the proteins (Amuc_0983, Amuc_1011 and Amuc_1945) were present in low amounts when the whole proteome of *A. muciniphila* was analyzed. Previously described RNA-seq data confirmed that transcripts of the genes for all the 17 proteins can be detected in our dataset of *A. muciniphila* grown on mucin or glucose (**Chapter 3**).

Table 2. Putative OM and membrane-associated extracellular proteins.

Locus tag	Annotation	Glucose (Log10 LFQ)	Mucin (Log10 LFQ)	Fold change
Amuc_0006	Putative uncharacterized protein	4.0	5.4	25.9
Amuc_0019	Putative uncharacterized protein	5.4	7.4	109.6
Amuc_0032	Putative uncharacterized protein	7.8	6.3	35.1
Amuc_0036	YD repeat protein	5.9	6.8	7.2
Amuc_0074	Putative uncharacterized protein	8.4	7.9	2.6
Amuc_0105	RND efflux system, outer membrane lipoprotein, NodT family	7.9	8.7	5.8
Amuc_0172	Putative uncharacterized protein	7.4	8.9	31.0
Amuc_0194	Putative uncharacterized protein	8.6	8.2	2.7
Amuc_0219	Efflux transporter, RND family, MFP subunit	6.4	6.2	1.7
Amuc_0294	30S ribosomal protein S17	8.4	8.8	2.5
Amuc_0301	50S ribosomal protein L2	8.3	8.7	2.7
Amuc_0304	50S ribosomal protein L3	8.3	8.5	1.7
Amuc_0308	30S ribosomal protein S12	8.0	8.7	5.4
Amuc_0336	TonB-dependent receptor	9.9	9.4	3.8
Amuc_0355	Putative uncharacterized protein	8.5	8.1	2.8
Amuc_0360	Putative uncharacterized protein	8.4	8.2	1.6
Amuc_0371	Two component regulator propeller domain protein	5.7	7.4	48.5
Amuc_0385	ErfK/YbiS/YcfS/YnhG family protein	6.6	8.1	31.0
Amuc_0392	Coagulation factor 5/8 type domain protein	7.6	7.8	1.7
Amuc_0394	Putative uncharacterized protein	7.6	7.8	1.7
Amuc_0433	Putative uncharacterized protein	7.5	7.4	1.3
Amuc_0435	Putative uncharacterized protein	8.4	8.8	2.4
Amuc_0438	Ribosomal protein S11	7.7	9.1	29.3
Amuc_0513	Putative uncharacterized protein	8.5	7.6	7.3
Amuc_0576	Peptidase M16 domain protein	7.5	8.8	19.7
Amuc_0584	Outer membrane autotransporter barrel domain protein	8.7	7.5	16.4
Amuc_0609	Putative uncharacterized protein	8.4	7.9	3.0
Amuc_0610	Tetratricopeptide TPR_2 repeat protein	8.6	8.5	1.4
Amuc_0682	OmpA/MotB domain protein	7.1	6.9	1.4
Amuc_0687	Outer membrane autotransporter barrel domain protein	9.6	8.5	15.2
Amuc_0735	YD repeat protein	4.0	5.8	67.4
Amuc_0789	Putative uncharacterized protein	7.2	4.0	1669.9
Amuc_0815	DNA polymerase III, beta subunit	6.5	6.2	1.9

Locus tag	Annotation	Glucose (Log10 LFQ)	Mucin (Log10 LFQ)	Fold change
Amuc_0820	Peptidoglycan-binding LysM	8.9	8.9	1.0
Amuc_0823	Putative uncharacterized protein	7.8	7.8	1.1
Amuc_0837	Putative uncharacterized protein	6.1	4.0	125.7
Amuc_0904	Aconitate hydratase 1	6.2	5.5	4.5
Amuc_0931	50S ribosomal protein L15	8.4	9.0	3.3
Amuc_0967	RNP-1 like RNA-binding protein	7.5	7.7	1.7
Amuc_1008	Glycoside hydrolase family 31	6.2	6.3	1.5
Amuc_1026	Peptidyl-prolyl cis-trans isomerase	6.7	4.0	542.7
Amuc_1039	Outer membrane autotransporter barrel domain protein	8.1	6.0	134.5
Amuc_1053	Outer membrane protein assembly complex, YaeT protein	9.4	9.2	1.5
Amuc_1061	Outer membrane protein-like protein	8.0	8.4	2.1
Amuc_1098	Type II and III secretion system protein	10.3	9.8	3.0
Amuc_1114	Outer membrane autotransporter barrel domain protein	6.5	7.0	3.4
Amuc_1283	Outer membrane autotransporter barrel domain protein	9.2	8.7	2.8
Amuc_1310	17 kDa surface antigen	10.0	9.6	2.2
Amuc_1333	Putative uncharacterized protein	7.0	4.0	931.8
Amuc_1382	Secretion protein HlyD family protein	7.0	6.4	3.6
Amuc_1412	Putative uncharacterized protein	9.5	9.2	1.9
Amuc_1420	Putative uncharacterized protein	5.5	5.7	1.6
Amuc_1434	Putative uncharacterized protein	7.4	7.5	1.1
Amuc_1439	Organic solvent tolerance protein OstA-like protein	8.6	8.3	2.1
Amuc_1500	Polysaccharide deacetylase	9.0	9.2	1.9
Amuc_1512	OmpA/MotB domain protein	6.1	5.8	2.2
Amuc_1514	Putative uncharacterized protein	6.7	6.4	1.8
Amuc_1525	Putative uncharacterized protein	6.6	6.7	1.2
Amuc_1537	Outer membrane autotransporter barrel domain protein	8.7	8.2	3.3
Amuc_1583	Putative uncharacterized protein	7.0	7.0	1.1
Amuc_1620	Outer membrane autotransporter barrel domain protein	7.6	6.3	20.7
Amuc_1638	Type II and III secretion system protein	9.4	8.9	3.7
Amuc_1656	Putative uncharacterized protein	9.8	9.5	2.0
Amuc_1684	TonB-dependent receptor	8.9	8.0	8.4
Amuc_1687	Putative uncharacterized protein	9.4	9.0	2.4

Locus tag	Annotation	Glucose (Log10 LFQ)	Mucin (Log10 LFQ)	Fold change
Amuc_1722	Outer membrane autotransporter barrel domain protein	9.2	8.0	15.0
Amuc_1723	Outer membrane autotransporter barrel domain protein	8.3	6.6	44.6
Amuc_1743	Putative uncharacterized protein	8.8	8.8	1.2
Amuc_1891	Putative uncharacterized protein	9.6	6.6	909.0
Amuc_1939	Efflux transporter, RND family, MFP subunit	4.0	5.9	79.6
Amuc_2043	RND efflux system, outer membrane lipoprotein, NodT family	9.2	9.2	1.1
Amuc_2077	Polysaccharide export protein	9.0	8.3	5.4
Amuc_2099	Putative uncharacterized protein	8.5	8.5	1.1
Amuc_2107	Putative uncharacterized protein	8.2	8.8	3.6
Amuc_2108	Glycoside hydrolase family 16	5.9	4.0	87.2
Amuc_2111	ErfK/YbiS/YcfS/YnhG family protein	7.8	7.8	1.0
Amuc_2127	Carbohydrate-selective porin OprB	9.1	8.5	4.4
Amuc_2152	YD repeat protein	6.3	6.8	3.2
Amuc_2165	Putative uncharacterized protein	5.9	4.0	76.7

Relative abundances of the proteins based on mass spectrometry analysis are presented on a log10 scale. Relative abundance of 4.0 represents proteins that were not detected or were under the detection limit. Proteins with a change of > 10-fold between bacteria grown on mucin and glucose are color-coded in grey.

Table 3. Putative OM proteins predicted and detected.

Locus tag	Annotation	Locus tag	Annotation
Amuc_0146	Alpha-L-fucosidase	Amuc_1011	Uncharacterized protein
Amuc_0173	Uncharacterized protein	Amuc_1123	Uncharacterized protein
Amuc_0356	Uncharacterized protein	Amuc_1143	YD repeat protein
Amuc_0434	Uncharacterized protein	Amuc_1145	Uncharacterized protein
Amuc_0480	Alpha-1,3-galactosidase B	Amuc_1340	Uncharacterized protein
Amuc_0599	Uncharacterized protein	Amuc_1350	Uncharacterized protein
Amuc_0892	OM autotransporter barrel domain protein	Amuc_1831	Uncharacterized protein
Amuc_0908	Uncharacterized protein	Amuc_1945	Uncharacterized protein
Amuc_0983	YD repeat protein		

A total of 17 proteins were predicted to be OM proteins but were not detected by mass spectrometry in the OM fractions. However, the proteins Amuc_0356 and Amuc_0480 were detected in the sucrose density-gradient samples and Amuc_0983, Amuc_1011 and Amuc_1945 were detected in the whole proteome samples.

Activity of mucin-degrading enzymes

The enzyme activity of four mucin-degrading enzymes (sialidase, fucosidase, N-acetyl-ß-glucosaminidase, GlcNAc-sulfatase) was measured from the OM fractions and the intracellular fractions (Figure 5). The results indicated higher enzymatic activity in the intracellular fractions than OM fractions, for both glucose and mucin-grown cells. Most of these enzymes in *A. muciniphila* are predicted to be periplasmic.

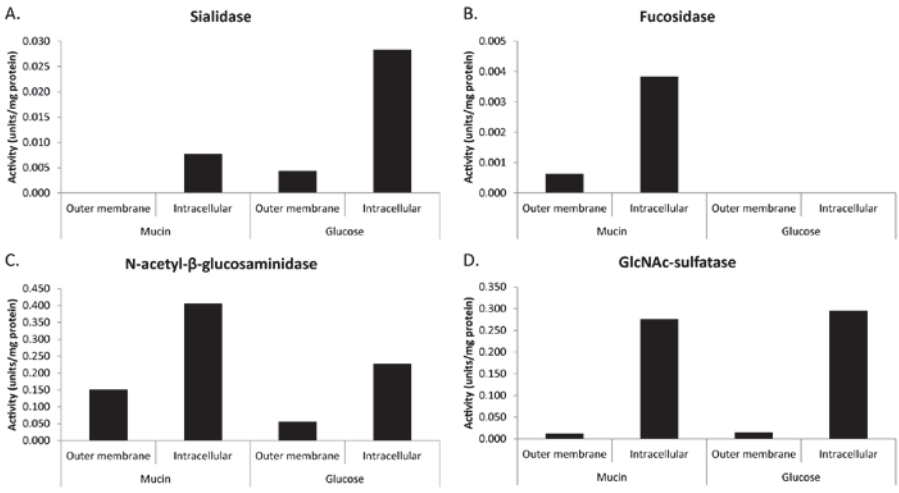


Figure 5. Activity of mucin-degrading enzymes in the OM and intracellular fractions. The enzyme activity of sialidase (A), fucosidase (B), N-acetyl-ß-glucosaminidase (C), and GlcNAc-sulfatase (D) was measured from OM and intracellular fractions of *A. muciniphila* grown on mucin or glucose. Average of duplicate measurements is shown.

Classification of *A. muciniphila* OM proteins

The proteins identified as potential OM proteins had multiple different functions (Figure 6). Nine proteins from both conditions were outer membrane proteins containing an autotransporter barrel domain. Four OM proteins of the mucin condition and three of the glucose condition were found to be involved in RND (Resistance-Nodulation-Division) efflux systems. Mucin-grown cells contained three and glucose-grown cells two YD repeat proteins. The function of many of the predicted OM proteins was unknown (26 proteins for mucin-grown and 29 for glucose-grown cells). BLAST searches were performed for all uncharacterized proteins with LFQ intensity above 8.5, but this did not reveal any closely related proteins.

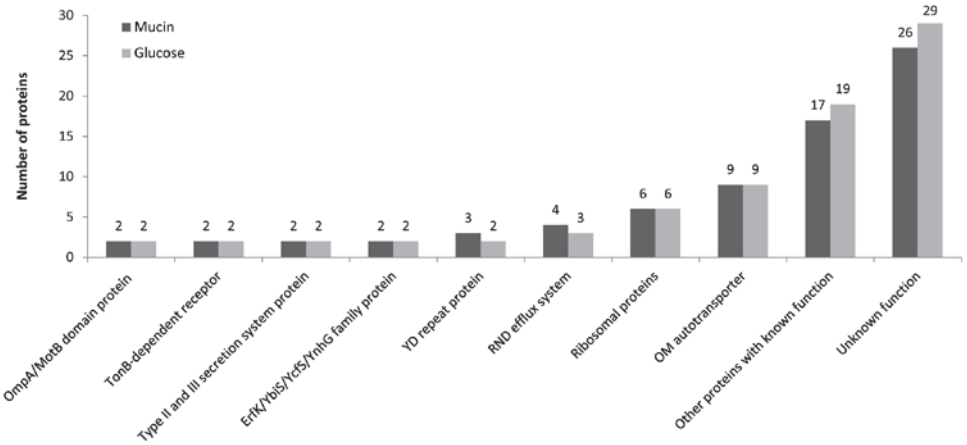


Figure 6. Classification of OM proteins. Proteins in the OM fraction of *A. muciniphila* grown on mucin (dark) or glucose (light) were identified by LC-MS/MS and classified based on analysis of their sequences.

Intriguingly, the four most abundant OM proteins were the same in *A. muciniphila* grown on mucin or glucose. Namely, type II and III secretion system protein (Amuc_1098), 17 kDa surface antigen (Amuc_1310), TonB-dependent receptor (Amuc_0336) and an uncharacterized protein (Amuc_1656).

BLAST searches revealed that the secretion system protein (Amuc_1098) showed similarities to a protein with the same predicted function in *Verrucomicrobium spinosum* (40 % identity, E-value: 0.0), and to hypothetical proteins from a *Verrucomicrobium* sp. (41 % identity, E-value 0.0) and *Haloferula* sp. (48 % identity, E-value 0.0) (Okazaki *et al.*, 2014; Ward *et al.*, 2000). Interestingly, Amuc_1098 also showed 31 % identity (E-value 4×10^{-86}) to a fimbrial assembly protein from *V. spinosum*, which is the closest described relative of *A. muciniphila* (92 % sequence similarity) (Derrien *et al.*, 2004). Furthermore, Amuc_1098 showed 27 % identity (E-value 5×10^{-62}) to the type 4 fimbrial biogenesis protein PilQ from *Lentisphaera araneosa*. A protein annotated as cell division protein FtsA (Amuc_1101), which is predicted to be in the same gene cluster with Amuc_1098, showed 32 % identity (E-value 8×10^{-78}) to type IV pilus assembly protein PilM, also from *V. spinosum*. Electron microscope (EM) images from *A. muciniphila* revealed fimbriae-like structures (Derrien et al 2009), and immuno-EM has confirmed the presence of some of the OM proteins in these fimbriae (Huuskonen et al, unpublished results).

BLAST analysis predicted the protein sequence of Amuc_1310 to belong to the Rick_17kDa_Anti superfamily, which includes several *Rickettsia* genus-specific 17 kDa surface antigen proteins. Also other 17 kDa surface antigens and (outer) membrane proteins from various bacteria showed significant (30–40 %) identities to Amuc_1310. The TonB-dependent receptor (Amuc_0336) showed low similarity (sequence identity < 25 %) to TonB-dependent receptors, which are canonical periplasmic-space spanning proteins involved in transport, and outer membrane proteins from *Pseudomonas aeruginosa*. Amuc_1656 showed no similarity to any protein with a known function.

Influence of environmental growth conditions on OM protein composition

From the identified OM proteins, 23 out of the 79 showed more than a 10-fold change between *A. muciniphila* grown on either mucin or glucose (Table 2). Six out of the nine proteins containing an autotransporter barrel domain were present at a higher level in glucose-grown as compared to mucin-grown cells. In addition, a two component regulator propeller domain protein (Amuc_0371) and a peptidase (Amuc_0576) were more abundant in mucin-grown cells, whilst a glycoside hydrolase family 16 protein (Amuc_2108) and peptidyl-prolyl cis-trans isomerase (Amuc_1026) were present in higher abundance in glucose-grown cells. Many uncharacterized proteins were among the ones with highest fold change differences between the two conditions. There were four uncharacterized proteins (Amuc_0789, Amuc_0837, Amuc_1333, Amuc_2165) exclusively found in the OM fraction extracted from glucose-grown cells and one uncharacterized protein (Amuc_0006), exclusively found in mucin-grown cells.

During the sucrose density-gradient centrifugation, the samples derived from mucin and glucose-grown *A. muciniphila* behaved differently from each other. For the mucin fractions the highest protein concentrations were measured from the top of the gradient where the sucrose concentration was lower, whereas for the glucose fractions the highest protein concentrations were at the bottom of the gradient where the sucrose concentration is higher (Figure 1).

The production of PEP-CTERM proteins is confirmed in *A. muciniphila*

BLAST searches and subsequent TIGRFAM and PFAM analyses revealed that three of the glucose-exclusive proteins in the OM fraction contained a PEP-CTERM domain (Amuc_0789, Amuc_1333, Amuc_2165). This domain was first identified in *V. spinosum* and is thought to be involved in protein sorting and cell surface localization (Haft *et al.*, 2006). The domain includes the motif Pro-Glu-Pro (PEP), which is considered a potential recognition or processing site, followed by a

predicted transmembrane helix and a cluster rich in basic amino acids. These target proteins are generally destined to transit cellular membranes during their biosynthesis and undergo further posttranslational modifications such as glycosylation.

So far PEP-CTERM domains have only been found in the genomes of bacteria that possess an inner membrane, a periplasm, an outer membrane, and the EpsH (exopolysaccharide locus protein H) gene. An EpsH-like gene is also found in the genome of *A. muciniphila* (Amuc_1470). We screened for the presence of PEP-CTERM proteins in the genome of *A. muciniphila* and found 23 proteins containing this domain (Table 4). The synthesis of nine of these was confirmed by mass spectrometry as they were present in at least one of the analyzed bacterial fractions. We did not detect any *A. muciniphila* EpsH peptides in our proteomic analysis. Also other studies have failed to show proteomic evidence for EpsH from organisms known to encode it in their genome (Haft *et al.*, 2012).

A. muciniphila OM proteins evoke a stronger TLR2 response than intracellular proteins

To study the impact of *A. muciniphila* OM proteins on the host immune response, Toll-like receptor 2 (TLR2) activation by the whole proteome, OM fraction and intracellular fraction was compared (Figure 7). All the tested protein fractions activated the NF-κB pathway through TLR2 receptor. The OM fraction of mucin-grown bacteria led to a higher TLR2 response than the intracellular fraction, whereas for glucose-grown bacteria there was no difference in the response between the fractions. This indicates that *A. muciniphila* OM proteins may play an important role in stimulating the immune system, but the response of more defined protein fractions should be tested to clarify the results.

Table 4. PEP-CTERM domain-containing proteins predicted and detected.

	Mucin			Glucose	
Locus tag	Whole proteome	Intracellular	Outer membrane	Whole proteome	Intracellular	Outer membrane
Amuc_0215	ND	ND	ND	ND	ND	ND
Amuc_0428	ND	ND	ND	ND	ND	ND
Amuc_0622	ND	ND	ND	ND	ND	ND
Amuc_0646	6.49	6.62	ND	6.49	7.32	6.21
Amuc_0688	ND	ND	ND	ND	ND	ND
Amuc_0789	6.93	7.82	ND	5.84	8.35	7.22
Amuc_0825	ND	ND	ND	ND	ND	ND
Amuc_0952	ND	6.97	ND	5.05	7.59	5.70
Amuc_0990	4.63	5.35	ND	5.57	6.78	ND
Amuc_0991	ND	ND	ND	ND	ND	ND
Amuc_1028	ND	ND	ND	ND	ND	ND
Amuc_1209	ND	ND	ND	ND	ND	ND
Amuc_1221	ND	ND	ND	ND	ND	ND
Amuc_1266	ND	ND	ND	ND	ND	ND
Amuc_1333	6.81	7.39	ND	6.61	7.94	6.97
Amuc_1451	ND	ND	ND	4.54	7.02	ND
Amuc_1560	ND	ND	ND	ND	ND	ND
Amuc_1978	ND	ND	ND	ND	ND	ND
Amuc_1982	ND	ND	ND	ND	ND	ND
Amuc_2045	ND	5.48	ND	6.16	6.68	ND
Amuc_2105	ND	ND	ND	ND	ND	ND
Amuc_2126	4.66	5.83	ND	5.51	6.84	ND
Amuc_2165	ND	7.35	ND	4.77	7.79	5.88

A total of 23 proteins were found to contain a PEP-CTERM domain in the genome of *A. muciniphila*. All the proteins are uncharacterized. Proteomic evidence was detected for nine of these. Relative abundance of the protein on a log10 scale in the fractions is shown. ND; not detected.

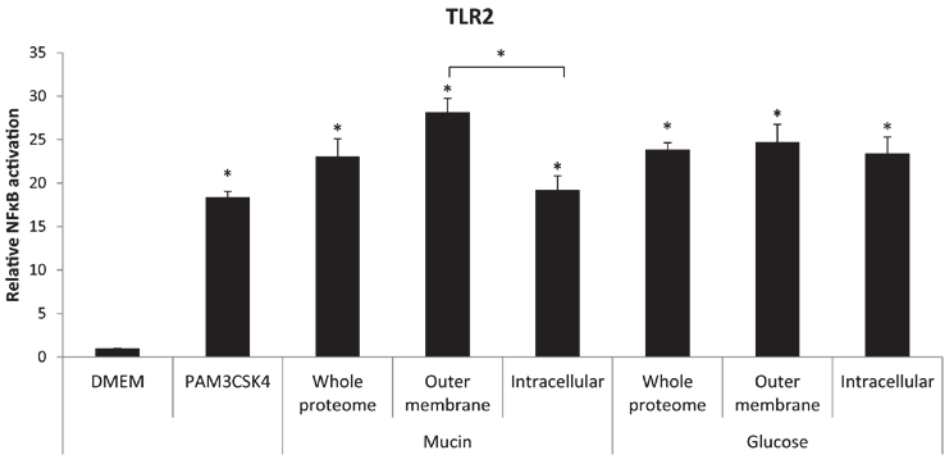


Figure 7. *A. muciniphila* protein fractions activating the TLR2 signaling pathway. TLR2 signaling in the HEK-Blue hTLR2 cell line by whole proteome, outer membrane fraction and intracellular fraction of *A. muciniphila* grown on mucin or glucose. 1 µg of protein/well was used for stimulation. DMEM; medium control, PAM3CSK4; positive control. *, P<0.05 compared to DMEM. Mean and standard deviation from three experiments are shown.

Discussion

Using a combination of experimental mass-spectrometry based identification and advanced bioinformatics, we identified 79 putative OM proteins in *A. muciniphila*, which comprises 3.6 % of the 2176 predicted protein-coding sequences in this mucus-degrading intestinal bacterium. This fraction is slightly higher than that found in other Gram-negative bacteria where it was estimated that 2-3 % of the genome codes for OM proteins (Molloy *et al.*, 2000; Wimley, 2003). As we included membrane-associated extracellular proteins in the analysis, the actual number of OM proteins in *A. muciniphila* is presumably somewhat lower. Extracellular proteins were included as they may be involved in communication with the host and are therefore of great interest. Seven different bioinformatics tools were employed to characterize *A. muciniphila* OM and extracellular proteins based on the predicted secretion signals, occurrence of transmembrane strands, presence of beta-barrel structures and attachment to the cell wall by lipid motifs.

Since it was discovered that sarkosyl can selectively solubilize cytoplasmic and inner membranes while conserving the integrity of the OM, this ionic detergent has been widely used in the purification of OM proteins in Gram-negative bacteria

(Brown *et al.*, 2010; Moumene *et al.*, 2015; Rhomberg *et al.*, 2004). In comparison to other methods, sarkosyl extraction leads to higher purity and better reproducibility of the OM extracts, while the basis for OM resistance to this detergent is not known (Filip *et al.*, 1973; Frankel *et al.*, 1991). In our study, using the sarkosyl-method led to enrichment of OM proteins, but many proteins from other cellular locations were also present. The high sensitivity of LC-MS/MS might be one of the reasons why so many proteins were identified from the fractions, and by optimizing the protocol, the abundance of non-OM proteins could be further reduced. Despite this we were able to identify a set of putative OM proteins from the fractions. However, due to the challenges of optimizing the extraction method for a new species, i.e. *A. muciniphila*, the mass spectrometry analysis was only performed for one sample from each growth condition and fractionation method. The extraction and subsequent mass spectrometry analysis should be repeated to confirm robustness of the results.

In addition to the OM proteins produced by *A. muciniphila*, we identified another 12 OM proteins from the genome using *in silico* analysis. These were not present in any of the studied fractions. Based on RNA-seq data all the genes encoding these proteins are transcriptionally active (**Chapter 3**), suggesting that their absence from the proteomics dataset could be due to a methodological reason, such as their high hydrophobicity. However, prediction of protein hydrophobicity did not reveal any of the 12 proteins to be particularly hydrophobic. It is also possible that these proteins were produced in such low amounts that they were not detected by LC-MS/MS. The sizes of these proteins ranged from 30 to 220 kDa, which should not pose a problem as both smaller and bigger proteins were identified by mass spectrometry.

Enzyme activity assays confirmed the activity of mucin-degrading enzymes in *A. muciniphila*, and showed higher activity in the intracellular fractions in comparison to OM fractions. This is most likely because the enzymes are transported from the cytoplasm to the periplasmic space, and are not bound to the membranes.

More insight into the OM proteome biosynthesis of *A. muciniphila* was generated through isolation and identification of proteins covering the most important pathways of OM biogenesis. The YaeT protein (Amuc_1053), which is an essential protein for OM protein biogenesis due to its role in insertion of beta-barrel proteins into the OM (Tokuda, 2009; Werner and Misra, 2005), was among the ten most highly abundant proteins in the extracted OM fractions. Furthermore, the organic solvent tolerance protein (Amuc_1439), known also as increased membrane permeability (Imp) protein, needed in LPS transport to the bacterial cell surface

(Bos *et al.*, 2004), was found in high abundance. Based on these observations, the presence of a Gram-negative OM in *A. muciniphila* is very likely, whereas evidence for the presence of an intracytoplasmic membrane is still scarce. Previously one study has observed partial compartmentalization in *A. muciniphila* cells using EM (Dubourg *et al.*, 2013). Another study identified a highly conserved genetic module preferentially present in compartmentalized PVC species, but this was not found in *A. muciniphila* (Kamneva *et al.*, 2012). In this study, the presence of the protein domain DUF1501 correlated with organisms known to contain intracellular membranes, including *V. spinosum*, the closest described relative of *A. muciniphila*. However, one *Planctomycetes* species, *Kuenenia stuttgartiensis*, known to possess intracellular membranes, did not have the domain.

The type II and III secretion system protein (Amuc_1098) was found to be the most highly produced OM protein in *A. muciniphila*, regardless of the environmental conditions. Another type II and III secretion system protein (Amuc_1638) was also found in high abundance in the OM fractions (Table 2). These types of proteins are involved in transport of many different substrates across the OM (Diepold and Wagner, 2014; Nivaskumar and Francetic, 2014). The substrates can also remain anchored to or bound at the surface of the bacteria after secretion. According to our recent analysis, the gene cluster that includes Amuc_1098, may be involved in formation of fimbriae (Huuskonen et al, unpublished results). This would make it an interesting candidate for immune signaling in *A. muciniphila*, as fimbriae and pili structures are known to be involved in the modulation of host immune responses (Craig and Li, 2008; Ehara *et al.*, 1993; von Ossowski *et al.*, 2013). However, more research is needed to identify the actual proteins that form the fimbriae in *A. muciniphila*.

Other groups of OM proteins found to be produced by *A. muciniphila* were autotransporters, RND multidrug efflux pumps and YD proteins. Autotransporters are an extensive family of proteins, which can be either secreted or cell-surface-exposed, and function for example as enzymes, adhesins, cytotoxins or mediate bacterial motility (Grijpstra *et al.*, 2013). They have mainly been studied for their virulence-related properties in pathogens but are also commonly found in non-pathogenic bacteria. RND multidrug efflux pumps exist in a tripartite form traversing both the OM and IM, and thus they are able to efficiently pump out drug molecules directly into extracellular space (Nikaido, 2011). The function of YD proteins is not very well studied but they have been suggested to be involved in carbohydrate binding and bacterial interactions with eukaryotic host cells (Jackson *et al.*, 2009; Koskiniemi *et al.*, 2013).

Among the highly abundant proteins in the OM fractions were several proteins annotated as cytoplasmic, including six ribosomal proteins. It is possible that these proteins were so abundant in the cultures that they ended up in high concentrations also in the OM fraction, but it may also be that they have secondary roles in the bacteria. Proteins that are able to perform two or more functions are called moonlighting proteins (Copley, 2012). Several glycolytic, housekeeping and ribosomal proteins are often found on the surface of bacteria where they develop other functions (Henderson and Martin, 2011; Sherry *et al.*, 2011).

The presence of specific OM proteins in Gram-negative bacteria is modulated by the available carbon and energy sources (Haussmann *et al.*, 2009; Papasotiriou *et al.*, 2008; Yang *et al.*, 2011). The abundance of one third of the putative *A. muciniphila* OM proteins was altered during growth on glucose in comparison to mucin. However, for most of these proteins, the presence of other proteins with the same function was found in both conditions. This suggests that the specific function is not missing, but is taken over by a different protein. Unfortunately, many of the proteins that were most affected by the difference in carbon sources remain uncharacterized, despite efforts to search for related proteins. This is a common issue when handling proteomics data, especially as *A. muciniphila* belongs to a phylum which has not been studied extensively yet. Further research into the function of these proteins could reveal important characteristics of *A. muciniphila.*

Differences between the two growth conditions were also seen during the sucrose density-gradient method, where the migration patterns differed markedly between samples derived from mucin or glucose-grown *A. muciniphila.* This may be a consequence of the higher protein concentration of the initial mucin sample that was laid on top of the gradient, in comparison to the glucose sample, or an indication of a higher ratio of low-density proteins in the mucin-grown *A. muciniphila.* No clear explanation for the different behavior was found from the proteomic analysis of the fractions. Additional experimental work is required to find out whether the difference is caused by a technical or a biological reason.

