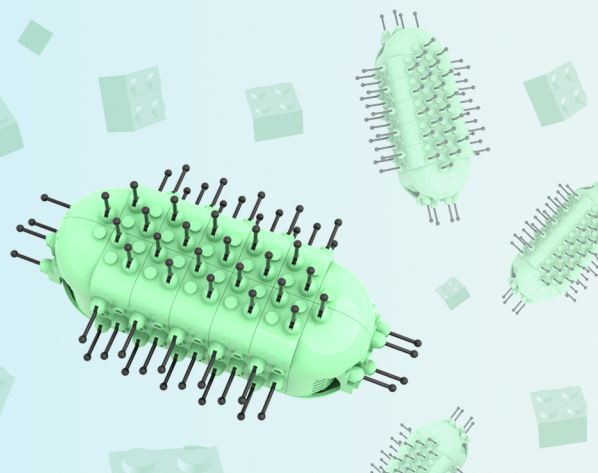




# MICROBES, MODELS & MUCINS

Modelling Host-Microbe  
Interactions at the  
Intestinal Mucus Layer  
*In Vitro*

**Janneke Elzinga**



# PROPOSITIONS

1. The use of animal-derived media components is a cultural problem.  
(this thesis)
2. Ironically, when studying the binding of an anti-obesity bacterium, fast food is a quick fix.  
(this thesis)
3. A focus on positive scientific results is blocking the road to Open Science.
4. The key to horizontal knowledge transfer is vertical integration of skills.
5. Using professional sports as a metaphor for academic life is like skating on thin ice.
6. Adding extra flavor to peanut butter requires very special grounds.

Propositions belonging to the thesis, entitled

Microbes, Models and Mucins: Modelling Host-Microbe Interactions at the Intestinal Mucus Layer *In Vitro*

Janneke Elzinga  
Wageningen, 31 May 2023







# **MICROBES, MODELS AND MUCINS:**

Modelling Host-Microbe Interactions at the Intestinal  
Mucus Layer *In Vitro*

Janneke Elzinga

## **Thesis committee**

### **Promotor**

Prof. Dr Hauke Smidt

Personal chair at the Laboratory of Microbiology  
Wageningen University & Research

### **Co-promotor**

Dr C. Belzer

Associate professor, Laboratory of Microbiology  
Wageningen University & Research

### **Other members**

Prof. Dr J.M. Wells, Wageningen University & Research

Dr H.J.M. Harmsen, University of Groningen

Dr E. van de Steeg, TNO, Leiden

Dr K. Strijbis, Utrecht University

**This research was conducted under the auspices of VLAG Graduate School  
(Biobased, Biomolecular, Chemical, Food, and Nutrition sciences)**





**MICROBES, MODELS AND MUCINS:**  
Modelling Host-Microbe Interactions at the Intestinal  
Mucus Layer *In Vitro*

Janneke Elzinga

**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 31 May 2023  
at 4 p.m. in the Omnia Auditorium

Janneke Elzinga

Microbes, Models and Mucins: Modelling Host-Microbe Interactions at the Intestinal  
Mucus Layer *In Vitro*

304 pages.

PhD thesis, Wageningen University, Wageningen, the Netherlands (2023)

With references, with summary in English and Dutch

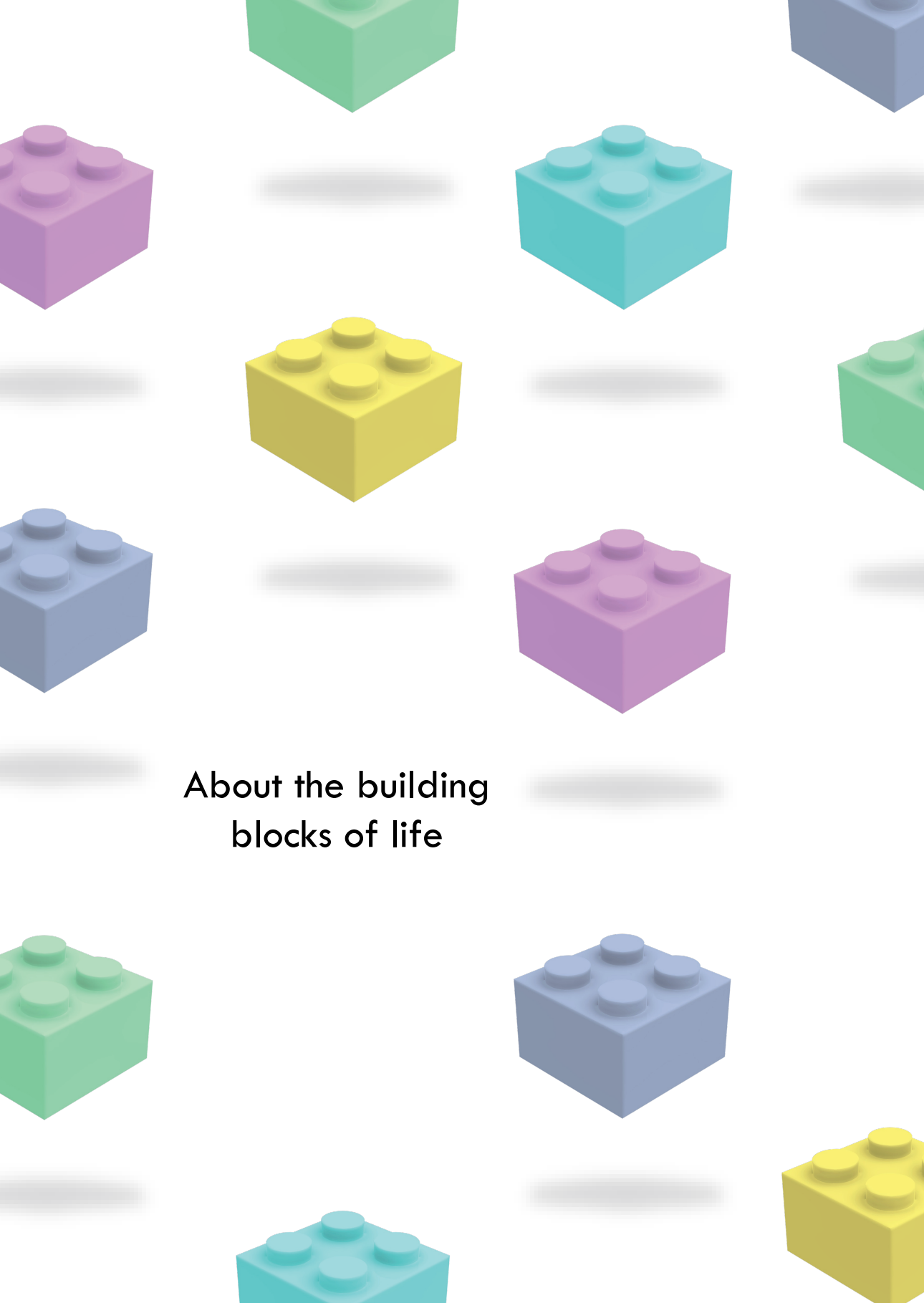
ISBN: 978-94-6447-664-4

DOI: 10.18174/628992

## Table of contents

<b>CHAPTER 1</b>	Introduction And Thesis Outline	<b>9</b>
<b>CHAPTER 2</b>	The Use of Defined Microbial Communities to Model Host-Microbe Interactions in the Human Gut	<b>33</b>
<b>CHAPTER 3</b>	Systematic Comparison of Transcriptomes of Caco-2 Cells Cultured Under Different Cellular and Physiological Conditions	<b>85</b>
<b>CHAPTER 4</b>	Characterization of Increased Mucus Production of HT29-MTX-E12 Cells Grown Under Semi-Wet Interface With Mechanical Stimulation	<b>115</b>
<b>CHAPTER 5</b>	The Effect of Mucus-Associated Bacteria on Intestinal Mucus Characteristics <i>In Vitro</i>	<b>155</b>
<b>CHAPTER 6</b>	Binding of <i>Akkermansia muciniphila</i> to Mucin Is O-Glycan Specific	<b>177</b>
<b>CHAPTER 7</b>	General Discussion	<b>207</b>
<b>APPENDICES</b>		<b>237</b>
	References	238
	English Summary	283
	Nederlandse Samenvatting	286
	Summary for Lego audience	290
	Samenvatting voor Legopubliek	292
	Abbreviations	294
	Acknowledgements	296
	About the Author	298
	List of Publications	299
	Overview of Completed Training Activities	301
	About the Cover	303
	Colophon	304





About the building  
blocks of life

# CHAPTER 1

## Introduction and Thesis Outline

## Introduction and Thesis Outline

Microbes are everywhere. All plants and animals, including humans, are host to a wide range of microbes, including bacteria, microeukaryotes and archaea, which are collectively called the microbiota. The whole theatre of activity and environment of these microbes, such as viruses, metabolites and other elements, is called the microbiome [1]. Given their wide genomic diversity, these microbes do not just form an additional organ<sup>1</sup> (i.e., a biological entity comprised of one organism's genome). Together with the eukaryotic host, they form a holobiont [2]. From an anthropocentric point of view, the human microbiota has long been considered an immunological threat but is now generally accepted as a crucial contributor to human health. This is reflected in the increasing interest in the role of the microbiota in human physiology and metabolism in health and disease [3]. Most of the research has focused on the body site where the majority of the microbial magic happens: the human gastrointestinal tract (GIT).

To get a grip on the numerous microbial inhabitants of the human GIT and their complex interplay with the human host, researchers would have to rely on invasive human studies or analysis of fecal samples. The latter, however, only provides a snapshot of what is exactly going on inside the GIT. As an alternative, most research has been conducted in other animals. Given the ethical and financial concerns regarding animal research and the low translatability from these animals to men, however, there is an urgent need for *in vitro* models representing host-microbe interactions across multiple levels: from the molecular interactions between host and microbe to the consequent system-wide effects. These models do not only advance our understanding of host-microbe interactions in the human GIT in health and disease but might also provide a way to cultivate and characterize the “unculturable”. Most importantly, these models provide platforms to evaluate novel therapeutic strategies for diseases associated with an aberrant microbiota, in a more reproducible and ethically responsible way.

This thesis describes the use of *in vivo* and *in vitro* models of the human GIT to study host-microbe interactions, with a specific focus on the colonic mucus layer. This chapter provides background information on the basic concepts within the field of host-microbe interactions in the human GIT. First, the human GIT and its intestinal inhabitants will be introduced. Next, commonly used *in vivo* and *in vitro* models to experimentally approach the host and/or microbes will be discussed. Finally, the aim and outline of this thesis will be presented.

---

<sup>1</sup> This is different from the semantics used in **Chapter 2** (published in 2018), demonstrating the rapid development of this field over the past five years.

## HUMAN INTESTINAL MICROBIOTA

### Where?

The human GIT encompasses the hollow organs from mouth to anus as well as several accessory glands and organs supplying secretions. The GIT is dedicated to the sequential processing of food. After mastication and lubrication in the mouth and oropharynx, digestion by mixing, protease activity and acid in the stomach, and adsorption of nutrients and fluid in the small intestine, the remaining chyme reaches the colon, where fluids and electrolytes are reabsorbed, before defecation via rectum and anus [4]. As the food proceeds from stomach to colon, it encounters an increasing microbial density and diversity [5]. These microbes assist in the digestion of food components which cannot be degraded by the host but are also involved in a wide range of other host processes.

### Who?

Due to, amongst other reasons, the anoxic conditions in a large part of the GIT, a considerable proportion of the gut microbiota is difficult to cultivate. Still, comprehensive studies of the human microbiota using a culture-independent, genome-based approach, have greatly expanded our knowledge on the microbial species inhabiting the human GIT, including Bacteria, Archaea and microeukaryotes [6-10]. Among Bacteria, twelve different phyla are currently distinguished [11] of which *Firmicutes*, *Proteobacteria*, *Actinobacteria*, *Bacteroidetes* and *Tenericutes* are predominant. [12] Note that these have recently been reclassified into *Bacillota*, *Pseudomonadota*, *Actinomycetota*, *Bacteroidota*, *Mycoplasmata* resp. [13]. Other phyla include *Chlamydiae*, *Deinococcus-Thermus*, *Fusobacteria*, *Spirochaetes*, *Synergistes*, *Lentisphaerae* and *Verrucomicrobia* [11]. Archaeal representatives in the GIT concern methane producers belonging to the taxon *Euryarchaeota* [14]. Lastly, microeukaryotes commonly found in the human intestinal tract include *Blastocystis* spp. and fungal genera [15]. The complete identification and characterization of all human intestinal microorganisms is still in progress, by improving cultivation strategies [16] and extension to non-Westernized populations [12, 17].





**Table 1:** Terminology used in this thesis.

holobiont	a unit of biological organization composed of a host and its microbiota [2]
human gastrointestinal tract	the tract of the digestive system from mouth to anus
microbiota	community of microorganisms
microbiome	a characteristic microbial community occupying a reasonably well-defined habitat which has distinct physio-chemical properties, not only referring to the microorganisms but also their theatre of activity (adapted from [1])
microorganism or microbe	microscopic-sized organism, including species from Bacteria, Archaea, Fungi, Protists and Algae
<i>in vivo</i>	(micro)organisms, cells, or biological molecules in their normal biological context, e.g., intestinal microbes in the GIT
<i>in vitro</i>	(micro)organisms, cells, or biological molecules outside their normal biological context, e.g., purified cultures of microbial strains and/or immortalized intestinal cell lines in a laboratory environment
<i>ex vivo</i>	(micro)organisms, cells or biological molecules taken directly outside their normal biological context, e.g., samples/biopsies intestinal microbes and/or tissue from human or other animal subjects. Note that fermentations using fecal inocula are often referred to as <i>in vitro</i> , but in some instances, could technically be considered <i>ex vivo</i> research
2D cell culture model	cells grown as monolayers on solid, impermeable surfaces or in uniform suspension (adapted from [18])
3D cell culture model	models ranging from monotypic cell cultures representing the minimum unit of the differentiated tissue <i>in vivo</i> to complex co-culture models that recapitulate the morphological (3D architecture) and functional features (multicellular complexity) of their <i>in vivo</i> parental tissues (adapted from [18])
Transwell / Thincert	brand name used for a (permeable) cell culture insert that is placed in a wells plate to generate a two-compartment cell culture environment

### How many?

As bacteria comprise the majority of the human intestinal microbiota, these microbes have been the most investigated and quantified. After several attempts to quantify bacterial cell numbers over the past five decades, the bacteria to human cell ratio is now estimated at 1:1.3-2.3, varying with age, gender, and body habitus. The vast majority resides in the lower small intestine (ileum) and the large intestine (colon), estimated at  $10^{11}$  and  $10^{14}$  bacterial cells per site, respectively [19, 20]. Colonization of the human GIT already starts at birth, after which bacterial composition co-matures with the growing child, dependent on a complex interplay of birth mode [21], diet [22-26], host genetics [27-33], sex and environmental factors [32, 34]. At the age of 1-2 years, the microbiota stabilizes into a composition resembling the adult microbiota [35], but remains subject to dynamic variations, amongst others due to dietary fluctuations [36], the circadian clock [37, 38], the use of antibiotics [39] and other drugs [40, 41]. An innumerable amount of studies have reported alterations in gut microbial composition in several intestinal and non-intestinal disorders, including inflammatory bowel disease (IBD) [42, 43], colorectal cancer (CRC) [44-47], obesity [48-50], Parkinson's disease [51, 52]<sup>2</sup>, although the cause-effect relationship needs to be determined [53]. In addition, there is spatial heterogeneity within the GIT, due to variations in chemical and physical parameters including pH, oxygen and nutrient resources [54]. The rest of this chapter and thesis will focus on the colonic bacteria because of their relatively high importance and density in the GIT.

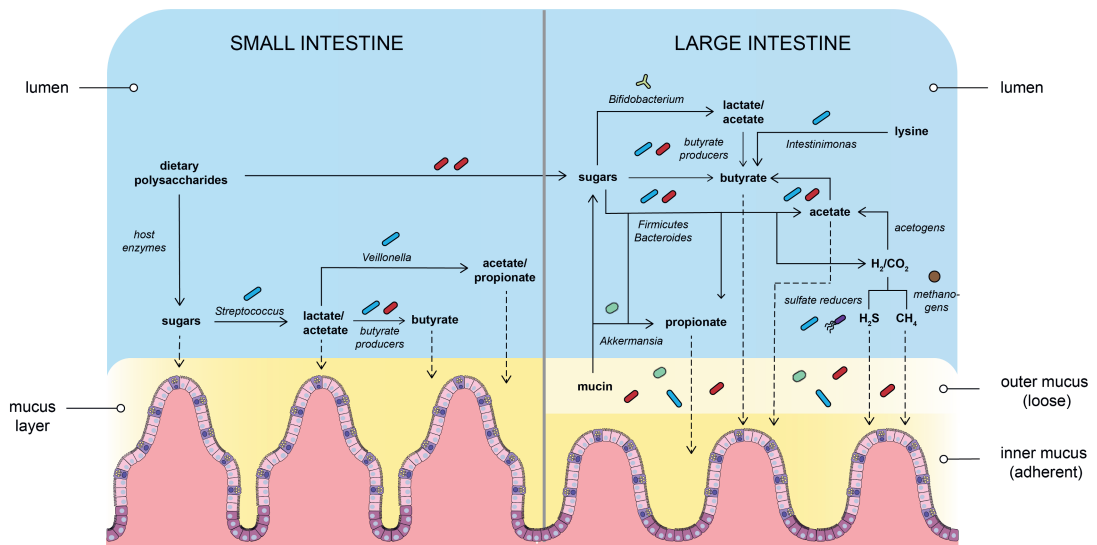
### What do they do?

The total genome content of the human intestinal microbiota has been estimated at > 22 million genes [55], encoding a wide range of biochemical and metabolic functions that complement host physiology [56]. Depending on the localization in the GIT, these functions include the (xeno)metabolism of compounds indigestible by host enzymes and the production of essential amino acids and vitamins [57], the development of the host's intestinal epithelium and immune system [58, 59], protection against invasion by opportunistic pathogens via colonization resistance [60], and maintenance of tissue homeostasis [61]. These effects are often not the action of a single microbe, but a complex interplay of whole microbial communities, including both microbe-microbe and bidirectional host-microbe interactions. Although there are numerous examples, I would like to highlight one of the most important functions of intestinal microbiota in the colon, which serves as a starting point for many host-microbe and microbe-microbe interactions: the production of SCFAs.

---

<sup>2</sup> Note that the studies mentioned are non-exhaustive





**Figure 1. The degradation of dietary polysaccharides and mucin by main bacterial groups in the small and large intestine.** SCFAs and other products are consumed by cross-feeding bacteria or the host intestinal epithelium. Note that the bacterial and intestinal cells are drawn at different scales. The different microbial shapes represent different phyla, but are non-exhaustive.

### *SCFA: A basis for many trophic chains*

Dietary fibers (polysaccharides, oligosaccharides and resistant starches) are barely degraded by host enzymes in the small intestine and continue to the colon to be broken down by specialized, microbial carbohydrate-active enzymes (CAZymes) [62, 63]. This results in the production of several compounds, of which SCFAs are the most studied. Acetate, propionate and butyrate comprise the majority of the total SCFA pool (> 95%) and have been quantified at a molar ratio of ~60:20:20 in the human colon [64]. SCFAs play a lot of different roles in host health and disease via the inhibition of histone deacetylases, the activation of G-coupled receptors or by serving as energy substrates for colonocytes or other microbes [65, 66]. In fact, the fermentation of dietary fiber provides the basis for many trophic chains in the human colon, including syntrophic (microbe-microbe) interactions [24, 67] (**Figure 1**). Next to dietary fiber, the host itself also produces a nutrient source for microbial fermentation in the form of mucus. The structure and function of the intestinal mucosal layer will be discussed in the next section.

## THE INTESTINAL MUCOSAL LAYER

### The intestinal epithelium

The whole GIT is lined by the mucosal layer, including a smooth muscle layer (*muscularis mucosae*), a loose connective tissue (*lamina propria*) and the intestinal epithelium, of which the latter is facing the intestinal lumen. The *lamina propria* contains a large number of immune cells, of both the innate and adaptive immune system. Additionally, both enteric and central nerves end into this layer. The overlying intestinal epithelial cells are joined together by tight junctions (incl. occludin, claudin, zonulin-1), forming an impermeable cell barrier, and differing in architectural structural and cellular composition along the GIT. In the colon, the epithelium is organized in crypts of Lieberkühn, invaginations into the underlying tissue [4]. The base of these crypts contain intestinal stem cells that give rise to progenitors that differentiate and mature while they travel up to the apex of the crypts [68, 69], where they are shed into the lumen. It is estimated that an entire crypt in the descending colon is replaced every 4-5 days [70, 71]. The main differentiated cell types found in the colon include colonocytes (responsible for nutrient and water absorption), (hormone-secreting) endocrine and goblet cells [4]. The latter produce mucus, a viscoelastic layer that covers the whole intestinal epithelium and serves as an important site for host-microbe interactions.

### The intestinal mucus layer

The intestinal mucus layer varies in composition and thickness along the GIT, being thickest in the colon [72-74], which is in line with the relatively high number of goblet cells (16% of epithelial cells) in this part of the GIT [74]. The mucus layer acts as a lubricant, facilitating the passage of food; provides a selective barrier, allowing the passage of low molecular weight components to the underlying epithelial cells; comprises a barrier protecting the underlying epithelium from both physical, chemical and (micro)biological damage; serves as a niche for bacteria that interact with mucin molecules [75]. The importance of the mucus layer has been demonstrated in the context of disease, such as colitis, CRC, and pathogen infection [76-78]. Although the mucus layer also contains lipids [79, 80] and defense-related proteins, including anti-microbial peptides [81], most of its functions can be explained by the molecular properties of its major structural component – mucin.

### Mucin structure

Mucins comprise a family of large proteins, consisting of a protein core containing a variable number of tandem repeats (VNTRs) of the amino acid residues proline, threonine and serine (also called PTS domain), which is unique for each mucin type [82, 83]. The threonine and serine residues are extensively O-GalNAc-glycosylated, resulting in a so-called “bottle brush”-like conformation [84]. Within the family of



mucins, there is general distinction between transmembrane and secreted mucins. Transmembrane mucins MUC1/3/4/12/13/15/17/20/21 are expressed on the cell membrane of enterocytes, with varying distribution and expression levels along the gastrointestinal tract (for an extensive overview, see e.g. [85]). Among the functions currently identified, transmembrane mucins contribute to the formation of the glycocalyx, serving an important function in cell-cell, cell-matrix and host-microbe interactions and signaling [86]. Although these mucins may be shed apically and integrated in the overlying mucus layer [87], the gel-forming, secreted mucins comprise the major component of the intestinal mucus layer. As opposed to transmembrane mucins, secreted mucins are characterized by the formation of disulfide bridges between the cysteine knot and von Willebrand C and D domains at the N- or C-termini of the mucin monomers [88, 89]. Examples include MUC5AC (stomach, duodenum and respiratory tissue) [90-94], MUC6 (stomach, gall bladder and duodenum) [92, 95, 96] and MUC2 (small intestine and colon) [92, 97, 98], of which MUC2 is the most abundant across the intestine [99] and most commonly studied so far.

### Mucin synthesis

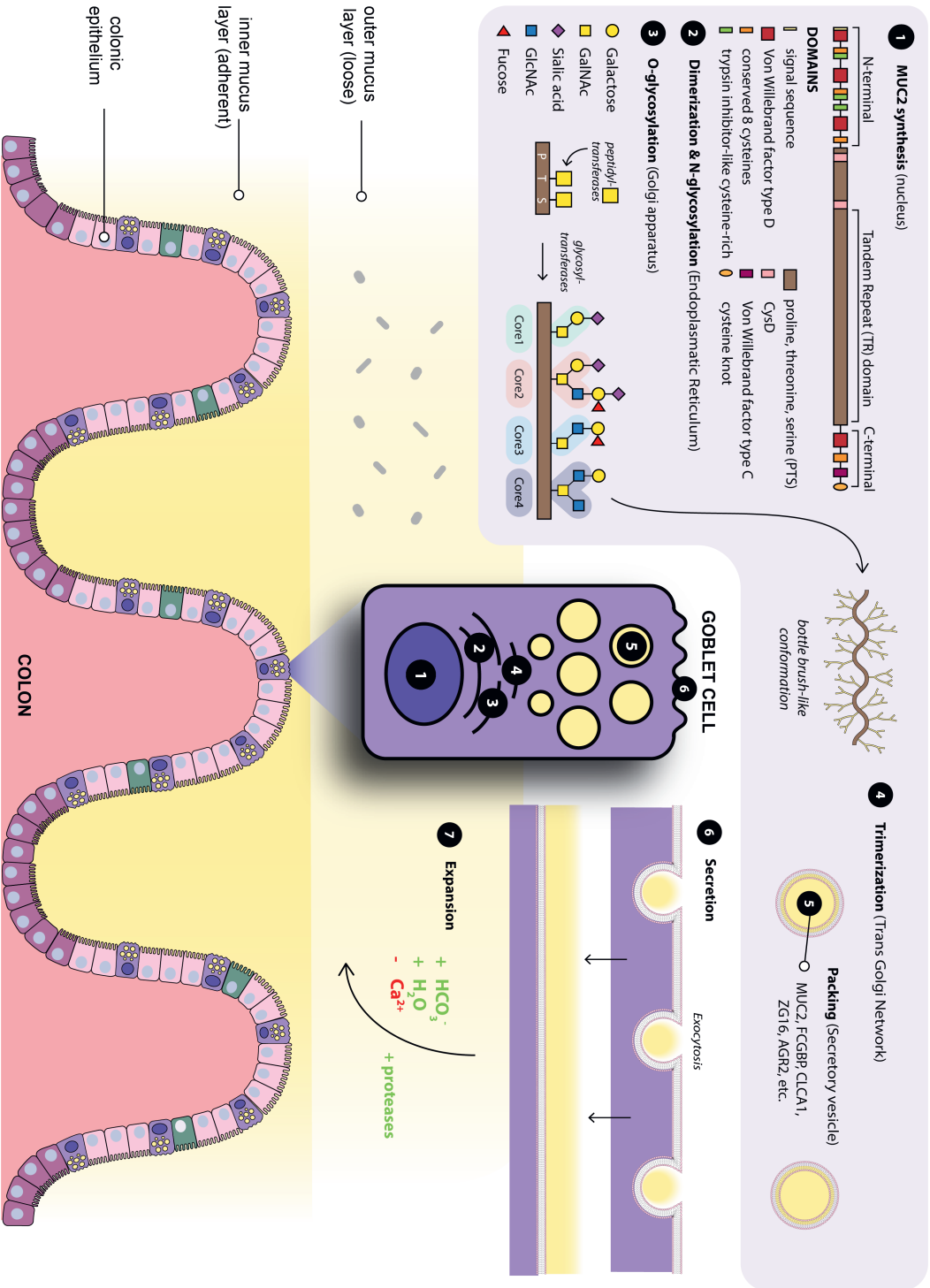
After expression in the nucleus of intestinal goblet cells (**Figure 2.1**), monomers of MUC2 dimerize through the formation of disulfide bridges in the endoplasmic reticulum [89, 100, 101]. Next, N-glycosylation of the dimers results in transfer to the Golgi apparatus [100] (**Figure 2.2**) where the mucins are O-glycosylated (see next section and **Figure 2.3**). After trimerization in the Trans Golgi Network [102] (**Figure 2.4**), the proteins are packed in secretory vesicles [103]. As the goblet cells migrate up from the crypt bottom, they fill their vesicles with MUC2 and other mucus components, such as the Fc fragment of IgG-binding protein, chloride channel accessory 1, zymogen granule protein 16 and anterior gradient homolog 2 [104, 105] (**Figure 2.5**). Secretion takes place through exocytosis (**Figure 2.6**), during which mucins are secreted in a pH and calcium-dependent manner [103]. The latter processes require the activity of CFTR (Cystic Fibrosis Transmembrane conductance Regulator) channels supplying bicarbonate ions [106]. Hydration of the mucin structures results in a volume expansion to a net-like, protective mucus barrier [107] (**Figure 2.7**). In the colon, the mucus layer is organized in two different layers: an inner and an outer layer. The former is continuously replenished with MUC2, firmly attached to the epithelium and is, at 200  $\mu\text{m}$  (in humans), converted to the outer mucus layer by endogenous proteases from both host (incl. CLCA1 [108]) and microbes colonizing the outer layer [84, 104]. This results in further expansion of the mucus layer up to 2-3 times [104] (**Figure 2.7**).

### Mucin O-glycosylation

Mucin glycans can make up 80% of the total mucin mass [82] and are the product of a broad collection of glycosyltransferases expressed by the host ([109] for a comprehensive overview). The synthesis of mucin glycans starts with the transfer of N-acetylgalactosamine (GalNAc) to a serine or threonine residue of the PTS-domain [110], resulting in the so-called Tn-antigen. Next, the glycans are extended by the addition of N-acetylglucosamine (GlcNAc), GalNAc or galactose, resulting in eight different types of core structures [111]. Core1-4 glycans (see **Figure 2.3**) are the most commonly found in human intestinal mucins, with core1 and 2 being predominant in the stomach and duodenum [112, 113], core3 in the small intestine [113], and core3 and 4 on colonic mucins [113-115]. These core structures are further extended by the addition of galactose, N-acetylglucosamine (GlcNAc) and GalNAc into large, branched chains with up to 20 monosaccharide units [111]. Finally, glycans are frequently terminated by fucose and sialic acid, of which occupation levels follow opposing gradients along the GIT: from stomach to colon, the number of sialylated mucin glycans increases, whereas the level of fucosylation decreases [113]. Sialic acids may be further modified by O-acetylation [116] and Gal and GlcNAc by sulfation [117]. The main structural variety in mucin glycans is provided by terminal sugar residues and can be partially attributed to blood group (e.g. the secretor gene [118]) and other genetic factors. Apart from inter-individual differences, more than 100 different O-glycan structures have been found on a single MUC2 protein [112].

**Figure 2 (next page). Mucin synthesis from the perspective of the goblet cell.** MUC2, the mucin type most dominantly expressed in the human colon, is shown as an example. MUC2 is synthesized in the nucleus (1), after which the protein backbone is N-glycosylated in the endoplasmic reticulum (2). Next, mucin is O-glycosylated in the Golgi Apparatus (3), resulting in a bottle brush-like formation. After trimerization in the Trans Golgi Network (4), the mucin proteins are packed in secretory vesicles (5). Via exocytosis, these vesicles are secreted (6) and expanded (7) in the colonic lumen. Parts of the figure are inspired by [119].





### Bacteria-mucin interactions

The division of the colonic layer into two parts facilitates a close relationship with intestinal bacteria: whereas the inner layer is impenetrable to bacteria [84, 104], protecting the underlying epithelium from invading pathogens, the outer mucus layer allows colonization by bacteria. Even though the division into an inner and outer colonic mucus layer is still present in germ-free (GF) animals [104], long-term microbial colonization is required for the inner layer to become impenetrable to bacteria and O-glycosylation patterns to mature [120]. The composition of this mucosa-associated microbial community was found to be significantly different from that of fecal or luminal microbiota [121-123]. The colonization by mucosa-associated microbes is further correlated with the radial oxygen gradient present from epithelium to lumen (5% to 0%, as measured in the murine colon) [124]. In addition, the broad repertoire of mucin glycans selects for the colonization by specific intestinal bacteria in the mucosal layer (e.g., [125-127]). *Vice versa*, (ex-)GF rodent models have demonstrated the influence of the microbiota on mucus thickness, compactness, content, and structure [128-130]. As a consequence, diseases with aberrant mucin O-glycosylation, including IBD (e.g. [131-134]) and CRC [135] are often associated with a deviant microbiota composition, although – also here – causal links still need to be established. The molecular bacteria-mucin interactions underlying these observations can be generally divided into three mechanisms: binding, degradation and indirect effects.

#### Mucin binding

Apart from getting trapped in the selective barrier of the mucin polymeric network [136], bacteria can also selectively bind to mucin molecules [127]. Mucin binding is generally thought to support the colonization by the binding microbe, and thus, provide the opportunity to exert effects (close) to the host cells. Moreover, binding of beneficial bacteria to mucin glycans inhibits binding of pathogenic species. Lastly, binding through carbohydrate binding modules (CBMs) encoded in glycosyl hydrolases (GHs), may facilitate mucin glycan degradation by mucolytic species (e.g., [137]). Both beneficial and pathogenic species have developed several strategies to bind intestinal mucus, including mucin-binding proteins (e.g. in *Lactobacilli*), flagella (e.g. in enteropathogenic and enterohemorrhagic *Escherichia coli*), pili (e.g. in *Lactocaseibacillus rhamnosus* GG), fimbriae (e.g. *E. coli* and *Salmonella enterica* Typhimurium) and moon-lighting proteins (expressing more than one function) (reviewed in [138]). The considerable proportion of mucin glycans in combination with their demonstrated effects on microbial colonization, suggest a crucial role for lectins, non-enzymatic carbohydrate binding proteins. Indeed, mucin-glycan binding lectins have been characterized, but mainly in pathogens, including *Helicobacter pylori*, *Campylobacter jejuni*, *Entamoeba histolytica* and *Vibrio cholerae* [138]. A recent





screening of human microbiota (meta)genome sequence data revealed a high prevalence of unique, uncharacterized sequences predicted to encode lectins, with likely various functions [139].

### *Mucin degradation*

Whereas mucin degradation was initially mostly studied in the context of pathogens, it is now considered part of the normal mucus turn-over. Similar to the broad repertoire of glycosyltransferases expressed by the host to generate mucin O-glycans, the mucosal intestinal microbiota expresses a diverse set of enzymes to break down these structures. Mucin glycan breakdown is often initiated by sulfatases and acetylases, after which GHs cleave the glycan structures, including  $\alpha$ -sialidases (GH33),  $\alpha$ -fucosidases (GH29, GH95), exo- and endo- $\beta$ -N-acetylglucosaminidases (GH84 and 85),  $\beta$ -galactosidases (GH2, GH20, GH42),  $\alpha$ -N-acetylglucosaminidases (GH89), endo-  $\beta$ 1,4-galactosidases (GH98) and  $\alpha$ -N-acetylglactosaminidases (GH101, GH129) (comprehensively described in the CAZy database [140]). Most GHs also contain a carbohydrate binding module, facilitating the adherence of the enzyme to the substrate, of which CBM families 32, 40, 47 and 51 have been reported to recognize mucin glycans (reviewed in [141]). Released mono- and oligosaccharides can be used by the degrading bacteria [142] or consumed by other beneficial as well as pathogenic bacteria present in the mucus layer [143-145]. After removal of the glycan structure, the mucin protein backbone can be cleaved by mucin-specific proteases (mucinases), including cysteine proteases, serine proteases and metalloproteases (reviewed in [146, 147]), or the relatively new class of O-glycoproteases requiring a combination of O-glycan and peptide sequence for recognition and cleavage (reviewed in [148]). Mucin degradation has been extensively characterized in *Bacteroides* species, which genomes contain specific polysaccharide-utilization loci (PULs), including genes encoding CAZymes that are able to degrade a wide range of polysaccharides [149]. Examples of such 'mucin generalists' include *Bacteroides thetaiotaomicron* and *Bacteroides fragilis* as well as *Bifidobacterium* spp. (e.g., *Bifidobacterium bifidum* and *Bifidobacterium longum* subsp. *longum*). In contrast, mucin specialists can only grow on mucin as the sole carbon source, of which examples include *Ruminococcus gnavus*, *Ruminococcus torques* and *Akkermansia muciniphila* (reviewed in [150]). The latter mucin degrader is introduced in more detail in **Box 1**. Note that strains from the same species can exhibit differences in mucin degradation capacity (e.g., [151]).

### *Indirect effects*

Next to direct interactions between microbes and mucin molecules, several bacterial effectors have been shown to modulate host mucus properties, such as pathogen-associated molecular patterns including lipopolysaccharide (LPS), peptidoglycan or flagellin, amongst others via activation of Toll-Like Receptors (TLRs) (reviewed in

[119]). Recently, a subset of “sentinel” goblet cells was identified in murine colonic crypts, which, upon contact with (microbial) TLR ligands, induced the secretion of Muc2 (murine variant of MUC2) by neighboring goblet cells [152]. In fact, next to producing mucus, goblet cells also play a critical role in antigen delivery to immune cells in the underlying lamina propria, therefore acting as critical immune surveillants, controlling immune homeostasis [153]. In addition, SCFAs, especially butyrate, can also enhance MUC2 gene expression (e.g. [154]). Lastly, some mucin-degrading bacteria are known to affect mucin expression and glycosylation via the secretion of peptide-like or other soluble molecules [155, 156].

#### BOX 1: THE CASE OF *AKKERMANSIA MUCINIPHILA*

*A. muciniphila* was originally isolated from the human intestine using mucin as a sole carbon and nitrogen source [157], but *Akkermansia*-like species have been identified in the GIT throughout a wide range of mammals and non-mammals [158]. A substantial proportion of the genome of this Gram-negative bacterium is solely dedicated to the degradation of mucin [159]. Moreover, *A. muciniphila* has the oxygen reduction capacity to survive at the oxic-anoxic interface of the mucus layer [160], further supporting its reputation as a professional mucin-degrader. In general, the abundance of *A. muciniphila* ranges between 1 and 4 % in the colon of healthy human beings [161] and is inversely correlated with several intestinal and metabolic disorders (comprehensively described in [162]). More importantly, the administration of both live and pasteurized *A. muciniphila* have been shown to influence metabolic parameters and reinforce the intestinal barrier of metabolic syndrome patients [163-165]. *A. muciniphila* is therefore considered a next-generation beneficial microbe. Beneficial effects of *A. muciniphila* have been attributed to, amongst others, the whole (pasteurized) bacterium, the pili-associated protein Amuc\_1100, SCFAs and protein 9 (reviewed in [166]). In addition, due to its role as mucin-degrader, liberating monosaccharides and SCFAs for other bacteria [167]), this bacterium is considered a key player in the intestinal ecosystem.



Cartoon representation of *A. muciniphila* used throughout this thesis. The hair-like structures on the outer membrane represent its pili.

## MODELLING HOST-MICROBE INTERACTIONS

Research on the intestinal microbiota in healthy or diseased human beings, has been mainly based on fecal samples, which, however, provide only snapshots of the intestinal microbiota composition and may not reflect community composition along

(transversally and longitudinally) the GIT [54, 123]. Alternatives such as biopsies have been limited, due to their invasiveness of sampling. Moreover, these approaches do not give full insight into the myriad of host-microbe interactions in the GIT, as they do not include a host component. To this end, researchers have relied mainly on *in vivo* animal models as surrogate for human beings or *in vitro* models representative of the human GIT, its intestinal microbiota or both.

### Modeling the human intestinal tract

#### *In vivo: From GF to humanized gnotobiotic animals*

The use of GF animals has taught scientists a lot about the importance of the microbiome, and dates back to the late 19<sup>th</sup> century, when several groups of researchers questioned the survival of animals in a bacteria-free environment [168]. In the 1940s, James Arthur Reyniers and Philip C. Trexler further developed GF technology, allowing larger-scale breeding of GF animals. Since then, GF rats and mice have been commonly used in biomedical sciences for different purposes [169]. More specifically regarding the GIT, it was recognized during the 1960s and '70s, that intestines of GF animals display aberrant histological, anatomical and physiological characteristics, compared to conventional laboratory animals [170]. The development of “the Schaedler cocktail” marked one of the first attempts to normalize GF mice, of which an altered version has been commonly used as a standardized gut microbiota by animal breeders and biomedical researchers ever since [171]. In more general, the controlled inoculation of GF animals (resulting in gnotobiotic mice), in combination with the potential to experimentally manipulate genetic background, diet and other factors, have allowed functional and mechanistic research on host-microbe interactions in the context of the whole organism.

Rodent models – specifically mice – are the most frequently used animal models in life sciences, as they are relatively small in size, reproduce fast, are easy and cheap to maintain in large numbers. Moreover, regarding the intestine, mice and men share many anatomical, histological and physiological features (reviewed in [172]). Clearly, however, there are also major differences (reviewed in [173, 174]), including, amongst others, intestinal pH [175], oxygen tension levels [176], mucin glycosylation [177], which have all been known to influence the intestinal microbiota. Consequently, 85% of the murine microbial genera and species are not found in humans, although humans and mice share a large qualitatively similar core microbiota [178, 179]. In this context, humanized gnotobiotic mice (GF mice inoculated with a human gut microbiota through fecal microbiota transplantation (FMT) [180, 181]) has definitely improved research into host-microbe interactions, although full translatability from mice to men has not been achieved yet [182].

Next to mice, rats have also been commonly used in the context of intestinal microbiota research. In fact, the intestinal microbiota of humanized rats more closely reflected the gut microbiota of human donors than mice [183]. Another increasingly

popular animal model to study intestinal microbiota is the pig, of which anatomy, physiology and immunology show high similarities to the human GIT (reviewed in [184]). Similarly, due to high similarity in gut morphology and digestive system functioning between humans and dogs, the canine microbiota has also been investigated (reviewed in [185]), which clustered more closely to the human than the murine fecal microbiota [186]. Finally, the intestinal microbiota of closely related, non-human primates showed high similarities with the human microbiota [187], but research on primates and dogs comes with more stringent ethical restrictions for experimentation and breeding difficulties, as opposed to mice.

### *The 3Rs: A call for non-animal technologies*

In general, the use of animals for research purposes has long been a matter of controversy. The use of animals in experiments is restricted by (inter)national laws (e.g. Artikel 9 in the Netherlands [188] and Directive 2010/63/EU in European Union [189]). These laws also explicitly mention the 3Rs to improve the welfare of animals used in research, proposed in the 1950s by Russell and Burch: 1) Replacement, *i.e.*, avoiding or replacing the use of animals where they otherwise would have been used, 2) Reduction, *i.e.* minimizing the number of animals used consistent with scientific aims and 3) Refinement, *i.e.* minimizing the pain, suffering and distress or lasting harm that research animals might experience [190]. Additionally, animal models are relatively expensive and time-consuming and, as discussed briefly above, do often not fully represent human biology. This is especially alarming in the development of drugs, in which animals serve as preclinical models, but often lead to mistranslation of safety, toxicity and efficacy of promising candidates to human trials and eventually, the patient [191]. Consequently, there has been an increasing interest in alternative models (non-animal technologies (NATs)), including *in vitro* and/or *ex vivo* models using established cell lines or human-derived tissues and cells, which will be discussed below in the context of the human colon.

### *In vitro: From static inserts to dynamic gut-on-a-chips*

Human colon carcinoma cell lines have been commonly used to represent the intestinal epithelial layer *in vitro*, of which Caco-2 cells are the best characterized. This cell line was originally isolated in the 1970s from a colorectal tumor [192]. As opposed to other isolated colon carcinoma cell lines [193], Caco-2 cells demonstrate spontaneous differentiation upon long-term culture leading to expression of several morphological and biochemical characteristics of small intestinal enterocytes [194, 195]. Other human intestinal cell lines have also been commonly used as a model of the human intestinal epithelium, including T84 [196], LS174T [197], DLD-1 [198] and HT29(-MTX) cells [199], each with their own advantages. In addition to traditional well plates, these cell lines have been commonly cultured in inserts (also known under its brand name Transwells or Thincerts), including an apical (“luminal”) and basolateral



(“abluminal”) compartment, allowing studies on intestinal barrier function and transport. Complexity of this model has been further increased by the combination of different intestinal cell types, e.g. Caco-2 and HT29-MTX (e.g. [200]), the introduction of non-intestinal cell lines (e.g. RajiB cells [201]) and the creation of an anoxic apical compartment allowing co-culture with intestinal bacteria [202]. Simpler alternatives allowing co-culture of intestinal and bacterial cells have also been developed [203]. These models, however, all concern static *in vitro* models, whereas intestinal epithelial cells are constantly exposed to different physical forces, including mechanical strain and shear stress [204]. This has been the basis for the development of so-called “gut-on-a-chips”, microfluidic devices in which cells are exposed to continuous flow providing shear stress and, moreover, continuous supply of nutrients [205-207]. This model was further improved by including a stretching and relaxing membrane, mimicking peristalsis-like motions. Caco-2 cells grown on this device were reported to differentiate into a small-intestine like epithelium – despite their colonic origin – including four distinct types of epithelial cells and the formation 3D villi-like structures. The inclusion of endothelial cells, immune cells and probiotic and pathogenic intestinal (aerobic) bacteria, provided the first proof-of-principle to use this device to model human intestinal health and disease [208-210]. Later representativity has been improved by inclusion of an anoxic apical compartment, co-culture with communities representative of the human intestinal microbiota and the incorporation of donor-derived cells (see below) [211, 212].

### *In vitro: From 2D monolayers to 3D intestine-like structures*

Although cell lines grown in regular, widely available well plates (2D) are able to form an intact monolayer, allowing treatment with (microbial) compounds in a reproducible and relatively cheap way, static monocultures often lack representativity of the host microenvironment *in vitro*. Alternatives include the use of human or other mammalian intestinal explants (reviewed in [213]) and precision-cut intestinal slices (reviewed in [214]), allowing *ex vivo* investigation of the intestinal epithelial barrier and host-microbe interactions (e.g. in Ussing chambers [215], everted gut sacs [216] or the inTESTine™ [217]). These models are, however, limited by low availability and short duration of viability. To recapitulate the 3D architecture of the *in vivo* intestinal tissue architecture *in vitro*, organoid (“organ-like”) models have been developed, which include cells and tissues embedded in or cultivated on top of extracellular matrix (ECM) scaffolds that allow cells to self-assemble into 3D structures [218]. Nowadays, however, the term “organoids” is more frequently used to describe the use of stem cells to regenerate the tissue of interest *in vitro*, which has recently accelerated the field of 3D modelling. More specifically, upon treatment with the right combination of growth factors, these stem cells develop into enteroids or adult stem cell-derived (ASC) organoids (derived from Lgr5-positive intestinal stem cells) or pluripotent stem cell-derived (PSC) organoids (derived from induced pluripotent stem

cells) [219, 220]. As opposed to colon carcinoma cell lines, of which some do not reveal all epithelial characteristics *in vitro* [193], stem cells are able to recapitulate the main types of epithelial cells observed *in vivo* and remain genetically stable for months to years [219-221]. Moreover, in the case of PSC organoids, the resulting tissue also includes mesenchymal cells, next to the intestinal epithelial cells [222]. Both ASC- and PSC-derived organoids have been successfully cultured as polarized monolayers in both static and dynamic *in vitro* models, facilitating personalized, functional studies including host-microbe interactions (amongst others [212, 223-225]).

## Modeling the intestinal microbiota

### *From dead to alive*

Despite recent successes in co-culturing intestinal epithelial cells with live (anaerobic) bacteria, a lot of (*in vitro*) research on host-microbe interactions has relied on the use of bacterial compounds, as opposed to the whole, living microorganism. These include bacterial cell-derived components (e.g., lipopolysaccharides (LPS), lipooligosaccharides (LOS) and exopolysaccharides (EPS), membrane fractions), metabolites (SCFAs, culture supernatant) or inactivated cells (e.g., heat-inactivated, pasteurized), which help to dissect the responsible mechanisms by which microbe-host effects are exerted. The collection of bacterial components to test is further limited by the cultivability of isolates in the lab. In fact, although metagenomics has advanced our insight into the diversity of the intestinal microbiota [226] and culturomics has enabled the cultivation of many members of “the unculturable” (reviewed in [227]), the majority of the identified intestinal species has not been cultivated yet [227, 228].

### *From single cells to (defined) communities*

As described earlier, intestinal microbes never operate alone, but are part of a community. Top-down approaches have been employed to obtain a representative model community, i.e., by using human fecal inocula for propagation *in vivo* (as discussed above) or *ex vivo* (see **Table 1**) using batch fermentations or GIT-mimicking chemostats. These facilitate a reproducible way to stably cultivate microbial communities, of which examples include the MacFarlane-Gibson three-stage continuous culture system [229], SHIME (Simulator of the Human Intestinal Microbial Ecosystem) [230], EnteroMix [231], the Lacroix model [232] and TIM-2 (TNO Intestinal Model) [67]. Complementarily, bottom-up approaches have also been applied, i.e., by selecting and combining known (cultivable) species based on their function, resulting in defined (or synthetic) communities modelling the human intestinal microbiota that are studied in batch or continuous fermentation experiments (*in vitro*) or in (gnotobiotic) animal models (*in vivo*).

Note that the term “synthetic community” has two additional meanings in the field of microbiology, as it can also refer to a combination of microbes not occurring



naturally *in vivo*, e.g. for biotechnological applications (reviewed in [233]). Alternatively, it can indicate the use of genetically modified microbes to investigate a wide range of hypotheses regarding microbe-microbe or host-microbe interactions. The latter does not only include “workhorse” *E.coli*, but also *B. thetaiotaomicron* as a well-characterized intestinal symbiont that has been subject to genetic manipulation (reviewed in [234]). As such, these bacteria are also considered model organisms.

### MODELING HOST-MICROBE INTERACTIONS AT THE COLONIC MUCOSAL LAYER *IN VITRO*

Despite being a crucial site for host-microbe interactions, the mucosal layer remains exceedingly difficult to address and, thus, study *in/ex vivo* or *in vitro*, for several reasons. From a host perspective, the mouse intestine does not entirely reflect the human intestine, as discussed previously. This also concerns mucus glycosylation of mice versus men: For instance, human colonic mucin is mostly core3- and core4-glycosylated [113], where murine colonic mucins carry largely core2-based glycans [120]. In addition, sialylation and fucosylation gradients along the human GIT are reversed in mice [177]. The investigation of mucus *in* or *ex vivo* is further complicated by collapse or shrinkage upon commonly used fixatives [235]. Recent alternatives include, amongst others, improved fixation strategies [235, 236] the use of (Combinatorial Labeling and Spectral Imaging) Fluorescent *In Situ* Hybridization ((CLASI-)FISH) [237, 238], and the use of fluorescent beads in *ex vivo* intestinal explants in which mucus thickness is preserved [239], which remain, however, invasive and time-consuming. As an alternative, several *in vitro* models have been developed, among which three different types can be distinguished: cellular, physiological and artificial models (adapted from [240]). All of these can be used to address (aspects of) host-microbe interactions.

#### Cellular models

Cellular models include *in vitro* models as discussed above, including 2D, 3D, static and dynamic models, which allow to study effects of microbes or microbial products on human cell response. As they do not represent goblet cells, Caco-2 cells fail to recapitulate the mucus layer [241], although they were reported to differentiate into, amongst others, goblet cells upon cultivation on a gut-on-a-chip [208]. As mentioned, HT29-MTX have been the primary model for mucus-producing goblet cells [242]. Under static conditions, these cells produce relatively more stomach-specific MUC5AC than MUC2, but upon culturing under Semi-Wet Interface with Mechanical Stimulation (SWMS) conditions, demonstrated increased MUC2 production and the formation of an adherent mucus layer [243]. In general, however, colon carcinoma cell lines demonstrate limited glycosylation capacities, including the lower abundance of core3 O-glycans [244]. Recently, the cultivation of 2D grown organoids have demonstrated the capacity to better represent the colonic mucus layer [224].

### Physiological models

Physiological models encompass isolated mucus from human and other animal subjects. Obviously, the former is most representative when interested in human mucus physiology, but is hampered by limited availability and volumes, lack of volunteers, high inter- and intravariability. Instead, mucus is often isolated from gastrointestinal mucosa of large animals, such as pigs and bovines. Apart from ethical concerns, these models do not offer the context of the (bilayered) colonic mucus and intestinal cell, and whether the obtained mucus is representative of human mucus physiology, including viscoelastic properties, net charge and mesh size, is further determined by intra-animal variability, the conservation method and the purification process (with or without lipids, proteins, DNA and cell debris) (reviewed in [240]). To partially overcome intra-animal variability and time constraints, commercially available mucus from pigs and bovines provides a commonly used source of mucin. Especially porcine gastric mucin (PGM) has been commonly used in research on host-microbe interactions. For instance, the incorporation of PGM in previously mentioned chemostats to recreate a mucosal environment, improved their representativity of *in vivo* microbial communities (e.g., [245, 246]). Moreover, PGM has been adopted in cellular mucus models: In the Host Microbiota Interaction module, Caco-2 cells were separated from microbe-containing SHIME effluent by a layer of PGM to study host-microbe interactions [247]. Similarly, the incorporation of a PGM-based mucus layer in a microfluidic device provided another method to co-culture Caco-2 cells with the anaerobic bacterium *Lactocaseibacillus rhamnosus* GG [248]. A drawback of these purified mucins, however, is the loss of other mucus components during extraction and purification [240]. In addition, PGM carries mucin glycans less abundant in the human colon (core1 and core2 [249]) and mucin types, due to its gastric origin (MUC5AC and MUC6) [250].

### Artificial models

A last category of mucin models is formed by synthetic mucins, which encompasses a wide range of strategies to mimic the broad range of biophysical properties of mucus *in vitro* (reviewed in [251]). Examples include non-peptide mucin substitutes (e.g., mucin-mimicking polymers such as hyaluronan, carboxymethyl cellulose), chemically synthesized polypeptide backbones and hydrogels. Some of these mucins can be expanded with (chemically synthesized) glycans, to study host-microbe interactions [251]. So far, the expression of complete mucin proteins including O-glycans has not been realized *in vitro*, but attempts have been made using structural fragments without [252, 253] or with control of O-glycosylation machinery [254, 255] in human cell lines and non-human cell lines.





### AIM OF THIS THESIS

This thesis aims to describe and critically evaluate currently used *in vivo* and *in vitro* models mimicking the human intestinal microbiota and/or the GIT to study bidirectional host-bacteria interactions, with a particular focus on the colonic mucosal layer. Specific attention will be on beneficial bacteria of the human GIT, as opposed to pathogenic species. The findings will be used to make recommendations for future *in vitro* models to investigate host-microbe interactions in the human GIT in a more representative manner, and replace, refine, and reduce the need for animal models in this research area.

### THESIS OUTLINE

This thesis can be roughly divided into two main parts. The first part of the thesis (**Chapters 2-3**) critically evaluates currently used *in vivo* and *in vitro* models mimicking the human intestinal microbiota, the human GIT, or their interactions (**Figure 3**). The second part of the thesis (**Chapters 4-6**) zooms in on host-microbe interactions at the colonic mucosal layer.

**Chapter 1** provides a general introduction to the human GIT and its microbial inhabitants, with a specific focus on the colonic mucosal layer. In addition, for each aspect, the available *in/ex vivo* and *in vitro* models used to investigate host, microbe and/or their interactions in the human GIT are presented.

**Chapter 2** reviews the use of defined microbial communities *in vivo* to study host-microbe interactions in the human GIT, as reported in literature. Based on these findings, we provide suggestions to improve current and future host-microbe interaction models, including more advanced *in vitro* models that are currently under development.

**Chapter 3** compares transcriptomes of the most used colonic carcinoma cell line – Caco-2 – cultivated under various cellular and physiological conditions, to systematically assess the effect of experimental and model parameters on cellular response. In addition to transcriptome data, we systematically extract and analyze model and experimental parameters as well as other study outcomes from selected studies. These parameters are used to uncover the main determinants for the cellular response of Caco-2 cells, to make recommendations for design of *in vitro* models of the human colon.

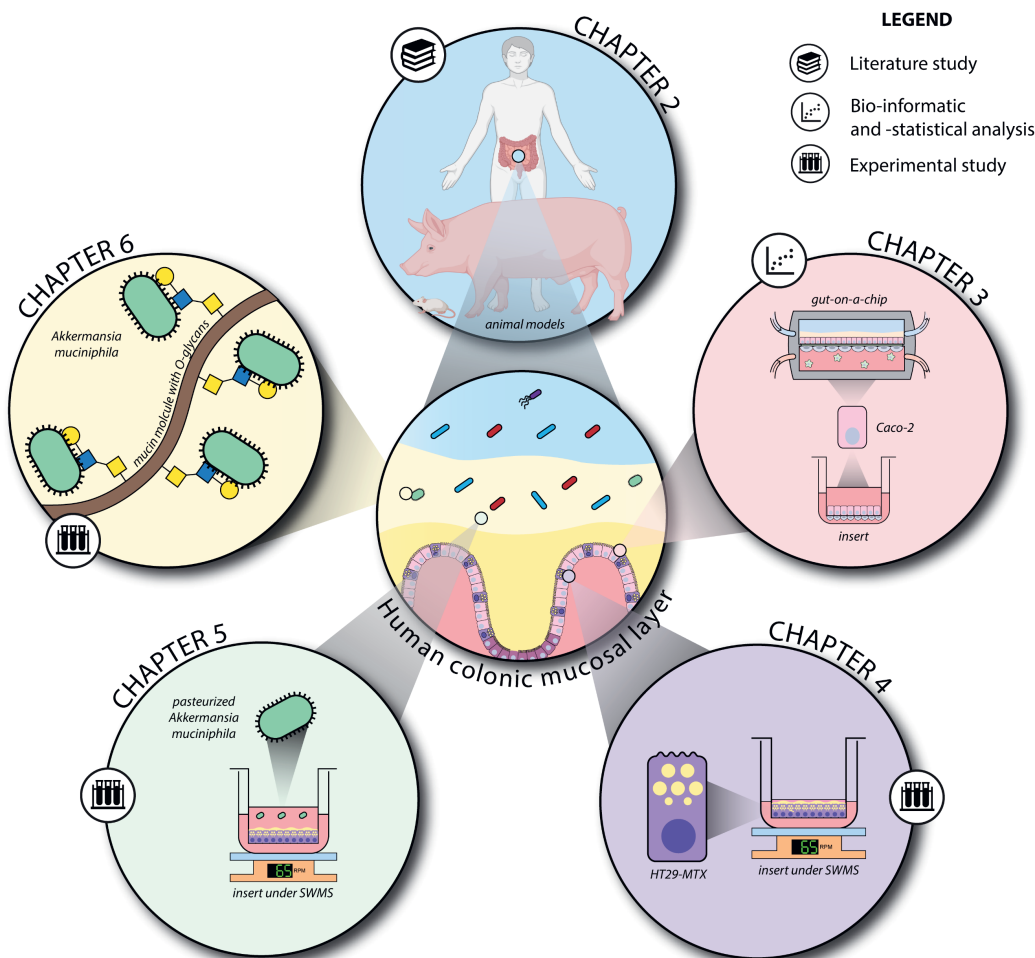
**Chapter 4** characterizes a previously established culture method – SWMS or Semi-Wet Interface with Mechanical Stimulation – which was demonstrated to enhance mucus stimulation in HT29-MTX cells. Using transcriptome analysis, combined with other experimental read-outs, we aim to reveal the molecular mechanisms underlying increased mucus production in HT29-MTX cells.

In **Chapter 5**, we apply the SWMS method to study the effect of mucin degrader *A. muciniphila* in its pasteurized form on the transcriptome of HT29-MTX cells. In addition, we discuss other *in vitro* models, including a gut-on-a-chip, which can be used to study microbe-mucus effects in the future.

**Chapter 6** takes a closer look at the binding of *A. muciniphila* to mucin. To dissect the molecular properties required for mucin binding, recombinant human mucin domains with tunable O-glycosylation are used as a model for human colonic mucin. This study aims to get more insight into the molecular cues required for mucin recognition and thus, potentially, colonization by a key mucin degrader.

Finally, **Chapter 7** discusses the recent developments in *in vitro* modelling of the human colonic mucosal layer and makes suggestions for future steps required to obtain *in vitro* models allowing the investigation of host-microbe interactions. This chapter also argues the role of slow science in this field, as well as in scientific practice and science communication in general.





**Figure 3. Contents of this thesis.** Schematic overview of the contents of this thesis, including figures that are used throughout this thesis. The human, pig and mouse were generated using Biorender.com.



“What I cannot create, I do not understand”

~ Richard Feynman



# CHAPTER 2

## The Use of Defined Microbial Communities to Model Host-Microbe Interactions in the Human Gut

Janneke Elzinga,<sup>1</sup> John van der Oost,<sup>1</sup> Willem M. de Vos,<sup>1,2</sup> Hauke Smidt<sup>1</sup>

<sup>1</sup>Laboratory of Microbiology, Wageningen University & Research, Wageningen, The Netherlands

<sup>2</sup>Research Programme Unit Human Microbiome, Faculty of Medicine, Helsinki University, Helsinki, Finland

Published in

*Microbiology and Molecular Biol Reviews* 2019;83(2): e00054-18

## The Use of Defined Microbial Communities to Model Host-Microbe Interactions in the Human Gut

### SUMMARY

The human intestinal ecosystem is characterized by a complex interplay between different microorganisms and the host. The high variation within the human population further complicates the quest towards adequate understanding of this complex system that is so relevant to human health and well-being. To study host-microbe interactions, defined synthetic bacterial communities have been introduced in gnotobiotic animals or in sophisticated *in vitro* cell models. This review reinforces that our limited understanding has often hampered appropriate design of defined communities that represent the human gut microbiota. On top of this, some communities have been applied to *in vivo* models that differ appreciably from the human host. In this review, the advantages and disadvantages of using defined microbial communities are outlined, and suggestions for future improvement of host-microbe interaction models are provided. With respect to the host, technological advances, such as the development of a gut-on-a-chip and intestinal organoids, may contribute to more accurate *in vitro* models of the human host. With respect to the microbiota, due to increasing availability of representative cultured isolates and their genomic sequences, our understanding and controllability of the human gut ‘core microbiota’ is likely to increase. Taken together, these advancements could further unravel the molecular mechanisms underlying the human-gut microbiota superorganism. Such a gain of insight would provide a solid basis for the improvement of pre-, pro- and synbiotics as well as the development of new therapeutic microbes.

## INTRODUCTION

Given its involvement in metabolic, nutritional, physiological and immunological processes, the human intestinal microbiome can be regarded as an essential organ of the human body [256]. Further strengthening its clinical relevance, the intestinal microbiome has been linked to numerous disease conditions, including metabolic and immune disorders, cancer and neurodegenerative diseases [257]. Apart from a remarkable increase in genome sequence data of the human gut microbiota, however, progress in functional insight has been hampered by its complexity: the existence of more than 1,000 prevalent species [258] combined with the high interpersonal variation within the human population in terms of genetics, environment and habits, results in a complex entity termed the human-microbiome superorganism [259]. The number of known host-microbe interactions has grown rapidly over the past decades, yet many aspects still remain obscure.

To solve this complexity, there is need for a reductionist approach in which both host and microbiome are simplified to the extent that experimental variables can be tightly controlled and deliberately manipulated. Regarding the microbiota, synthetic or defined communities have been proposed as useful models to study microbial ecology [260]. In recent years, the number of cultivable gastrointestinal microbial species has rapidly expanded [258] by the use of sophisticated or brute force culturomics approaches [16, 261]. These strategies have allowed for the design of defined communities that are representative of the normal human intestinal microbiota. With respect to the human host, laboratory animals, notably mice, have proven valuable models for developing human medicine. The colonization of germ-free (GF) animals with defined bacterial communities, resulting in gnotobiotic animals, has already been applied for decades. During the 1960s and 1970s, it was recognized that the intestines of GF animals display aberrant histological, anatomical and physiological characteristics compared to conventional laboratory animals [170]. The development of the Schaedler cocktail for colonization of the murine gut [262] marked one of the first attempts to normalize GF mice. An altered version has been widely adopted as a standardised gut microbiota by animal breeders and biomedical researchers ever since. Over time, various other defined communities have been designed to generate gnotobiotic animals for purposes beyond standardisation; they have proven a valuable *in vivo* tool to study microbial ecology (e.g., microbial invasion, microbe-microbe interactions, and metabolism) and host-microbe interactions. However, mice and other animal models have various limitations that hamper their use as models for the human microbiome, as has been recently reviewed [263, 264]. Interesting alternatives concern the development of sophisticated *in vitro* models, such as organ-on-chip systems and organoids.

This review summarizes existing models of host-microbe interactions in which defined communities, as models of the (human) gut microbiota, were applied. We aim





## CHAPTER 2

to present all *in vivo* studies that used defined microbial communities representing the intestinal microbiota of healthy individuals and in which host parameters were considered. The design of these model communities, as well as the selection of its host, are compared and critically evaluated. The potential use of defined communities in *in vitro* (cellular) models, as a surrogate host, are outlined as well. We conclude by discussing the increased value, opportunities and possible obstacles when applying defined communities in to-be-developed *in vitro* host-microbe interaction models.

## MAIN TEXT

**Defined Communities Mimicking the Normal Intestinal Microbiota *In Vivo***

A number of recent studies addressed host-microbe interactions *in vivo* by using defined communities representative of the healthy human gut microbiota (**Table 1a-c**). These include various mouse studies with more or less defined intestinal microbiota that are summarized below. Studies in which animals were antibiotic-treated before bacterial colonization are excluded from our analysis as their reproducibility and gnotobiology cannot be reassured [265]. The following section first discusses the specifically named defined communities applied in rodents (**Table 1a**,  $n = 31$ ), followed by non-specifically named communities in rodents (**Table 1b**,  $n = 16$ ). Finally, the defined communities administered to non-rodent models are discussed (**Table 1c**,  $n = 6$ ).

*(Altered) Schaedler flora*

In 1965, Russel W. Schaedler colonized GF mice with a defined microbial community composed of strains isolated from normal mice, to study the fate of the bacteria in the gastrointestinal tract (GIT) and their effect on caecum size. With respect to these parameters, it turned out that the Schaedler flora (SF) was able to, at least partially, normalize the caecum size of the GF size in comparison with animals raised under conventional conditions [262]. The defined microbial population was supplied to animal vendors to serve as a community that could limit the infection of ex-GF rodents with opportunistic pathogens. Schaedler developed several different bacterial cocktails over time. In 1978, Roger P. Orcutt set out to standardize and improve the SF flora, but in view of the monitoring costs, the total number of bacterial species was limited to eight. Orcutt made a selection of bacterial species (Altered Schaedler Flora (ASF)) based on their representation and stable colonization in the murine gut, their ease of identification (morphologically) and their presence in or interference with isolator contaminants. For instance, the cocci and spore-forming, blunt-ended rods were eliminated, which represented the majority of isolator contaminants. Also, the amount of facultative anaerobes was limited, as they outgrew aerobic isolator contaminants and thus, impeded the ability to detect the latter [266]. The ASF consists of six Firmicutes (*Clostridium* species (ASF356), *Lactobacillus intestinalis* or *acidophilus* (ASF360), *Lactobacillus murinus* or *salivarius* (ASF361), *Eubacterium plexicaudatum* (ASF492), *Pseudoflavonifractor* sp. (ASF500) and *Clostridium* sp. (ASF502)), one Bacteroidetes (*Parabacteroides distasonis* (ASF519)) and one Deferribacteres (*Mucispirillum schaedleri* (ASF457)).

The ASF has been used multiple times as a reference or minimal defined microbiota, and its applications were extensively reviewed elsewhere [267]. Several studies involving ASF in mice (or other animals) reported its effect on host parameters (**Table 1a-c**). The list is probably not exhaustive, given the wide application of ASF

mice as control or minor population in studies, which makes these studies harder to identify. The applications of ASF in rodents varied from wild-type strains (mostly C57BL/6, but also C3H/HeN and Swiss-Webster mice) to models prone to diseases including IBD [268-270], type I diabetes [271] or colorectal cancer [272]. The ASF lacks Proteobacteria, a phylum shared by mice and humans, whereas some researchers did introduce Proteobacteria to ASF mice, such as *Oxalobacter formigenes* [273] and *Escherichia coli* [274]. Other studies included only selected members of the ASF, because not all were found to successfully colonize the murine caecum [271] or to test the level of colonization resistance of different combinations of ASF members [275]. Overall, the application of ASF to study host-microbe interactions has been quite diverse, regarding host strain, gut region of interest and host parameters studied. Although the ASF has been used multiple times as a reference microbiota and has aided in the establishment of other defined microbiota, such as Oligo-MM and the Bristol Microbiota, its representability of the normal gut microbiota has been criticized [276], as discussed later in this review.

### *Oligo-MM*

Another defined community of murine microbiota, Oligo-MM<sup>12</sup>, was constructed in an attempt to provide full colonization resistance against *Salmonella enterica* serovar Typhimurium (S. Tm) [275]. Twelve strains were selected to represent the five most prevalent and abundant phyla of the laboratory mouse intestine, i.e. Firmicutes: 'Acutalibacter muris', *Flavonifractor plautii*, *Clostridium clostridioforme*, *Blautia coccoides*, *Clostridium innocuum*, *Lactobacillus reuteri*, *Enterococcus faecalis*; Bacteroidetes: 'Bacteroides caecimuris', 'Muribaculum intestinale'; Actinobacteria: *Bifidobacterium longum* subsp. *animalis*; Proteobacteria: 'Turicimonas muris' and Verrucomicrobia: *Akkermansia muciniphila*. Colonization resistance of ASF mice or mice colonized with Oligo-MM<sup>12</sup> and/or (a subset of) ASF strains, were compared to conventional mice. ASF was used as a reference, because of its wide usage in gnotobiotic mouse research. Oligo-MM<sup>12</sup> mice conferred increased, but not full, resistance compared to mice colonized with a subset of ASF strains with and without Oligo-MM. Functional genomic analysis of Oligo-MM and whole ASF revealed that both consortia together cover 66.6% of the KEGG modules of a conventional mouse microbiota. Addition of three facultative anaerobes (*E. coli*, *Streptococcus danieliae* and *Staphylococcus xylosus*), underrepresented in Oligo-MM<sup>12</sup>, increased coverage and furthermore, conferred full colonization resistance [275]. C57Bl/6 mice stably colonized with Oligo-MM<sup>12</sup> have been designated stable Defined Moderately Diverse Microbiota mice (sDMDMm2). The designers of Oligo-MM<sup>12</sup> stressed the importance of expanding the amount of available mouse-derived strains, as initiated recently [277], in favour of the design of functionally defined and simplified microbial consortia for application in gnotobiotic animals [275]. Because Oligo-MM<sup>12</sup> found to lack the enzymatic pathway to carry out 7 $\alpha$ -dehydroxylation, an important bile acid

transformation, the addition of *Clostridium scindens* (a 7 $\alpha$ -dehydroxylating bacterium) was tested in another study. This modification normalized large intestinal bile acid composition in mice, which was accompanied by colonization resistance against *Clostridium difficile* and decreased intestinal pathology [278]. Finally, Oligo-MM<sup>12</sup> served as a defined reference microbiota to verify the significant difference between the bacterial composition in the large intestinal outer mucus layer and the lumen [279], but host parameters were not assessed. Note that the latter two studies that applied of Oligo-MM<sup>12</sup> left out the three additional facultative anaerobes that were found to be crucial for full colonization resistance.

### SIHUMI(x)

Because ASF was found to poorly represent the dominant intestinal bacteria and ASF mice hardly differed from GF mice in a key set of microbial biochemical activities [276] (Midtvedt criteria, see below), a simplified human intestinal microbiota (SIHUMI) was established in rats to provide a highly standardized animal model to study host-microbe interactions. Species were selected according to their prevalence in humans, their fermentative capacity, the availability of their genomic sequence and their ability to stably colonize the rodent gut. SIHUMI(x) includes four Firmicutes (*Anaerostipes caccae*), *Lactobacillus plantarum*, *Blautia producta* and *Clostridium ramosum*), one Bacteroidetes (*Bacteroides thetaiotaomicron*), one Actinobacterium (*B. longum*) and one proteobacterium (*E. coli*). All seven members successfully colonized the rat intestinal tract and total bacterial numbers in faecal samples did not differ from those in human faeces. The amount of short-chain fatty acids (SCFAs) produced, however, was dramatically lower compared to humans, probably owing to the smaller number of species. An eighth species was added to the consortium (SIHUMIx), *Clostridium butyricum*, which led to increased butyrate production. All members of the SIHUMIx were successfully transferred to offspring. Dietary interventions varying in fibre and fat content resulted in responses (partially) reflecting those observed in mice and humans [280].

In other studies, SIHUMIx served as a resident community to study the effect of the addition or removal of species. For instance, inclusion of *A. muciniphila*, a mucin-degrading microbe, was found to worsen intestinal inflammation induced by *S. typhimurium* Tm in mice [281]. The same researchers recently showed, however, that in a colitis-prone mouse model colonized with SIHUMI, *A. muciniphila* did not induce or exacerbate intestinal inflammation [282]. In two other studies, the polyamine-producing *Fusobacterium varium* was added to the low polyamine-producing SIHUMIx in mice, which disclosed that gut morphology was neither affected by increased putrescine concentrations [283], nor by higher levels of other polyamines and SCFAs [284]. Additionally, the mechanism underlying the obesogenic potential of *C. ramosum* in a SIHUMIx-associated animal model was further investigated by including or excluding this bacterium in SIHUMIx-associated mice fed a high- or a low-fat diet.

The increased body fat deposition in the presence of *C. ramosum* was suggested to be due to the upregulation of small intestinal glucose and fat transporters [285]. It should be noted that, although SIHUMI was originally established in rats, all other studies applied the community in mice.

### *Towards a normal model gut microbiota*

Since the generation of the Schaedler flora in the 1960s, other defined gut microbiotas have been developed in an attempt to normalize GF animals or generate animal models harbouring a bacterial community representative of the human gut microbiome. During the 1970s, Syed *et al.* aimed to normalize GF mice with respect to caecum size, caecal numbers of *E. coli*, histology of the intestinal tract, and the development of a mucosa-associated microbiota in stomach and large intestine [286]. A mixture of 50 strictly anaerobic (later designated 'N-strains' [287]) and 70 facultative anaerobes ('F-strains') were found to generate a normal mouse phenotype, whereas less complex bacterial communities led to intermediate phenotypes with respect to the parameters studied, including caecum size, caecal *E. coli* levels, GIT histology and development of a mucosa-associated microbiota in stomach and large intestine [286]. The exact taxonomic classification of the species within the F- and N-strains was limited by lack of characterization at that time [286]. It was considered likely that a number of the isolates used were identical. Based on morphology and fatty acid production, the total of number of different strains was estimated to be rather in the order of 35 (N-strains) and 60 (F-strains) [287]. The N-strains alone could not control the *E. coli* population and caecum size when associated with mice fed on a crude instead of refined diet, but this could be restored by additional association with the F-strains [287]. The F-strains were exploited as an indigenous gut microbiota to investigate *E. coli* plasmid transfer *in vivo* [288], but other studies using the N- or F-strains could not be identified.

At the end of the 1970s, the use of the UW-GL (University of Wisconsin Gnotobiotic Laboratory) flora was reported, which was used as the intestinal microbiota of heterozygous athymic mice [289]. This defined bacteriome consisted of nine Gram-positive species from the genera *Lactobacillus*, *Bacillus*, *Clostridium* and *Corynebacterium* [290] and additionally, two Gram-negative species that were not further specified [289]. It was used to study its colonization resistance against *Candida albicans* [290] and *Clostridium botulinum* [289]. The latter study compared UW-GL with other defined microbiotas including ASF and a partial UW-GL. Whereas death rates significantly dropped compared to GF mice, only complete UW-GL fully prevented *C. botulinum* infection [289]. The use of the UW-GL microbiota has not been reported since.

Logically, the conception of a healthy or 'normal' microbiota is dependent on the available knowledge on conventional animals and/or healthy human subjects, and thus the composition varied per study. While testing the effect of bacterial species on

intestinal IgA immune system development, Moreau *et al.* paid specific attention to communities of *Clostridium* species, which was considered a dominant microbiota of the digestive tract of adult conventional mice [291]. In studies using defined communities with human-derived gut bacteria, species were selected based on their prevalence in (healthy) human faeces [292, 293] and/or their representation of the major three or four dominant phyla of the human gut microbiota [293-295]. Next to the designers of Oligo-MM<sup>12</sup>, only few studies acknowledged the presence of five phyla (including *Verrucomicrobia*) of the human gut microbiota. A recently designed 14-membered synthetic microbiota that collectively possessed important core metabolic capabilities was applied to study *in vivo* foraging of host-derived mucus glycoproteins during fibre deprivation [296]. Similarly, other studies took into account the functional capabilities of species. For instance, one study included species that are able to break down complex dietary polysaccharides not accessible to the host (*B. thetaiotaomicon*, *Bacteroides ovatus*, *Bacteroides caccae*), to consume oligosaccharides and simple sugars (*Eubacterium rectale*, *Marvinbryantia formatexigens*, *Collinsella aerofaciens* and *E. coli*), to ferment amino acids (*Clostridium symbiosum* and *E. coli*) or to remove the end products of fermentation by reducing sulfate (*Desulfovibrio piger*) or generate acetate (*Blautia hydrogenotrophica*) [294]. This community has been frequently exploited to study host-microbe interactions or microbe-microbe interactions by the same research group or adopted by others, albeit in different combinations ranging from eight to 15 species [293, 295, 297-303]. Recently, a more diverse, complex defined community comprising not less than 92 species was developed [304]. The consortium consisted of phylogenetically diverse, human-derived bacterial strains, which had previously been cultured and sequenced. It also included strains representing species that were demonstrated to be age- and/or growth-discriminatory in models of microbiota development during the first years of life. Of all strains, 44 comprised a core group that could be detected in faecal samples of all colonized mice, independent of dietary intervention [304]. No host parameters, however, were assessed in this study.

Remaining inclusion criteria for defined communities are the availability of the genomic sequence and the cultivability of the species. Obviously, both criteria make each individual species more easily traceable. If the entire genetic repertoire of the defined community is known, gene expression of the whole community as well as its individual members can easily be assessed [281, 293] and their function can be more precisely predicted. Interestingly, although ASF has been used for over 50 years, publications on replication of the four extremely-oxygen sensitive ASF members on a defined medium, is still lacking [267].

#### *Defined communities in non-rodents*

Previously discussed defined microbiota were either isolated from rodents or applied to them. Laycock *et al.* stressed the need for a well- established intestinal colonization



microbiota for pigs, given the higher representability of these animal models in early immune development studies [305]: in pigs, there is no transfer of maternal immunoglobulin G *in utero* [306, 307], and a poorly developed mucosal system in neonates [308]. Furthermore, pigs are genetically more similar to humans than mice [309], and their digestive physiology is comparable to ours [310]. Colonization of germ-free piglets with ASF members turned out to be largely unsuccessful and only the most consistently colonizing ASF member (*Parabacteroides* sp.) was incorporated in the novel 'Bristol' microbiota. Additional strains were selected based on their representation of the major phylogenetic groups in gut sections of 12–18-week-old pigs, and either their ability to grow on a wide range of metabolic carbohydrate structures (*Roseburia intestinalis*) or their presence in unweaned pigs (*Clostridium glycolicum* and *Lactobacillus amylovorus*). Except for *R. intestinalis*, the novel microbiota successfully colonized the GIT after administration to germ-free piglets, with high clinical safety and an expected increase in immunoglobulin serum levels [305]. The Bristol microbiota was exploited by other researchers as a simplified starter microbiota to study additional effects of a complex microbiota on early life microbiota development [311], the intestinal expression of a butyrate-sensing olfactory receptor [312] and on the gastric transcriptome [313]. Note that in the latter three studies, the piglets were not maintained in a sterile environment, hampering comparison of the effect of the Bristol microbiota on host parameters between studies. A different ten-membered porcine gut microbiota, originally designed as a competitive exclusion culture for pigs, was used to investigate antibody repertoire development in ex-germ-free newborn piglets [314]. Another 'defined commensal microflora' (DMF) included seven porcine bacterial species and was similar in composition to ASF. Species were originally isolated from the caecal contents of six-week-old healthy pigs and administered to germ-free pigs to evaluate the interactions between intestinal commensals, antibiotics, probiotics, and human rotavirus. This model was primarily applied as a model commensal gut microbiota of neonates [315, 316].

### Other Defined Communities *In Vivo*

Apart from the defined communities as model for the normal (human) gut microbiome to study host-microbe interactions, other kinds of communities have been composed for application in gnotobiotic animals. These communities, however, are not listed in **Table 1a-c** and their application goes beyond the scope of this review, as they did not aim to represent the 'normal' microbiota. For instance, these include disease-specific consortia, e.g., IBD-related [268, 317-320]. Others are age-specific, such as the Human Baby Microbiota [321-323], DMF [315, 316] and a recently developed *Bifidum*-dominated model consortium [324]. Lastly, some communities were developed for therapeutic or probiotic purposes. A well-studied and globally marketed multispecies probiotic is the bacterial cocktail VSL#3, which was recently

characterized at the genomic level and has been used to treat various gastro-intestinal disorders [325-327]. Other communities were designed to treat infections (amongst others, *C. difficile* infection (CDI) [328-330] and colitis [331]), or to facilitate recovery of cholera [332]. Two remarkable applications of defined communities, which were not per se meant to model the normal human gut microbiota, are discussed in more detail below.

### *Therapeutic communities*

Although the concept is not new and pioneered already 30 years ago [328], the interest in faecal transplantations has recently increased and the avenue of synthetic microbiotas as stool substitutes has been suggested [333]. A particular example of such a stool substitute is Microbial Ecosystem Therapeutic 1 (MET-1), designed as a synthetic stool mixture to treat recurrent CDI. Sixty-two species were recovered from the stool of a healthy 41-year-old female donor, of which 33 species were selected that were sensitive to a range of antimicrobials and were easy to culture. Two CDI patients that were 'rePOOPulated' with MET-1 returned to their normal bowel pattern within a few days and remained symptom-free for at least six months. The use of a synthetic stool mixture has several advantages over conventional stool transplants: (i) the bacterial composition is known, controllable and reproducible, (ii) a pure consortium is more stable than stool, (iii) the formulation is safe, owing to the lack of viruses and pathogens, and (iv) the administered organisms can be selected based on their sensitivity to antimicrobials, which further enhances safety [330]. Some of these benefits also strengthen the use of defined communities in host-microbe interaction research. Notably, the application of MET-1 as a defined community in GF animals, instead of antibiotic-treated animals, was limited to one study, in which it was used as a healthy, Firmicutes-rich microbiota to study colitis susceptibility and host immune responses [334].

In contrast to the use of a defined synthetic community, the anaerobically cultivated human intestinal microflora (ACHIM) has been derived from a fecal sample from a healthy Western donor that has been maintained in anaerobic culture for more than 20 years now and has been applied in faecal microbiota transplantation [335]. Although the microbiota is regularly checked for the absence of pathogenic organisms and multiple CDI patients have been treated successfully with this cultured microbiota transplant from a single donor [335], its composition is not controllable.

Instead of starting with a certain disease or phenotype and generating a defined community to treat this condition, as true for MET-1 and ACHIM, researchers recently tested different defined bacterial communities to generate various phenotypes in mice and to identify the strains responsible for the observed phenotypic variation. By administering GF mice with one of 94 different, defined bacterial consortia of species randomly drawn from the culture collection, strains were identified that modulated adiposity, intestinal metabolite composition and the immune



system. According to the authors, a similar approach could be applied to identify and characterize next-generation probiotics or combinations of pre- and probiotics [336].

### *Minimal communities*

Another category of defined communities is formed by minimal communities. Essentially, all defined microbial communities are minimal in the sense that they are not as complex as microbiota *in vivo*. Nonetheless, some studies exploited even more simplified defined consortia, *i.e.*, with a limited number of species or clearly lacking certain functions, to study host-microbe interactions in general. This is exemplified by bi-association studies involving single members of (dominant) phyla. In a recent study GF mice were colonized with *B. thetaiotaomicron*, as a prominent member of the adult human gut microbiota, plus one of three probiotic strains (*B. longum*, *B. animalis* or *Lactobacillus casei*) to study microbe-microbe and host-microbe interactions [337]. In the same lab, gnotobiotic mice were colonized with bacteria from the two dominant phyla in the adult human distal gut microbiota – Firmicutes and Bacteroidetes. Based on their prominence in culture-independent surveys in the distal human gut, the pattern of representation of carbohydrate active enzymes in their glycobomes and *E. rectale*'s ability to generate butyrate as a major end product of fermentation, a 'marriage was arranged' between *E. rectale* and *B. thetaiotaomicron*. This reductionist approach provided information on microbe-microbe interactions, the microbial response to host diet and the microbial effects on host physiology (*e.g.* the upregulation of production of (mucin) glycans by the host) [338].

Despite the value of minimal communities for studying microbe-microbe and host-microbe interactions, a study into mice colonized with another simplified microbiota (*B. thetaiotaomicron* and *B. longum*) clearly demonstrated that the simple microbiota could not reconstitute the metabolomic complexity of a humanized microbiota, *i.e.* derived from human donors [339]. Nevertheless, **Table 1b-c** includes some minimal communities, because of their representation of major phyla of the human gut microbiota or relevant application to study host parameters.

### **Critical Evaluation of Defined Communities *In Vivo*: The Microbiota**

In the preceding sections, we provided an objective description of defined microbial communities that have been applied in *in vivo* models to study host-microbe interactions. The next section discusses the representability of these communities, focussing on their design criteria and source (murine vs. human). Additionally, a comparison is made between simple versus complex, and bottom-up versus top-down constructed communities. Suggestions for future design of defined communities representing the normal intestinal microbiota are provided as well.

#### *How representative are defined microbiota models of a normal microbiota?*

The development of defined communities representative of the human gut microbiota raises the issue: "What defines a normal microbiota?". Among the included studies

that aimed to design a representative gut microbiota, different selection criteria were used. The representation of the major phyla and various metabolic capacities have been frequently put forward. A meta-analysis was performed comparing the composition of the core mouse gut microbiome (based on five different mouse models, *i.e.* varying in age, phenotype and sampling site) with the human gut microbiome (based on 16 individuals) [340]. Apart from the differences within the mouse microbiota, Bacteroidetes and Firmicutes were clearly the most dominant phyla in all samples (together 87-97%). [340] The same is true for the composition of well-established defined communities ASF, SIHUMI(x) and Oligo-MM<sup>12</sup> (75-87.5%). Similar to most murine microbiota included in the meta-analysis, however, ASF and SIHUMI(x) lack Verrucomicrobia, which was found among the five most abundant phyla in human and some murine samples [340]. In that sense, Oligo-MM<sup>12</sup>, originally designed to represent the murine microbiota, is compositionally more complete than SIHUMI(x), which was meant to represent the human microbiota. The frequently used ASF also lacks Actinobacteria and Proteobacteria, which are abundant in both murine and human samples [340-342]. Similarly, a large part of the other defined communities discussed here (**Table 1a-c**) did not include representatives of all five major phyla of the human microbiota, some not even one of the two most prominent phyla. Note that species selection has been mostly based on microbiota composition of Western individuals.

Further, community design has been limited by availability of genomes and cultivability of strains. In the case of ASF, the number of species was limited for financial reasons, *i.e.*, taking into account the monitoring costs. Nevertheless, this community has been frequently used in gnotobiotic animal models. The assumption that ASF mice can be regarded as conventional mice with respect to their gut microbiota, has been criticized [276]. Several functional activities in faecal materials from ASF mice were analysed and compared to samples from GF and conventional rodents and other mammalian species, including humans. The five biomarkers investigated, the so-called Midtvedt criteria (*i.e.* conversion of cholesterol to coprostanol, conversion of bilirubin to urobilinogens, degradation of  $\beta$ -aspartylglycine, degradation of mucin, and the absence of fecal tryptic activity [276]) are claimed to reflect host-bacterial interactions, independent of the intestinal localization of the bacteria involved and the kind of species. With regard to these criteria, faecal samples from ASF mice showed patterns more resembling GF rather than conventional mice [276], which complemented previous results demonstrating an abnormal microbiota in mice raised under specific pathogen free (SPF) conditions [343]. Although this could be due to one of the limitations of ASF, *i.e.* its low diversity, ASF mice were shown to be immunologically, reproductively and metabolically similar to conventional mice [276]. The Midtvedt criteria were also used to assess the suitability of SIHUMI(x) as a model microbiome. SIHUMI(x)-associated rats shared



four criteria with conventional rats, of which three were, however, less pronounced [280].

A major difference between ASF and a consortium such as SIHUMI(x), is the fact that the latter involves human-derived bacterial strains. Most members of recently developed communities, except for Oligo-MM, are of human origin as well. This may be obvious, given the fact that, although their microbiota is similar at the division (superkingdom) level, 85% of the microbial genera and species detected in mice are not found in humans [344]. Although qualitatively, humans and mice share a largely similar core, their intestinal microbiota is quantitatively very different [340]. On the other hand, the development of small intestinal immune maturation was found to be host-specific, with humanized mice resembling more closely GF mice than mice associated with a murine microbiota [345]. This host-specificity might also, at least partially, explain the unsuccessful colonization of piglets with ASF [305]. Additionally, humanized rodent models were claimed to have been utilized mainly for short-term biomedical research studies [267]. The question remains how human-derived bacteria would adapt during long-term colonization and vertical transmission in murine hosts [267, 346, 347], and thus, which kind of microbiota would be most reliable to study host-microbe interactions when using murine hosts. The maximum colonization time reported in the studies discussed here (**Table 1a-c**) was less than one year. With respect to vertical transmission, stability after transfer to offspring has been addressed mainly for murine microbiota only (ASF [348] and Oligo-MM [275]). Within the humanized defined communities, SIHUMI(x) is an exception, of which bacterial concentrations in caecum were verified between founder rats as well as their offspring. At the age of eight weeks, SIHUMI(x)-rats harboured similar bacterial levels as their founders, but not at two weeks (except for *E. coli*) [280].

### *Simplified versus complex communities*

The distinction between minimal communities, with two or three members, and larger defined communities is not black-and-white. For instance, ASF, initially used as a microbiota to standardize mouse models, slowly adopted the role of a minimal community, instead of one representing the normal microbiota of mice. Nonetheless, the simplicity of a defined community also has some advantages over more complex communities. The limited nature of ASF should, as proposed by Brand *et al.*, allow investigators to evaluate the *in vivo* effect of the removal or addition of bacterial species on mucosal homeostasis and colonization dynamics, or potentially, factorial interactions of the community [267]. Indeed, some of the studies discussed here (**Table 1a-c**) used only a subset of the ASF species or added species to already established defined communities, including ASF and SIHUMI(x). Additionally, one- and two-member communities could be applied to model aspects of a more complete microbiota, such as depletion of certain dietary compounds or metabolites [339].

Finally, as already discussed, a simplified consortium makes each species traceable, as opposed to a very complex community [281, 293].

On the other hand, complex communities might more closely resemble the normal human gut microbiota and are more likely to confer colonization resistance to opportunistic pathogens, which has been a frequently mentioned criterion in the studies described above. In the 1980s, Freter and co-workers formulated the nutrient-niche theory, which states that a certain bacterium can only successfully colonize if it is able to use a specific limiting nutrient more efficiently than its competitors [349]. This implies that colonization resistance correlates with community complexity, as supported by several studies [275, 289, 350]. Freter's theory was corroborated in a recent study in which the relative abundance of each species of a ten-membered community was correctly predicted based on the concentration of individual dietary ingredients [294]. The theory assumes, however, an environment in which bacterial growth is balanced and nutrients are perfectly mixed, whereas in reality bacteria are metabolically flexible (*i.e.* they have the ability to switch nutrient source) and nutrient levels in the gut are spatiotemporally heterogeneous (reviewed in [54, 351]). Metabolic flexibility was hardly addressed in the studies discussed in this review. Some researchers did assure the inclusion of species in a defined community that, as a whole, was able to thrive on a wide range of nutrients. Once established *in vivo*, however, the behaviour of the community was seldomly addressed or only for a single species. This could be due to the fact that most of the included studies focused primarily on the effects of the whole microbiota or a subset of species on the host (host-microbe), rather than the exact nutrient niche occupation by its separate species (microbe-microbe interactions). Exceptional is a recent study, which quantified the *in vivo* response of both mucin-specialists (*A. muciniphila* and *Barnesiella intestinihominis*) and -generalists (*B. caccae* and *B. thetaiotaomicron*) upon fibre deprivation [296]. A fibre-deficient diet stimulated the expansion and activity of the mucus-degrading bacteria, promoting epithelial access and pathogen-induced colitis [296].

With respect to spatiotemporal heterogeneity, Oligo-MM<sup>12</sup> was used to verify that the bacterial compositions in the large intestinal outer mucus layer and the intestinal lumen are significantly different [279]. Due to extensive mucus shedding and mixing in the lumen, however, the differences may be relatively small [351]. Indeed, it was recently shown that, at microscale level, the proximal colon should be viewed as a partially mixed bioreactor rather than a clearly compartmentalized gut section with spatially segregated communities. A next step would be to quantify the distribution of nutrients and metabolites and the role of host factors such as diet, gut motility and mucus composition [301]. Vice versa, it would be interesting to study the effect of spatial organization on relevant host parameters, which were unfortunately not addressed in the latter study. The authors did admit that the 15-membered community used may not be complex enough to demonstrate stronger spatial



associations with food particles, host cells and mucus [301], reinforcing, all in all, the need for more complex communities.

Both metabolic flexibility and spatiotemporal heterogeneity allow for increased community diversity, which is thought to be crucial for ecosystem robustness [351]. Defined communities enable the precise investigation of both concepts, but, on the other hand, the question remains whether they can be made sufficiently complex to properly address these issues.

### *Bottom-up versus top-down approaches*

One way to obtain a more complex model community is to start with a complex sample, e.g. human stool, and narrowing the amount of species down via one or more enrichment steps, e.g. by culturing on selective media (top-down approach [352]) or using fermentation models. **Table 1a-c** includes only a few examples with regard to normal microbiota (Oligo-MM<sup>12</sup> [275], [293]). The majority of the studies listed in **Table 1a-c** used a bottom-up approach, in which single, previously cultured and characterized strains are combined into a synthetic bacterial community, e.g., based on selection criteria previously mentioned, and administered to germ-free animals. An advantage of the latter method is the known composition of the microbiota, as previously emphasized. A drawback, however, is formed by the risk that the desired phenotype (in this case a normalized host) cannot be entirely recapitulated [352].

### *Future design*

A probably more important question is whether a normal microbiota actually exists. In the 1970s, Freter *et al.* concluded that significant fluctuations occur in the normal microbiota and that there is “no such a thing as a reproducible and precisely definable ‘normal enteric flora’”. Instead, they considered the F-strains collection most optimal to use as a microbiota representing a “state which is sometimes found in ‘normal’ individuals” [287]. Clearly, the concept of the normal microbiota has changed over time and has evolved with the development of techniques to sequence the human gut microbiome, with increased insight into its composition, dynamics, and function. Recently, researchers aimed to draw the compositional functional core of the human gut microbiota, or the core microbiome. They emphasized that the gut microbiome should be considered as a complex landscape, with both common and individual characteristics, and alternative stable states with respect to composition, structure and function [353]. They listed a top set of 50 bacterial genus-like taxa that are part of the phylogenetic core, a common core of bacterial taxa shared by the majority of (adult Western) human individuals, based on data from previous studies [353-355]. This core may include keystone species, whose role are crucial for ecosystem structure and function, for instance the breakdown carbon sources to support the growth of other core members [356, 357]. Mapping this core including its key stone species, and comparing it with diseased microbiota, could increase our understanding of a

normal microbiota and facilitate the design of a defined community representative of a healthy human gut microbiota. Next to the phylogenetic core, increased insight into the minimal intestinal metagenome [358] and the active functional core [359] within the human gut ecosystem might provide new criteria for assessing the ‘normality’ of a designed defined community. The paradigm seems to shift from rather black-box-like measures, such as the Midtvedt-criteria, to actually understanding the function of the gut microbiota and the contribution of its individual species. Subsequently, this approach could allow a more thorough comprehension and more accurate design of age-, region- and disease-specific defined communities.

Although this review primarily focusses on bacterial communities, it should be mentioned that the human (gut) microbiome also includes fungi, archaea, microeukaryotes and many viruses, mainly bacteriophages. A study from 1980 included a ‘yeast fungus’ in a defined hexaflora, but the specific role of this microbe was not addressed [360]. One of the few studies in this area addressed the interaction between the murine host, an archaeon (*Methanobrevibacter smithii*) and a bacterium (*B. thetaiotaomicron*) [361]. In addition, the same research group designed a gnotobiotic animal model with a simplified defined gut community to study phage-bacterial host dynamics [298]. In parallel with the healthy gut microbiome, researchers recently mapped the healthy gut phageome [362], but this field is still in its infancy. It is reasonable to assume that, with increasing insight into the role of non-bacterial gut microbes in host-microbe interactions, the design of defined microbial communities becomes more representative of the whole human gut microbiome.

### Critical Evaluation of Defined Communities *In Vivo*: The Host

Next to the discussion on the exact composition of the defined microbial community, the selection of the host animal to study host-microbe interactions is critical. Rodents are the most commonly used mammalian models in which defined communities have been applied. The suitability of rodents as model for the human host was extensively reviewed elsewhere [263] and goes beyond the scope of this review. In summary, murine intestines are anatomically, histologically, and physiologically very similar to human intestines, but size, metabolic rates and dietary habits differ largely, leading to qualitative and quantitative differences in microbial composition [263]. With respect to the gnotobiotic models discussed in this review, there are some additional discrepancies to be mentioned. The high value of using gnotobiotic animals as models of humans, *i.e.*, their known composition and controllability, seem to be weakened by poor control of host parameters known to influence the human gut microbiome, such as diet, genotype, sex, part of the gut studied, age and the immune system.

#### *Host parameters influencing the microbiota*

Diet is a complex and strong determinant of gut microbiota composition (reviewed in [363, 364]). The individual species levels were assessed of a ten-membered defined



community in mice fed with diets systematically varying in protein, fat, polysaccharides, and simple sugars, to develop a model to predict the variation in species abundance. Next, the model was validated with 48 random combinations and concentrations of four ingredients selected from a set of eight human baby foods. Approximately 60% of the variation in species abundance could be explained by the known concentrations of pureed foods [294]. This study exemplified the application of defined communities to systematically assess the response of individual gut members to various food components, which are, moreover, typical for the human diet. Clearly, a standardized diet of a laboratory animal is different from that of humans, which varies per region, season, individual taste and even per day. Some studies listed in **Table 1a-c** incorporated a previously developed prototypic “Western style” diet [280, 285, 292, 295, 299, 338], containing high amounts of saturated and unsaturated fats and carbohydrates commonly used as human food additives (i.e. sucrose, maltodextrin and/or corn starch). A lack of standardization in lab animal feeding protocols, however, has been emphasized previously for instance with respect to diet composition and texture [365] and indeed, diets used by studies discussed here are highly variable (**Table 1a-c**). Moreover, in ~40% of the studies, the diet was not clearly defined or not even reported, which is alarming given the large impact of diet on the gut microbiome.

The choice for mouse genotype also varied per study (**Table 1a-c**), although an effect of host genotype on microbiota composition was established within species [366-370]. These results were corroborated by studies with defined communities such as ASF [371] and SIHUMI(x) [317]. Additionally, colonization of different mouse strains with SIHUMI(x) demonstrated host-specific caecal levels of polyamines and SCFAs [284]. In mice associated with *B. longum* and *B. thetaiotaomicron*, host genetic background was found to affect the overall transcriptome of the latter bacterium, but not the expansion of the bacterial substrate range of this bacterium [337]. Obviously, defined communities allow the careful investigation of such host-dependent effects, but validation of host-microbe interactions in a wide range of host strains seems crucial before drawing conclusions and extrapolation to humans.

Although reports on the effect of gender have been contradictory [358, 369, 372-376] it might be a crucial determinant in gut microbiota composition and/or behaviour. In turn, commensal microbiota was shown to affect sex hormone levels [377, 378]. Sex differences in gut microbiota composition were, recently, comprehensively investigated in 89 common inbred mouse strains. After excluding confounding by host genetics, diet, age or cage effects, the researchers detected gender-specific differences in taxa abundances and diet responses. These differences could be partially explained by sex hormones [379]. Among the studies discussed here (**Table 1a-c**), one reported differences in metabolic profiles in urine and plasma between both sexes, but no explanation was put forward [292]. In an older study, male mice were found more susceptible to death after *C. botulinum* infection, which

could be explained by their coprophagic behaviour or a more general higher susceptibility to disease [289]. In contrast, other studies reported an absence of gender-specific effects on, for instance, levels of *Oxalobacter formigenes* colonized in ASF mice [273] or assembly of a synthetic microbiota [296]. Whereas some studies discussed here (**Table 1**) reported to have used a gender-mixed population, others included only one gender (n = 12 of 53 studies), in which male more often than female (nine vs. three) animals were used. Remarkably, the establishment of SIHUMI(x) was verified in both genders, whereas the effect of dietary fibre was tested in male and the effect of high-fat diet was investigated in female rats [280]. A similar discrepancy was found in a study that assessed the effect of five fermented milk product strains in human female twins, but male gnotobiotic animals. Although microbiota responses were more or less similar in both species [293], such a gender-mismatch may complicate translation. Lastly, not all studies clearly reported the gender used per experiment, and approximately half of the studies did not report animal gender at all. This too, may hinder data reproduction and, more importantly, translation.

Defined communities allow the quantitative comparison of microbial compositions along the GIT, within and between models. ASF-associated mice were used to quantitatively demonstrate that the microbiota of the colon is poorly reflected in faecal samples [348]. Relative abundance of species were also different between faeces (rectal swabs) and colon in pigs colonized with a defined microbiota [316]. In rats colonized with SIHUMI(x), however, bacterial concentrations of caecum, colon and faeces were similar [280]. Additionally, increases in relative abundances of mucin-degrading bacteria in caecum and colon upon switching to a fibre-free diet, were reflected in faeces [296]. In a mouse model associated with a 12-membered community, individual bacterial levels were also similar between faeces and caecum [299]. These conflicting results could be explained by various factors, including host, community composition and sampling time. Irrespective of the actual difference between GIT sites, it is disappointing that some other studies relied solely on faecal bacterial content. In a study applying a 92-membered community, for instance, not even half of the members could be detected in faeces. Other species may have established themselves in different regions of the gut, but this was beyond the scope of the paper [304]. Nevertheless, due the invasiveness of sampling, systematic studies comparing colonic and faecal bacterial content are lacking in humans as well [54, 364]. The variation in GIT sites looked at by the studies included in **Table 1a-c**, makes it hard to compare the colonization pattern of the defined communities to natural colonization. Apart from differences along the GIT, capturing the transversal heterogeneity within one compartment may be crucial for properly modelling and understanding host-microbe interactions, as discussed above.

The age at which animals are colonized was quite variable among the studies, including animals bred with the desired defined community as opposed to GF animals colonized with the community of interest to create a gnotobiotic animal model. In the





latter case, animals are inoculated at various time points among studies, whereas timing of microbial colonization was demonstrated to impact, amongst others, immune maturation [380, 381], mucosal homeostasis [382] and gut-brain axis communication in mice [383]. Moreover, as previously discussed, colonization time of animals in studies discussed here (**Table 1a-c**) was limited. Nevertheless, some studies confirmed the stability of their defined community of interest over time and even over generations, which should be sufficient to draw conclusions within a specific colonization time window. This does, however, not allow to infer any information on the long-term effects of colonization.

A last factor determining gut microbiota composition and behavior is the immune system, which in turn is influenced by, amongst others, aforementioned factors and the gut microbiota itself. Looking at the studies discussed here (**Table 1a-c**), several researchers investigated immunological parameters such as serum immunoglobulin levels and the presence of (subsets of) immunological cells in the gut. Nevertheless, due to the complexity of the immune system, it is hard to quantify and compare the model hosts used with respect to immunological parameters. The key findings on the interactions of gut microbiota members and their products with the immune system have been recently reviewed elsewhere [352]. The authors emphasized the value of minimal microbiomes and subsequent standardized (animal) models. Determining the effects of specific gut microbiota on the host, could help to identify host-microbe interactions that shape the immune system [352]. Most studies discussed in this review did not make a distinction between the contributions of each specific microbe to immunological effects observed.

The advantages and the levels of controllability of gnotobiotic research, as well as its pitfalls in practice, as outlined above, are summarized in **Table 2**.

### *Validation of in vivo models*

As emphasized earlier, differences exist between humans and animals, not only limited to their intestinal microbiota. In line with the question what a normal or healthy intestinal microbiota defines, one could ask: “When is the animal model sufficiently representative of the human situation?” With regard to the studies discussed here (**Table 1a-c**), diverse host criteria are applied. For the models exploiting a murine microbiota, validation is relatively easy. Most researchers aimed to normalize GF hosts to conventionally raised animals, thereby focusing on host parameters such as caecal size or weight [262, 275, 286, 287]. With respect humanized mice, validation is more complicated, but some studies made an effort. For instance, total bacterial numbers in feces and fecal SCFA levels between humans and SIHUMI(x) rats were evaluated, and a previously reported increase of Erysipelotrichaceae upon high-fat diet in humans was mirrored in SIHUMI(x) animals [280]. Other host parameters (e.g., immune system or other systemic parameters) were, however, not taken into account.

Similarly, validation was lacking in other studies applying SIHUMI(x), in which, moreover, mice were used instead of rats [281, 283-285].

A better example was recently described in a study in which the effect was tested of a fermented milk product in both humans and gnotobiotic mice humanized with a 15-membered microbiota. The proportional representation of the intestinal bacterial species and genes and metabolic changes upon introduction of the probiotic strains, were hardly different between mice and men, but the researchers also acknowledged the limitations of their gnotobiotic animal model with respect to translatability [293]. In most other studies (**Table 1**), control groups were limited to conventionally raised and GF animals or animals with a control treatment, for which translatability of the results to the human situation remains speculation.

### Defined Communities *In Vitro*

As opposed to *in vivo* models, the use of defined communities to study host-microbe interactions *in vitro* has been limited, so far, although the development of sophisticated *in vitro* model systems is advancing rapidly. In this section we discuss *in vitro* models in which defined communities have been applied or could be applied to study host-microbe interactions. A distinction is made between models focused on the microbiota (e.g., composition and characteristics), and those that were designed to realistically represent the human host *in vitro*. **Figure 1** summarizes all existing *in vitro* models of the human host and microbiota, illustrating how their interactions can be studied combining advanced *in vitro* cell based systems with defined communities. Ultimately, the goal is to combine best of both worlds.

#### *Modelling the intestinal microbiota in vitro*

The use of fermentation models has proven successful in modelling the intestinal microbiota *in vitro*, ranging from short-term batch incubations to multi-compartmental continuous systems. As discussed already, most defined communities applied *in vivo* (**Table 1a-c**) were constructed bottom-up, by selecting species based on their function, prevalence, or other criteria. Alternatively, communities can be composed top-down by inoculating GIT-mimicking chemostats with human faeces. Well-known examples of these chemostats, such as the MacFarlane/Gibson three-stage continuous culture system, (M-)SHIME, EnteroMix, Lacroix Model and TIM-2, have been extensively reviewed elsewhere [384-386]. The high reproducibility, stability, and complexity of bacterial communities cultured in chemostats [387, 388] has allowed the development and application of representative communities of the human intestinal microbiota *in vitro*. Most of these models, however, did not include a host component. The HMI™ module comprised a promising exception in which first, faeces from a healthy volunteer was fed into an adapted SHIME system, with fluid compartments mimicking the stomach, small intestine, and ascending colon. Subsequently, the SHIME-effluent was exposed to an artificial mucus layer, separated by a semi-permeable membrane



from a compartment containing Caco-2 cells. This module allowed the co-culture of bacteria with enterocytes up to 48 h [389], which is discussed in more detail below.

### *Modelling the host in vitro*

With respect to well-established defined communities, the probiotic cocktail VSL#3 and the faecal transplant substitute MET-1 have been tested on various human or animal intestinal cell lines (Caco-2, T84 and HT-29) (e.g., [390-392]). In most studies, however, the use of bacterial lysates or conditioned media was preferred over live bacteria (e.g., [325, 393-396]), because the – mainly anaerobic – gut bacteria cannot survive under the aerobic conditions needed for intestinal cell culture. In these 2D models, the interaction with the immune system or other tissues, cannot be studied. Although the direct effect of VSL#3 was tested on spleen and dendritic cells [397, 398], tissue-tissue interactions were lacking in these models. This problem can be (partially) solved in Transwell co-culture models, in which bacteria, mucosal immune cells and intestinal epithelial cells can be studied together [399]. A Transwell model with an apical anaerobic compartment enabled the co-culture of an anaerobe bacterium with an intestinal cell line to study host-microbe interactions [400]. Still, these cell lines lacked their tissue-specific context, including all major types of epithelial cells (e.g., Goblet cells, enterocytes, enter endocrine and Paneth cells) organized in crypts and villi. Moreover, as cell lines are tumor-derived, their epithelial characteristics are affected. These issues have been overcome by the development of gut organoids, self-organizing 3D epithelial structures derived from intestinal stem cells [401] or human pluripotent stem cells [402]. The use of organoids to study host-microbe interactions was reviewed elsewhere [403]. The closed structure of organoids, in which the lumen is sealed with epithelial cells and a mucus layer, may facilitate the establishment of hypoxia in the core lumen [403]. The anaerobic pathogen *C. difficile* survived up to 12 h within organoids, but luminal oxygen levels still ranged from 5-15%, which may be tolerated by specific strains of *C. difficile* only [404]. More recently, researchers developed an organ culture system for the mouse intestine, in which the stromal and hematopoietic components of the normal intestine were preserved *ex vivo*. The device supported the survival and growth of both anaerobic and aerobic microbiota, allowing the investigation of their effects on neuronal parameters [405]. The co-culture of defined microbial communities with human cells in Transwells, organoids or organ culture systems has been limited, probably owing to the static nature of these models. More advanced *in vitro* models to study host-microbe interactions have been developed (as recently reviewed in [385]) of which only a few have hitherto allowed the co-culture of multiple bacteria with intestinal cells or cell lines.

Organ-on-a-chip technology is an emerging concept within biomedical research, to replace conventional cell culture and animal testing. Organ-on-chips are microfluidic devices in which cells are cultured with organ-relevant spatiotemporal

chemical gradients and dynamical mechanical cues, thereby aiming to reconstitute the structural tissue arrangements and functional complexity of living organs *in vitro* [406]. Several gut-on-a-chips have already been developed [407-410], only one in which multiple intestinal bacteria were successfully cultured [410]. In this device, two channels simulating the gut lumen and a blood vessel are separated by a membrane coated with extracellular matrix and Caco-2 cells [410]. As opposed to cell monolayers and organoids, the gut-on-a-chip is a dynamic model: shear stress and gut peristalsis are mimicked by continuous medium flow and stretching/relaxing of the membrane, respectively. Interestingly, these environmental cues stimulated Caco-2 cells to undergo differentiation into four types of intestinal epithelial cells, organized in 3D villi-like structures [411]. Also, the successful incorporation of endothelial cells and peripheral blood mononuclear cells, was demonstrated [412]. The authors claimed the successful cultivation of a single bacterium ‘on chip’ (*Lactobacillus rhamnosus*) for more than one week [410] and the eight-membered VSL#3 for at least 96 h [412]. The viability of the probiotic bacteria was, however, solely based on imaging, and which species exactly succeeded in ‘colonizing’ the crypts, was not exactly determined. The growth of anaerobic bacteria in this device has not yet been reported. In contrast, another recent study reported the successful co-culture of strictly anaerobic bacteria, *B. caccae*, with *L. rhamnosus* and Caco-2 cells. In their microfluidic-based model mimicking the human gut, HuMiX, bacteria were grown in a separate, anoxic compartment [413]. Similarly, the HMI™ module allowed the investigation of bacteria for up to 48 h under microaerophilic conditions. FISH analysis revealed the presence of strict anaerobic bifidobacteria in the upper part of the mucus layer and the positioning of *F. prausnitzii* at the oxic-anoxic interphase [389]. In both the HuMiX and HMI™ module, however, a mucin-coated attachment membrane prevented direct or natural contact between host and microbe. Moreover, as opposed to the gut-on-a-chip, gut peristalsis was not mimicked and the formation of the main epithelial cell types or crypts, were not reported in these models [389, 413]. A promising development in gut-on-a-chip technology is the incorporation of 2D organoids, which grow in a plane rather than in clumps, in the chip device [414], combining the advantages of organoids (tissue differentiation) with those of gut-on-a-chip technology (controllable flow, mechanical cues and tissue-tissue interaction). To date, the cultivation of a defined intestinal microbiota in this device, has not yet been reported.

### *Validation of in vitro models*

In comparison with animal models, validation of *in vitro* models is even more challenging. The cellular processes studied in Transwells, organoids or gut-on-a-chips, cannot be readily validated in human subjects. On the other hand, however, such sophisticated *in vitro* models enable the investigation of processes that cannot be readily studied in humans, increasing our understanding of the molecular mechanisms



## CHAPTER 2

of certain bacterial compounds or products. Furthermore, they allow the elimination of potentially confounding factors present in *in vivo* models, such as the immune system. At the same time, this is also one of the major drawbacks of aforementioned *in vitro* models: as opposed to *in vivo* models, they lack a systemic component, whereas the impact of the gut microbiota on human health extends beyond the GIT. The emergence of organ-on-chip technologies has led to the concept of a 'human-on-a-chip' [415], but its implementation in research is still at an early stage. Nevertheless, the road to such a human-on-a-chip may be just as interesting. 'Rebuilding' the human body through assembly of its separate parts (lung-on-a-chip, gut-on-a-chip, kidney-on-a-chip, etc.), might increase our understanding of these building blocks and their contribution to the whole.

## CONCLUSIONS AND FUTURE OUTLOOK

Our understanding of the human gut microbiome has rapidly grown over the past decades, which has definitely supported the design of defined communities representative of the human gut microbiome. Whereas defined communities were initially aimed to normalize germ-free hosts to conventionalized mice, they could be a valuable tool to study host-microbe interactions, because of their controllability and traceability. For the same reasons, defined communities have a high potential for therapeutic application. In this review, however, we showed that these rationally designed consortia have been applied in *in vivo* models that are not entirely representative of the human host environment. Next to the obvious and frequently discussed differences between mice and men, we also discussed the power of gnotobiotic animals has been further undermined by poor control of the host parameters known to affect gut microbiota composition and behaviour.

Simultaneously with the increasing knowledge on the human gut microbiota, the implementation of more advanced *in vitro* models of the human gut is accelerating, with the development of stem-cell derived organoids and gut-on-a-chip approaches. Although the research is still in its infancy, these systems might partially replace the use of animal models. This development is beneficial not only for ethical and – on the long-term – financial reasons, but also from a scientific perspective. Human-inspired *in vitro* systems allow us to model and capture host-microbe interactions at a more fundamental and controlled level.

Both the design of defined communities and *in vitro* models of the gut have not yet reached their plateau. The former can be improved, via either bottom-up or top-down approaches. Key is to further expand our knowledge about the intestinal microbiome in health and disease, in which the NIH Human Microbiome Project and the European MetaHit project have played a crucial role [358, 416] (bottom-up). The characterization of gut microbiota and genome sequences facilitates the *in silico* prediction of host-microbe interactions through constraint-based genome-scale metabolic modelling [417] or other types of mathematical modelling [418] and, subsequently, the *in silico* design of representative defined communities (bottom-up). Further exploring our whole microbiome, including phages, fungi, and archaea, will revolutionize the design of microbial communities as well (bottom-up). Lastly, the increased ability to reproducibly culture the microorganisms in human faeces *in vitro* using well-established fermentation technologies [387, 419] may open the avenue to study human faeces-derived, functionally enriched defined communities at a more personalized level (top-down). In this way, both health- and disease-related microbiota can be easily reproduced. The same level of personalization can be obtained on the host side. For instance, the implementation of 2D organoids from patient-derived induced pluripotent stem cells in *in vitro* systems, such as the gut-on-a-chip with, can lead to highly personalized screening devices.

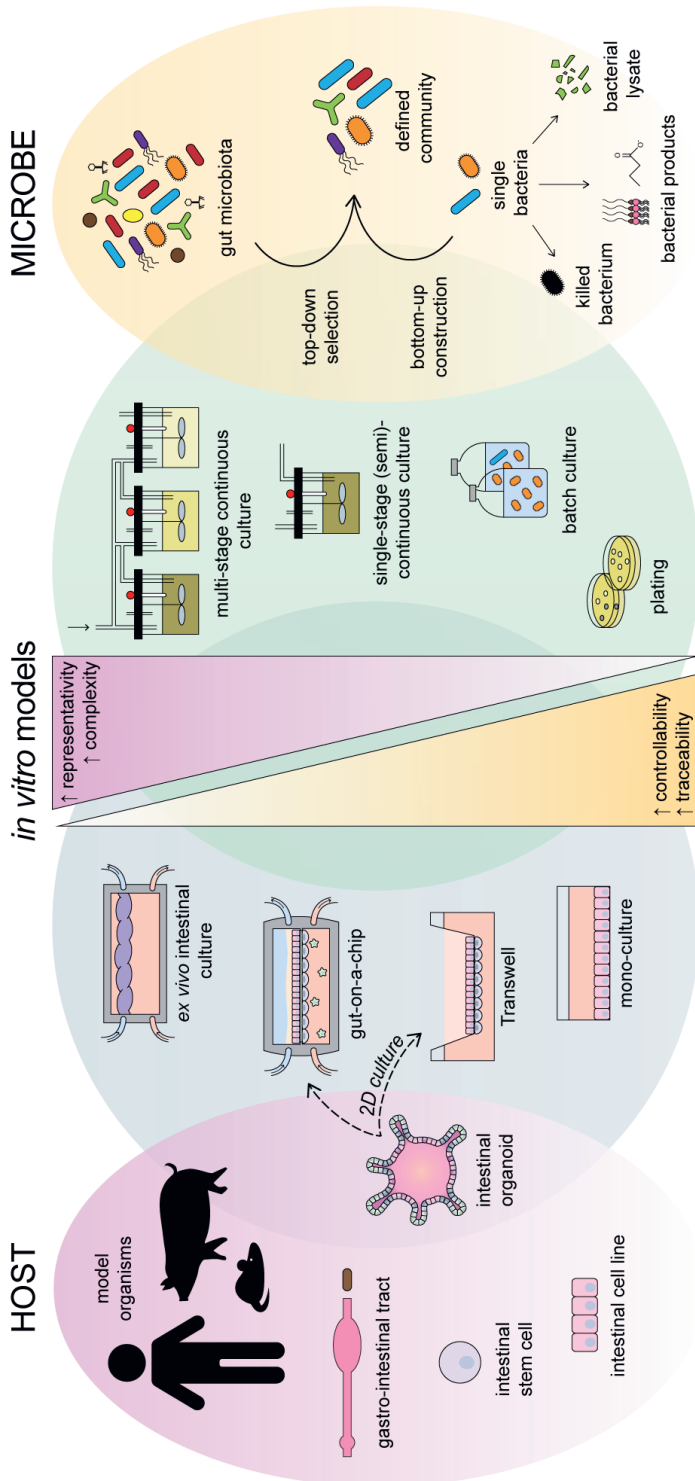


## CHAPTER 2

All in all, these models will provide a basis for the rational development and screening of novel therapies targeting intestinal diseases, ranging from anti-, pre- and probiotics to manipulate existing gut microbiota, to therapeutic microbes [420], faecal microbiota transplantation [421] and stool substitutes [330].