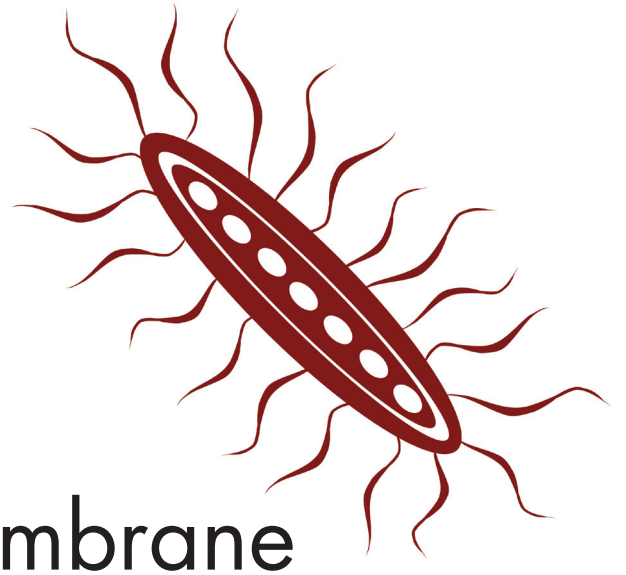
To our knowledge, this is the first proteomic evidence for the presence of PEP-CTERM proteins in bacteria. Previous research failed to detect peptides from proteins containing this domain in four *Cyanothece* strains, and only described peptides from the archaeal PGF-CTERM system (Haft *et al.*, 2012). *A. muciniphila* encodes 23 PEP-CTERM proteins and we detected nine of these by mass spectrometry. The proteins were present in relatively low abundance in the whole proteome samples, which may explain the difficulty in detecting them. Higher levels of PEP-CTERM proteins were detected in the intracellular fractions compared to outer membrane fractions. These proteins may be heavily

posttranslationally modified, including removal of N-terminal and C-terminal transmembrane domains and extensive glycosylation (Haft *et al.*, 2012). This complicates the identification of these peptides by mass spectrometry, implying that they are more easily detected in the intracellular fraction where the modifications may have not taken place yet. Thus, we cannot conclude whether the proteins were more abundantly present in the cytoplasm as opposed to the outer membrane, or if the observed lower levels were due to difficulties in assigning the proteomic data. Addition of a deglycosylation step to the sample preparation protocol might enhance the identification of PEP-CTERM proteins. In conclusion, these findings suggest that *A. muciniphila* may be a well-suited model organism for studying the PEP-CTERM/exosortase system in bacteria, as relatively little is known about it thus far.

Finally, we showed that *A. muciniphila* proteins activate the intestinal TLR2 receptor through the NF-κB pathway *in vitro*. The results suggested that OM proteins may lead to a higher immune response than intracellular proteins. The role of *A. muciniphila* OM proteins in host immune response is further studied in **Chapter 6**.

In summary, this study provides the first proteomic characterization of *A. muciniphila* OM proteins. We have identified highly abundant proteins involved in secretion and transport, as well as many uncharacterized proteins. These results form a valuable dataset for further studies on the exact roles of individual proteins, especially in respect to the immunomodulatory properties of *A. muciniphila*.

Chapter 6



Outer membrane proteins of *Akkermansia muciniphila* drive its immunostimulatory properties

Noora Ottman, Justus Reunanen, Marjolein Meijerink, Taija E. Pietilä, Veera Kainulainen, Judith Klievink, Laura Huuskonen, Steven Aalvink, Mikael Skurnik, Sjef Boeren, Reetta Satokari, Annick Mercenier, Airi Palva, Hauke Smidt, Willem M. de Vos, and Clara Belzer

Submitted for publication

Summary

The abundance of *Akkermansia muciniphila*, a microbe colonizing the gut mucosa, is reduced in obesity and inflammatory diseases, but its immunological properties are unexplored. Here *A. muciniphila* MucT showed immunostimulatory capacities in relevant *in vitro* models by inducing production of cytokines and activating Toll-like receptor (TLR) 2 and TLR4. Levels of IL-10 induced by live *A. muciniphila* were similar to those induced by *Faecalibacterium prausnitzii* A2-165 and *Lactobacillus plantarum* WCFS1. A set of *A. muciniphila* cell-envelope proteins was found to be enriched for TLR2-signaling capacity and candidate proteins were purified. This resulted in the identification of a 30 kDa protein as the main candidate for activating TLR2 and cytokine response as well as increasing transepithelial resistance. Localization studies confirmed this protein to be located in the outer membrane. These results indicate that outer membrane protein composition of *A. muciniphila* prompts immunological homeostasis in the gut mucosa by keeping the immune system alert and by improving gut barrier function.

Introduction

The human gastrointestinal (GI) tract provides a living environment for the complex and diverse microbiota, which is involved in many metabolic, nutritional, physiological and immunological interactions with the host (Sommer and Backhed, 2013).

The host immune system plays an important role in distinguishing between commensal and pathogenic bacteria. On one hand, the immune system needs to stay alert to recognize potential pathogens, and on the other hand, it has to tolerate the commensal bacteria inhabiting the gut (Rescigno, 2013). This homeostasis is achieved through pattern recognition receptor (PRR) families expressed in immune cells. PRRs, such as Toll-like receptors (TLRs) and nucleotide binding and oligomerization domain-like receptors (NLRs), identify microbe-associated molecular patterns (MAMPs). MAMPs are molecules associated with both commensal and pathogenic microorganisms. Another important component of the mucosal immune system are the secretory immunoglobulins, such as IgA and IgG, which are secreted by plasma cells and function by excluding bacteria from the epithelium (Belzer *et al.*, 2011; Strugnell and Wijburg, 2010). Defining the immunomodulatory capacity of members of the microbiota is essential in understanding their involvement in the establishment of mucosal tolerance and balanced intestinal immune responses. There is also growing evidence about the influence of the gut microbiota on the systemic immune system, and consequently, the development of autoimmune diseases (Kuhn and Stappenbeck, 2013).

One of the key players in the colonic mucus-associated microbiota is *Akkermansia muciniphila*, which colonizes a considerable part of the human population and comprises 1–4 % of the fecal microbiota of healthy adults (Derrien *et al.*, 2008; Derrien *et al.*, 2004). This bacterium is highly adapted to its living environment as it is capable of using mucin as the sole carbon and nitrogen source. Levels of *A. muciniphila* have been shown to be inversely correlated with several disorders, such as inflammatory bowel diseases (IBD) (Png *et al.*, 2010; Rajilic-Stojanovic *et al.*, 2013), appendicitis (Swidsinski *et al.*, 2011), obesity (Karlsson *et al.*, 2012) and diabetes (Zhang *et al.*, 2013), but not much is known about its immunological mechanism of action.

The impact of *A. muciniphila* on the host has been studied in mono-associated mice and organoids, where most of the genes affected by the bacteria were implicated in immune and metabolic responses (Derrien *et al.*, 2011; Lukovac *et al.*, 2014). The induction of immune response-associated genes was most obvious in the colon of *A. muciniphila* colonized mice, where over 60 genes, including 16 genes encoding

cluster of differentiation (CD) antigen markers and 10 genes encoding immune cell membrane receptors were up-regulated upon exposure. The impact on host metabolism is in line with the fact that *A. muciniphila* can have an inhibiting effect on obesity and diabetes development. The abundance of *A. muciniphila* decreased in obese and type 2 diabetic mice, and treatment with the bacteria reversed high-fat diet induced metabolic disorders, such as adipose tissue inflammation (Everard *et al.*, 2013). This was confirmed in a later study where *A. muciniphila*-administered high-fat diet fed mice showed improved glucose tolerance and an increase in the number of goblet cells and adipose tissue-resident CD4+ Foxp3+ regulatory T cells (Shin *et al.*, 2014). Conversely to these studies implying a protective effect of *A. muciniphila* on intestinal barrier function and immune stimulation, several mouse studies have reported increased numbers of these mucosal bacteria in dextran sodium sulfate (DSS)-induced colitis (Berry *et al.*, 2012; Hakansson *et al.*, 2014; Kang *et al.*, 2013). This could be explained by a simple outgrowth of *A. muciniphila* in response to the thickening of the mucus layer during DSS-induced colitis. A similar explanation can rationalize the observation that *A. muciniphila* administration in a minimal community appeared to aggravate *Salmonella enterica* Typhimurium-induced gut inflammation in a gnotobiotic mouse model (Ganesh *et al.*, 2013).

The aim of this study was to characterize the immunomodulatory properties of *A. muciniphila* MucT by measuring cytokine production in human derived peripheral blood mononuclear cells (PBMCs) and activation of inflammatory pathways on reporter cell lines expressing either TLR2/4/5/9 or NOD2-receptor. The immune response of *A. muciniphila* was compared to two other commensals, *Faecalibacterium prausnitzii* A2-165 and *Lactobacillus plantarum* WCFS1. A proteomics approach was used to identify candidate signaling molecules from bacterial fractions, and a collection of these proteins were purified from overproducing *Escherichia coli* clones. These proteins were tested for their capacity to induce TLR2-signaling, cytokine production and to affect transepithelial resistance (TER) in Caco-2 model system. Localization of specific proteins was studied with immunofluorescence labeling using specific antibodies.

Results

***A. muciniphila* stimulates species-specific cytokine patterns when compared to other commensals.** Stimulation of peripheral blood mononuclear cells (PBMCs) with *A. muciniphila* MucT resulted in induction of both anti- and proinflammatory cytokines (IL-1ß, IL-6, IL-8, IL-10 and TNF-α). This induction was seen for live cells as well as for heat-killed cells and supernatant (Table 1). Among the measured cytokines, IL-10, IL-8, IL-6 and TNF-α were the highest induced.

A. muciniphila immune stimulation in PBMCs was compared to two other established beneficial gut microbes, *Faecalibacterium prausnitzii* A2-165 and *Lactobacillus plantarum* WCFS1. The stimulation of PBMCs led to a microbe-specific pattern of all tested microorganisms. In comparison to *F. prausnitzii*, *A. muciniphila* induced higher levels of IL-6 and IL-8, while levels of IL-10 were similar and levels of TNF-α were lower (Table 1, Figure 1A-D). Compared to *L. plantarum*, *A. muciniphila* induced higher levels of IL-6, IL-8 and IL-10 but less TNF-α (Table 1, Figure 1A-D). PBMCs were also stimulated with the supernatants of these bacteria leading to a markedly different cytokine response for IL-10 and TNF-α, whereas IL-8 and IL-6 showed more similar patterns compared to live bacteria (Figure 1E-H). Remarkably, *A. muciniphila* supernatant induced significantly more IL-10 than *F. prausnitzii* supernatant and more IL-10 and TNF-α than *L. plantarum* supernatant (Figure 1E, F).

Table 1. Effect of bacteria and bacterial fractions on cytokine production of human PBMCs.

Cytokine (pg/ml)	Live *A. muciniphila*	Heat-killed *A. muciniphila*	Supernatant *A. muciniphila*	Live *F. prausnitzii*	Live *L. plantarum*
IL-1ß	894 ± 298	392 ± 71	1650 ± 510	870 ± 301	894 ± 298
IL-6	18029 ± 309	13477 ± 2014	23225 ± 8102	10178 ± 2648	7028 ± 2812
IL-8	60018 ± 18229	54230 ± 9030	86171 ± 32298	48354 ± 10526	33085 ± 16760
IL-10	823 ± 310	638 ± 118	1221 ± 310	834 ± 388	333 ± 215
TNF-α	1920 ± 349	957 ± 568	2095 ± 1249	5545 ± 2615	5459 ± 2830
IL-12p70	< 2	< 2	< 2	< 2	253 ± 293
Supernatant filtrates					
Cytokine (pg/ml)	10 kDa filtrate	30 kDa filtrate	100 kDa filtrate	300 kDa filtrate	1000 kDa filtrate
IL-1ß	< 8	< 8	62 ± 29	678 ± 350	949 ± 522
IL-6	22 ± 23	569 ± 430	2556 ± 1049	18697 ± 7097	18717 ± 7326
IL-8	4321 ± 4530	11928 ± 7128	29040 ± 14973	74208 ± 28348	73695 ± 36442
IL-10	< 4	6 ± 4	43 ± 28	937 ± 287	1215 ± 379
TNF-α	< 4	37 ± 48	247 ± 255	1476 ± 849	1188 ± 930
IL-12p70	< 2	< 2	< 2	< 2	< 2
Sucrose density-gradient separated fractions and purified Amuc_1100					
Cytokine (pg/ml)	Fraction 1 (4.5 µg/ml)	Fraction 2 (4.5 µg/ml)	Amuc_1100 (4.5 µg/ml)		
IL-1ß	437 ± 225	< 8	504 ± 227		
IL-6	9312 ± 3329	414 ± 364	12508 ± 2362		
IL-8	64877 ± 19528	13339 ± 5039	45432 ± 12507		
IL-10	941 ± 404	21 ± 18	526 ± 180		
TNF-α	2165 ± 883	60 ± 35	1317 ± 885		
IL-12p70	< 2	< 2	< 2		

Data are presented as mean ± SD, n = 3 donors

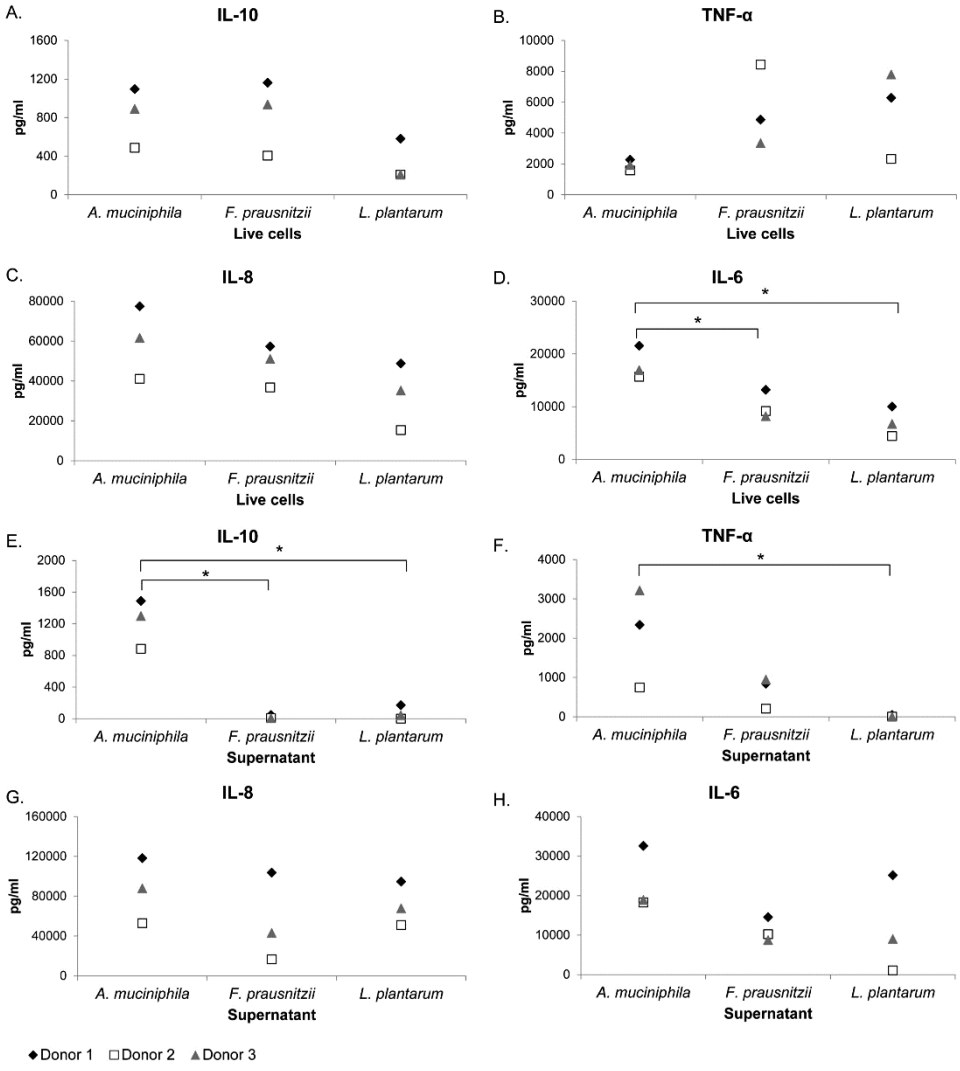


Figure 1. Effect of *A. muciniphila*, *F. prausnitzii* and *L. plantarum* on cytokine production of human PBMCs. IL-10 (A), TNF-α (B), IL-8 (C) and IL-6 (D) responses of human PBMCs (n = 3 donors) stimulated with *A. muciniphila*, *F. prausnitzii* and *L. plantarum* live cells. IL-10 (E), TNF-α (F), IL-8 (G) and IL-6 (H) responses of human PBMCs (n = 3 donors) stimulated with *A. muciniphila*, *F. prausnitzii* and *L. plantarum* supernatant. *, $P<0.05$.

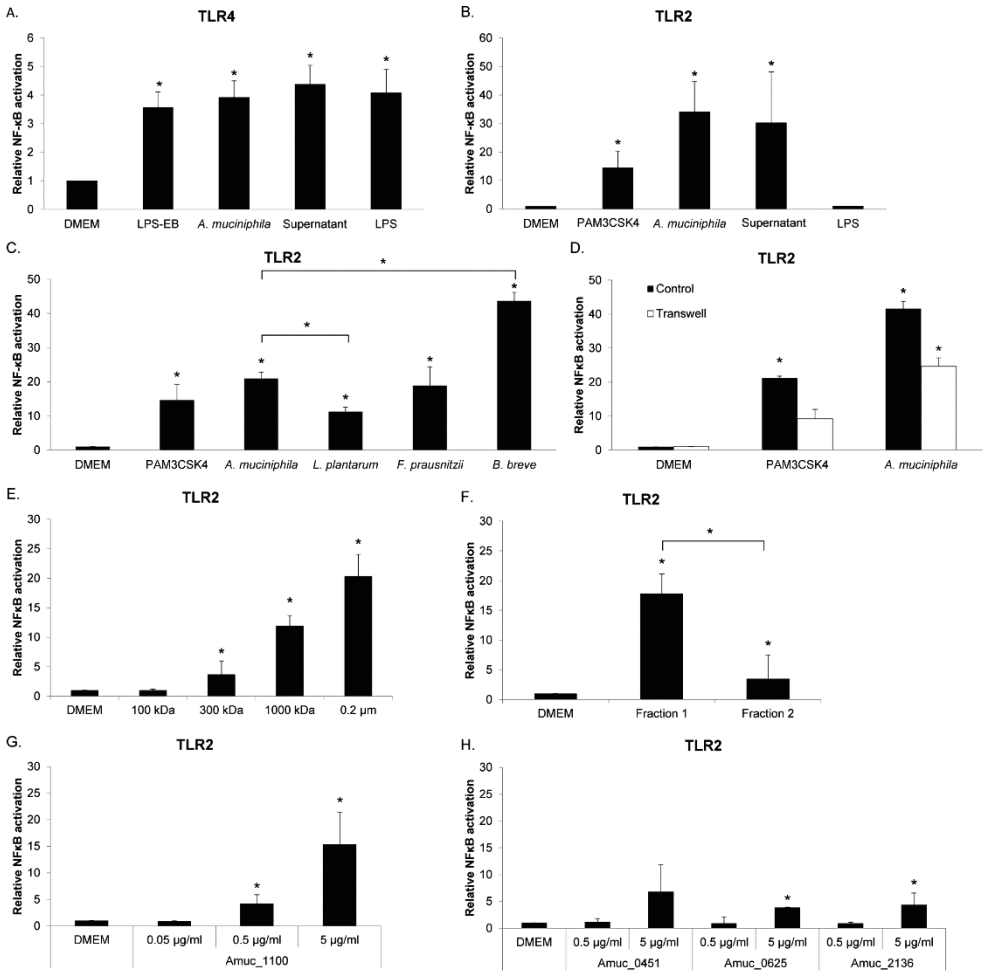


Figure 2. *A. muciniphila* activates signaling pathways through TLR2 and TLR4. (A) TLR4 signaling by live *A. muciniphila* (~10^7 bacteria/well), *A. muciniphila* supernatant and LPS isolated from *A. muciniphila* (concentration corresponds to amount of LPS in ~10^7 *A. muciniphila* cells). DMEM: medium control, LPS-EB: positive control (concentration corresponds to amount of LPS in ~10^7 *E. coli* cells). (B) TLR2 signaling by live *A. muciniphila* (~10^7 bacteria/well), *A. muciniphila* supernatant and LPS. DMEM; medium control, PAM3CSK4; positive control. (C) TLR2 signaling by live *A. muciniphila*, *L. plantarum*, *F. prausnitzii* and *B. breve* (~10^6 bacteria/well). (D) TLR2 signaling in a Transwell system compared to control (i.e. samples not separated from the cell line by a membrane). (E) TLR2 signaling by filtrated supernatant signaling molecules. (F) TLR2 signaling by *A. muciniphila* bacterial fractions (1 µg of protein/well). (G) TLR2 signaling by *A. muciniphila* purified protein Amuc_1100 (0.01, 0.1 and 1 µg of protein/well). (H) TLR2 signaling by *A. muciniphila* purified proteins Amuc_0451, Amuc_0625 and Amuc_2136 (0.1 and 1 µg of protein/well). DMEM; medium control, *, $P<0.05$ compared to DMEM. See also Figure S1.

***A. muciniphila* activates the NF-κB pathway through TLR4 and TLR2 receptors.** To determine which intestinal receptors are involved in immune stimulation of *A. muciniphila*, reporter cell lines expressing TLR2, TLR4, TLR5, TLR9 or NOD2 receptors were employed. The strongest activation was seen on TLR4 (Figure 2A) and TLR2 (Figure 2B). *A. muciniphila* did not activate TLR5 and TLR9, and only minor activation was seen for NOD2 (data not shown).

TLR2 responses were higher in the presence of *A. muciniphila* compared to *L. plantarum* (Figure 2C), but there was no significant difference between the TLR2 response induced by *F. prausnitzii* and *A. muciniphila* (Figure 2C). However, the TLR2 response towards *A. muciniphila* was lower than the induction by the Gram-positive *Bifidobacterium breve* DSM-20213 (Figure 2C).

TLR4 is an important receptor for recognizing Gram-negative sensitive lipopolysaccharide (LPS). We first verified the presence of LPS in *A. muciniphila* by using a protocol to extract LPS from bacterial cells and applying silver staining to visualize it on a gel (Figure S1). In these reporter cell line-experiments both live bacteria and *A. muciniphila* LPS significantly stimulated NF-κB dependent secreted embryonic alkaline phosphatase (SEAP) production via TLR4 (Figure 2A). On top of this, *A. muciniphila* LPS induced production of IL-8, IL-6 and low amounts of IL-10 and TNF-α in PBMCs (data not shown). As expected, *A. muciniphila* LPS did not induce a TLR2 response (Figure 2B).

30 kDa outer membrane protein (Amuc_1100) is a strong TLR2 activator and induces cytokines in PBMCs. *A. muciniphila* supernatant activated TLR2 as the NF-κB activity persisted while bacteria were separated from the cell line by a membrane in a Transwell assay (Figure 2D). This indicates *A. muciniphila* can activate TLR2 with both cell derived fragments and extracellular molecules. The main fermentation products in the supernatant of *A. muciniphila* are acetate and propionate, but no effect on NF-κB activity was observed for these fatty acids at a concentration of 1 mM (data not shown). Using centrifugal membrane filters of differing pore sizes, we could demonstrate that for TLR2 and cytokine induction the size of signaling molecules from the supernatants had to be larger than 100 kDa (Figure 2E). The production of IL-8, IL-1β, IL-6, IL-10 and TNF-α by PBMCs increased along with the molecule size of the filtrated supernatant (Table 1.)

Next we separated bacterial membranes from intracellular proteins with sucrose density-gradient centrifugation. LC-MS/MS-analysis was done on four different fractions along the gradient to identify the proteins and determine their relative amounts. Two samples (Fraction 1 & 2) separated with sucrose density-gradient centrifugation were found to differ notably in the relative amount of cell envelope

proteins. In total 117 cell envelope proteins were detected in these fractions, with 84 of them exclusively found in Fraction 1 (Figure 3A). Fraction 1 was especially enriched for proteins involved in protein transport and secretion, in comparison to Fraction 2. Fraction 2 had a more heterogeneous protein content based on protein function, but overall lower diversity of proteins. Among the most abundant proteins in Fraction 2 were mucin-degrading enzymes (glycosyl hydrolase, beta-galactosidase, N-acetylhexosaminidase) and other enzymes (alanine-glyoxylate transaminase, hyaluronon glucosaminidase). These fractions were also tested on the TLR2 cell line and PBMCs in equal protein concentrations. Fraction 1 induced higher TLR2 activity than Fraction 2 (Figure 2F). Fraction 1 also induced higher cytokine production in PBMCs as compared to Fraction 2 (Table 1).

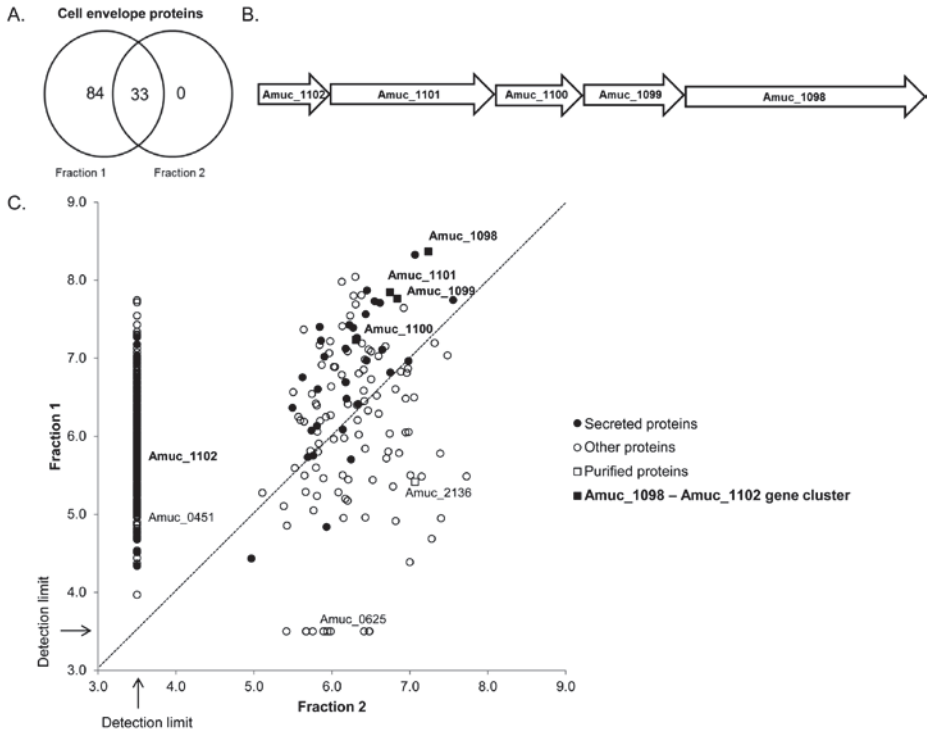


Figure 3. **Proteins encoded by the gene cluster Amuc_1098 to Amuc_1102 are found abundantly in a fraction enriched for membrane and cell-envelope proteins.** (A) Number of *A. muciniphila* membrane and cell-envelope proteins detected with LC-MS/MS in two different fractions from a sucrose density-gradient separation method. (B) Amuc_1098 to Amuc_1102 gene cluster. (C) Abundance of proteins found in Fraction 1 vs. Fraction 2. Relative abundances of the proteins are presented on a log10 scale. A log10 relative abundance of 3.5 represents proteins that were not detected or were under the detection limit.

Based on proteomics analysis with LC–MS/MS we discovered that the relative abundance of all proteins that were encoded by a gene cluster corresponding to locus tags Amuc_1098 to Amuc_1102 (Figure 3B), was at least ten times higher in Fraction 1 compared to Fraction 2 (Figure 3C). Amuc_1098 is predicted to encode a type II and type III secretion system protein and Amuc_1101 is predicted to encode a cell division protein FtsA. The other three genes (Amuc_1099, Amuc_1100, Amuc_1101) are annotated as hypothetical proteins. All the genes, except for Amuc_1101, have a signal sequence at the N terminus, indicating they are destined towards the secretory pathway. We used immunofluorescence labeling to localize the proteins in the bacteria, and successfully identified Amuc_1100 (32 kDa) as an outer membrane protein (Figure 4B). However, some variation was observed in the intensity of the labeling of the *A. muciniphila* cells with the anti-Amuc_1100 antibodies (Figure 4B). Similarly, the labeling intensity also varied, but in a less extensive way, with the whole cell antibody (Figure 4A).

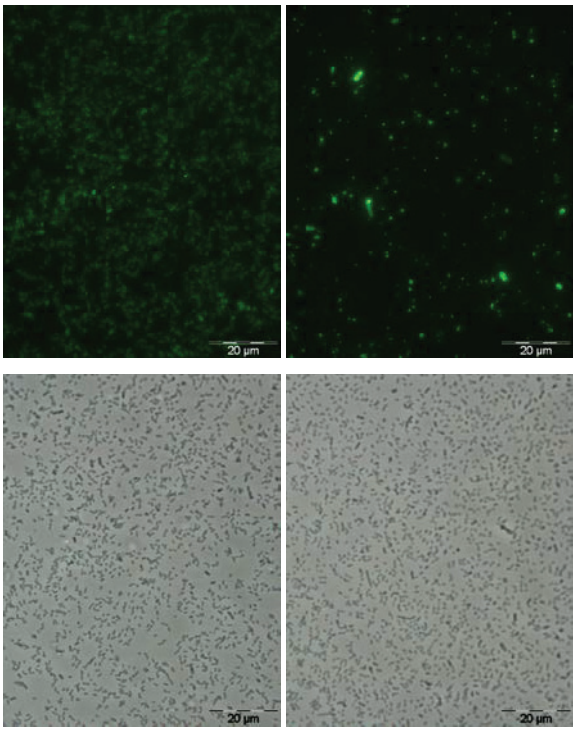


Figure 4. Amuc_1100 is located on the outer membrane of *A. muciniphila*. Immunofluorescence staining of *A. muciniphila* cells with whole cell antiserum (A) or anti-Amuc_1100 (B) and Alexa-488-conjugated secondary IgG. Phase-contrast images of the same microscopic fields are shown below.