### ACKNOWLEDGEMENTS

This research was partly funded by the Netherlands Organisation for Scientific Research (NWO) in the framework of the [Building Blocks of Life](#) programme (737.016.003), the Gravitation grant (SIAM 024.002.002) and the National Roadmap for Large-Scale Research Facilities (NRGWI.obrug.2018.005). The authors declare no relevant conflicting financial interests.



**Figure 1.** *In vitro* models of the human gut and gut microbiota. Models are organized from bottom to top, with the most representative and complex at the top and the most controllable and traceable - with respect to host parameters or microbial species - at the bottom.





**Table 1a. Studies using defined communities to study host-microbe interactions in vivo: Specifically named communities (n = 30)**

The following study characteristics are listed: microbial consortium name (if applicable), taxonomic affiliation, strain source, host species and strain, part of the gut studied, no. of animals per experimental group, diet, sex, age, and study outcomes reported.

\* Two different strains tested are counted as one species. Strains were not always reported. Pathogenic species, in case of an infection model, are not included.

\*\*The colonization time includes the time from colonization (0 in case of transfer of microbiota to offspring) till and including the time of sacrifice or end of experimental (e.g., dietary) manipulations, in case this is clearly stated in the paper. If age is given and animals are colonized at birth, the age is included in colonization time.

\*\*\* Study outcomes are only reported for the animals colonized with the defined community of interest




Abbreviations: LP = lamina propria; MLN = mesenteric lymph nodes; MPO = myeloperoxidase; NR = not reported; SCFA = short-chain fatty acids; Treg = regulatory T-cell





Ref.	Name consortium (no. of species*)	Phylum division	Strain source	Host species (strain)	Part of the gut studied	No. of animals/group	Chow	Sex (M/F, both, or NR)	Age (col. time**)	Study outcomes***
[262]	<b>Schaedler flora (5)</b> 2 <i>Lactobacillus</i> sp., anaerobic streptococcus sp. (group N), <i>Bacteroides</i> strain, <i>Enterococcus</i> sp., coliform strain		Mouse	Mouse (NR)		20	NR	NR	4 wk (3 wk - 4 mo)	Colonization pattern; caecal size
<b>Altered Schaedler Flora (ASF)</b>										
[289]	<b>ASF (8)</b> ASF356: <i>Clostridium</i> species ASF360: <i>Lactobacillus</i> <i>intestinalis</i> or <i>acidophilus</i>		Mouse	Mouse (HA/ICR)		30	NR	Both	Adult (14- 56 d)	Death after <i>C. botulinum</i> infection; fecal <i>C.</i> <i>botulinum</i>

	ASF361: <i>Lactobacillus murinus</i> or <i>salivarius</i> ASF457: <i>Mucispirillum schaedleri</i> ASF492: <i>Eubacterium plexicaudatum</i> ASF500: <i>Pseudoflavonifractor</i> sp. ASF502: <i>Clostridium</i> sp. ASF519: <i>Parabacteroides distasonis</i>							
[422]	Mouse  Rat (F344)		1-5	Stress (Cholesterol, Rivet, Libitum)	M	NR (2 wk)	toxin excretion; colonization pattern of <i>C. botulinum</i> Hepatic genotoxicity mononitrotoluene isomers; metabolic activation of 2NT by intestinal bacteria; caecal bacterial content	
[269]	Mouse	Mouse (scid C.B-17)		4-6	Autodigestion, pellitete diet ad libitum	NR	NR (8-12 wk post reconstitution CD4+ T-cells)	+ (see row below)
	+: After <i>H. hepaticus</i> infection) necropsy; colonic inflammation score; grossly thickened colon, cecum, and rectum on		Rectal prolapse; clinically severe disease; grossly thickened colon, cecum, and rectum on necropsy; colonic inflammation score; histopathology					
[268]	Mouse	Rat (HLA-B27 on 33-3/F344)		7-11	NR	at least M	2 mo (1 mo)	Gross gut score, levels MPO and IL-1B in caecal tissue; histologic












CHAPTER 2

[423]	Mouse	Mouse (C3H/HeN)		4-8	Irradiated diet (Hartman Teklad)	NR	6-8 wk (9-14 wk)	After <i>H. bilis</i> /B. hydrophila colonization: Caecal pathological gross and histological scores; serum IgG1+IgG2a antibody response	Inflammatory score caecum and antrum
	Mouse	Mouse (C3H/HeN)		7-10	Irradiated diet (Hartman Teklad)	NR	6-8 wk (10 wk)	Faecal bacterial contents; (after <i>H. bilis</i> infection); caecal pathological scores; caecal histological changes; serum Ig	
	Mouse	Mouse (SW)		2-5	NR	NR	6-9 wk (NR)	Presence of Th17 cells and Foxp3+ regulatory cells in LP of small intestine	





[426]						Mouse (C57BL/6)	Mouse		NR	NR	NR	NR	Total intestinal IgA and intestinal IgA anti- CBir1; proliferation of splenic CBir1 TgT- cells after CBir1 gavage
[427]						Mouse (B6.Rag-/-)	Mouse		NR	NR	F	8-10 wk (10 d)	Homeostatic and spontaneous proliferation of TCR Tg T- cells in LP
[350]						Mouse (C57BL/6)	Mouse		5-8	Aut ocl ave d cho w	NR	8 wk (at least 3 dpi)	+ (see row below)
+After infection: <i>S. Typhimurium</i> levels in mesenteric lymph nodes, spleen, caecum, and feces; caecal pathology score; caecal microbiota density; bacterial content and microbiota complexity in feces													
[274]	ASF (8,9) 8: ASF 9: ASF + <i>Escherichia coli</i> HA108 or HA107	ASF (9) 	Mouse (C57BL/6)	3	NR	NR	Mouse		NR	NR	NR	NR	No. of IgA plasma cells per intestinal villus in duodenum, jejunum, ileum, and colon; IgA- bacterial binding in intestine;

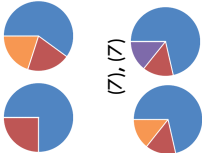



[428]	ASF (8)		Mouse	Mouse (NMRI, C57BL/6, BALB/c, NIH Swiss, SW, NMRI, MyD88-/-, Ticam1-/-, SMARTA, C57BL/6.C D45.1 +)		3-10	NR	NR	NR	anti- <i>E. coli</i> IgA titre + (see row below)
+: Caecal bacterial contents; colonic Treg cell response and relative IL-10 expression in spleen, MLN, Peyer's patches, colonic and small intestinal LP, thoracic duct lymph; IL-17 production; relative abundance of strains; microscopic localization in colon and small intestine										
[270]			Mouse	Mouse (Nod1 -/- and Nod2-/- on C57BL/6)		NR	NR	NR	6-9 wk (NR)	Caecal bacterial contents; + (row below)
+ intestinal tissue conductance and Cr-EDTA-flux; E-cadherin protein expression and RegIII-gamma mRNA expression in colon; survival, colitis disease severity, histology score and myeloperoxidase activity after DSS-induction; colonic IL-6, IL-10, MCP-1, IFN-γ, TNF-α, IL-12p70 levels										
[429]			Mouse	Mouse (C57BL/6)		NR	Aut ocl ave d food	Both	8-12 wk (8-12 wk)	RegIII-gamma RNA and protein expression ileum and colon
[430]			Mouse	Mouse (C57BL/6 and C57BL/6 TSLPR-/-)		3-5	NR	NR	NR (28 d)	+ (row below)



[431]	Expression of thymic stromal lymphopoietin mRNA in intestinal epithelial cell or colonic LP (LP); percentage of CD4+ T cells secreting IL-17A and IFN gamma in the colonic LP and MLN; expansion of colonic Treg cells in colonic LP and MLN; expression of receptor for TSLP by CD4+ and regulatory T-cells		4	NR	NR	3 d (3 d)	+
[334]	Structure of myenteric plexus, nerve density, average no. of HuC/D-positive myenteric neurons per ganglion, cell body size and average no. of nNOS-positive neurons per myenteric ganglion in duodenum, jejunum, and ileum; small intestinal motility (frequency and amplitude of muscle contractions) in duodenum, jejunum, and ileum before and after general neural or specific nitric blockade		5-14 per group	+	Both	6-12 wk (3 wk)	++ (row below)
[279]	+ Autoclaved mouse breeder's diet (Harlan), unlimited access; ++ Colonic histology, inflammatory (MPO) activity, enteropathy (presence of faecal albumin) and cytokine expression; faecal microbiota profiles; colonic gene expression; proportion of T-cell subtypes in colonic LP and other mucosal and systemic immune compartments		3 (ASF), 5-23 (Oligo-MM)	NR	Both	NR	ASF: Thickness colon total & inner mucus; mucus turnover time; Oligo-MM: alpha diversity colon & cecum
[273]	ASF (8) Oligo-MM <sup>12</sup> was also used, but no host parameters were assessed		4-7 per group	+	M (no gender effect observed)	3-9 mo (3-9 mo + 6 wk)	++ (row below)
[273]	ASF (8,9) 8: ASF 9: ASF + Oxalobacter formigenes		4-7 per group	+	M (no gender effect observed)	3-9 mo (3-9 mo + 6 wk)	++ (row below)



+ LW-485 autoclavable rodent diet, free access; ++ Bacterial levels in stomach, caecum, proximal colon and caecal mucosa; body weight; dietary oxalate intake; caecal and faecal oxalate levels; urine volume; urinary metabolite levels; caecal wet weight; caecal water metabolites										
Partial ASF										
[271]	Partial ASF (6) ASF 356, 361, 492, 502, 519 and 500 ASF 360 and 457 not colonized		Mouse	Mouse (NOD.MyD88KO)	None	9-23	NR	Both	NR (up to 30 wk)	Incidence of diabetes; histological scores of pancreatic islet destruction
[272]	Partial ASF (4,5) 4: ASF360, ASF361, ASF457, ASF519 5: 4 + Butyrivibrio fibrisolvens (type I, ATCC 19171 and type II, ATCC 51255)	(4)  (5) 	Mouse and bovine	Mouse (BALB/c)		4-5	+	NR	NR (2.5-5 mo after colorectal cancer induction)	++ (row below)
+ Autoclaved low-fiber diet (SSRZ, 1813680) or high-fiber diet (SSVL, 1813901) or tributyrin diet (5AVC 1814961); ++ Colorectal tumor multiplicity, tumor size and tumor grade; levels of LDHA, lactate, butyrate, H3ac and total H3 in colonic tissue and tumors; luminal SCFA levels; H3ac and expression levels of Fas, p21 and p27 genes in colonic tissue and tumors; apoptosis and cell proliferation levels in colonic tissue and tumors										

[275]	<b>Partial ASF (4,5,7,7)</b> 4: ASF356, ASF360, ASF361, ASF519, 5: ASF360, ASF361, ASF457, SB2 [ASF502], ASF519, 7: ASF356, ASF360, ASF361, ASF457, ASF500, SB2 [ASF502] and ASF519, 7: 4+E, coli M1B1, <i>Streptococcus danielae</i> ERD01G, <i>Staphylococcus xylo</i> us 33-ERD13C		Mouse	Mouse (C57BL/6)		4-6	NR	Both	0 or 8-12 wk (8-12 wk or 40 d)	Faecal bacterial content; bacterial load of <i>S. typhimurium</i> in feces, caecum, and MLN; relative caecal weight; functional genomic analysis of bacteria
-------	--	---	-------	-----------------	---	-----	----	------	--------------------------------	--

Oligo-MM

[275]	<b>Oligo-MM (12, 15, 17)</b> 12: Oligo-MM: 'Acutalibacter muris' KB18, <i>Flavonifractor plautii</i> YL31, <i>Clostridium clostridioforme</i> YL32, <i>Blautia coxoides</i> YL58, <i>Clostridium innocuum</i> I46, <i>Lactobacillus reuteri</i> I49, <i>Enterococcus faecalis</i> KB1, 'Bacteroides caecimuris' I48, 'Muribaculum intestinale' YL27, <i>Bifidobacterium</i>		Mouse	Mouse (C57BL/6)		4-6	NR	Both	0 (8-12 wk)	Faecal bacterial content; bacterial load of <i>S. typhimurium</i> in feces, caecum, and MLN; relative caecal weight; functional genomic
-------	--	--	-------	-----------------	---	-----	----	------	-------------	---

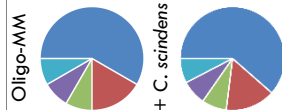




*longum* subsp. *animalis*  
YL12, *Turkionema muris*  
YL2, *Akkermansia*  
*muscipila* YL44, 15; 12  
+ 3 facultative  
anaerobes (*E. coli*  
M11B1, *Streptococcus*  
*danielae* ERD01G,  
*Staphylococcus xylosus*  
33-ERD13C), 17: 12 +  
5 ASF (ASF360,  
ASF361, ASF457, SB2  
ASF502), ASF519)

Oligo-MM (12,13)

**Oligo-MM (12,13)**  
12: Oligo-MM  
13: 12 + *Clostridium*  
*scindens* ATCC35704



Mouse

Mouse  
(C57BL/6)

5-8

NR

NR





0  
(6-12  
wk)analysis of  
bacteria


Faecal and  
caecal  
bacterial  
contents;  
caecal levels  
of lipocalin-  
2;  
calprotectin  
expression in  
caecal tissue;  
histopatholog  
y of caecum;  
caecal bile  
acid  
metabolome



SIHUMI (x)

[280]	<b>SIHUMI(x) (7,8)</b> <i>Anaerostipes caccae</i> DSM(Z) 14662 or 14667 <i>Bacteroides thetaiotaomicron</i> DSM(Z) 2079 <i>B. longum</i> NCC 2705 <i>Blautia producta</i> DSM(Z) 2950 <i>Clostridium ramosum</i> DSM(Z) 1402 <i>E. coli</i> K-12 MG1655 <i>Lactobacillus plantarum</i> DSM(Z) 20174 (x) <i>Clostridium butyricum</i> DSM(Z) 10702		Human	Rat (Sprague-Dawley)		3-21	+	Both	0-3 mo (2-38 wk)	Stability of microbiota in offspring; SCFA concentrations and pH in caecum, colon, and feces; bacterial counts in caecum, colon, and feces; Midtvedt criteria	
+ Sterilized standard chow (g/kg: 225 protein, 50 crude fat, 65 ash, 135 moisture, 480 N-free extract), fermentable-fiber-free diet, inulin diet, pectin diet, high- & low-fat diet											
[281]	<b>SIHUMI(x) (8,9)</b> 8: SIHUMI(x) 9: 8 + <i>A. muciniphila</i> ATCC BAA-835		Human	Mouse (C3H)		5-10	NR	NR	12 wk (5-15 d)	A) (see three rows below)	
[283]	<b>SIHUMI(x) (8,9)</b> 8: SIHUMI(x)		Human	Mouse (C3H/He Oul)		12	+	F	0 (8 wk)	B) (see two rows below)	

[284]	9: 8 + <i>Fusobacterium varium</i> ATCC 8501		Human	Mouse (Prm/Alt, C3H/He)		12-13	++	F	0 (56 ± 1 d)	C) (see row below)
<b>+ Irradiated standard chow R03-40;</b> ++ Sterilized pelleted standard chow R03-40; A) Bacterial cell numbers and proportions in caecum and colon; caecal and colonic histopathology score; expression of pro-inflammatory cytokines in caecal and colonic mucosa; serum protein levels of pro-inflammatory cytokines; cell number of <i>S. Typhimurium</i> in MLN and spleen; size of MLN; macrophage infiltration in caecal tissue; localization of <i>A. muciniphila</i> and <i>S. Typhimurium</i> ; mucin formation, mucus thickness, mucus composition and number of mucin-filled cells; B) Body weight; dry mass of caecum and colon; bacterial content caecum and colon; polyamine concentrations in caecum and colon; SCFA concentrations in caecum and colon; histology of caecum and distal colon (thickness of crypt depth, epithelial layer, mucosa, submucosa, muscularis externa); mitosis and apoptosis of caecal and distal colonic tissue; C) Length of small, large, and whole intestine; thickness of muscle, crypt and villi in proximal and small intestine and colon; faecal and caecal microbial content; caecal concentrations of SCFAs and polyamines										
[285]	<b>SIHUMI(x) (7.8)</b>  7: SIHUMI(x) without <i>C. ramosum</i>  8: SIHUMI(x)		Human	Mouse (C3H/He Oul)		3-9	Irra diat ed low -fat or hig h- fat diet ad libit um	M	0 (16 wk)	D) (see two rows below)

[282]	SIHUMI(x) (8,9) 8: SIHUMI(x) 9: 8 + <i>A. muciniphila</i> ATCC BAA-835	(9) 	Human	Mouse (C57BL/6 .129P2- Il10 <sup>tm1Cgn</sup> )		5-6	+ Row below	M	0 or 8 wk (3 wk)	E) (see row below)
<p>+ Irradiated standard chow (Altromin fortified type 1310; Altromin, Lage, Germany) ad libitum</p> <p>D) Body weight; body fat percentage; adipose tissue weight (epididymal, mesenteric and subcutaneous); energy intake; food efficiency; digestibility of high-fat diet; digestible energy; caecal and colonic bacterial content per species; blood glucose; leptin gene expression in epididymal tissue; liver weight; liver triglyceride levels; liver glycogen contents; expression of genes involved in lipid transport, lipid synthesis, cholesterol synthesis and lipid catabolism; gene expression of proteins involved in small intestinal glucose uptake; SCFA formation in caecum, colon and portal vein plasma; gene expression of SCFA-related proteins in colonic mucosa; gene expression of lipid transport and storage proteins in ileum; parameters of intestinal permeability and low-grade inflammation</p> <p>E) Body weight; histopathology score in submucosa, LP, surface epithelium, lumen; colon length; relative mRNA levels of <i>Tnfa</i>, <i>Ifng</i>, <i>Reg3g</i>; fecal lipocalin-2 concentration; fecal and caecal bacterial levels; caecal histology; number of goblet cells per 100 epithelial cells in caecum and colon; mucus layer thickness in colon; relative <i>Muc2</i> mRNA levels in distal small intestine, caecum, and colon</p>										



The following study characteristics are listed: microbial consortium name (if applicable), taxonomic affiliation, strain source, host species and strain, part of the gut studied, no. of animals per experimental group, diet, sex, age, and study outcomes reported.



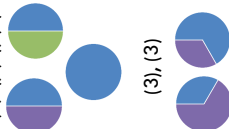

\* Two different strains tested are counted as one species. Strains were not always reported. Pathogenic species, in case of an infection model, are not included.

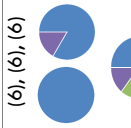










\*\*\*The colonization time includes the time from colonization (0 in case of transfer of microbiota to offspring) till and including the time of sacrifice or end of experimental manipulations, in case this is clearly stated in the paper. If age is given and animals are colonized at birth, the age is included in colonization time. (e.g., dietary






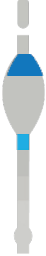


\*\*\* Study outcomes are only reported for the animals colonized with the defined community of interest

Abbreviations: LP = lamina propria; MLN = mesenteric lymph nodes; MPO = myeloperoxidase; NR = not reported; SCFA = short-chain fatty acids; Treg = regulatory T-cell




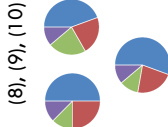



Ref.	Name consortium (no. of species*)	Phylum division	Strain source	Host species (strain)	Part of the gut studied	# of animals/group	Chow	Sex (M/F, both, or NR)	Age (col. time**)	Study outcomes***
[286]	<b>NA, F- and N-strains (2,9,11,41, 130)</b> 2: <i>E. coli</i> C25 + <i>Lactobacillus</i> , 9: 2 + enterococcus + <i>Lactobacillus</i> + <i>Candida</i> + 4 morphologically different strains of gram-negative anaerobes, 11: 9 + 2 strains of gram-negative anaerobes with fusiform morphology, 41: 11 + 30 additional strains of gram-	■ Firmicutes ■ Bacteroidetes ■ Actinobacteria ■ Proteobacteria ■ Verrucomicrobia ■ Other	Mouse	Mouse (CD-1)		4-57	Autoclaved Lobund diet L-356 or pelleted sterile diet from Charles River Mouse Farms	NR	NR (1 - 60 d)	Caecal number of <i>E. coli</i> C25; caecal size; histopathology of stomach, small intestine, caecum, and colon



	negative anaerobes, 130: 50 strains of gram-negative strict anaerobes (N) + 80 facultative anaerobes (F)	Not specified	Mouse	Mouse (CD-1)		5-7.5	Sterilized Lobound diet L-356, Charles River Formula 7RF, Lobound diet L-485 or Purina Breeder Chow	NR	NR (4 wk)	Caecal size; caecal levels of fatty acids; caecal levels of E. coli; pH of caecal contents
[287]	<b>N- and F-strains (60,96,96,97)</b> 60: N-strains + 14 facultative anaerobes + E. coli C25 96: F-strains + E. coli C25 or E. coli 40T or Shigella 97: F-strains + E. coli C25 + Shigella or E. coli 40T									
[432]	<b>na (4)</b> Lactobacillus sp. 1 and 2, Bacteroides sp., Streptococcus group N		Rat?	Rat (Sprague-Dawley)	None	2	Autoclaved standard diet (Ref 7) supplemented with caffeic acid	NR	NR	Urinary metabolites of caffeic acids
[291]	<b>na(2,2,2,2,3,3,4,5,6,6,6,8,8,9,13,15,17)</b> 2: Actinobacillus s3+Streptococcus s1, 2: Bacteroides s8+Actinobacillus s3, 2: Eubacterium S10+Micrococcus s6, 2: Clostridium C1+C2, 3: Bacteroides s8+Actinobacillus s3+E. coli s7, 3: Shigella flexneri+C5+C6, 4: C1-C4, 4: S. flexneri+C3-C5, 6:	(2,2), (2), (2)  (3), (3)	Human and mouse	Mouse (CD-1)		≥ 2	Sterilized commercial diet (Usine d'Alimentati on Rationnelle) ad libitum	Both	2-5 mo (4 wk after last inoculation)	Number of IgA plasmacytes in duodenum




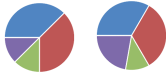

	C1-C6, 6: <i>Actinobacillus</i> s3+ <i>Streptococcus</i> s1+ <i>Lactobacillus</i> s4+ <i>Corynebacterium</i> s5+ <i>Micrococcus</i> s6+ <i>Streptococcus</i> s2, 6: <i>S. flexneri</i> +C5-C9, 8: 6 ( <i>Actinobacillus</i> , etc)+ <i>Bacteroides</i> s8+E. coli s7, 8: <i>S. flexneri</i> +C3-C9, 9: C1-C9, 13: C1-C13, 15: C1-15, 17: 8 ( <i>Actinobacillus</i> , etc)+C1-9	(6), (6), (6)  (8), (8)  (4,9,13,15),  (17) 	Mouse, rat, human	Mouse (C3H)		2-6	Autoclaved commercial diet	NR	Adult (up to 51 d)	Faecal bacterial counts; (mucosal) histology of stomach, jejunum, ileum, caecum, colon
[433]	na (2,2,2,2,2,2,2,3) 2,2: <i>Clostridium</i> E or P with <i>E. coli</i> K-12 2 (x 6): <i>Clostridium</i> E + <i>E. coli</i> S, <i>Proteus mirabilis</i> , <i>Klebsiella pneumoniae</i> , <i>Bacteroides (Alistipes)</i> <i>putredinis</i> , <i>Veillonella</i> <i>alcalescens</i> or <i>Clostridium</i> <i>perfringens</i> 3: <i>Clostridium</i> E and P + <i>E. coli</i> K-12	(2,2,2,2,2), (2)  (2,2), (3)  	Mouse, rat, human	Mouse (Balb/c)		Total of 3	Sterilized Ralston Purina 5010C	Both	0 (60- 90 d)	Caecal levels bacteria and <i>Candida albicans</i> ; histology tongue and stomach
[290]	UW-GL (9) Genera <i>Lactobacillus</i> , <i>Bacillus</i> , <i>Clostridium</i> and <i>Corynebacterium</i> . Species not defined.	UW-GL 	NR	Mouse (Balb/c)						



[360]	na (6) <i>Streptococcus (Enterococcus) faecalis</i> , <i>Lactobacillus brevis</i> , <i>Aerobacter aerogenes</i> , <i>Staphylococcus epidermidis</i> , <i>Bacteroides spurius</i> (?), a yeast fungus		NR	Mouse (Balb/c /ABOM f)	None		3-6	Sterilized food (two different procedures)	NR	0 (14 wk)	Serum levels of IgG1, IgG2, IgM and IgA
[289]	Partial or complete UW-GL (2,3,9) 2: <i>Lactobacillus</i> + <i>Clostridium</i> 3: 2 + <i>Bacillus</i> 9: UW-GL	Partial UW-GL 	NR and mouse	Mouse (HA/ICR)			10-48	NR	Both	Adult (14-56 d)	Death after <i>C. botulinum</i> infection; faecal toxin excretion; colonization pattern <i>C. bot</i>
[361]	na (2) <i>B. thetaiotaomicron</i> VPI-5482 + <i>Desulfovibrio piger</i> ATCC 29098		Human	Mouse (NMRI/KI)			4-5	Autoclaved polysacchari de-rich (B&K) ad libitum	M (sub et)	Adult or 12 wk (14-28 d)	Bacterial content of caecum and distal colon; bacterial gene expression; glycan levels in caecum; SCFA production in caecum; serum acetate; liver triglycerides; epididymal fat pad
[434]	na (2, 6, 10) 2: <i>Staphylococcus epidermidis</i> + <i>Veillonella parvula</i> 6, 10: anaerobic strains isolated from a conventional male mouse (not specified)	(2) 	Mouse	Mouse (B10.BR)			45-73	Sterilized ST1 (Institute of Physiology AS CR)	M	21 d (12 mo)	Occurrence of ankylosing enthesopathy of the ankle; colon histology; bacterial content in ileum and colon





[338]	na (2) <i>B. thetaotaomicron</i> + <i>Eubacterium rectale</i>		Human	Mouse (NMRI-KI)		4-5	Irradiated standard low-fat, plant polysaccharide-rich diet (diet 2018 from Harland Teklad) + (row below)	M	11 wk (14 d)	Bacterial gene expression; caecal colonization levels; fermentation efficiency in caecum; colonic gene expression; protein expression in caecum
+: or high-fat, "high-sugar" Western-type diet (Harlan Teklad 96132) or low-fat, "high-sugar" (Harland Teklad 03317)										
[292]	na (3,8,9,10) 3: <i>E. coli</i> HS, <i>B. vulgatus</i> DSM1447, <i>B. thetaotaomicron</i> DSM2079, 8: 3 + <i>B. longum</i> NCC2705, <i>Blautia hansenii</i> DSM20583, <i>C. scindens</i> DSM5676, <i>Eubacterium ventriosum</i> DSM3988, <i>Lactobacillus rhamnosus</i> NCC4007, 9: 8 + <i>Collinsella aerofaciens</i> DSM3979 (colonized most mice), 10: 9 + <i>Faecalibacterium prausnitzii</i> DSM17677 (not colonized)	(3)  (8), (9), (10) 	Human	Mouse (C3H/HeN)		15 in total	Sterile standard chow diet or high-fat diet ad libitum	Both	7 wk (70 d after 1st inoculation)	Faecal and caecal bacterial cell counts; body weight; metabolites in urine and plasma
[293]	na (15, 19) 15: <i>Bacteroides caccae</i> , <i>Bacteroides ovatus</i> , <i>B. thetaotaomicron</i> , <i>B. uniformis</i> , <i>B. vulgatus</i> , <i>Bacteroides WH2</i> , <i>C. scindens</i> , <i>Clostridium</i>	(15) 	Human	Mouse (C57Bl/6J)		5	Autoclaved low fat, plant polysaccharide-rich diet (B&K rat and mouse	M	6-8 wk (42 d)	Faecal and caecal bacterial content; bacterial gene expression; urinary metabolites

[435]	<p><i>spiroforme, C. aerofaciens, Dorea longicatena, E. rectale, F. prausnitzii, Parabacteroides distasonis, Ruminococcus obeum, R. torques</i> (strain info NA), 19: 15 + <i>Bifidobacterium animalis</i> subsp. <i>lactis</i> CNCM I-2494, <i>Lactobacillus delbrueckii</i> subsp. <i>bulgaricus</i> CNCM I-1632+CNCM I-1519, <i>Lactococcus lactis</i> subsp. <i>cremoris</i> CNCM I-1631, <i>Streptococcus thermophilus</i> CNCM I-1630</p>		<p>Human</p>	<p>Rat (F344)</p>		<p>6-16</p>	<p>Irradiated polysaccharide-rich diet (R03, SAFE)</p>	<p>M</p>	<p>&lt; 3 months (30 d after inoculation F. prausnitzii)</p>	<p>Host gene expression: colonic epithelium; SCFA caecal concentrations; oxidoreduction potential caecal contents; colonic crypt depth; total cells/crypt in colon; expression of differentiation proteins of secretory lineage (KLF-4, ChgA); MUC2 production in colonic epithelium; colonic mucin glycosylation</p>
-------	---	---	--------------	-------------------	---	-------------	--	----------	--	---

[339]	na (2) <i>B. thetaiotaomicron</i> VPI-5482 + <i>B. longum</i> NCC2705		Human	Mouse (SW)		3	Standard diet (Purine LabDiet 5K67)	NR	NR (10 d)	Faecal bacterial content; metabolites in feces and urine
[295]	na (2,8,9) 2: <i>B. thetaiotaomicron</i> + <i>D. piger</i> 8: <i>B. thetaiotaomicron</i> , <i>B. cacciae</i> , <i>B. ovatus</i> , <i>E. rectale</i> , <i>Marvinbryantia formatexigens</i> , <i>C. aerofaciens</i> , <i>E. coli</i> , <i>Clostridium symbiosum</i> 9: 8 + <i>D. piger</i>	(2)  (8), (9) 	Human	Mouse (NMRI)		4-20	Irradiated LF/high-plant polysaccharides or HF/high-simple sugars ad libitum or the HF/HS diet with modified sulfate concentrations (600-fold range) or HF/HS diet supplemented with chondroitin sulfate	M	7-8 wk (2 wk)	Faecal bacterial relative abundance; faecal metatranscriptome; gene expression of <i>D. piger</i> ; gene expression of mouse proximal colon; caecal metabolites

[298]	<p><b>na (14) + virus-like particles</b></p> <p><i>C. aerofaciens</i> ATCC 25986, <i>B. catcae</i> ATCC 43185, <i>B. ovatus</i> ATCC 8483, <i>B. thetaiotaomicron</i> VPI-5482+7330, <i>Bacteroides uniformis</i> ATCC 8492, <i>Bacteroides vulgatus</i> ATCC 8482, <i>Bacteroides cellulosilyticus</i> WH2, <i>Parabacteroides distasonis</i> ATCC 8503, <i>C. scindens</i> ATCC 35704, <i>C. symbiosum</i> ATCC 14940, <i>C. spiroforme</i> DSM 1552, <i>D. longicatena</i> DSM 13814, <i>E. rectale</i> ATCC 33656, <i>R. obeum</i> ATCC 29174</p>		Human	Mouse (C57BL/6J)		5	Autoclaved low-fat/high-plant polysaccharide diet (B&K) ad libitum	NR	8 wk (46 d)	Gut barrier and immune function; overall health status; body weight and adiposity; number of CD4+ and CD8+ T-cells in spleens and MLN; faecal bacterial content and viral abundance; genetic changes upon viral attack (phage resistance); bacterial content of proximal and distal small intestine, caecum, colon; prophase activation
-------	---	---	-------	------------------	---	---	--	----	-------------	---








[296]	<b>No name (14)</b> <i>B. ovatus</i> DSM 1896, <i>Bacteroides uniformis</i> DSM 8492, <i>B. thetaiotaomicron</i> DSM 2079, <i>B. caccae</i> DSM 19024, <i>Barnesiella</i> <i>intestinihominis</i> YIT11860, <i>Roseburia intestinalis</i> 14610, L1-82, <i>E. rectale</i> DSM 17629, A1-86, <i>F. prausnitzii</i> DSM 17677, A2-165, <i>Marvinbryantia formatigenes</i> DSM 14469, I-52, <i>C. symbiosum</i> DSM 934, <i>C. aerofaciens</i> DSM 3979, <i>E. coli</i> HS, <i>A. muciniphila</i> DSM 22959 <i>Muc, D. piger</i> ATC29098		Human	Mouse (SW)		Total of 51	Autoclaved standard fiber-rich (15% dietary fiber), fiber-free or prebiotic (addition of purified soluble glycans) ad libitum	Both	8-9 wk (54 d)	+ (row below)
	+ : Microbial composition in feces, caecum, colonic lumen and mucus layer; bacterial CAZyme expression in caecum; mucin specific transcript in <i>B. caccae</i> , <i>A. muciniphila</i> and <i>B. thetaiotaomicron</i> caecal microbial enzyme activity; levels of SCFA and organic acids; colonic mucus layer thickness; colonic expression of mucus-production related genes; number of goblet cells in colon; histopathology; body weight; faecal lipocalin; colon length; caecal transcriptome; after infection with <i>C. rodentium</i> : histological score of caecum and colon, area of inflamed tissue in caecum, survival, ascending and descending colon and rectum, adherent <i>C. rodentium</i> in colon									

**Table 1c. Studies using defined communities to study host-microbe interactions in vivo: Communities in non-rodents (n = 6).** The following study characteristics are listed: microbial consortium name (if applicable), taxonomic affiliation, strain source, host species and strain, part of the gut studied, no. of animals per experimental group, diet, sex, age, and study outcomes reported. \* Two different strains tested are counted as one species. Strains were not always reported. Pathogenic species, in case of an infection model, are not included. \*\*The colonization time includes the time from colonization (0 in case of transfer of microbiota to offspring) till and including the time of sacrifice or end of experimental (e.g., dietary) manipulations, in case this is clearly stated in the paper. If age is given and animals are colonized at birth, the age is included in colonization time.

\*\*\* Study outcomes are only reported for the animals colonized with the defined community of interest Abbreviations: LP = lamina propria; MLN = mesenteric lymph nodes; MPO = myeloperoxidase; NR = not reported; SCFA = short-chain fatty acids; Treg = regulatory T-cell

Ref.	Name consortium (no. of species*)	Phylum division	Strain source	Host species (strain)	Part of the gut studied	No. of animals/group	Chow	Sex (M/F, both, or NR)	Age (col. time**)	Study outcomes****
[305]	<b>Bristol (3,4), Modified ASF (6,7,7)</b> 3: <i>Lactobacillus amylovorus</i> DSM 16698T, <i>Clostridium glycolicum</i> and <i>Parabacteroides</i> sp. (ASF519),		Pig	Pig (commercial hybrid and Babrahma m)		2-6	Evaporated milk	NR	0-17 d (14-21 d after 1st inoculation)	Presence of bacteria and mean total bacterial content in proximal and distal jejunum, terminal ileum, caecum, and colon; serum immunoglobulin concentrations
			4: 3 + <i>R. intestinalis</i> , 6: <i>Clostridium</i> sp. (ASF356), <i>Lactobacillus</i> sp. (ASF360), <i>Lactobacillus animalis</i> (ASF361), <i>E. plexicaudatus</i> (ASF492), <i>Parabacteroides</i> sp. (ASF519) and <i>Propionibacterium</i> sp. (ASF519),							
[312]	<b>Bristol (3)</b>		Pig	Pig ((Great York x Pie) x 'Dalland' cross)		6	+ (row below w)	NR	Neonates (26-37 d)	Relative ORS1E1 expression in jejunum



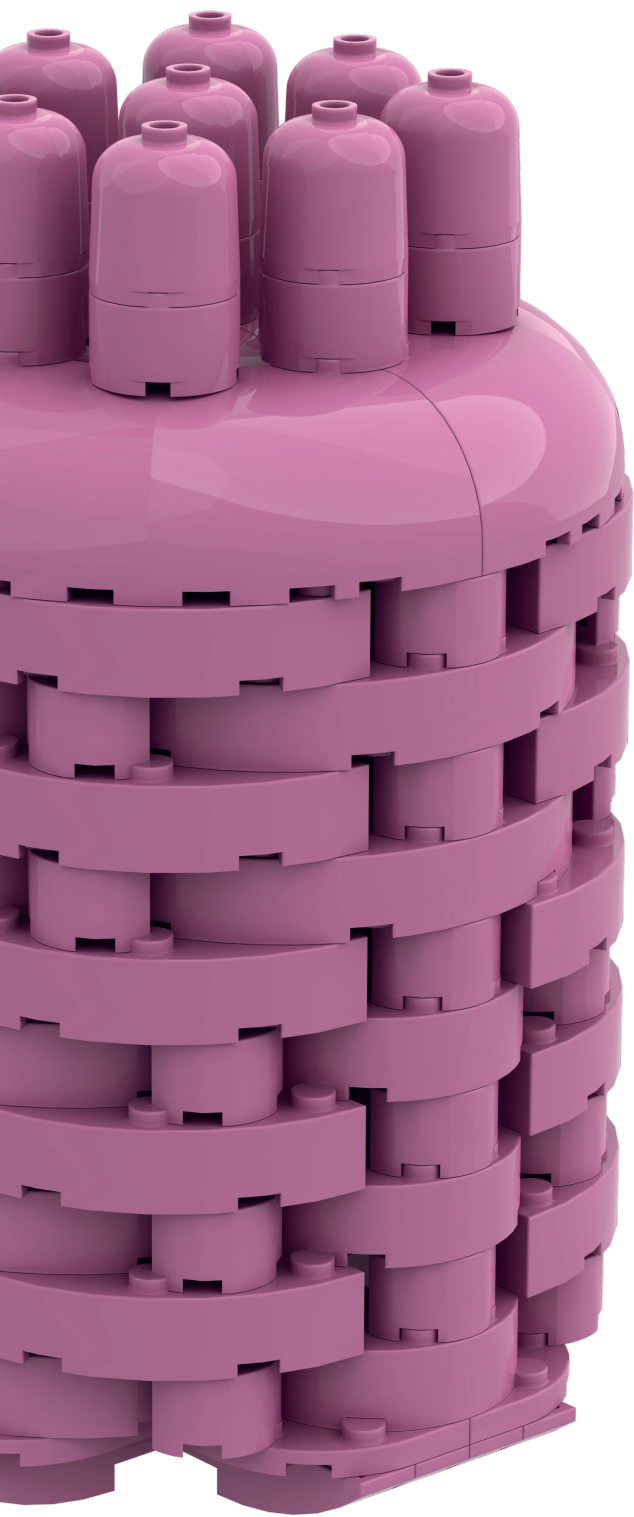
[313]	<b>Bristol (3)</b>		+ Pasteurized sow colostrum (first hrs), an <i>ad libitum</i> milk replacer diet, (day 0-4), a moist diet (remaining)						
			Pig (Great York x Pietrain) x 'Dalland' cross		6	+	NR	1 d (2-3 wk)	Oxyntic mucosa transcriptome
			+: Sow serum or pasteurized sow colostrum, followed by <i>ad libitum</i> milk replacer diet (day 0-4), followed by a control diet or medium chain fatty acid diet						
[315]	<b>DMF (7,8)</b> <i>Bifidobacterium</i> <i>adolescentis</i> , <i>B. longum</i> , <i>B.</i> <i>thetaiotaomicron</i> , <i>E.</i> <i>faecalis</i> , <i>L. brevis</i> , <i>S. bovis</i> and <i>C. clostridioforme</i> 8: DMF + <i>E. coli</i> Nissle	DMF  (8) 	Pig (Landrac e x Yorks hire x Duroc cross- bred)		3-6	NR	NR	7 d (35 d)	Faecal virus shedding; mean duration diarrhea; diarrhea severity; percentage of diarrhea; gene expression levels of CgA, MUC2, PCNA, SOX9, villin in jejunal intestinal epithelial cells
[316]			Pig (NR)		3-5	NR	NR	5 d (14- 35 d)	Bacterial content in rectum, duodenum, jejunum, ileum, colon, and feces/rectal swabs; diarrhoea and virus shedding after virulent human rotavirus challenge

**Table 2. Advantages and pitfalls of gnotobiotic animal models** in comparison with human research, with respect to the factors influencing intestinal microbiota composition or behaviour. Based on studies listed in **Table 1a-c** and literature.

Factor	Advantage (vs. human research)	Pitfalls in practice
Inoculum (defined community)	Controllable composition  <i>Healthy vs. diseased microbiota (e.g., missing key stone species), human- vs. animal-derived</i>	Animal microbiome $\neq$ human microbiome Difficulties in defining a healthy or normal microbiota Host-specific selection of microbiota
Diet	Controllable composition, timing, amount <i>Tailored to human diet (region, age, season, etc.)</i>	Lack of standardization in lab animal feeding protocols Not always reported (Table 1a-c)
Host genotype	Controllable – genetic changes possible  <i>Ability to introduce disease</i>	Validation of HMIs in multiple strains needed before extrapolation to humans Animal genotype $\neq$ human genotype
Sex	Controllable	Only one gender investigated (Table 1a-c) Not always reported (Table 1a-c)
Part of the gut	Ability to measure bacterial levels in virtually all intestinal parts Ability to capture transversal heterogeneity	Anatomy and physiology different from humans Variations in relative abundance per gut region different per model (Table 1a-c) Focus on specific gut regions or faeces only (Table 1a-c)
Colonization time	Controllable	Long-term effects not studied (Table 1a-c) Animals not always colonized starting at birth (Table 1a-c) Stability over generations not always confirmed (Table 1a-c)
Immune system	Controllable at start/birth	Uncontrollable in long-term studies, especially locally Complex, determined by in- and external factors Not quantified or quantifiable (Table 1a-c)







*“Garbage in,  
garbage out”*

# CHAPTER 3

## Systematic Comparison of Transcriptomes of Caco-2 Cells Cultured Under Different Cellular and Physiological Conditions

Janneke Elzinga<sup>1</sup>, Menno Grouls<sup>2</sup>, Guido J.E.J. Hooiveld<sup>3</sup>, Meike van der Zande<sup>4</sup>, Hauke Smidt<sup>1</sup>, Hans Bouwmeester<sup>2</sup>

1. Laboratory of Microbiology, Wageningen University & Research, Wageningen, The Netherlands
2. Division of Toxicology, Wageningen University & Research, Wageningen, The Netherlands
3. Nutrition, Metabolism and Genomics Group, Division of Human Nutrition and Health, Wageningen University & Research, Wageningen, The Netherlands
4. Wageningen Food Safety Research, Wageningen University & Research, Wageningen, The Netherlands

Published in

*Archives of Toxicology*, 2023, 97(3): 737-75

## Systematic Comparison of Transcriptomes of Caco-2 Cells Cultured Under Different Cellular and Physiological Conditions

### ABSTRACT

There is a need for standardized *in vitro* models emulating the functionalities of the human intestinal tract to study human intestinal health without the use of laboratory animals. The Caco-2 cell line is a well-accepted and highly characterized intestinal barrier model, which has been intensively used to study intestinal (drug) transport, host-microbe interactions and chemical or drug toxicity. This cell line has been cultured in different *in vitro* models, ranging from simple static to complex dynamic microfluidic models. We aimed to investigate the effect of these different *in vitro* experimental variables on gene expression. To this end, we systematically collected and extracted data from studies in which transcriptome analyses were performed on Caco-2 cells grown on permeable membranes. A collection of 13 studies comprising 100 samples revealed a weak association of experimental variables with overall as well as individual gene expression. This can be explained by the large heterogeneity in cell culture practice, or the lack of adequate reporting thereof, as suggested by our systematic analysis of experimental parameters not included in the main analysis. Given the rapidly increasing use of *in vitro* cell culture models, including more advanced (micro)fluidic models, our analysis reinforces the need for improved, standardized reporting protocols. Additionally, our systematic analysis serves as a template for future comparative studies on *in vitro* transcriptome and other experimental data.

## INTRODUCTION

After (partial) digestion of food and absorption of fluid, nutrients, and drugs in the upper part of the human gastrointestinal tract (GIT), the chyme reaches the colon, where fluid and electrolytes are (re)absorbed. Whereas the small intestine is most important for the uptake of food related chemicals and nutrients, the colon has other essential functions related to human health [436]. The colon abundantly contains microorganisms, estimated to reach a total of  $\sim 10^{12}$  microorganisms [437] which aid in the transformation of food components, e.g. yet undigested complex carbohydrates, to compounds such as short-chain fatty acids and vitamins, which contribute to host health [438]. Moreover, intestinal microorganisms have demonstrated a significant impact on drug transformation [41]. The role of the chemical exposure and human intestinal microbiota in various diseases, ranging from intestine-related diseases, including Inflammatory Bowel Disease [439-441], neuronal diseases such as Parkinson's and Alzheimer's [442-444], and cancer [44, 445, 446], to metabolic and psychological disorders [447-450], has sparked the interest in human intestinal health. In this context, there is an increasing need for *in vitro* and *in vivo* models reliably mimicking the human GIT, to investigate intestinal barrier integrity, host-microbe interactions and the toxicological effects of food related chemicals, food components, bacteria-derived metabolites, and drugs.

Murine and porcine *in vivo* models have been commonly used [451-453] to answer a wide range of scientific questions related to human intestinal health. Although those models allow for experiments in the context of the whole organism, they lack translational value in terms of human (intestinal) physiology [191, 454], anatomy [455, 456], and microbiology [457-459]. This further strengthens the existing ethical concerns about the use of animals for safety and efficacy testing of compounds of human interest [460]. More than ever, there is a call for increased insight in existing *in vitro* models mimicking the human intestinal tract as well as for improved *in vitro* models. This may not only help to refine protocols of dedicated animal studies, but also reduce the number of animals sacrificed for science. Eventually and more importantly, the use of *in vitro* models might partially replace the use of animal models [461, 462].

The immortalized cell line Caco-2 is a well-accepted and highly characterized model for the human intestinal epithelium. This cell line was originally isolated in the 1970s from a colorectal tumor [192]. As opposed to other isolated colon carcinoma cell lines [193], Caco-2 cells demonstrated spontaneous differentiation upon long-term culture leading to expression of several morphological and biochemical characteristics of small intestinal enterocytes [194, 195]. Growth and differentiation of Caco-2 cells on permeable membranes allows investigation of the transport properties of the cells [463], and this model has been extensively applied and reported in transport studies for toxicological or pharmaceutical research [463-



468]. Additionally, from the parental Caco-2 cell line, several clones have been generated over the years and selected based on characteristics of interest (reviewed in Sambuy *et al.* 2005) [469]. Besides Caco-2 cells, other human intestinal cell lines have also been commonly used as a model of the human intestinal epithelium, including T84 [196] and HT29(-MTX) cells [199, 470-473].

Despite the extensive use of the Caco-2 cell line in commonly used simple cell culture inserts (e.g. Transwell, ThinCert), its representativity of the human intestinal epithelium has been debated [474, 475]. In this respect, advanced *in vitro* techniques including the use of microfluidic devices [411, 476] and co-culture with other human cell types or (anaerobic) bacteria [477-479] have been applied to this cell line to mimic the intestinal tract more accurately in terms of physiology, cell differentiation, drug transport and/or host-microbe interactions [411, 476-479]. Simultaneously, primary epithelial cell cultures, including enteroids and adult and induced pluripotent stem cell derived intestinal models have been developed [219, 220, 480-483] and may be used as alternative to more complex *in vitro* intestinal models depending on the research question. Although these more complex models allow the development of personalized models of the human intestinal tract, the power of Caco-2 cells grown on permeable membranes lies in their culture simplicity, reproducibility, and the considerable number of studies available for comparison. Consequently, in theory, its widespread use should allow systematic comparison of the effects of different culturing parameters on the intestinal cellular response. Such a comparison would not only help to assess the reproducibility of *in vitro* models using Caco-2 cells, but also provide suggestions for adjustments of current *in vitro* techniques to improve their functionality and to better conform to the OECD Guidance Document on Good *In Vitro* Method Practices (GIVIMP) [484].

In this study, we aimed to compare cellular responses of different Caco-2 cell-based *in vitro* models based on gene expression, in which *in vitro* “model” is specified as “the physical and cellular conditions under which the cells are cultured”. We focused only on studies in which transcriptome analysis was performed on Caco-2 grown on permeable membranes, since this outcome allows for a comprehensive description of the cell response and serves as a starting point for investigating other outcomes. We collected published studies on Caco-2 cells cultured as cell layers in cell culture inserts or in devices, such as microfluidic chips, as well as studies with more biologically complex models in which Caco-2 cells were cultured as spheres, co-cultured with other cell types in a different compartment or exposed to human intestinal bacteria or their products. Based on the collected studies (2007-2021), we defined eight relevant experimental variables and utilized a bioinformatics approach to analyze and interpret the effect of these variables on transcriptomic responses. We followed an unbiased approach to explore the contribution of the defined variables to the overall transcriptome profiles. Subsequently, we zoomed in on specific genes and corresponding pathways and biological processes, of which

regulation of expression could to a significant extent be explained by one of the variables. Additionally, to complement transcriptome data, we carefully extracted other experimental parameters and evaluated several functional experimental outcomes. Of these, only Trans Epithelial Electrical Resistance (TEER) turned out to be commonly reported and thus was compared between respective studies. Overall, our study comprises a systematic and critical data analysis of *in vitro* models using Caco-2 cells grown on permeable membranes.



## MATERIALS & METHODS

### Study collection

A schematic overview of the study selection and the corresponding number of series can be found in **Figure 1**. The NCBI Gene Expression Omnibus (GEO) was searched for the term “Caco-2 OR Caco2” in May 2021, which resulted in 330 series with unique GEO Series identifier (GSEid) (**Online Resource 1**). Title, accession display and/or full-text paper of each series were manually screened to select for data series in which transcriptomic analysis was performed on Caco-2 cells which were cultured in cell culture inserts and on more advanced *in vitro* models with an apical and a basolateral compartment, including adapted inserts that introduce alterations such as flow or an anoxic compartment, gut-on-a-chips, and 3D spheric cell models. If in doubt, the full text of publication(s) linked to the series was screened. Series were excluded for which (as primary reason) a) no cell culture insert or advanced *in vitro* model was used; b) transcriptomics was performed on or including other cell types than Caco-2 cells (*i.e.* the mRNA would not only be derived from Caco-2 cells) c) no proper control condition was included (*i.e.* not commonly used medium); d) data had been taken from a previously deposited GEO DataSet and/or; e) study details could not be retrieved (*e.g.* studies were not published). For series of which the description in the database pointed towards the use of a cell culture insert or advanced *in vitro* model, but which were not linked to a publication in GEO, potential corresponding publications were actively searched using PubMed, Scopus, and Google. Additional databases (SRA-database from NCBI, as well as Array Express from EMBL-EBI) were searched using the same strategy but did not retrieve additional studies that were not already present in GEO. Lastly, we included data series from our own work, which had been submitted to NCBI but were only released after the date of the database search (GSE158620 and GSE173729). Next, array platforms on which <15,000 unique genes were analyzed, were excluded from further analysis (**Online Resource 1**). Series were divided per platform manufacturer, distinguishing between Affymetrix, Illumina, Agilent, and “others”. Because of the low representation of models in the latter three categories (thus resulting in a low statistical power), only studies performed on Affymetrix platforms were included for further analysis.

### Data extraction

From all selected data series, only data of samples encompassing Caco-2 cells grown under proper control conditions (*i.e.*, regular cell medium) or Caco-2 cells exposed to non-pathogenic intestinal bacteria, or their bacterial products were selected and downloaded. Additional data was manually extracted from full-text papers. Information on experimental set-up was extracted primarily from method-sections or from the Supplementary Information. Some papers referred to previous publications for the used experimental procedures, which then were assessed as well. In case the

experimental parameter of interest could not be retrieved, it was considered not reported (“NR”), except for culture temperature and atmosphere (assuming this was 37 °C at 5% CO<sub>2</sub>). Additionally, the seeding area of cell culture inserts was based on standard sizes, in case number of wells or other information was provided. Coating of membranes was considered not applicable (“NA”) if not reported, because this was not included in standardized Caco-2 insert protocols [467, 485]. Details for each study can be found in **Online Resource 1**. Data on TEER was manually extracted from the text and/or extracted from figures using a digital, on-screen ruler (Measura X, Gekar Tech). Values that had been normalized to a control were excluded. Each model identified was further categorized based on “GSEid” (GEO Series identifier), “Microbiome” (exposure to non-pathogenic bacteria or their bacterial compounds), “Culture Time” (in which short (< 9 days), medium (≥ 9, but ≤ 17 days) or long (> 17 days) were distinguished), “Oxygen” status in apical compartment (anoxic or oxic), “Flow” (static, dynamic or partially dynamic), “Cell System” (Caco-2 only or co-culture), “Device” (insert or chip) and “Platform” (type of array platform used). Partially dynamic refers to models where flow was applied for the majority of the entire culture time (i.e., > 80%) and/or only either to the apical or basolateral side of the cells. In our dataset, co-cultures consisted of Caco-2 cells cultured in same device with human leukemia monocytic cell line (THP-1), peripheral blood mononuclear or endothelial cells, of which mRNA was extracted from the Caco-2 cells separately.

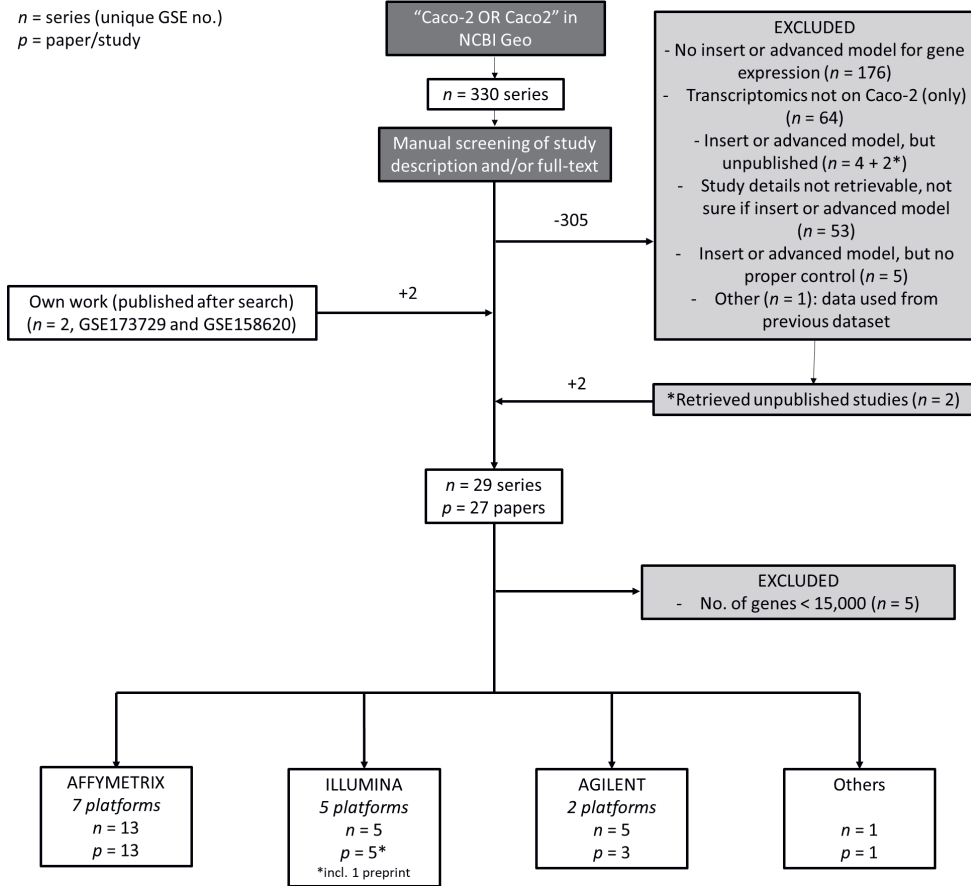
### Transcriptome analysis

All samples were integrated according to an established workflow described before [486]. Briefly, for each experiment raw data files were downloaded from GEO, which were then subjected to background correction and probe-to-probeset (gene) summarization according to the robust multiarray (RMA) algorithm [487]. Since samples were analysed on multiple Affymetrix array platforms, only those genes were kept that were probed for on all array platforms. This resulted in the inclusion of 11,203 unique shared genes. Per array resulting non-normalized expression estimates of these 11,203 genes were then transformed into rank percentile values, in which the gene with the highest expression estimate was assigned the value of 1 and the lowest expression estimate was set to 0. All expression estimates in between were given a value based on the ranking of expression i.e. 0.01, 0.02 etc. with the steps in between adjusted to the number of genes, and genes with the same expression level were given the same value based on the average rank if they were not tied (i.e. tied for the value of 0.01 would give both a value of 0.015 in case each step was 0.01) [486]. The dataset analysed in this study consisted of 100 samples, and from each sample expression data of 11,203 genes was extracted. On this shared transcriptome, three different analyses were performed. 1) A multi-level principal component analysis, in which array type was used as blocking variable, performed using the Bioconductor package PCAtools (version 2.6.0) [488]. Based on





the Elbow method [489] the relevant number of PCs were determined 2) The top 10% most variable genes were visualized in a heatmap using the package pheatmap (v1.0.12) 3) To quantify and interpret sources of variation the package variancePartition (version 1.26.0) was used [490]. This package uses a linear mixed model (LMM) to quantify variation in gene expression attributable to biological or technical variables. To fit the normality assumption of an LMM, the rank percentile values were first transformed using the probit function [486]. The genes of which the variance was explained for at least 40% by one of the variables were related to biologically meaningful changes using gene set overrepresentation analysis (ORA) applying a one-sided Fisher's exact test [491]. Gene sets were retrieved from the expert-curated Kyoto Encyclopedia of Genes and Genomes (KEGG) database [492] or Gene Ontology: Biological Processes (GOBP) [493, 494]. ORA was performed using the package clusterProfiler (v4.3.3) [495]. The whole code is available via <https://zenodo.org/record/7525836#.Y77gohWZO3A>



**Figure 1. Flow chart of study selection process.** The NCBI Geo Database was searched for "Caco-2" or "Caco2" in May 2021, which retrieved 330 data series. Exclusion of 306 data series and inclusion of two (own) data series resulted in a final selection of 29 series linked to 27 unique research papers. After exclusion of five data series because of limited numbers of analyzed genes, studies were divided per manufacturer, of which the Affymetrix platform comprised the largest group with thirteen data series corresponding to 13 research papers

## RESULTS

### Study collection pipeline retrieves 27 unique studies, of which 13 used Affymetrix platforms

Our search strategy to identify studies that performed transcriptome analysis on Caco-2 cells retrieved 330 GEO Series (*i.e.*, unique GSEids) (**Figure 1** and **Online Resource 1**). Next, 176 series were excluded as in these studies regular wells were used as these Caco-2 models lack the presence of a basolateral compartment, which limits the investigation of transport and (anoxic) host-microbe interactions [469, 496]. Other reasons to exclude series were the following: transcriptomics was performed on or including other cell types, *e.g.* immune cells and microbial cells not separated from Caco-2 ( $n = 64$ ); no proper control condition was included ( $n = 5$ ); data had been taken from a previously deposited dataset ( $n = 1$ ) or data had not been published (yet) and study details could not be retrieved ( $n = 59$ ). Of the latter category, the description of six series in NCBI pointed at the use of inserts or advanced *in vitro* models, of which two could be retrieved via other search strategies. We included two series of our own (NCBI-submitted) work (GSE158620 and GSE173729), resulting in a total number of unique series corresponding to 27 studies or papers. Across series, different array manufacturers and platforms were used. Five series had to be excluded because of a small number of unique genes analyzed by the platform used ( $g < 15,000$ ) (**Online Resource 1**). Affymetrix platforms comprised the largest group, including 13 series across seven different platforms. Illumina and Agilent platforms were used in five series each (five and two different platforms, *resp.*) and one series was analyzed on the Stanford SHCU platform ("Others"). For our analysis, we decided to only continue with the series analyzed on Affymetrix platforms, because this group comprised a wider range of *in vitro* models with multiple conditions per model.

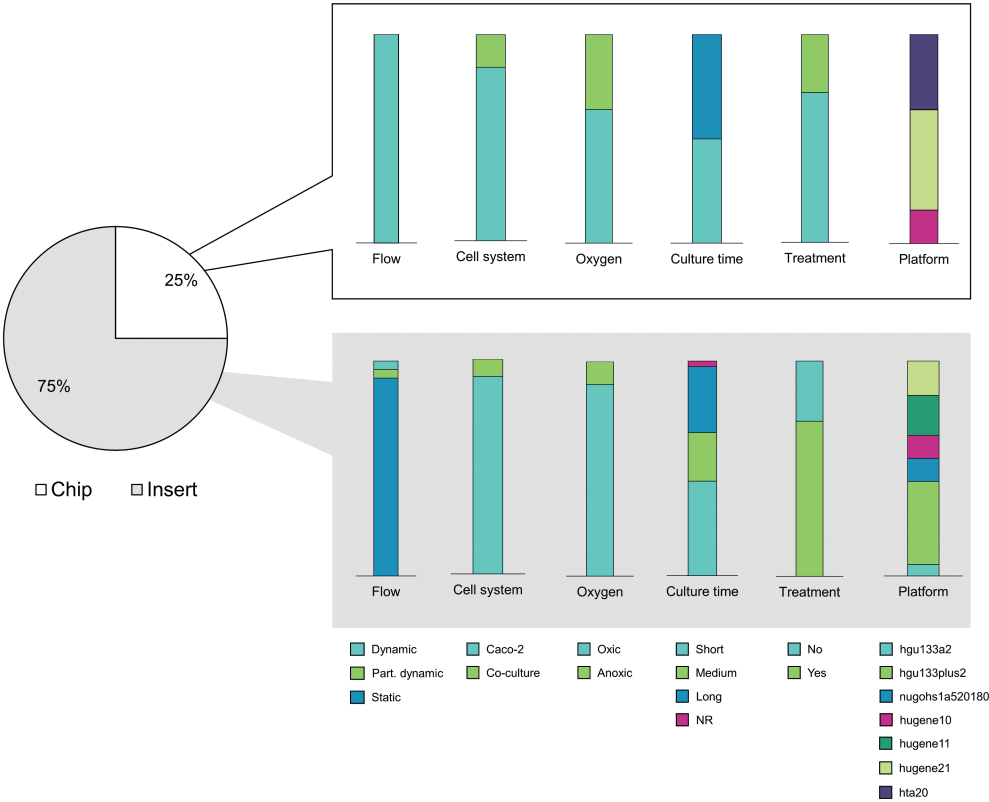
**Table 1. Overview of Affymetrix studies used for final transcriptome analysis.** Studies are presented in chronological order. Each row represents a separate model. # = No. of samples

GSEid	Ref	Platform	#	Device	Cell-system	Flow	Oxygen	Culture time	Microbiome? If yes, species and number
7259	[446]	Human Genome U133A 2.0 Array	4	Insert	Caco-2	Static	Oxic	Short; Medium	No
15636	[497]	Human Genome U133 Plus 2.0 Array	17	Insert	Caco-2	Static	Oxic	Short	<i>L. acidophilus</i> NCFM (Bacterium 1, strain B and SN), <i>B. lactis</i> (Bacterium 4 and SN), <i>L. salivarius</i> (Bacterium 5 and SN)
17625	[498]	Human Genome U133 Plus 2.0 Array	6	Insert	Co-culture	Static	Oxic	Medium	No
21976	[499]	NuGO array (human) NuGO_Hs1a520180	8	Insert	Caco-2	Static	Anoxic apical	Short	<i>B. bifidum</i> PRL2010 (Bacterium 7)
30292	[500]	Human Genome U133 Plus 2.0 Array	3	Insert	Caco-2	Static	Oxic	Long	No
30364	[501]	Human Genome U133 Plus 2.0 Array	3	Insert	Caco-2	Static	Oxic	Medium	No
65790	[479]	Human Gene 1.0 ST Array	2	Insert	Caco-2	Static	Oxic	NR	No
		Human Gene 1.0 ST Array	4	Chip	Co-culture	Dynamic	Oxic	Short	VSL#3 (Bacterial community 1)
79383	[476]	Human Transcriptome Array 2.0	9	Chip	Caco-2	Dynamic	Anoxic apical	Short	<i>L. rhamnosus</i> GG (Bacterium 7), LGG + <i>B. cacciae</i> (Bacterium 7 and 8), separated by PC membrane
81867	[502]	Human Gene 1.0 ST Array	3	Insert	Caco-2	Static	Oxic	Short	No
		Human Gene 1.0 ST Array	3	Insert	Caco-2	Part. Dynamic	Oxic	Short	No
115022	[503]	Human Gene 1.1 ST Array	8	Insert	Caco-2	Static	Oxic	Long	<i>L. acidophilus</i> W37 (Bacterium 1, strain A), <i>L. brevis</i> W63 (2), <i>L. casei</i> W56 (3)
156269	[504]	Human Gene 2.1 ST array	4	Insert	Caco-2	Static	Oxic	Long	No
158620	[505]	Human Gene 2.1 ST array	4	Chip	Caco-2	Dynamic	Oxic	Long	No
		Human Gene 2.1 ST Array	8	Chip	Caco-2	Dynamic	Oxic	Long	No
		Human Gene 2.1 ST Array	8	Insert	Caco-2	Static	Oxic	Long	No
173729	[470]	Human Gene 1.1 ST Array	3	Insert	Caco-2	Static	Oxic	Medium	No
		Human Gene 1.1 ST Array	3	Insert	Caco-2	Dynamic	Oxic	Medium	No



Data extraction results in 100 samples, of which 75% were derived from inserts

From the collected papers, we manually extracted the experimental set-up of each study (**Online Resource 1**). After selecting only samples encompassing Caco-2 cells grown under proper control conditions (*i.e.*, regular cell medium) or Caco-2 cells exposed to non-pathogenic intestinal bacteria or their bacterial products, we ended up with 100 different samples, including replicates (**Online Resource 1**). Of all 100 samples, 75% were derived from Caco-2 cells grown on inserts and 25% from Caco-2 grown on chips (**Figure 2**). Note that the model used in GSE8187 (a 96-wells insert with flow) was classified as insert, as opposed to the semantics used by the authors [502]. None of the studies that used an Affymetrix platform cultured Caco-2 as spheres. Across all 100 samples different culture times were applied, ranging from 4-21 days (**Online Resource 2** for full experimental set-up per study), which were further divided into short, medium, or long. One study on inserts did not report the timepoint of analysis. Samples were further categorized based on “GSEid”, “Microbiome”, “Oxygen”, “Flow”, “Cell system”, “Device” and “Platform” (**Table 1**).

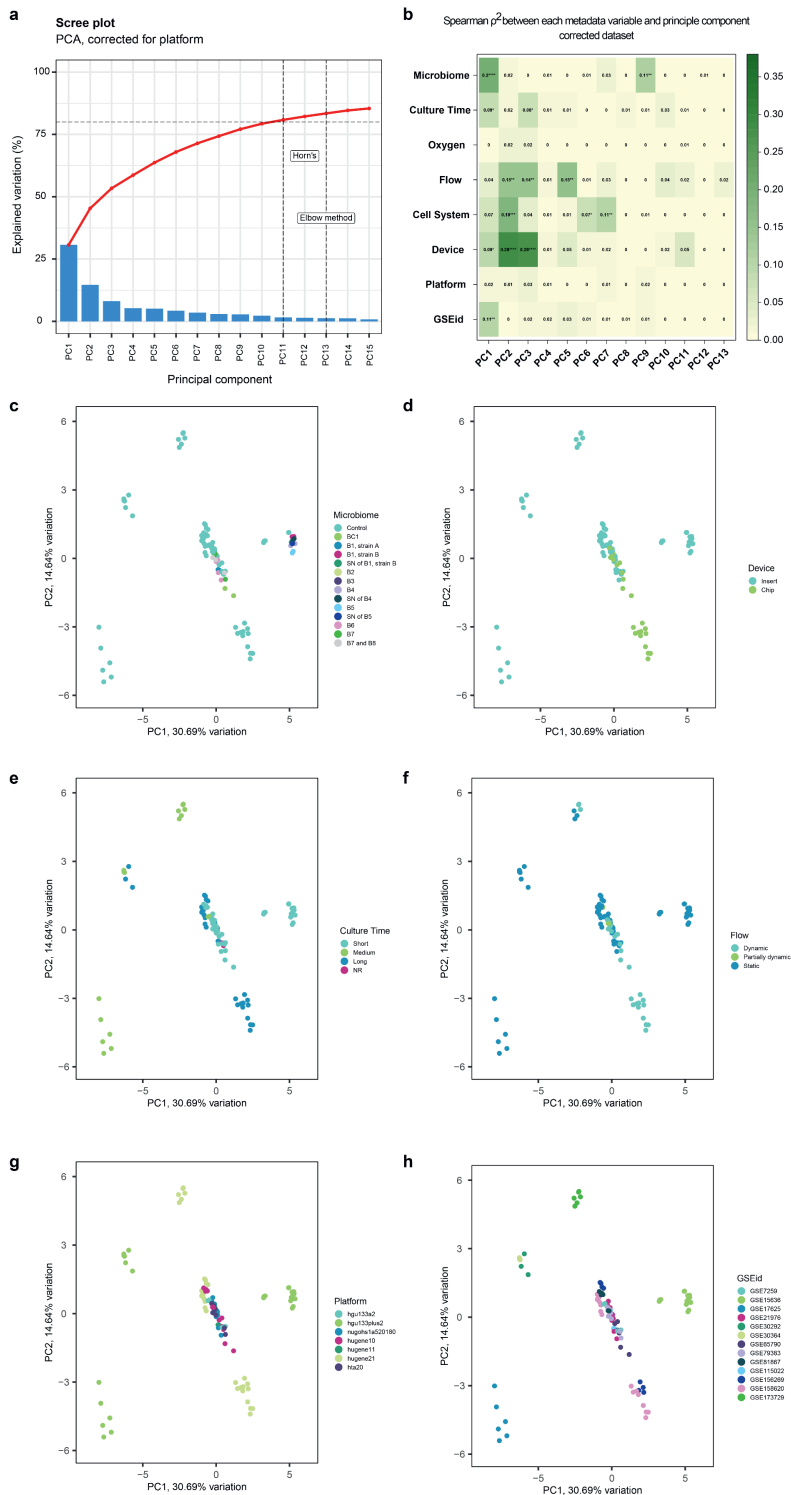


**Figure 2. Overview of samples per model.** Pie chart and bar chart represent no. of samples as percentage of total samples ( $s = 100$ ) analyzed on an Affymetrix platform. For simplicity, the variable “Microbiome” is only divided into “Yes” and “No”. NR = not reported; hta20 = Human Transcriptome Array 2.0; hugene21 = Human Gene 2.1 ST Array; hugene11 = Human Gene 1.1 ST Array; hugene10 = Human Gene 1.0 ST Array; nugohs1a520180 = NuGO array (human) NuGO\_Hs1a520180; hgu133plus2 = Human Genome U133 Plus 2.0 Array; hgu133a2 = Human Genome U133A 2.0 Array



**Figure 3 (next page) Multilevel principal component analysis at model level.** After correction for platform, a principal component analysis was performed on a total of 11,203 shared genes, distinguishing eight experimental variables. a) Scree Plot visualizing the variation explained by 18 PCs. b) Spearman correlation  $\rho^2$  per model variable for the first 13 PCs. PCA plots of the first two components are provided and labelled by c) microbiome (further defined in **Table 1**), d) culture time, e) device, f) model, g) array platform and h) GSEid. PCA plots of oxygen and cell system can be found in **Online Resource 3**. \*  $p < 0.05$ ; \*\*  $p < 0.01$ ; \*\*\*  $p < 0.001$ ; \*\*\*\*  $p < 0.0001$ . B# = Bacterium #; BC# = Bacterial Community #; SN = supernatant; NR = not reported; hta20 = Human Transcriptome Array 2.0; hugene21 = Human Gene 2.1 ST Array; hugene11 = Human Gene 1.1 ST Array; hugene10 = Human Gene 1.0 ST Array; nugohs1a520180 = NuGO array (human) NuGO\_Hs1a520180; hgu133plus2 = Human Genome U133 Plus 2.0 Array; hgu133a2 = Human Genome U133A 2.0 Array

## CHAPTER 3



### Multi-level PCA reveals weak correlation of experimental variables with shared transcriptome

First, we looked at the contribution of the eight pre-defined experimental variables to gene expression profiles, after controlling for the different Affymetrix array platforms. A multi-level PCA was performed (**Figure 3**) on the maximum number of genes shared by all array platforms ( $g = 11,203$  genes), referred to as the “shared transcriptome”. Based on the Elbow method [489], we only considered the first 13 principal components (PCs) (**Figure 3a-b, Online Resource 3**), of which PC1 accounted for 31% and PC2 explained 15% of the variation in the dataset. Overall, correlation coefficients were low, indicating weak to moderate correlation [506]. We report all variables separately below, starting with the highest correlations on PC1, 2 and 3.

“Microbiome” was the variable that contributed most significantly to the variation explained by PC1 ( $p = 0.20$ ,  $p < 0.0001$ , **Figure 3b**). Visualization of PC1 against PC2 demonstrated a separate cluster formed by three different bacterial species and their supernatant on PC1. These microbial exposures (See **Table 1** for specifications) were all applied within one study [497], but also clustered together with the control condition of that respective study (**Figure 3c**, see GSE15636 in **Figure 3h**). The variable “Device” contributed significantly to the variation explained by PC2 and PC3 ( $p^2 = 0.28$  and  $0.28$ , resp.,  $p < 0.0001$ , **Figure 3b**). Visualization of PC1 against PC2 demonstrated a grouping of the chips, while the inserts were less congruent (**Figure 3d**). Additionally, within the cluster of chips, a separation was observed between the short- and long-term cultured Caco-2 cells on chip (**Figure 3b and e**), whereas in general, the variable “Culture Time” contributed significantly, but relatively weakly to the variation explained by PC1 and PC3 ( $p^2 = 0.09$  and  $p^2 = 0.08$ ,  $p < 0.05$ ). Co-culture with other cell types contributed significantly to multiple PCs, mostly to PC2 ( $p^2 = 0.18$ ,  $p < 0.001$ ) (**Figure 3b, Online Resource 3**). The (partial) presence or absence of flow contributed significantly to the variation explained by multiple PCs, with a similar contribution to PC2, PC3 and PC5 ( $p^2 = 0.15$ ,  $0.14$  and  $0.15$  resp.,  $p < 0.01$ ). Visualization of the first two principal components, showed that the (partially) dynamic conditions clustered together, except for one study using Semi-Wet interface with Mechanical Stimulation (**Figure 3f**, see GSE173729 in (**Figure 3h**). Interestingly, none of the PCs were significantly associated with the variable “Oxygen” (**Figure 3b, Online Resource 3**). Array platform did not contribute significantly to any principal component (**Figure 3b and g**), demonstrating the successful correction of inter-platform differences (**Online Resource 3** for uncorrected data). Finally, GSEid contributed significantly but weakly to the variation explained by PC1 (**Figure 3b and 3h**).

Complementary to the multi-level PCA, a clustered heatmap was generated based on the 10% most variable genes of the shared transcriptome (**Figure 4**). This analysis did not show clear groupings according to one of the variables. The clustered



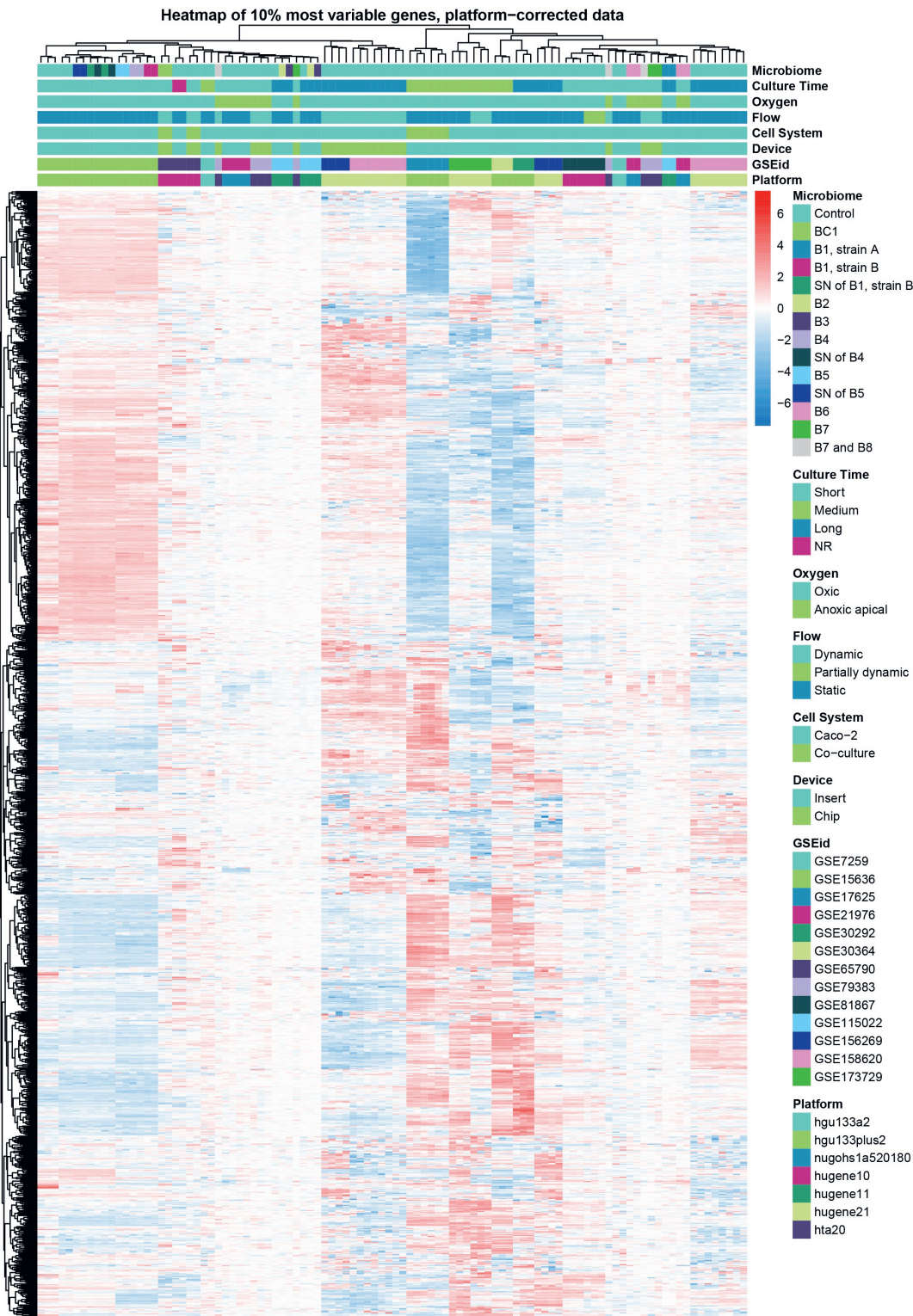


heatmap shows that the samples from two comparative studies between inserts and chips from Kulthong *et al.* (“GSEid”: GSE156269 and GSE158620) [504, 505] clustered based on the device, in line with what was found by PCA. Moreover, the clustered heatmap confirmed separate clustering of the long- from the short-term cultured cells on chip.

### Variance partition analysis reveals high contribution of “GSEid” to individual gene expression

Next, we focused specifically on the genes of which variance of expression was explained by one of the eight specified experimental variables, by variance partition analysis [490]. Without correction for “Platform”, the analysis revealed that the variance per gene was explained to the largest extent by “Platform” with an average of 43% across all genes ( $g = 11,203$ ) (**Online Resource 4**). After correction for “Platform”, the variance of genes was explained mostly by “GSEid” (average of 38%), followed by residual, undefined parameters (25%) and “Device” (8%, **Figure 5a**). Because of the relatively high contribution of “GSEid”, a technical parameter, we decided to focus only on genes of which variance was explained by one of the variables for more than 40%. The number of genes fulfilling this criterium varied from 0 (for “Oxygen”) to 5,198 (for “GSEid”) (**Figure 5b**). The variance in the relatively high number of genes explained by residual parameters, can be explained by other model parameters that we extracted from the respective studies, but could not be included in the variance partition analyses, for instance used cell passages, membrane on which cells were seeded, membrane pore size, seeding area and seeding density. Overall, these model parameters were heterogenous across and within *in vitro* models or were not reported at all by studies (**Online Resource 4**).

**Figure 4 (next page). Clustering of samples ( $s = 100$ ) based on the expression of the top 10% most variable genes shared between samples ( $g = 1122$ ).** Gene names are left out for readability. Heatmap represents relative gene expression varying from low (blue) to high (red). B# = Bacterium #; BC# = Bacterial Community #; SN = supernatant; hta20 = Human Transcriptome Array 2.0; hugene21 = Human Gene 2.1 ST Array; hugene11 = Human Gene 1.1 ST Array; hugene10 = Human Gene 1.0 ST Array; nugohs1a520180 = NuGO array (human) NuGO\_Hs1a520180; hgu133plus2 = Human Genome U133 Plus 2.0 Array; hgu133a2 = Human Genome U133A 2.0 Array

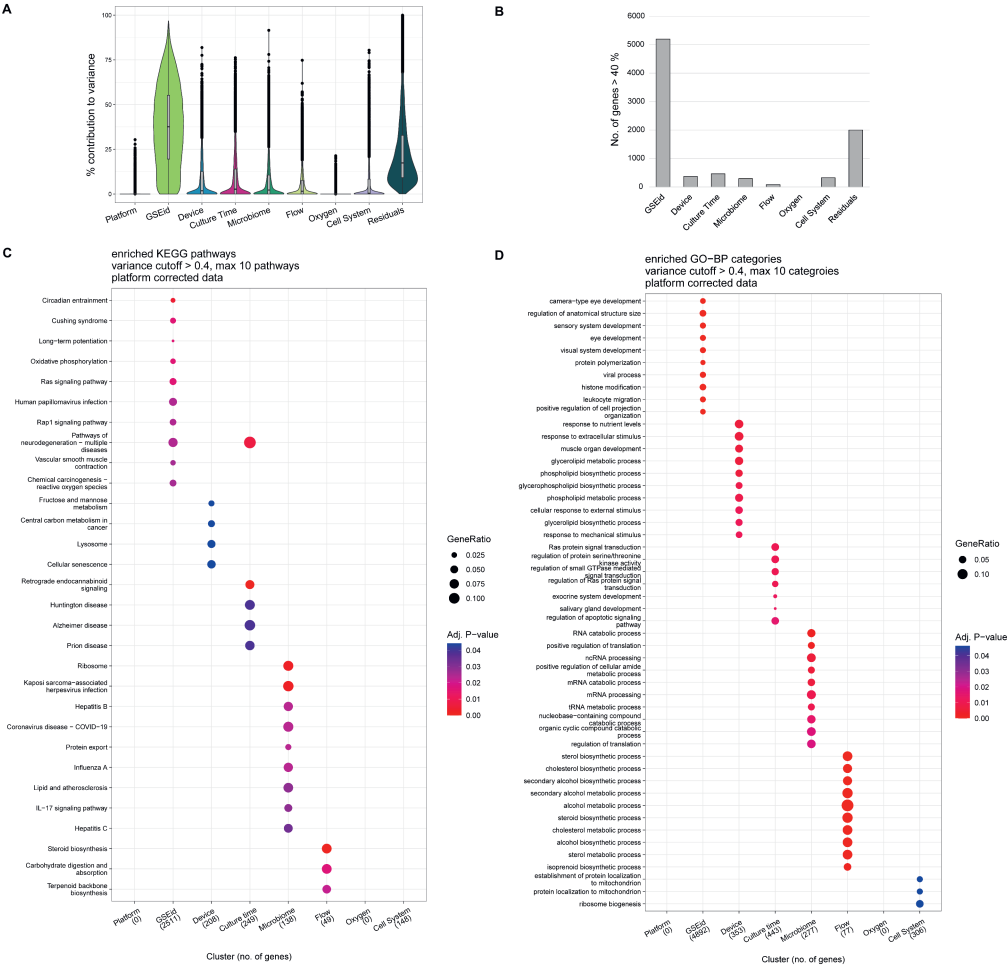


Next, we checked per variable which biological processes were enriched among the identified genes, *i.e.*, to which pathway(s) these genes were mapped more often than would be expected by chance. We used the KEGG and GOBP databases to retrieve the enriched pathways and biological processes, respectively. Only four variables resulted in significant overrepresentation of KEGG pathways, *i.e.*, "GSEid", "Culture Time", "Microbiome" and "Flow". Overrepresentation analysis of genes of which variance was explained for more than 40% by "GSEid", resulted in a heterogeneous mix of non-intestine related KEGG pathways, which was the same for "Culture time" (**Figure 5c**). Although the microbial exposures included in our dataset concerned non-pathogenic bacteria, overrepresentation analysis of the genes of which the variance was explained to a considerable extent by "Microbiome" revealed pathways related to host-pathogen interactions, such as Kaposi sarcoma-associated herpesvirus infection, COVID-19, hepatitis B and C, Influenza A and IL-17 signaling pathway. Genes of which the variance was explained mainly by "Flow" were enriched in the pathways involved in steroid and terpenoid backbone biosynthesis, but also in the breakdown and absorption of carbohydrates indicating an effect on energy homeostasis.

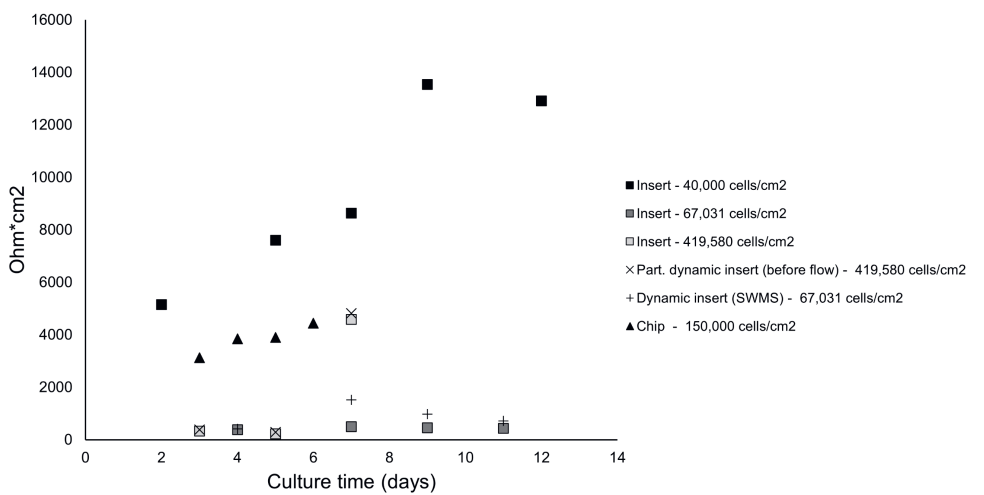
A similar approach was taken for Gene Ontology Biological Process (GO-BP) categories (**Figure 5d**), which resulted in overrepresentation of categories for six variables ("GSEid", "Device", "Culture time", "Microbiome", "Flow" and "Cell system"). Biological processes overrepresented in "GSEid" included a wide variety of processes, with some clearly unrelated to the intestine as they pertain to the development of other organs. Among the rest there was a focus on stress and adjustments in the cell via for example "histone modification" and "regulation of apoptotic signaling pathway" and two immune-related processes, "viral process" and "leukocyte migration". Among "Device", "Microbiome", and "Flow" a range of different metabolic and biosynthetic processes showed up, many related to lipid metabolism. Process overrepresentation analysis for "Culture time" and "Cell System" resulted in several processes unrelated to the intestine or expected effects.

### Comparison of TEER values reveals heterogeneity in values and reporting quality

Complementary to the transcriptome data, we collected additional functional experimental data from the identified studies to further characterize the used Caco-2 cell models. Among the identified studies, TEER was the only commonly reported outcome, which is a common measure of epithelial barrier integrity in *in vitro* studies using epithelial cell layers. TEER was reported in six Affymetrix studies comprising eight different models, of which one only reported the percentage of change in TEER, disabling comparison with other studies. Another study did not report seeding area, limiting calculations from Ohms/cm<sup>2</sup> to Ohms \* cm<sup>2</sup>. We extracted TEER values from six different models from four different studies across time points ranging from 2-12 days (**Figure 6**), demonstrating a wide range of values.



**Figure 5. Variance partition analysis of all genes shared between samples ( $s = 100$ ,  $g = 11,203$ ). a) Violin plot shown the percentage contribution of each variable to the expression of all genes. Based on the uncorrected data ([Online Resource 4](#)), a cut-off of 40% was chosen. b) Number of genes of which contribution of respective variable was more than 40%. The overrepresentation of pathways within these genes in c) KEGG Pathways and d) GO-BP are displayed per variable**



**Figure 6. TransEpithelial Electrical Resistance across models.** Only studies have been included which reported absolute TEER values ( $n = 4$ ). Model conditions are presented individually, including initial seeding density. All conditions are ox. Values were extracted from graphs using a digital ruler. Each data point represents the mean of 2-4 technical or biological replicates.

## DISCUSSION

In this study, we systematically compared transcriptomes of Caco-2 cells grown on permeable membranes using different *in vitro* systems modelling the human intestinal tract. We used the most frequently applied model, Caco-2 grown on inserts, as baseline for the comparison to other models incorporating an apical and basolateral compartment, allowing transport. Our shared transcriptome analysis indicated that of the studied parameters, the following had a significant, albeit relatively weak, influence on gene expression: the device in which Caco-2 cells were cultured; the presence of flow; and the exposure to non-pathogenic bacteria or their bacterial compounds. When looking at the individual gene level, however, variance in expression was mostly determined by the study (*i.e.*, GSEid). This points at a large heterogeneity in cell culture practices in *in vitro* models, which is supported by our analysis of experimental parameters from the respective studies (Online Resource 1), including variables such as passage number, Caco-2 sub-clone and protocol used, all of which can influence the results [469, 507-509]. Because these variables cannot be expressed as a number or concern an ambiguous range without a common starting point (*e.g.*, for cell culture passage), these could not be included in the analysis. Additionally, we demonstrated a lack of proper reporting of experimental variables, which has been stressed previously [484]. Our systematic analysis sets the scene for similar future analyses of (-omics) data between *in vitro* models, and therefore is a next step towards transparent reporting of *in vitro* studies to achieve an increased acceptance of non-laboratory animal study-derived data in life sciences.

Our multi-level PCA demonstrated that the variables “Flow” and “Device” had a similar effect. This is probably because dynamic samples were often derived from chips, except for three samples derived from inserts. These three concerned Caco-2 cells grown under Semi-Wet interface with Mechanical Stimulation, in which the cells are cultured under minimal liquid in the apical compartment and put on an orbital shaker [510]. Interestingly, visualization of PC1 against PC2 revealed that these samples clustered separately from the other dynamic samples (chips), but the same was true for the static control (insert) of this study [470], indicating that other factors explain the separate clustering of these samples. Overall, differences in (micro)fluidic design of included dynamic samples complicated inter-model comparison [511, 512]. Although most of the samples still clustered together, it would be informative to quantify shear stress or the resulting shear stress on the cells, as opposed to distinguishing only between static, dynamic and partially dynamic. Shear stress values were, however, not reported for all studies (except three [479, 504, 505] or could, in some cases, not be calculated with the available information.

“Microbiome” was relatively strongly associated with (shared) gene expression, despite the heterogeneity in bacterial treatments tested in the included studies. For microbiome-specific effects of the tested bacteria, we refer to the





individual papers corresponding to the studies (Species and corresponding references in **Table 1**). To facilitate co-culture of Caco-2 with intestinal (anaerobic) bacteria an *in vitro* model that is (at least) partially anoxic is required. Remarkably, our analyses demonstrated that the partial lack of oxygen did not seem to influence Caco-2 gene expression in the included *in vitro* models at all. For the studies that included microbes, the actual concentration of oxygen in the apical compartment was not always quantified. Therefore, the question remains whether the lack of association of oxygen with gene expression is due to failure of the respective studies to create an anoxic atmosphere in the apical compartment or that the absence of oxygen on one side simply does not affect gene expression of Caco-2 cells (e.g. because oxygen supply via the other side is sufficient). It would be advantageous for future studies to monitor and report O<sub>2</sub> concentrations, as has been done already in inserts [202] as well as more advanced *in vitro* models [477].

Because we observed a large heterogeneity in culture time between studies, we decided to define three groups based on cell culture time ranges. The definition of total culture time varied between studies. For instance, Dehal *et al.* accounted for the preconfluent phase of cells once seeded, by starting to count from day 2 after seeding [446], whereas other studies considered the time of seeding as the starting point. In the case of “Culture time” there is a consensus in literature that there is a direct relation with differentiation and thus, gene expression. Although the exact time until a plateau is reached differs for each differentiation marker, the generally accepted culture time for full differentiation is 21 days [467, 469, 485, 513]. In our dataset, total culture times differed between studies using inserts, which could depend on the study aim. For instance, for studying barrier properties, fifteen days of culture was shown to be sufficient [508]. However, we observed no strong influence of the variable “Culture time” on gene expression in our PCA. This could be due to the inclusion of chips and studies with adjusted protocols, both reported to change the relationship between culture time and differentiation [411, 514, 515]. Amongst others, the shear stress that cell experience on chips have been described to enhance differentiation of intestinal cell lines thereby reducing the total culturing time [411, 514]. This became apparent within the group of chips, where the short-cultured cells were separated from the long-term cultured cells in both the PCA (PC1 against PC2) and the heatmap cluster analysis. Within inserts the lack of association between “Culture time” and gene expression is likely due to the inclusion of studies such as GSE30292, in which cells were maintained at low density by subpassaging cells at 50% confluence, instead of the density prescribed by ATCC (between 80-90%) [500, 516]. Subsequently, these cells were cultured for a long period on inserts, e.g. three weeks showing a profound effect on gene expression over time. Interestingly, both our PCA and heatmap cluster analysis demonstrated that these samples clustered together with medium-term cultured cells, and not with other long-term cultured cells (all maintained at normal density before seeding). It was already shown previously

that low-density cells, although expressing the same level of differentiation markers as high-density (90%) grown cells, have a slower exit from cell cycle, including a delay in downregulation of cyclin A and increase of differentiation marker sucrase [515]. Although low-density grown cells should have differentiated to the same extent as high-density grown cells after three weeks [515, 516], our data suggests that low-density maintained cells grown for a long period on inserts demonstrate a transcriptome profile more similar to high-density maintained cells grown for a medium period on inserts. This reinforces the need for clear reporting on cell maintenance practice. Overall, the effects of, amongst others, flow and seeding density on the interaction between culture time and gene expression, have potentially obscured the expected effect of culture time on gene expression. For future analyses, these interactions could be included as a separate variable in the PCA. This, however, requires proper quantification of these interactions, which is now hampered by model heterogeneity and incomplete reporting.

Our variance partition analysis provided an overview of which pathways or biological processes were associated with each of the selected variables. In the case of the KEGG pathways only four variables contributed to pathways. The pathways associated with "Microbiome" were immune system-related, though the microorganisms included in our study are non-pathogenic bacteria, while the pathways are associated with infectious diseases. This is likely because both groups of microbiota act on Toll-like and other pattern recognition receptors [517, 518], but with probably opposing downstream effects. We did not find similar processes associated in the GOBP analysis, where "Microbiome" was mostly associated with RNA and translation processes, which are rather broad. The parameter "Flow" was only associated with two pathways, i.e., steroid and terpenoid backbone synthesis, and these were supported by the GOBP analysis, where "Flow" associated with many processes related to these pathways. Similar GOBPs were enriched in the genes mainly determined by "Device", which is likely due to the overlap between these two variables, as discussed previously. Both the variable "Culture time" and "Cell System" were associated with several RNA related processes but did not show a clear direction of effect in the other processes. Better harmonization of *in vitro* models or more accurate categorization of models (including more variables, see below), would probably point at more specific categories of pathways and processes, with a higher overlap between the KEGG and GOBP analysis.

The relatively low correlation coefficients in the PCA, as well as the high number of genes of which variance was explained to a substantial extent by residual variables, indicated that a substantial number of variables is still missing in our analysis. Future analyses would be stronger by including other experimental variables like passage number, seeding density, and seeding membrane. Our systematic extraction of these parameters revealed a large heterogeneity between studies, making it impossible to reach significant outcomes when these parameters are used





as input. We extracted other parameters as well, such as cell medium composition in terms of fetal calf serum, antibiotics, and other supplements; specific clones used and medium refreshment frequency, demonstrating similar heterogeneity. All these parameters have been reported to affect Caco-2 cell behavior [469], and therefore should be reported when publishing data, as also suggested by the MIAME- and MINSEQE guidelines (outlining the Minimum Information About a Microarray Experiment or Minimum Information About a Next-generation Sequencing Experiment that should be included when describing a microarray or sequencing study) [519, 520]. In general, the degree to which data was deposited in MIAME- or MINSEQE-compliant public data repositories, like NCBI Geo and ArrayExpress, limited the availability of our transcriptome-centered approach. For instance, we encountered studies that had not deposited their data in these repositories [514, 521] and vice versa, data series which had not been linked to the correct study (see number of retrieved studies **Figure 1**). We cannot determine how many relevant, unpublished studies we missed, since the description in NCBI on the exact *in vitro* model used was not always complete. In the context of the increasing global interest in Open Science, the importance of depositing open data in public repositories was recently stressed [522]. Specifically for microarray gene expression analyses, researchers demonstrated limited repeatability of published microarray studies, which was due to inadequate reporting on the used methods and other factors like software unavailability, or unclear reporting of the results [523]. The quality and ability to reuse data from other end points has been complicated by reporting issues as well. This is exemplified by TEER measurements, the most reported outcome in our dataset (other than gene expression). We concluded, however, that even the reporting quality of TEER was low, *i.e.*, in terms of number of replicates used; culture area; normalization to blank inserts and temperature at which the measurement was performed. Note that in a few studies [519] TEER was solely monitored as a quality measure of the monolayer, although the reported required minimum varied between studies (150-330  $\Omega \text{ cm}^2$ ) or was not defined. Similarly, a standardized method to measure TEER on chip devices has not been established yet [501, 503]. Our data reinforces the need for standardized TEER protocols and reporting guidelines for inserts, chips and other devices that are currently being developed.

All variables taken together, the data used for this study reiterates the need for a universally accepted Caco-2 cell culturing method, which also includes proper reporting of all variables and read-outs. The lack of adequate reporting is a commonly known problem in *in vitro* research [524], which limits the reproducibility and translatability of animal-free methods. Moreover, our study demonstrates that poor reporting quality of (meta)data also limits integration of existing *in vitro* data in systematic analyses across models or studies, re-emphasizing the need for Findable, Accessible, Interoperable, Reusable (FAIR) data guiding principles [525]. Additionally, the need to apply novel-approach methods in, for instance, chemical risk

assessment as well as in fundamental and clinical research is increasing. In this context, the development of an *in vitro* critical appraisal (IV-CAT) tool to improve the peer-review as well as the quality of published *in vitro* research, as proposed by De Vries and Whaley [526] is highly appreciated.

In summary, our study aimed to compare transcriptome responses of Caco-2 cells in different *in vitro* models as systematically as possible. Our analysis comprised both a transcriptome-wide and gene-specific approach, has the potential to find associations of predefined experimental variables with gene expression and uncover biological pathways associated with these variables. We complemented this analysis with manual extraction of other data, including model parameters and functional outcomes such as TEER. In this way, this paper can serve as an example for future comparison to *in vitro* models. Currently, controls are designed with only their own experiment in mind, failing to consider variables that are important to allow for comparison to other studies. More importantly, the results show the need for standardization and benchmarking of both current and future *in vitro* models. This should start with proper reporting of model parameters [527], as only in this way research can be reproduced and compared. We acknowledge that benchmarking of an *in vitro* model depends largely on the research question, e.g., whether the *in vitro* model is used for risk assessment, drug development or uncovering fundamental biological knowledge. Current approaches might still function in cases where you compare a potent exposure to controls across studies, but it fails to allow for extraction of more subtle effects of (underreported) model and experimental variables, reducing potential value of data. For current *in vitro* methods, the OECD is already working towards improving models via GIVIMP, setting standards on models and reporting [484]. At the same time, there is also a push to apply (part of) this knowledge in organ-on-chip technology in (to-be-) developed models [512, 528] by standardizing chip design. But application in the development of these models should go further than just technical design. A continued push in this direction is key and responsibility lies with not only the researchers that should execute experiments according to existing guidelines [524], but also the funding agents deciding to invest in the application or development of *in vitro* models as well as journal editors and peer reviewers critically evaluating the work. Only in this way, researchers will be able to unlock the full potential of *in vitro* models, to eventually reduce, refine, and replace the need for animal testing.

## ACKNOWLEDGEMENTS

For this work JE and MG were supported by a grant funded by the Dutch Research Council, a [Building Blocks of Life](#) project (No. 737.016.003). Work on this project by MvdZ was supported by the Dutch Ministry of Agriculture, Nature, and Food Quality (Grant: KB-37-002-020). The authors acknowledge the constructive comments on this work provided by Ivonne M. C. M. Rietjens and Clara Belzer.



### CONFLICT OF INTEREST

The authors declare that they have no conflict of interest.

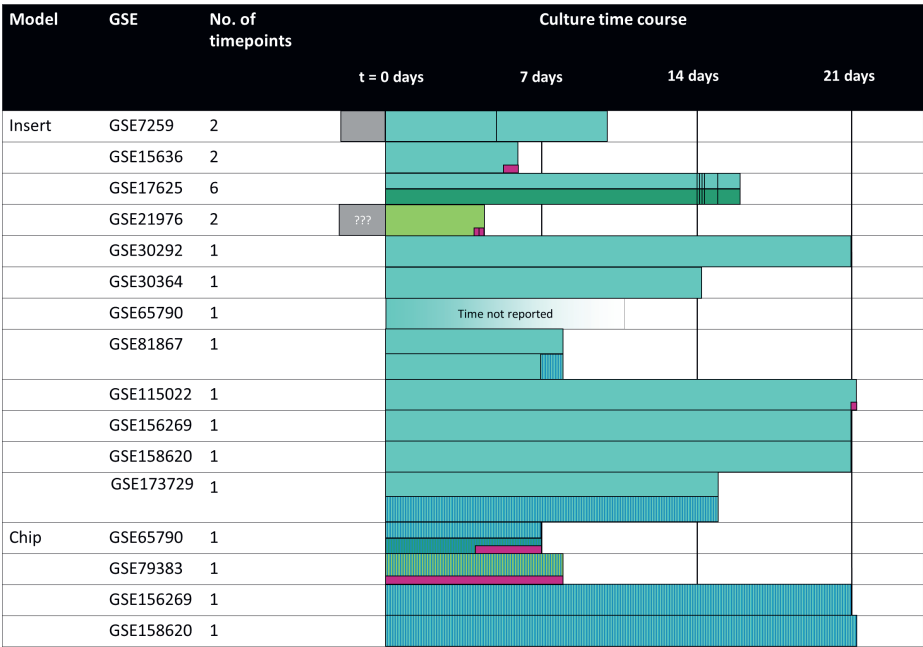
### DATA AVAILABILITY STATEMENT

The datasets analyzed during the current study have been generated by others and are available in the NCBI Gene Expression Omnibus (GEO), <https://www.ncbi.nlm.nih.gov/geo/>. The code to analyze the data is deposited online on Zenodo.org, <https://zenodo.org/record/7525836#.Y77gohWZO3A>.

SUPPLEMENTARY INFORMATION

**Online Resource 1** (Available online via <https://link.springer.com/article/10.1007/s00204-022-03430-y#Sec13>)

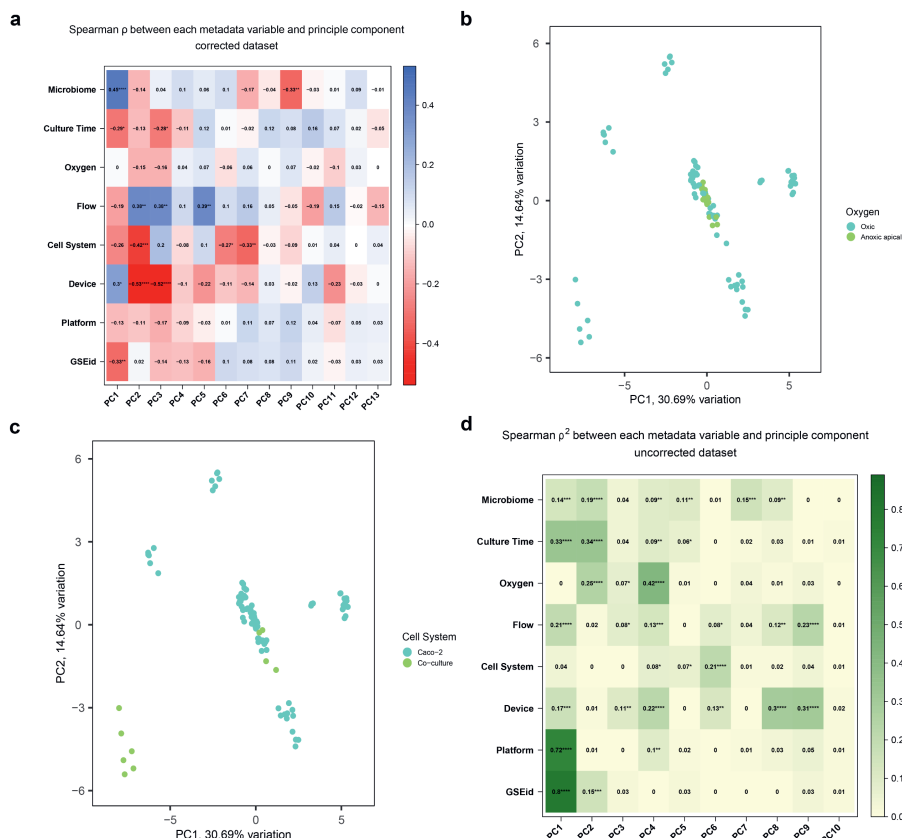
Summary of data regarding study collection and data extraction, including a) Retrieved 330 GEO Series (n = 330) in the GEO Database, including excluding criteria b) number of genes analyzed per array platform, including excluded platforms c) data extraction from all studies using Affymetrix platforms, including model specifications and experimental set-up d) sample overview of all samples (n = 100) analyzed on Affymetrix platforms



Legend: Grey = reported as pre-confluent, White = static vs. dynamic, Teal = oxic vs. anoxic apical compartment, Green = co-culture with (eukaryotic) cells, Pink = exposure to microbiota

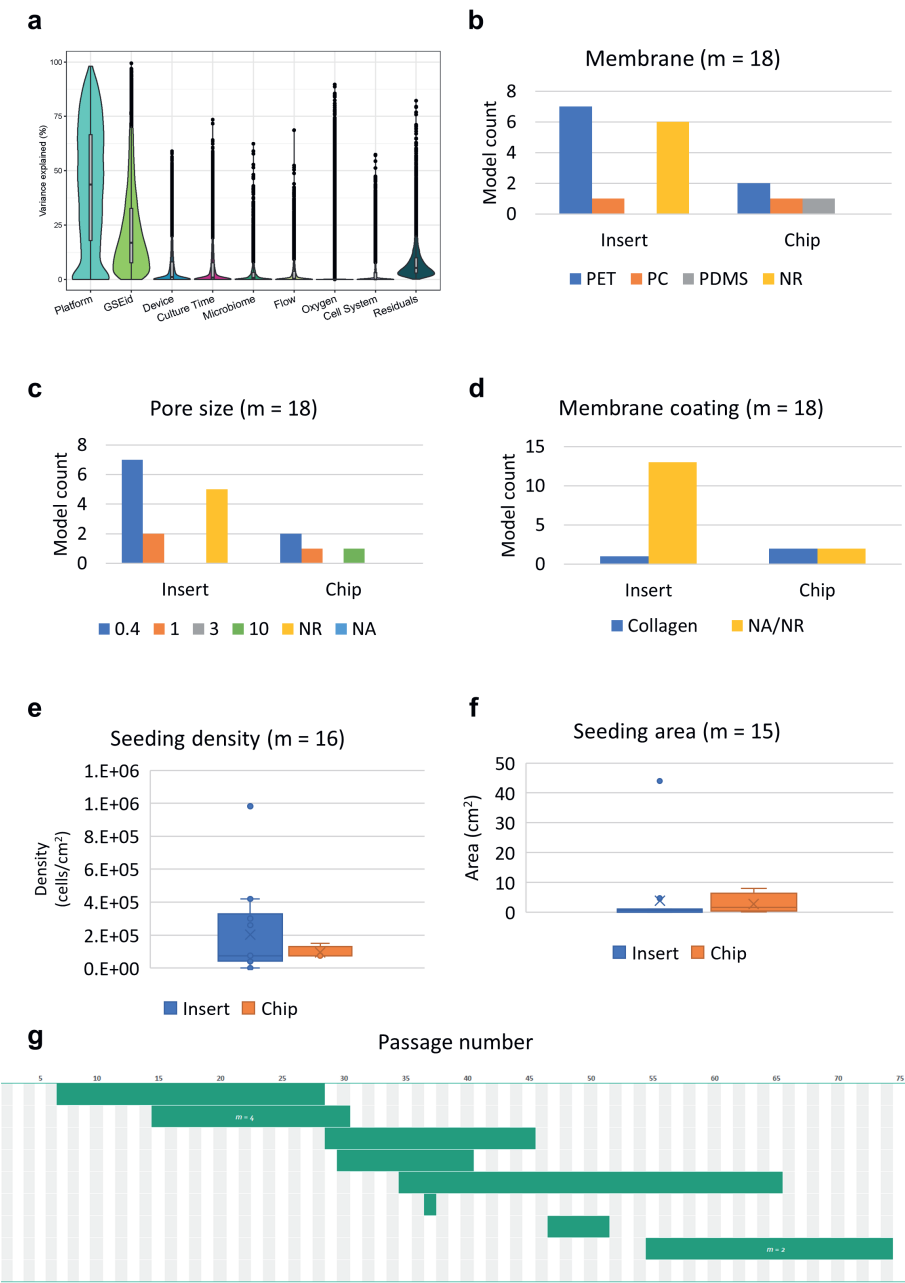
**Online Resource 2 Overview of experimental set-up per model analyzed on Affymetrix platforms**

## CHAPTER 3

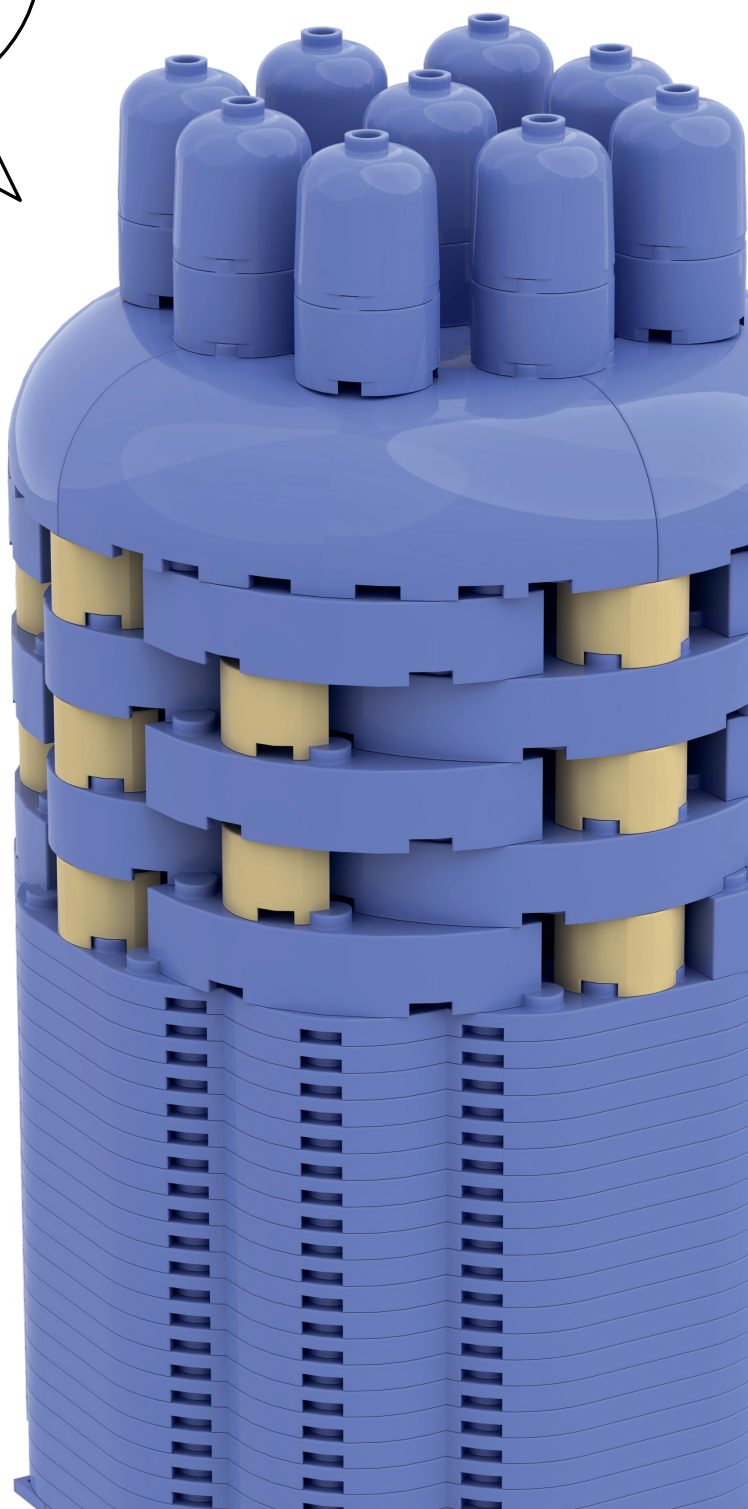


**Online Resource 3.** Multilevel principal component analysis at model level. After correction for platform, a principal component analysis was performed on a total of 11,203 shared genes, distinguishing eight experimental variables. a) Spearman correlation PC per model variable for the first 13 PCs. PCA plots of the first two components are provided and labelled by b) oxygen c) cells system. d) Spearman correlation  $\rho^2$  per model variable for the first 13 PCs without correction for platform. \*  $p < 0.05$ ; \*\*  $p < 0.01$ ; \*\*\*  $p < 0.001$ ; \*\*\*\*  $p < 0.0001$

**Online Resource 4 (next page).** Variance partition analysis of all genes shared between samples ( $s = 100$ ,  $g = 11,203$ ). **a**) Violin plot shown the percentage contribution of each variable to the expression of all genes without correction for platform. Additional data was extracted per model, including **b**) membrane on which the cells were seeded, **c**) pore size of the membrane, **d**) membrane coating, **e**) seeding density, **f**) seeding area and **g**) passage number use



Urgh! These  
scientists give  
me the shakes.



# CHAPTER 4

## Characterization of Increased Mucus Production of HT29-MTX-E12 Cells Grown Under Semi-Wet Interface With Mechanical Stimulation

Janneke Elzinga<sup>1\*¶</sup>, Benthe van der Lugt<sup>2¶</sup>, Clara Belzer<sup>1&</sup>, Wilma T. Steegenga<sup>2&</sup>

<sup>¶</sup>These authors contributed equally to this work

<sup>&</sup>These authors also contributed equally to this work

<sup>1</sup>Laboratory of Microbiology, Wageningen University and Research, Wageningen, The Netherlands

<sup>2</sup>Division of Human Nutrition and Health, Wageningen University and Research, Wageningen, The Netherlands

Published in

*PLoS ONE* (2021), 16(12): e02611915B5B



## Characterization of Increased Mucus Production of HT29-MTX-E12 Cells Grown Under Semi-Wet Interface With Mechanical Stimulation

### ABSTRACT

The intestinal mucus layer plays a crucial role in human health. To study intestinal mucus function and structure *in vitro*, the mucus-producing intestinal cell line HT29-MTX-E12 has been commonly used. However, this cell line produces only low amounts of the intestine-specific MUC2. It has been shown previously that HT29-MTX-E12 cells cultured under Semi-Wet interface with Mechanical Stimulation (SWMS) produced higher amounts of MUC2, concomitant with a thicker mucus layer, compared to cells cultured conventionally. However, it remains unknown which underlying pathways are involved. Therefore, we aimed to further explore the cellular processes underlying the increased MUC2 production by HT29-MTX-E12 cells grown under SWMS conditions. Cells grown on Transwell membranes for 14 days under static and SWMS conditions (after cell seeding and attachment) were subjected to transcriptome analysis to investigate underlying molecular pathways at gene expression level. Caco-2 and LS174T cell lines were included as references. We characterized how SWMS conditions affected HT29-MTX-E12 cells in terms of epithelial barrier integrity, by measuring transepithelial electrical resistance, and cell metabolism, by monitoring pH and lactate production per molecule glucose of the conditioned medium. We confirmed higher MUC2 production under SWMS conditions at gene and protein level and demonstrated that this culturing method primarily stimulated cell growth. In addition, we also found evidence for a more aerobic cell metabolism under SWMS, as shown previously for similar models. In summary, we suggest different mechanisms by which MUC2 production is enhanced under SWMS and propose potential applications of this model in future studies.