Purified recombinant Amuc_1100 protein gave specific induction of TLR2 (Figure 2G) and was able to induce IL-1ß, IL-6, IL-8, IL-10 and TNF-α production in PBMCs (Table 1). As the genes Amuc_1101 and Amuc_1102 did not lead to overproduced soluble proteins, we were not able to test the effect of these proteins on the immune response. As a control to Amuc_1100, secreted enzymes that are involved in *A. muciniphila* mucin degradation were tested. It was found that the enzymes Amuc_0451 (sulfatase), Amuc_0625 (exo-alpha-sialidase), and Amuc_2136 (glycoside hydrolase) are abundantly produced by *A. muciniphila* and dose-dependent TLR2-signaling was detected for each of them (Figure 2H). However, the TLR2-signaling response tended to be lower for these periplasmic enzymes in comparison to Amuc_1100.

Additionally, we tested the impact of *A. muciniphila* and the purified recombinant proteins on the development of the integrity of epithelial cell layer by determining the TER of Caco-2 monolayers. *A. muciniphila* showed a significantly increased TER after 24 h of co-cultivation with the Caco-2 cells as reported previously (Reunanen *et al.*, 2015) (Figure 5). In our assay also Amuc_1100 showed a significantly increased TER at 24 h (Figure 5), whereas the other purified proteins (Amuc_2136, Amuc_0625 and Amuc_0451) did not significantly increase TER (data not shown). *E. coli* is known to have adverse effects on epithelial cell monolayer integrity and decreased the TER (Figure 5) (Geens and Niewold, 2010).

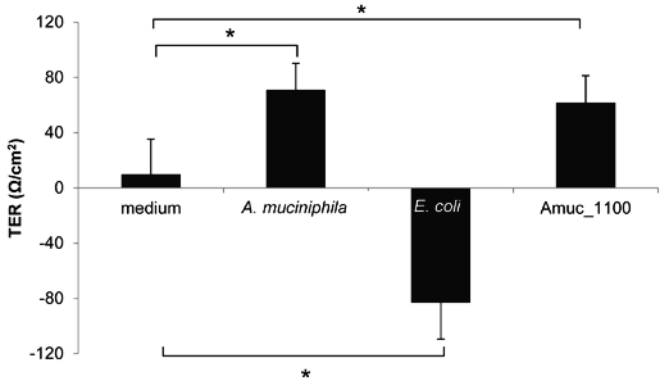


Figure 5. *A. muciniphila* and outer membrane protein Amuc_1100 increase the development of transepithelial electrical resistance. The impact of *A. muciniphila*, purified protein Amuc_1100 (0.05 µg/ml) or *E. coli* on the TER development of Caco-2 monolayer after 24 h of stimulation. Mean and standard deviations from three parallel wells are shown. Significant differences ($p < 0.05$) in the TER values as compared to control (growth medium without bacteria) at 24 h are indicated with an asterix.

Discussion

A. muciniphila MucT was capable of inducing a wide range of immunomodulatory responses *in vitro*. The immune modulatory capacity was traced back to large molecule complexes with a molecular size of over 100 kDa. Subsequently, we showed that a bacterial fraction enriched in cell envelope proteins, presumably including large structures exposed to the surface of the bacteria, induced high TLR2 signaling and cytokine production in PBMCs. We showed this fraction to be highly enriched in a set of proteins encoded by the gene cluster Amuc_1098 - Amuc_1102 that could constitute the fimbriae-like structures observed in electron microscope images (Derrien *et al.*, 2004). Amuc_1100 is part of this gene cluster, and its 30-kDa product could be overproduced in *E. coli* and found to induce production of IL-6, IL-8 and IL-10 in PBMCs. Moreover, using immunofluorescence microscopy we could localize the Amuc_1100 protein at the outside of *A. muciniphila* cells, compatible with its location in fimbriae-like structures. Some heterogeneity in the immunolabeling was observed, possibly reflecting different amounts of fimbriae per cell. As *A. muciniphila* is located in the mucus layer, not far from the epithelial cells, it may benefit from these types of appendages when interacting with the host.

In addition, a few secreted mucin-degrading enzymes also activated PBMCs and TLR2. Previously it was reported that *A. muciniphila* extracellular vesicles (EV) stimulate IL-6 secretion, but pre-treatment of a colon epithelial cell line with these EV before treating them with *E. coli* EV, decreased IL-6 production as compared to treatment with *E. coli* EV alone (Kang *et al.*, 2013). It is unclear which proteins are present in the *A. muciniphila* EV that trigger the response, and whether mucin fragments from the culture medium may have confounded the results. In our study *A. muciniphila* was grown on non-mucus medium for all the immune assays to prevent compounds of hog gastric mucin from interfering with the immune response.

IL-8, IL-6, IL-1ß, IL-10 and TNF-α were induced by *A. muciniphila* in human derived PBMCs, indicating it cannot be strictly defined as anti- or proinflammatory, but may instead have a more complex role in preserving the balance of the gut ecosystem. Interestingly, live bacteria induced significantly higher IL-1ß production than heat-killed bacteria, suggesting that cell derived fragments as well as secreted compounds have an important role in the immune signaling. The immunomodulatory outer membrane structures may have also been damaged during the heat treatment. Stimulation of PBMCs with *A. muciniphila* led to production of proinflammatory IL-8. However, a recent study showed that *A. muciniphila* induced IL-8 production in enterocytes at 100-fold higher cell

concentrations as compared to *E. coli*, suggesting a fairly low inflammatory potential in the gut (Reunanen *et al.*, 2015).

The immune modulatory effects observed for *A. muciniphila* were distinct from the microbiota members *F. prausnitzii* A2-165 and *L. plantarum* WCFS1. Differential host response towards *A. muciniphila* and *F. prausnitzii* also became apparent in a study conducted with mouse derived organoids (Lukovac *et al.*, 2014). Therein, *A. muciniphila* triggered mostly regulation of metabolic markers as compared to *F. prausnitzii*. In the model systems used in the present study *A. muciniphila* and *F. prausnitzii* live cells induced very similar amounts of IL-10. Interestingly, *A. muciniphila* supernatant induced high amounts of IL-10, whereas *F. prausnitzii* supernatant barely stimulated any IL-10 production. This would indicate greater anti-inflammatory capacity for *A. muciniphila* secreted products and metabolites in comparison to *F. prausnitzii*. As Amuc_1100 was shown to be located on the outer membrane of *A. muciniphila*, it is plausible that this protein is also liberated to the supernatant and therefore contributes to stimulation of IL-10. Another possibly important difference between the secreted compounds produced by these two bacteria is the production of mucin-degrading enzymes, which are abundantly present in *A. muciniphila*, but not in *F. prausnitzii*.

The cytokine response of human PBMCs was consistently lower for *L. plantarum* compared to *A. muciniphila*, except for induction of TNF-α. This is in line with the comparison of mouse transcriptional response to colonization with *A. muciniphila* or *L. plantarum*, which revealed that *A. muciniphila* induces relatively higher up-regulation of genes participating in immune response signaling and ERK/MAPK signaling (Derrien *et al.*, 2011). Despite greater immune response, the *A. muciniphila*-colonized mice did not develop microscopically visible inflammation or show any sign of discomfort. The diverse immune stimulation caused by *A. muciniphila* could thus be an indication of gut immune tolerance towards this commensal. The differences between *A. muciniphila* and *L. plantarum* were also demonstrated in obese and type 2 diabetic mice, where treatment with *A. muciniphila* improved the metabolic outcome, whereas treatment with *L. plantarum* did not at all have this beneficial effect (Everard *et al.*, 2013).

The differential immune response between *A. muciniphila*, *F. prausnitzii* and *L. plantarum* could be an indication of their physiological or metabolic differences, or the fact that in the gut these bacteria colonize separate niches. As a mucin-degrader, *A. muciniphila* is in closer contact with the host, in contrast to bacteria colonizing the lumen. Our study further strengthens the findings made in the aforementioned studies (Derrien *et al.*, 2011; Everard *et al.*, 2013; Lukovac *et al.*, 2014) on the variances of host response between these commensal bacteria.

A. muciniphila LPS gave a strong response with TLR4 and is most likely the activating molecule for this receptor in *A. muciniphila*. Recently it has been reported that the position of the phosphate in the lipid A of bacterial LPS may play an important role in separating bacterial-host innate immune system interactions into either symbiotic or pathogenic relationships (Coats *et al.*, 2011). Determining the molecular structure of *A. muciniphila* LPS would be valuable in understanding its immunostimulatory role in the gut. Previously, it was proposed that LPS of *A. muciniphila* is inflammatory in a model of experimental alcoholic liver disease in mice, as the levels of *A. muciniphila* were higher after chronic intragastric alcohol feeding, and lower in $Muc2^{-/-}$ mice, along with lower plasma LPS concentration (Hartmann *et al.*, 2013).

TLR2 is best known for recognizing lipoteichoic acid (LTA) from Gram-positive bacteria, but some Gram-negative bacteria containing non-classical LPS have also been shown to signal through TLR2 (Alhawi *et al.*, 2009; Erridge *et al.*, 2004). As the lipid structure of *A. muciniphila* membranes is not characterized in detail, we evaluated the ability of *A. muciniphila* LPS to stimulate TLR2. Even though live bacteria and the supernatant induced a strong TLR2 response, purified LPS did not.

A. muciniphila did not activate TLR5, the intestinal receptor for flagellin, reflecting the notion that there are no flagellin genes found in the genome of *A. muciniphila* (van Passel *et al.*, 2011). Only high concentrations of *A. muciniphila* (10^8 bacteria/well) induced a minor response of the TLR9 receptor, which recognizes unmethylated CpG sequences in DNA molecules. Isolated *A. muciniphila* DNA did not induce any TLR9 response. The genome of *A. muciniphila* has a lower than median frequency (255 vs. 401) of GTCGTT hexamers, in comparison with 59 other bacterial species (Kant *et al.*, 2014), which may explain the observed low activation of TLR9. Another explanation may be the intracellular location of TLR9, which makes it more difficult for the ligands to reach it, especially in an *in vitro* setting.

The localization of *A. muciniphila* in the mucus layer, close to the epithelial layer, probably has had a great impact on the mechanisms of immune modulation this bacterium has developed. As *A. muciniphila* is diminished in many inflammatory diseases, it could be speculated that the absence of *A. muciniphila* in case of inflammation prevents immune suppression at the mucosal epithelial border. Cross-talk between *A. muciniphila* and the host might affect immunological tolerance and homeostasis within the gut, possibly by keeping the immune system alert for potential disruptions. *A. muciniphila* has been shown to restore mucus layer thickness and to increase intestinal endocannabinoids in diet-induced obese mice (Everard *et al.*, 2013), suggesting at the same time it contributes to improving

gut barrier function. Here we showed that *A. muciniphila* and the outer membrane protein Amuc_1100 increased the development of TER in Caco2-cells, which also indicates strengthening of the epithelial barrier function.

Altogether, these results could partially explain the positive correlation between levels of *A. muciniphila* and gut health. Several studies have reported depletion of *A. muciniphila* in the fecal microbiota of ulcerative colitis patients, both in remission (James *et al.*, 2014; Rajilic-Stojanovic *et al.*, 2013) and in clinically active disease (Vigsnaes *et al.*, 2012). In addition to fecal microbiota, reduced levels of *A. muciniphila* have also been found in biopsies of intestinal mucosa from IBD-patients in comparison to healthy controls (Png *et al.*, 2010). The *in vitro* results presented here should be confirmed in an *in vivo* model, and special focus should be put on the identified proteins to verify their role in *A. muciniphila* immune signaling.

In conclusion, this study revealed the diverse immunostimulatory capacities of *A. muciniphila* and identified candidate bacterial products that mediate this stimulation. We have also shown that established and next-generation probiotics have a wide range of species-specific immune stimulatory properties, which should be taken into consideration when developing new applications and interventions.

Experimental procedures

Bacterial growth conditions. *Akkermansia muciniphila* MucT (ATTC BAA-835) was grown in a basal medium as described previously (Derrien *et al.*, 2004). The medium was supplemented with either hog gastric mucin (0.5 %, Type III; Sigma), a mix of sugars (D-glucose, L-fucose, N-acetylglucosamine, N-acetylgalactosamine; 2.5 mM each) or glucose (10 mM). The medium without mucin was supplemented with tryptone (8 g/l) and L-threonine (2 mM). Incubations were performed in serum bottles sealed with butyl-rubber stoppers at 37ºC under anaerobic conditions provided by a gas phase of 182 kPa (1.5 atm) N_2/CO_2 (80/20 ratio). Growth was measured by spectrophotometer as optical density at 600 nm (OD600).

Faecalibacterium prausnitzii A2-165 was grown anaerobically at 37ºC in YCFA medium supplemented with 33 mM glucose (Duncan *et al.*, 2002). *Lactobacillus plantarum* WCFS1 was grown aerobically and *Bifidobacterium breve* DSM-20213 anaerobically at 37 ºC in Difco™ Lactobacilli MRS broth (Becton, Dickinson and Company, Sparks, USA).

Peripheral blood mononuclear cells assay. Peripheral blood of three healthy donors was received from the Sanquin Blood Bank, Nijmegen, The Netherlands. Peripheral blood mononuclear cells (PBMCs) were separated from the blood of healthy donors using Ficoll-Paque Plus gradient centrifugation according to the manufacturer's protocol (Amersham biosciences, Uppsala, Sweden). After centrifugation the mononuclear cells were collected, washed in Iscove's Modified Dulbecco's Medium (IMDM) + Glutamax (Invitrogen, Breda, The Netherlands) and adjusted to 0.5×10^6 cells/ml in IMDM + Glutamax supplemented with penicillin (100 U/ml) (Invitrogen), streptomycin (100 µg/ml) (Invitrogen), and 10% heat inactivated Fetal Bovine Serum (FBS, Lonza, Basel, Switzerland). PBMCs (0.5×10^6 cells/well) were seeded in 48-well tissue culture plates.

Per donor a negative control (just medium), and two positive controls (LPS from *E. coli* 1 µg/ml) and (LTA from *Staphylococcus aureus* 5 µg/ml) in duplicate were used. PBMCs were stimulated with live bacteria or bacterial fractions. For the heat-killed cells the bacterial culture was kept at 99°C for 10 min. The ratio of PBMC to bacteria was 1:10. Cells were incubated for 24 hours and the culture supernatants were collected for cytokine analysis. Cytokine levels of IL-6, IL-8, IL-10, TNF-α, IL-1ß and IL-12p70 were measured using multiplex analysis (Human inflammation CBA kit, Becton and Dickinson) according to the manufacturer's protocol on a FACS CantoII (Becton Dickinson) and analyzed using BD FCAP software (Becton Dickinson). The detection limits according to the manufacturer were as follows: 3.6

pg/ml IL-8, 7.2 pg/ml IL-1ß, 2.5 pg/ml IL-6, 3.3 pg/ml IL-10, 3.7 pg/ml TNF-α, 1.9 pg/ml IL-12p70.

***In vitro* culture and stimulation of human HEK-Blue hTLR2/4/5/9/NOD2 cell lines.** For the inflammatory pathway analysis HEK-Blue hTLR2, hTLR4, hTLR5, hTLR9 and hNOD2 cell lines (Invivogen, CA, USA) were used. Stimulation of the receptors with the corresponding ligands activates NF-κB and AP-1, which induces the production of secreted embryonic alkaline phosphatase (SEAP), the levels of which were measured by spectrophotometer (Spectramax). All cell lines were grown and subcultured up to 70–80% of confluency using as a maintenance medium Dulbecco's Modified Eagle Medium (DMEM) supplemented with 4.5 g/l D-glucose, 50 U/ml penicillin, 50 µg/ml streptomycin, 100 µg/ml Normocin, 2 mM L-glutamine, and 10% (v/v) of heat-inactivated FBS. For each cell line, an immune response experiment was carried out by seeding HEK-blue cells in flat-bottom 96-well plates and stimulating them by addition of 20 µl bacterial suspensions. The 96-well plates were incubated for 20–24 h at 37°C in a 5% CO2 incubator. Receptor ligands Pam3CSK4 (10 ng/ml for hTLR2), LPS-EB (50 ng/ml for hTLR4), RecFLA-ST (0.1 ng/ml for hTLR5), ODN 2006 (50 µM for hTLR9) and L18-MDP (0.1 ng/ml for hNOD2) were used as positive control whereas maintenance medium without any selective antibiotics was used as negative control. SEAP secretion was detected by measuring the OD600 at 15 min, 1 h, 2 h, and 3 h after addition of 180 µL of QUANTI-Blue (Invivogen, CA, USA) to 20 µL of induced HEK-Blue hTLR2/4/5/9/NOD2 supernatant. The data shown here are from the 1 h measurement point. Experiments for TLR2 and TLR4 were performed in triplicate, and experiments for TLR5, TLR9 and NOD2 in duplicate.

Transwell assay. To study the TLR2 signaling activity of secreted molecules, the bacteria were separated from the cell line using Transwell (Corning, USA) cell culture membrane inserts (0.4 µm pore size). Bacterial suspension (100 µl) was added either directly to wells containing HEK-TLR2 cells or first to the Transwell insert compartment, which was then inserted into the wells with HEK-TLR2 cells. The plates were incubated and SEAP secretion measured as described above.

Bacterial filtrates. To study the size of the signaling molecules, the bacterial supernatant was filtered using filters of different pore sizes and molecular mass cut off sizes. The following filters were used: 0.45 µm and 0.2 µm polyethersulfone syringe filters (Advanced Microdevices, Ambala Cantt., India), 1000 kDa Vivaspin 20 Polyethersulfone ultrafiltration unit (Sartorius, Goettingen, Germany), 3K and 300K Pall Nanosep® centrifugal device with Omega membrane (Pall corporation, Ann Arbor, Michigan), 10K Amicon® Ultra regenerated cellulose centrifugal filter (Merck Millipore Ltd., Cork, Ireland), Vivaspin 500 with 30,000 MWCO

(Polyethersulfone) Membrane Concentrator (Vivascience, Sartorius Group, Hannover, Germany). 500 µl of supernatant was passed through the filter and the filtrate was used in the assays.

Bacterial fractionation method. The membranes of *A. muciniphila* were isolated from cultures grown with glucose as the carbon source with sucrose density-gradient centrifugation, as described previously (Hobb *et al.*, 2009). The samples were stored in 2 ml low binding tubes (Eppendorf, Hamburg, Germany) at −20°C. Qubit® Protein Assay Kit (Life technologies, Oregon, USA) was used according to the manufacturer's instructions to determine the protein content of cell extracts. Samples were loaded on a 10 % acrylamide separation gel (Product number 25201, Precise™ Protein Gels, Thermo Scientific, Rockford, IL, USA) using the mini-PROTEAN 3 cell (Bio-Rad Laboratories, Hercules, CA, USA). The electrophoresis procedure was according to the manufacturer's instructions. Gels were stained using CBB R-250 as indicated in the protocol of the mini-PROTEAN 3 system.

In-gel digestion of proteins and purification of peptides were done following a modified version of the protocol described by Rupakula et al., 2013. Disulfide bridges in proteins were reduced by covering whole gels with reducing solution (10 mM dithiothreitol, pH 7.6, in 50 mM NH_4HCO_3), and the gels were incubated at 60°C for 1 h. Alkylation was performed for 1 h by adding 25 mL of iodoacetamide solution (10 mM iodoacetamide in 100 mM Tris-HCl, pH 8.0). Gels were thoroughly rinsed with dd H_2O water in between steps. Each lane of SDS-PAGE gels was cut into one slice, and the slices were cut into approximately 1 mm^3 cubes and transferred to separate 0.5 ml protein LoBind tubes (Eppendorf, Hamburg, Germany). Enzymatic digestion was done by adding 50 µl of trypsin solution (5 ng/µl trypsin in 50 mM NH_4HCO_3) to each tube, and by incubating at room temperature overnight with gentle shaking. Extraction of peptides was performed with manual sonication in an ultrasonic water bath for 1 s before the supernatant was transferred to a clean protein LoBind tube. Trifluoroacetic acid (10 %) was added to the supernatant to reach a pH between 2 and 4. The supernatant was used for LC-MS/MS analysis. Samples were measured by nLC–MS/MS with a Proxeon EASY nLC and a LTQ-Orbitrap XL mass spectrometer as previously described (Lu *et al.*, 2011).

LC–MS data analysis was performed as described previously (Rupakula *et al.*, 2013), with false discovery rates (FDRs) set to 0.01 on peptide and protein level, and additional result filtering (minimally 2 peptides necessary for protein identification of which at least one is unique and at least one is unmodified). To analyze the abundance of proteins in the fractions, their label-free quantification

(LFQ) intensities were compared (Cox *et al.*, 2014). Non-existing LFQ intensity values due to not enough quantified peptides were substituted with a value lower than the LFQ intensity value for the least abundant, detected peptide.

Plasmid constructs and protein production. The genes Amuc_0451, Amuc_0625, Amuc_1100, and Amuc_2136 were amplified by PCR without their signaling sequence, with primers as specified in supplemental experimental procedures. PCR products of the genes were cloned into either pET-24d or pET-26b vectors (Novagen®, Merck Millipore, MA, USA). For genes Amuc_0451, Amuc_0625, and Amuc_1100 pET-26b was used with restriction sites NdeI and XhoI. For Amuc_2136 gene, pET-24d was used and the PCR product containing PciI and XhoI restriction sites was cloned at the NcoI and XhoI sites of the vector.

E. coli XL1Blue or TOP10 cells were transformed with constructed plasmids by electroporation or heat shock, respectively. Cells with kanamycin resistance were selected by plating the transformed cells on Luria agar plates containing 50 μg/ml kanamycin. Plasmids isolated from colonies on these plates were checked for having the right insert length by PCR and subsequently, isolated plasmids were sequenced to confirm the right insert.

E. coli BL21(DE3) cells were transformed with the right plasmid for protein expression. Luria broth containing kanamycin (50 μg/ml) was inoculated with overnight culture and grown with shaking at 220 rpm at 37°C until exponential phase, and protein production was induced by adding IPTG up to 1 mM. After three hours of induction, cells were pelleted by centrifuging 10 min at 5000 g and cell pellets stored at –20°C until lysis.

Cell pellets were resuspended and lysed using lysozyme and sonification (Sonifier 450, Branson Ultrasonics Corporation, Danbury, CT). Supernatant was collected after centrifugation and proteins were His-tag purified by metal affinity purification under native conditions using Ni-NTA His•Bind Resin (Novagen®, Merck Millipore, MA, USA). Elution buffer was exchanged for a 50 mM Tris-HCl, 50 mM NaCl, pH of 7.4 buffer by using 5 ml resin bed Zeba spin columns (Pierce, Rockford, IL, USA). After buffer exchange, protein content was measured by BCA assay (Pierce, Rockford, IL, USA) and proteins were stored at -20°C.

Immunofluorescence microscopy. Rabbit antibodies were raised against the purified recombinant Amuc_1100 protein and these were used in immunomicroscopic analysis of its location in *A. muciniphila.* Total antibodies raised against *A. muciniphila* whole cells were used as a control. The immunization

was done in Eurogentec (Seraing, Belgium) and the Laboratory Animal Centre of University of Helsinki as described previously (Kainulainen *et al.*, 2013).

Immunofluorescence staining was used to confirm the presence of Amuc_1100 on the surface of *A. muciniphila* as described previously (Kainulainen *et al.*, 2012). Briefly, *A. muciniphila* cells were cultivated for 24 h with glucose as the carbon source, washed with phosphate-buffered saline (PBS), and fixed with 3.5% (w/v) paraformaldehyde in PBS prior to labeling with *A. muciniphila* whole cell antiserum or anti-Amuc_1100 pre-immune serum as the primary antiserum and Alexa-488 (Invitrogen)-conjugated anti-rabbit IgG (1 µg/ml) as the secondary antibody. Bacteria were then examined with an epifluorescence microscope (Leica DM 4000B) equipped with a filter for the Alexa-488 label (excitation, 450 to 490 nm; emission, 515 nm), and images were digitally recorded using CellP imaging software for life sciences microscopy (Soft Imaging System GmbH).

Extraction of LPS. *A. muciniphila* LPS was extracted using the hot phenol-water extraction method as described previously (Zhang and Skurnik, 1994), with minor modifications. Briefly, bacterial cells from 5 ml overnight cultures were collected by centrifugation, washed once with water and resuspended into 500 µl of ultrapure water. The bacterial suspensions were warmed up at 65°C and then mixed with an equal volume of water-saturated phenol preheated to 65°C. The mixture was incubated at 65°C for 10 min and then transferred to ice to cool down. After centrifugation at 4°C for 5 min, the aqueous layer was carefully transferred to a new Eppendorf tube and the incubation with an equal volume of hot phenol was repeated twice. After this two volumes of acetone were added to the aqueous layer to precipitate LPS. The suspension was incubated at -20°C for two hours after which it was centrifuged at 4°C for 10 minutes and the pellet was dissolved in 50 µl LPS-free water. The quality of the LPS was checked by silver staining (Figure S1.) and quantity of the LPS was checked by EndoLISA® Endotoxin Detection Assay (Hyglos GmbH, Bernried am Starnberger See, Germany) according to manufacturer's protocol.

Transepithelial electrical resistance (TER) assay. Caco-2 cells (5 x 10^4 cells/insert) were seeded in Millicell cell culture inserts (3 µm pore size; Millipore) and grown for 8 days. The growth conditions of Caco-2 cells were as described previously (Kainulainen *et al.*, 2015). Bacterial cells were washed once with RPMI 1640, and applied onto the inserts at OD600 of 0.25 in RPMI 1640. Purified *A. muciniphila* proteins Amuc_0451, Amuc_0625, Amuc_1100, and Amuc_2136 were applied onto the inserts at concentrations of 0.05, 0.5 and 5 µg/ml. The transepithelial resistance was determined with a Millicell ERS-2 TER meter

(Millipore) from cell cultures at time points 0 h, 24 h, and 48 h after addition of bacterial cells or proteins.

Statistical analysis. Data are expressed as means ± standard deviation. Statistical analysis of the results from the HEK-Blue cell lines and cytokine analysis was performed by one-way analysis of variance (ANOVA) followed by Tukey's HSD if homogeneity of variance was met or Games-Howell if variance was unequal. IBM SPSS software (IBM SPSS Statistics 22) was used for analysis; p values < 0.05 were considered significant.

Acknowledgments

Zhilin Li is thanked for the help with LPS analysis. This work was supported by Advanced Research Grant 250172 (MicrobesInside) of the European Research Council to WMdV, the Gravity (SIAM) and Spinoza grants of the Netherlands Organization for Scientific Research (NWO) to WMdV, the European Community's Seventh Framework Programme (FP7/2007-2013) under grant agreement n°215553-2 to NO, and the Finland Academy of Sciences (138902 & 258439 to RS, 252803 to JR, 141130 to WMdV). The authors declare no conflict of interest.

Supplemental Figures

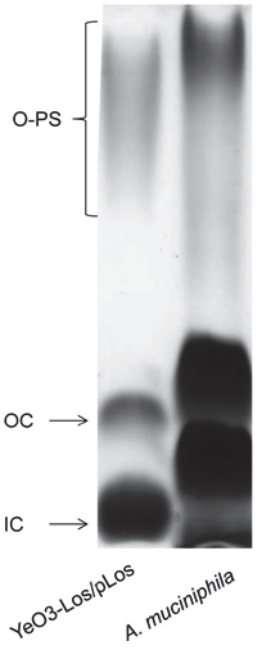


Figure S1. Silver staining of *A. muciniphila* LPS. Related to Figure 2A. *A. muciniphila* activates signaling pathways through TLR2 and TLR4. *Yersinia enterocolitica* (strain YeO3-Los/pLos) LPS was used as a control (Pinta *et al.*, 2012). O-PS; O-antigen, OC; outer core, IC; inner core. *A. muciniphila* LPS migrates slower than YeO3-Los/pLos LPS, suggesting the OC and IC bands have in their structure more sugar residues than the corresponding YeO3-Los/pLos bands.

Supplemental Experimental Procedures

Gene	Forward primer sequence (5' - 3')*	Reverse primer sequence (5' - 3')*
Amuc_0451	gcagcgcatatgaagccggctgcccggtctgtcaaaact	ggacgcctcgagcctggattcctt
Amuc_0625	gggcagcatatggaagagaaaaccggtttccctacg	ggctttcctctcgagcttgagaacaggagc
Amuc_1100	gggtaccatatgatcgtcaattccaaacgc	ccttggctcgagatcttcagacggttcctg
Amuc_2136	gatgacatgtccgacgccaagcagattgcggattccctt	catctcgagcaacggctgtacgttcat

* Restriction enzyme recognition sequences are underlined

PCR-primers used for plasmid construction.

Chapter 7

General discussion

The health impact of the microbes inhabiting our intestinal tract is becoming increasingly clear as a rapidly growing amount of research is performed in this field. The specific mutualistic interactions between the gut microbiota and the host may vary according to microbial community composition (Schluter and Foster, 2012). Even more important than the microbial composition is the functional capacity of the microbiota. Different microbial species are capable of performing the same tasks, which implies that harboring and expressing an optimal set of bacterial genes is more important to the health of the host than the abundance of specific species (Turnbaugh *et al.*, 2009a; Zhang *et al.*, 2015b). This conserved set of encoded functions, referred to as the core gut microbiome, has been shown to be shared among individuals (Turnbaugh and Gordon, 2009).

Formerly considered as potentially harmful to the host due to its association with pathogens, mucin degradation by intestinal bacteria is now considered an essential process in replenishing the mucosal layer (Tailford *et al.*, 2015a). Numerous bacteria have been identified to take part in this process, and the mucosa-associated microbiota can modulate the health of the host by producing favorable metabolites, enhancing colonization resistance and stimulating the immune response (Ouwerkerk *et al.*, 2013). The production rates of mucins and antimicrobial molecules present in the mucus layer are regulated by both commensal microbiota and host inflammatory signals (McGuckin *et al.*, 2011). Mucin degradation also offers ecological benefits to other bacteria, which scavenge the products, such as oligosaccharides or sulfate, released as a result of the enzymatic activity of mucin-degrading bacteria (Derrien *et al.*, 2010). One of the key players in the mucosa-associated microbiota is *Akkermansia muciniphila*, the presence of which is associated with a healthy gut. The work described in thesis aimed at characterizing the metabolic capacities and immunostimulatory properties of *A. muciniphila*. The research approaches included transcriptomic and proteomic analyses, as well as *in vitro* immunological assays. The insights gained from this research are discussed in the following sections.