## INTRODUCTION

The surface of the human gastro-intestinal tract (GIT) tract is covered by a layer of mucus, protecting the host from pathogens, harmful chemical or biological substances and physical damage [529-532]. Defects in this mucus layer have been implicated in several intestinal pathologies. For instance, *Muc2* knock-out mice were shown to develop spontaneous colitis [533] and colorectal cancer [534], emphasizing the important protective role of mucus. Along the GI tract, the mucus layer is thickest in the colon [535], where the number of intestinal bacteria is also highest [536]. Interestingly, the mucus layer does not only protect the underlying epithelium from these high bacterial numbers, but also provides a binding site and nutrition-rich niche for residing intestinal bacteria, such as mucin-degrading bacterial species which liberate short-chain fatty acids for cross-feeding bacteria and the host [143, 537, 538]. As a nutrient source, mucin has been shown to be a major driver of intestinal microbiota composition *in vitro* [539].

The dual role of colonic mucus can be explained by the existence of two layers: The inner layer is densely packed, firmly attached to the intestinal epithelium and devoid of bacteria, whereas the outer layer is loose, constantly removed and colonized by bacteria [104, 535]. Colonic mucus is mainly composed of the gel-forming mucin type 2 (MUC2) [540, 541], a heavily O-glycosylated protein secreted by intestinal goblet cells [542-544]. Additionally, mucus contains salts, lipids [545] and defense-related proteins, such as antimicrobial peptides, growth factors, trefoil factors, immunoglobulins and other proteins [104].

Our understanding of colonic mucus structure, function and composition has been largely dependent on *in vivo* models, such as rodents and pigs (reviewed in Etienne-Mesmin *et al.* [451]) or *ex vivo* techniques using human mucosal biopsies (first established by Browning *et al.* [546]) and murine intestinal explants [405, 547]. These models, however, show large heterogeneity between and within subjects or, in the case of animal models, have a poor translational value. Additionally, these models are often expensive, require specialized experience and pose ethical concerns. Attempts have been made to recapitulate the intestinal mucosal layer – including host cells – *in vitro*, varying from simple to more advanced systems. Relatively simple cell models include the use of mucus-excreting colonic cancer cell lines, such as HT29-MTX and LS174T cells [242, 548], but both examples have limitations. HT29-MTX cells have successfully been grown in confluent monolayers [549], but this cell type predominantly secretes MUC5AC, a mucin that is present in the stomach and airways, while producing only a limited amount of colon-specific MUC2 [242]. On the other hand, LS174T cells do produce MUC2 [548], but are not capable of growing in an organized and adherent cell layer [510]. More advanced models include culturing cell lines in innovative models, such as gut-on-a-chips [410, 477] or 3D scaffolds [550-552]; or the use of human colonoids [553, 554] and human intestinal organoids [222];



or a combination [555], which all demonstrated increased MUC2 production *in vitro* compared to conventional cell culture. Although these models resemble the *in vivo* colonic mucosal layer more closely in terms of mucus composition, they are expensive and require specialized expertise [451].

To obtain a more physiologically relevant mucus barrier, simpler yet effective alternative strategies have been shown to further increase mucus production and/or secretion in intestinal cancer cell lines, using biochemical compounds, such as prostaglandin E2 [556] and Notch  $\gamma$ -secretase inhibitors [557], or bacteria-derived compounds (e.g. sodium butyrate [558] and LPS [559]). Physical strategies have also been applied, e.g. growing intestinal porcine epithelial cells at an air-liquid interface (ALI) [560, 561], stimulating cells mechanically, or a combination of both. For instance, Navabi and colleagues managed to create polarized, functional, crypt-forming intestinal cell layers with an adherent mucus layer, when growing HT29-MTX-E12 and other intestinal cell lines on Transwell membranes in semi-wet interfaces with mechanical stimulation (SWMS). SWMS conditions include decreased apical and basolateral medium volumes and continuous shaking on a rocking platform. Importantly, these cells demonstrated increased expression of MUC2, both absolute and relative to MUC5AC [510]. HT29-MTX-E12 cells grown under these conditions produced a thicker layer compared to static conditions. The positive effect of mechanical stimulation was also shown in more recent *in vitro* models, showing a positive effect on MUC2 production [410]. However, it remains unexplored what molecular mechanisms are involved.

In our study, we aimed to further explore the cellular processes underlying the increased production of MUC2 by HT29-MTX-E12 cells grown under SWMS conditions. To this end, cells were subjected to transcriptome analysis after 15 days of culture to investigate underlying molecular pathways involved. As control cell lines for the transcriptomic analysis, we included Caco-2 and LS174T cells, a non-mucus producing and MUC2-producing cell line, respectively. Next, we further characterized the HT29-MTX-E12 monolayer, by measuring transepithelial electrical resistance and quantifying cell density. Additionally, as similar (semi-wet only) models have shown a more aerobic cell metabolism [562-565], we quantified pH and lactate production per molecule glucose of the conditioned medium.

In brief, our study confirms the upregulation of MUC2 in HT29-MTX-E12 cells cultured under SWMS. Additionally, our study demonstrates upregulation of cell cycle processes, downregulation of KLF4, differential regulation of ion transporters and increased aerobic metabolism of cells cultured under SWMS versus static conditions. In overall, we attempted to gain more insight into the potential mechanisms underlying increased MUC2 production in HT29-MTX-E12 cells grown under SWMS conditions.

## MATERIALS & METHODS

### Cell culture

HT29-MTX-E12 cells (ECACC) were obtained from Sigma-Aldrich (Darmstadt, Germany). Caco-2 (ATCC HTB-37) and LS174T (ATCC CL-188) cells were purchased at LGC Standards (Wesel, Germany). All cell types were cultured in Dulbecco's Modified Eagle Medium with 4.5g/L glucose, 110 mg/L sodium pyruvate and 584 mg/L L-glutamine (Corning, NY, USA) supplemented with 10% Fetal Bovine Serum and 1% penicillin/streptomycin. When cells reached 80-90% confluency, they were counted and seeded. Passage numbers between 15 and 27 were used for Caco-2 cells and between 3 and 21 for HT29-MTX-E12 cells. LS174T cells were used between passage 17 and 19. Caco-2 and HT29MTX-E21 cells were seeded (day 0) at a density of 273,000 cells/mL in 275  $\mu$ L per well on 12 mm 0.4  $\mu$ m-pore polyester Transwell membranes (Corning 3460). A volume of 1 mL medium was added to the basolateral compartment. One day after seeding (day 1), media of all Transwells was refreshed and Semi-Wet conditions with Mechanical Stimulation (SWMS) were applied [510, 552]. To these Transwells, 75  $\mu$ L and 850  $\mu$ L of medium was added to the apical and basolateral compartments, respectively, and plates were put on a CO<sub>2</sub>-resistant shaker (Thermo Fisher Scientific, Breda, The Netherlands) at 65 rpm. Media volumes of Transwells grown under static conditions were unchanged. LS174T cells were seeded at similar seeding density on regular 12-well cell culture plates. This cell line was cultured under static conditions in a regular wells plate only, since these cells do not form a continuous monolayer of cells, but grow in a rather irregular manner [510]. Medium was refreshed every Monday, Wednesday and Friday and cells were harvested 15 days after seeding ( $t = 15$  days). Pictures of the cells on Transwell membranes were taken with a Leica DFC450C microscope camera using Leica Application Suite X software. A 20x magnification was used.

### RNA isolation

Cells were washed with ice-cold PBS twice, trypsinized and RNA was isolated using the Maxwell® 16 LEV simplyRNA Cells Kit (Promega, cat. no. AS1270) and the Maxwell® 16 MDx Instrument (Promega), following the manufacturer's instructions.

### Microarray

RNA isolate from three independent biological replicate experiments was used for microarray analysis. Total RNA yield was measured using photometry (DeNovix, USA). RNA quality was determined on an Agilent 2100 Bioanalyzer (Agilent Technologies, Amsterdam, The Netherlands). RNA was only used when the RNA integrity number (RIN) exceeded 8.0. One hundred nanogram of RNA was converted to cDNA and labelled (Ambion WT expression kit, Life Technologies, Bleiswijk, The Netherlands). Samples were hybridized to an Affymetrix Human Gene 1.1 ST array



plate according to the standard Affymetrix instructions (Affymetrix, Santa Clara, CA, USA). The robust multi-array average (RMA) pre-processing algorithm in the Bioconductor library AffyPLM was used to obtain normalized expression estimates [566]. Probe sets were defined and assigned as described by Dai *et al.* [567]. Differences in gene expression between static and SWMS conditions per cell type were analyzed using the Intensity Based Moderated T statistics (IBMT) [568], using *p* values <0.05 as threshold. The Venn diagram was created using Venny 2.1 [569]. Microarray data has been submitted to the Gene Expression Omnibus (GEO) at the NCBI (GSE173729).

### TEER measurements

The transepithelial electrical resistance (TEER) was measured with an EVOM2 Volt/Ohm meter using STX2 electrodes (World Precision Instruments) at day 4, 7, 9, 11 and 14. For each biological replicate (= independent experiment, *n* = 3), two technical replicates (= wells) were measured. One hour prior to measuring, medium of the Transwells was refreshed and equal medium volumes were applied in all wells (275  $\mu$ L apical and 1 mL basolateral). Before TEER measurements were performed, Transwells were put at room temperature for 5 min to allow temperature equilibration. After TEER measurements, the medium volumes in the wells were adapted again to volumes of the respective conditions. The background value (i.e., TEER value of an empty Transwell) was subtracted from the total TEER values.

### FITC-dextran assays

The passage of Fluorescein isothiocyanate–dextran with an average molecular weight of both 4 kDa (FITC-D4) and 40 kDa (FITC-D40) (Sigma-Aldrich, FD4 and FD40) from the apical to the basolateral compartment was measured. FITC-D40 was dissolved in DMEM without phenol red (Gibco, 21063029) at a concentration of 1 mg/mL. A control was added with cells from which the mucus layer was removed. To this end, cells on Transwells were treated for 1 hour with 60 mM N-Acetyl-L-cysteine (Sigma, A9165). After washing with PBS, 275  $\mu$ L FITC-D40 was added to the apical membrane and 1000  $\mu$ L DMEM without phenol red to the basolateral membrane. After 3 h incubation at 37 °C/5% CO<sub>2</sub>, 100  $\mu$ L of medium was taken from the basolateral compartment and measured on a Spectramax M2 fluorescence plate reader at 490/520 nm (excitation/emission). An empty Transwell served as a control for complete permeability. For each biological replicate (= independent experiment, *n* = 3), two technical replicates (= wells) were measured.

### Protein isolation

Cells were washed twice with PBS and lysed in RIPA Lysis and Extraction Buffer (ThermoFisher Scientific), supplemented with the protease and phosphatase inhibitors PhosSTOP and cComplete (Roche Diagnostics, Almere, The Netherlands). Lysates were

incubated on ice for 20 min following centrifugation for 10 min at 13,000 x g. Supernatant was collected and protein concentrations were measured using a bicinchoninic acid assay (Thermo Fisher Scientific).

### Western Blot

Protein lysates (20 µg of protein/lane) were loaded onto 4-15% Mini-PROTEAN TGX Precast Protein Gels (Bio-Rad, Veenendaal, The Netherlands). Next, proteins were transferred onto a polyvinylidene difluoride membrane (Trans-Blot Turbo Midi 0.2 µm PVDF Transfer Packs, Bio-Rad) using the Transblot Turbo System (Bio-Rad). Membranes were blocked for 1 hour and incubated overnight at 4 °C with rabbit anti-KLF4 (Sigma-Aldrich, catalogue no. SAB1300678) and rabbit anti-HSP90 (Cell Signaling Technology, cat. no. 4874). Antibodies were used in 1:500 and 1:5000 dilutions for KLF4 and HSP90, respectively. Membranes were incubated for 1 hour with goat anti-rabbit (GenScript, cat. no. A00098) diluted 1:5000. Blocking and incubation of primary and secondary antibodies were done in TBS with 0.1% Tween 20 (TBS-T) and 5% (w/v) skimmed milk. In between, membranes were washed in TBS-T. Signals were quantified using the ChemiDoc MP system (Bio-Rad) and Clarity ECL substrate (Bio-Rad).

### Dot Blot

Proteins were diluted to 68.2 µg/mL and 6x serially diluted 1:2 in PBS. In total 7 dilutions and one PBS-control were blotted per condition. Of each dilution, 50 µL of sample was blotted on a Pierce 0.2 µm nitrocellulose membrane (Thermo Scientific) in a Dot Blot device (The Convertible, cat. series 1055. Gibco BRL) connected to a vacuum-pump. Next, membranes were blocked, incubated and imaged as described for Western Blot. Primary antibodies were diluted 1:2000 and 1:1000 for mouse anti-MUC2 (Abcam, cat. no. ab11197) and -MUC5AC (Sigma-Aldrich, cat. no. WH0004586M7) respectively. Secondary goat anti-mouse antibody (Genscript, cat. no. A00160) was diluted at 1:2500. To check protein quantity on the membranes, a separate, identical membrane was blotted and incubated for 10 min in Ponceau Red (Honeywell Fluka). Next, the blot was washed in demi-water and signals were quantified using the ChemiDoc MP system. Dot Blots with anti-MUC2 and -MUC5AC were quantified by measuring the density of the first six dots (rows) using ImageJ software. A density-based linear trendline was calculated from the second to the sixth dot. The coefficients (per µg/mL) were corrected for Ponceau Red density. Ponceau Red signals were quantified by measuring density of the first dot (first row) only. Biological replicates per condition were  $n = 3$ .



### pH, lactate and glucose measurements

Every time medium of Transwells was refreshed, conditioned medium was collected from HT29-MTX-E12 cells and stored at -20 °C until further processing. Medium of apical and basolateral compartments were collected separately, but compartments were pooled per time point and condition. Samples were spun down and 200 µL of supernatant was transferred to a 96-wells plate. After stabilization, absorbance was measured at 415 and 560 nm at 5% CO<sub>2</sub> in a Synergy Neo2 Hybrid Multi-Mode Reader with a CO<sub>2</sub> and O<sub>2</sub> gas controller (1210013) (BioTek Instruments). Aliquots of growth medium (ca. 5 mL) were used for preparation of a standard curve with pH values ranging from 2.6 to 9.9. Lactate and glucose were quantified with a Shimadzu LC2030C-Plus high-performance liquid chromatography (HPLC) system equipped with a Shodex SH1821 column kept at 45 °C and running 0.01 N sulfuric acid as eluent (1 mL/min). Compounds were detected by determining the refractive index and identified using pure lactate and glucose as external standards and crotonate as an internal standard.

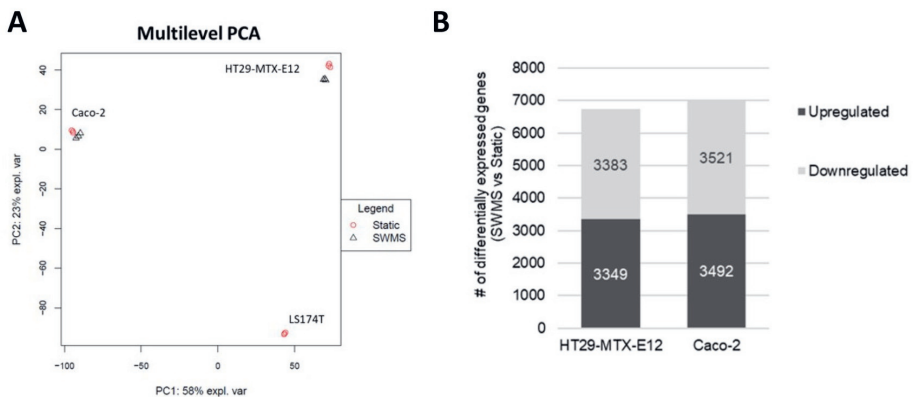
### Statistical analysis

Statistical analysis of microarray data is described above. Data distribution for other (continuous) outcomes was tested with the D'Agostino-Pearson omnibus normality test using GraphPad Prism (San Diego, CA, USA). Significance of differences between two conditions was analysed using a two-tailed, unpaired Student's t-test. For non-normally distributed data, a Mann-Whitney U test was used. The data are presented as mean  $\pm$  standard deviation. A *p*-value of  $\leq 0.05$  was considered significant. All experiments were performed in three biological replicates.

## RESULTS

### Gene expression changes in HT29-MTX-E12, Caco-2 and LS174T cells

To investigate SWMS-specific effects leading to higher mucus production of the SWMS culturing method, HT29-MTX-E12 cells were grown for 14 days under static and SWMS conditions (after cell seeding and attachment), and microarray analysis was performed on mRNA of the cells. We included Caco-2 cells, a non-mucus producing cell line, grown on Transwell membranes under static and SWMS conditions for 14 days as a reference control. The LS174T cell line was used as a control for a mucus-producing cell line. Multilevel principal component analysis performed on the 500 most variable genes showed a strong clustering of the biological replicates per cell type with a clear separation of the three cell lines (**Figure 1A**). A number of 6,732 and 7,013 genes were differentially regulated between SWMS and static conditions in HT29-MTX-E12 and Caco-2 cells, respectively ( $p < 0.05$ ) (**Figure 1B** and **S1A-B Figure**).



**Figure 1. Overview microarray results in HT29-MTX-E12, Caco-2 and LS174T cells grown under static and/or SWMS conditions. A)** Multilevel principal component analysis (PCA) of the 500 most variable genes. **B)** Number of differentially expressed genes between SWMS and static conditions in HT29-MTX-E12 and Caco-2 cells (IBMT  $p < 0.05$ ). Microarray results are based on  $n = 3$  biological replicates per condition.



### Increased MUC2/MUC5AC ratio in SWMS-cultured HT29-MTX-E12 cells

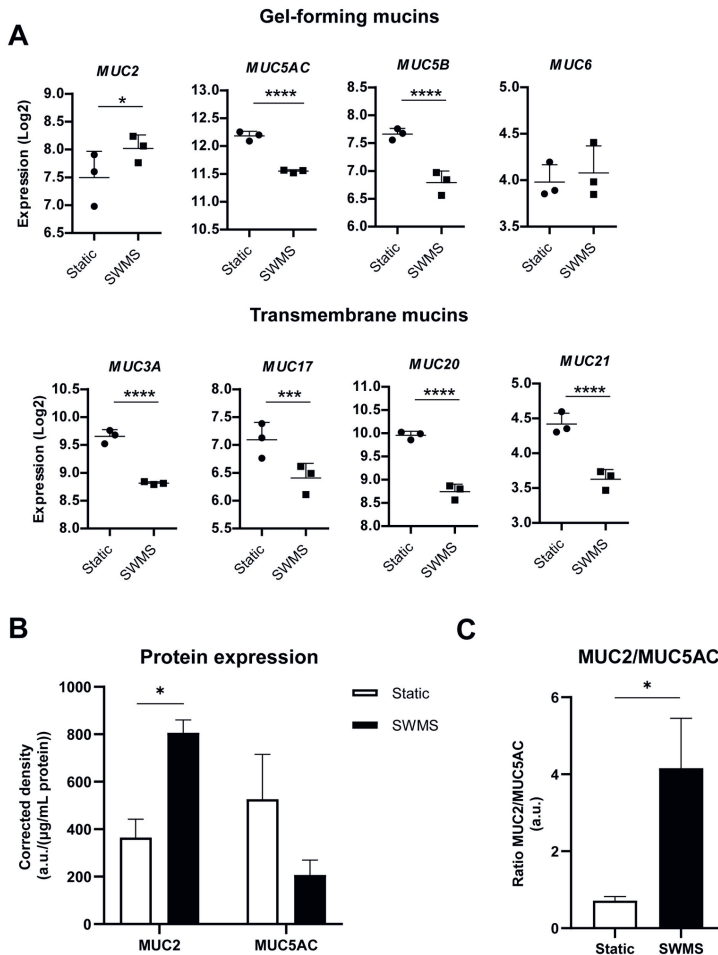
We investigated the expression of genes encoding for mucins in HT29-MTX-E12 cells, since we know from previous studies that culturing this cell type under SWMS conditions led to increased MUC2 production [510]. Indeed, in our study, SWMS conditions resulted in a significant 1.42-fold upregulation of MUC2 ( $p < 0.05$ ) (**Figure 2A**). Expression of gel-forming mucins MUC5AC and MUC5B was decreased in response to SWMS (FC = -1.56 and -1.84, respectively,  $p < 1 \cdot 10^{-4}$ ) (**Figure 2A, S1 File**). Transmembrane mucins were also significantly lower expressed after SWMS conditions (**Figure 2A**). MUC3A, 13, 17, 20 and 21 showed FC values between -1.52 and -2.32 (**S1 File**). Eight out of 19 mucins (MUC4, 6, 7, 12, 15, 16, 19 and 22) displayed very low expression levels under both static and SWMS conditions (RMA < 4) and were not significantly differentially expressed (**S1 File**). Taken together, these data indicate that SWMS conditions resulted in a change in expression of mucin encoding genes with an increased expression of MUC2, while other gel-forming as well as transmembrane mucins were significantly lower or not differentially expressed.

Dot Blot data supported microarray data for MUC2, as protein levels also showed a significant upregulation ( $p < 0.05$ ) under SWMS versus static conditions (**Figure 2B, S2A and C Figure**). Less MUC5AC was detected under SWMS conditions, however, variation between replicates was rather high and differences were not significant ( $p = 0.20$ ) (**Figure 2B and S2B and C Figure**). Overall, HT29-MTX-E12 cells grown under SWMS conditions showed a significantly increased MUC2/MUC5AC-ratio (**Figure 2C**,  $p < 0.05$ ), confirming previous findings [510]. For LS174T, which we included as a control cell line reported to produce predominantly MUC2, a relatively high MUC2/MUC5AC ratio was calculated (**S2D-E Figure**). In contrast, Caco-2 cells, which are not reported to produce mucus, showed low RMA values (< 5) for MUC2 and MUC5AC and protein expression of these mucins was below detection level in both conditions (**S2F-G Figure**). Furthermore, similar to HT29-MTX-E12 cells, SWMS conditions resulted in significant downregulation of MUC20, MUC3A and MUC13 in this cell line (**S1 File**).

### Upregulation of genes and pathways related to cell cycle, cell growth and cell proliferation in cells cultured under SWMS conditions

As both HT29-MTX-E12 and Caco-2 cell lines were subjected to SWMS conditions and grown under static conditions, we compared the genes up- or downregulated in both cell lines. Of the 2,844 significantly differentially expressed genes in both cell lines, 353 genes (12.4 % of shared total significantly regulated genes) were upregulated by 1.5-fold or higher in both cell types due to the SWMS conditions. A total of 180 genes (6.3 %) were significantly downregulated by 1.5-fold or higher in both cell lines and a total of 98 genes showed opposite effects when comparing the two cell lines (**S3 Figure**). Interestingly, among the shared upregulated genes, a

number of cell cycle related genes was included. Moreover, the top 10 most upregulated genes of HT29-MTX-E12 cells under SWMS versus static conditions was dominated by genes related to cell growth, cell motility and cell proliferation, i.e., *KIF14*, *KIF20A*, *DEPDC1*, *DLGAP5* and *SPC25* (**Table 1**).



**Figure 2. Mucin gene and protein expression in HT29-MTX-E12 cells grown under static and SWMS conditions. A)** Microarray gene expression values (Log2) of a panel of secretory and transmembrane mucins **B)** Protein expression of MUC2 and MUC5AC, expressed as density (a.u.) per  $\mu\text{g/mL}$  protein blotted, after correction of Ponceau Red density. Blot data in S2 Figure) **C)** Ratio of MUC2 and MUC5AC protein expression. \*  $p < 0.05$ ; \*\*\*  $p < 0.001$ ; \*\*\*\*  $p < 0.0001$ ;  $n = 3$  biological replicates.

**Table 1.** Top 10 strongest up- and downregulated genes in HT29-MTX-E12 cells grown under SWMS versus static conditions.**UPREGULATED**

Gene	FC	p-value	Gene name
RNY4P23	10.80	5.40E-06	RNY4 pseudogene 23
SUCNR1	8.78	2.55E-11	succinate receptor 1
KIF14	8.00	2.45E-12	kinesin family member 14
KIF20A	7.99	1.65E-08	kinesin family member 20A
DEPDC1	7.93	6.11E-11	DEP domain containing 1
DLGAP5	7.91	5.15E-12	DLG associated protein 5
HMMR	7.82	1.03E-15	hyaluronan mediated motility receptor
H4C1	7.56	7.27E-07	H4 clustered histone 1
IL33	7.44	2.98E-13	interleukin 33
SPC25	7.41	4.47E-13	SPC25 component of NDC80 kinetochore complex

**DOWNREGULATED**

Gene	FC	p-value	Gene name
TFF2	-15.75	2.62E-13	trefoil factor 2
SCGB2A1	-8.91	1.33E-10	secretoglobin family 2A member 1
ST8SIA6	-4.56	1.06E-08	ST8 alpha-N-acetyl-neuraminide alpha-2,8-sialyltransferase 6
DUOX2	-4.53	4.52E-09	dual oxidase 2
TFF1	-4.44	1.37E-16	trefoil factor 1
APOL1	-4.41	1.29E-06	apolipoprotein L1
ADGRF1	-3.96	7.06E-08	adhesion G protein-coupled receptor F1
DUOXA2	-3.88	1.67E-09	dual oxidase maturation factor 2
UPK3B	-3.80	4.29E-09	uroplakin 3B
SPRR1B	-3.63	2.31E-07	small proline rich protein 1B

Gene Set Enrichment Analysis (GSEA) also revealed that most significantly enriched upregulated pathways were dominated by cell cycle and DNA replication related pathways (**Table 2** and **S2 File**). Furthermore, marker of proliferation Ki-67 (*MKI67*) was 5.03-fold upregulated by SWMS culturing conditions ( $p < 1 \cdot 10^{-10}$ ) in HT29-MTX-E12 and 2.33-fold in Caco-2 cells ( $p < 1 \cdot 10^{-5}$ ) (**S4A Figure**). Additionally, we found higher cell counts at day 15 under SWMS compared to static conditions in both cell lines (**S4B Figure**), confirming that SWMS conditions lead to an increase in cell proliferation. Closely related to cell cycle genes are cytoskeleton-related genes, some of which also showed differential regulation under SWMS (**S1 File**). Altogether, these data indicate that SWMS conditions result in the activation of cell proliferation pathways compared to static conditions, independent of cell type.

**Table 2.** Top 10 most significantly enriched upregulated pathways in HT29-MTX-E12 cells induced by SWMS conditions.

Pathway entry	Enriched upregulated pathways	NES*	FDR q-value
HSA04110	Cell Cycle	2.856	0.000
HSA03460	Fanconi Anemia Pathway	2.637	0.000
HSA03030	DNA Replication	2.596	0.000
HSA05322	Systemic Lupus Erythematosus	2.539	0.000
HSA03440	Homologous Recombination	2.506	0.000
HSA04114	Oocyte Meiosis	2.329	0.000
HSA03420	Nucleotide Excision Repair	2.315	0.000
HSA03430	Mismatch Repair	2.312	0.000
HSA00100	Steroid Biosynthesis	2.306	0.000
HSA04914	Progesterone Mediated Oocyte Maturation	2.268	0.000

\*NES: Normalized enrichment score.

### Expression of target genes involved in Notch- and Atoh-key pathways did not point towards a favored cell differentiation state by SWMS conditions

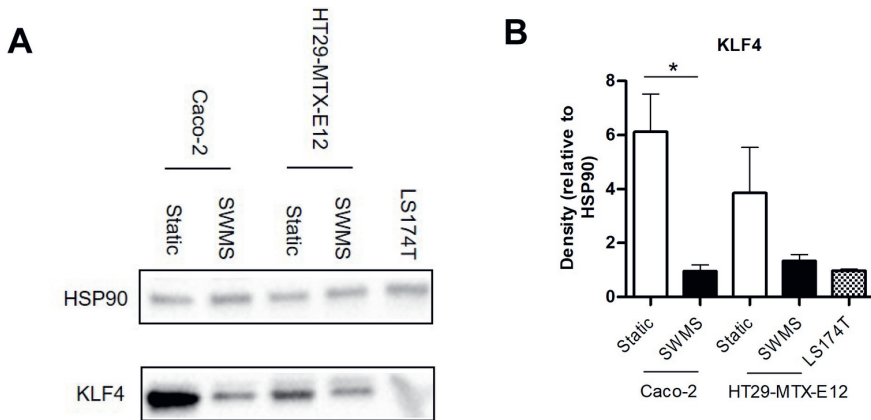
Next to the marked effects of SWMS conditions on cell proliferation in both HT29-MTX-E12 and Caco-2 cells, we took a closer look at the effects on intestinal cell differentiation. Key regulators in this process belong to the Notch/Atoh signaling pathway [570, 571]. When it comes to cell differentiation, Notch and Atoh have opposing roles: Notch activation promotes absorptive cell differentiation, while Atoh activation favors differentiation into secretory cell types [570-572]. Interestingly, in our transcriptomic dataset, both *NOTCH1* and *ATOH1* were significantly upregulated by SWMS conditions in HT29-MTX-E12 cells ( $FC = 1.61$ ,  $p < 0.001$  and  $FC = 1.64$ ,  $p < 0.01$ , respectively) and the latter also in Caco-2 cells ( $FC = 2.09$ ,  $p < 0.001$ ). The stem cell marker *LGR5*, which is known to be part of a positive feedback loop regulated by Notch [573], was significantly upregulated by SWMS conditions in



HT29-MTX-E12 cells ( $FC = 2.33$ ,  $p < 1 \cdot 10^{-5}$ ), although RMA values were below 5. In Caco-2 cells, *LGR5* had higher RMA values and was even stronger upregulated ( $FC = 4.37$ ,  $p < 1 \cdot 10^{-7}$ ). Further downstream the Notch pathway, the target genes Hairy and enhancer of split (HES) family members *HES1* and *HES6* were both significantly upregulated by SWMS conditions in HT29-MTX-E12 cells ( $FC = 1.24$ ,  $p < 0.05$  and  $FC = 2.08$ ,  $p < 1 \cdot 10^{-5}$ , respectively). Apart from Cyclin D1 (*CCND1*) ( $FC = -1.51$ ,  $p < 1 \cdot 10^{-6}$ ) no other Notch-target genes were significantly differentially expressed. Regarding the Atoh pathway, the downstream target gene Neurogenin 3 (*NEUROG3*) was 2.07-fold and 2.41-fold higher expressed ( $p < 0.001$ ) in HT29-MTX-E12 and Caco-2 cells, respectively. Other Atoh target genes, such as *SPDEF* and *GFI1* were not differentially expressed between SWMS and static conditions in both cell types. Altogether, these transcriptomic data show that some target genes of both the Notch and Atoh pathway were differentially expressed in HT29-MTX-E12 and Caco-2 cells. However, these data do not point towards a favored differentiation state (absorptive versus secretory cell fate).

### Downregulation of KLF4 at both gene and protein level under SWMS conditions in HT29-MTX-E12 and Caco-2 cells

Another Notch-target involved in differentiation of progenitor cells into goblet cells is KLF4 [574]. Given the secretory, goblet cell-like phenotype of HT29-MTX-E12 cells grown under SWMS conditions, the significant downregulation of *KLF4* ( $FC = -1.86$ ,  $p < 0.0001$ ) is interesting. Acting as both a transcriptional repressor and activator in the gastrointestinal epithelium, this zinc-finger transcription factor plays a critical role in the decision between proliferation and cell cycle arrest/differentiation [575]. We validated the downregulation of KLF4 under SWMS conditions at protein level using Western Immunoblotting. Indeed, when cultured under SWMS conditions, the expression of KLF4 protein was lower compared to the static conditions in both cell lines (**Figure 3A-B** and **S5 Figure**). We next investigated whether this downregulation of KLF4 resulted in differential expression of target genes of KLF4, as identified in literature [576]. Among the most strongly regulated are genes related to cell-cycle control or essential for cell proliferation and differentiation, including cyclin B1 ( $FC = 6.59$ ,  $p < 1 \cdot 10^{-13}$ ) [577], cyclin E2 ( $FC = 2.03$ ,  $p < 0.001$ ) [578, 579], ornithine decarboxylase ( $FC = 1.80$ ,  $p < 1 \cdot 10^{-7}$ ) [580], cyclin E1 ( $FC = 1.62$ ,  $p < 1 \cdot 10^{-4}$ ) [578, 579], cyclin D1 ( $FC = -1.51$ ,  $p < 1 \cdot 10^{-6}$ ) [581, 582] and cyclin D2 ( $FC = 1.85$ ,  $p < 0.001$ ) [583]. A non-exhaustive list of genes (in)directly related to KLF4 is provided in **Supplementary File 1**. In contrast to *KLF4*, other reported goblet-cell markers such as *AGR*, *CDX2*, *MEP1B* and *FCGBP*, amongst others, were not significantly regulated in HT29-MTX-E12 cells grown under SWMS. A non-exhaustive list of “Goblet cell markers” [584] is given in **S1 File**.



**Figure 3. Protein expression of KLF4 in HT29-MTX-E12, Caco-2 and LS174T cells grown under static and/or SWMS conditions.** **A)** Western Blot results of KLF4 and HSP90 (house-keeping protein) in Caco-2, HT29-MTX-E12 and LS174T-cells grown under static and SWMS conditions or static only (LS174T). Of each condition, one representative biological replicate is shown. See **S5 Figure** for all replicates. **B)** KLF4 protein quantity expressed as the protein band density relative to HSP90. \*  $p < 0.05$ ,  $n = 3$  biological replicates.

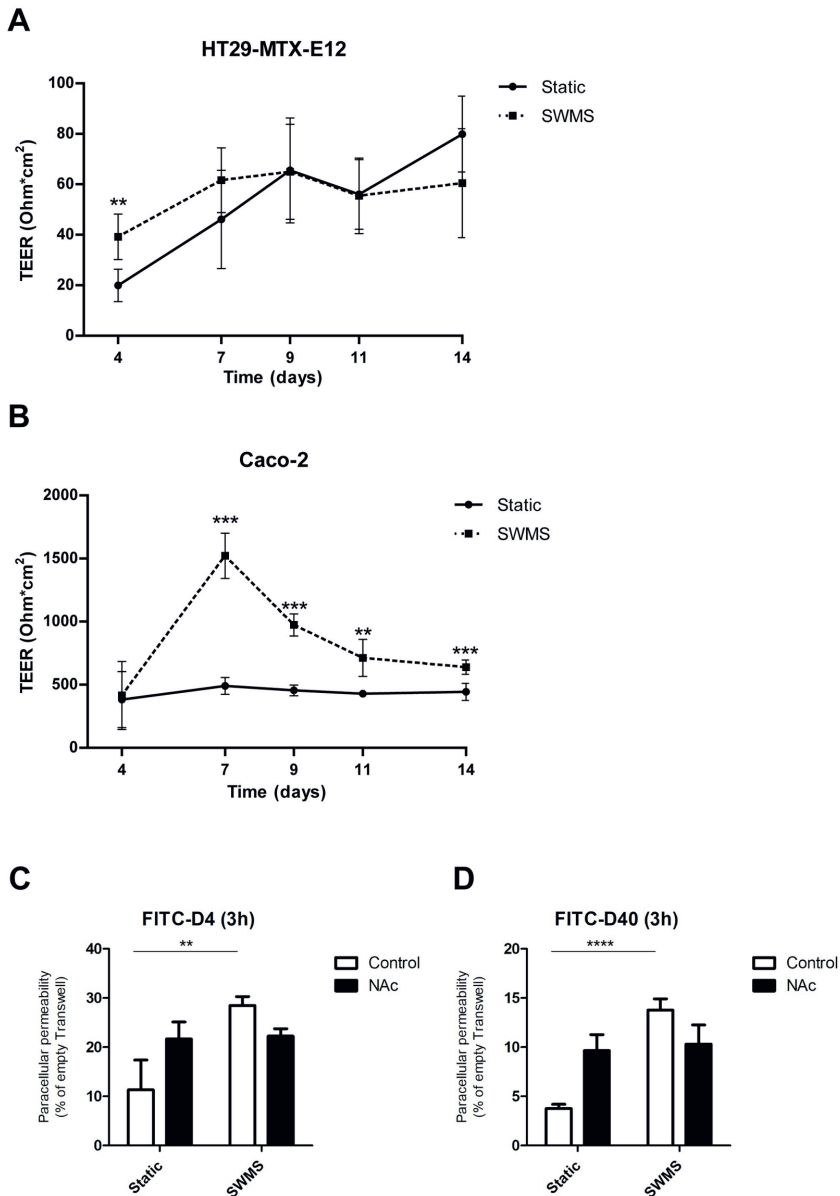
#### Minor effect on transepithelial electrical resistance, but increased paracellular permeability of HT29-MTX-E12 cells grown under SWMS conditions

To investigate the effects of SWMS conditions on intestinal epithelial barrier integrity, transepithelial electrical resistance (TEER) was measured at multiple time points during culturing. TEER values of HT29-MTX-E12 cells increased steadily over time and no clear effects were observed between static and SWMS conditions (**Figure 4A**). For Caco-2 cells, the TEER of the cells grown on SWMS conditions increased steeply during the first seven days of culturing and decreased gradually afterwards, while the TEER of cells grown on static conditions was relatively stable (**Figure 4B**). For HT29-MTX-E12, we complemented TEER data with assays using fluorescent probes (FITC-dextran of 4 and 40 kDa) to quantify paracellular permeability. Interestingly, at  $t = 14$  days, HT29-MTX-E12 cells grown under SWMS showed higher permeability than static conditions for both probe sizes (**Figure 4C-D**). After washing off the mucus with a mucolytic reagent (N-Acetyl-L-cysteine (NAC)), however, no differences in paracellular permeability could be observed. Based on available literature, we compiled a panel of genes responsible for epithelial barrier integrity (**Supplementary File 1**). Except for a 3.33-fold increase in expression of Catenin alpha like 1 (CTNNAL1) in HT29-MTX-E12 cells under SWMS conditions ( $p < 1 \cdot 10^{-9}$ ), only a low number of significantly differentially expressed genes was found for both HT29-MTX-E12 and Caco-2 cells (**S1 File**). Together, these results indicate that SWMS conditions had

negligible effects on TEER in HT29-MTX-E12 cells, which was reflected by the relatively low number of significantly differentially expressed genes related to cell integrity but increased paracellular permeability. Interestingly, a clear difference in cell appearance on the Transwell membranes was observed for both cell types at the time cells were harvested, as cells grown under SWMS conditions seemed to concentrate in the center of the membrane (**S6A-B Figure**).

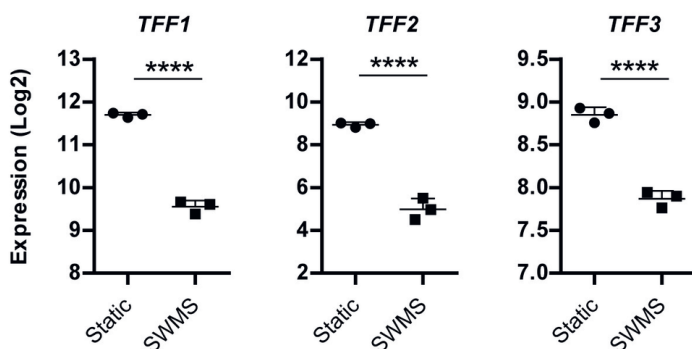
#### **Downregulation of Trefoil Factor genes in HT29-MTX-E12 cells grown under SWMS conditions**

When focusing on the most significantly downregulated genes in HT29-MTX-E12 grown under SWMS compared to static conditions, Trefoil Factor 2 (*TFF2*) was most strongly downregulated ( $FC = -15.75$ ,  $p < 1 \cdot 10^{-12}$ ) (**Table 1**). Although to a lesser extent, *TFF1* and *TFF3* were also significantly downregulated (**Figure 5**). Trefoil factors are small cysteine rich peptides and form a family of mucin-associated secretory molecules involved in many physiological processes to maintain and restore gastrointestinal mucosal homeostasis (Reviewed in Aihara *et al.* [585]). TFF peptides are known to auto- and cross-regulate their expression via the epidermal growth factor receptor *in vitro* [585, 586]. However, *EGF1* was not significantly differentially expressed between culture conditions in HT29-MTX-E12 cells in our dataset. On the contrary, the gene encoding Epidermal growth factor receptor (*EGFR*) was significantly downregulated by 1.42-fold ( $p < 0.0001$ ).



**Figure 4. Transepithelial electrical resistance (TEER) and paracellular permeability of HT29-MTX-E12 and/or Caco-2 cells grown under static and SWMS conditions up to 14 days.** TEER was measured at  $t = 4, 7, 9, 11$  and 14 days and expressed in  $\text{Ohm}/\text{cm}^2$  in **A)** HT29-MTX-E12 and **B)** Caco-2 cells,  $n = 3$  biological replicates. Paracellular permeability was quantified by incubating for 3 h with FITC-dextran of size **C)** 4 kDa and **D)** 40 kDa. Values were normalized to an empty Transwell,  $n = 3$  biological replicates. Control = with mucus, NAc = treated with N-acetyl-L-cysteine to remove mucus; \*  $p < 0.05$ ; \*\*  $p < 0.01$ ; \*\*\*  $p < 0.001$ .





**Figure 5. TFF gene expression in HT29-MTX-E12 cells grown under static and SWMS conditions.** Microarray expression values (Log2) of *TFF1*, 2 and 3 and the corresponding fold change (FC). \*\*\*\*  $p < 0.0001$ ;  $n = 3$  biological replicates.

### Different regulation of ion transporters under SWMS conditions in both HT29-MTX-E12 and Caco-2 cells

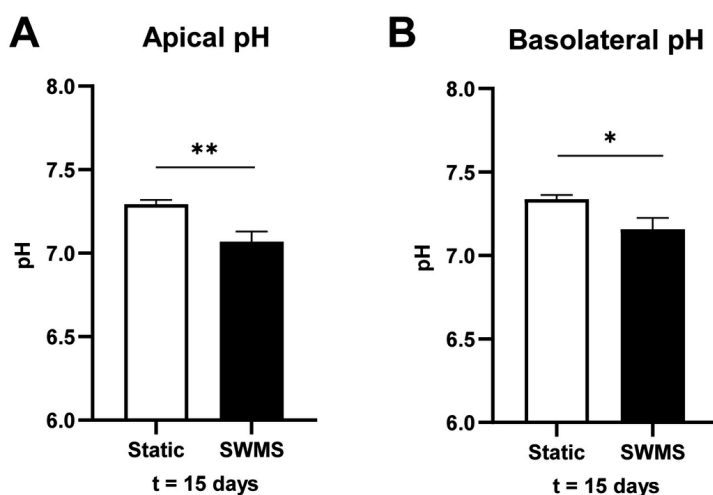
Since ion transport has proven crucial in (intestinal) mucus production [587], we explored the effect of SWMS conditions on the expression of ion transporters and exchangers. The gene encoding Cystic Fibrosis transmembrane regulator (CFTR), a chloride and bicarbonate transporter expressed on the apical side of (intestinal) epithelial cells and demonstrated to be indispensable for normal mucus production [588], was significantly upregulated in HT29-MTX-E12 cells grown under SWMS conditions (FC = 1.68,  $p < 0.00001$ ). Other ion transporters that were significantly regulated under SWMS conditions in HT29-MTX-E12 and Caco-2 cells include *NBCe1/SLC4A4*, encoding a basolateral  $\text{Na}^+/\text{HCO}_3^-$  importer (FC = -1.95,  $p < 0.01$  in HT29-MTX-E12) and *NHE1/SLC9A1*, encoding a basolateral  $\text{Na}^+/\text{H}^+$  exchanger (FC = -1.59,  $p < 1 \cdot 10^{-5}$ ). Additionally, the significant downregulation of *DRA/SLC26A3*, encoding an apical  $\text{Cl}^-/\text{HCO}_3^-$  exchanger (FC = -1.25,  $p < 0.05$ ) under SWMS conditions is interesting, given its importance in intestinal salt and fluid absorption [589]. This gene was, however, not highly expressed in HT29-MTX-E12 cells (RMA < 3). Additionally, the major regulator of intestinal pH, NHE3 (SLC9A3) [590-592], was however, relatively low expressed in static HT29-MTX-E12 cells (RMA < 6) and only slightly altered under SWMS conditions. Whereas GSEA revealed that the pathway “Mineral absorption” was not significantly regulated in HT29-MTX-E12, interestingly, this pathway was the most enriched among all downregulated pathways in Caco-2 cells (NES = -2.58, FDR q-value = 0.00, **S2 File**). These cells also showed relatively strong and significant regulation of aforementioned ion transporters (FC = 1.78,  $p < 1 \cdot 10^{-8}$  for *CFTR*, FC = -2.03,  $p < 1 \cdot 10^{-6}$  for *NBCe1*, FC = -1.33,  $p < 0.01$  for *NHE1* and FC = -3.18,  $p < 1 \cdot 10^{-8}$  for *DRA*). All in all, culturing under SWMS resulted in significant regulation of several key ion transporters in both HT29-MTX-E12 and Caco-2 cells.

### Lower glucose consumption and lactate production per cell in HT29-MTX-E12 cells grown under SWMS conditions

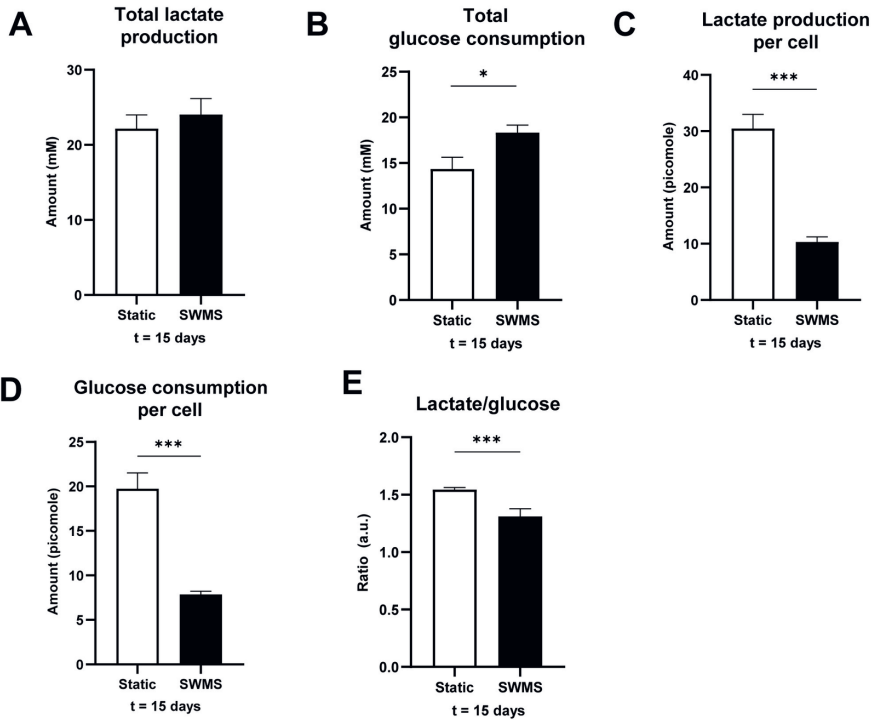
The observed visible color difference of cell medium in HT29-MTX-E12 cells, accompanied by significant regulation of  $H^+$  and  $HCO^-$  transporters, could indicate a difference in cell medium pH between growth conditions. The pH of the medium samples collected during every refreshment remained similar between the apical and basolateral compartment in both conditions. During the first 13 days, pH decreased to a similar extent in both conditions. Under SWMS conditions, the pH remained relatively stable over time after one week, whereas static conditions showed a decreased pH at  $t = 14$  days in the apical compartment (difference of 0.2) (**S7A-B Figure**). At  $t = 15$  days, one day after the last medium refreshment, trends reversed and medium of static conditions showed a significantly higher pH compared to SWMS conditions in both compartments (difference of 0.2) (**Figure 6A-B** and **S7A, C-D Figure**). A decrease in pH could be explained by increased lactate production and subsequent acidification of the medium. Indeed, pH of cell medium showed opposite trends to the amount of lactate measured in the medium, i.e., lactate concentrations increased with decreasing pH values and remained stable for SWMS conditions (**S8A Figure**). As medium volumes were, however, different between static and SWMS conditions, we calculated total lactate production and glucose consumption per well. Cell medium collected from cells grown under static conditions showed significantly higher lactate production at  $t = 7-14$  days (**S8B Figure**), accompanied with no difference in total glucose consumption per well (**S8C Figure**), indicating a lower amount of lactate produced per mole glucose under SWMS conditions. At  $t = 15$  days, one day after the last medium change (at which not all glucose had been consumed yet), lactate production was similar between conditions (**Figure 7A**). On the contrary, cells grown under SWMS conditions had consumed significantly more glucose (**Figure 7B**), again resulting in a lower lactate-per-glucose ratio under SWMS conditions. After correction for cell count at  $t = 15$  days, glucose consumption and lactate production per cell were still significantly lower under SWMS conditions (**Figure 7C-D**,  $p < 0.001$ ), resulting in a significantly lower lactate-per-glucose ratio (**Figure 7E**). Lower glucose consumption under SWMS conditions coincided with a significantly lower expression of *GLUT1/SLC2A1*, encoding a transmembrane glucose transporter ( $FC = -1.46$ ,  $p < 1 \cdot 10^{-7}$ ) and *HK2*, encoding the glycolytic enzyme hexokinase 2 ( $FC = -2.18$ ,  $p < 1 \cdot 10^{-8}$ ). GSEA revealed, however, no significant enrichment of the “Carbohydrate absorption pathway” or “Glycolysis pathway” in HT29-MTX-E12 cells cultured under SWMS conditions. Interestingly, however, both pathways were among the top-10 enriched downregulated in Caco-2 cells (NES “Carbohydrate absorption” = -2.35,  $p = 0.00$  and NES “Glycolysis pathway” = -2.12,  $p < 1 \cdot 10^{-3}$ , respectively). As shown previously for similar models, these data could point towards a more aerobic cell metabolism under SWMS conditions [561-565]. This is further supported by a significant downregulation of the “HIF1-signalling



pathway” in both HT29-MTX-E12 and Caco-2 cells, as revealed by GSEA (NES = -1.79, FDR q-value = < 0.05 and NES = -2.30, FDR q-value = 0.01, respectively).



**Figure 6. pH values of medium collected at t = 15 days from HT29-MTX-E12 cells grown under static and SWMS conditions.** Medium pH of **A)** apical and **B)** basolateral compartments of HT29-MTX-E12 cells grown under static and SWMS conditions at t = 15 days. \*  $p < 0.05$ ; \*\*  $p < 0.01$ ,  $n = 3$  biological replicates.



**Figure 7. Lactate production and glucose consumption by HT29-MTX-E12 grown under static and SWMS conditions.** **A)** Total lactate (micromole) produced per well in medium collected from apical and basolateral compartments of HT29-MTX-E12 cells grown under static or SWMS conditions, at  $t = 15$  days. **B)** Total glucose (micromole) consumed per well from medium collected from apical and basolateral compartments of HT29-MTX-E12 grown under static and SWMS conditions, at  $t = 15$  days. **C)** Glucose consumption (picomole) per cell by HT29-MTX-E12 grown under static and SWMS conditions at  $t = 15$  days. **D)** Lactate production (picomole) per cell by HT29-MTX-E12 grown under static and SWMS conditions at  $t = 15$  days. **E)** Ratio of lactate produced glucose consumed per cell in HT29-MTX-E12 grown under static and SWMS conditions at  $t = 15$  days. \*  $p < 0.05$ ; \*\*  $p < 0.01$ ; \*\*\*  $p < 0.001$ ,  $n = 3$  biological replicates.

## DISCUSSION

In the present study, we aimed to further characterize the potential mechanisms underlying the increased MUC2 production by the SWMS culture method as described by Navabi *et al.* [510]. To this end, we cultured HT29-MTX-E12 cells under both static and SWMS conditions and performed microarray analysis to investigate changes in gene expression by taking both a targeted and untargeted approach. First, we aimed to validate the increased MUC2 production under SWMS conditions. Indeed, SWMS conditions induced higher MUC2 expression in HT29-MTX-E12 cells, which was also reflected at the protein level. It seemed that the increase in MUC2 occurred at the expense of MUC5AC, since both gene and protein expression of this mucin was decreased. This resulted in a significantly increased MUC2/MUC5AC ratio, confirming previous findings [510]. An overview of the most prominent changes observed at  $t = 15$  days is graphically summarized (**Figure 8**).

The transcriptomic analysis of HT29-MTX-E12 cells revealed that SWMS conditions resulted in a strong upregulation of genes and pathways related to cell cycle regulation. In line with this finding, we found higher cell numbers after SWMS culturing. Although the effect size was lower, similar results were found in Caco-2 cells. Similar to our results, a previous study using ALI – which is similar to the semi-wet interface part of SWMS conditions – also showed an increase in cell number and cell layer thickness in intestinal porcine epithelial cells [560], and increased cell proliferation in intestinal organoids [593]. Our experiments can, however, not answer whether the higher cell count is due to an increased height and/or a columnar shape of cells grown under SWMS, as seen in both ALI and SWMS cultures [510, 560] or due to stacking of the cells, although the concentration of cells in the center of the well point to the latter. In any case, the shared upregulation of cell cycle-related pathways between colon carcinoma cell lines that are so different at the transcriptomic level [594], emphasizes the SWMS-specific effect on cell growth, independent of the cell type. As cell culturing conditions may also strongly influence cell morphology and tight junction formation [595], we explored if the SWMS method resulted in an increased resistance over the epithelial membrane by performing TEER measurements. TEER values of HT29-MTX-E12 cells were comparable to results from earlier studies [549, 596]. No differences were found between SWMS and static conditions for HT29-MTX-E12 cells during the culturing period, which was supported by the low number of differentially regulated genes related to intestinal barrier integrity. These findings are opposing the results from Navabi *et al.*, as they found a slight, but significant, increase in TEER values of HT29-MTX-E12 cells cultured under SWMS conditions. However, apart from the fact that measurements were performed at 21 days, Navabi *et al.* also used a different device (Ussing chamber) for TEER measurements. Furthermore, the higher concentration of cells in the center of the Transwell as a result of SWMS conditions may have resulted in an underestimation of TEER values, as the

chopstick electrodes do not measure the resistance over the whole membrane. In addition to TEER measurements, we also carried out FITC-dextran assays with two probe sizes. The significantly increased paracellular permeability of cells grown under SWMS compared to static conditions is remarkable. Since we do not know whether the mucus layer can interfere with FITC-dextran, we decided to also wash of the mucus layer with the mucolytic agent NAc. This pre-treatment resulted in an unchanged paracellular permeability between both conditions. Nevertheless, the seemingly increased mucus permeability under SWMS, as quantified using FITC-dextran probes, is interesting, but require further investigation with more robust probes to study mucus penetrability, e.g. fluorescent beads [239] that do not chemically interact with mucus.

In parallel to cell proliferation, we also investigated the potential change in cell differentiation pathways as a result of the SWMS method. We focused on the Notch/Atoh1 pathways, since these are key pathways in the decision between epithelial cell development into either absorptive or secretory cell types [572]. Given the MUC2-promoting effects of SWMS conditions in HT29-MTX-E12 cells, we hypothesized that SWMS conditions inhibited Notch and promoted Atoh1, thereby favoring the differentiation of secretory goblet cells. The interplay between Notch and Atoh in the context of goblet cell differentiation was underscored by the effect of the  $\gamma$ -secretase inhibitor DAPT, which further enhanced mucus production in HT29-MTX-E12 cells by indirectly inhibiting Notch, and thus promoting goblet cell differentiation [510]. Although we found a number of significantly differentially expressed target genes of both Notch and Atoh1, the results did not point towards one particular overrepresented pathway. Based on these results, we suppose that the increased MUC2 production in HT29-MTX-E12 cells was not the result of a change towards a favored secretory cell fate. This was further supported by a significant downregulation of the goblet cell marker KLF4 at both protein and gene level. The downregulation of *KLF4*/KLF4, identified as a cell cycle checkpoint protein and negative regulator of cell growth [597, 598], matches with the observed increased cell cycle regulation under SWMS conditions. However, based on our results, we cannot identify cause-effect relations and point at the exact trigger that led to decreased *KLF4*/KLF4 expression.

Proper mucus production depends on the activity of ion transport [587] and ALL-models with other cell types have shown changes in ion transport [563, 599-601]. Therefore, the significant regulation of several key ion transporters in both HT29-MTX-E12 and Caco-2 cells in our microarray analysis, is interesting. It should be considered, however, that the microarray results are limited to one time point; in this case at which a lower pH is measured under SWMS versus static conditions. Therefore, we cannot conclude whether the change in mucus phenotype can be (partially) explained by differential regulation of ion transporters, or that this regulation is a consequence of the microclimate at the timepoint of analysis. Moreover, our method



to measure pH did not allow us to measure potential local pH differences, as seen *in vivo* [602], which could further clarify the observed changes.

It has been demonstrated previously that both the semi-wet conditions/ALI and the mechanical stimulation part of SWMS separately, support a more aerobic cell culture environment in non-intestinal cell lines [562-565]. More recent studies confirmed increased oxygen supply and oxidative phosphorylation, concomitant with suppressed glycolysis in intestinal epithelial cells from porcine origin [560, 561]. Similar to the study by Klasvagt *et al.* [561], we measured significantly lower lactate levels and decreased glucose consumption in HT29-MTX-E12 cells grown under SWMS conditions. Interestingly, as opposed to Klasvagt *et al.*, we did not find a downregulation of the HIF-1 $\alpha$  gene, but GSEA revealed significant enrichment of the HIF-1 signalling pathway among the downregulated pathways in both HT29-MTX-E12 and Caco-2 cells. This, together with the significantly decreased lactate-per glucose ratio, suggests that under SWMS conditions, cells switch to a more aerobic cell metabolism. Increased mucus production has been observed in other ALI-models (without mechanical stimulation), such as murine gastric surface mucous cells [565] and airway cells from different origins [603].

Mechanical stimulation as part of the SWMS conditions, induced by continuous shaking at 65 rpm, led to a continuous exposure to shear stress. In a model highly similar to SWMS, the shear stress value was calculated to be  $1.6 \times 10^{-2} \pm 4.7 \times 10^{-3}$  dyne/cm<sup>2</sup> [604]. In other *in vitro* and *ex vivo* models, shear stress values between  $1.3$  and  $2.0 \times 10^{-2}$  dyne/cm<sup>2</sup> were sufficient to increase MUC2 protein or gene expression in Caco-2 cells, colon organoids or enteroids compared to static conditions [410, 605]. Additionally, a recent study demonstrated increased mucus thickness of HT29-MTX cells cultured under physiologically relevant shear stress ( $0.009$  dyne/cm<sup>2</sup> [410]), with changes in relative MUC2 and MUC5AC expression compared to static conditions over time [606]. Although we were not able to replicate MUC2-inducing effects in Caco-2 and our model cannot be directly compared to aforementioned models in terms of mechanical stimulation and culture time, we cannot exclude the possibility that shear stress played a role in the observed increase in mucus production under SWMS. Interestingly, next to its role in cell differentiation and cell cycle arrest [607], KLF4 is also known as a mechanosensitive transcription factor in vascular endothelial cells [608], osteoblasts [609] and a dermal cell line [610], in a context-dependent manner. To the best of our knowledge, there is no literature available providing evidence for a link between shear stress and KLF4 in intestinal tissues or models. Moreover, the relation between KLF4 and significantly regulated cytoskeleton-related and mechanosensitive genes (panels listed in S1 File) remains to be further investigated.

An important limitation of our study is that our microarray was restricted to a single time point, whereas our study aimed to capture an overview of the cellular

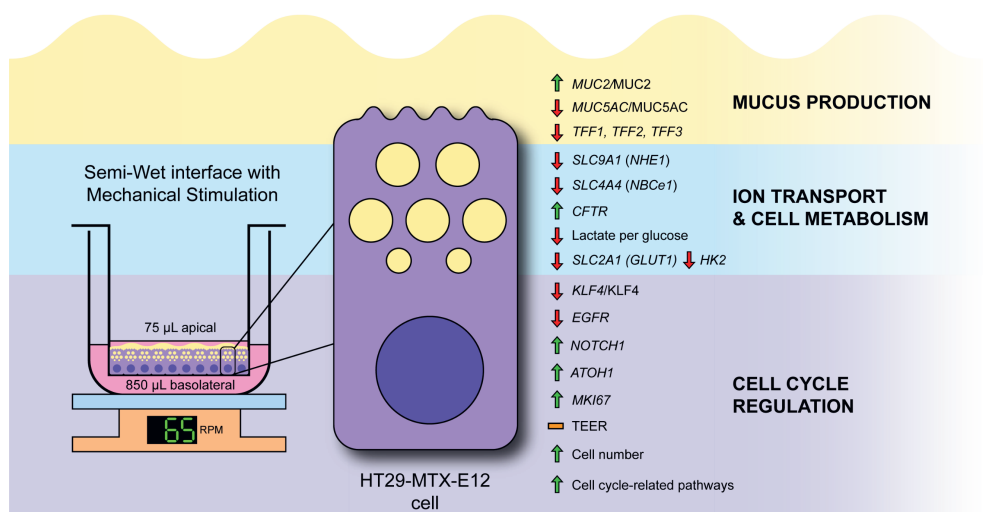
processes affected by the SWMS culture method. To increase insight in gene expression over time (e.g., with regard to cell proliferation or mucus production), future studies should include multiple time points. Another limitation is that gene expression does not always match with protein expression or activity, though protein levels of MUC2, MUC5AC and KLF4 supported the changes observed at gene levels in both HT29-MTX-E12 and Caco-2 cells. Apart from technical limitations of our study, we demonstrate that HT29-MTX-E12 cells grown under SWMS conditions, as a model developed to better represent the *in vivo* intestine in terms of MUC2 and cell morphology [510] still has its limitations. For instance, the downregulation of genes encoding other mucus-associated proteins, such as TFF3, suggest a less representative mucus layer in terms of whole mucus composition. Additionally, although we focused primarily on secreted mucins, transmembrane mucins also contribute to a healthy mucosal barrier and prevent the invasion of pathogens [611], amongst a myriad of other functions (reviewed in amongst others [612]. In our transcriptomic analysis, however, *MUC1* was not significantly regulated under SWMS conditions and *MUC4* was only expressed at low levels in HT29-MTX-E12 cells. Though, protein expression of these mucins, as well as their contribution to the model, would be worthwhile studying in future experiments. Still, the function of the SWMS-produced mucus layer as a physical barrier was demonstrated by its interference with the genotoxic activity of colibactin, a toxin produced by certain *Escherichia coli* strains [552]. In a similar fashion, the model could be applied to evaluate the diffusion of drugs and other particles [613]. Furthermore, the model could be useful to study interactions with all kinds of intestinal bacteria, but would further require quantification of the mucin glycans present, as these have demonstrated to play a crucial role in mucin-microbe interactions [614].

Altogether, we confirm the usefulness of SWMS cell culture conditions to improve the *in vivo* representation of the mucus layer *in vitro*, with regard to the increase in intestinal mucin MUC2. Our study provides insight in potential underlying processes, which might ultimately lead to a step-by-step improvement of the representativeness of the *in vitro* mucus layer. For instance, our study demonstrates upregulation of cell cycle regulation, downregulation of KLF4, differential regulation of ion transporters and increased aerobic metabolism of cells cultured under SWMS versus static conditions. Further research should also focus on the qualitative aspects of the *in vitro* mucus layer, for example with regard to mucin glycosylation, disulphide bonds that assure firmness of the mucus, and the contribution of other proteins present in the mucus layer, such as TFFs, as observed *in vivo*.





## Effects SWMS vs. static at t = 15 days



**Figure 8. Graphical summary of effects observed in HT29-MTX-E12 cells grown under SWMS compared to static conditions at t = 15 days.** HT29-MTX-E12 cells were grown on Transwell under static and SWMS conditions. The most important changes between static and SWMS at t = 15 days are depicted. MUC2/MUC2 = mucin 2; MUC5AC/MUC5AC = mucin 5; TFF1/2/3 = trefoil factor 1/2/3; SLC9A1 (NHE1) = Na<sup>+</sup>/H<sup>+</sup> Exchanger 1; SLC4A4 (NBCe1) = Na<sup>+</sup>/HCO<sub>3</sub><sup>-</sup> cotransporter 1; CFTR = Cystic Fibrosis transmembrane conductance regulator; lactate:glucose = amount of lactate (mole) produced per mole of glucose consumed in conditioned medium; SLC2A1 (GLUT1) = Glucose Transporter Type 1; HK2 = Hexokinase 2; KLF4/KLF4 = Kruppel like factor 4; EGFR = epidermal growth factor receptor; NOTCH1 = Notch receptor 1; ATOH1 = atonal BHLH transcription factor 1; MKI67 = marker of proliferation Ki-67; TEER = transepithelial electrical resistance.

## ACKNOWLEDGEMENTS

JE has been funded by a [Building Blocks of Life](#) grant from the Netherlands Organization for Scientific Research (NWO), grant no. 737.016.003. BvdL has been funded by the Nutricia Research Foundation, grant no. 2018-25 and the NWO Graduate Programme on Food Structure, Digestion and Health, grant no. 022.006.009.

## CONFLICT OF INTEREST

Authors declare no potential conflict of interest.

## SUPPLEMENTARY INFORMATION

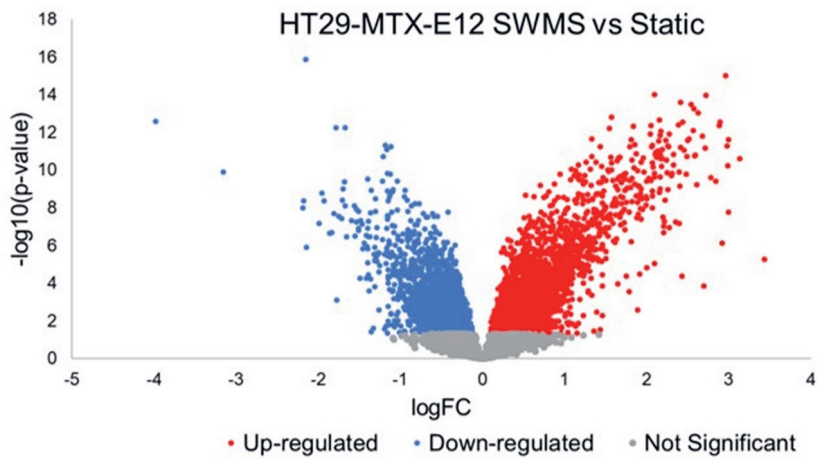
Available online (via <https://doi.org/10.1371/journal.pone.0261191>):

**S1 File. Panels of genes, including mucins, goblet cell markers, epithelial barrier-related genes, KLF4-target genes, ion transporters, cytoskeleton and mechanosensitive genes in HT29-MTX-E21 cells.** The RMA values of the three independent biological replicates are given, together with the corresponding fold changes and p-values between SWMS and static conditions.

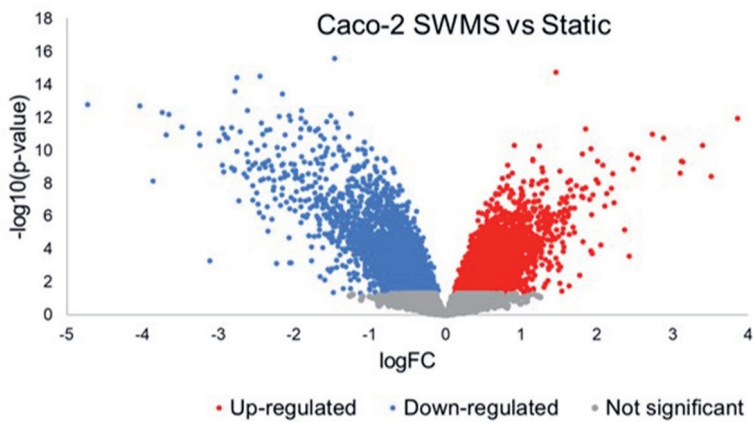
**S2 File. Up- and downregulated pathways between SWMS and static conditions in HT29-MTX-E12 and Caco-2 cells,** obtained by Gene Set Enrichment Analysis (GSEA). NES = Normalized Enrichment Score; FDR = False Discovery Rate. (XLSX)



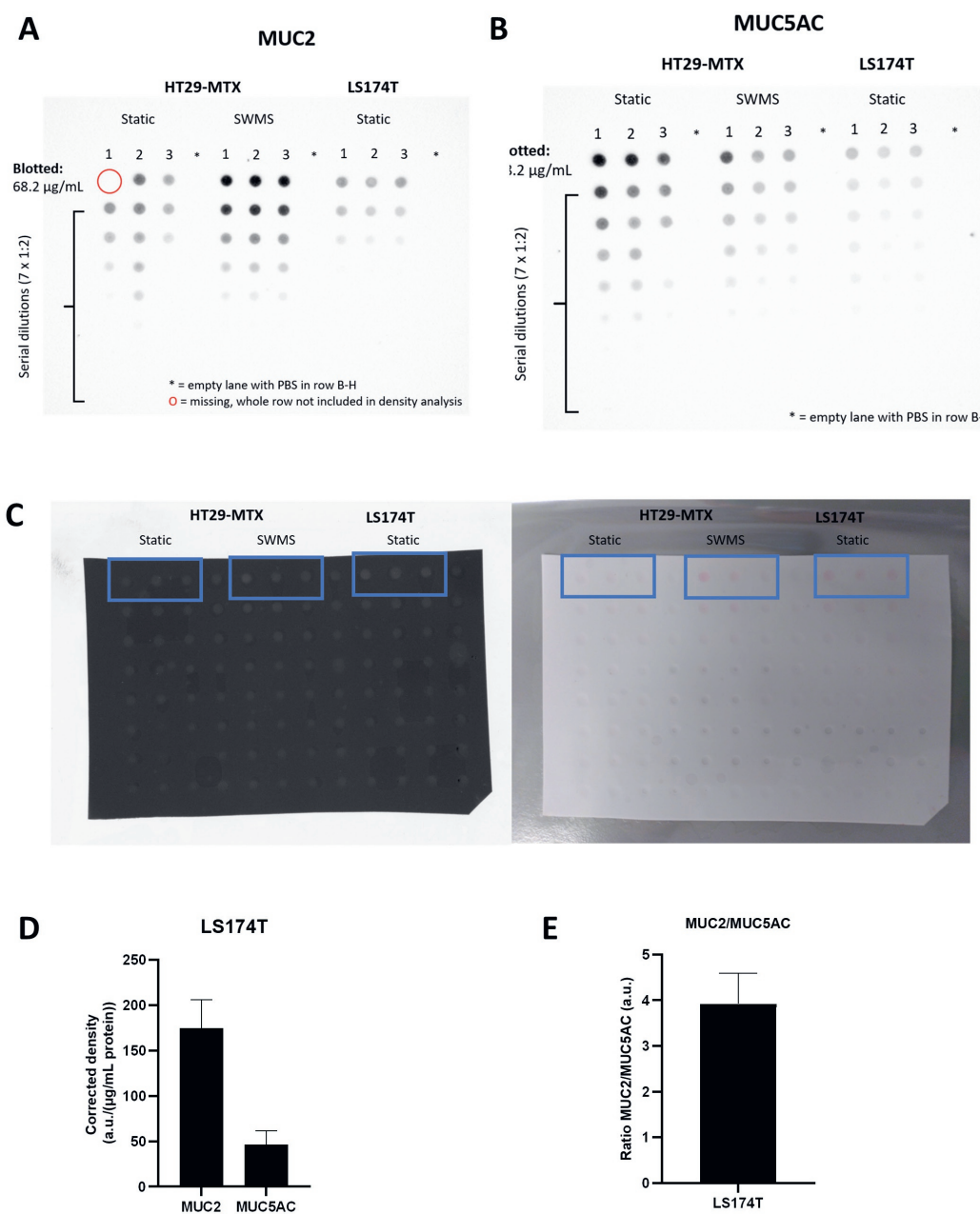
**A**



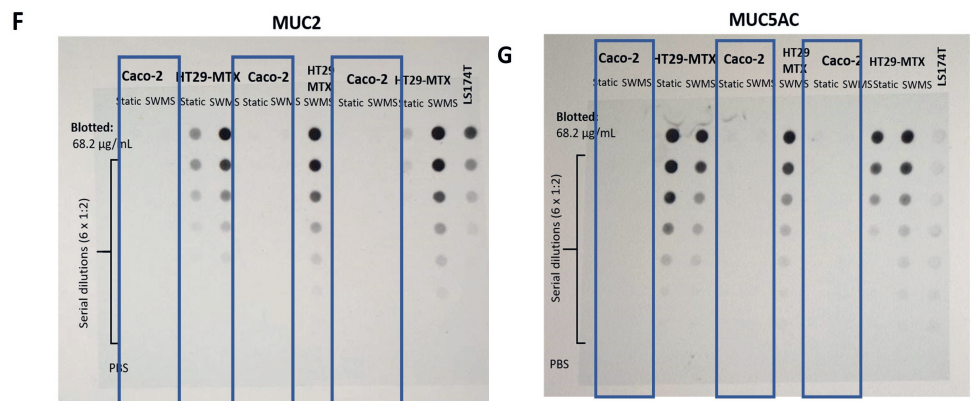
**B**



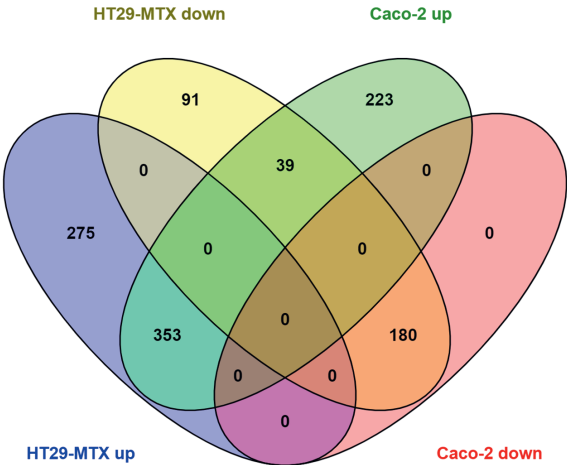
**S1 Figure. Overview microarray results in HT29-MTX-E12, Caco-2 and LS174T cells grown under static and/or SWMS conditions.** A) Volcano plot highlighting the Log Fold Change (logFC) on the x-axis and the corresponding p-values ( $-\log(10)$ ) on the y-axis for the comparison SWMS versus static conditions in HT29-MTX-E12 cells and B) Caco-2 cells.  $n = 3$ .



(cont. on next page)



**S2 Figure. Overview of Dot Blot results for A) MUC2 and B) MUC5AC in HT29-MTX-E12 cells cultured under static and SWMS conditions and LS174T cells.** A concentration of 68.2 µg/mL was blotted and seven times serially diluted 1:2. In column 4, 8 and 12, PBS was used as a negative control. C) Images of Ponceau Red staining (colorimetric and photographic) that were used as reference for total protein content. Protein density was based on the colorimetric image. D) Protein expression of MUC2 and MUC5AC in LS174T cells, expressed as density (a.u.) per ug/mL protein blotted, after correction of Ponceau Red density (n = 3) E) Ratio of MUC2 and MUC5AC protein expression in LS174T cells (n = 3). Original Dot Blots for Caco-2 are displayed in F) for MUC2 and G) for MUC5AC. Caco-2 data are indicated with a blue box. HT29-MTX-E12 and LS174T samples have been repeated including additional replicates for figure A and B.



Up = FC ≥ 1.5  
Down = FC ≤ -1.5

Top 20 upregulated genes shared between HT29-MTX and Caco-2

Gene name	FC in HT29-MTX	p-value	FC in Caco-2	p-value	Gene function
<i>RNY4P23</i>	10.80	0.00	2.12	0.05	RNY4 pseudogene 23
<i>KIF14</i>	8.00	0.00	1.52	0.01	kinesin family member 14
<i>KIF20A</i>	7.99	0.00	1.76	0.01	kinesin family member 20A
<i>DLGAP5</i>	7.91	0.00	2.04	0.00	DLG associated protein 5
<i>HMMR</i>	7.82	0.00	1.52	0.00	hyaluronan mediated motility receptor
<i>H4C1</i>	7.56	0.00	2.86	0.00	H4 clustered histone 1
<i>SPC25</i>	7.41	0.00	2.06	0.00	SPC25 component of NDC80 kinetochore complex
<i>H3C8</i>	7.20	0.00	2.17	0.00	H3 clustered histone 8
<i>TOP2A</i>	6.89	0.00	1.79	0.00	DNA topoisomerase II alpha
<i>CCNB1</i>	6.59	0.00	1.75	0.00	cyclin B1
<i>CENPF</i>	6.57	0.00	1.82	0.00	centromere protein F
<i>NDC80</i>	6.42	0.00	1.90	0.00	NDC80 kinetochore complex component
<i>H1-5</i>	6.04	0.00	2.57	0.00	H1.5 linker histone, cluster member
<i>ASPM</i>	6.02	0.00	2.02	0.00	abnormal spindle microtubule assembly
<i>PRR11</i>	5.97	0.00	1.52	0.00	proline rich 11
<i>CENPE</i>	5.95	0.00	1.88	0.00	centromere protein E
<i>NUSAP1</i>	5.82	0.00	1.72	0.00	nucleolar and spindle associated protein 1
<i>FOXM1</i>	5.72	0.00	1.81	0.00	forkhead box M1

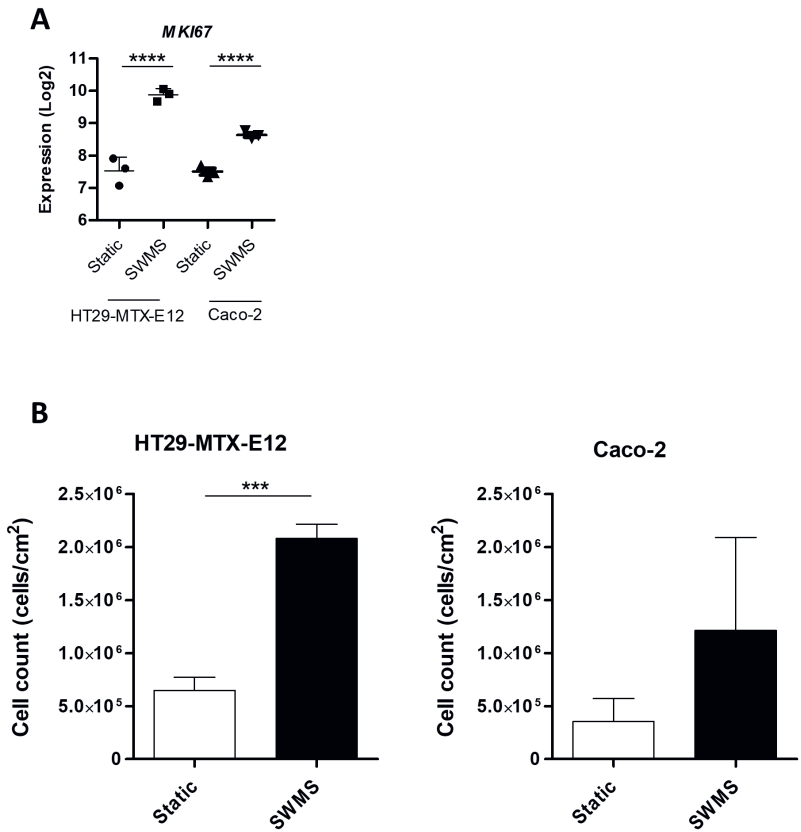
(cont. on next page)



## Top 20 downregulated genes shared between HT29-MTX and Caco-2

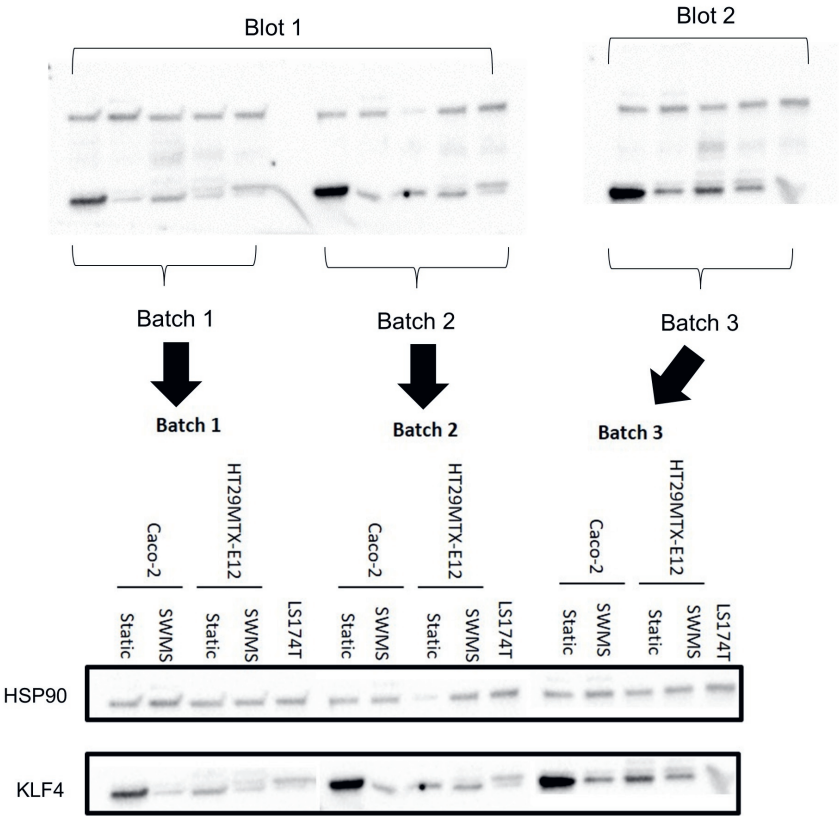
Gene name	FC in HT29-MTX	p-value	FC in Caco-2	p-value	Gene function
<i>SLC6A10P</i>	-1.51	0.01	-2.29	0.00	solute carrier family 6 member 10, pseudogene
<i>CD55</i>	-1.51	0.00	-2.11	0.00	CD55 molecule (Cromer blood group)
<i>HLA-C</i>	-1.51	0.00	-2.07	0.00	major histocompatibility complex, class I, C
<i>SLC46A3</i>	-1.51	0.00	-1.97	0.00	solute carrier family 46 member 3
<i>NRAD1</i>	-1.51	0.00	-1.64	0.00	non-coding RNA in the aldehyde dehydrogenase 1A pathway
<i>BCAT1</i>	-1.52	0.00	-2.60	0.00	branched chain amino acid transaminase 1
<i>ELF3-AS1</i>	-1.52	0.00	-1.58	0.00	ELF3 antisense RNA 1
<i>ARL14</i>	-1.52	0.01	-2.08	0.00	ADP ribosylation factor like GTPase 14
<i>MIR4268</i>	-1.52	0.01	-1.50	0.03	microRNA 4268
<i>MUC13</i>	-1.52	0.00	-3.20	0.00	mucin 13, cell surface associated
<i>ZNF625</i>	-1.52	0.04	-1.66	0.03	zinc finger protein 625
<i>VSIR</i>	-1.53	0.00	-1.71	0.00	V-set immunoregulatory receptor
<i>UACA</i>	-1.53	0.00	-1.53	0.00	uveal autoantigen with coiled-coil domains and ankyrin repeats
<i>TSPAN18</i>	-1.53	0.00	-3.25	0.00	tetraspanin 18
<i>EGFL7</i>	-1.53	0.00	-1.57	0.00	EGF like domain multiple 7
<i>PHKA2-AS1</i>	-1.53	0.01	-2.82	0.00	PHKA2 antisense RNA 1
<i>EPB41L1</i>	-1.54	0.00	-1.89	0.00	erythrocyte membrane protein band 4.1 like 1
<i>PTPRH</i>	-1.54	0.00	-1.60	0.00	protein tyrosine phosphatase receptor type H
<i>PDK1</i>	-1.55	0.00	-2.79	0.00	pyruvate dehydrogenase kinase 1
<i>LRP1</i>	-1.55	0.00	-1.66	0.00	LDL receptor related protein 1

**S3 Figure. Venn diagram showing the number of shared and unique differentially expressed up- and downregulated genes ( $-1.5 \leq \text{Fold change} \leq 1.5$ ) between HT29-MTX-E12 and Caco-2 cells cultured under static and SWMS conditions. The top 20 up- and downregulated genes shared between HT29-MTX-E12 and Caco-2 cells is given in the tables.**

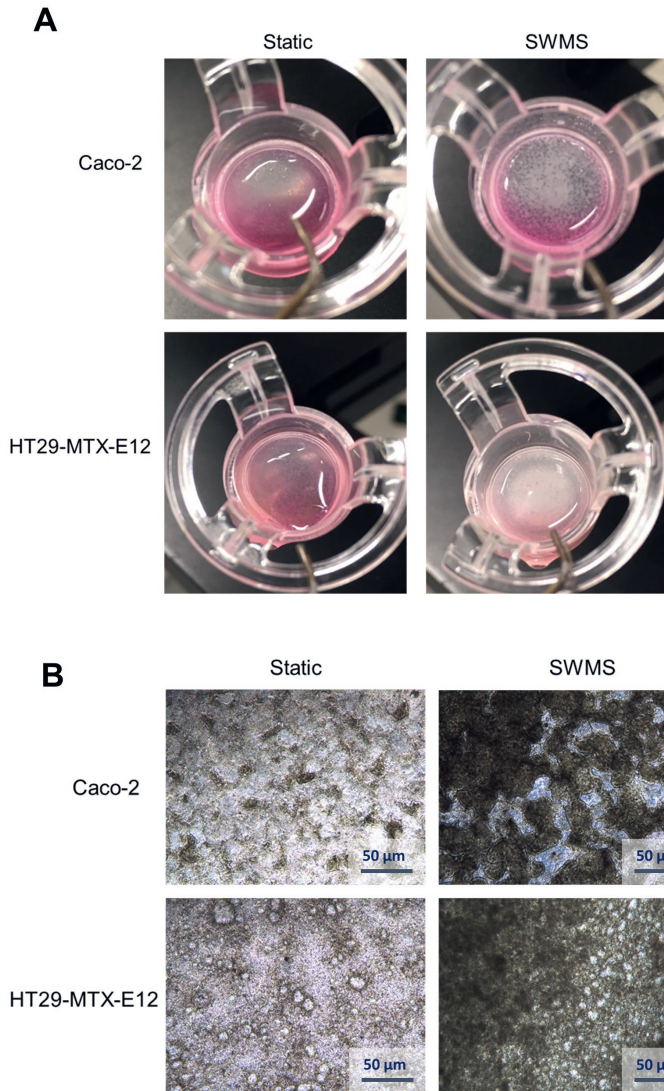


**S4 Figure. Overview of cell-cycle regulation parameters in HT29-MTX-E12 and Caco-2 cells under static and SWMS conditions.** A) Microarray gene expression values (Log2) of MKI67 in HT29-MTX-E12 and Caco-2 cells cultured under static and SWMS conditions. B) Cell count after  $t = 15$  days, expressed as cells per cm<sup>2</sup>, of HT29-MTX-E12 and Caco-2 cells cultured under static and SWMS conditions. \*\*\*\*  $p < 0.0001$ .



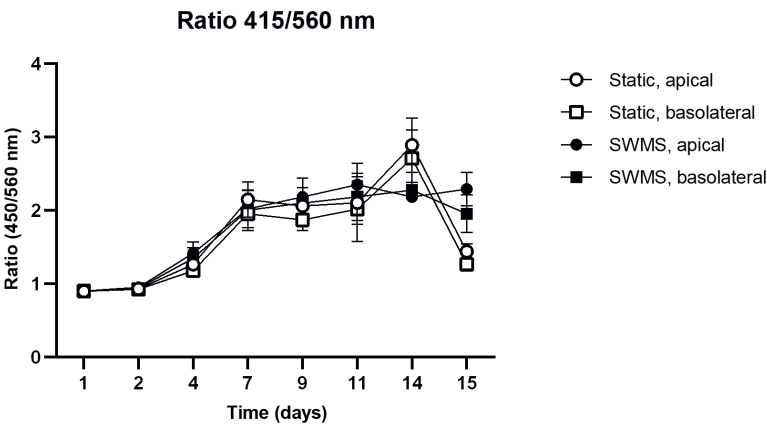


**S5 Figure. Western immunoblotting results of KLF4**, including all three biological replicates (batch 1, 2 and 3) in Caco-2, HT29-MTX-E12 and LS174T-cells grown under static and SWMS conditions or static only (LS174T). HSP90 was used as house-keeping protein.

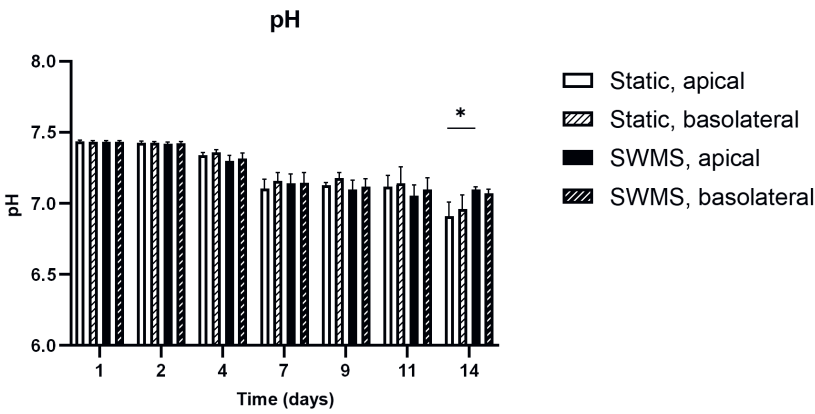


**S6 Figure. Images of HT29-MTX-E12 and Caco-2 cells grown under static and SWMS conditions.** A) Pictures of HT29-MTX-E12 and Caco-2 cells grown under static and SWMS conditions at  $t = 15$  days. B) Bright-field microscopy pictures (20x) of HT29-MTX-E12 and Caco-2 cells grown under static and SWMS condition, focussed on the centre of the Transwell membranes.

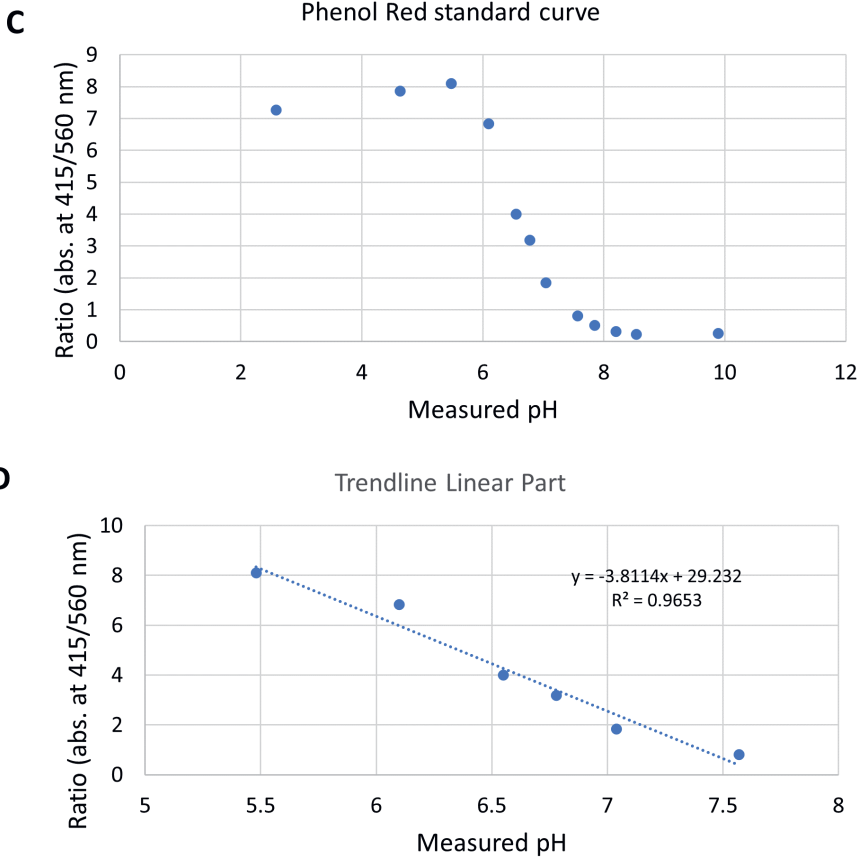
**A**



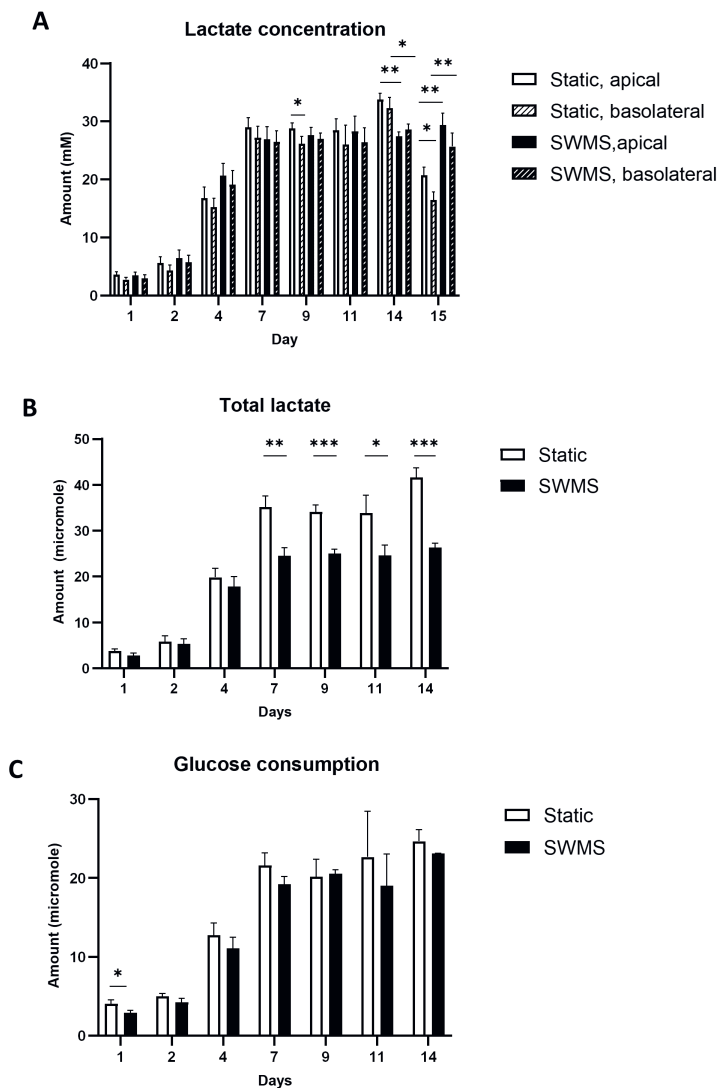
**B**



(cont. on next page)

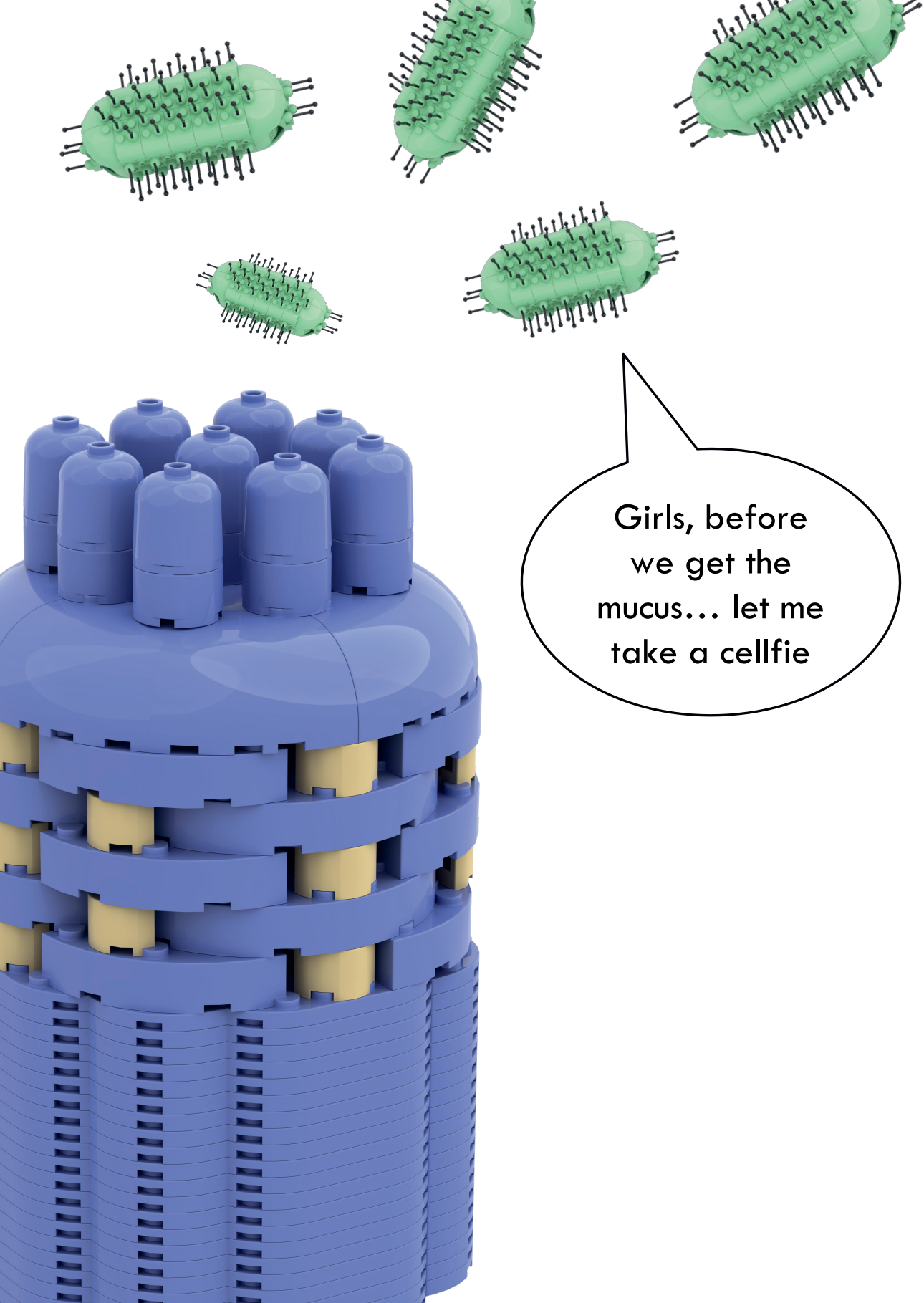


**S7 Figure. pH data and calibration of HT29-MTX-E12 and Caco-2 cells under static and SWMS conditions.** A) Ratio of Absorbance of cell culture medium (415 and 560 nm) of HT29-MTX-E12 and Caco-2 cells under static and SWMS conditions measured at 5% CO<sub>2</sub>. B) Medium pH of apical and basolateral compartments of HT29-MTX-E12 cells grown under static and SWMS conditions at  $t = 1-14$  days. \*  $p < 0.05$ ; \*\*  $p < 0.01$ ,  $n = 3$  C) Standard curve of pH values and absorbance values of cell culture medium measured at 415/560 nm at 5% CO<sub>2</sub>. D) The linear part the standard curve, including trendline.



**S8 Figure. Lactate and glucose concentrations of HT29-MTX-E12 cells grown under static and SWMS conditions.** A) Lactate concentration (mM) in cell culture medium of HT29-MTX-E12 cells collected during every medium refreshing moment. B) Total lactate (micromole) produced per well in medium collected from apical and basolateral compartments of HT29-MTX-E12 grown under static or SWMS conditions, at  $t = 1-14$  days. C) Total glucose (micromole) consumed per well from medium collected from apical and basolateral compartments of HT29-MTX-E12 grown under static and SWMS conditions, at  $t = 1-14$  days





Girls, before  
we get the  
mucus... let me  
take a cellfie

# CHAPTER 5

## The Effect of Mucus-Associated Bacteria on Intestinal Mucus Characteristics

### *In Vitro*

Janneke Elzinga<sup>1</sup>, Sharon Geerlings<sup>1</sup>, Esmée van Eck<sup>1</sup>, Guido Hooiveld<sup>2</sup>,  
Elsbeth Bossink<sup>3</sup>, Benthe van der Lugt<sup>2</sup>, Mathieu Odijk<sup>3</sup>, Loes Segerink<sup>3</sup>,  
Hauke Smidt<sup>1</sup>, Wilma Steegenga<sup>2</sup>, Clara Belzer<sup>1</sup>

<sup>1</sup>Laboratory of Microbiology, Wageningen University and Research,  
Wageningen, The Netherlands

<sup>2</sup>Nutrition, Metabolism and Genomics Group, Division of Human Nutrition and  
Health, Wageningen University & Research, Wageningen, The Netherlands

<sup>3</sup>BIOS Lab on a Chip Group, MESA+ Institute, Technical Medical Center, Max  
Planck Institute for Complex Fluid Dynamics, University of Twente, Enschede,  
The Netherlands

Manuscript in preparation



## The Effect of Mucus-associated Bacteria on Intestinal Mucus Characteristics *In Vitro*

### ABSTRACT

The human gastrointestinal tract (GIT) is covered by a mucus layer, which plays an important role in maintaining intestinal health. The mucus layer is thickest in the colon and is mainly composed of mucin, a heavily O-glycosylated protein. Amongst other functions, mucin provides an attachment site and nutrient source for mucus-associated bacteria, such as *Akkermansia muciniphila*. *A. muciniphila* is a next-generation beneficial microorganism, which is able to grow on mucin as the sole source of carbon and nitrogen, and its abundance in the human GIT has been correlated with a healthy intestine. Although the impact of this intestinal symbiont on the mucus layer has been studied *in vivo*, the exact effects and underlying mechanisms of this bacterium on mucus production and glycosylation remain understudied. In this study, we used HT29-MTX-E12 cells cultured under Semi-Wet Interface with Mechanical Stimulation to test the effect of pasteurized cells of *A. muciniphila* on mucus production *in vitro* by transcriptome analysis. In brief, we detected little to no effects of incubation with *A. muciniphila* under the studied conditions. This study provides a first step towards additional experiments to investigate microbe-mucus effects in different *in vitro* models of the human colonic mucus layer, including a gut-on-a-chip and cultivation in flasks. Potentially, these methods could serve as screening methods before investigation in more complex *in vitro* or *in vivo* models.

## INTRODUCTION

The human gastrointestinal tract (GIT) is covered by a mucus layer, which plays an important role in human intestinal health [78, 84]. In a healthy intestine, the mucus layer is thickest in the colon [72, 73], where it acts as a lubricant for food and protects the underlying intestinal epithelium from chemical, physical and biological hazards, while at the same time providing a colonization niche for bacteria [75, 104]. This dual role of the colonic mucus layer can be explained by its division into two layers: a densely packed inner layer, which is firmly attached to the intestinal epithelium and free of bacteria, and a loose outer layer, which is continuously replenished and provides a colonization site for certain bacteria [81, 84]. Several intestinal diseases have been linked to a compromised mucus layer [78]. Thus, the colonic mucus layer comprises an important site for host-microbe interactions.

Apart from lipids and defense-related proteins, colonic mucus consists mainly of the gel-forming mucin type 2 (MUC2), which is produced by intestinal goblet cells [79, 81, 97]. Like the other proteins in the family of mucins, MUC2 contains a domain of Tandem Repeats (TRs) of proline, threonine and serine, which are heavily O-glycosylated [82, 83]. Mucin O-glycosylation takes place in the Golgi and endoplasmic reticulum (ER) of intestinal goblet cells by the concerted action of glycosyltransferases: after addition of an O-linked N-acetylgalactosamine (GalNAc), mucin glycan chains are elongated with galactose, GalNAc and N-acetylglucosamine (GlcNAc), after which they are terminated with fucose and sialic acid [111]. Mucin O-glycans are generally accepted to play a major role in host-microbe interactions. Research on (ex-)germ-free mice demonstrated that bacterial colonization is required for mature intestinal mucin O-glycosylation and the development of an impenetrable inner layer [120]. Also, diseases associated with aberrant mucin O-glycosylation, including Inflammatory Bowel's Disease (IBD) [131-133] and colorectal cancer [135], have been linked to a deviant microbiota composition, but cause and effect still need to be disentangled. Amongst others, mucin glycans provide an attachment site for mucus-associated bacteria [138] and serve as a nutrient source for a select group of mucin-degrading bacteria that express a wide range of glycosyl hydrolases (GHs) and proteases [150, 615]. By degrading mucin glycans as well as the protein backbone, these bacteria liberate short-chain fatty acids for cross-feeding bacteria and the host [142, 145, 167]. In this way, these mucin-degrading bacterial species form the basis of many trophic chains in the human colon. A prominent example of a key mucin degrader is *Akkermansia muciniphila*, of which a large part of the genome is solely dedicated to mucin degradation [157, 159]. Moreover, its abundance is correlated with a healthy intestine [161, 162], and administration of both live and pasteurized *A. muciniphila* have been shown to influence metabolic parameters and reinforce the intestinal barrier [163-165]. Most importantly, the proven effectivity of



pasteurized *A. muciniphila* cells in mice and human interventions, allows these cells to be commercialized as they are safe for human consumption [616, 617].

The effect of individual mucus-associated species on mucus production has been investigated *in vivo*. Colonization of *A. muciniphila* was shown to promote mucus thickness in mouse models, but the exact molecular mechanisms remain to be elucidated [618, 619]. Additionally, acetate produced by *Bacteroides thetaiotaomicron* was shown to increase fucosylation and sialylation, while decreasing sulfated and neutral mucins in mono-associated rats [620]. Peptide-like molecules secreted by *Ruminococcus gnavus* increased the presence of GlcNAc and alpha-1,2-fucose residues in mono-associated mice [156]. Some of these effects have also been replicated *in vitro* using HT29-MTX cells [155, 156, 620]. This colon carcinoma cell line has been the model of choice to represent the colonic mucus layer *in vitro* [199]. Under static conditions it produces mainly stomach-specific MUC5AC, instead of MUC2 [621], but the cultivation of HT29-MTX under Semi-Wet interface with Mechanical Stimulation (SWMS) demonstrated enhanced MUC2 production *in vitro* [243]. Recently, we further characterized this model at the transcriptome level (**Chapter 4** of this thesis), but to our knowledge, the effect of mucus-associated bacteria on mucus production and/or glycosylation, have not yet been investigated in this model.

In this study, we aimed to investigate the effect of pasteurized *A. muciniphila* on mucus production and glycosylation *in vitro*, using HT29-MTX-E12 cells grown under SWMS. Overall, we are the first to investigate effects of pasteurized *A. muciniphila* on gene expression in this *in vitro* model. In addition, we make suggestions for future *in vitro* models to study microbe-mucus interactions, including mucin O-glycosylation.

## MATERIALS & METHODS

### Bacterial culture

*A. muciniphila* (ATCC BAA-835) was cultivated on minimal medium supplemented with threonine and a mixture of GlcNAc and glucose. The medium was composed as follows: resazurin (0.5 mg/L),  $\text{KH}_2\text{PO}_4$  (0.4 g/L),  $\text{Na}_2\text{HPO}_4 + 2 \text{H}_2\text{O}$  (0.669 g/L),  $\text{NH}_4\text{Cl}$  (0.3 g/L),  $\text{NaCl}$  (0.3 g/L),  $\text{MgCl}_2 \cdot 6 \text{H}_2\text{O}$  (0.1 g/L), L-threonine (6 g/L), trace elements in acid (1 mL/L of the solution containing 50 mM HCl, 1 mM  $\text{H}_3\text{BO}_3$ , 0.5 mM  $\text{MnCl}_2 \cdot 4\text{H}_2\text{O}$ , 7.5 mM  $\text{FeCl}_2 \cdot 4\text{H}_2\text{O}$ , 0.5 mM  $\text{CoCl}_2$ , 0.1 mM  $\text{NiCl}_2$  and 0.5 mM  $\text{ZnCl}_2$ , 0.1 mM  $\text{CuCl}_2 \cdot 2\text{H}_2\text{O}$ ) and trace elements in alkaline (1 mL/L of the solution containing 10 mM NaOH, 0.1 mM  $\text{Na}_2\text{SeO}_3$ , 0.1 mM  $\text{Na}_2\text{WO}_4$  and 0.1 mM  $\text{Na}_2\text{MoO}_4$ ). After autoclaving, the medium was supplemented with 1% of filter-sterilized vitamin solution (11 g/L  $\text{CaCl}_2$ , 20 mg biotin, 200 mg nicotinamide, 100 mg p-aminobenzoic acid, 200 mg thiamin (vitamin B1), 100 mg panthothenic acid, 500 mg pyridoxamine, 100 mg cyanocobalamin (vitamin B12), and 100 mg riboflavin), 25 mM of GlcNAc and glucose and 1 mL of reducing agent stock solution (80 g/L  $\text{NaHCO}_3$ , 4.8 g/L  $\text{Na}_2\text{S} \cdot 9 \text{H}_2\text{O}$  and 10 g/L L-cysteine-HCl).

*B. fragilis* (ATCC 25285) and *R. gnavus* (ATCC 29149) were cultivated on Brain heart infusion-supplemented (BHIS) medium. BHIS medium contained the following components: brain heart infusion broth (37 g/L), yeast extract (5 g/L), resazurin (1 mg/L), L-cysteine-HCl (0.5 g/L), hemin solution (10 mL/L) and vitamin K1 solution (0.2 mL/L). Hemin solution was prepared by adding 50 mg of hemin and 1 mL of NaOH to 100 mL distilled water. In addition, vitamin K1 solution was prepared by adding 0.15 mL of vitamin K1 to 30 mL of 95% EtOH.

All strains were cultivated anoxically in serum bottles with a headspace of mixed gas consisting of 80/20  $\text{CO}_2/\text{N}_2$ . The different media were inoculated with 1% (v/v) of glycerol stocks containing *A. muciniphila*, *R. gnavus* or *B. fragilis*. Inoculated bottles were then incubated at 37 °C (non-shaking) for 40 h. After the incubation, supernatant was separated by centrifuging the cultures repeatedly at 10,000 x g for 20 min at 4 °C until the supernatant was clear. Cells were resuspended in PBS at an OD corresponding to  $10^9$  CFU/mL (of live equivalent). Subsequently, the cells were pasteurized at 70 °C for 30 min. Bacteria were further diluted 10 x and 100 x (corresponding to an OD of  $10^8$  and  $10^7$  CFU/mL, respectively) in PBS and stored at -20 °C until use in experiments. Colony-forming units were based on OD<sub>600</sub>: For *A. muciniphila*,  $\text{OD}_{600} 1 = 3.6 \cdot 10^8$ . For *B. fragilis* and *R. gnavus*  $\text{OD}_{600} 1 = 2.4 \cdot 10^9$ . Effectivity of pasteurization was checked by plating samples on BHI/mucin (*A. muciniphila*) or BHIS (*B. fragilis* and *R. gnavus*) agar in duplicate, and incubating plates anoxically at 37 °C.



### Human cell culture

HT29-MTX-E12 cells (ECACC) were obtained from Sigma-Aldrich (Darmstadt, Germany). Cells were used between passage 5-20 and cultured in Dulbecco's Modified Eagle Medium with 4.5 g/L glucose, 110 mg/L sodium pyruvate and 584 mg/L L-glutamine (Corning, NY, USA) supplemented with 10% Fetal Bovine Serum and 1% penicillin/streptomycin at 37 °C at 5 % CO<sub>2</sub> until use in respective experiments. After reaching 80-90% confluence, cells were harvested using 10% trypsin/PBS solution. Cells were counted using the TC20 automated cell counter device (Bio-Rad, Veenendaal, The Netherlands).

### Transwell under SWMS

HT29-MTX-E12 were cultured under SWMS as described previously ([243] and **Chapter 4**). In short, 12 mm 0.4 µm-pore polyester Transwell membranes (Corning 3460) were seeded at 273,000 cells/mL, with 275 µL medium apically and 1 mL basolaterally ( $t = 0$ ). After one day, medium was replaced with 75 µL apically and 850 µL basolaterally, and plates were placed on an orbital shaker in a CO<sub>2</sub> incubator at 65 rpm. Medium was refreshed three times per week. At  $t = 15$  days, medium was replaced with 275 µL PBS with pasteurized bacteria ( $10^7$ - $10^9$  CFU/mL) or without bacteria. For one batch of Transwells exposed to *B. fragilis* and *R. gnavus*, cells were exposed at  $t = 14$  days. After 24 h of incubation, basolateral medium was collected and pooled. After washing twice with ice-cold PBS, cells were harvested for RNA isolation. Transwell experiments were performed three independent times (biological replicates, *i.e.*, cell passages). For each biological replicate, 1-3 technical replicates (*i.e.*, Transwells) were performed, but only one technical replicate was included for RNAseq.

### HPLC

Metabolites of intestinal cell lines were quantified using High-Performance Liquid Chromatography (HPLC) as described previously (**Chapter 4**). In short, medium incubated with cells was centrifuged for 15 min at  $> 10.000 \times g$ . Lactate and glucose were quantified with a Shimadzu LC2030C-Plus HPLC system equipped with a Shodex SH1821 column kept at 45 °C and 0.01 N sulfuric acid as eluent (1 mL/min). Compounds were detected by determining the refractive index and identified using pure compounds as external standards.

### pH measurement

pH of undiluted cell culture medium was measured as described previously (**Chapter 4**). In short, medium incubated with cells was centrifuged for 15 min at  $> 10.000 \times g$ . Next, 200 µL was transferred to a 96-well plate, and absorbance was measured at 415/560 nm at 5% CO<sub>2</sub> in a microplate reader (Biotek™ Synergy™ Neo2) after

calibration for > 1.5 h or until a stable signal was reached. As standard curve we used the one described in **Chapter 4**.

### RNA isolation

RNA was isolated with the RNeasy Mini Kit (Qiagen, cat nr. 74104) following the manufacturer's protocol. For Transwells, 200 µL RLT solution as provided with the kit was added to the cells and left for 5 min. Before the final spin-down, the column was incubated 1 min at 37 °C with Nuclease-free water. This step was repeated with flow-through. RNA yield was measured using the Qubit™ RNA Broad Range Assay kit (ThermoFisher). Samples were checked for DNA contamination using the Qubit™ DNA Broad Range Assay kit (ThermoFisher). RNA quality was determined with a Qsep100 Bio-Fragment Analyzer (Bioptic Inc., Jiangsu). RNA was only used for transcriptome analysis (see below) when the RNA integrity number (RIN) exceeded 8.0.

### Illumina Next Generation Sequencing

Total RNA was sequenced using Illumina sequencing technology (GenomeScan BV, Leiden, The Netherlands). The NEBNext Ultra II Directional RNA Library Prep Kit for Illumina was used to process the samples. Sample preparation was performed according to the protocol “NEBNext Ultra II Directional RNA Library Prep Kit for Illumina” (NEB#E7760S/L). Briefly, 1) rRNA was depleted from total RNA using the Qiagen fast select kit (#334376 and #335927). After fragmentation of the rRNA-depleted RNA, a cDNA synthesis was performed. 2) mRNA was isolated from total RNA using the oligo-dT magnetic beads. After fragmentation of the mRNA, a cDNA synthesis was performed. Products from 1) and 2) were used for ligation with the sequencing adapters and PCR amplification of the resulting product. The quality and yield after sample preparation was measured with the Fragment Analyzer. Clustering and DNA sequencing using the NovaSeq6000 was performed according to manufacturer's protocol. NovaSeq control software NCS v1.8 was used. Image analysis, base calling and quality check was performed with the Illumina data analysis pipeline RTA3.4.4 and Bcl2fastq v2.230. Quality-filtered sequence tags were provided as raw data.

### Transcriptome analysis

The quality of the raw data was checked using various metrics implemented in the RSeQC package (version v5.0.1) [622]. Reads were mapped to the GRCh38.p13 human genome assembly-based transcriptome sequences as annotated by the GENCODE consortium (release 41) [623] using Salmon version (1.9.0) [624]. The obtained transcript abundance estimates and lengths were imported in R using the package tximport (version 1.26.1) [625], scaled by average transcript length and library size, and summarized at the gene level. Differential gene expression was determined using limma (version 3.54.1) [626] utilizing the obtained scaled gene-



level counts. Briefly, before statistical analyses, non-specific filtering of the count table was performed to increase detection power [627], based on the requirement that a gene should have an expression level greater than 10 counts per million reads mapped, for at least six libraries across samples. Differences in library size were adjusted by the trimmed mean of M-values normalization method [628], implemented in the package edgeR (version 3.40.2) [629]. Counts were transformed to  $\log_2(\text{cpm})$  values and associated precision weights and entered into the limma analysis pipeline [630]. Differentially expressed genes were identified by using generalized linear models that incorporate empirical Bayesian methods [626, 631]. Genes were defined as significantly up- or downregulated when  $p/\text{FDR} < 0.05$  and fold change  $> 1.5$  or  $< -1.5$ . Volcano plots were generated using the package Enhanced Volcano (version 1.16.0). An overrepresentation analysis and gene set enrichment analysis (GSEA) were performed using KEGG (Kyoto Encyclopedia of Genes and Genomes) gene sets and Gene Ontology (GO) categories with clusterProfiler 4.0 [495]. The whole code will be made publicly accessible upon publication of the manuscript.