Adaptation of *A. muciniphila* to the consumption of host-derived glycans

A. muciniphila was isolated using mucin as the sole carbon and nitrogen source, and further analysis has proven it to be highly specialized in consumption of this host-derived glycan. Determination of mucin degradation by periodic acid–Schiff staining and high-performance anion-exchange chromatography patterns of the culture supernatant showed *A. muciniphila* to degrade up to 85 % of the mucin molecule (Derrien, 2007). A major benefit of utilizing mucin in the gut is that it provides a stable source of energy, whereas the concentrations of diet-derived substrates are constantly changing. Even more advantageous is the ability to

switch between consumption of host and diet-derived glycans depending on the availability of these nutrients, as in the case of *Bacteroides thetaiotaomicron* (Sonnenburg *et al.*, 2005).

A. muciniphila was able to individually utilize all the mucin-derived sugars, except for sialic acid, in a medium supplemented with a protein source (**Chapter 3**). The *A. muciniphila* genome encodes four sialidases (Amuc_0623, Amuc_0625, Amuc_1547, Amuc_1835), the function of which is to cleave sialic acid off from sialylated glycans, such as mucin. It was recently shown that two of the sialidases (Amuc_0625, Amuc_1835) released sialic acid from sialic acid lactose and 6'-sialyllactose (Tailford *et al.*, 2015b). In the case of mucin, sialic acid residues may prevent the action of other glycosyl hydrolases as the residues are located in the non-reducing ends of the mucin molecule. *A. muciniphila* grows efficiently on mucin, and it is therefore likely that it is also able to cleave off the sialic acid residues from mucin, thereby exposing the mucin sugars, even if it is not able to utilize the sialic acid for energy. Enzymatic assays showed sialidase activity in the *A. muciniphila* fractions extracted in **Chapter 5**, encouraging further research on the ability of *A. muciniphila* to use sialylated compounds.

The uptake of galactose and fucose was markedly increased in the presence of mucin, indicating that additional mucin-derived components are needed for optimal growth. This was also evident when *A. muciniphila* was grown on the non-mucin sugar glucose, as glucose was not depleted in the absence of mucin or mucin-derived monosugars. Transcriptome analysis comparing the gene expression between *A. muciniphila* grown on mucin or glucose showed upregulation of genes involved in protein biosynthesis and energy metabolism in the mucin condition. As expected, the majority of the hydrolases involved in mucin degradation were upregulated in *A. muciniphila* grown on mucin in comparison to glucose. *A. muciniphila* grown on glucose showed a considerable stress response, and some specific genes were upregulated, such as Amuc_1094 which was identified to encode a glucokinase induced by glucose. Proteome analysis correlated well with the transcriptome, and thus confirmed the response observed for the two substrates.

Due to the extensive mucin-degrading capacity of *A. muciniphila*, and its presence in early life intestine, the ability of *A. muciniphila* to grow on human milk and human milk oligosaccharides (HMOs) was investigated in **Chapter 4**. *A. muciniphila* demonstrated metabolic activity on human milk, lactose and 2'-fucosyllactose, but growth was limited in comparison to mucin. Transcriptome analysis of *A. muciniphila* grown on human milk showed upregulation of genes involved in lactose degradation, carbohydrate transport and translational activity in comparison to mucin. However, the possible confounding effect of different

growth phases between the conditions needs to be considered when interpreting the results. Furthermore, human milk from various donors, possibly with different histo-blood group systems, and other HMOs besides the relatively simple 2'-fucosyllactose and 3'-sialyllactose should be tested to confirm the ability of *A. muciniphila* to persist in the human milk environment.

Fermentation of mucin by *A. muciniphila* leads to production of the short-chain fatty acids (SCFA) acetate and propionate, both of which have various physiological effects on the host (Macfarlane and Macfarlane, 2012). In addition to providing energy to the epithelial cells, propionate, and to some extent acetate, have shown positive effects in prevention and treatment of conditions such as inflammatory bowel diseases (IBD), obesity and colorectal cancer (Kasubuchi *et al.*, 2015; Louis *et al.*, 2014). Several mechanisms for these actions have been suggested. Propionate inhibits the activity of histone deacetylases (HDACs) in colonocytes and immune cells, which has consequences for gene expression and cellular differentiation, including downregulation of proinflammatory cytokines (Schilderink *et al.*, 2013). Propionate also induces the differentiation of regulatory T cells involved in controlling intestinal inflammation (Arpaia *et al.*, 2013; Smith *et al.*, 2013). Acetate and propionate interact with G protein-coupled receptors of host cells, and interaction with GPR43 in particular is proposed to play a role in inducing their anti-inflammatory effects via the modulation of regulatory T cells (Maslowski *et al.*, 2009). Interestingly, propionate increases the production of leptin, a satiety hormone, through GPR41, and is therefore involved in host energy balance regulation (Samuel *et al.*, 2008; Xiong *et al.*, 2004).

The effect of *A. muciniphila* metabolites, including acetate and propionate, on host epithelium was recently studied using an *ex vivo* model based on mouse organoids (Lukovac *et al.*, 2014). Organoids are small crypt-villus structures grown from ileal intestinal stem cells, which can be used as a physiologically relevant surrogate system to study interactions of desired compounds with the intestine (Grabinger *et al.*, 2014). *A. muciniphila* fermentation products induced regulation of host epithelial genes involved in cellular lipid metabolism and growth. Additionally, *A. muciniphila* metabolites significantly decreased expression of GPR43 and Pparγ, and increased expression of Hdac3 and Hdac5, implying regulation of host transcriptional response via histone acetylation modifications. These results, together with the transcriptome data from mice mono-colonized with *A. muciniphila* (Derrien *et al.*, 2011), suggest that metabolic activity of *A. muciniphila* in the gut leads to physiological responses, and may therefore have consequences on the health of the host.

However, in the intestinal environment bacteria are never present alone, but as members of complex communities that influence the activity of individual species (El Aidy *et al.*, 2013). In the case of *A. muciniphila*, other bacteria residing close to the mucus layer may use its fermentation products and the oligosaccharides it releases from mucin. Several common members of the gut microbiota, such as *Eubacterium hallii* and *Anaerostipes caccae* are able to convert acetate to butyrate (Louis and Flint, 2009), suggesting that the presence of *A. muciniphila* in the gut may indirectly increase the levels of butyrate, which in turn is beneficial for the host (Guilloteau *et al.*, 2010). Cross-feeding between a limited number of bacterial species can be studied *in vitro* with simple co-cultures, but to study the general dynamics of the gut microbiota more sophisticated models are needed (Fritz *et al.*, 2013). Examples of *in vitro* models for the gastrointestinal (GI) tract are the Simulator of the Human Intestinal Microbial Ecosystem (SHIME) (Molly *et al.*, 1993), the TIM1 model for the stomach and small intestine (Minekus *et al.*, 1995), and TIM2 model for the large intestine (Minekus *et al.*, 1999). Recently a mucus layer has been integrated in the SHIME model (M-SHIME), allowing improved simulation of the mucosal and luminal microbiota (Van den Abbeele *et al.*, 2012). Another way of studying microbial interactions is to use *in silico* models, which can resolve complex interactions at the molecular level, and as such complement *in vitro* and *in vivo* models (Heinken and Thiele, 2015). Knowledge gained from these models can be used to design defined, synthetic bacterial communities (Grosskopf and Soyer, 2014), which can even be aimed at treating specific disorders in the gut. Using defined microbial ecosystems as therapeutics is still in its infancy, but this approach is a viable alternative to conventional fecal microbiota transplantation (Petrof and Khoruts, 2014). A pilot study described successful treatment of two patients suffering from recurrent *Clostridium difficile* infection with a mixture of 33 carefully selected strains of bacteria, isolated from the stool of one healthy donor (Petrof *et al.*, 2013). To find out more about the activity and function of *A. muciniphila* in bacterial communities, experiments using the abovementioned techniques should be conducted.

The role of *A. muciniphila* outer membrane proteins in immune stimulation

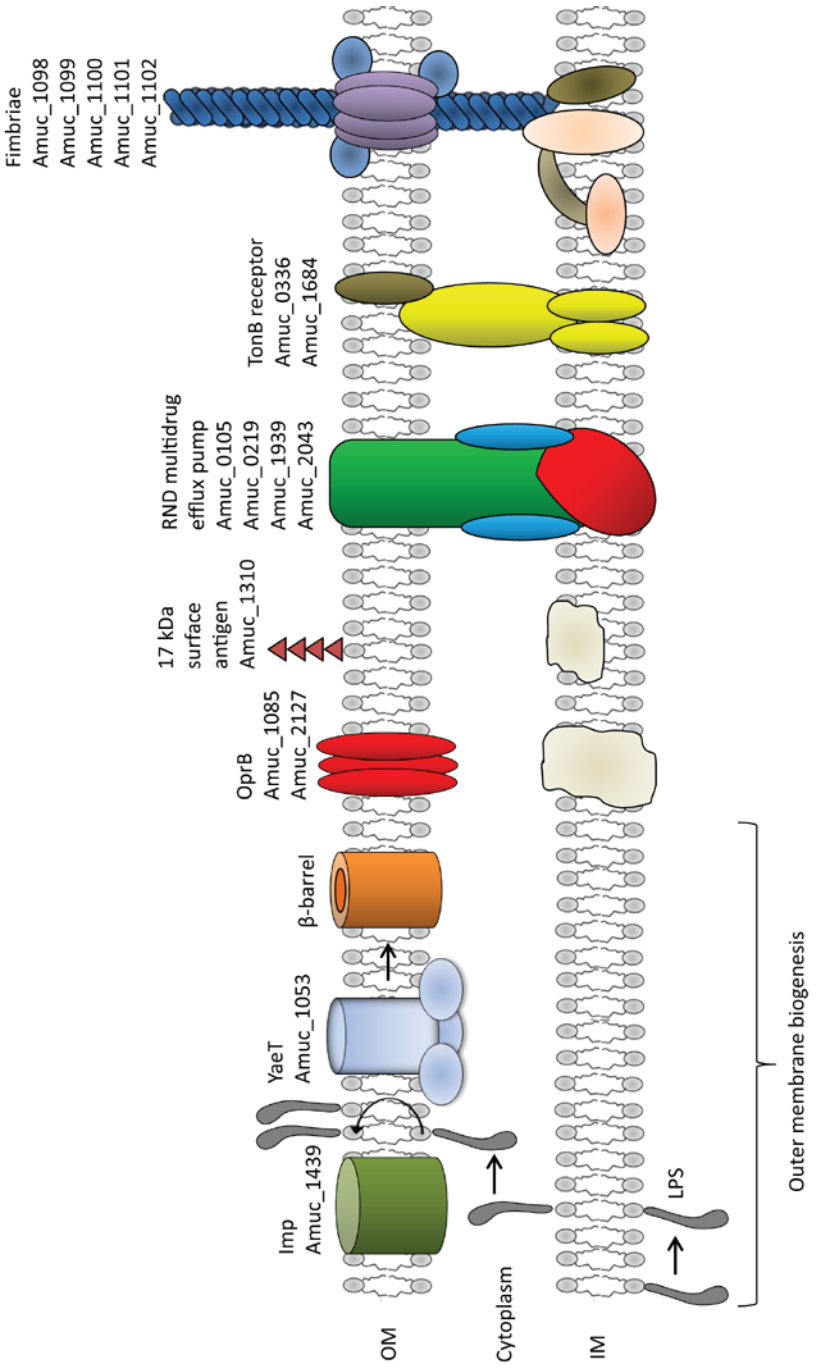
The cell envelope structures of both Gram-positive and Gram-negative bacteria have a great influence on their immunostimulatory capacities. Many components of the cell envelope, including lipopolysaccharides (LPS), peptidoglycans and lipoteichoic acids are involved in immune response in the gut by stimulating intestinal receptors (Bron *et al.*, 2012). Also surface polysaccharides, such as polysaccharide A from *Bacteroides fragilis* (Surana and Kasper, 2012), surface-

layer proteins from *Lactobacillus acidophilus* NCFM and *Propionibacterium freudenreichii* (Konstantinov *et al.*, 2008; Le Marechal *et al.*, 2015) as well as cell surface-associated exopolysaccharides from *Bifidobacterium breve* (Fanning *et al.*, 2012) have been shown to modulate various immune functions.

Akkermansia muciniphila was previously shown to modulate the mucosal immune response of mono-colonized mice (Derrien *et al.*, 2011). Predicting that *A. muciniphila* outer membrane (OM) proteins play a role in this immune stimulation, the OM proteome was enriched and characterized by mass spectrometry (**Chapter 5**), and candidate proteins were purified (in some cases after heterologous overproduction in *Escherichia coli*), and tested for their immunomodulatory capacities *in vitro* (**Chapter 6**). Proteins involved in secretion, transport and biogenesis of the Gram-negative membranes were found to be abundant in the OM of *A. muciniphila* (Figure 1). A fraction enriched for *A. muciniphila* cell envelope proteins induced a higher Toll-like receptor (TLR) 2 response and higher cytokine production than a fraction enriched for enzymes.

The most abundant protein in the OM enriched fractions was a type II and III secretion system protein (Amuc_1098). Further analysis revealed that this protein belongs to a gene cluster (Amuc_1098 to Amuc_1102), which contains proteins resembling fimbrial biogenesis proteins. A hypothetical protein Amuc_1100 from the identified cluster was shown to be localized on the OM of *A. muciniphila*. Purified Amuc_1100 induced production of IL-6, IL-8 and IL-10, activated TLR2 and increased transepithelial electrical resistance (TER). Fimbriae-like structures have been observed in electron microscopy images of *A. muciniphila*, and immuno-EM has confirmed the presence of some of the proteins in the identified gene cluster in these fimbriae (Huuskonen et al, unpublished results). The genome of *A. muciniphila* also encodes for several other genes possibly involved in formation of fimbriae. Combined these results make a strong case for the presence of fimbriae in *A. muciniphila*, but their exact structure and function is still under investigation.

Figure 1. *A. muciniphila* cell envelope with attention to outer membrane proteins. The presence of all the named proteins, except for Amuc_1085, was confirmed by mass spectrometry analysis. The increased membrane permeability (Imp) protein (Amuc_1439) is needed in transport of lipopolysaccharide (LPS) to the bacterial cell surface. The YaeT protein (Amuc_1053) is involved in insertion of ß-barrel proteins into the OM. Carbohydrate-selective porin OprB (Amuc_1085, Amuc_2127) is a substrate-selective channel for a variety of different sugars, but its exact role in *A. muciniphila* is unknown. Also the function of the 17 kDa surface antigen (Amuc_1310) is unclear. RND multidrug efflux pumps (Amuc_0105, Amuc_0219, Amuc_1939, Amuc_2043) catalyze the active efflux of many toxic compounds that may also include antibiotics and chemotherapeutic agents. TonB receptors (Amuc_0336, Amuc_1684) are involved in uptake of specific substrates. The gene cluster Amuc_1098 to Amuc_1102 is predicted to be involved in the formation of fimbriae in *A. muciniphila*. OM; outer membrane, IM; inner membrane. Schematic representation, not in scale.


Fimbriae
Amuc_1098
Amuc_1099
Amuc_1100
Amuc_1101
Amuc_1102
TonB receptor
Amuc_0336
Amuc_1684
RND multidrug
efflux pump
Amuc_0105
Amuc_0219
Amuc_1939
Amuc_2043
17 kDa
surface
antigen
Amuc_1310
OprB
Amuc_1085
Amuc_2127
β-barrel
YaeT
Amuc_1053
Imp
Amuc_1439
LPS
OM
Cytoplasm
IM
Outer membrane biogenesis

Based on the immune assays *A. muciniphila* OM proteins and Amuc_1100 play a role in the immune response. Fimbriae would be the first contact site of *A. muciniphila* to the extracellular space, which would make them likely to be important in host-bacterial interactions. One of the ways to study the function of the proteins, such as Amuc_1100, would be to construct a knock-out strain that does not contain the chosen gene. However, a genetic system for *A. muciniphila* does not currently exist. Another way would be to express the gene in a model organism, such as *E. coli*, and for example compare the immunomodulatory or adhesion properties of the modified *E. coli* to the non-modified strain. In this case, the appropriate *E. coli* strain would need to be chosen carefully to avoid confounding results. Also other models can be used to study the host interactions of bacteria or their products. For example, recently mouse gut organoids were used successfully to compare the transcriptional response of *A. muciniphila* metabolites to *Faecalibacterium prausnitzii* metabolites (Lukovac *et al.*, 2014). In addition to the metabolites, also the transcriptional response of the host to *A. muciniphila* OM proteins and fimbrial proteins could be studied in these and other organoid systems.

Already in the study with *A. muciniphila* mono-colonized mice, bacterial cell envelope components were suggested to play a role in the host response (Derrien *et al.*, 2011). Colonization by *A. muciniphila* led to increased expression of genes involved in immune responses, whereas this was not seen in mice mono-associated with the Gram-positive *Lactobacillus plantarum*. In addition to dissimilar cell envelope structures, these two bacteria also colonize different niches in the gut. *A. muciniphila* was found to generate most of the host response in the colon while that of *L. plantarum* was mainly found in the ileum, which may reflect their natural habitat (Derrien et al 2011). Moreover, *A. muciniphila* was found to reside in the mucus layer close to the epithelial cells while *L. plantarum* was found in the lumen. Despite the pronounced immune response to *A. muciniphila* colonization, which included upregulated pathways involved in antigen presentation, B and T cell receptor signaling and leukocyte signaling, there were no strong indications that upregulation of these genes took part in driving a proinflammatory response. It was concluded that colonization by *A. muciniphila* led to a non-inflammatory, commensal interaction and to intestinal tolerance (Derrien *et al.*, 2011). The highly abundant OM proteins characterized in **Chapter 6** probably played a role in the mouse immune response, and it would be of interest to find out whether the purified proteins, or fimbriae alone, can stimulate the immune system also *in vivo*.

Akkermansia muciniphila – a next-generation probiotic?

The rapid development of omics-based techniques and other culture-independent microbiological methods has made it possible to identify potential new health-promoting bacteria from the comparison of healthy and diseased populations. This was the case for example with *F. prausnitzii*, which raised interest after it was discovered that Crohn's disease patients had significantly reduced levels of this species in comparison to healthy controls (Sokol *et al.*, 2008). Since then, many studies have shown an association between the relative abundance of *F. prausnitzii* and a number of other diseases (Bruzzese *et al.*, 2014; Machiels *et al.*, 2014; Remely *et al.*, 2014; Remely *et al.*, 2015b; Stoll *et al.*, 2014). Anti-inflammatory properties of *F. prausnitzii* on the host have also been demonstrated in mouse models (Martin *et al.*, 2015; Miquel *et al.*, 2015; Sokol *et al.*, 2008). Numerous metabolites have been associated with the protective effect of *F. prausnitzii* in a mouse model of acute colitis, but only salicylic acid reduced IL-8 production in a cell line (Miquel *et al.*, 2015). The levels of salicylic acid were positively correlated to presence of *F. prausnitzii* in the mouse gut, but *F. prausnitzii* has not been shown to produce this compound. However, some *Fusobacterium prausnitzii* (former nomenclature for *F. prausnitzii*) strains are able to ferment salicin, an alcoholic glucoside that can be converted into salicylic acid (Miquel *et al.*, 2015). Butyrate was also shown to reduce IL-8 production, but not the butyrate specifically produced by *F. prausnitzii*, indicating that other, yet unknown, compounds in the *F. prausnitzii* supernatant may be responsible for the observed anti-inflammatory effects.

In order to accurately study the characteristics of the identified bacteria, they need to be culturable, which may be an issue with strictly anaerobic gut bacteria. For instance, segmented filamentous bacteria (SFB) have shown important immunostimulatory properties in mice, but resisted *in vitro* culturing until very recently making it difficult to study their interactions with the host (Schnupf *et al.*, 2015). Similarly, they may be present in animal models but need to be also found in human systems. This is doubtful for SFB as these are not, or almost not detected in human (Sczesnak *et al.*, 2011), except for children under the age of three years (Yin *et al.*, 2013). In contrast, *A. muciniphila* is abundantly present in human as well as model systems, culturable, and negatively correlated with several diseases (Belzer and de Vos, 2012). Moreover, the mechanisms by which *A. muciniphila* could work have been studied *in vitro* (**Chapter 6**) and *in vivo* (Everard *et al.*, 2013). These qualities and evidence make *A. muciniphila* a promising candidate for a next-generation probiotic. However, it is a long road from being a candidate to being on the market (Kumar *et al.*, 2015).

Probiotics are defined as 'live microorganisms that, when administered in adequate amounts, confer a health benefit on the host' (Hill *et al.*, 2014). The majority of currently used probiotics are from the genera *Lactobacillus* and *Bifidobacterium*, and most research also exists on these food-grade bacteria. For a bacterium to be a successful probiotic, it must fill numerous criteria: the bacteria must have a beneficial effect on the host, be able to survive passage through the GI tract, be safe for human consumption and have good technological properties (Fontana *et al.*, 2013; Ouwehand *et al.*, 2002). Additional scientific information requirements for novel probiotics have been suggested recently (Kumar *et al.*, 2015). These include genome announcement, antibiotic resistance profile, selection of proper *in vivo* model, toxicological studies, and definition of target population.

In humans the evidence of the beneficial effect of *A. muciniphila* comes so far only from association studies, whereas in mice feeding with *A. muciniphila* has shown direct health effects (Everard *et al.*, 2013) (**Chapter 1**). Pilot studies in humans to discover the transcriptional response of the host to exposure to *A. muciniphila* are on-going, but before larger scale human trials can be conducted, studies into certain technical and safety-related issues need to be undertaken. *A. muciniphila* grows most efficiently on mucin, but for manufacturing and safety reasons a defined medium, which supports high growth and is free of animal-derived compounds needs to be designed. The information about the metabolic capacities of *A. muciniphila* described in this thesis (**Chapter 3**) can be used as a basis for further research on this topic. The survival rate of *A. muciniphila* during passage through the GI tract has not been thoroughly examined. Several *in vitro* models exist to study the resistance of bacteria to the harsh conditions in the stomach and small intestine, where low pH and presence of bile and digestive enzymes can influence the viability of bacteria (Dunne *et al.*, 2001; Geirnaert *et al.*, 2014). In **Chapter 4**, evidence for an active acid resistance system was found in the transcriptome analysis of *A. muciniphila* grown on human milk. This system, relying on transport of L-glutamine and its conversion to L-glutamate, may aid *A. muciniphila* in surviving passage through the GI tract (Figure 2). A similar system has also been described in *Lactobacillus reuteri* (Teixeira *et al.*, 2014).

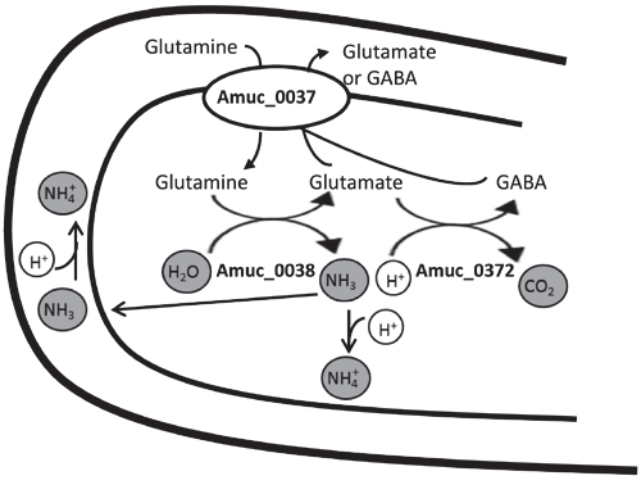


Figure 2. Proposed mechanism of the acid resistance system in *A. muciniphila*. Glutamine is transported into the cell through the glutamate/γ-aminobutyric acid (GABA) antiporter (Amuc_0037). Subsequently, glutamine is converted to glutamate by a glutaminase (Amuc_0038). This leads to release of ammonia, which neutralizes a proton, resulting in elevated intracellular pH in acidic environments. In addition, the decarboxylation of glutamate to GABA by glutamate decarboxylase (Amuc_0372) results in fixation of a proton. The genes for Amuc_0037, Amuc_0038 and Amuc_0372 are all upregulated during growth of *A. muciniphila* on mucin in comparison to human milk (**Chapter 4**). Figure adapted from (Lu *et al.*, 2013).

Correct packaging is also an important consideration when evaluating the survival of the probiotic product. In the case of strict anaerobes, like *A. muciniphila*, extra caution has to be followed to keep the bacteria viable. However, it was recently shown that *A. muciniphila* tolerates low amounts of oxygen (Reunanen *et al.*, 2015; van der Ark *et al.*, 2015), suggesting that the anaerobic nature of this bacterium may not jeopardize its application though further research is encouraged on the topic.

The complete genome sequence of *A. muciniphila* is available and can be used for example for screening of transferrable antibiotic-resistance genes and virulence-related genes. The *A. muciniphila* genome contains potential beta-lactamase genes and a gene coding for a 5-nitroimidazole antibiotic resistance protein but no signs of genetically transferable elements have been described in the vicinity of these genes (van Passel *et al.*, 2011). Moreover, *A. muciniphila* has been shown to be susceptible to imipenem, piperacillin/tazobactam and doxycycline, but resistant to vancomycin, metronidazole and penicillin (Dubourg *et al.*, 2013). It is known that

many lactic acid bacteria, including well-known probiotics, are resistant to vancomycin because of absence of its cell-envelope target, and similar mechanisms may explain the resistance in *A. muciniphila* (Salminen *et al.*, 1998). Several studies have reported increased levels of *A. muciniphila* after antibiotic treatment, particularly vancomycin treatment in model animals and human (Ferrer *et al.*, 2014; Hansen *et al.*, 2012; Russell *et al.*, 2012; Vrieze *et al.*, 2014). Notably, broad-spectrum antibiotic treatment was associated with high-level colonization by *A. muciniphila* in two patients, with up to 85 % of all the sequences associated with this species (Dubourg *et al.*, 2013). However, neither of the patients presented significant gastrointestinal disorders. Furthermore, no adverse effects were seen in mice mono-colonized with *A. muciniphila* (Derrien *et al.*, 2011). Even though *A. muciniphila* was isolated from human feces, and is a common member of the gut microbiota, its safety has to be assessed carefully with *in vitro* and *in vivo* studies before it can be used as a probiotic. Authorities such as the European Food Safety Authority (EFSA) use the qualified presumed safety (QPS) status to assess safety of probiotic strains. However, it is not clear yet which safety studies, if any, have to be performed to obtain this status for next-generation probiotics when these are already present in the normal population at high levels.

Before bringing a probiotic to the market, one of the key regulatory questions is whether the probiotic is categorized as a food supplement or a medicinal product. Food-related applications are controlled by EFSA, whereas EMA (European Medicines Agency) is responsible for evaluation of medicinal products. EMA guidelines are made with drug molecules in mind, which poses a problem with probiotics as they are living microorganisms. However, if a bacterial molecule, such as a specific protein, is demonstrated to confer a health benefit on its own, it could be categorized as a medicinal product. Probiotic companies are currently struggling with these issues, and the future of the regulatory framework for probiotics remains to be seen, notably in the food and supplements space, as this is where these companies are active.

One of the major considerations for a probiotic is whether it will be used to prevent or treat diseases. This also influences the definition of a target population. In the case of *A. muciniphila,* patients with metabolic disorders or IBD would be an obvious choice, as most evidence exists on the beneficial association between these diseases and *A. muciniphila*. However, causal links need to be established between the presence of *A. muciniphila* and specific disorders in order to define the therapeutic potential of this bacterium. While efficacy is the key factor for the use of a probiotic, it is of interest to know the mechanism of action that induces a beneficial effect on the host. This is of specific relevance in case optimization or improvement of probiotic strains is desired. Even for the currently used probiotics

this is not known and, interestingly, also often not for pharmaceutical products. Modulation of the host immune system is one of the most commonly proposed health benefits attributed to the consumption of probiotics. *A. muciniphila* was shown to induce production of a wide range of cytokines in human derived peripheral blood mononuclear cells (PBMCs) and to activate TLR2 and TLR4 receptors in reporter cell lines (**Chapter 6**). The levels of IL-10 induced by live *A. muciniphila* were similar to those induced by *F. prausnitzii* and *L. plantarum*, but *A. muciniphila* supernatant induced higher IL-10 production compared to these two gut commensals. Induction of high IL-10 levels is suggested to be one of the mechanisms behind the anti-inflammatory role of *F. prausnitzii* (Sokol *et al.*, 2008). In addition, *A. muciniphila* was shown to increase the TER of Caco-2 monolayer (**Chapter 6**) (Reunanen *et al.*, 2015), and to restore mucus layer thickness in mice that were fed a high-fat diet (Everard *et al.*, 2013). The positive effect on the mucus layer was also observed in another mouse trial, where administration of *A. muciniphila* increased the number of mucin-producing goblet cells and adipose tissue-resident CD4 Foxp3 regulatory T cells (Tregs) (Shin *et al.*, 2014). These observations indicate strengthening of the epithelial barrier function, which is often considered as a key effect of probiotic therapy (Mennigen and Bruewer, 2009). Adherence is important for probiotic colonization and antagonism against pathogens by competitive exclusion, but it can also be a risk of translocation, especially in patient groups with compromised intestinal barrier function. *A. muciniphila* was shown to adhere to human colonic cell lines but not to colonic mucus (Reunanen *et al.*, 2015). However, as *A. muciniphila* produces an active repertoire of mucus-degrading enzymes (**Chapter 3**), it cannot be excluded that initial binding of *A. muciniphila* takes place followed by rapid release. Moreover, exclusion of potential pathogens may also occur via competition for substrates, and because of its high affinity to mucin, this is a likely mechanism by which *A. muciniphila* remains established in the mucosal layer that has large dynamics (de Vos, 2015).