## RESULTS

### **Incubation with pasteurized mucus-associated bacteria has little effect on medium parameters of HT29-MTX-E12 grown under SWMS**

To investigate the effect of *A. muciniphila* on mucus production *in vitro*, HT29-MTX-E12 cells were grown on Transwells under SWMS, which has been shown to increase gene and protein expression of MUC2 [243]. After 15 days, cells were incubated apically for 24 h in the presence of  $10^7$ – $10^9$  CFU/mL pasteurized *A. muciniphila* in PBS or PBS only (control) under static conditions. As a comparison, 14–15-day-old HT29-MTX-E12 cells were incubated with pasteurized cells of two other mucin-degrading bacteria (*B. fragilis* and *R. gnavus*). After 24 h of incubation, spent cell culture medium was collected to quantify changes in pH and lactate/glucose [470]. No effect on lactate/glucose ratio (**Figure 1a**) or pH (**Figure 1b**) was observed upon co-incubation with any of the three bacterial species.

### **Incubation with pasteurized *A. muciniphila* has little effect on gene expression of HT29-MTX-E12 grown under SWMS**

Whole transcriptome analysis was performed on HT29-MTX-E12 cells incubated for 24 h with *A. muciniphila*. Bacteria were removed prior to RNA isolation. Overall, quality control of the individual samples revealed that only a low proportion (up to 33.8%) of the reads mapped to protein coding regions (**Supplementary Figure 1a**), suggesting contamination of samples with genomic DNA. This level of gDNA was, however, not expected to interfere with analysis of actual gene expression or only of genes with low counts [632], which were filtered out in the downstream analysis. Thus, we decided to continue with all samples, except those derived from HT29-MTX-E12 cells incubated with  $10^9$  CFU/mL *A. muciniphila* because of one outlying biological replicate with only 1.3% of reads mapping to coding regions of the human genome (**Supplementary Figure 1a**). PCA analysis as well as hierarchical cluster analysis revealed scattering of individual biological replicates, without clear clustering per condition (**Figure 1c**, **Supplementary Figure 1b**). These results suggest that pasteurized *A. muciniphila* has no to little biological effect on gene expression of HT29-MTX-E12 cells or that the experimental variation between biological batches was larger than any biological effect.

### **Significantly regulated genes show little overlap of genes and pathways between incubations with *A. muciniphila***

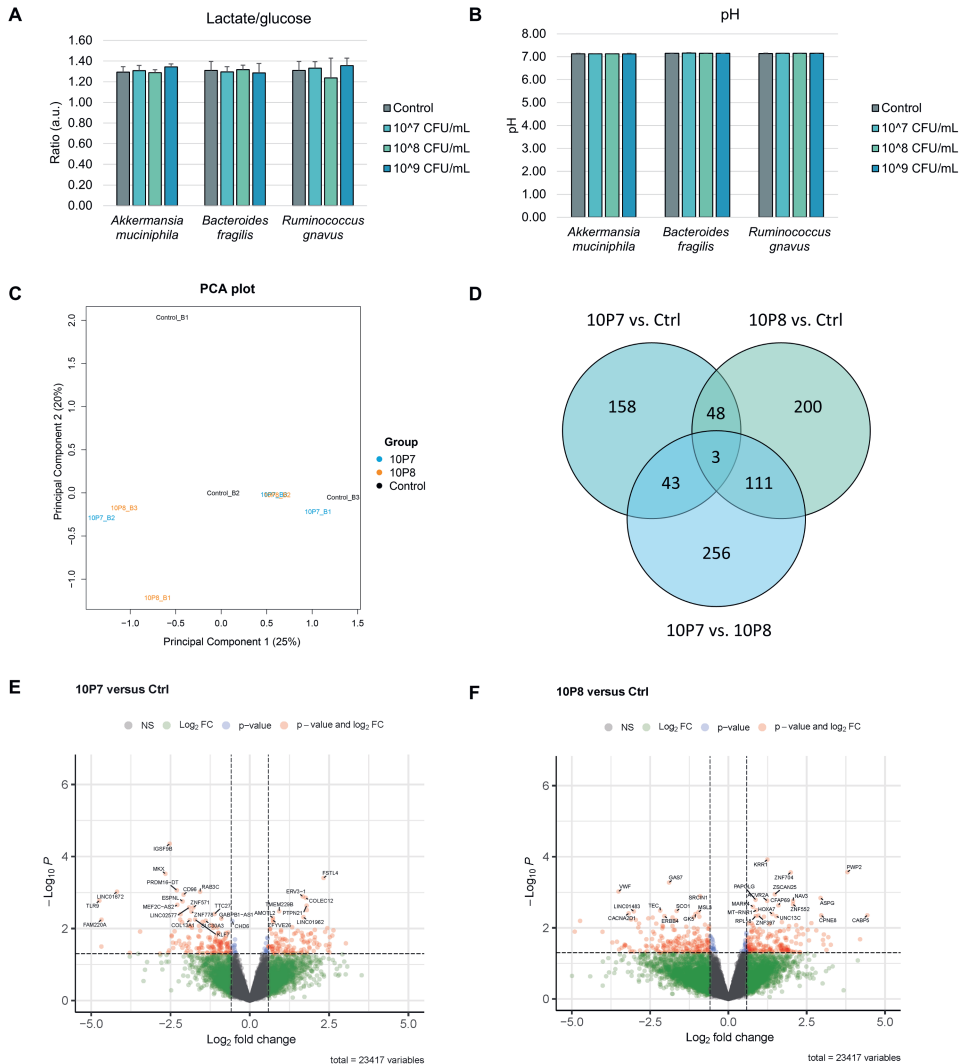
Despite large heterogeneity between biological replicates, for exploratory reasons, we decided to take a closer look at the genes significantly up- and downregulated ( $p < 0.05$ ) after incubation with  $10^7$  ("10P7") or  $10^8$  ("10P8") CFU/mL pasteurized *A. muciniphila*. Compared to the control condition ("Ctrl"), 252 and 282 genes were significantly regulated upon incubation with  $10^7$  and  $10^8$  CFU/mL, respectively, of which 51 shared by both concentrations (**Figure 1c-e**). Overrepresentation analysis





## CHAPTER 5

of the significantly regulated genes revealed very little overlap of KEGG-pathways or GO-Biological Processes (GOBP) categories between intestinal cells incubated with  $10^7$  or  $10^8$  CFU/mL *A. muciniphila* versus control conditions (**Figure 2** and **3**). None of the activated or suppressed KEGG-pathways or GO-BP categories were particularly associated with mucus production or glycosylation.



**Figure 1. Medium parameters and gene expression analysis of HT29-MTX-E12 incubated with pasteurized *A. muciniphila*.** **A)** Ratio of lactate and glucose and **B)** pH of medium obtained from HT29-MTX-E12 cells cultured under SWMS and incubated with 10<sup>7-9</sup> CFU/mL *A. muciniphila*, *B. fragilis* or *R. gnnavus*. Bars represent mean and SD of three biological replicates. **C-E** Transcriptome analysis of HT29-MTX-E12 cells incubated with *A. muciniphila*. **C)** PCA plot (PC1 and 2) of individual biological replicates of each condition tested. **D)** Venn Diagram of significantly regulated genes (FC > 1.5 or < -1.5 and p < 0.05) and Volcano plots showing log-fold changes in expression of individual genes, comparing incubation with 10<sup>7</sup> ("10P7") (**E**) and 10<sup>8</sup> ("10P8") (**F**) CFU/mL pasteurized *A. muciniphila* versus the negative control ("Ctrl"). Significantly regulated genes are marked in red.



**Figure 2. Overrepresentation analysis of significantly regulated genes of HT29-MTX-E12 incubated with pasteurized *A. muciniphila*.** KEGG pathways activated or suppressed upon incubation with 10<sup>7</sup> ("10P7") and 10<sup>8</sup> ("10P8") CFU/mL pasteurized *A. muciniphila* versus the negative control ("Ctrl").



**Figure 3. Overrepresentation analysis of significantly regulated genes of HT29-MTX-E12 incubated with pasteurized *A. muciniphila*.** GOBP categories activated or suppressed upon incubation with 10<sup>7</sup> ("10P7") and 10<sup>8</sup> ("10P8") CFU/mL pasteurized *A. muciniphila* versus the negative control ("Ctrl").

## DISCUSSION

In this study, we investigated the effect of a well-characterized mucus-associated bacterium on gene expression in an *in vitro* model of the human colonic epithelium, including a mucus layer. When exposing HT29-MTX-E12 cells under SWMS to pasteurized *A. muciniphila*, we observed large heterogeneity among biological replicates. The absence of a clear effect of *A. muciniphila* could be explained by the technical variation on both the side of the microbe and host, or the lack of effect of pasteurized bacteria on this specific cell line and/or these specific conditions. For instance, the contact with the bacteria might either have been too short or too long, or be prevented by the presence of the mucus layer. Lastly, the contamination of genomic DNA, probably due to the omission of DNase treatment in the RNA isolation protocol, could have obscured small biological effects of pasteurized *A. muciniphila* on HT29-MTX-E12.

Similar to our previous work (**Chapter 4**), the ratio of lactate per glucose and pH were quantified, as these were changed upon cultivation under SWMS versus static conditions [470]. Although bacteria were added under static conditions for 24 h, which might (partially) reverse the previously observed changes in pH and lactate/glucose, we thought these experimental outcomes would be specifically interesting in the context of *A. muciniphila*, since (live) bacteria were shown to enhance glucose metabolism in Caco-2 cells [633]. We did not observe an effect of any of the bacterial species examined here, but both outcomes are still a relatively easy method to further characterize the effects of bacterial factors, and, in some instances, might serve as quality check of an *in vitro* model of the human GIT.

Despite increased production of MUC2 and a thicker mucus layer [243], the SWMS method knows several limitations. First of all, cancer cell lines such as HT29-MTX(-E12) are known to produce only low amounts of core 3- and core 4-glycosylated mucins [634], which are prevalent in the healthy human intestine [113]. Secondly, the current model does not allow co-culture with live anaerobic bacterial species, limiting the investigation of the effect of “colonization” of bacteria of interest and their products on mucus production. HT29-MTX under SWMS have been employed to co-cultivate with (aerobic) *Escherichia coli* Nissle 1917, but only up to 4 h [635]. The investigation of live bacteria, however, is not a prerequisite, as pasteurized *A. muciniphila* is as beneficial as the live bacterium *in vivo* [163, 164] and is currently applied as a food supplement. In the near future, microfluidic devices are expected to facilitate long-term co-cultivation of mucus-associated bacteria with intestinal cells. The experiments described here are part of a bigger study, including a gut-on-a-chip on a rocking platform. The design of this device, including methods, are presented in the **Supplementary Methods**. In this set-up, a gut-on-a-chip is placed on a rocking platform with fluid-filled pipette tips in the in- and outlets of the upper and lower channel, representing the intestinal lumen containing epithelial cells and

blood compartment, respectively. By rocking the platform at a set interval, the fluid flows from one side to the other driven by gravity. In this way, the cells are not continuously exposed to shear stress, but to a peak shear stress, as reported previously [636]. The peak shear stress in the upper channel was calculated to be 10.18 dyne/cm<sup>2</sup> in the absence of cells. It would be interesting to compare transcriptomes of HT29-MTX-E12 cells grown on-chip versus under SWMS conditions, including the effect of pasteurized *A. muciniphila* and other mucus-associated species or their metabolites and (secreted) compounds. Similar to the SWMS method, the gut-on-a-chip on a rocking platform does not allow accurate control of shear stress values, but the model can be expanded by applying continuous (apical and basolateral) flow using a syringe pump and tubing. In addition, due its small size, this model has the advantage over the SWMS method that it can be used to study co-cultivation of HT29-MTX-E12 with live *A. muciniphila* under anoxic conditions. The (technical) application and challenges of this particular chip device and its implementation for host-microbe interaction studies, has been discussed in previous work as well [637].

The current study is limited to quantification of gene expression, which does not always reflect actual patterns of (posttranslational) mucin O-glycosylation. Although expression of glycosyltransferases was shown to match corresponding glycan profiles in murine embryonic stem cells [638], glycosylation capacities of individual cell types remain to be explored, stressing the need to capture the transcriptome, proteome and glycome at the same time [109, 639]. We aimed to quantify mucin glycosylation indirectly using lectin staining in cells seeded on 96-well plates, but the incubation of intestinal cells with *A. muciniphila* and subsequent binding of the bacteria (see **Chapter 6**) limits probing with lectins. Moreover, lectin staining does not allow fully quantitative description of the monosaccharides present. To this end, more advanced analyses, such as mass spectrometry, are required. In general, however, mucin purification from tissue or cell lines is a complex procedure and requires large amounts of mucus. To increase mucus yield *in vitro* for quantification of the glycan composition, HT29-MTX-E12 have been grown in flasks [640]. We slightly adapted this method, by cultivating HT29-MTX-E12 in flasks on an orbital shaker and exposing them to concentrations of mucus-associated bacteria ranging from 10<sup>6</sup> – 10<sup>8</sup> CFU/mL. The analysis of the mucus glycans collected from these intestinal cells is still ongoing.

In summary, this study aimed to assess the effect of a key mucin-degrader on mucus production, starting with whole transcriptome analysis of previously characterized HT29-MTX-E12 under SWMS, incubated with pasteurized *A. muciniphila*. This experiment provides a first step towards additional experiments to investigate the effects of mucus-associated bacteria in different *in vitro* models of the human colonic mucus layer, including a gut-on-a-chip and cultivation in flasks (ongoing). Potentially, these methods could serve as screening methods before investigation in more complex *in vitro* or *in vivo* models.



## SUPPLEMENTARY INFORMATION

## Supplementary Methods

*Fabrication of a gut-on-a-chip*

A gut-on-a-chip was fabricated based on previously described methods [637]. Two poly(methylmethacrylate) (PMMA) molds for top and bottom channels were designed in 3D-CAD software (SolidWorks®, 2018) and micromilled (Datron Neo, Germany). PDMS base and curing agent were mixed (10:1 %w/w, Sylgard 184 Silicone elastomer kit, Dow Corning) and degassed. PDMS was cast on the two PMMA molds and cured for 4 h at 60 °C. In- and outlets were punched in the PDMS top layer with a 1 mm biopsy punch (Ted Pella Inc, US). PDMS/toluene mortar of 0.7 g PDMS base agent, 0.07 g curing agent (Sylgard 184 Silicone elastomer kit, Dow Corning), and 270 µL of toluene (Sigma-Aldrich) was thoroughly mixed and spin coated (1500 rpm, 60 seconds) on a glass slide. Top and bottom PDMS surfaces were treated with oxygen plasma (Femto Science Cute, 40 seconds, 50 Watt) and then placed in the mortar. Three pre-cut PETE membranes (GVS Filter technology) were placed between the top and bottom channels, and the mortar was cured for 4 h at 60 °C. A picture of the gut-on-a-chip with a schematic overview of one combination of top and bottom channel is shown in **Supplementary Figure 2**. The top channel is 1 mm (h) × 1 mm (w) × 20.4 mm (l), whereas the bottom channel is 0.2 mm (h) × 1 mm (w) × 29.0 mm (l).

*Seeding and culture on the gut-on-a-chip*

Seeding and cultivation of the chip was adapted from a protocol described previously [637]. Before seeding, PBS and DMEM were pre-equilibrated at 37 °C at 5% CO<sub>2</sub>. Chips were treated with oxygen plasma for 2 min, after which channels were rinsed with 70% ethanol, followed by PBS. Channels were checked for air bubbles using bright field microscopy. Channels were filled with cell culture medium and incubated for at least 2 h at 37 °C at 5% CO<sub>2</sub>. HT29-MTX-E12 (passage 5-7) were harvested at 80-90% confluency using Trypsin-EDTA (0.05%), harvested and counted and diluted to 500.000 cells/mL. Cells were seeded on a chip by pipetting 5 µL of DMEM to the outlet of the channel and 35 µL of the cell suspension in the inlet of the channel. Chips were incubated for 1 hour at 37 °C at 5% CO<sub>2</sub>, after which medium was replaced using gravity-driven flow by replacing the pipet tips at the outlets with a pipet filled with 5 µL cell culture medium and a pipet with 200 µL at the inlets. The chip was incubated at 37 °C at 5% CO<sub>2</sub>. After one day, medium was replaced as described and the chip was agitated a rocking platform at 37 °C at 5% CO<sub>2</sub> with an angle of 25° and rocking at an interval of 8 min (OrganoFlow®, MIMETAS BV). Cells were cultured on chip for 7 days, while refreshing medium once per day as described. Cells on chips were checked regularly using brightfield microscopy. Medium in the upper channel was replaced by PBS for 24 h. After washing 2x with ice-cold PBS, cells were harvested for dual RNA and protein isolation.

Medium of top and bottom channels were pooled and collected for HPLC. Chip experiments were performed three independent times (biological replicates, i.e. cell passages). Since the cells were not continuously exposed to flow, a peak shear stress was calculated, as reported previously (e.g., [636]). The peak shear stress in the top channel was calculated at 1.018 Pa (= 10.18 dyne/cm<sup>2</sup>) and that in the bottom channel 0.34456 Pa (= 3.4456 dyne/cm<sup>2</sup>). The calculations are provided in Appendix A (without the presence of cells).

### Calculation of peak shear stress

#### Pressure difference

The pressure difference between the tips in inlet versus outlet can be calculated by

$$G \cdot \rho \text{ (density liquid)} \cdot \text{distance between inlets} \cdot \sin(\text{angle}).$$

$$G = 9.81 \text{ m/s}^2$$

$$\rho = 997 \text{ kg/m}^3$$

$$\text{Angle} = 25^\circ$$

$$\text{Distance top channel} = 20 \text{ mm} = 0.02 \text{ m}$$

$$\text{Distance bottom channel} = 29 \text{ mm} = 0.029 \text{ m}$$

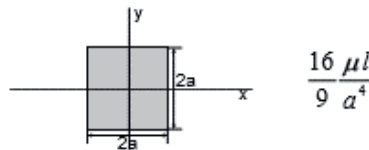
$$\text{Pressure difference top} = 9.81 \text{ m/s}^2 \cdot 997 \text{ kg/m}^3 \cdot (\sin(25)) \cdot 0.020 \text{ m} = 82.6 \text{ kg/(m} \cdot \text{s}^2) = \mathbf{82.6 \text{ Pa}}$$

$$\text{Pressure difference bottom} = 9.81 \text{ m/s}^2 \cdot 997 \text{ kg/m}^3 \cdot (\sin(25)) \cdot 0.029 \text{ m} = 82.6 \text{ kg/(m} \cdot \text{s}^2) = \mathbf{119.9 \text{ Pa}}$$

#### Resistance

The resistance of the square top channel can be calculated by

square



[proefschrift Edwin Oosterbroek; ISBN 90-36513464]

In which

$$\mu_{\text{medium}} = 7.8 \cdot 10^{-4} \text{ Pa/s}$$

$$l = \text{length} = 20.4 \text{ mm} = 2.04 \text{ cm}$$

$$\text{height (h)} = \text{width (w)} = 1 \text{ mm} = 0.1 \text{ cm}$$

$$a = 0.1 \text{ cm} / 2 = 0.05 \text{ cm}$$

$$\mathbf{R \text{ top channel}} = 16/9 \cdot (7.8 \times 10^{-4} \cdot 2.04 / 0.05^4) = \mathbf{452.6 \text{ [(Pa*s)/cm}^3\text{]}}$$

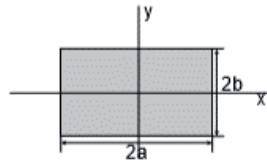
The resistance of the rectangle bottom channel can be calculated by





## CHAPTER 5

rectangle  
first order  
approximation



$$\frac{4\mu l}{ab^3} \left\{ \frac{16}{3} - 3.36 \frac{b}{a} \left( 1 - \frac{b^4}{12a^4} \right) \right\}^{-1}$$

$$\mu_{\text{medium}} = 7.8 \cdot 10^{-4} \text{ Pa}\cdot\text{s}$$

$$l = 2.9 \text{ cm}$$

$$h = 0.2 \text{ mm} = 0.02 \text{ cm}$$

$$w = 1 \text{ mm} = 0.1 \text{ cm}$$

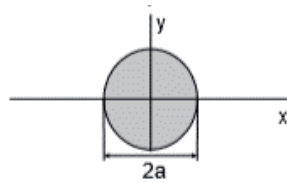
$$a = 0.2 \text{ cm} / 2 = 0.01 \text{ cm}$$

$$b = 0.1 \text{ cm} / 2 = 0.05 \text{ cm}$$

$$\mathbf{R \text{ bottom channel} = 38820.8 \text{ [(Pa} \cdot \text{s)/cm}^3\text{]}}$$

Resistance of pipette tip: can be calculated by

circle (full)



$$\frac{8\mu l}{\pi a^4}$$

*[proefschrift Edwin Oosterbroek; ISBN 90-36513464]*

Smallest diameter pipette tip: 0.5 mm

When taking a tip with diameter of 2 mm, resistance is approximately **100 (Pa · s)/cm<sup>3</sup>**

*Obtained peak flow rate*

$$P = R \cdot Q$$

$$Q \text{ maximal flow} = P/R =$$

$$Q_{\text{max top}}: 82.6 \text{ [Pa]} / 452.6 \text{ [(Pa} \cdot \text{s)/cm}^3\text{]} = 0.1825 \text{ [cm}^3\text{/s]} \text{ (1 cm}^3\text{ = 1 mL)}$$

$$\mathbf{Q_{\text{max top}} = 0.1835 \text{ mL/s} = 11.01 \text{ mL/min}}$$

$$Q_{\text{bottom}}: 119.9 \text{ [Pa]} / 38820.8 \text{ [(Pa} \cdot \text{s)/cm}^3\text{]} = 0.00308855 \text{ [cm}^3\text{/s]}$$

$$\mathbf{Q_{\text{max bottom}} = 0.00308855 \text{ mL/s} = 0.19 \text{ mL/min}}$$

*Obtained peak shear stress*

$$\tau = \frac{6\mu Q}{wh^2} \left( 1 + \frac{h}{w} \right) f^* \left( \frac{h}{w} \right)$$

*Formula from Van der Helm, 2016*

Shear stress top channel:

Symbol	Name	#	Unit
$\mu_{\text{medium}}$	dynamic viscosity	7.80E-04	Pa/s
Q	volumetric flow rate	0.1835	mL/s
w	width channel	0.1	cm
h	height channel	0.1	cm
f* gut	infinite summation series. Son et al for $(h/w) = 1$	0.5928	- For $(h/w)=1$

Peak shear stress top channel = **1.018 Pa = 10.18 dyne/cm<sup>2</sup>**

Shear stress bottom channel:

Symbol	Name	#	Unit
$\mu_{\text{medium}}$	dynamic viscosity	7.80E-04	Pa/s
Q	volumetric flow rate	0.00308855	mL/s
w	width channel	0.1	cm
h	height channel	0.02	cm
f* blood	infinite summation series. Son et al for $(h/w) = 0.2$	0.7946	- For $(h/w)=0.2$

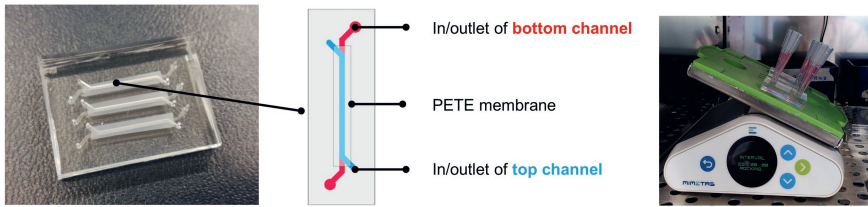
Peak shear stress bottom channel = **0.34456 Pa = 3.4456 dyne/cm<sup>2</sup>**



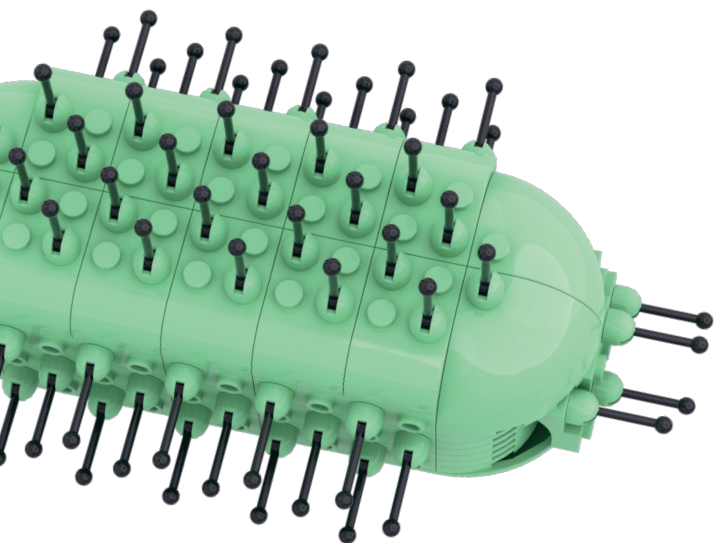
Supplementary Figures



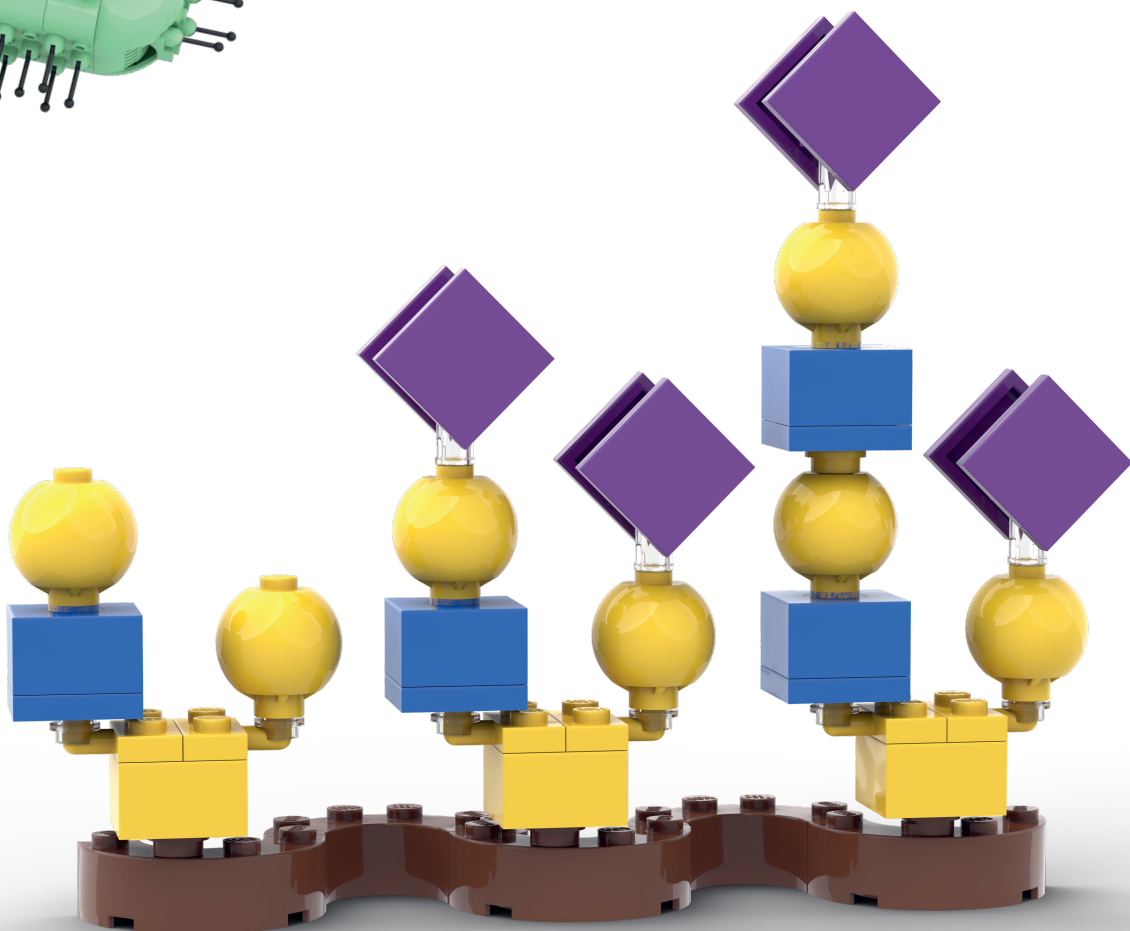
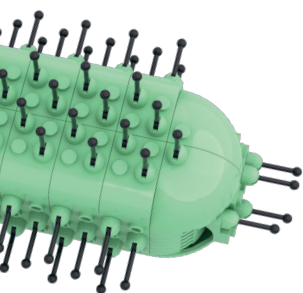
**Supplementary Figure 1. Sequencing analysis of HT29-MTX-E12 samples exposed to pasteurized *A. muciniphila*.** **A)** Assignment of bases (percentage) to regions in the reference genome. 10P7/8/9 represent the different concentrations of *A. muciniphila*. B1/2/3 represent the biological batches. **B)** Scree Plot visualizing the variation explained by 8 PCs.



**Supplementary Figure 2. Design of the gut-on-a-chip on a rocking platform.** Picture and schematic figure of the gut-on-a-chip device used in this study, including the final experimental set-up on a rocking platform.



Digital schematic  
reconstruction of  
*Akkermansia  
muciniphila*  
looking for a  
sweet spot on an  
O-glycosylated  
mucin molecule.



# CHAPTER 6

## Binding of *Akkermansia muciniphila* to Mucin is O-glycan Specific

Janneke Elzinga<sup>1</sup>, Yoshiki Narimatsu<sup>2</sup>, Noortje de Haan<sup>2</sup>, Henrik Clausen<sup>2</sup>, Willem M. de Vos<sup>1,3,\*</sup>, Hanne L.P. Tytgat<sup>1,4,\*</sup>

<sup>1</sup>Laboratory of Microbiology, Wageningen University & Research, Wageningen, The Netherlands

<sup>2</sup>Copenhagen Center for Glycomics, Department of Cellular and Molecular Medicine, Faculty of Health Sciences, University of Copenhagen, Copenhagen, Denmark

<sup>3</sup>Human Microbiome Research Program, Faculty of Medicine, University of Helsinki, Helsinki, Finland

<sup>4</sup>Current address: Nestlé Institute of Health Sciences, Nestlé Research, Lausanne, Switzerland

\*Shared last authors

Manuscript in preparation

## Binding of *Akkermansia muciniphila* to Mucin is O-glycan Specific

### ABSTRACT

The intestinal bacterium *Akkermansia muciniphila* is specialized in the degradation of mucins, i.e. the heavily O-GalNac-glycosylated proteins forming the major components of the mucus lining the intestine. This bacterium is the start of many trophic chains in the human intestinal ecosystem, as it can degrade both the mucin glycans and protein backbone. Little is known about the mucin-binding properties of *A. muciniphila* despite adhesion to mucins being crucial for its persistence within the human gastrointestinal tract. In this study, we showed that mucus binding of the anaerobic *A. muciniphila* is independent of the environmental oxygen concentrations, and we found both live and pasteurized cells to bind similarly to mucus. We further dissected binding of pasteurized *A. muciniphila* to mucins using recombinantly expressed human mucin tandem repeat (TR) reporter modules carrying distinct structures and patterns of O-glycans. Using these mucin reporters, we demonstrated that pasteurized *A. muciniphila* selectively recognized the LacNAc (Gal01-4GlcNAc $\beta$ 1-R) disaccharide on mucin O-glycans, which is abundantly found in core2 and core3 O-glycans, common in human colonic mucins. We further showed that desialylation by endogenous *A. muciniphila* sialidase activity promoted binding. In summary, our study provides novel insights into microbe-mucin interactions important for recognition and colonization by a key mucin-foraging bacterium.

## INTRODUCTION

The human gastrointestinal tract (GIT) is lined by mucus, which plays a key role in maintenance of intestinal health [78]. This barrier is mainly composed of mucins, a family of large, heavily O-GalNac-glycosylated proteins (*i.e.* sugars linked to serine (Ser) or threonine (Thr) residue) that are either expressed on the intestinal cell membrane (transmembrane mucins, including MUC1, 3, 4, 12, 13, 17) or secreted in the intestinal lumen (most commonly MUC2, but also MUC5AC, 5B and 6) [85]. When secreted, mucins bind water and form a protective, mesh-like structure expanding 100-1000 times in volume [107]. Consequently, this layer acts as a lubricant for food passage and serves as a physical barrier against pathogens and chemical damage of the intestinal epithelium. At the same time, especially in the colon, it provides a niche for certain commensal bacteria, which bind to mucins and mucin glycans [75]. O-glycans make up > 80% of the total mass of mucins [82], and they can serve as attachments and a nutrient source for bacteria, having a critical impact on their colonization [127, 142, 150].

The majority of mucin O-glycans is found on Ser/Thr residues in tandem repeated regions (TRs) comprised of PTS-rich (Proline-Threonine-Serine) domains, and these TRs vary in length, number, patterns of Ser/Thr O-glycosites among the different mucins [82, 83]. Mucin-type O-glycosylation takes place in the Golgi apparatus of mucus-producing goblet cells through the concerted action of host glycosyltransferases (GTs) (comprehensive overview in [109]). O-glycosylation is initiated by the addition of N-Acetylgalactosamine (GalNAc) to the protein backbone, which is then further elongated via the addition of galactose (Gal), GalNAc and N-Acetylglucosamine (GlcNAc), sialic acid and/or fucose, sometimes in a branched manner, resulting in complex glycan chains [111]. The exact glycan composition is dependent on the location in the GIT [112, 113] and host genetics (amongst others, blood group [118]). Moreover, diseases, such as colorectal cancer and inflammatory bowel disease, have been associated with aberrant glycan composition of mucin glycans [131-133, 135].

Commensal bacteria interact with mucins and mucin glycans provide a nutrient source for a select group of commensal, so-called mucolytic bacteria, like *Bifidobacterium* spp., *Ruminococcus gnavus*, *Bacteroides thetaiotaomicron* and *Akkermansia muciniphila* (reviewed in [150, 615]). Of these, *A. muciniphila* is recognized as a mucin-degrading specialist. This intestinal symbiont is not only able to cleave and degrade mucin glycans, but also the protein core of mucins, due to the expression of both a large set of glycoside hydrolases (GHs) and proteases [150, 615]. *A. muciniphila* plays a key role in the intestinal host-microbiome ecosystem at the mucosal interface by liberating monosaccharides from mucin glycans for itself or for cross-feeding with other bacteria, resulting in the production of short-chain fatty acids, and as such *A. muciniphila* provides the start of many trophic chains in the human GIT [167]. Moreover, the abundance of *A. muciniphila* in the human GIT has been





negatively correlated with a wide range of disorders, including obesity, diabetes, cardiometabolic diseases and low-grade inflammation (reviewed in [162]). This Gram-negative bacterium has also been shown to reinforce the mucosal barrier in mice and men by increasing the mucus layer thickness [618] and administration of both live and pasteurized *A. muciniphila* reversed high-fat diet-induced metabolic disorders [163-165]. Therefore, *A. muciniphila* is considered a promising next-generation beneficial microorganism, which is clinically substantiated by its ability to improve metabolic health of overweight and obese subjects [162, 166]. The shown equivalence in effectivity between live and pasteurized *A. muciniphila* in both mice and human interventions, further supports the probiotic potential of *A. muciniphila* as pasteurized cells are safe for human consumption [616, 617].

The mucin-binding properties of *A. muciniphila* have been studied using human colonic mucins obtained from healthy and diseased donors [641, 642] or mucin-producing cell lines [641, 643]. The mucin glycans in these models, however, were not characterized or were not representative of the mucins in the main colonization site of *A. muciniphila* – the human colon – where MUC2 mainly carries core3 O-glycans [113, 126]. Porcine Gastric Mucin (PGM) preparations are commonly used in studies because of its commercial availability. PGM, however, is not entirely representative of human colonic mucins due to its high proportion of core1 and core2 O-glycans low levels of sialic acids, and presence of MUC5AC and 6 [113, 126, 249, 250]. This gap between native human mucins and available animal mucins to characterize microbe-mucin binding properties, was recently bridged by the development of a cell-based platform for the display and production of representative mucin TR reporters with defined O-glycans [254]. These mucin TR domains (150-200 amino acids) are expressed in human embryonic kidney (HEK293) cells with engineered O-glycosylation capacities, conveying most of the informational content of the secreted or transmembrane mucin of interest [254]. Previous work demonstrated essentially complete occupancy of potential O-glycosites for most of the secreted and transmembrane mucins [644] and has showcased the potential to use these glycodomains to investigate the binding specificities of microbial and viral adhesins as well as substrate preferences of microbial glycopeptidases [254, 644, 645].

In this study, we aimed to investigate the binding requirements of *A. muciniphila*, using different TR domains with distinct glycan structures. Relying on the versatility, applicability and availability of these mucin TR glycodomains, we expanded our knowledge on host colonization, mucin recognition and mucin degradation by a symbiont that plays a key role in the intestinal host-microbe ecosystem.

## MATERIALS & METHODS

### Bacterial culture and pasteurization

*A. muciniphila* MucT (ATCC BAA-835) was cultured anaerobically in a basal medium as previously described [157]. The trace element and vitamin solutions were prepared as described previously [646]. The basal medium was supplemented with 20 g/L Tryptone (Oxoid™, ThermoFisher Scientific™), 4 g/L L-threonine (Sigma-Aldrich), 0.25% (w/v) Glc and 0.275% (w/v) GlcNAc (~25 mM total, Sigma-Aldrich) [647]. Incubations were done in serum bottles sealed with butyl-rubber stoppers at 37 °C under anoxic conditions provided by a gas phase of 182 kPa (1.8 atm) N<sub>2</sub>/CO<sub>2</sub> (80:20, v/v). Pilot experiments demonstrated slightly higher binding of bacteria in the stationary phase compared to the exponential phase (**Supplementary Figure S1**), so bacteria were harvested at end-exponential or stationary phase (~1 · 10<sup>9</sup> CFU/mL, optical density at 600 nm (OD<sub>600</sub>) of 2.8) by pelleting using multiple rounds of centrifugation. Supernatant was discarded and pellets were resuspended in 1 x Phosphate Buffered Saline (PBS) at 1 · 10<sup>9</sup> CFU/mL (OD<sub>600</sub> = 2.8). Bacterial cells were directly used for subsequent experiments ("live" fraction) or pasteurized for minimally 30 min at 70 °C in a temperature-controlled water bath [164]. Viability was tested by plating live and pasteurized fraction on Brain Heart Infusion agar supplemented with 0.5% crude PGM. If not directly proceeding to experiments, cells were stored at -20 °C.

### Production and purification of recombinant mucin TR reporters

A panel of HEK293-6E<sup>WT</sup> knock-out/knock-in (KO/KI) glycoengineered isogenic cells with different O-glycosylation capacities were used to produce secreted mucin TR reporters designed to include representative sequences (150-200 amino acids) from human MUC1, MUC2, MUC5AC and MUC7 TRs [254]. The following glycoengineered cells were used to produce different glycoforms: core2 (WT), T (KO GCNT1, KO ST6GALNAC2/3/4, KO ST3GAL1/2), Tn (KO C1GALT1), DiST (KO GCNT1) and core3 (KO COSMC, KI B3GNT6) (**Figure 2a**). Briefly, cells stably expressing mucin reporters were seeded at a density of 2.5 · 10<sup>5</sup> cells/mL in serum-free F17 culture media (Invitrogen) supplemented with 0.1% Kolliphor P188 (Sigma) and 4 mM GlutaMax at 37 °C and 5% CO<sub>2</sub> under constant agitation (120 rpm) and cultured for 5 days. Culture medium was spun down twice (1000 x g, 5 min and 3000 x g, 10 min), and mixed 3:1 (v/v) with 4x binding buffer (100 mM sodium phosphate, pH 7.4, 2 M NaCl), and run through a nickel-nitrilotriacetic acid (Ni-NTA) affinity resin column (Qiagen), pre-equilibrated with washing buffer (25 mM sodium phosphate, pH 7.4, 500 mM NaCl, 20 mM imidazole). After extensive washing with washing buffer mucin TR reporters were eluted with binding buffer containing 200 mM imidazole. Eluted fractions were analyzed by SDS-PAGE and fractions containing the mucin TR reporter were desalted followed by buffer exchange to MilliQ using Zeba



spin columns (ThermoFisher Scientific). Purified mucin TR reporters were quantified using a Pierce™ BCA Protein Assay Kit (ThermoFisher Scientific) following the manufacturer's instructions and evaluated by NuPAGE Novex Bis-Tris (4–12%, ThermoFisher Scientific) Coomassie blue analysis.

### Binding of *A. muciniphila* to mucins by ELISA

#### *Experiments using PGM*

ELISAs were performed as described previously [254] and adapted for evaluation of *A. muciniphila* binding. MaxiSorp 96-well plates (Nunc) coated with up to 500 ng of dilutions of ethanol-purified and dialyzed PGM type III (Sigma-Aldrich, [648]) or at least 100 ng/mL purified mucin TR reporters and incubated overnight at 4 °C in 50 µL carbonate-bicarbonate buffer (pH 9.6). Plates were blocked with 100 µL PLI-P buffer (0.5 M NaCl, 0.003 M KCl, 0.0015 M KH<sub>2</sub>PO<sub>4</sub>, 0.0065 M Na<sub>2</sub>HPO<sub>4</sub> · 2H<sub>2</sub>O, 1% Triton-X100, 1% BSA, pH 7.4) for 1 h at RT and incubated with live or pasteurized *A. muciniphila* up to 1·10<sup>9</sup> CFU/mL for 1 h at 4° C, unless described otherwise. After extensive washing with PBS containing 0.05% Tween-20 (PBS-T), plates were incubated with rabbit anti-*A. muciniphila* serum (kind gift of Dr J. Reunanen (University of Helsinki), 1:1000 in PLI-P) for 1 h at 4° C, followed by extensive washing and incubation with 1 µg/mL HRP-conjugated polyclonal goat anti-rabbit IgG (H+L) (Invitrogen) for 1 h at 4° C. The ELISA was then developed by addition of TMB substrate (ThermoFisher Scientific™) and stopped with 0.5 M H<sub>2</sub>SO<sub>4</sub> followed by measurement of absorbance at 450 nm (Agilent BioTek, Gen5 software). To test the effect of different temperatures (4 °C, RT and 37 °C), all incubation steps with bacteria or antibodies were carried out at this temperature. To test the effect of oxygen (at 37 °C), an anoxic environment created in a box with a Oxoid™ AnaeroGen™ sachet (Thermo Scientific). To test the effect of other compounds in solution during binding, *A. muciniphila* was resuspended in PBS with 0.1–1% PGM, 1% PEG 100 or 600 kDa (Merck) [649].

#### *Experiments using mucin TR reporters*

After optimization with PGM, ELISAs were repeated as described above, in which purified mucin TR reporters were coated up to 25 ng and incubated with pasteurized 5 · 10<sup>8</sup> CFU/mL *A. muciniphila*. All incubation steps were performed at 4 °C. The following lectins were used as binding references: 0.05 µg/mL *Ricinus Communis* Agglutinin I (RCA-I, Vector Laboratories), 0.1 µg/mL Peanut Agglutinin (PNA) or 2.0 µg/mL Pan-Specific Lectenz (Lectenz Bio) as detection probe and 1 µg/mL HRP-conjugated streptavidin (Merck) for signal development. For quantification of coating efficiency of the mucin TR reporters with FLAG tags 0.1 µg/mL anti-FLAG M2-Peroxidase-HRP-conjugated mAb (Sigma) was used. For experiments using sialidases, the proteins Amuc\_0625 and Amuc\_1835 were produced with C-terminal His-Tag as described previously [650]. Sialidase activity was quantified in phosphate buffer, pH

6.0 using 1 mM MU-NANA (2-(4-Methylumbelliferyl)- $\alpha$ -D-N-acetylneuraminic acid) (Sigma-Aldrich) and a standard curve of 0.25-0.625 mM MU (4-methylumbelliferone) (Sigma-Aldrich) (**Supplementary Figure S2**). Mucins were incubated with 2.5 mU sialidase from *Clostridium perfringens* (Sigma-Aldrich), 1.4 mU Amuc\_0625 or 0.5 mU Amuc\_1835 in 20 mM sodium acetate buffer, pH 6.0 for 4 h, or buffer only.

### Glycoprofiling by MS

MUC2 and fetuin O-glycans were released, derivatized, purified and analyzed by C18 nanoflow liquid chromatography (LC) coupled to mass spectrometry (MS) as described previously [651]. For each sample, 1 to 2  $\mu$ g protein was mixed with 25  $\mu$ L release reagent (20% hydroxylamine and 20% 1,8-diazabicyclo(5.4.0)undec-7-ene (DBU)) and incubated for 1 h at 37 °C. Released O-glycans were enriched by hydrazide beads and labeled with 50  $\mu$ L 2-aminobenzamide (2-AB) reagent (500 mM 2-AB, 116 mM 2-methylpyridine borane complex (PB) in 45:45:10 methanol:water:acetic acid) for 2.5 h at 50 °C. Labeled glycans were purified by HILIC and porous graphitized carbon (PGC) SPE. Samples were resolved in 20  $\mu$ L water for MS analysis.

For each sample, 2  $\mu$ L was injected for nanoLC-MS/MS analysis, using a single analytical column setup. The analytical column was prepared using a PicoFrit Emitter (New Objectives, 75  $\mu$ m inner diameter), packed with Reprosil-Pure-AQ C18 phase (Dr. Maisch, 1.9  $\mu$ m particle size, 22-25 cm column length). The emitter was interfaced to an Orbitrap Fusion Lumos mass spectrometer (Thermo Fisher Scientific) via a nanoSpray Flex ion source. Samples were eluted in a 30 min method with a gradient from 3% to 45% of solvent B in 15 min, from 45% to 100% B in the next 5 min and 100% B for the last 10 min at 200 nL/min (solvent A: 0.1% formic acid in water; solvent B: 0.1% formic acid in 80% ACN). A precursor MS scan ( $m/z$  200-1700, positive polarity) was acquired in the Orbitrap at a nominal resolution of 120,000, followed by Orbitrap higher-energy C-trap dissociation (HCD)-MS/MS at a nominal resolution of 50,000 of the 10 most abundant precursors in the MS spectrum (charge states 1 to 4). A minimum MS signal threshold of 30,000 was used to trigger data-dependent fragmentation events. HCD was performed with an energy of 27%  $\pm$  5%, applying a 20 s dynamic exclusion window with a mass tolerance of 25 ppm. Structural annotation and relative quantification of the O-GalNAc glycans was performed as described before using the Minora Feature Detector node in Thermo Proteome Discoverer 2.2.0.388 (Thermo Fisher Scientific Inc.), GlycoWorkbench 2.1 (build 146) [652] and the Thermo Xcalibur qual browser 3.0.63. Glycan structure annotation was based on literature [651] and MS/MS analysis.

### Cell-binding assays of *A. muciniphila* to cells transiently expressing mucin TR reporters

Transmembrane GFP-tagged mucin TR reporters were transiently expressed in engineered HEK293-6E cells and used for flow cytometry as described previously



[254]. All WT and isogenic HEK293 cells (WT, Tn (KO *C1GALT1*), Core3 (KO *COSMC*, KI *B3GNT6*) and  $\Delta B4GALT1/2/3/4$  (KO *B4GALT1/2/3/4*) were cultured in DMEM (Sigma-Aldrich) supplemented with 10% heat-inactivated fetal bovine serum (Gibco) and 2 mM GlutaMAX (Gibco) in a humidified incubator at 37 °C and 5% CO<sub>2</sub>. Cells were seeded on 24-wells (NUNC) and transfected with 0.5 µg of plasmids using Lipofectamine 3000. At 24 h post-transfection, the cells were incubated with 10 mU sialidase of *C. perfringens* or PBA only for 1 h at 37 °C and further probed with  $0.625 - 5.0 \cdot 10^8$  CFU/mL *A. muciniphila* on ice or at 4 °C for 1 h, followed by incubation with anti-*A. muciniphila* serum (1:1000) and cross-absorbed Alexa Fluor™ 647 -conjugated goat anti-rabbit IgG (2 µg/mL, Invitrogen) for 1 h. To check desialylation, cells were incubated with 2.0 µg/mL biotinylated SiaFind™ Pan-Specific Lectenz® (Lectenz Bio) pre-incubated with 2 µg/mL Alexa Fluor™ 647-conjugated streptavidin. For normalization, cells were incubated with 0.2 µg/mL APC-conjugated rat IgG2a λ anti-FLAG (Biolegend). All cells were resuspended in PBA for flow cytometry analysis (SONY SA3800). Median fluorescent intensity (MFI) for all cell populations was quantified using FlowJo software (FlowJo LLC).

### Statistical analysis

Data are represented as the mean of three biological replicates, each representing at least two technical replicates, unless stated otherwise. For binding of *A. muciniphila* to purified mucins, a biological replicate represents one independent culture of *A. muciniphila*. For flow cytometry experiments, a biological replicate represents a cell population derived from an independent cell passage. A student's t-test was performed to assess differences in means between conditions, unless stated otherwise. Graph and bar chart figures were generated using GraphPad Prism 5.0 and Excel.

## RESULTS

### Pasteurized *A. muciniphila* binds to PGM in a concentration-dependent manner

To measure the binding of *A. muciniphila* to mucin, we established and optimized an ELISA-based assay using PGM. *A. muciniphila* was grown on a minimal medium supplemented with Glc and GlcNAc, which was previously shown to be indispensable for rapid growth [647]. Binding was tested in oxic conditions at different temperatures (4 °C, RT and 37 °C) as well as anoxic conditions at 37 °C to mimic the environment of the human colon. *A. muciniphila* was found to bind to PGM in a concentration-dependent manner, plateauing between 50 and 500 ng PGM. The presence of oxygen did not affect binding of *A. muciniphila*, which is in accordance with earlier results showing that *A. muciniphila* can survive in certain oxic conditions up to at least 10 h [160]. Interestingly, binding was stronger at 4 °C and RT compared to 37 °C (**Figure 1a**). To further exclude signal reduction of primary binding due to mucin degradation, binding of pasteurized *A. muciniphila* was tested at varying concentrations of PGM and bacteria, which showed a similar binding efficiency, kinetics and concentration-response as live cells (**Figure 1b**, **Supplementary Figure S3a**).

We tested if PGM in solution could scavenge binding of *A. muciniphila* to mucin. Indeed, already at 0.1% (w/v) PGM could inhibit binding of *A. muciniphila* cells to coated PGM. This effect was not observed for compounds with higher biomass (PEG 100 and 600 kDa) or similar viscosity (PEG 600 kDa) [649] (**Figure 1c**) and was replicated for live bacteria (**Supplementary Figure S3b**). Consequently, as a mucin-based medium thus could potentially interfere with the assay, all subsequent ELISAs were performed with pasteurized bacteria grown on minimal medium supplemented with Glc and GlcNAc, and carried out under an oxic atmosphere at 4 °C.

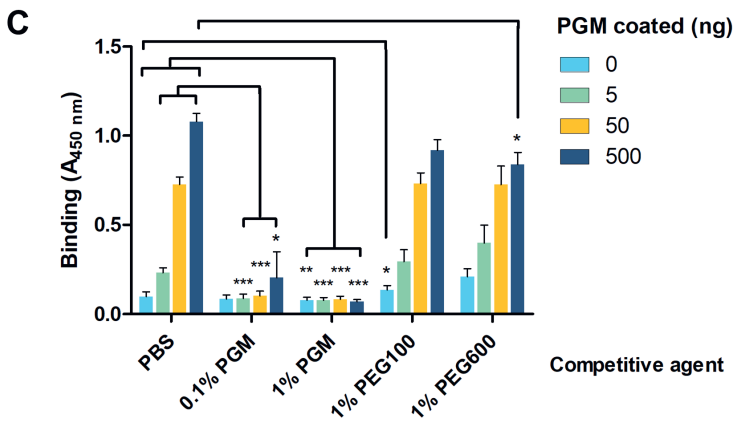
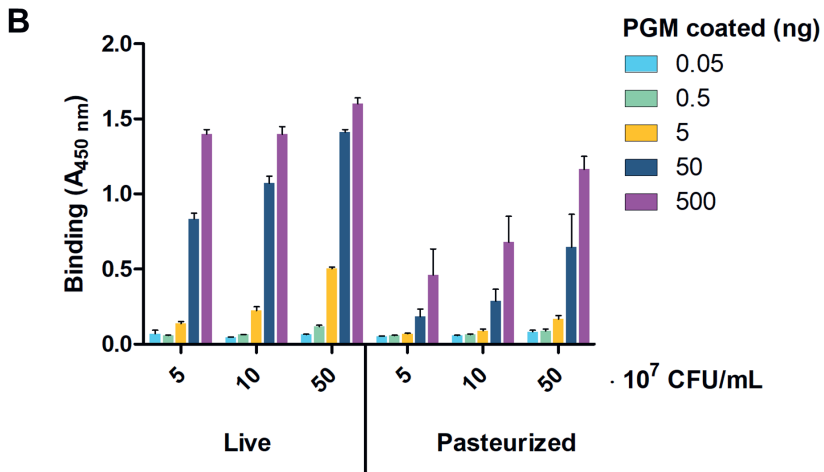
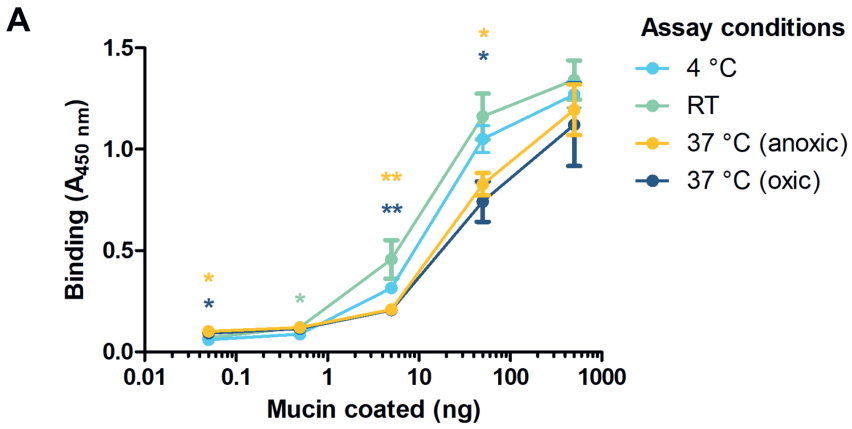
### Binding of *A. muciniphila* to mucin depends on the presence of LacNAc

To rationally dissect the binding of *A. muciniphila* to human mucins, we used a previously developed Mucin Display platform for the display and production of human TR mucin domains with defined O-glycans [254]. We used the glycoengineered HEK293 cells designated Tn (KO C1GALT1), T (KO GCNT1, ST3GAL1/2, ST6GALNAC2/3/4), DiST (KO GCNT1, KI ST6GALNAC2/3/4), core3 (KO COSMC, KI B3GNT6), as well as WT which produce a mixture of mono- and disialylated core1 and core2 structures (**Figure 2a**), as these three types represent the most relevant O-glycan structures in the human intestine [113-115]. The stable expression and purification of mucin TR reporters (MUC1, MUC2, MUC5AC and MUC7) in these isogenic HEK293 cells allow the production of core1, 2 and 3 glycans of varying complexity with different clustered patterns of O-glycans [254]. Probing purified mucin TR reporters with pasteurized *A. muciniphila* showed strong preference for



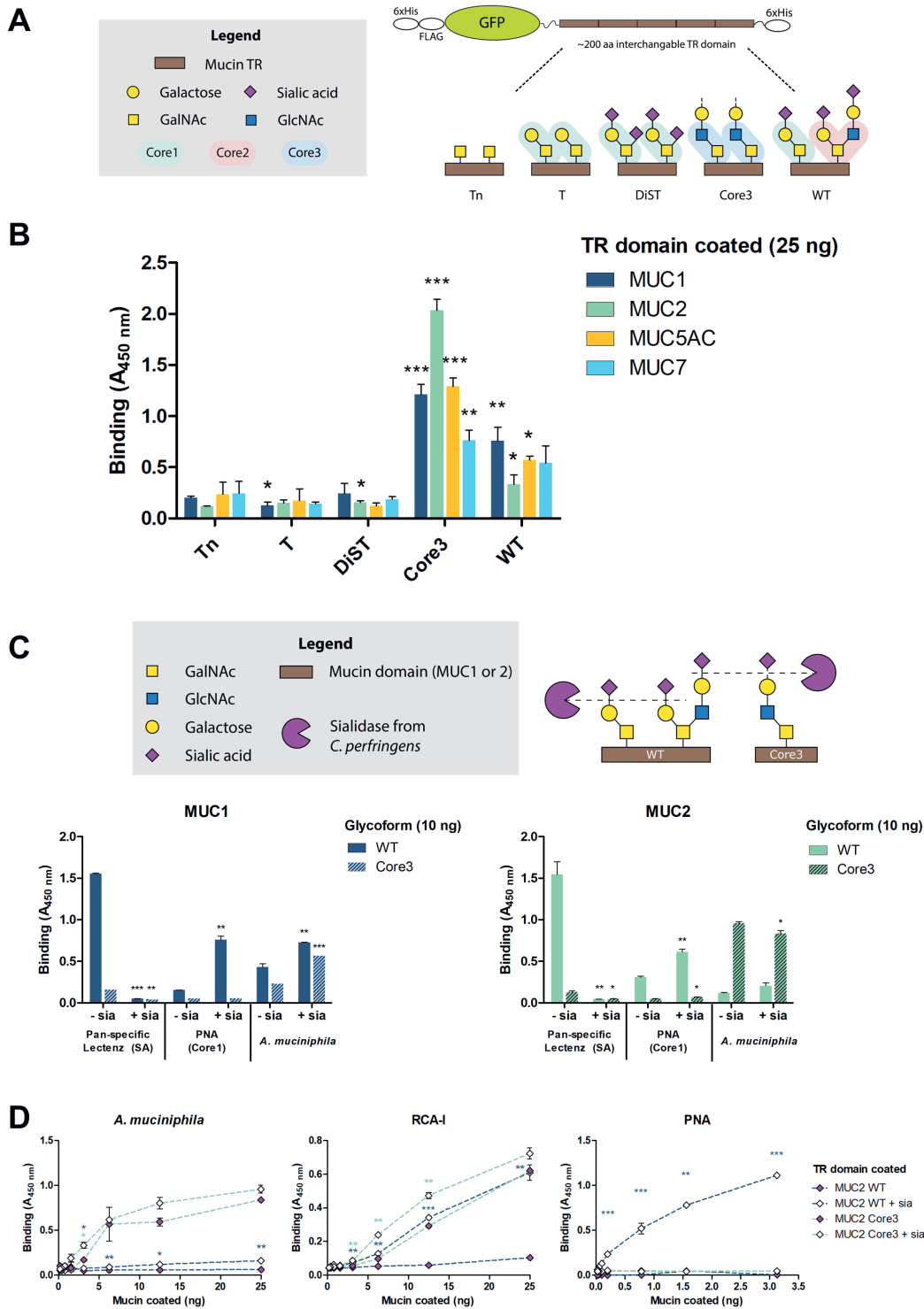
core2 and core3 O-glycosylated reporters without clear preferences for distinct mucin TRs (**Figure 2b**, **Supplementary Figure 4a,b**), which suggests the binding is mainly directed to the O-glycan structure and not particular patterns or clusters of O-glycans. Treatment of MUC1 and MUC2 WT reporters with a *Clostridium perfringens* sialidase further increased binding to MUC1, but not for MUC2 (**Figure 2b, c**). This low or absence of binding to MUC2 WT may be explained by a relatively low abundance of core2 glycans on MUC2 WT [254], further supporting the preference for the LacNAc (Gal $\beta$ 1-4GlcNAc) epitope (**Figure 2c**). We also probed LacNAc with *Ricinus Communis* Agglutinin I (RCA-I, [653]), which revealed a concentration-dependent, sialic acid-sensitive signal for MUC2 core3, similar to *A. muciniphila* binding patterns (**Figure 2d**), thus confirming a binding preference of *A. muciniphila* for non-sialylated LacNAc. Furthermore, desialylation of the core3 glycoforms of mucin reporters produced only minor changes in binding, which is consistent with poor sialylation of core3 O-glycans in HEK293 cells, as previously reported (**Supplementary Figure 5a-b**).