Not all the studies on *A. muciniphila* have reported beneficial health effects, and these need to be considered as well when evaluating its probiotic properties. Mouse studies have reported increased levels of *A. muciniphila* in a model of dextran sodium sulfate (DSS)-induced colitis (Berry *et al.*, 2012; Hakansson *et al.*, 2014). This is in contrast to human studies, where abundance of *A. muciniphila* is reduced in ulcerative colitis patients in comparison to healthy controls (James *et al.*, 2014; Png *et al.*, 2010; Rajilic-Stojanovic *et al.*, 2013; Vigsnaes *et al.*, 2012), and a mouse study where *A. muciniphila* extracellular vesicles ameliorated DSS colitis (Kang *et al.*, 2013). The DSS colitis model resembles acute inflammation-induced injury more than human ulcerative colitis (Berry *et al.*, 2012). DSS exerts a cytotoxic

effect on the gut epithelium, which results in thickening of the mucus layer and increased intestinal permeability (Schreiber *et al.*, 2013). Conversely to this, in human chronic colitis the mucus layer is thinner as compared to healthy controls (Strugala *et al.*, 2008; Swidsinski *et al.*, 2007). Thickening of the mucus layer in the DSS-model could explain the overgrowth of *A. muciniphila,* as mucin is its main source of energy in the gut, and therefore it is not clear whether *A. muciniphila* is actually involved in the pathogenesis of colitis in these models. Another link between the host immune system and *A. muciniphila* was described in a study investigating the gut microbiota of immunodeficient $Rag1^{-/-}$ mice, which lack all mature lymphocytes, and thus have no adaptive immune system (Zhang *et al.*, 2015a). *A. muciniphila* was enriched in these immunodeficient mice, and the enrichment was suppressed when $Rag1^{-/-}$ mice received bone marrows from wild-type mice leading to restoration of $CD4^+$ T cells and mature B cells. These findings indicate that the adaptive immune system can directly alter the gut microbiota, but the components responsible for this effect are not known. The study is yet another demonstration of the cross-talk between *A. muciniphila* and the host immune system, but the implications of this interaction in healthy or immunocompromised hosts is still unclear. An important lesson from these studies is that simple conclusions on the function of a specific bacterium, such as *A. muciniphila,* cannot be easily made. In all association studies confounding factors should be taken into account, in mouse models the different effects of the treatments and knock-outs should be considered, and finally care should be taken to extrapolate mouse studies to man (see below).

In vitro assays, such as the ones used in this thesis to measure intestinal receptor activity or cytokine production in PBMCs (**Chapter 6**), are useful for screening candidate bacteria for their stimulatory ligands or anti- and proinflammatory properties. However, these models have certain limitations, which must be taken into consideration when interpreting the results. For example, the assay with PBMCs only involves one type of immune cells and ignores the complexity of the *in vivo* communication between different cell types. Also the lack of other, competing bacteria in these assays is a drawback, as in the intestine members of the microbiota are in constant interaction with each other. *In vivo* experiments with animal models are the next step in confirming the findings of the *in vitro* assays, but it must be kept in mind that for example the mouse gut differs from the human gut both in terms of physiology and composition of the microbiota (Krych *et al.*, 2013). This makes it sometimes difficult to extrapolate the results from mouse studies to the human situation. Therefore, also in the case of *A. muciniphila*, experiments with humans are eventually needed to validate the findings of the *in vitro* assays and animal models.

Concluding remarks and future perspectives

Since the isolation and initial characterization of *A. muciniphila* a little more than ten years ago (Derrien *et al.*, 2004), this bacterium has been associated with numerous diseases, most often with reduced abundance in patients in comparison to healthy subjects. However, relatively few studies have described the physical or functional properties of *A. muciniphila* in detail. In this thesis the substrate utilization capacities concerning host-derived glycans mucin and human milk, and the immunostimulatory profile of *A. muciniphila* and its products were investigated. In depth transcriptome and proteome analyses were used to complement the available genomic information. In addition to the outcomes described here, concerning metabolic pathways and OM proteome among other things, these datasets can be further utilized in interpretation of future and to some degree also past studies.

A. muciniphila is the only currently described human intestinal bacterium that belongs to the phylum *Verrucomicrobia*. Limited amount of studies are available on the other members of this phylum, and due to the diverse living environments of these bacteria, ranging from seawater to termite gut, comparison of data is not always reasonable. The cell envelope structure of *Verrucomicrobia* has gained interest recently but the proteomic analysis of *A. muciniphila* OM described in this thesis is the first experimental protein-level characterization of *Verrucomicrobia* membranes. On the upside, metatranscriptomic and metaproteomic datasets can easily be assessed for *A. muciniphila*, as *Verrucomicrobia* is a deeply rooted phylum (Lange *et al.*, 2015; Rooijers *et al.*, 2011). Studying novel bacteria offers great possibilities, and fortunately the number of cultured species from the human gut environment is ever increasing with over a thousand cultured isolates presently described in detail (Rajilic-Stojanovic and de Vos, 2014). This is of substantial interest as culturing the intestinal microbiota has rather been neglected after the development of fast DNA-based approaches. A culturing renaissance is taking place now, as it turns out that insight in cultures is needed when trying to develop new therapies.

A. muciniphila provides a relevant target for studying host-microbial communication because few commensal bacteria live in such close contact with the epithelial cells in the intestine. According to the findings presented in **Chapter 6**, *A. muciniphila* is able to evoke diverse immunological responses *in vitro*. Whether this is the case also *in vivo* and what are the consequences of the immunostimulation to the host remains to be discovered. However, as *A. muciniphila* is commonly associated with a healthy gut, it is likely that it is involved in beneficial interactions with the host, possibly by regulating immune

homeostasis and keeping the immune system alert for potential disruptions. Future studies should focus on determining the exact mechanisms of *A. muciniphila* immunomodulation and establishing causal relationships between disease development and presence of this species in the gut. This can be achieved by using *in vitro* assays, for example by observing T-cell differentiation in a co-culture of *A. muciniphila* and immune cells. Furthermore, *in vivo* mouse studies where several immune parameters can be measured simultaneously are required to get a more complete overview of the immunomodulatory properties. Previously published studies with positive effects of *A. muciniphila* on the metabolic profile of mice should be followed up by studying the possible protective effect of *A. muciniphila* on the development of other diseases, such as IBD and metabolic diseases. After the technical and safety considerations described above have been addressed, and if the results derived from animal studies consistently point to a beneficial effect of *A. muciniphila*, human trials should be conducted. Ultimately, the comprehensive characterization of individual members of the gut microbiota leads to better understanding of the whole microbial community, and thereafter improved knowledge of human health.

References

Aagaard, K., Ma, J., Antony, K.M., Ganu, R., Petrosino, J., and Versalovic, J. (2014). The placenta harbors a unique microbiome. *Science translational medicine 6*, 237ra265.

Abe, F., Muto, M., Yaeshima, T., Iwatsuki, K., Aihara, H., Ohashi, Y., and Fujisawa, T. (2010). Safety evaluation of probiotic bifidobacteria by analysis of mucin degradation activity and translocation ability. *Anaerobe 16*, 131-136.

Aguilera, L., Toloza, L., Gimenez, R., Odena, A., Oliveira, E., Aguilar, J., Badia, J., and Baldoma, L. (2014). Proteomic analysis of outer membrane vesicles from the probiotic strain Escherichia coli Nissle 1917. *Proteomics 14*, 222-229.

Alhawi, M., Stewart, J., Erridge, C., Patrick, S., and Poxton, I.R. (2009). Bacteroides fragilis signals through Toll-like receptor (TLR) 2 and not through TLR4. *Journal of medical microbiology 58*, 1015-1022.

Allen-Vercoe, E. (2013). Bringing the gut microbiota into focus through microbial culture: recent progress and future perspective. *Current opinion in microbiology 16*, 625-629.

Allen, L.H. (2012). B vitamins in breast milk: relative importance of maternal status and intake, and effects on infant status and function. *Advances in nutrition 3*, 362-369.

Angelberger, S., Reinisch, W., Makristathis, A., Lichtenberger, C., Dejaco, C., Papay, P., Novacek, G., Trauner, M., Loy, A., and Berry, D. (2013). Temporal bacterial community dynamics vary among ulcerative colitis patients after fecal microbiota transplantation. *The American journal of gastroenterology 108*, 1620-1630.

Anhe, F.F., Roy, D., Pilon, G., Dudonne, S., Matamoros, S., Varin, T.V., Garofalo, C., Moine, Q., Desjardins, Y., Levy, E.*, et al.* (2014). A polyphenol-rich cranberry extract protects from diet-induced obesity, insulin resistance and intestinal inflammation in association with increased Akkermansia spp. population in the gut microbiota of mice. *Gut*.

Arpaia, N., Campbell, C., Fan, X., Dikiy, S., van der Veeken, J., deRoos, P., Liu, H., Cross, J.R., Pfeffer, K., Coffer, P.J.*, et al.* (2013). Metabolites produced by commensal bacteria promote peripheral regulatory T-cell generation. *Nature 504*, 451-455.

Arrieta, M.C., Stiemsma, L.T., Amenyogbe, N., Brown, E.M., and Finlay, B. (2014). The intestinal microbiome in early life: health and disease. *Frontiers in immunology 5*, 427.

Artis, D. (2008). Epithelial-cell recognition of commensal bacteria and maintenance of immune homeostasis in the gut. *Nature reviews Immunology 8*, 411-420.

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

Azad, M.B., Konya, T., Maughan, H., Guttman, D.S., Field, C.J., Chari, R.S., Sears, M.R., Becker, A.B., Scott, J.A., Kozyrskyj, A.L., *et al.* (2013). Gut microbiota of healthy Canadian infants: profiles by mode of delivery and infant diet at 4 months. *CMAJ : Canadian Medical Association journal = journal de l'Association medicale canadienne 185*, 385-394.

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

Baxter, N.T., Zackular, J.P., Chen, G.Y., and Schloss, P.D. (2014). Structure of the gut microbiome following colonization with human feces determines colonic tumor burden. *Microbiome 2*, 20.

Belzer, C., and de Vos, W.M. (2012). Microbes inside--from diversity to function: the case of Akkermansia. *The ISME journal 6*, 1449-1458.

Belzer, C., Liu, Q., Carroll, M.C., and Bry, L. (2011). THE ROLE OF SPECIFIC IgG AND COMPLEMENT IN COMBATING A PRIMARY MUCOSAL INFECTION OF THE GUT EPITHELIUM. *European journal of microbiology & immunology 1*, 311-318.

Ben-Amor, K., Heilig, H., Smidt, H., Vaughan, E.E., Abee, T., and 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*, 4679-4689.

Bendtsen, J.D., Kiemer, L., Fausboll, A., and Brunak, S. (2005). Non-classical protein secretion in bacteria. *BMC microbiology 5*, 58.

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

Bergwitz, C., and Juppner, H. (2011). Phosphate sensing. *Advances in chronic kidney disease 18*, 132-144.

Berry, D., Schwab, C., Milinovich, G., Reichert, J., Ben Mahfoudh, K., Decker, T., Engel, M., Hai, B., Hainzl, E., Heider, S., *et al.* (2012). Phylotype-level 16S rRNA analysis reveals new bacterial indicators of health state in acute murine colitis. *The ISME journal 6*, 2091-2106.

Berry, D., Stecher, B., Schintlmeister, A., Reichert, J., Brugiroux, S., Wild, B., Wanek, W., Richter, A., Rauch, I., Decker, T., *et al.* (2013). Host-compound foraging by intestinal microbiota revealed by single-cell stable isotope probing. *Proceedings of the National Academy of Sciences of the United States of America 110*, 4720-4725.

Berven, F.S., Flikka, K., Jensen, H.B., and Eidhammer, I. (2004). BOMP: a program to predict integral beta-barrel outer membrane proteins encoded within genomes of Gram-negative bacteria. *Nucleic acids research 32*, W394-399.

Biagi, E., Nylund, L., Candela, M., Ostan, R., Bucci, L., Pini, E., Nikkila, J., Monti, D., Satokari, R., Franceschi, C*., et al.* (2010). Through ageing, and beyond: gut microbiota and inflammatory status in seniors and centenarians. *PloS one 5*, e10667.

Biasucci, G., Benenati, B., Morelli, L., Bessi, E., and Boehm, G. (2008). Cesarean delivery may affect the early biodiversity of intestinal bacteria. *The Journal of nutrition 138*, 1796S-1800S.

Bidart, G.N., Rodriguez-Diaz, J., Monedero, V., and Yebra, M.J. (2014). A unique gene cluster for the utilization of the mucosal and human milk-associated glycans galacto-N-biose and lacto-N-biose in Lactobacillus casei. *Molecular microbiology 93*, 521-538.

Bode, L. (2012). Human milk oligosaccharides: every baby needs a sugar mama. *Glycobiology 22*, 1147-1162.

Boesten, R., Schuren, F., Ben Amor, K., Haarman, M., Knol, J., and de Vos, W.M. (2011). Bifidobacterium population analysis in the infant gut by direct mapping of genomic hybridization patterns: potential for monitoring temporal development and effects of dietary regimens. *Microbial biotechnology 4*, 417-427.

Booijink, C.C., Boekhorst, J., Zoetendal, E.G., Smidt, H., Kleerebezem, M., and de Vos, W.M. (2010a). Metatranscriptome analysis of the human fecal microbiota reveals subject-specific expression profiles, with genes encoding proteins involved in carbohydrate metabolism being dominantly expressed. *Applied and environmental microbiology 76*, 5533-5540.

Booijink, C.C., El-Aidy, S., Rajilic-Stojanovic, M., Heilig, H.G., Troost, F.J., Smidt, H., Kleerebezem, M., De Vos, W.M., and Zoetendal, E.G. (2010b). High temporal and inter-individual variation detected in the human ileal microbiota. *Environmental microbiology 12*, 3213-3227.

Bos, M.P., Tefsen, B., Geurtsen, J., and Tommassen, J. (2004). Identification of an outer membrane protein required for the transport of lipopolysaccharide to the bacterial cell surface. *Proceedings of the National Academy of Sciences of the United States of America 101*, 9417-9422.

Bottacini, F., Medini, D., Pavesi, A., Turroni, F., Foroni, E., Riley, D., Giubellini, V., Tettelin, H., van Sinderen, D., and Ventura, M. (2010). Comparative genomics of the genus Bifidobacterium. *Microbiology 156*, 3243-3254.

Bron, P.A., van Baarlen, P., and Kleerebezem, M. (2012). Emerging molecular insights into the interaction between probiotics and the host intestinal mucosa. *Nature reviews Microbiology 10*, 66-78.

Brown, R.N., Romine, M.F., Schepmoes, A.A., Smith, R.D., and Lipton, M.S. (2010). Mapping the subcellular proteome of Shewanella oneidensis MR-1 using sarkosyl-

based fractionation and LC-MS/MS protein identification. *Journal of proteome research 9*, 4454-4463.

Brugman, S., Perdijk, O., van Neerven, R.J., and Savelkoul, H.F. (2015). Mucosal Immune Development in Early Life: Setting the Stage. *Archivum immunologiae et therapiae experimentalis*.

Bruzzese, E., Callegari, M.L., Raia, V., Viscovo, S., Scotto, R., Ferrari, S., Morelli, L., Buccigrossi, V., Lo Vecchio, A., Ruberto, E.*, et al.* (2014). Disrupted intestinal microbiota and intestinal inflammation in children with cystic fibrosis and its restoration with Lactobacillus GG: a randomised clinical trial. *PloS one 9*, e87796.

Cammarota, G., Ianiro, G., Bibbo, S., and Gasbarrini, A. (2014). Fecal microbiota transplantation: a new old kid on the block for the management of gut microbiota-related disease. *Journal of clinical gastroenterology 48 Suppl 1*, S80-84.

Candela, M., Rampelli, S., Turroni, S., Severgnini, M., Consolandi, C., De Bellis, G., Masetti, R., Ricci, G., Pession, A., and Brigidi, P. (2012). Unbalance of intestinal microbiota in atopic children. *BMC microbiology 12*, 95.

Carey, H.V., Walters, W.A., and Knight, R. (2013). Seasonal restructuring of the ground squirrel gut microbiota over the annual hibernation cycle. *American journal of physiology Regulatory, integrative and comparative physiology 304*, R33-42.

Caricilli, A.M., Picardi, P.K., de Abreu, L.L., Ueno, M., Prada, P.O., Ropelle, E.R., Hirabara, S.M., Castoldi, A., Vieira, P., Camara, N.O.*, et al.* (2011). Gut microbiota is a key modulator of insulin resistance in TLR 2 knockout mice. *PLoS biology 9*, e1001212.

Cartron, M.L., Maddocks, S., Gillingham, P., Craven, C.J., and Andrews, S.C. (2006). Feo--transport of ferrous iron into bacteria. *Biometals : an international journal on the role of metal ions in biology, biochemistry, and medicine 19*, 143-157.

Castanys-Munoz, E., Martin, M.J., and Prieto, P.A. (2013). 2'-fucosyllactose: an abundant, genetically determined soluble glycan present in human milk. *Nutr Rev 71*, 773-789.

Chapman, M.A., Grahn, M.F., Boyle, M.A., Hutton, M., Rogers, J., and Williams, N.S. (1994). Butyrate oxidation is impaired in the colonic mucosa of sufferers of quiescent ulcerative colitis. *Gut 35*, 73-76.

Chomczynski, P. (1993). A reagent for the single-step simultaneous isolation of RNA, DNA and proteins from cell and tissue samples. *BioTechniques 15*, 532-534, 536-537.

Claesson, M.J., Cusack, S., O'Sullivan, O., Greene-Diniz, R., de Weerd, H., Flannery, E., Marchesi, J.R., Falush, D., Dinan, T., Fitzgerald, G.*, et al.* (2011). Composition, variability, and temporal stability of the intestinal microbiota of the elderly. *Proceedings of the National Academy of Sciences of the United States of America 108 Suppl 1*, 4586-4591.

Clamp, J.R., and Ene, D. (1989). The gastric mucosal barrier. *Methods and findings in experimental and clinical pharmacology 11 Suppl 1*, 19-25.

Coats, S.R., Berezow, A.B., To, T.T., Jain, S., Bainbridge, B.W., Banani, K.P., and Darveau, R.P. (2011). The lipid A phosphate position determines differential host Toll-like receptor 4 responses to phylogenetically related symbiotic and pathogenic bacteria. *Infection and immunity 79*, 203-210.

Collado, M.C., Derrien, M., Isolauri, E., de Vos, W.M., and Salminen, S. (2007). Intestinal integrity and Akkermansia muciniphila, a mucin-degrading member of the intestinal microbiota present in infants, adults, and the elderly. *Applied and environmental microbiology 73*, 7767-7770.

Cone, R.A. (2009). Barrier properties of mucus. *Advanced drug delivery reviews 61*, 75-85.

Copley, S.D. (2012). Moonlighting is mainstream: paradigm adjustment required. *BioEssays : news and reviews in molecular, cellular and developmental biology 34*, 578-588.

Costello, E.K., Gordon, J.I., Secor, S.M., and Knight, R. (2010). Postprandial remodeling of the gut microbiota in Burmese pythons. *The ISME journal 4*, 1375-1385.

Cox, J., Hein, M.Y., Luber, C.A., Paron, I., Nagaraj, N., and Mann, M. (2014). Accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction, termed MaxLFQ. *Molecular & cellular proteomics : MCP 13*, 2513-2526.

Craig, L., and Li, J. (2008). Type IV pili: paradoxes in form and function. *Current opinion in structural biology 18*, 267-277.

Creecy, J.P., and Conway, T. (2015). Quantitative bacterial transcriptomics with RNA-seq. *Current opinion in microbiology 23*, 133-140.

Cummings, J.H., and Macfarlane, G.T. (1997). Role of intestinal bacteria in nutrient metabolism. *JPEN Journal of parenteral and enteral nutrition 21*, 357-365.

De Filippo, C., Cavalieri, D., Di Paola, M., Ramazzotti, M., Poullet, J.B., Massart, S., Collini, S., Pieraccini, G., and Lionetti, P. (2010). Impact of diet in shaping gut microbiota revealed by a comparative study in children from Europe and rural Africa. *Proceedings of the National Academy of Sciences of the United States of America 107*, 14691-14696.

De Preter, V., Hamer, H.M., Windey, K., and Verbeke, K. (2011). The impact of pre- and/or probiotics on human colonic metabolism: does it affect human health? *Molecular nutrition & food research 55*, 46-57.

de Vos, W.M. (2015). Microbial biofilms and the human intestinal microbiome. *Biofilms and Microbiomes 1, 15005*.

Dekker, J., Rossen, J.W., Buller, H.A., and Einerhand, A.W. (2002). The MUC family: an obituary. *Trends in biochemical sciences 27*, 126-131.

Derrien, M. (2007). Mucin utilisation and host interactions of the novel intestinal microbe Akkermansia muciniphila. *PhD thesis, Wageningen University*.

Derrien, M., Collado, M.C., Ben-Amor, K., Salminen, S., and de Vos, W.M. (2008). The Mucin degrader Akkermansia muciniphila is an abundant resident of the human intestinal tract. *Applied and environmental microbiology 74*, 1646-1648.

Derrien, M., Van Baarlen, P., Hooiveld, G., Norin, E., Muller, M., and de Vos, W.M. (2011). Modulation of Mucosal Immune Response, Tolerance, and Proliferation in Mice Colonized by the Mucin-Degrader Akkermansia muciniphila. *Frontiers in microbiology 2*, 166.

Derrien, M., van Passel, M.W., van de Bovenkamp, J.H., Schipper, R.G., de Vos, W.M., and Dekker, J. (2010). Mucin-bacterial interactions in the human oral cavity and digestive tract. *Gut Microbes 1*, 254-268.

Derrien, M., Vaughan, E.E., Plugge, C.M., and de Vos, W.M. (2004). Akkermansia muciniphila gen. nov., sp. nov., a human intestinal mucin-degrading bacterium. *International journal of systematic and evolutionary microbiology 54*, 1469-1476.

Dethlefsen, L., Eckburg, P.B., Bik, E.M., and Relman, D.A. (2006). Assembly of the human intestinal microbiota. *Trends in ecology & evolution 21*, 517-523.

Devos, D.P. (2014). PVC bacteria: variation of, but not exception to, the Gram-negative cell plan. *Trends in microbiology 22*, 14-20.

Dicksved, J., Halfvarson, J., Rosenquist, M., Jarnerot, 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. *The ISME journal 2*, 716-727.

Diepold, A., and Wagner, S. (2014). Assembly of the bacterial type III secretion machinery. *FEMS microbiology reviews 38*, 802-822.

Dill-McFarland, K.A., Neil, K.L., Zeng, A., Sprenger, R.J., Kurtz, C.C., Suen, G., and Carey, H.V. (2014). Hibernation alters the diversity and composition of mucosa-associated bacteria while enhancing antimicrobial defence in the gut of 13-lined ground squirrels. *Molecular ecology 23*, 4658-4669.

Doerfel, L.K., and Rodnina, M.V. (2013). Elongation factor P: Function and effects on bacterial fitness. *Biopolymers 99*, 837-845.

Dominguez-Bello, M.G., Costello, E.K., Contreras, M., Magris, M., Hidalgo, G., Fierer, N., and Knight, R. (2010). Delivery mode shapes the acquisition and structure of the initial microbiota across multiple body habitats in newborns. *Proceedings of the National Academy of Sciences of the United States of America 107*, 11971-11975.

Drell, T., Larionova, A., Voor, T., Simm, J., Julge, K., Heilman, K., Tillmann, V., Stsepetova, J., and Sepp, E. (2015). Differences in Gut Microbiota Between Atopic and Healthy Children. *Current microbiology*.

Dubourg, G., Lagier, J.C., Armougom, F., Robert, C., Audoly, G., Papazian, L., and Raoult, D. (2013). High-level colonisation of the human gut by Verrucomicrobia following broad-spectrum antibiotic treatment. *International journal of antimicrobial agents 41*, 149-155.

Duerr, C.U., and Hornef, M.W. (2012). The mammalian intestinal epithelium as integral player in the establishment and maintenance of host-microbial homeostasis. *Seminars in immunology 24*, 25-35.

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

Dunne, C., O'Mahony, L., Murphy, L., Thornton, G., Morrissey, D., O'Halloran, S., Feeney, M., Flynn, S., Fitzgerald, G., Daly, C.*, et al.* (2001). In vitro selection criteria for probiotic bacteria of human origin: correlation with in vivo findings. *Am J Clin Nutr 73*, 386S-392S.

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

Ehara, M., Ichinose, Y., Iwami, M., Utsunomiya, A., Shimodori, S., Kangethe, S.K., Neves, B.C., Supawat, K., and Nakamura, S. (1993). Immunogenicity of Vibrio cholerae O1 fimbriae in animal and human cholera. *Microbiology and immunology 37*, 679-688.

El Aidy, S., Van den Abbeele, P., Van de Wiele, T., Louis, P., and Kleerebezem, M. (2013). Intestinal colonization: how key microbial players become established in this dynamic process: microbial metabolic activities and the interplay between the host and microbes. *BioEssays : news and reviews in molecular, cellular and developmental biology 35*, 913-923.

El Aidy, S., van den Bogert, B., and Kleerebezem, M. (2015). The small intestine microbiota, nutritional modulation and relevance for health. *Current opinion in biotechnology 32*, 14-20.

Elhenawy, W., Debelyy, M.O., and Feldman, M.F. (2014). Preferential packing of acidic glycosidases and proteases into Bacteroides outer membrane vesicles. *mBio 5*, e00909-00914.

Ellekilde, M., Krych, L., Hansen, C.H., Hufeldt, M.R., Dahl, K., Hansen, L.H., Sorensen, S.J., Vogensen, F.K., Nielsen, D.S., and Hansen, A.K. (2014). Characterization of the gut microbiota in leptin deficient obese mice - Correlation to inflammatory and diabetic parameters. *Research in veterinary science 96*, 241-250.

Engfer, M.B., Stahl, B., Finke, B., Sawatzki, G., and Daniel, H. (2000). Human milk oligosaccharides are resistant to enzymatic hydrolysis in the upper gastrointestinal tract. *Am J Clin Nutr 71*, 1589-1596.

Erridge, C., Pridmore, A., Eley, A., Stewart, J., and Poxton, I.R. (2004). Lipopolysaccharides of Bacteroides fragilis, Chlamydia trachomatis and Pseudomonas aeruginosa signal via toll-like receptor 2. *Journal of medical microbiology 53*, 735-740.

Everard, A., Belzer, C., Geurts, L., Ouwerkerk, J.P., Druart, C., Bindels, L.B., Guiot, Y., Derrien, M., Muccioli, G.G., Delzenne, N.M.*, et al.* (2013). Cross-talk

between Akkermansia muciniphila and intestinal epithelium controls diet-induced obesity. *Proceedings of the National Academy of Sciences of the United States of America 110*, 9066-9071.

Everard, A., Lazarevic, V., Derrien, M., Girard, M., Muccioli, G.G., Neyrinck, A.M., Possemiers, S., Van Holle, A., Francois, P., de Vos, W.M.*, et al.* (2011). Responses of gut microbiota and glucose and lipid metabolism to prebiotics in genetic obese and diet-induced leptin-resistant mice. *Diabetes 60*, 2775-2786.

Fagarasan, S., Kawamoto, S., Kanagawa, O., and Suzuki, K. (2010). Adaptive immune regulation in the gut: T cell-dependent and T cell-independent IgA synthesis. *Annual review of immunology 28*, 243-273.

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

Faith, J.J., McNulty, N.P., Rey, F.E., and Gordon, J.I. (2011). Predicting a human gut microbiota's response to diet in gnotobiotic mice. *Science 333*, 101-104.

Fanning, S., Hall, L.J., Cronin, M., Zomer, A., MacSharry, J., Goulding, D., Motherway, M.O., Shanahan, F., Nally, K., Dougan, G.*, et al.* (2012). Bifidobacterial surface-exopolysaccharide facilitates commensal-host interaction through immune modulation and pathogen protection. *Proceedings of the National Academy of Sciences of the United States of America 109*, 2108-2113.

Favier, C.F., de Vos, W.M., and Akkermans, A.D. (2003). Development of bacterial and bifidobacterial communities in feces of newborn babies. *Anaerobe 9*, 219-229.

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

Feehily, C., and Karatzas, K.A. (2013). Role of glutamate metabolism in bacterial responses towards acid and other stresses. *Journal of applied microbiology 114*, 11-24.

Fernandez, L., Langa, S., Martin, V., Maldonado, A., Jimenez, E., Martin, R., and Rodriguez, J.M. (2013). The human milk microbiota: origin and potential roles in health and disease. *Pharmacological research : the official journal of the Italian Pharmacological Society 69*, 1-10.

Ferrer, M., Martins dos Santos, V.A., Ott, S.J., and Moya, A. (2014). Gut microbiota disturbance during antibiotic therapy: a multi-omic approach. *Gut Microbes 5*, 64-70.

Filip, C., Fletcher, G., Wulff, J.L., and Earhart, C.F. (1973). Solubilization of the cytoplasmic membrane of Escherichia coli by the ionic detergent sodium-lauryl sarcosinate. *Journal of bacteriology 115*, 717-722.

Finn, R.D., Bateman, A., Clements, J., Coggill, P., Eberhardt, R.Y., Eddy, S.R., Heger, A., Hetherington, K., Holm, L., Mistry, J.*, et al.* (2014). Pfam: the protein families database. *Nucleic acids research 42*, D222-230.

Fischbach, M.A., and Sonnenburg, J.L. (2011). Eating for two: how metabolism establishes interspecies interactions in the gut. *Cell host & microbe 10*, 336-347.

Flamholz, A., Noor, E., Bar-Even, A., and Milo, R. (2012). eQuilibrator--the biochemical thermodynamics calculator. *Nucleic acids research 40*, D770-775.

Fontana, L., Bermudez-Brito, M., Plaza-Diaz, J., Munoz-Quezada, S., and Gil, A. (2013). Sources, isolation, characterisation and evaluation of probiotics. *The British journal of nutrition 109 Suppl 2*, S35-50.

Forde, B.M., and O'Toole, P.W. (2013). Next-generation sequencing technologies and their impact on microbial genomics. *Briefings in functional genomics 12*, 440-453.

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

Frankel, S., Sohn, R., and Leinwand, L. (1991). The use of sarkosyl in generating soluble protein after bacterial expression. *Proceedings of the National Academy of Sciences of the United States of America 88*, 1192-1196.

Fritz, J.V., Desai, M.S., Shah, P., Schneider, J.G., and Wilmes, P. (2013). From meta-omics to causality: experimental models for human microbiome research. *Microbiome 1*, 14.

Galdiero, S., Falanga, A., Cantisani, M., Tarallo, R., Della Pepa, M.E., D'Oriano, V., and Galdiero, M. (2012). Microbe-host interactions: structure and role of Gram-negative bacterial porins. *Current protein & peptide science 13*, 843-854.

Ganesh, B.P., Klopfleisch, R., Loh, G., and Blaut, M. (2013). Commensal Akkermansia muciniphila exacerbates gut inflammation in Salmonella Typhimurium-infected gnotobiotic mice. *PloS one 8*, e74963.

Gardner, S.G., Johns, K.D., Tanner, R., and McCleary, W.R. (2014). The PhoU protein from Escherichia coli interacts with PhoR, PstB, and metals to form a phosphate-signaling complex at the membrane. *Journal of bacteriology 196*, 1741-1752.

Geens, M.M., and Niewold, T.A. (2010). Preliminary Characterization of the Transcriptional Response of the Porcine Intestinal Cell Line IPEC-J2 to Enterotoxigenic Escherichia coli, Escherichia coli, and E. coli Lipopolysaccharide. *Comparative and functional genomics 2010*, 469583.

Geirnaert, A., Steyaert, A., Eeckhaut, V., Debruyne, B., Arends, J.B., Van Immerseel, F., Boon, N., and Van de Wiele, T. (2014). Butyricicoccus pullicaecorum, a butyrate producer with probiotic potential, is intrinsically tolerant to stomach and small intestine conditions. *Anaerobe 30*, 70-74.

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

Gill, S.R., Pop, M., Deboy, R.T., Eckburg, P.B., Turnbaugh, P.J., Samuel, B.S., Gordon, J.I., Relman, D.A., Fraser-Liggett, C.M., and Nelson, K.E. (2006). Metagenomic analysis of the human distal gut microbiome. *Science 312*, 1355-1359.

Goel, A., Eckhardt, T.H., Puri, P., de Jong, A., Santos, F.B., Giera, M., Fusetti, F., de Vos, W.M., Kok, J., Poolman, B.*, et al.* (2015). Protein costs do not explain evolution of metabolic strategies and regulation of ribosomal content. *Molecular microbiology*.

Gonzalez, R., Klaassens, E.S., Malinen, E., de Vos, W.M., and Vaughan, E.E. (2008). Differential transcriptional response of Bifidobacterium longum to human milk, formula milk, and galactooligosaccharide. *Applied and environmental microbiology 74*, 4686-4694.

Gosalbes, M.J., Durban, A., Pignatelli, M., Abellan, J.J., Jimenez-Hernandez, N., Perez-Cobas, A.E., Latorre, A., and Moya, A. (2011). Metatranscriptomic approach to analyze the functional human gut microbiota. *PloS one 6*, e17447.

Gouyer, V., Gottrand, F., and Desseyn, J.L. (2011). The extraordinarily complex but highly structured organization of intestinal mucus-gel unveiled in multicolor images. *PloS one 6*, e18761.

Grabinger, T., Luks, L., Kostadinova, F., Zimberlin, C., Medema, J.P., Leist, M., and Brunner, T. (2014). Ex vivo culture of intestinal crypt organoids as a model system for assessing cell death induction in intestinal epithelial cells and enteropathy. *Cell death & disease 5*, e1228.

Graham, C., McMullan, G., and Graham, R.L. (2011). Proteomics in the microbial sciences. *Bioeng Bugs 2*, 17-30.

Grijpstra, J., Arenas, J., Rutten, L., and Tommassen, J. (2013). Autotransporter secretion: varying on a theme. *Research in microbiology 164*, 562-582.

Grosskopf, T., and Soyer, O.S. (2014). Synthetic microbial communities. *Current opinion in microbiology 18*, 72-77.

Grzeskowiak, L., Collado, M.C., Mangani, C., Maleta, K., Laitinen, K., Ashorn, P., Isolauri, E., and Salminen, S. (2012). Distinct gut microbiota in southeastern African and northern European infants. *Journal of pediatric gastroenterology and nutrition 54*, 812-816.

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

Guilloteau, P., Martin, L., Eeckhaut, V., Ducatelle, R., Zabielski, R., and Van Immerseel, F. (2010). From the gut to the peripheral tissues: the multiple effects of butyrate. *Nutrition research reviews 23*, 366-384.

Gum, J.R., Jr. (1992). Mucin genes and the proteins they encode: structure, diversity, and regulation. *American journal of respiratory cell and molecular biology 7*, 557-564.

Gupta, R.S., Bhandari, V., and Naushad, H.S. (2012). Molecular Signatures for the PVC Clade (Planctomycetes, Verrucomicrobia, Chlamydiae, and Lentisphaerae) of

Bacteria Provide Insights into Their Evolutionary Relationships. *Frontiers in microbiology 3*, 327.

Haft, D.H., Paulsen, I.T., Ward, N., and Selengut, J.D. (2006). Exopolysaccharide-associated protein sorting in environmental organisms: the PEP-CTERM/EpsH system. Application of a novel phylogenetic profiling heuristic. *BMC biology 4*, 29.

Haft, D.H., Payne, S.H., and Selengut, J.D. (2012). Archaeosortases and exosortases are widely distributed systems linking membrane transit with posttranslational modification. *Journal of bacteriology 194*, 36-48.

Hakansson, A., Tormo-Badia, N., Baridi, A., Xu, J., Molin, G., Hagslatt, M.L., Karlsson, C., Jeppsson, B., Cilio, C.M., and Ahrne, S. (2014). Immunological alteration and changes of gut microbiota after dextran sulfate sodium (DSS) administration in mice. *Clinical and experimental medicine*.

Hamer, H.M., Jonkers, D.M., Loof, A., Vanhoutvin, S.A., Troost, F.J., Venema, K., Kodde, A., Koek, G.H., Schipper, R.G., van Heerde, W.L.*, et al.* (2009). Analyses of human colonic mucus obtained by an in vivo sampling technique. *Digestive and liver disease : official journal of the Italian Society of Gastroenterology and the Italian Association for the Study of the Liver 41*, 559-564.

Hansen, C.H., Krych, L., Nielsen, D.S., Vogensen, F.K., Hansen, L.H., Sorensen, S.J., Buschard, K., and Hansen, A.K. (2012). Early life treatment with vancomycin propagates Akkermansia muciniphila and reduces diabetes incidence in the NOD mouse. *Diabetologia 55*, 2285-2294.

Hartmann, P., Chen, P., Wang, H.J., Wang, L., McCole, D.F., Brandl, K., Starkel, P., Belzer, C., Hellerbrand, C., Tsukamoto, H.*, et al.* (2013). Deficiency of intestinal mucin-2 ameliorates experimental alcoholic liver disease in mice. *Hepatology 58*, 108-119.

Haussmann, U., Qi, S.W., Wolters, D., Rogner, M., Liu, S.J., and Poetsch, A. (2009). Physiological adaptation of Corynebacterium glutamicum to benzoate as alternative carbon source - a membrane proteome-centric view. *Proteomics 9*, 3635-3651.

Heinken, A., and Thiele, I. (2015). Anoxic Conditions Promote Species-Specific Mutualism between Gut Microbes In Silico. *Applied and environmental microbiology 81*, 4049-4061.

Henderson, B., and Martin, A. (2011). Bacterial virulence in the moonlight: multitasking bacterial moonlighting proteins are virulence determinants in infectious disease. *Infection and immunity 79*, 3476-3491.

Hildebrand, F., Ebersbach, T., Nielsen, H.B., Li, X., Sonne, S.B., Bertalan, M., Dimitrov, P., Madsen, L., Qin, J., Wang, J.*, et al.* (2012). A comparative analysis of the intestinal metagenomes present in guinea pigs (Cavia porcellus) and humans (Homo sapiens). *BMC Genomics 13*, 514.

Hildebrandt, M.A., Hoffmann, C., Sherrill-Mix, S.A., Keilbaugh, S.A., Hamady, M., Chen, Y.Y., Knight, R., Ahima, R.S., Bushman, F., and Wu, G.D. (2009). High-fat

diet determines the composition of the murine gut microbiome independently of obesity. *Gastroenterology 137*, 1716-1724 e1711-1712.

Hill, C., Guarner, F., Reid, G., Gibson, G.R., Merenstein, D.J., Pot, B., Morelli, L., Canani, R.B., Flint, H.J., Salminen, S., *et al.* (2014). Expert consensus document. The International Scientific Association for Probiotics and Prebiotics consensus statement on the scope and appropriate use of the term probiotic. *Nature reviews Gastroenterology & hepatology 11*, 506-514.

Hobb, R.I., Fields, J.A., Burns, C.M., and Thompson, S.A. (2009). Evaluation of procedures for outer membrane isolation from Campylobacter jejuni. *Microbiology 155*, 979-988.

Hooper, L.V., Littman, D.R., and Macpherson, A.J. (2012). Interactions between the microbiota and the immune system. *Science 336*, 1268-1273.

Hooper, L.V., and Macpherson, A.J. (2010). Immune adaptations that maintain homeostasis with the intestinal microbiota. *Nature reviews Immunology 10*, 159-169.

Hooper, L.V., Xu, J., Falk, P.G., Midtvedt, T., and 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 Sciences of the United States of America 96*, 9833-9838.

Hoskins, L.C., Agustines, M., McKee, W.B., Boulding, E.T., Kriaris, M., and Niedermeyer, G. (1985). Mucin degradation in human colon ecosystems. Isolation and properties of fecal strains that degrade ABH blood group antigens and oligosaccharides from mucin glycoproteins. *The Journal of clinical investigation 75*, 944-953.

Isanapong, J., Sealy Hambright, W., Willis, A.G., Boonmee, A., Callister, S.J., Burnum, K.E., Pasa-Tolic, L., Nicora, C.D., Wertz, J.T., Schmidt, T.M., *et al.* (2013). Development of an ecophysiological model for Diplosphaera colotermitum TAV2, a termite hindgut Verrucomicrobium. *The ISME journal 7*, 1803-1813.

Jackson, A.P., Thomas, G.H., Parkhill, J., and Thomson, N.R. (2009). Evolutionary diversification of an ancient gene family (rhs) through C-terminal displacement. *BMC Genomics 10*, 584.

Jakobsson, H.E., Rodriguez-Pineiro, A.M., Schutte, A., Ermund, A., Boysen, P., Bemark, M., Sommer, F., Backhed, F., Hansson, G.C., and Johansson, M.E. (2015). The composition of the gut microbiota shapes the colon mucus barrier. *EMBO reports 16*, 164-177.

Jalanka-Tuovinen, J., Salonen, A., Nikkila, J., Immonen, O., Kekkonen, R., Lahti, L., Palva, A., and de Vos, W.M. (2011). Intestinal microbiota in healthy adults: temporal analysis reveals individual and common core and relation to intestinal symptoms. *PloS one 6*, e23035.

James, S.L., Christophersen, C.T., Bird, A.R., Conlon, M.A., Rosella, O., Gibson, P.R., and Muir, J.G. (2014). Abnormal fibre usage in UC in remission. *Gut*.

Jansson, J., Willing, B., Lucio, M., Fekete, A., Dicksved, J., Halfvarson, J., Tysk, C., and Schmitt-Kopplin, P. (2009). Metabolomics reveals metabolic biomarkers of Crohn's disease. *PloS one 4*, e6386.

Johansson, M.E., Larsson, J.M., and Hansson, G.C. (2011). The two mucus layers of colon are organized by the MUC2 mucin, whereas the outer layer is a legislator of host-microbial interactions. *Proceedings of the National Academy of Sciences of the United States of America 108 Suppl 1*, 4659-4665.

Johansson, M.E., Phillipson, M., Petersson, J., Velcich, A., Holm, L., and Hansson, G.C. (2008). The inner of the two Muc2 mucin-dependent mucus layers in colon is devoid of bacteria. *Proceedings of the National Academy of Sciences of the United States of America 105*, 15064-15069.

Johansson, M.E., Sjovall, H., and Hansson, G.C. (2013). The gastrointestinal mucus system in health and disease. *Nature reviews Gastroenterology & hepatology 10*, 352-361.

Juncker, A.S., Willenbrock, H., Von Heijne, G., Brunak, S., Nielsen, H., and Krogh, A. (2003). Prediction of lipoprotein signal peptides in Gram-negative bacteria. *Protein science : a publication of the Protein Society 12*, 1652-1662.

Kainulainen, V., Loimaranta, V., Pekkala, A., Edelman, S., Antikainen, J., Kylvaja, R., Laaksonen, M., Laakkonen, L., Finne, J., and Korhonen, T.K. (2012). Glutamine synthetase and glucose-6-phosphate isomerase are adhesive moonlighting proteins of Lactobacillus crispatus released by epithelial cathelicidin LL-37. *Journal of bacteriology 194*, 2509-2519.

Kainulainen, V., Reunanen, J., Hiippala, K., Guglielmetti, S., Vesterlund, S., Palva, A., and Satokari, R. (2013). BopA does not have a major role in the adhesion of Bifidobacterium bifidum to intestinal epithelial cells, extracellular matrix proteins, and mucus. *Applied and environmental microbiology 79*, 6989-6997.

Kainulainen, V., Tang, Y., Spillmann, T., Kilpinen, S., Reunanen, J., Saris, P.E., and Satokari, R. (2015). The canine isolate Lactobacillus acidophilus LAB20 adheres to intestinal epithelium and attenuates LPS-induced IL-8 secretion of enterocytes in vitro. *BMC microbiology 15*, 4.

Kamada, N., Seo, S.U., Chen, G.Y., and Nunez, G. (2013). Role of the gut microbiota in immunity and inflammatory disease. *Nature reviews Immunology 13*, 321-335.

Kamneva, O.K., Knight, S.J., Liberles, D.A., and Ward, N.L. (2012). Analysis of genome content evolution in pvc bacterial super-phylum: assessment of candidate genes associated with cellular organization and lifestyle. *Genome biology and evolution 4*, 1375-1390.

Kang, C.S., Ban, M., Choi, E.J., Moon, H.G., Jeon, J.S., Kim, D.K., Park, S.K., Jeon, S.G., Roh, T.Y., Myung, S.J., *et al.* (2013). Extracellular vesicles derived from gut microbiota, especially Akkermansia muciniphila, protect the progression of dextran sulfate sodium-induced colitis. *PloS one 8*, e76520.

Kang, S., Denman, S.E., Morrison, M., Yu, Z., Dore, J., Leclerc, M., and McSweeney, C.S. (2010). Dysbiosis of fecal microbiota in Crohn's disease patients

as revealed by a custom phylogenetic microarray. *Inflammatory bowel diseases 16*, 2034-2042.

Kant, R., de Vos, W.M., Palva, A., and Satokari, R. (2014). Immunostimulatory CpG motifs in the genomes of gut bacteria and their role in human health and disease. *Journal of medical microbiology 63*, 293-308.

Karlsson, C.L., Onnerfalt, J., Xu, J., Molin, G., Ahrne, S., and Thorngren-Jerneck, K. (2012). The microbiota of the gut in preschool children with normal and excessive body weight. *Obesity 20*, 2257-2261.

Karlsson, F.H., Tremaroli, V., Nookaew, I., Bergstrom, G., Behre, C.J., Fagerberg, B., Nielsen, J., and Backhed, F. (2013). Gut metagenome in European women with normal, impaired and diabetic glucose control. *Nature 498*, 99-103.

Karp, P.D., Paley, S.M., Krummenacker, M., Latendresse, M., Dale, J.M., Lee, T.J., Kaipa, P., Gilham, F., Spaulding, A., Popescu, L.*, et al.* (2010). Pathway Tools version 13.0: integrated software for pathway/genome informatics and systems biology. *Briefings in bioinformatics 11*, 40-79.

Kasubuchi, M., Hasegawa, S., Hiramatsu, T., Ichimura, A., and Kimura, I. (2015). Dietary gut microbial metabolites, short-chain fatty acids, and host metabolic regulation. *Nutrients 7*, 2839-2849.

Khan, J., Iiboshi, Y., Cui, L., Wasa, M., Sando, K., Takagi, Y., and Okada, A. (1999). Alanyl-glutamine-supplemented parenteral nutrition increases luminal mucus gel and decreases permeability in the rat small intestine. *JPEN Journal of parenteral and enteral nutrition 23*, 24-31.

Klaassens, E.S., Ben-Amor, K., Vriesema, A., Vaughan, E.E., and de Vos, W. (2011). The fecal bifidobacterial transcriptome of adults: a microarray approach. *Gut Microbes 2*, 217-226.

Klaassens, E.S., Boesten, R.J., Haarman, M., Knol, J., Schuren, F.H., Vaughan, E.E., and de Vos, W.M. (2009). Mixed-species genomic microarray analysis of fecal samples reveals differential transcriptional responses of bifidobacteria in breast- and formula-fed infants. *Applied and environmental microbiology 75*, 2668-2676.

Klaassens, E.S., de Vos, W.M., and Vaughan, E.E. (2007). Metaproteomics approach to study the functionality of the microbiota in the human infant gastrointestinal tract. *Applied and environmental microbiology 73*, 1388-1392.

Knol, J., Scholtens, P., Kafka, C., Steenbakkers, J., Gro, S., Helm, K., Klarczyk, M., Schopfer, H., Bockler, H.M., and Wells, J. (2005). Colon microflora in infants fed formula with galacto- and fructo-oligosaccharides: more like breast-fed infants. *Journal of pediatric gastroenterology and nutrition 40*, 36-42.

Kobata, A. (2010). Structures and application of oligosaccharides in human milk. *Proceedings of the Japan Academy Series B, Physical and biological sciences 86*, 731-747.

Koenig, J.E., Spor, A., Scalfone, N., Fricker, A.D., Stombaugh, J., Knight, R., Angenent, L.T., and Ley, R.E. (2011). Succession of microbial consortia in the

developing infant gut microbiome. *Proceedings of the National Academy of Sciences of the United States of America 108 Suppl 1*, 4578-4585.

Kolmeder, C.A., de Been, M., Nikkila, J., Ritamo, I., Matto, J., Valmu, L., Salojarvi, J., Palva, A., Salonen, A., and de Vos, W.M. (2012). Comparative metaproteomics and diversity analysis of human intestinal microbiota testifies for its temporal stability and expression of core functions. *PloS one 7*, e29913.

Kolmeder, C.A., and de Vos, W.M. (2014). Metaproteomics of our microbiome - developing insight in function and activity in man and model systems. *Journal of proteomics 97*, 3-16.

Konstantinov, S.R., Smidt, H., de Vos, W.M., Bruijns, S.C., Singh, S.K., Valence, F., Molle, D., Lortal, S., Altermann, E., Klaenhammer, T.R.*, et al.* (2008). S layer protein A of Lactobacillus acidophilus NCFM regulates immature dendritic cell and T cell functions. *Proceedings of the National Academy of Sciences of the United States of America 105*, 19474-19479.

Koropatkin, N.M., Cameron, E.A., and Martens, E.C. (2012). How glycan metabolism shapes the human gut microbiota. *Nature reviews Microbiology 10*, 323-335.

Koskiniemi, S., Lamoureux, J.G., Nikolakakis, K.C., t'Kint de Roodenbeke, C., Kaplan, M.D., Low, D.A., and Hayes, C.S. (2013). Rhs proteins from diverse bacteria mediate intercellular competition. *Proceedings of the National Academy of Sciences of the United States of America 110*, 7032-7037.

Kovatcheva-Datchary, P., Egert, M., Maathuis, A., Rajilic-Stojanovic, M., de Graaf, A.A., Smidt, H., de Vos, W.M., and Venema, K. (2009). Linking phylogenetic identities of bacteria to starch fermentation in an in vitro model of the large intestine by RNA-based stable isotope probing. *Environmental microbiology 11*, 914-926.

Krogh, A., Larsson, B., von Heijne, G., and Sonnhammer, E.L. (2001). Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *Journal of molecular biology 305*, 567-580.

Krych, L., Hansen, C.H., Hansen, A.K., van den Berg, F.W., and Nielsen, D.S. (2013). Quantitatively different, yet qualitatively alike: a meta-analysis of the mouse core gut microbiome with a view towards the human gut microbiome. *PloS one 8*, e62578.

Kuhn, K.A., and Stappenbeck, T.S. (2013). Peripheral education of the immune system by the colonic microbiota. *Seminars in immunology 25*, 364-369.

Kumar, H., Salminen, S., Verhagen, H., Rowland, I., Heimbach, J., Banares, S., Young, T., Nomoto, K., and Lalonde, M. (2015). Novel probiotics and prebiotics: road to the market. *Current opinion in biotechnology 32*, 99-103.

Kurokawa, K., Itoh, T., Kuwahara, T., Oshima, K., Toh, H., Toyoda, A., Takami, H., Morita, H., Sharma, V.K., Srivastava, T.P.*, et al.* (2007). Comparative metagenomics revealed commonly enriched gene sets in human gut microbiomes. *DNA Res 14*, 169-181.

Lacroix, C., de Wouters, T., and Chassard, C. (2015). Integrated multi-scale strategies to investigate nutritional compounds and their effect on the gut microbiota. *Current opinion in biotechnology 32C*, 149-155.

Lange, K., Hugenholtz, F., Jonathan, M.C., Schols, H.A., Kleerebezem, M., Smidt, H., Muller, M., and Hooiveld, G.J. (2015). Comparison of the effects of five dietary fibers on mucosal transcriptional profiles, and luminal microbiota composition and SCFA concentrations in murine colon. *Molecular nutrition & food research*.

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

Larsen, N., Boye, M., Siegumfeldt, H., and Jakobsen, M. (2006). Differential expression of proteins and genes in the lag phase of Lactococcus lactis subsp. lactis grown in synthetic medium and reconstituted skim milk. *Applied and environmental microbiology 72*, 1173-1179.

Le Marechal, C., Peton, V., Ple, C., Vroland, C., Jardin, J., Briard-Bion, V., Durant, G., Chuat, V., Loux, V., Foligne, B*., et al.* (2015). Surface proteins of Propionibacterium freudenreichii are involved in its anti-inflammatory properties. *Journal of proteomics 113*, 447-461.

Lee, H., and Ko, G. (2014). Effect of metformin on metabolic improvement and gut microbiota. *Applied and environmental microbiology 80*, 5935-5943.

Lee, K.C., Webb, R.I., Janssen, P.H., Sangwan, P., Romeo, T., Staley, J.T., and Fuerst, J.A. (2009). Phylum Verrucomicrobia representatives share a compartmentalized cell plan with members of bacterial phylum Planctomycetes. *BMC microbiology 9*, 5.

Ley, R.E., Backhed, F., Turnbaugh, P., Lozupone, C.A., Knight, R.D., and Gordon, J.I. (2005). Obesity alters gut microbial ecology. *Proceedings of the National Academy of Sciences of the United States of America 102*, 11070-11075.

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

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., and Genome Project Data Processing, S. (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics 25*, 2078-2079.

Liu, X., Fan, H., Ding, X., Hong, Z., Nei, Y., Liu, Z., Li, G., and Guo, H. (2014). Analysis of the gut microbiota by high-throughput sequencing of the V5-V6 regions of the 16S rRNA gene in donkey. *Current microbiology 68*, 657-662.

Lopetuso, L.R., Scaldaferri, F., Franceschi, F., and Gasbarrini, A. (2014). The gastrointestinal microbiome - functional interference between stomach and intestine. *Best practice & research Clinical gastroenterology 28*, 995-1002.

Louis, P., and Flint, H.J. (2009). Diversity, metabolism and microbial ecology of butyrate-producing bacteria from the human large intestine. *FEMS microbiology letters 294*, 1-8.

Louis, P., Hold, G.L., and Flint, H.J. (2014). The gut microbiota, bacterial metabolites and colorectal cancer. *Nature reviews Microbiology 12*, 661-672.

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

Lu, J., Boeren, S., de Vries, S.C., van Valenberg, H.J., Vervoort, J., and Hettinga, K. (2011). Filter-aided sample preparation with dimethyl labeling to identify and quantify milk fat globule membrane proteins. *Journal of proteomics 75*, 34-43.

Lu, P., Ma, D., Chen, Y., Guo, Y., Chen, G.Q., Deng, H., and Shi, Y. (2013). L-glutamine provides acid resistance for Escherichia coli through enzymatic release of ammonia. *Cell research 23*, 635-644.

Lukovac, S., Belzer, C., Pellis, L., Keijser, B.J., de Vos, W.M., Montijn, R.C., and Roeselers, G. (2014). Differential modulation by Akkermansia muciniphila and Faecalibacterium prausnitzii of host peripheral lipid metabolism and histone acetylation in mouse gut organoids. *mBio 5*.

Lyra, A., Forssten, S., Rolny, P., Wettergren, Y., Lahtinen, S.J., Salli, K., Cedgard, L., Odin, E., Gustavsson, B., and Ouwehand, A.C. (2012). Comparison of bacterial quantities in left and right colon biopsies and faeces. *World journal of gastroenterology : WJG 18*, 4404-4411.

Ma, D., Lu, P., Yan, C., Fan, C., Yin, P., Wang, J., and Shi, Y. (2012). Structure and mechanism of a glutamate-GABA antiporter. *Nature 483*, 632-636.

Macfarlane, G.T., and Macfarlane, S. (2011). Fermentation in the human large intestine: its physiologic consequences and the potential contribution of prebiotics. *Journal of clinical gastroenterology 45 Suppl*, S120-127.

Macfarlane, G.T., and Macfarlane, S. (2012). Bacteria, colonic fermentation, and gastrointestinal health. *Journal of AOAC International 95*, 50-60.

Machiels, K., Joossens, M., Sabino, J., De Preter, V., Arijs, I., Eeckhaut, V., Ballet, V., Claes, K., Van Immerseel, F., Verbeke, K.*, et al.* (2014). A decrease of the butyrate-producing species Roseburia hominis and Faecalibacterium prausnitzii defines dysbiosis in patients with ulcerative colitis. *Gut 63*, 1275-1283.

Makivuokko, H., Tiihonen, K., Tynkkynen, S., Paulin, L., and Rautonen, N. (2010). The effect of age and non-steroidal anti-inflammatory drugs on human intestinal microbiota composition. *The British journal of nutrition 103*, 227-234.

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

Mann, E.R., Landy, J.D., Bernardo, D., Peake, S.T., Hart, A.L., Al-Hassi, H.O., and Knight, S.C. (2013). Intestinal dendritic cells: their role in intestinal inflammation, manipulation by the gut microbiota and differences between mice and men. *Immunology letters 150*, 30-40.

Marcobal, A., Barboza, M., Froehlich, J.W., Block, D.E., German, J.B., Lebrilla, C.B., and Mills, D.A. (2010). Consumption of human milk oligosaccharides by gut-related microbes. *Journal of agricultural and food chemistry 58*, 5334-5340.

Marcobal, A., Barboza, M., Sonnenburg, E.D., Pudlo, N., Martens, E.C., Desai, P., Lebrilla, C.B., Weimer, B.C., Mills, D.A., German, J.B., *et al.* (2011). Bacteroides in the infant gut consume milk oligosaccharides via mucus-utilization pathways. *Cell host & microbe 10*, 507-514.

Mariat, D., Firmesse, O., Levenez, F., Guimaraes, V., Sokol, H., Dore, J., Corthier, G., and Furet, J.P. (2009). The Firmicutes/Bacteroidetes ratio of the human microbiota changes with age. *BMC microbiology 9*, 123.

Marietta, E.V., Gomez, A.M., Yeoman, C., Tilahun, A.Y., Clark, C.R., Luckey, D.H., Murray, J.A., White, B.A., Kudva, Y.C., and Rajagopalan, G. (2013). Low incidence of spontaneous type 1 diabetes in non-obese diabetic mice raised on gluten-free diets is associated with changes in the intestinal microbiome. *PloS one 8*, e78687.

Martens, E.C., Chiang, H.C., and Gordon, J.I. (2008). Mucosal glycan foraging enhances fitness and transmission of a saccharolytic human gut bacterial symbiont. *Cell host & microbe 4*, 447-457.

Martin, R., Miquel, S., Chain, F., Natividad, J.M., Jury, J., Lu, J., Sokol, H., Theodorou, V., Bercik, P., Verdu, E.F., *et al.* (2015). Faecalibacterium prausnitzii prevents physiological damages in a chronic low-grade inflammation murine model. *BMC microbiology 15*, 67.

Martin, R., Nauta, A.J., Ben Amor, K., Knippels, L.M., Knol, J., and Garssen, J. (2010). Early life: gut microbiota and immune development in infancy. *Beneficial microbes 1*, 367-382.

Maslowski, K.M., Vieira, A.T., Ng, A., Kranich, J., Sierro, F., Yu, D., Schilter, H.C., Rolph, M.S., Mackay, F., Artis, D., *et al.* (2009). Regulation of inflammatory responses by gut microbiota and chemoattractant receptor GPR43. *Nature 461*, 1282-1286.

Matamoros, S., Gras-Leguen, C., Le Vacon, F., Potel, G., and de La Cochetiere, M.F. (2013). Development of intestinal microbiota in infants and its impact on health. *Trends in microbiology 21*, 167-173.

Mattam, A.J., and Yazdani, S.S. (2013). Engineering E. coli strain for conversion of short chain fatty acids to bioalcohols. *Biotechnology for biofuels 6*, 128.

McGuckin, M.A., Linden, S.K., Sutton, P., and Florin, T.H. (2011). Mucin dynamics and enteric pathogens. *Nature reviews Microbiology 9*, 265-278.

McLean, M.H., Dieguez, D., Jr., Miller, L.M., and Young, H.A. (2015). Does the microbiota play a role in the pathogenesis of autoimmune diseases? *Gut 64*, 332-341.

McNulty, N.P., Yatsunenko, T., Hsiao, A., Faith, J.J., Muegge, B.D., Goodman, A.L., Henrissat, B., Oozeer, R., Cools-Portier, S., Gobert, G., *et al.* (2011). The impact of a consortium of fermented milk strains on the gut microbiome of

gnotobiotic mice and monozygotic twins. *Science translational medicine 3*, 106ra106.

Mellitzer, G., Beucher, A., Lobstein, V., Michel, P., Robine, S., Kedinger, M., and Gradwohl, G. (2010). Loss of enteroendocrine cells in mice alters lipid absorption and glucose homeostasis and impairs postnatal survival. *The Journal of clinical investigation 120*, 1708-1721.

Mennigen, R., and Bruewer, M. (2009). Effect of probiotics on intestinal barrier function. *Annals of the New York Academy of Sciences 1165*, 183-189.

Miller, R.S., and Hoskins, L.C. (1981). Mucin degradation in human colon ecosystems. Fecal population densities of mucin-degrading bacteria estimated by a "most probable number" method. *Gastroenterology 81*, 759-765.

Minekus, M., Marteau, P., Havenaar, R., and Huis in't Veld, J.H. (1995). A Multicompartmental Dynamic Computer controlled Model Simulating the Stomach and Small Intestine. *ATLA 23*, 197–209.

Minekus, M., Smeets-Peeters, M., Bernalier, A., Marol-Bonnin, S., Havenaar, R., Marteau, P., Alric, M., Fonty, G., and Huis in't Veld, J.H. (1999). A computer-controlled system to simulate conditions of the large intestine with peristaltic mixing, water absorption and absorption of fermentation products. *Applied microbiology and biotechnology 53*, 108-114.

Miquel, S., Leclerc, M., Martin, R., Chain, F., Lenoir, M., Raguideau, S., Hudault, S., Bridonneau, C., Northen, T., Bowen, B., *et al.* (2015). Identification of Metabolic Signatures Linked to Anti-Inflammatory Effects of Faecalibacterium prausnitzii. *mBio 6*.

Molloy, M.P., Herbert, B.R., Slade, M.B., Rabilloud, T., Nouwens, A.S., Williams, K.L., and Gooley, A.A. (2000). Proteomic analysis of the Escherichia coli outer membrane. *European journal of biochemistry / FEBS 267*, 2871-2881.

Molly, K., Vande Woestyne, M., and Verstraete, W. (1993). Development of a 5-step multi-chamber reactor as a simulation of the human intestinal microbial ecosystem. *Applied microbiology and biotechnology 39*, 254-258.

Monira, S., Nakamura, S., Gotoh, K., Izutsu, K., Watanabe, H., Alam, N.H., Endtz, H.P., Cravioto, A., Ali, S.I., Nakaya, T., *et al.* (2011). Gut microbiota of healthy and malnourished children in bangladesh. *Frontiers in microbiology 2*, 228.

Monsen, A.L., Refsum, H., Markestad, T., and Ueland, P.M. (2003). Cobalamin status and its biochemical markers methylmalonic acid and homocysteine in different age groups from 4 days to 19 years. *Clinical chemistry 49*, 2067-2075.

Moumene, A., Marcelino, I., Ventosa, M., Gros, O., Lefrancois, T., Vachiery, N., Meyer, D.F., and Coelho, A.V. (2015). Proteomic profiling of the outer membrane fraction of the obligate intracellular bacterial pathogen Ehrlichia ruminantium. *PloS one 10*, e0116758.

Mowat, A.M. (2003). Anatomical basis of tolerance and immunity to intestinal antigens. *Nature reviews Immunology 3*, 331-341.

Muegge, B.D., Kuczynski, J., Knights, D., Clemente, J.C., Gonzalez, A., Fontana, L., Henrissat, B., Knight, R., and Gordon, J.I. (2011). Diet drives convergence in gut microbiome functions across mammalian phylogeny and within humans. *Science 332*, 970-974.

Mueller, N.T., Bakacs, E., Combellick, J., Grigoryan, Z., and Dominguez-Bello, M.G. (2015). The infant microbiome development: mom matters. *Trends in molecular medicine 21*, 109-117.

Murphy, E.F., Cotter, P.D., Hogan, A., O'Sullivan, O., Joyce, A., Fouhy, F., Clarke, S.F., Marques, T.M., O'Toole, P.W., Stanton, C.*, et al.* (2012). Divergent metabolic outcomes arising from targeted manipulation of the gut microbiota in diet-induced obesity. *Gut*.

Nadal, I., Santacruz, A., Marcos, A., Warnberg, J., Garagorri, M., Moreno, L.A., Martin-Matillas, M., Campoy, C., Marti, A., Moleres, A.*, et al.* (2009). Shifts in clostridia, bacteroides and immunoglobulin-coating fecal bacteria associated with weight loss in obese adolescents. *Int J Obes (Lond) 33*, 758-767.

Nanthakumar, N.N., Fusunyan, R.D., Sanderson, I., and Walker, W.A. (2000). Inflammation in the developing human intestine: A possible pathophysiologic contribution to necrotizing enterocolitis. *Proceedings of the National Academy of Sciences of the United States of America 97*, 6043-6048.

Neef, A., and Sanz, Y. (2013). Future for probiotic science in functional food and dietary supplement development. *Current opinion in clinical nutrition and metabolic care 16*, 679-687.

Nicholson, J.K., Holmes, E., Kinross, J., Burcelin, R., Gibson, G., Jia, W., and Pettersson, S. (2012). Host-gut microbiota metabolic interactions. *Science 336*, 1262-1267.

Niers, L., Martin, R., Rijkers, G., Sengers, F., Timmerman, H., van Uden, N., Smidt, H., Kimpen, J., and Hoekstra, M. (2009). The effects of selected probiotic strains on the development of eczema (the PandA study). *Allergy 64*, 1349-1358.

Nikaido, H. (2011). Structure and mechanism of RND-type multidrug efflux pumps. *Advances in enzymology and related areas of molecular biology 77*, 1-60.

Nivaskumar, M., and Francetic, O. (2014). Type II secretion system: a magic beanstalk or a protein escalator. *Biochim Biophys Acta 1843*, 1568-1577.

Nylund, L., Satokari, R., Salminen, S., and de Vos, W.M. (2014). Intestinal microbiota during early life - impact on health and disease. *Proc Nutr Soc 73*, 457-469.

O'Connell Motherway, M., Zomer, A., Leahy, S.C., Reunanen, J., Bottacini, F., Claesson, M.J., O'Brien, F., Flynn, K., Casey, P.G., Munoz, J.A.*, et al.* (2011). Functional genome analysis of Bifidobacterium breve UCC2003 reveals type IVb tight adherence (Tad) pili as an essential and conserved host-colonization factor. *Proceedings of the National Academy of Sciences of the United States of America 108*, 11217-11222.

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

O'Neill, L.A., and Bowie, A.G. (2007). The family of five: TIR-domain-containing adaptors in Toll-like receptor signalling. *Nature reviews Immunology 7*, 353-364.

Okazaki, K., Iino, T., Kuroda, Y., Taguchi, K., Takahashi, H., Ohwada, T., Tsurumaru, H., Okubo, T., Minamisawa, K., and Ikeda, S. (2014). An assessment of the diversity of culturable bacteria from main root of sugar beet. *Microbes and environments / JSME 29*, 220-223.

Otto, A., Becher, D., and Schmidt, F. (2014). Quantitative proteomics in the field of microbiology. *Proteomics 14*, 547-565.

Ouwehand, A.C., Salminen, S., and Isolauri, E. (2002). Probiotics: an overview of beneficial effects. *Antonie van Leeuwenhoek 82*, 279-289.

Ouwerkerk, J.P., de Vos, W.M., and Belzer, C. (2013). Glycobiome: bacteria and mucus at the epithelial interface. *Best practice & research Clinical gastroenterology 27*, 25-38.

Palframan, R.J., Gibson, G.R., and Rastall, R.A. (2003). Carbohydrate preferences of Bifidobacterium species isolated from the human gut. *Current issues in intestinal microbiology 4*, 71-75.

Palmer, C., Bik, E.M., DiGiulio, D.B., Relman, D.A., and Brown, P.O. (2007). Development of the human infant intestinal microbiota. *PLoS biology 5*, e177.

Papasotiriou, D.G., Markoutsa, S., Meyer, B., Papadioti, A., Karas, M., and Tsiotis, G. (2008). Comparison of the membrane subproteomes during growth of a new pseudomonas strain on lysogeny broth medium, glucose, and phenol. *Journal of proteome research 7*, 4278-4288.

Pelaseyed, T., Bergstrom, J.H., Gustafsson, J.K., Ermund, A., Birchenough, G.M., Schutte, A., van der Post, S., Svensson, F., Rodriguez-Pineiro, A.M., Nystrom, E.E., *et al.* (2014). The mucus and mucins of the goblet cells and enterocytes provide the first defense line of the gastrointestinal tract and interact with the immune system. *Immunological reviews 260*, 8-20.

Penders, J., Thijs, C., Vink, C., Stelma, F.F., Snijders, B., Kummeling, I., van den Brandt, P.A., and Stobberingh, E.E. (2006). Factors influencing the composition of the intestinal microbiota in early infancy. *Pediatrics 118*, 511-521.

Petersen, T.N., Brunak, S., von Heijne, G., and Nielsen, H. (2011). SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nature methods 8*, 785-786.

Petrof, E.O., Gloor, G.B., Vanner, S.J., Weese, S.J., Carter, D., Daigneault, M.C., Brown, E.M., Schroeter, K., and Allen-Vercoe, E. (2013). Stool substitute transplant therapy for the eradication of Clostridium difficile infection: 'RePOOPulating' the gut. *Microbiome 1*, 3.

Petrof, E.O., and Khoruts, A. (2014). From stool transplants to next-generation microbiota therapeutics. *Gastroenterology 146*, 1573-1582.

Pflughoeft, K.J., and Versalovic, J. (2012). Human microbiome in health and disease. *Annu Rev Pathol 7*, 99-122.

Pinta, E., Li, Z., Batzilla, J., Pajunen, M., Kasanen, T., Rabsztyn, K., Rakin, A., and Skurnik, M. (2012). Identification of three oligo-/polysaccharide-specific ligases in Yersinia enterocolitica. *Molecular microbiology 83*, 125-136.

Png, C.W., Linden, S.K., Gilshenan, K.S., Zoetendal, E.G., McSweeney, C.S., Sly, L.I., McGuckin, M.A., and Florin, T.H. (2010). Mucolytic bacteria with increased prevalence in IBD mucosa augment in vitro utilization of mucin by other bacteria. *The American journal of gastroenterology 105*, 2420-2428.

Puertollano, E., Kolida, S., and Yaqoob, P. (2014). Biological significance of short-chain fatty acid metabolism by the intestinal microbiome. *Current opinion in clinical nutrition and metabolic care 17*, 139-144.

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

Qin, J., Li, Y., Cai, Z., Li, S., Zhu, J., Zhang, F., Liang, S., Zhang, W., Guan, Y., Shen, D., *et al.* (2012). A metagenome-wide association study of gut microbiota in type 2 diabetes. *Nature 490*, 55-60.

Quinlan, A.R., and Hall, I.M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics 26*, 841-842.

Rajilic-Stojanovic, M. (2013). Function of the microbiota. *Best practice & research Clinical gastroenterology 27*, 5-16.

Rajilic-Stojanovic, M., and de Vos, W.M. (2014). The first 1000 cultured species of the human gastrointestinal microbiota. *FEMS microbiology reviews 38*, 996-1047.

Rajilic-Stojanovic, M., Heilig, H.G., Molenaar, D., Kajander, K., Surakka, A., Smidt, H., and de Vos, W.M. (2009). Development and application of the human intestinal tract chip, a phylogenetic microarray: analysis of universally conserved phylotypes in the abundant microbiota of young and elderly adults. *Environmental microbiology 11*, 1736-1751.

Rajilic-Stojanovic, M., Shanahan, F., Guarner, F., and de Vos, W.M. (2013). Phylogenetic analysis of dysbiosis in ulcerative colitis during remission. *Inflammatory bowel diseases 19*, 481-488.

Rausch, P., Rehman, A., Kunzel, S., Hasler, R., Ott, S.J., Schreiber, S., Rosenstiel, P., Franke, A., and Baines, J.F. (2011). Colonic mucosa-associated microbiota is influenced by an interaction of Crohn disease and FUT2 (Secretor) genotype. *Proceedings of the National Academy of Sciences of the United States of America 108*, 19030-19035.

Reichardt, N., Duncan, S.H., Young, P., Belenguer, A., McWilliam Leitch, C., Scott, K.P., Flint, H.J., and Louis, P. (2014). Phylogenetic distribution of three pathways for propionate production within the human gut microbiota. *The ISME journal 8*, 1323-1335.

Remely, M., Aumueller, E., Jahn, D., Hippe, B., Brath, H., and Haslberger, A.G. (2014). Microbiota and epigenetic regulation of inflammatory mediators in type 2 diabetes and obesity. *Beneficial microbes 5*, 33-43.

Remely, M., Hippe, B., Geretschlaeger, I., Stegmayer, S., Hoefinger, I., and Haslberger, A. (2015a). Increased gut microbiota diversity and abundance of Faecalibacterium prausnitzii and Akkermansia after fasting: a pilot study. *Wiener klinische Wochenschrift*.

Remely, M., Tesar, I., Hippe, B., Gnauer, S., Rust, P., and Haslberger, A.G. (2015b). Gut microbiota composition correlates with changes in body fat content due to weight loss. *Beneficial microbes*, 1-9.

Rescigno, M. (2013). Mucosal immunology and bacterial handling in the intestine. *Best practice & research Clinical gastroenterology 27*, 17-24.

Reunanen, J., Kainulainen, V., Huuskonen, L., Ottman, N., Belzer, C., Huhtinen, H., de Vos, W.M., and Satokari, R. (2015). Akkermansia muciniphila adheres to enterocytes and strengthens the integrity of epithelial cell layer. *Applied and environmental microbiology*.

Rhomberg, T.A., Karlberg, O., Mini, T., Zimny-Arndt, U., Wickenberg, U., Rottgen, M., Jungblut, P.R., Jeno, P., Andersson, S.G., and Dehio, C. (2004). Proteomic analysis of the sarcosine-insoluble outer membrane fraction of the bacterial pathogen Bartonella henselae. *Proteomics 4*, 3021-3033.

Richard, H., and Foster, J.W. (2004). Escherichia coli glutamate- and arginine-dependent acid resistance systems increase internal pH and reverse transmembrane potential. *Journal of bacteriology 186*, 6032-6041.

Ringel, Y., Maharshak, N., Ringel-Kulka, T., Wolber, E.A., Sartor, R.B., and Carroll, I.M. (2015). High throughput sequencing reveals distinct microbial populations within the mucosal and luminal niches in healthy individuals. *Gut Microbes*, 0.

Robbe, C., Capon, C., Coddeville, B., and Michalski, J.C. (2004). Structural diversity and specific distribution of O-glycans in normal human mucins along the intestinal tract. *The Biochemical journal 384*, 307-316.

Robinson, M.D., McCarthy, D.J., and Smyth, G.K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics 26*, 139-140.

Roediger, W.E., Moore, J., and Babidge, W. (1997). Colonic sulfide in pathogenesis and treatment of ulcerative colitis. *Digestive diseases and sciences 42*, 1571-1579.

Roeselers, G., Mittge, E.K., Stephens, W.Z., Parichy, D.M., Cavanaugh, C.M., Guillemin, K., and Rawls, J.F. (2011). Evidence for a core gut microbiota in the zebrafish. *The ISME journal 5*, 1595-1608.

Rolfe, M.D., Rice, C.J., Lucchini, S., Pin, C., Thompson, A., Cameron, A.D., Alston, M., Stringer, M.F., Betts, R.P., Baranyi, J., *et al.* (2012). Lag phase is a distinct growth phase that prepares bacteria for exponential growth and involves transient metal accumulation. *Journal of bacteriology 194*, 686-701.

Rooijers, K., Kolmeder, C., Juste, C., Dore, J., de Been, M., Boeren, S., Galan, P., Beauvallet, C., de Vos, W.M., and Schaap, P.J. (2011). An iterative workflow for mining the human intestinal metaproteome. *BMC Genomics 12*, 6.

Rosendale, D.I., Blatchford, P.A., Sims, I.M., Parkar, S.G., Carnachan, S.M., Hedderley, D., and Ansell, J. (2012). Characterizing kiwifruit carbohydrate utilization in vitro and its consequences for human faecal microbiota. *Journal of proteome research 11*, 5863-5875.

Ruas-Madiedo, P., Gueimonde, M., Fernandez-Garcia, M., de los Reyes-Gavilan, C.G., and Margolles, A. (2008). Mucin degradation by Bifidobacterium strains isolated from the human intestinal microbiota. *Applied and environmental microbiology 74*, 1936-1940.

Rupakula, A., Kruse, T., Boeren, S., Holliger, C., Smidt, H., and Maillard, J. (2013). The restricted metabolism of the obligate organohalide respiring bacterium Dehalobacter restrictus: lessons from tiered functional genomics. *Philosophical transactions of the Royal Society of London Series B, Biological sciences 368*, 20120325.

Russell, S.L., Gold, M.J., Hartmann, M., Willing, B.P., Thorson, L., Wlodarska, M., Gill, N., Blanchet, M.R., Mohn, W.W., McNagny, K.M.*, et al.* (2012). Early life antibiotic-driven changes in microbiota enhance susceptibility to allergic asthma. *EMBO reports 13*, 440-447.

Salminen, S., von Wright, A., Morelli, L., Marteau, P., Brassart, D., de Vos, W.M., Fonden, R., Saxelin, M., Collins, K., Mogensen, G.*, et al.* (1998). Demonstration of safety of probiotics -- a review. *International journal of food microbiology 44*, 93-106.

Salonen, A., and de Vos, W.M. (2014). Impact of diet on human intestinal microbiota and health. *Annual review of food science and technology 5*, 239-262.

Samuel, B.S., Shaito, A., Motoike, T., Rey, F.E., Backhed, F., Manchester, J.K., Hammer, R.E., Williams, S.C., Crowley, J., Yanagisawa, M.*, et al.* (2008). Effects of the gut microbiota on host adiposity are modulated by the short-chain fatty-acid binding G protein-coupled receptor, Gpr41. *Proceedings of the National Academy of Sciences of the United States of America 105*, 16767-16772.

Santacruz, A., Collado, M.C., Garcia-Valdes, L., Segura, M.T., Martin-Lagos, J.A., Anjos, T., Marti-Romero, M., Lopez, R.M., Florido, J., Campoy, C.*, et al.* (2010). Gut microbiota composition is associated with body weight, weight gain and biochemical parameters in pregnant women. *The British journal of nutrition 104*, 83-92.

Santacruz, A., Marcos, A., Warnberg, J., Marti, A., Martin-Matillas, M., Campoy, C., Moreno, L.A., Veiga, O., Redondo-Figuero, C., Garagorri, J.M.*, et al.* (2009). Interplay between weight loss and gut microbiota composition in overweight adolescents. *Obesity 17*, 1906-1915.

Santarella-Mellwig, R., Franke, J., Jaedicke, A., Gorjanacz, M., Bauer, U., Budd, A., Mattaj, I.W., and Devos, D.P. (2010). The compartmentalized bacteria of the planctomycetes-verrucomicrobia-chlamydiae superphylum have membrane coat-like proteins. *PLoS biology 8*, e1000281.

Sato, T., van Es, J.H., Snippert, H.J., Stange, D.E., Vries, R.G., van den Born, M., Barker, N., Shroyer, N.F., van de Wetering, M., and Clevers, H. (2011). Paneth cells constitute the niche for Lgr5 stem cells in intestinal crypts. *Nature 469*, 415-418.

Schellenberger, J., Que, R., Fleming, R.M., Thiele, I., Orth, J.D., Feist, A.M., Zielinski, D.C., Bordbar, A., Lewis, N.E., Rahmanian, S.*, et al.* (2011). Quantitative prediction of cellular metabolism with constraint-based models: the COBRA Toolbox v2.0. *Nature protocols 6*, 1290-1307.

Schilderink, R., Verseijden, C., and de Jonge, W.J. (2013). Dietary inhibitors of histone deacetylases in intestinal immunity and homeostasis. *Frontiers in immunology 4*, 226.

Schluter, J., and Foster, K.R. (2012). The evolution of mutualism in gut microbiota via host epithelial selection. *PLoS biology 10*, e1001424.

Schmieder, R., and Edwards, R. (2011). Quality control and preprocessing of metagenomic datasets. *Bioinformatics 27*, 863-864.

Schnupf, P., Gaboriau-Routhiau, V., Gros, M., Friedman, R., Moya-Nilges, M., Nigro, G., Cerf-Bensussan, N., and Sansonetti, P.J. (2015). Growth and host interaction of mouse segmented filamentous bacteria in vitro. *Nature 520*, 99-103.

Scholtens, P.A., Oozeer, R., Martin, R., Amor, K.B., and Knol, J. The early settlers: intestinal microbiology in early life. *Annu Rev Food Sci Technol 3*, 425-447.

Scholtens, P.A., Oozeer, R., Martin, R., Amor, K.B., and Knol, J. (2012). The early settlers: intestinal microbiology in early life. *Annual review of food science and technology 3*, 425-447.

Schrager, J. (1970). The chemical composition and function of gastrointestinal mucus. *Gut 11*, 450-456.

Schreiber, O., Petersson, J., Walden, T., Ahl, D., Sandler, S., Phillipson, M., and Holm, L. (2013). iNOS-dependent increase in colonic mucus thickness in DSS-colitic rats. *PloS one 8*, e71843.

Schwartz, S., Friedberg, I., Ivanov, I.V., Davidson, L.A., Goldsby, J.S., Dahl, D.B., Herman, D., Wang, M., Donovan, S.M., and Chapkin, R. (2012). A metagenomic study of diet-dependent interaction between gut microbiota and host in infants reveals differences in immune response. *Genome biology 13*, R32.

Schwiertz, A., Taras, D., Schafer, K., Beijer, S., Bos, N.A., Donus, C., and Hardt, P.D. (2010). Microbiota and SCFA in lean and overweight healthy subjects. *Obesity 18*, 190-195.

Sczesnak, A., Segata, N., Qin, X., Gevers, D., Petrosino, J.F., Huttenhower, C., Littman, D.R., and Ivanov, II (2011). The genome of th17 cell-inducing segmented filamentous bacteria reveals extensive auxotrophy and adaptations to the intestinal environment. *Cell host & microbe 10*, 260-272.

Sekirov, I., Russell, S.L., Antunes, L.C., and Finlay, B.B. (2010). Gut microbiota in health and disease. *Physiological reviews 90*, 859-904.

Sela, D.A., Chapman, J., Adeuya, A., Kim, J.H., Chen, F., Whitehead, T.R., Lapidus, A., Rokhsar, D.S., Lebrilla, C.B., German, J.B., *et al.* (2008). The genome sequence of Bifidobacterium longum subsp. infantis reveals adaptations for milk utilization within the infant microbiome. *Proceedings of the National Academy of Sciences of the United States of America 105*, 18964-18969.

Selengut, J.D., Haft, D.H., Davidsen, T., Ganapathy, A., Gwinn-Giglio, M., Nelson, W.C., Richter, A.R., and White, O. (2007). TIGRFAMs and Genome Properties: tools for the assignment of molecular function and biological process in prokaryotic genomes. *Nucleic acids research 35*, D260-264.

Sherry, A.E., Inglis, N.F., Stevenson, A., Fraser-Pitt, D., Everest, P., Smith, D.G., and Roberts, M. (2011). Characterisation of proteins extracted from the surface of Salmonella Typhimurium grown under SPI-2-inducing conditions by LC-ESI/MS/MS sequencing. *Proteomics 11*, 361-370.

Shin, N.R., Lee, J.C., Lee, H.Y., Kim, M.S., Whon, T.W., Lee, M.S., and Bae, J.W. (2013). An increase in the Akkermansia spp. population induced by metformin treatment improves glucose homeostasis in diet-induced obese mice. *Gut*.

Shin, N.R., Lee, J.C., Lee, H.Y., Kim, M.S., Whon, T.W., Lee, M.S., and Bae, J.W. (2014). An increase in the Akkermansia spp. population induced by metformin treatment improves glucose homeostasis in diet-induced obese mice. *Gut 63*, 727-735.

Shrivastava, R., Basu, B., Godbole, A., Mathew, M.K., Apte, S.K., and Phale, P.S. (2011). Repression of the glucose-inducible outer-membrane protein OprB during utilization of aromatic compounds and organic acids in Pseudomonas putida CSV86. *Microbiology 157*, 1531-1540.

Sjogren, Y.M., Jenmalm, M.C., Bottcher, M.F., Bjorksten, B., and Sverremark-Ekstrom, E. (2009). Altered early infant gut microbiota in children developing allergy up to 5 years of age. *Clinical and experimental allergy : journal of the British Society for Allergy and Clinical Immunology 39*, 518-526.

Smith, P.M., Howitt, M.R., Panikov, N., Michaud, M., Gallini, C.A., Bohlooly, Y.M., Glickman, J.N., and Garrett, W.S. (2013). The microbial metabolites, short-chain fatty acids, regulate colonic Treg cell homeostasis. *Science 341*, 569-573.

Snoeck, V., Goddeeris, B., and Cox, E. (2005). The role of enterocytes in the intestinal barrier function and antigen uptake. *Microbes and infection / Institut Pasteur 7*, 997-1004.

Sokol, H., Pigneur, B., Watterlot, L., Lakhdari, O., Bermudez-Humaran, L.G., Gratadoux, J.J., Blugeon, S., Bridonneau, C., Furet, J.P., Corthier, G., *et al.* (2008). Faecalibacterium prausnitzii is an anti-inflammatory commensal bacterium identified by gut microbiota analysis of Crohn disease patients. *Proceedings of the National Academy of Sciences of the United States of America 105*, 16731-16736.

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

Sonnenburg, J.L., Xu, J., Leip, D.D., Chen, C.H., Westover, B.P., Weatherford, J., Buhler, J.D., and Gordon, J.I. (2005). Glycan foraging in vivo by an intestine-adapted bacterial symbiont. *Science 307*, 1955-1959.

Sonoyama, K., Fujiwara, R., Takemura, N., Ogasawara, T., Watanabe, J., Ito, H., and Morita, T. (2009). Response of gut microbiota to fasting and hibernation in Syrian hamsters. *Applied and environmental microbiology 75*, 6451-6456.

Sonoyama, K., Ogasawara, T., Goto, H., Yoshida, T., Takemura, N., Fujiwara, R., Watanabe, J., Ito, H., Morita, T., Tokunaga, Y.*, et al.* (2010). Comparison of gut microbiota and allergic reactions in BALB/c mice fed different cultivars of rice. *The British journal of nutrition 103*, 218-226.

Speth, D.R., van Teeseling, M.C., and Jetten, M.S. (2012). Genomic analysis indicates the presence of an asymmetric bilayer outer membrane in planctomycetes and verrucomicrobia. *Frontiers in microbiology 3*, 304.

Stoll, M.L., Kumar, R., Morrow, C.D., Lefkowitz, E.J., Cui, X., Genin, A., Cron, R.Q., and Elson, C.O. (2014). Altered microbiota associated with abnormal humoral immune responses to commensal organisms in enthesitis-related arthritis. *Arthritis research & therapy 16*, 486.

Strugala, V., Dettmar, P.W., and Pearson, J.P. (2008). Thickness and continuity of the adherent colonic mucus barrier in active and quiescent ulcerative colitis and Crohn's disease. *International journal of clinical practice 62*, 762-769.

Strugnell, R.A., and Wijburg, O.L. (2010). The role of secretory antibodies in infection immunity. *Nature reviews Microbiology 8*, 656-667.

Surana, N.K., and Kasper, D.L. (2012). The yin yang of bacterial polysaccharides: lessons learned from B. fragilis PSA. *Immunological reviews 245*, 13-26.

Swidsinski, A., Dorffel, Y., Loening-Baucke, V., Theissig, F., Ruckert, J.C., Ismail, M., Rau, W.A., Gaschler, D., Weizenegger, M., Kuhn, S.*, et al.* (2011). Acute appendicitis is characterised by local invasion with Fusobacterium nucleatum/necrophorum. *Gut 60*, 34-40.

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

Tailford, L.E., Crost, E.H., Kavanaugh, D., and Juge, N. (2015a). Mucin glycan foraging in the human gut microbiome. *Frontiers in genetics 6*, 81.

Tailford, L.E., Owen, C.D., Walshaw, J., Crost, E.H., Hardy-Goddard, J., Le Gall, G., de Vos, W.M., Taylor, G.L., and Juge, N. (2015b). Discovery of intramolecular trans-sialidases in human gut microbiota suggests novel mechanisms of mucosal adaptation *Nature communications (in press).*

Tasse, L., Bercovici, J., Pizzut-Serin, S., Robe, P., Tap, J., Klopp, C., Cantarel, B.L., Coutinho, P.M., Henrissat, B., Leclerc, M.*, et al.* (2010). Functional metagenomics to mine the human gut microbiome for dietary fiber catabolic enzymes. *Genome research 20*, 1605-1612.

Tedelind, S., Westberg, F., Kjerrulf, M., and Vidal, A. (2007). Anti-inflammatory properties of the short-chain fatty acids acetate and propionate: a study with relevance to inflammatory bowel disease. *World journal of gastroenterology : WJG 13*, 2826-2832.

Teixeira, J.S., Seeras, A., Sanchez-Maldonado, A.F., Zhang, C., Su, M.S., and Ganzle, M.G. (2014). Glutamine, glutamate, and arginine-based acid resistance in Lactobacillus reuteri. *Food microbiology 42*, 172-180.

Thurl, S., Henker, J., Siegel, M., Tovar, K., and Sawatzki, G. (1997). Detection of four human milk groups with respect to Lewis blood group dependent oligosaccharides. *Glycoconjugate journal 14*, 795-799.

Tokuda, H. (2009). Biogenesis of outer membranes in Gram-negative bacteria. *Bioscience, biotechnology, and biochemistry 73*, 465-473.

Tseng, T.T., Tyler, B.M., and Setubal, J.C. (2009). Protein secretion systems in bacterial-host associations, and their description in the Gene Ontology. *BMC microbiology 9 Suppl 1*, S2.

Tsujikawa, Y., Nomoto, R., and Osawa, R. (2013). Difference in Degradation Patterns on Inulin-type Fructans among Strains of Lactobacillus delbrueckii and Lactobacillus paracasei. *Bioscience of microbiota, food and health 32*, 157-165.

Turnbaugh, P.J., and Gordon, J.I. (2009). The core gut microbiome, energy balance and obesity. *The Journal of physiology 587*, 4153-4158.

Turnbaugh, P.J., Hamady, M., Yatsunenko, T., Cantarel, B.L., Duncan, A., Ley, R.E., Sogin, M.L., Jones, W.J., Roe, B.A., Affourtit, J.P.*, et al.* (2009a). A core gut microbiome in obese and lean twins. *Nature 457*, 480-484.

Turnbaugh, P.J., Ley, R.E., Mahowald, M.A., Magrini, V., Mardis, E.R., and Gordon, J.I. (2006). An obesity-associated gut microbiome with increased capacity for energy harvest. *Nature 444*, 1027-1031.

Turnbaugh, P.J., Quince, C., Faith, J.J., McHardy, A.C., Yatsunenko, T., Niazi, F., Affourtit, J., Egholm, M., Henrissat, B., Knight, R.*, et al.* (2010). Organismal, genetic, and transcriptional variation in the deeply sequenced gut microbiomes of identical twins. *Proceedings of the National Academy of Sciences of the United States of America 107*, 7503-7508.

Turnbaugh, P.J., Ridaura, V.K., Faith, J.J., Rey, F.E., Knight, R., and Gordon, J.I. (2009b). The effect of diet on the human gut microbiome: a metagenomic analysis in humanized gnotobiotic mice. *Science translational medicine 1*, 6ra14.

Turroni, F., Peano, C., Pass, D.A., Foroni, E., Severgnini, M., Claesson, M.J., Kerr, C., Hourihane, J., Murray, D., Fuligni, F.*, et al.* (2012). Diversity of Bifidobacteria within the Infant Gut Microbiota. *PloS one 7*, e36957.

Urbaniak, C., Cummins, J., Brackstone, M., Macklaim, J.M., Gloor, G.B., Baban, C.K., Scott, L., O'Hanlon, D.M., Burton, J.P., Francis, K.P.*, et al.* (2014). Microbiota of human breast tissue. *Applied and environmental microbiology 80*, 3007-3014.

Ursell, L.K., Clemente, J.C., Rideout, J.R., Gevers, D., Caporaso, J.G., and Knight, R. (2012). The interpersonal and intrapersonal diversity of human-associated microbiota in key body sites. *The Journal of allergy and clinical immunology 129*, 1204-1208.

Vaishampayan, P.A., Kuehl, J.V., Froula, J.L., Morgan, J.L., Ochman, H., and Francino, M.P. (2010). Comparative metagenomics and population dynamics of the gut microbiota in mother and infant. *Genome biology and evolution 2*, 53-66.

Van den Abbeele, P., Roos, S., Eeckhaut, V., MacKenzie, D.A., Derde, M., Verstraete, W., Marzorati, M., Possemiers, S., Vanhoecke, B., Van Immerseel, F.*, et al.* (2012). Incorporating a mucosal environment in a dynamic gut model results in a more representative colonization by lactobacilli. *Microbial biotechnology 5*, 106-115.

van den Berg, B. (2012). Structural basis for outer membrane sugar uptake in pseudomonads. *The Journal of biological chemistry 287*, 41044-41052.

van den Bogert, B., Erkus, O., Boekhorst, J., de Goffau, M., Smid, E.J., Zoetendal, E.G., and Kleerebezem, M. (2013). Diversity of human small intestinal Streptococcus and Veillonella populations. *FEMS microbiology ecology 85*, 376-388.

van der Ark, K.C.H., Ouwerkerk, J.P., Robert Finestra, T., Davids, M., Belzer, C., and de Vos, W.M. (2015). How to handle oxygen: the case of the mucosal gut anaerobe Akkermansia muciniphila. *Poster at Scientific Spring Meeting KNVM & NVMM*.

Van der Meulen, R., Avonts, L., and De Vuyst, L. (2004). Short fractions of oligofructose are preferentially metabolized by Bifidobacterium animalis DN-173 010. *Applied and environmental microbiology 70*, 1923-1930.

van Gelder, A.H., Aydin, R., Alves, M.M., and Stams, A.J. (2012). 1,3-Propanediol production from glycerol by a newly isolated Trichococcus strain. *Microbial biotechnology 5*, 573-578.

van Nood, E., Vrieze, A., Nieuwdorp, M., Fuentes, S., Zoetendal, E.G., de Vos, W.M., Visser, C.E., Kuijper, E.J., Bartelsman, J.F., Tijssen, J.G.*, et al.* (2013). Duodenal infusion of donor feces for recurrent Clostridium difficile. *N Engl J Med 368*, 407-415.

van Passel, M.W., Kant, R., Zoetendal, E.G., Plugge, C.M., Derrien, M., Malfatti, S.A., Chain, P.S., Woyke, T., Palva, A., de Vos, W.M.*, et al.* (2011). The genome of Akkermansia muciniphila, a dedicated intestinal mucin degrader, and its use in exploring intestinal metagenomes. *PloS one 6*, e16876.

Verberkmoes, N.C., Russell, A.L., Shah, M., Godzik, A., Rosenquist, M., Halfvarson, J., Lefsrud, M.G., Apajalahti, J., Tysk, C., Hettich, R.L.*, et al.* (2009). Shotgun metaproteomics of the human distal gut microbiota. *The ISME journal 3*, 179-189.

Vigsnaes, L.K., Brynskov, J., Steenholdt, C., Wilcks, A., and Licht, T.R. (2012). Gram-negative bacteria account for main differences between faecal microbiota from patients with ulcerative colitis and healthy controls. *Beneficial microbes 3*, 287-297.

Vijay-Kumar, M., Aitken, J.D., Carvalho, F.A., Cullender, T.C., Mwangi, S., Srinivasan, S., Sitaraman, S.V., Knight, R., Ley, R.E., and Gewirtz, A.T. (2010). Metabolic syndrome and altered gut microbiota in mice lacking Toll-like receptor 5. *Science 328*, 228-231.

von Ossowski, I., Pietila, T.E., Rintahaka, J., Nummenmaa, E., Makinen, V.M., Reunanen, J., Satokari, R., de Vos, W.M., Palva, I., and Palva, A. (2013). Using recombinant Lactococci as an approach to dissect the immunomodulating capacity of surface piliation in probiotic Lactobacillus rhamnosus GG. *PloS one 8*, e64416.

Vrieze, A., Out, C., Fuentes, S., Jonker, L., Reuling, I., Kootte, R.S., van Nood, E., Holleman, F., Knaapen, M., Romijn, J.A., *et al.* (2014). Impact of oral vancomycin on gut microbiota, bile acid metabolism, and insulin sensitivity. *Journal of hepatology 60*, 824-831.

Vrieze, A., Van Nood, E., Holleman, F., Salojarvi, J., Kootte, R.S., Bartelsman, J.F., Dallinga-Thie, G.M., Ackermans, M.T., Serlie, M.J., Oozeer, R., *et al.* (2012). Transfer of intestinal microbiota from lean donors increases insulin sensitivity in individuals with metabolic syndrome. *Gastroenterology 143*, 913-916 e917.

Wacklin, P., Makivuokko, H., Alakulppi, N., Nikkila, J., Tenkanen, H., Rabina, J., Partanen, J., Aranko, K., and Matto, J. (2011). Secretor genotype (FUT2 gene) is strongly associated with the composition of Bifidobacteria in the human intestine. *PloS one 6*, e20113.

Wacklin, P., Tuimala, J., Nikkila, J., Sebastian, T., Makivuokko, H., Alakulppi, N., Laine, P., Rajilic-Stojanovic, M., Paulin, L., de Vos, W.M., *et al.* (2014). Faecal microbiota composition in adults is associated with the FUT2 gene determining the secretor status. *PloS one 9*, e94863.

Walker, A.W., Duncan, S.H., Louis, P., and Flint, H.J. (2014). Phylogeny, culturing, and metagenomics of the human gut microbiota. *Trends in microbiology 22*, 267-274.

Wang, L., Christophersen, C.T., Sorich, M.J., Gerber, J.P., Angley, M.T., and Conlon, M.A. (2011). Low relative abundances of the mucolytic bacterium Akkermansia muciniphila and Bifidobacterium spp. in feces of children with autism. *Applied and environmental microbiology 77*, 6718-6721.

Wang, M., Li, M., Wu, S., Lebrilla, C.B., Chapkin, R.S., Ivanov, I., and Donovan, S.M. (2015). Fecal Microbiota Composition of Breast-fed Infants is Correlated with Human Milk Oligosaccharides Consumed. *Journal of pediatric gastroenterology and nutrition*.

Wang, Z.Q., Lawson, R.J., Buddha, M.R., Wei, C.C., Crane, B.R., Munro, A.W., and Stuehr, D.J. (2007). Bacterial flavodoxins support nitric oxide production by Bacillus subtilis nitric-oxide synthase. *The Journal of biological chemistry 282*, 2196-2202.

Ward, N.L., Rainey, F.A., Hedlund, B.P., Staley, J.T., Ludwig, W., and Stackebrandt, E. (2000). Comparative phylogenetic analyses of members of the order Planctomycetales and the division Verrucomicrobia: 23S rRNA gene sequence

analysis supports the 16S rRNA gene sequence-derived phylogeny. *International journal of systematic and evolutionary microbiology 50 Pt 6*, 1965-1972.

Weir, T.L., Manter, D.K., Sheflin, A.M., Barnett, B.A., Heuberger, A.L., and Ryan, E.P. (2013). Stool microbiome and metabolome differences between colorectal cancer patients and healthy adults. *PloS one 8*, e70803.

Wells, J.M., Loonen, L.M., and Karczewski, J.M. (2010). The role of innate signaling in the homeostasis of tolerance and immunity in the intestine. *International journal of medical microbiology : IJMM 300*, 41-48.

Wen, T., Oussenko, I.A., Pellegrini, O., Bechhofer, D.H., and Condon, C. (2005). Ribonuclease PH plays a major role in the exonucleolytic maturation of CCA-containing tRNA precursors in Bacillus subtilis. *Nucleic acids research 33*, 3636-3643.

Werner, J., and Misra, R. (2005). YaeT (Omp85) affects the assembly of lipid-dependent and lipid-independent outer membrane proteins of Escherichia coli. *Molecular microbiology 57*, 1450-1459.

Willing, B.P., Dicksved, J., Halfvarson, J., Andersson, A.F., Lucio, M., Zheng, Z., Jarnerot, G., Tysk, C., Jansson, J.K., and Engstrand, L. (2010). A pyrosequencing study in twins shows that gastrointestinal microbial profiles vary with inflammatory bowel disease phenotypes. *Gastroenterology 139*, 1844-1854 e1841.

Wilson, M.M., Anderson, D.E., and Bernstein, H.D. (2015). Analysis of the outer membrane proteome and secretome of Bacteroides fragilis reveals a multiplicity of secretion mechanisms. *PloS one 10*, e0117732.

Wimley, W.C. (2003). The versatile beta-barrel membrane protein. *Current opinion in structural biology 13*, 404-411.

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

Wopereis, H., Oozeer, R., Knipping, K., Belzer, C., and Knol, J. (2014). The first thousand days - intestinal microbiology of early life: establishing a symbiosis. *Pediatric allergy and immunology : official publication of the European Society of Pediatric Allergy and Immunology 25*, 428-438.

Wright, D.P., Knight, C.G., Parkar, S.G., Christie, D.L., and Roberton, A.M. (2000). Cloning of a mucin-desulfating sulfatase gene from Prevotella strain RS2 and its expression using a Bacteroides recombinant system. *Journal of bacteriology 182*, 3002-3007.

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

Wu, S., Tao, N., German, J.B., Grimm, R., and Lebrilla, C.B. (2010). Development of an annotated library of neutral human milk oligosaccharides. *Journal of proteome research 9*, 4138-4151.

Xiong, Y., Miyamoto, N., Shibata, K., Valasek, M.A., Motoike, T., Kedzierski, R.M., and Yanagisawa, M. (2004). Short-chain fatty acids stimulate leptin production in adipocytes through the G protein-coupled receptor GPR41. *Proceedings of the National Academy of Sciences of the United States of America 101*, 1045-1050.

Xu, X., Wang, Z., and Zhang, X. (2015). The human microbiota associated with overall health. *Critical reviews in biotechnology 35*, 129-140.

Yang, J.N., Wang, C., Guo, C., Peng, X.X., and Li, H. (2011). Outer membrane proteome and its regulation networks in response to glucose concentration changes in Escherichia coli. *Mol Biosyst 7*, 3087-3093.

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

Yin, Y., Wang, Y., Zhu, L., Liu, W., Liao, N., Jiang, M., Zhu, B., Yu, H.D., Xiang, C., and Wang, X. (2013). Comparative analysis of the distribution of segmented filamentous bacteria in humans, mice and chickens. *The ISME journal 7*, 615-621.

Yu, C.S., Lin, C.J., and Hwang, J.K. (2004). Predicting subcellular localization of proteins for Gram-negative bacteria by support vector machines based on n-peptide compositions. *Protein science : a publication of the Protein Society 13*, 1402-1406.

Yu, N.Y., Wagner, J.R., Laird, M.R., Melli, G., Rey, S., Lo, R., Dao, P., Sahinalp, S.C., Ester, M., Foster, L.J., *et al.* (2010). PSORTb 3.0: improved protein subcellular localization prediction with refined localization subcategories and predictive capabilities for all prokaryotes. *Bioinformatics 26*, 1608-1615.

Zackular, J.P., Baxter, N.T., Iverson, K.D., Sadler, W.D., Petrosino, J.F., Chen, G.Y., and Schloss, P.D. (2013). The gut microbiome modulates colon tumorigenesis. *mBio 4*, e00692-00613.

Zaura, E., Nicu, E.A., Krom, B.P., and Keijser, B.J. (2014). Acquiring and maintaining a normal oral microbiome: current perspective. *Frontiers in cellular and infection microbiology 4*, 85.

Zeng, B., Han, S., Wang, P., Wen, B., Jian, W., Guo, W., Yu, Z., Du, D., Fu, X., Kong, F., *et al.* (2015). The bacterial communities associated with fecal types and body weight of rex rabbits. *Scientific reports 5*, 9342.

Zhang, H., DiBaise, J.K., Zuccolo, A., Kudrna, D., Braidotti, M., Yu, Y., Parameswaran, P., Crowell, M.D., Wing, R., Rittmann, B.E., *et al.* (2009). Human gut microbiota in obesity and after gastric bypass. *Proceedings of the National Academy of Sciences of the United States of America 106*, 2365-2370.

Zhang, H., Sparks, J.B., Karyala, S.V., Settlage, R., and Luo, X.M. (2015a). Host adaptive immunity alters gut microbiota. *The ISME journal 9*, 770-781.

Zhang, J., Guo, Z., Xue, Z., Sun, Z., Zhang, M., Wang, L., Wang, G., Wang, F., Xu, J., Cao, H., *et al.* (2015b). A phylo-functional core of gut microbiota in healthy young Chinese cohorts across lifestyles, geography and ethnicities. *The ISME journal*.

Zhang, L., and Skurnik, M. (1994). Isolation of an R- M+ mutant of Yersinia enterocolitica serotype O:8 and its application in construction of rough mutants utilizing mini-Tn5 derivatives and lipopolysaccharide-specific phage. *Journal of bacteriology 176*, 1756-1760.

Zhang, X., Shen, D., Fang, Z., Jie, Z., Qiu, X., Zhang, C., Chen, Y., and Ji, L. (2013). Human gut microbiota changes reveal the progression of glucose intolerance. *PloS one 8*, e71108.

Zivkovic, A.M., German, J.B., Lebrilla, C.B., and Mills, D.A. (2011). Human milk glycobiome and its impact on the infant gastrointestinal microbiota. *Proceedings of the National Academy of Sciences of the United States of America 108 Suppl 1*, 4653-4658.

Zoetendal, E.G., Booijink, C.C., Klaassens, E.S., Heilig, H.G., Kleerebezem, M., Smidt, H., and de Vos, W.M. (2006). Isolation of RNA from bacterial samples of the human gastrointestinal tract. *Nature protocols 1*, 954-959.

Zoetendal, E.G., Raes, J., van den Bogert, B., Arumugam, M., Booijink, C.C., Troost, F.J., Bork, P., Wels, M., de Vos, W.M., and Kleerebezem, M. (2012a). The human small intestinal microbiota is driven by rapid uptake and conversion of simple carbohydrates. *The ISME journal 6*, 1415-1426.

Zoetendal, E.G., Raes, J., van den Bogert, B., Arumugam, M., Booijink, C.C., Troost, F.J., Bork, P., Wels, M., de Vos, W.M., and Kleerebezem, M. (2012b). The human small intestinal microbiota is driven by rapid uptake and conversion of simple carbohydrates. *ISME J*.

Zoetendal, E.G., Rajilic-Stojanovic, M., and de Vos, W.M. (2008). High-throughput diversity and functionality analysis of the gastrointestinal tract microbiota. *Gut 57*, 1605-1615.

Zoetendal, E.G., von Wright, A., Vilpponen-Salmela, T., Ben-Amor, K., Akkermans, A.D., and 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*, 3401-3407.

Appendices

Abbreviations

2'-FL	2′-fucosyllactose
3-FL	3-fucosyllactose
3'-SL	3′-sialyllactose
6'-SL	6′-sialyllactose
ANOVA	One-way analysis of variance
APC	Antigen presenting cell
BCFA	Branched-chain fatty acids
BLAST	Basic local alignment search tool
BTTM	Basal Tryptone Threonine Medium
CAZymes	Carbohydrate-active enzymes
CD	Crohn's disease
CFU	Colony forming unit
COG	Clusters of Orthologous Groups
CRC	Colorectal cancer
DGE	Differential gene expression
DSS	Dextran sulfate sodium
EFSA	European Food Safety Authority
EPS	Exopolysaccharide
EV	Extracellular vesicle
FBS	Fetal Bovine Serum
FDR	False discovery rate
FMP	Fermented milk product
FMT	Fecal microbiota transplantation
GABA	Glutamate/γ-aminobutyric acid
GalNAc	N-acetylgalactosamine
GALT	Gut-associated lymphoid tissue
GI	Gastrointestinal
GlcNAc	N-acetylglucosamine
GO	Gene ontology
HDAC	Histone deacetylase
HMO	Human milk oligosaccharide
HPLC	High-performance liquid chromatography
IBD	Inflammatory bowel disease
IgA	Immunoglobulin A
IMDM	Iscove's Modified Dulbecco's Medium
KEGG	Kyoto encyclopaedia of genes and genomes
LC-MS/MS	Liquid chromatography-tandem mass spectrometry
LFQ	Label-free quantification
LPS	Lipopolysaccharide
LTA	Lipoteichoic acid

MAMP	Microbe-associated molecular pattern
MAP	Mitogen-activated protein
N.D.	Not detected
NLR	Nucleotide binding and oligomerization domain-like receptors
OD	Optical density
OM	Outer membrane
PBMC	Peripheral blood mononuclear cell
PEP	Pro-Glu-Pro
PRR	Pattern recognition receptor
PTS	Phosphotransferase
PUL	Polysaccharide utilization loci
PVC	*Planctomycetes-Verrucomicrobia-Chlamydiae*
QPS	Qualified presumed safety
RLR	Retinoic acid-inducible gene I (RIG-I)-like receptors
RND	Resistance-Nodulation-Division
SCFA	Short-chain fatty acids
SD	Standard deviation
SFB	Segmented filamentous bacteria
SHIME	Simulator of the Human Intestinal Microbial Ecosystem
SIgA	Secretory IgA
T2D	Type 2 diabetes
TER	Transepithelial electrical resistance
TLR	Toll-like receptors
TMM	Trimmed mean of M-values
TOF	Time of flight
UC	Ulcerative colitis

Co-author affiliations

Steven Aalvink[1]
Clara Belzer[1]
Sjef Boeren[2]
Mark Davids[3]
Laura Huuskonen[4]
Veera Kainulainen[4]
Judith Klievink[5]
Jan Knol[6]
Rocio Martin[6]
Vitor A. P. Martins dos Santos[3]
Marjolein Meijerink[7]
Annick Mercenier[7]
Airi Palva[4]
Taija E. Pietilä[4]
Justus Reunanen[4,8]
Reetta Satokari[4]
Peter J. Schaap[3]
Mikael Skurnik[5,9]
Hauke Smidt[1]
Willem M. de Vos[1,4,5]

[1]Laboratory of Microbiology, Wageningen University, Wageningen, 6703 HB, The Netherlands
[2]Laboratory of Biochemistry, Wageningen University, Wageningen, 6700 ET, The Netherlands
[3]Laboratory of Systems and Synthetic Biology, Wageningen University, Wageningen, 6703 HB, The Netherlands
[4]Department of Veterinary Biosciences, University of Helsinki, Helsinki, 00014, Finland
[5]Department of Bacteriology and Immunology, and Research Programs Unit, Immunobiology, University of Helsinki, Helsinki, 00014, Finland
[6]Nutricia Research, Uppsalalaan 12, Utrecht Science Park, 3584 CT Utrecht, The Netherlands
[7]Host-Microbe Interactomics, Animal Sciences, Wageningen University, Wageningen, 6700 AH, The Netherlands
[8]Microbiology and Biotechnology, Department of Food and Environmental Sciences, University of Helsinki, 00014, Finland
[9]Helsinki University Central Hospital Laboratory Diagnostics, Helsinki, 00029, Finland

Acknowledgements

'Lusikka kauniiseen käteen
Takaisin satulaan
Vannon kiven ja kannon kautta
Teen kaiken uudestaan'

'I'll swallow my pride
Get back on the horse
I cross my heart and hope to die
I'll do it all again'

At times during my PhD project it felt like every day was a new struggle to overcome the disappointment of a failing experiment or a manuscript full of corrections. However, after encountering these difficulties often enough I realized that the confidence gained from them made each hardship easier to surpass. Besides the difficulties, I of course also remember the good times: the overbearing joy of successful experiments, the thrill of presenting new results to supervisors and conference audiences, and the pleasure of sharing all this with great colleagues. Most of all I will miss the empty corridors of the Microbiology building at nights when my research was all that mattered.

The work presented in this thesis was made possible by the contribution of numerous people, and I would like to take the opportunity to thank all of you. First of all, Clara Belzer, thank you for guiding me throughout my PhD, from teaching me how to run a protein gel on my very first week to submitting an article with me on my last week. It was great to see how you built up and managed the Akkermansia-team and later took on several other projects. You taught me to have strong opinions, to spice up my writing, and that sometimes doing things the 'messy way' is ok.

Willem de Vos, thank you for all your time and effort to this project. Every meeting with you gave me endless new ideas and challenged me to push harder to achieve the best results. The initial shock of the fast-paced meetings with you soon changed into me always preparing as well as I could and, if possible, trying to talk as fast as you do. I really appreciate you giving me the opportunity to visit Finland frequently for collaborations, during which I made valuable connections for the future.

Hauke Smidt, I am thankful for your input to my project during the past years and especially for your thorough comments to my manuscripts. You always encouraged me and made me feel like things were on track even when I was not so sure of it.

Members of the thesis committee, thank you for your time and critical assessment of this thesis. I also want to acknowledge all the co-authors of my manuscripts. Mark Davids, I could always come to you with questions about the metabolism of *Akkermansia* or bioinformatics in general, and I really appreciate that. This thesis

wouldn't have been the same without your input. Taija Pietilä, thank you for teaching me the secrets of cell culturing and for helping me understand immunology. Many thanks also to everyone else in Finland who contributed to my research: Reetta Satokari, Laura Huuskonen, Veera Kainulainen, Judith Klievink, Airi Palva, Justus Reunanen and Mikael Skurnik. Marjolein Meijerink, it was great working with you on the immunology experiments, thank you for all your help! Sjef Boeren, you were always happy to meet me when I had trouble with my proteomics experiments, thanks a lot for that! Thank you also to Rocio Martin, Jan Knol, Annick Mercenier, Peter Schaap and Vitor Martins dos Santos for your input.

I want to thank everyone who is or was a part of the Akkermansia-team during my PhD. By studying the same bacterium from different aspects, we all learned a lot from each other. Steven, you are excellent at your job and I really appreciate you always taking the time to answer my questions (even now that I'm not in Wageningen anymore). Kees, it was great to have you join the team and provide us with some new perspective, thanks for all the good discussions about *Akkermansia* and science in general. Romy, Klaudyna, Floor, Gerben, Lennart, Tom, Leo: even though you didn't study the amazing *Akkermansia*, you helped me out a lot with different scientific issues, and I am very thankful for that.

I really enjoyed my time at Microbiology because of the great atmosphere created by the people from all over the world. I was happy to get to know people from MolEco, MicFys, BacGen and Systems biology and spend time with you both in the lab and after work. I had a great time participating in the Veluwe loops, lab trips, WE-days, barbecues and so on. Thank you everyone for your help in the lab and for making the cold Finn feel like a part of the group!

Coline and Sebastian, it was great having you as my office mates during the first years. Thanks for all the advice and stupid jokes! Floor, Teunke, Corina, Bas, Javi, Vicente, it was a lot of fun to organize the PhD-trip with you. Thanks for sharing the experience with me!

Special thanks to Carolien and Anja for always answering my questions and helping me out with practical issues. Wim, thank you for patiently dealing with my computer problems. Ineke, Hans and Philippe, thank you for helping me out with orders and any other issues I would come across while working in the lab.

Alex, Pim and Annelieke, I am very grateful to have had you as Master's students. I learned a great deal by teaching you and all of you made valuable contributions to my PhD.

In the beginning of my PhD I had the pleasure of joining the CrossTalk-meetings all over Europe. I am grateful for this opportunity and happy to have met all the people involved.

During the years in Wageningen I also valued the time spent with people from outside the lab. Jess and Jen, thank you for being awesome friends! Borek, thank you for encouraging me to apply for the PhD in the Netherlands and for your support in the beginning of this journey. Thank you Selle for keeping me and my shoulders in shape, especially during the writing process. Sven, thank you for always being there in the parties to capture the memories and for taking photos during my defense. Corina, Armando, Ale, Kiira, Youri, Pelin, Johannes and Emilio: thank you for taking my mind off work and sharing some great moments with me!

Thank you to my aunt Raija, who kindly helped me with the layout of this thesis and designed the cover. I am very pleased with the outcome!

Janneke, even though I had my doubts about you when you first joined our team (you re-arranged everything in the lab!), you turned out to be the best colleague I could ever wish for. We got to know the ups and downs of working with *Akkermansia* and everything that came along with it. I learned a lot from you and I appreciate your attitude towards science and life. It was a lot of fun sharing it all with you and I wish the best for you in the future!

Angie, having you around was an essential part of why my period in Wageningen was the best time of my life. From the very beginning you inspired me to be a better scientist and to also enjoy life outside the lab. I am sure I wouldn't have survived through the four years without frequently sitting down with you over ~~one~~ several bottles of wine and discussing everything that was going on in our professional and personal lives (though sometimes it was difficult to separate the two). Thank you for always being there for me!

Finally, I would like to thank the ones nearest and dearest to me. Vicente, kiitos myöhäisistä illallisista, rauhoittavista sanoista ja siitä, että jaksoit kärsivällisesti kuunnella minua kun elämääni ei mahtunut muuta kuin väitöskirja. Enempää en olisi voinut toivoa. Äiti ja Iskä, kiitos tuesta ja tsemppauksesta. Kiitos myös siitä, että aina Suomeen tullessa saatoin hetkeksi heittäytyä teille täyshoitoon ja unohtaa huolet. Joonas, kiitos kiinnostavista keskusteluista ja erityisesti avusta propositioiden kanssa. Mirka, kiitos erinäisistä viisaista sanoista. Viimeisenä kiitos Mummolle, joka jaksaa aina uskoa minuun.

About the author

Noora Annika Ottman was born on the 3rd of December, 1985, in Espoo, Finland. She began her bachelor's degree in Health Biosciences at the University of Turku in 2004. After receiving her BSc, she continued with her master studies following the specialization Food Development. She performed her MSc thesis research at the Department of Microbiology in Danisco Finland Oy on the topic 'Synbiotic activity of *Bifidobacterium lactis* and polydextrose in a human colon simulator' under the supervision of Dr. Arthur Ouwehand and Dr. Kirsi Laitinen. After obtaining her MSc degree in 2009, Noora spent six months at the School of Nutrition and Health Sciences, Taipei Medical University, Taiwan. In January 2011, she moved to The Netherlands and joined the Molecular Ecology group at the Laboratory of Microbiology of Wageningen University. Her PhD research about the gut bacterium *Akkermansia muciniphila* was part of the Marie Curie Initial Training Network "Cross-Talk". The work was done under the supervision of Dr. Clara Belzer, Prof. Dr Hauke Smidt and Prof. Dr Willem M. de Vos, and the results of the PhD research are presented in this thesis. Noora currently works as a post-doctoral researcher at the Finnish Institute of Occupational Health on the topic of skin microbiota and allergy.

List of publications

Reunanen J, Kainulainen V, Huuskonen L, **Ottman N**, Belzer C, Huhtinen H, de Vos WM, Satokari R. (2015) *Akkermansia muciniphila* adheres to enterocytes and strengthens the integrity of epithelial cell layer. Appl Environ Microbiol. Mar 20.

Ottman N, Smidt H, de Vos WM, Belzer C. (2012) The function of our microbiota: who is out there and what do they do? Front Cell Infect Microbiol. Aug 9;2:104.

Mäkeläinen H, **Ottman N**, Forssten S, Saarinen M, Rautonen N, and Ouwehand AC. (2010) Synbiotic effects of galacto-oligosaccharide, polydextrose and *Bifidobacterium lactis* Bi-07 in vitro. International Journal of Probiotics and Prebiotics 5(4): 203-210

Laitinen K, **Ottman N**, Isolauri E. (2009) Turhat rajoitukset pois pikkulasten ruokavalioista. Suomen Lääkärilehti 64: 4442-4443. (Finnish Medical Journal, title in English: Unnecessary restrictions should be removed from the diets of small children.)

Overview of completed training activities

Discipline specific activities

Meetings & workshops

- 13th Gut Day Symposium
    - Wageningen, 2011
- Cross-Talk Workshop 'Lessons from host-pathogens interactions'
    - Milan, 2011
- Cross-Talk Network Meeting & Fall School
    - Wageningen, 2011
- 8th Joint INRA-Rowett Symposium on Gut Microbiology 'Gut Microbiota: friend or foe?'
    - Clermont-Ferrand, 2012
- Host-Microbes Cross-Talk: From animal models to human patients
    - Oslo, 2012
- Cross-Talk Final Symposium
    - Paris, 2012
- 6th Finnish Gut Day
    - Helsinki, 2013
- Scientific Spring Meeting KNVM & NVMM
    - Arnhem, 2013
- Danone International Scientific Symposium
    - Doorwerth, 2013
- 7th Finnish Gut Day
    - Helsinki, 2014
- Scientific Spring Meeting KNVM & NVMM
    - Arnhem, 2014
- 9th Joint Rowett-INRA Symposium on Gut Microbiology 'From sequence to function'
    - Aberdeen, 2014
- 8th Finnish Gut Day
    - Helsinki, 2015

Courses

- Advanced proteomics
    - Wageningen, 2011
- Functional metagenomics of the intestinal tract and food-related microbes
    - Helsinki, 2011

General courses

- VLAG PhD week
    - Baarlo, 2011
- PhD competence assessment
    - Wageningen, 2011
- Project and time management
    - Wageningen, 2011
- Voice matters - Voice and presentation skills Training
    - Wageningen, 2012
- Bionformatics – A user's approach
    - Wageningen, 2012
- Techniques for writing and presenting a scientific paper
    - Wageningen, 2012
- Career Orientation
    - Wageningen, 2014

Optionals

- Preparation of PhD research proposal
- PhD/Post-Doc meetings
- Molecular Ecology group meetings
- Microbiology PhD study trip
    - USA & Canada, 2013
- Organization of PhD study trip

COLOPHON

The research described in this thesis was financially supported by the European Community's Seventh Framework Programme (FP7/2007-2013) under grant agreement n°215553-2, and the Advanced Research Grant 250172 (MicrobesInside) of the European Research Council.

Cover design: Raija Sandqvist

Layout design: Raija Sandqvist & Noora Ottman

Printed by: Gildeprint, The Netherlands