**Figure 1 (next page). ELISA binding assays of *Akkermansia muciniphila* with porcine gastric mucin (PGM).** **A**, Binding of live *A. muciniphila* ( $5 \cdot 10^8$  CFU/mL in PBS) under oxic conditions at different temperatures (4°C, RT or 37°C) and under anoxic conditions at 37°C to different amounts of PGM coated (0-500 ng). \*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ , comparing binding to the same amount of coated PGM at 4 °C. **B**, Binding of live and pasteurized *A. muciniphila* ( $5 \cdot 10^8$  CFU/mL) at 4 °C under oxic conditions. **C**, Binding inhibition of pasteurized *A. muciniphila* ( $5 \cdot 10^8$  CFU/mL) with PEG polymers (100 and 600 kDa) or PGM on binding at 4 °C under oxic conditions. Plates were coated with PGM as indicated. Bars and data points represent the mean  $\pm$  SD of three biological replicates. \*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ , comparing binding to the same concentration of coated PGM between different conditions, with binding in PBS as a reference.





CHAPTER 6



**Figure 2 (previous page). ELISA binding assays of *Akkermansia muciniphila* with glycoengineered human mucin TR reporters.** **A**, Graphic depiction of the mucin reporters with different glycoforms. **B**, Binding of pasteurized *A. muciniphila* ( $5 \cdot 10^8$  CFU/mL) to isolated mucin TR reporters (25 ng coated). TR reporters (MUC1, MUC2, MUC5AC, MUC7) with five different glycoforms (Tn, T, DiST, core3 and WT) were coated and incubated with pasteurized *A. muciniphila*. Bars represent the mean  $\pm$  SD of three biological replicates. \*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ , comparing the respective glycoform with Tn of same TR domain. **C**, Probing of MUC1 and MUC2 reporters (WT or Core3) (10 ng coated) with or without *Clostridium perfringens* sialidase pretreatment (20 mU overnight with *A. muciniphila* ( $1E+9$  CFU/mL), Pan-specific Lectenz (2  $\mu$ g/mL) and PNA (0.1  $\mu$ g/mL). One representative experiment is shown, with bars representing the mean  $\pm$  SD of 2 technical replicates. \*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ , comparing sialidase-treated with the untreated condition. **D**, Titrations of MUC2 WT and Core3 reporters pretreated with and without *C. perfringens* sialidase and probed with pasteurized *A. muciniphila* ( $5 \cdot 10^8$  CFU/mL), RCA-I (0.05  $\mu$ g/mL), and PNA (0.1  $\mu$ g/mL). One representative experiment is shown, representing the mean  $\pm$  SD of 2-3 technical replicates. \*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ , comparing sialidase-treated with untreated condition at same concentration of mucin. For clarity, statistic significance is only depicted for the four highest concentrations.

### *A. muciniphila* binds to O-glycans on the cell surface of HEK293 cells

Since the LacNAc disaccharide is a common terminal structure found on most types of glycoconjugates, including glycolipids and N-glycoproteins, we probed binding of pasteurized *A. muciniphila* to glycoengineered HEK293 isogenic cells by flow cytometry analysis. *A. muciniphila* was found to only bind to WT cells after sialidase treatment and binding was eliminated by loss of elongated O-glycans (KO C1GALT1), which suggests that *A. muciniphila* preferentially binds LacNAc on O-glycans. Moreover, binding was also lost after elimination of LacNAc synthesis (KO B4GALT1/2/3/4), further supporting LacNAc-specific binding (**Figure 3a**). Interestingly, *A. muciniphila* did not bind significantly to HEK293 cells engineered to express core3 O-glycans (KO C1GALT1 KI B3GNT6), except when the MUC2 TR reporter was expressed transiently on the cell surface, regardless of sialidase pretreatment (**Figure 3b**). Strikingly, this increase in binding was not observed when expressing the MUC1 TR reporter on the cell surface (**Figure 3b**).

### Endogenous sialidase activity of *A. muciniphila* also contributes to enhanced binding

The genome of *A. muciniphila* encodes two main exo-sialidases Amuc\_0625 and 1835 [159, 654]. The first one is expressed higher during growth on mucin [654] and has been extensively characterized for its optimal pH, temperature and substrate specificity [650, 655]. Here, we sought to determine whether this *A. muciniphila*

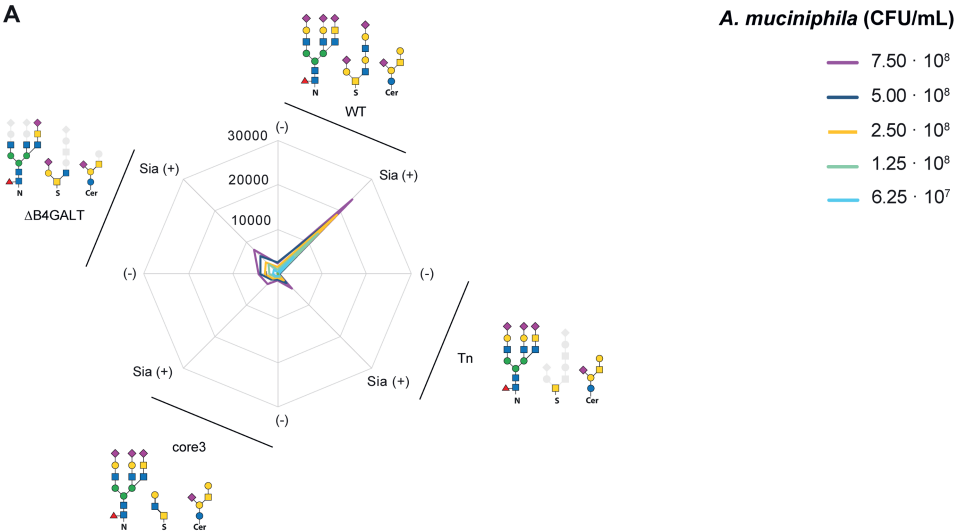


sialidase Amuc\_0625 could, similarly to the *C. perfringens* sialidase, uncover the binding epitopes. To this end, we treated mucin TRs with purified Amuc\_0625, heterologously expressed in *E. coli* [650]. Given the generally low binding to MUC2 WT and relatively low sialic acid content in core3 reporters, we included a sialylated mucin TR domain to which binding of *A. muciniphila* was highest (MUC1 WT) (**Figure 2b**). Staining by a sialic acid-specific probe demonstrated successful desialylation by Amuc\_0625 and enhanced binding by *A. muciniphila*, similar to the effect observed after desialylation with a sialidase from *C. perfringens* (**Figure 4a**). A similar, but weaker effect was demonstrated for the heterologously expressed sialidase Amuc\_1835 (**Supplementary Figure 7a**), probably owing to variations in substrate specificity, metal requirements and pH optima or possibly a lower activity of this sialidase [655]. The sialidase activity of pasteurized *A. muciniphila* itself was low, as the sialidase was not heat stable (**Supplementary Figure S7b**). Overall, these data show that *A. muciniphila* cells contain the enzymatic machinery to promote their own binding to mucus.

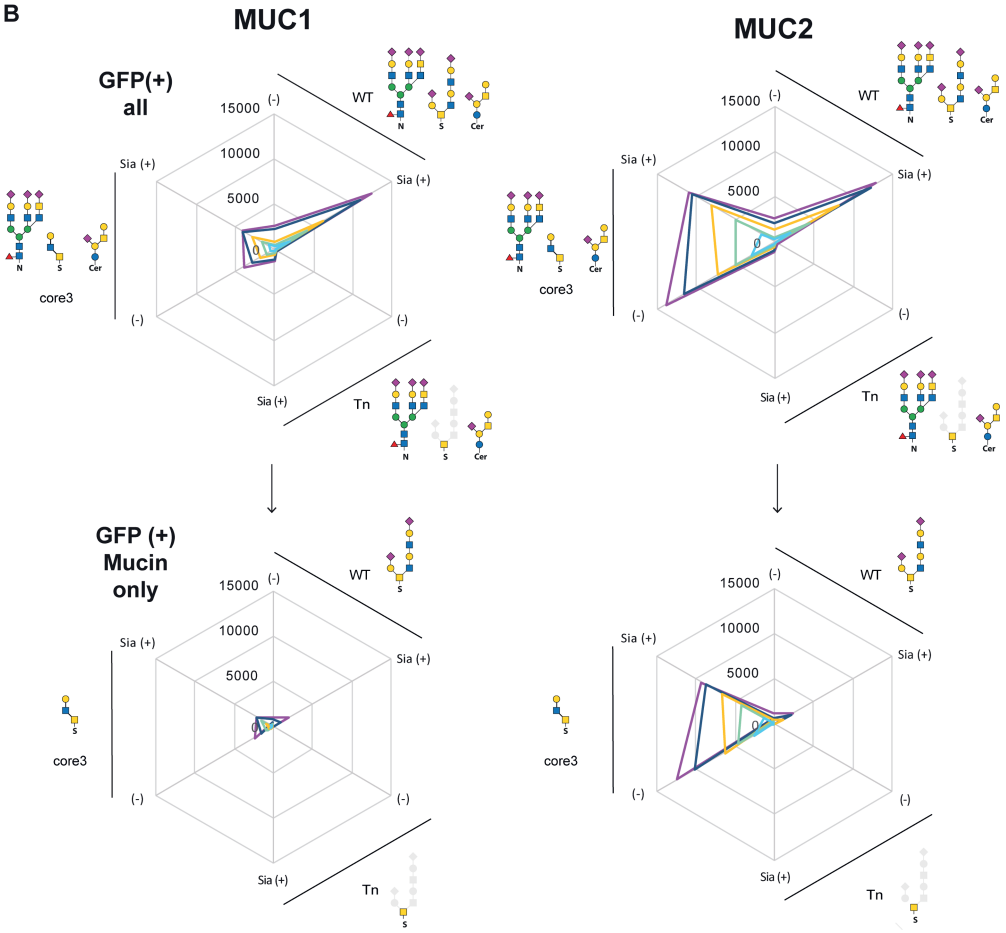
**Figure 3 (next page). Flow cytometry analysis of *A. muciniphila* binding with glycoengineered HEK293-cells. A**, Binding of pasteurized *A. muciniphila* (up to  $7.5 \cdot 10^8$  CFU/mL) to HEK293-WT, -Tn (KO C1GALT1), -core3 (KO COSMC, KI B3GNT6) and - $\Delta$ B4GALT (KO B4GALT1/2/3/4) with (Sia +) or without (-) *Clostridium perfringens* sialidase treatment. Data points represent the average median fluorescence intensity (MFI) of two biological replicates. **B**, Binding of pasteurized *A. muciniphila* to MUC1 and MUC2 mucin reporters transiently expressed in HEK293-WT, -Tn (KO C1GALT1), -core3 (KO COSMC, KI B3GNT6). Mucin expression was monitored by GFP expression and binding was estimated as median fluorescence intensity (MFI) of gating cells expressing mucin-GFP reporter (GFP+ all) or subtracted binding levels to cells without GFP expression (GFP+ Mucin only). One representative experiment is shown.

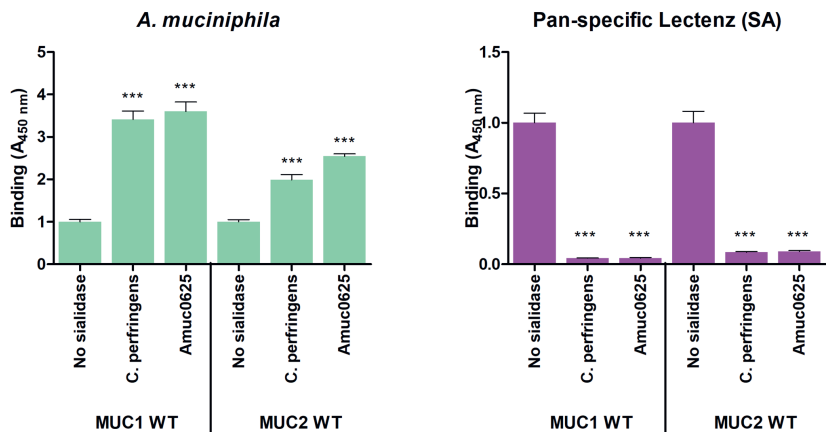
BINDING OF *A. MUCINIPHILA* TO MUCIN

A

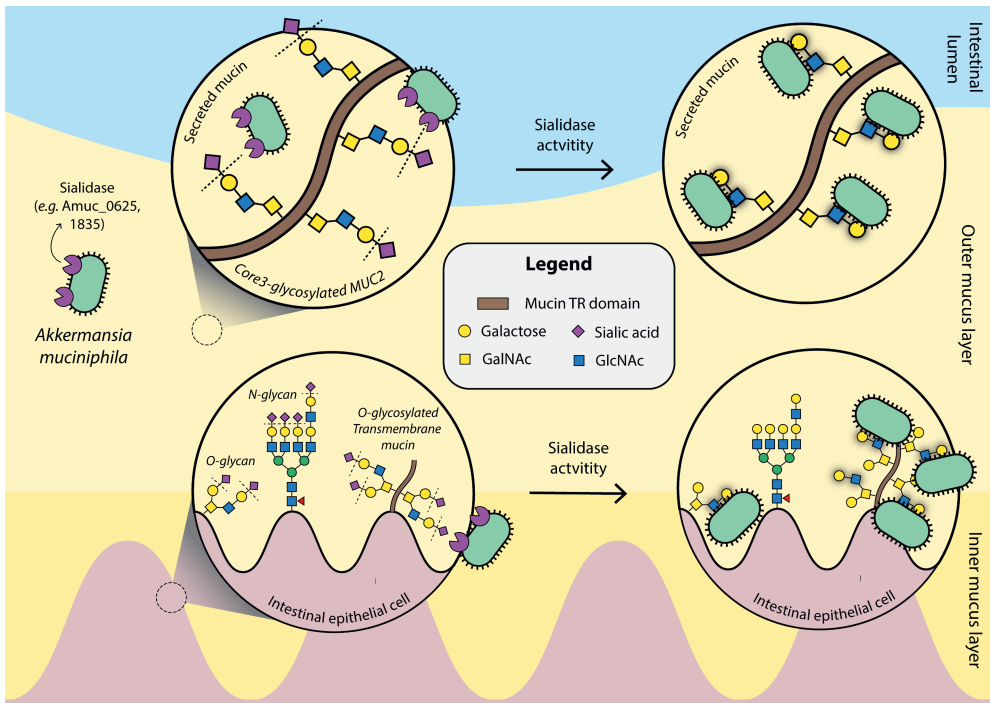


B





**Figure 4. Function of endogenous sialidase from *A. muciniphila* with human mucin TR reporters.** ELISA assay with MUC1 and MUC2 WT (12.5 ng coated) pretreated with endogenous sialidase Amuc\_0625, probed with pasteurized *A. muciniphila* ( $5 \cdot 10^8$  CFU/mL) and Pan-Specific Lectenz (2.0  $\mu$ g/mL). Mucin reporters were pretreated with 1.4 mU sialidase overnight at 37 °C. Plates were read at 450 nm and normalized to untreated control. One representative experiment is shown, representing the mean  $\pm$  SD of 2-3 technical replicates. \*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ , comparing sialidase-treated with untreated condition.



**Figure 5. Proposed binding preferences of *A. muciniphila*.** Schematic overview of suggested glycan binding by *A. muciniphila* *in vivo* as demonstrated in this study. A sialidase of *A. muciniphila* (Amuc\_0625) is expressed by *A. muciniphila* to remove sialic acid present at terminal positions of mucin glycans (top panel) and glycans expressed on cell surface (bottom panel) Subsequently, *A. muciniphila* binds to LacNAc structures exposed after desialylation.

## DISCUSSION

*A. muciniphila* is a well-known mucus-adapted intestinal symbiont with unique mucin-binding properties. Here we dissected the binding of *A. muciniphila* to mucin. We showed that mucus binding of the anaerobic *A. muciniphila* is independent of oxygen concentration in the environment and that both live and pasteurized cells bind similarly to mucins. The use of pasteurized *A. muciniphila* cells allowed us to uncouple the metabolic and enzymatic activity from the mucus binding activity. Using glycoengineered cells and TR reporters representing human mucins with tunable O-glycan structures, we demonstrated that binding of *A. muciniphila* is rather indiscriminate for the type of mucins and mainly dependent on LacNAc terminated O-glycans, which may be exposed by desialylation by endogenous sialidases *A. muciniphila*'s, like Amuc\_0625 (summarized in **Figure 5**).

Our study investigated the binding of *A. muciniphila*, under different assay conditions (temperature, oxygen) and molecular cues (glycosylation, mucin TR sequence). Optimization was carried out using commercially available PGM, which has been commonly used in mucin research and has been extensively characterized [249], but carries different O-glycans than human colonic MUC2 [113, 126]. Differences in protein and glycan content may explain why binding patterns of several probiotics to PGM did not reflect binding to human intestinal mucins [640]. Our study demonstrates that this is not the case for *A. muciniphila*, due to its glycan-specific, rather than protein-specific binding, justifying the use of PGM for optimization of binding parameters in this case.

Our experiments with PGM demonstrated highest binding of live *A. muciniphila* at temperatures lower than human body temperature, which could be explained by remaining mucin-degrading activity at 37 °C. This could also explain previous reports of low binding to human mucin at 37°C [641]. Incubation at 37 °C, however, abolished the concentration-dependent binding of pasteurized *A. muciniphila*, suggesting the need to discern mucin binding versus activity at 37 °C.

Experiments with mucin TR reporters with tunable O-glycans, revealed that *A. muciniphila* binds to LacNAc. The preference for core3 is in line with the mucin glycans present in the colon, one of the main *in vivo* colonization niches of this bacterium [113, 126], suggesting that intestinal mucin glycans promote the colonization of *A. muciniphila*. We can, however, not exclude that the presence of GlcNAc in the growth medium results in a bias towards this sugar.

Previous work suggested that binding of *A. muciniphila* to human colonic mucin from healthy tissue was relatively poor. The degree of glycosylation of these mucins, however, was not characterized [641]. Moreover, the use of live bacteria in this study could have compromised the binding assay, as the enzymatic activity of the cells could have resulted in degradation rather than binding. Additionally, given the use of surgically obtained materials, it is not unlikely that the mucin glycans were already

partially degraded resulting in the loss of critical binding epitopes (**Figure 2b**). In contrast to these results, a recent study did find considerable binding of live *A. muciniphila* to human colonic mucins [642]. Differences in binding between different studies could be explained by variations in sulfation [656] or sialic acid O-acetylation [642], which was beyond the scope of our study. Binding of *A. muciniphila* to Caco-2 and HT29 cells was also demonstrated [641, 643], but these colon carcinoma cells exhibit relatively low expression of mucins [241, 621], which makes these models less representative of the human *in vivo* colonization niche. These cells do, however, express core2 glycans and I-branched (containing additional poly-LacNAc) glycans on their cell surface [634], which matches with the preferred binding requisites of *A. muciniphila* identified in our study.

The results of our study strongly points towards the involvement of a lectin-like adhesin in mucus binding of *A. muciniphila*, however, little is known about microbial adhesins from intestinal bacteria [138, 139]. A recent metagenomic screening of human intestinal microbiota revealed thousands of sequences predicted to encode lectins [139]. The most common carbohydrate-binding domain in human intestinal microbiota was found to be a domain previously described as Bacteroidetes-Associated Carbohydrate-binding Often N-terminal domain (BACON). Of interest, this BACON-like domain is also present in some proteins of *A. muciniphila* [139, 159, 657]. This domain has not been extensively characterized but is suggested to be involved in the binding of mucin glycans [657-659]. Additionally, as annotated in the Cazy-database, genomes of *A. muciniphila* strains reveal several types of other carbohydrate-binding modules (CBMs) found on e.g. glycoside hydrolases [140]. Lastly, the role of the (underexplored) type IV pili of *A. muciniphila* [660] in accommodating binding to mucin glycans also remains to be elucidated.

Next to uncovering the glycans preferred by *A. muciniphila* as substrates for binding, our data provides valuable insights into the role of bacterial binding in mucin glycan degradation. Increased binding after removal of terminal sialic acids by endogenous exo-sialidases suggests that activity of sialidases is an important initial step for *A. muciniphila* to gain access to the underlying mucin glycans for further degradation (**Figure 4b**), as has been suggested previously [661]. In fact, it was shown recently, that – after removal of sialic acids – mucin glycan degradation by *A. muciniphila* is continued by endo-acting O-glycanases (Amuc\_0724, 0875 and 2108) targeting LacNAc structures [661], which hints at the initial binding recognition site identified in this study. Additionally, two M60-like peptidases involved in mucin degradation by *A. muciniphila* (Amuc\_1514 and Amuc\_0627) [662, 663] were earlier shown to be limited by the presence of sialic acid during cleavage [645, 663]. Recent work even demonstrated a preference of Amuc\_0627 [645] and metallopeptidase Amuc\_1438 [664] for the truncated T and Tn O-glycan structures, suggesting that mucinases of commensal bacteria, as opposed to mucinases from pathogenic species such as StcE [665], start degrading the protein core only after





glycans have been trimmed down by GHs [645]. The lack of binding of *A. muciniphila* to Tn- and T-mucin TRs fits well within this theory and suggests that after glycoside hydrolysis, binding is no longer required, and the bacterium continues with degradation of the protein core.

Our study further showcases the efficacy of recombinant human TR reporters to elucidate binding mechanisms of a crucial commensal microbe. A next step would be to use these reporters to further elucidate the microbial adhesin of *A. muciniphila* responsible for binding. Moreover, given the previously observed differences in binding between human isolates of *A. muciniphila* [643], it would be interesting to extend these experiments to all four *A. muciniphila* phylogroups identified in human beings and other mammalian hosts [666-669]. The presence of different core structures in the murine GIT as well as the expression of a different repertoire of GHs by mouse-colonizing strains of *A. muciniphila* [140], could imply that these bacteria also express different mucin-binding ligands. Furthermore, to improve representability of the human intestinal tract *in vivo*, binding experiments could be extended to a more dynamic environment, e.g. using microfluidics [649], or with more complex O-glycosylation epitopes, including sulfated [670], O-acetylated, fucosylated and more complex branched glycan structures. It would also be interesting to study *A. muciniphila* binding in the context of a microbial community with microbes competing for the same binding sites. Finally, our experiments indicate that *A. muciniphila* binds to O-glycans presented on the cell surface or transmembrane mucins, suggesting its potential involvement in direct microbe-host cell signaling.

In summary, our study is the first to demonstrate that *A. muciniphila* binds to LacNAc present in O-glycans on mucins and cell surface molecules, and that it is enzymatically equipped to promote its own access to mucin glycans for binding and degradation. These results pave the way for future research into colonization, mucin recognition and mucin degradation by an intestinal symbiont that plays a crucial role in the mucosal host-microbe ecosystem.

## FUNDING STATEMENT

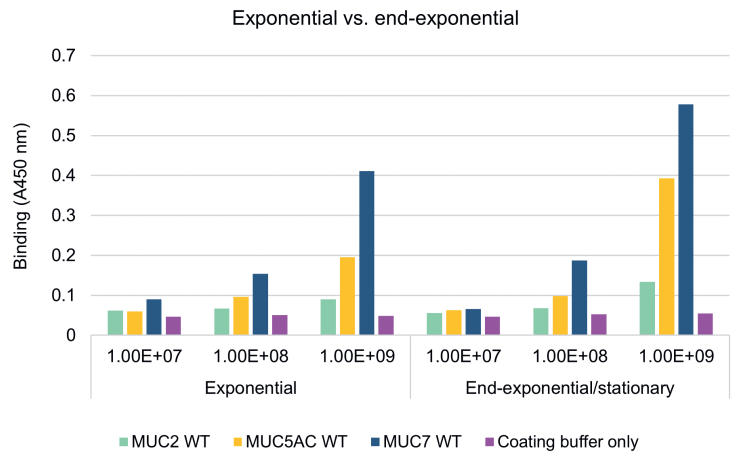
JE was funded by a [Building Blocks of Life](#) grant from the Dutch Research Council (NWO) (grant no. 737.016.003) and a FEMS Research and Training Grant (FEMS-GO-2021-069). JE and HPLT were supported by the Spinoza Award and SIAM Gravity Grant 024.002.002 of the Netherlands Organization for Scientific Research (NWO) of WMdV. HC and YN were supported by the Novo Nordisk Foundation and Lundbeck Foundation. NdH has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (GlycoSkin H2020-ERC; 772735).

## COMPETING INTEREST

WMdV is a co-founder, inventor on patents and shareholder of The Akkermansia Company that commercializes pasteurized *A. muciniphila*.

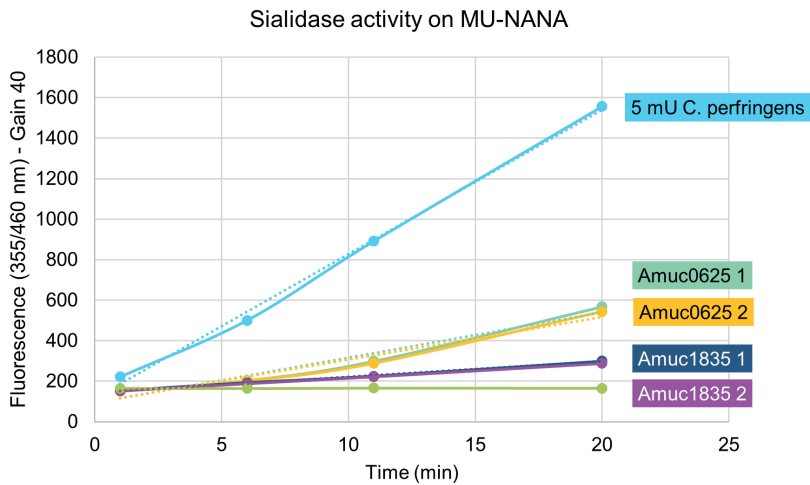


SUPPLEMENTARY INFORMATION



**Supplementary Figure 1. ELISA binding assays of *Akkermansia muciniphila* with mucin TRs at different growth phases.** Binding of live *A. muciniphila* ( $1 \cdot 10^7$ - $9$  CFU/mL in PLI-P) under oxic conditions at 4 °C to 10 ng of MUC2, MUC5AC and MUC7 WT or coating buffer only. One representative biological replicate is shown.

A



B

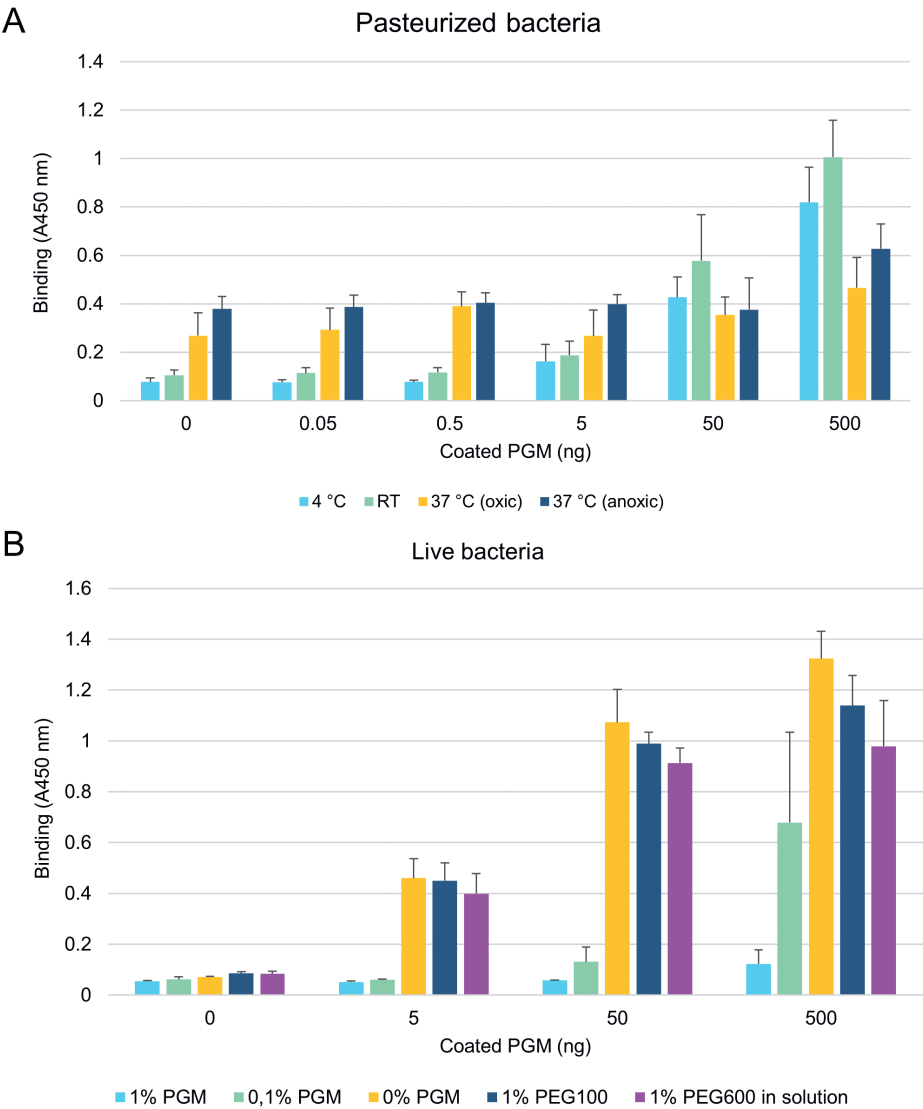
Enzyme	Slope (intensity/min)	Units (μMole MU/min)	mU in reaction (in 100 μL)	Stock concentration (U/mL)*
5 mU C. perfringens	71,3	0,004	4,5	4**
Amuc0625	21,8	0,001	1,4	5.5
Amuc1835	7,4	0,0005	0,5	1.9

\*Amuc0625 and 1835 were diluted 400x with 1 mM MU-NANA in a final reaction of 100 μL. The average of two aliquots is shown.

\*\* Sialidase from *C. perfringens* was bought commercially and diluted to a final concentration of 5 mU (~4.5 mU in this assay).

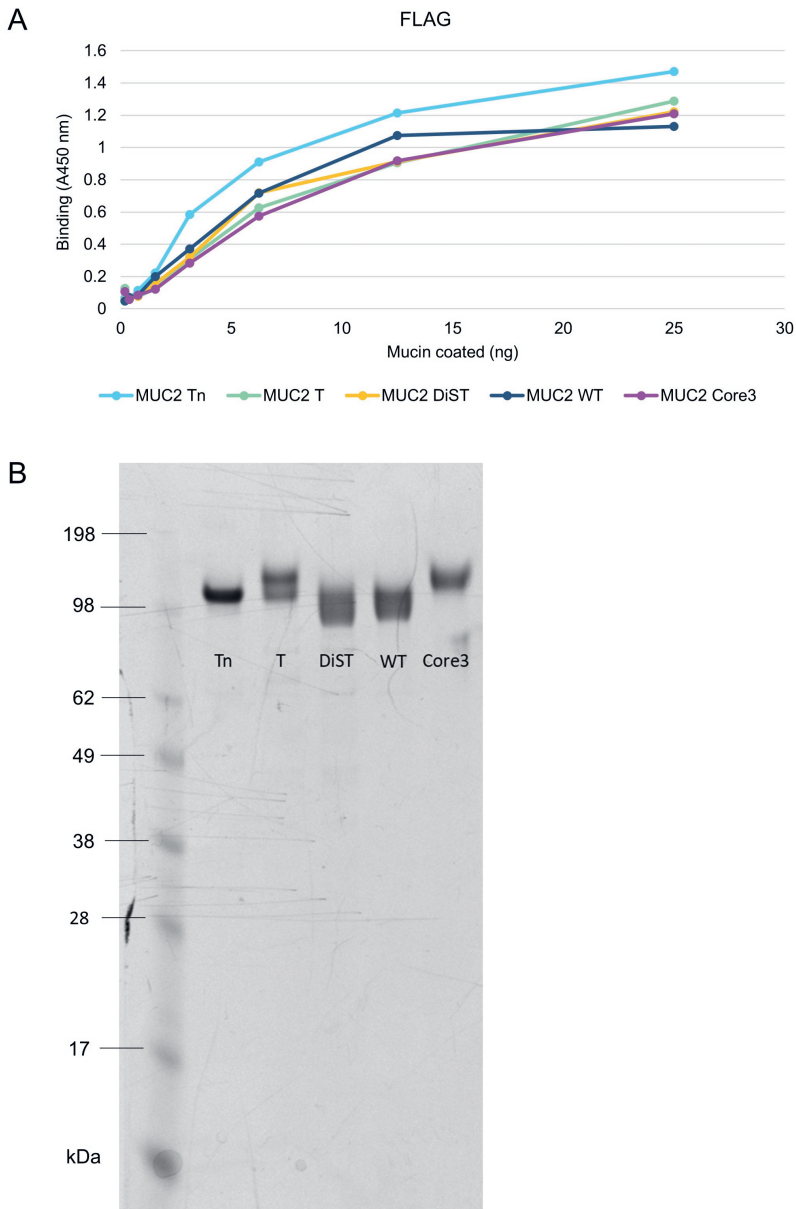
**Supplementary Figure 2. Quantification of activity of *C. perfringens* and *A. muciniphila* sialidases using MU-NANA assay.** **A**, Released 4-methylumbelliferone (MU) after incubation of 5 mU *C. perfringens* or two different aliquots of heterologously expressed Amuc\_0625 and 1835 with 1 mM MU-NANA (2'-(4-Methylumbelliferyl)-α-D-N-acetylneuraminic acid). **B**, Summary of kinetic parameters for each sialidase. All experiments were performed at pH 6.0 at 37 °C.





**Supplementary Figure 3. ELISA binding assays of *Akkermansia muciniphila* with porcine gastric mucin (PGM).** **A**, Binding of pasteurized *A. muciniphila* (5 · 10<sup>8</sup> CFU/mL in PBS) under oxic conditions at different temperatures (4 °C, RT or 37 °C) and under anoxic conditions at 37 °C to different amounts of PGM coated (0-500 ng). **B**, Inhibition with polymers PEG (100 and 600 kDa) or PGM on binding of live *A. muciniphila* (5 · 10<sup>8</sup> CFU/mL) at 4 °C under oxic conditions. Plates were coated with PGM as indicated. Bars and data points represent the mean ± SD of three biological replicates.

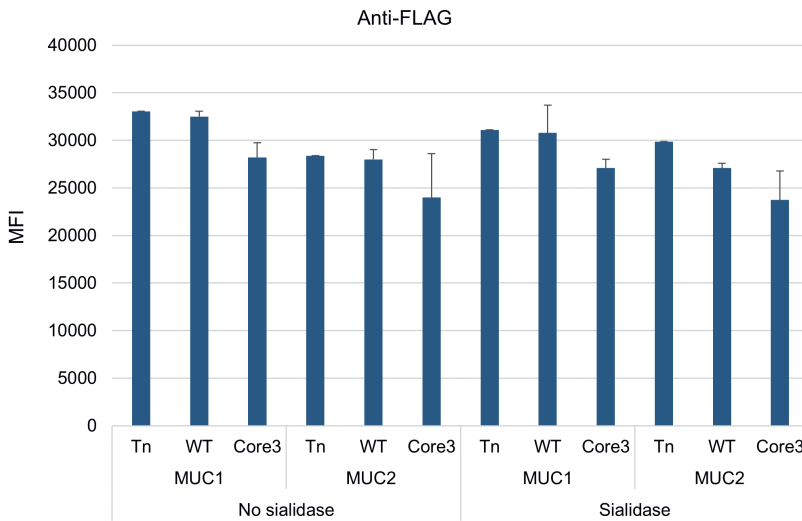
# BINDING OF *A.MUCINIPHILA* TO MUCIN



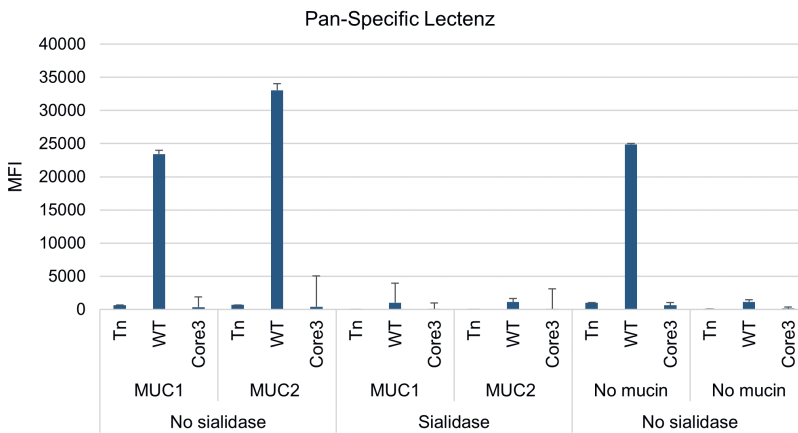
**Supplementary Figure 4. Quantification of protein load by FLAG-tag and Coomassie blue analysis. A,** Probing of MUC2 reporters (Tn, T, DiST, WT and core3) up to 25 ng with anti-FLAG (0.1  $\mu\text{g/ml}$ ). **B,** Coomassie blue analysis of MUC2 reporters (1  $\mu\text{g}$  per lane).



A



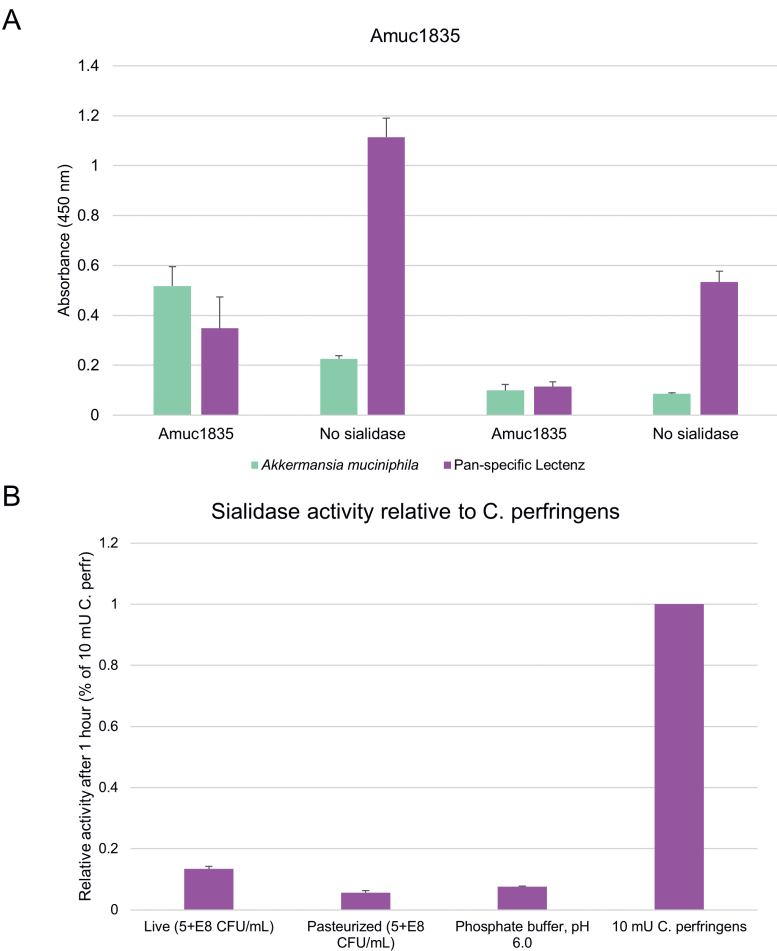
B



**Supplementary Figure 6. Flow cytometry analysis by FLAG-tag and Pan-Specific Lectenz.** Probing of MUC1 and MUC2 mucin reporters transiently expressed in HEK293-WT, -Tn (KO *C1GALT1*), -core3 (KO *COSMC*, KI *B3GNT6*) with anti-FLAG (A) and Pan-Specific Lectenz (B), with and without sialidase treatment. Data points represent the average median fluorescence intensity (MFI) of two biological replicates.







**Supplementary Figure 7. Sialidase activity of Amuc\_1835 and pasteurized or live *A. muciniphila*.** ELISA assay with MUC1 and MUC2 WT (12.5 ng coated) pretreated with endogenous sialidase Amuc\_1835 were probed with pasteurized *A. muciniphila* ( $5 \cdot 10^8$  CFU/mL) and Pan-Specific Lectenz (2.0  $\mu$ g/mL). 0.5 mU sialidase was added and incubated for overnight at 37 °C. Plates were read at 450 nm. Bars represent the mean  $\pm$  SD of 2-3 technical replicates of one representative biological replicate. **B**, Relative activity of live and pasteurized *A. muciniphila* ( $5 \cdot 10^8$  CFU/mL) compared to 10 mU *C. perfringens* for 1 hour at pH 6.0 at 37 °C under oxic conditions.



*Science, never a doll  
moment*



# CHAPTER 7

## General Discussion

## General Discussion

### MODELS: THE PAST, THE PRESENT AND THE FUTURE

The myriad of host-microbe interactions taking place in the human gastrointestinal tract (GIT) are extremely complicated to model, which, at the same time, also reinforces the need to model them to improve our understanding of the human-microbe intestinal ecosystem. To solve this puzzle, a standardized and systematic approach is required for both *in vivo* and *in vitro* models of host and microbe, as reinforced by **part 1** of this thesis. Moreover, although, the activity of intestinal microbes has been shown to extend beyond the human GIT, **part 2** showcases that modelling a small piece of the puzzle, at the site where host and bacterium are in direct contact – the colonic mucus layer – is already a complex scientific endeavor. Meanwhile, the development of *in vitro* models of the human colonic mucus layer has not been standing still. In this chapter, I discuss the *in vitro* models described in this thesis (past) as well as contemporary trends in the field (present). Based on the current caveats in these *in vitro* models, I make suggestions for to-be-developed *in vitro* models of the colonic mucus layer (future). Apart from discussing the scientific basis, I also provide suggestions on which directions current (academic) scientific practice and education should take towards more sustainable science in the broadest sense of this term.

### THE PAST: A SEPARATION OF MICROBES AND MUCINS

Both **Chapter 2** and **3** critically evaluate the historical use of common *in vivo* and *in vitro* models of the human GIT (**Chapter 2** and **3**) or its microbial community (**Chapter 2**), with the common observation that the use of these models requires better standardization in terms of experimental design and reporting. This has been (re-)emphasized by others recently, not only in the context of *in vitro* models in general ([524] and **Chapter 3**), but also specifically with respect to the intestinal microbiota in animal models [182, 671]. Obviously, it is easy to criticize models established decades ago, with contemporary scientific knowledge, however, we should remember that, for instance, the (A)SF was designed in an era in which large-scale anaerobic cultivation of human gut microbiota as well as identification and quantitation of its members by 16S rRNA gene sequencing were not even developed yet. In addition, it should be acknowledged that the application of gnotobiotic animals or Caco-2 cell lines in the studies reported in **Chapter 2** and **3** respectively, depend heavily on the aims of the respective studies, which are likely different from those of the chapters in this thesis. In fact, the contribution of these studies to the advancement of *in vivo* and *in vitro* models to mimic the human GIT and its microbiota more closely, should not be underestimated, as they do showcase the power of defined microbial communities as well as standardized (animal or cell) models. More specifically, we demonstrate the power of cell lines in **Chapter 3**, **4** and **5** by cultivating the commonly used Caco-2

and HT29-MTX on permeable membranes. These commercially available cell lines provide a cost-effective, high-throughput and reproducible method to study the colonic epithelial layer and its response to bacterial compounds. In **Chapter 3**, we employed these features by comparing the wide range of applications of Caco-2 cells under different *in vitro* conditions, with varying model parameters, including oxygen levels, flow and co-cultures with other cell types or microbiota. In this analysis, we did not find a major effect of any of the selected model and experimental parameters on mucus-related genes. However, this could also depend mainly on the cell line, as in **Chapter 4**, we demonstrated a mucus-enhancing effect of SWMS in HT29-MTX-E12 cells, but not in Caco-2 cells. Cultivation under SWMS (**Chapter 4** and **5**) provided a relatively cheap method to study the introduction of model parameters (orbital shaking, semi-wet interface, and presence of mucus-associated bacteria) on mucus parameters. Moreover, transcriptomic analyses allowed a comprehensive characterization of the cellular response in two different cell lines (**Chapter 4**) and the effect of a key mucus-associated bacterium (**Chapter 5**). In addition, we introduced gut-on-a-chip model, which is believed to be a promising model in future NATs to model the human GIT (see next section). Nevertheless, carcinoma cell lines have several limitations in the context of host-bacteria interactions, including deviant mucin O-glycosylation compared to the *in vivo* human colon. On the contrary, in this context, another relatively simple, but non-intestinal, cell line should not be ignored: HEK-cells. These kidney-derived cell lines have been the “work horse” for heterologous expression of proteins or pathways of interest [672]. This is exemplified in **Chapter 6**, in which we employed HEK293T cells to heterologously express mucin TR domain reporters that are more defined and reminiscent of human colonic mucins in terms of O-glycosylation, as opposed to commonly used PGM or mucins derived from HT29-MTX cells. Using these reporters, we uncovered that binding of *A. muciniphila* is O-glycan specific and confirmed previous assumptions that sialic acid needs to be removed before this bacterium can access the underlying glycan epitopes. Lastly, HEK293-cells have been commonly used as reporter cell lines to test the activation of pattern recognition receptors and cytokine secretion upon exposure to microbial products, as also demonstrated by research affiliated with this thesis (**Box 1**).

The downside of the simplicity of *in vitro* models discussed in this thesis, is the obligate separation of intestinal bacteria and the colonic mucus layer, restricting experiments to the use of non-living bacteria or their compounds and thus, unidirectional microbe-host interactions. Luckily, there are several *in vitro* models under development, which do include a more representative mucus layer and/or a living intestinal microbiota, which allow modelling of bidirectional host-microbe interactions. These will be discussed in the next section.



## THE PRESENT: MICROBES AND MUCINS COMING CLOSER

Organ-on-a-chip systems are considered the next generation of *in vitro* models to mimic the *in vivo* tissue as well as to eventually replace preclinical models in drug development [673]. The field has been rapidly evolving recently, also in the context of the human intestine: After the first report on the co-culture of Caco-2 cell lines with live, aerobic bacteria in a “gut-on-a-chip” in 2012 [208], this model has been recently expanded to the use of organoids with or without anaerobic microbes or communities (“intestine-on-a-chip<sup>3</sup>”) [211, 212]. In addition, different *in vitro* methods allowing co-culture of host and bacteria are available today, including the HoxBan module [203], Organoplate™ [674], HMI Module [247], 3D intestine model [675], HuMiX model [248], Intestinal Hemi-Anaerobic Co-culture System (iHACS [225]), GuMI physiome platform [676] and the O<sub>2</sub> gradient cassette [677]. Although different in design, they share similar features. Below I discuss the features most relevant in the context of host-bacteria interactions at the colonic mucus layer: 1) oxygen levels; 2) the application of mechanical and shear stress; 3) host cells; 4) the application of 3D culture and 5) the bacteria investigated. For each, I will discuss the current state of art as well as obstacles to be tackled, when it comes to modelling microbe-mucus interactions.

### Oxygen levels

One of the basic requirements for the co-culture of intestinal cells with bacteria – and at the same time, the biggest challenge – is the creation of an oxygen gradient to facilitate the growth of anaerobic species [54], while providing “physiological hypoxia” [678] to the eukaryotic host cells. This has been addressed in more conventional, static insert models [202, 677, 679] and the relatively simple HoxBan module [203], facilitating co-culture with single anaerobes up to a few days. Oxygen gradients are naturally present in 3D-grown organoids [680] and allowed co-culture with anaerobic bacteria up to a few days in the presence of a mucus layer [681]. The need for 3D organoids to be passaged every 7-14 days, however, limits stable co-culture for longer times [682]. In contrast, dynamic models, with continuous supply of nutrients, would theoretically allow long-term co-culture without any compromise. Examples include the HMI™ module [247] and the HuMiX model [248], which contain an anoxic compartment and allowed co-culture with multiple bacteria. These models, however, did not allow “natural” colonization of the intestinal mucus layer, due to the presence of a membrane and exogenous mucus layer separating the bacteria from the intestinal cells. In the recently developed Intestine-on-chip, an aerobic basolateral

---

<sup>3</sup> “Gut-on-a-chip”, “intestine-on-a-chip”, “Intestine Chip”, “Colon Chip” are distinct terms and names used by the Ingber group. In this thesis, I will consistently use “gut-on-a-chip” to indicate any microfluidic device in which intestinal cells are cultivated, unless describing a specific model from the Ingber group

compartment (~20%) was maintained while creating an anoxic apical compartment that supported the growth of an obligate anaerobe as well as a complex human intestinal microbiota, in which host cells were separated from the bacteria by its own mucus layer [212]. Similarly, the GuMI physiome platform (insert under flow with anoxic apical environment) allowed 48-hour co-cultivation of colon-derived organoids with an oxygen-sensitive bacterium (*F. prausnitzii*) separated by a 15  $\mu\text{m}$ -thick mucus layer, but *B. thetaiotaomicron* overgrew and killed the epithelial monolayer [676].

It is good to note that the *in vivo* colon does not provide a clear separation of an anoxic and oxic compartment. In fact, there is a steep oxygen gradient in the colon, from the anoxic lumen, across the epithelium to the vascularized subepithelial mucosa [124]. Although this gradient may be naturally present in aforementioned *in vitro* models, only a few models have addressed or monitored this gradient. For instance, the HMI module clearly displayed a difference in colonization site of *F. prausnitzii* versus *Bifidobacterium* spp. relative to the (exogenously added) mucus layer [247]. An oxygen gradient was applied using an oxygen permeable plug in inserts and crypt-like scaffolds (See “3D culture”), demonstrating that such a gradient, as opposed to a fully oxic or anoxic model, maintained a viable and properly polarized epithelium, while at the same time facilitating the growth of facultative and obligate anaerobes [679]. A luminal oxygen gradient was also addressed in the 3D intestine model (See “3D culture”), which was seeded with Caco-2 and HT29-MTX cells of which oxygen consumption kinetics and metabolic activity contributed to the radial gradient. The use of patterned lumens, increasing the seeding area of the cells, further decreased luminal oxygen concentrations to < 0.1% [675]. An adopted version of this model was used to successfully facilitate co-culture of an enteroid-derived epithelium with both anaerobic and aerobic species from a human fecal sample for up to 48 h. Interestingly, while the luminal oxygen concentration without microbes was not lower than ~1%, this model still supported the growth of obligate anaerobes, suggesting that the initial colonization and oxygen consumption by aerobic and facultative anaerobic microbes, promoted the subsequent colonization by anaerobes [683], similar to what happens in the developing human intestine [684]. Thus, this model employed both the contribution of host and bacteria to create an oxygen environment resembling the *in vivo* situation [685]. This model, however, did not address the spatial distribution of bacterial species [683], which is, of course, also due to technical limitations in mucus preservation as mentioned in **Chapter 1**.

Oxygen also has an impact on the mucus layer, directly and indirectly. Directly, as disulfide bonds that contribute to mucin network formation are also dependent on the presence of oxygen. Therefore, it could be that the anoxic intestinal lumen contributes to the formation of the outer mucus layer, which is physically different from the inner layer. Indeed, it was demonstrated that oxygen concentrations impact the viscosity of PGM containing PEG-modified nanoparticles [407], but the direct contribution of oxygen to mucus barrier formation *in vivo* remains





to be elucidated. Oxygen has an indirect impact on mucus production. In two static *in vitro* models using donor-derived colonic epithelial cells, the application of an oxygen gradient (basolateral to apical) did not reveal an effect on MUC2 expression when compared to a fully oxic environment [677, 679], whereas a completely anoxic environment resulted in lower MUC2 expression [679]. These data suggested that oxygen supply is needed for intestinal supplies to properly differentiate into goblet cells. Looking at the molecular level, the Hypoxia-inducible factor (HIF) 1 is not only a global regulator of oxygen homeostasis in several cell types, including intestinal epithelial cells (reviewed in [686]), but is also involved in the transcription of several barrier protective genes, including MUC3 [687], trefoil factor [688], MUC5AC [689] and human  $\beta$ -defensin-1 [690]. Indeed, in **Chapter 4**, we observed an overall downregulation of the HIF-signaling pathway in both HT29-MTX and Caco-2 cells under SWMS (probably due to a more oxygenated medium). Lastly, (physiological) hypoxia also has an influence on mucin O-glycosylation. Amongst others, mucus sialylation was recently shown to be affected by oxygen tension; disulfide formation of sialyltransferases, required for their activity, was found dependent on the oxidase Quiescin sulphydryl oxidase 1 (QSOX1) [691, 692]. Remarkably, when revising the data of **Chapter 4**, we did find significant regulation of QSOX1 (FC = -1.8) as well as ST6GAL1 (encoding a sialyltransferase, FC = 2.8). The fact that both genes are significantly regulated, suggests that mucin O-glycosylation is definitely worth to characterize further in HT29-MTX-E12 under SWMS, supporting the work described in the discussion of **Chapter 5**.

In summary, the recapitulation of the oxic-anoxic interface in an *in vitro* model of the colon is not only critical to facilitate co-culture with (strictly) anaerobic bacteria, but also relevant for the production of a representative mucus layer, an interplay that should be taken into account when developing *in vitro* models of the human colon incorporating an oxic-anoxic interface.

### Mechanical and shear stress

As opposed to conventional static *in vitro* models, several recent models have integrated mechanical and shear stress to mimic the dynamics of the human GIT *in vivo*. Mechanical strain can be induced in several ways; shear stress is a product of fluid flow and viscosity. As elaborately discussed in **Chapter 4**, next to the Semi-Wet part, Mechanical Stimulation could also be (partially) responsible for the increased expression of MUC2 and the formation of a thicker mucus layer in HT29-MTX cell lines. Moreover, whereas Caco-2 under SWMS failed to show increased MUC2 production (**Chapter 4**), a combination of flow and mechanical strain in the first gut-on-a-chips induced the expression of MUC2 in Caco-2 cells, which were not previously reported to produce mucus *in vitro* [208, 209].

The effect of shear stress has also been studied in organoids-on-chips, in which flow induced spontaneous goblet cell differentiation, as opposed to the same

organoids grown on inserts or in spheroids, without the need for a differentiation medium [224]. Moreover, this study was the first to report the presence of a bilayered mucus layer *in vitro*, with a total thickness of 500-600  $\mu\text{m}$  [224], similar to the one observed in the human colon *in vivo* [239]. Colon-derived organoids grown on Organoplate™ also showed increased expression of MUC2 under microfluidic as opposed to static conditions [693]. Both models would be interesting to study in the context of intestinal bacteria.

Next to shear stress, mechanical strain mimicking intestinal peristalsis has also been incorporated in *in vitro* models, for instance by cyclic stretch of the seeding membrane (amongst others [208-210, 212]). In the context of mucus production, it was shown that only flow, and not the combination of mechanical strain and flow, significantly increased MUC2 expression in jejunal organoids [694].

As said, shear stress is not only a product of flow, but also solution viscosity. Thus, the presence of a viscous mucus layer may affect the resulting shear stress on the underlying epithelial layer. Additionally, shear stress not only impacts host cell behavior, but has also been long known to affect bacterial adhesion, physiology, pathogenicity and biofilm formation (reviewed in [695]). Therefore, microfluidic models have also opened up opportunities to study the effect of shear stress on bacterial cell response in a systematic, high-throughput manner.

Therefore, as shear and mechanical stress influence both mucus production and bacteria, they are key factors to consider when studying host-microbe interactions in *in vitro* models of the human colon.

### Host cells

The choice for intestinal cell source is important for the representativity of the intestinal cell layer *in vitro*. Early studies on Caco-2 already showed a more enterocyte-like (instead of colonocyte-like) phenotype of this cell line in conventional static *in vitro* models [195, 696], and cultivation of Caco-2 cells on a gut-on-a-chip resulted in further differentiation into four different cell types, which were transcriptionally more similar to the human small intestine than the colon [209]. A more recent single-cell transcriptomic characterization uncovered the transformation of Caco-2 cells on-chip into heterogenous cellular clusters as opposed to a homogenous cell population on inserts, and even reprogramming of cancer-associated genes [697]. Although cultivation under dynamic conditions may transcriptionally reprogram Caco-2-cells, there is a clear trend towards the use of stem-cell derived intestinal cells instead of colon cancer cell lines. An important reason is the ability to recreate the epithelial layer from the GIT site of interest, by varying the isolation source or differentiation medium. For instance, this has resulted in a small-intestine-on-a-chip [211] and a colon-on-a-chip [224, 693]. Based on a selection of colon-specific markers, the colon-on-a-chip was found to be transcriptionally more representative of the colon than Caco-2 cells grown on a chip [693]. Moreover, the “Colon chip” developed by Sontheimer-



Phelps *et al.* is the only *in vitro* model so far that reported a bilayered mucus layer, similar to the one observed *in vivo* [224]. Additionally, as opposed to cancer cell lines, organoids derived from a healthy donor would probably result in more complete mucin O-glycosylation, but despite its crucial role in host-microbe interactions, this has not been addressed comprehensively in organoids yet.

In the context of host cells, other non-epithelial cell types should not be ignored. These include circulating blood cells and immune cells (e.g., peripheral blood mononuclear cells, vascular endothelial cells and mesenchymal cells that (in)directly communicate with the intestinal epithelium and contribute to the cellular complexity of the colonic mucosal layer. Some of the aforementioned models indeed incorporated additional cell types [208-210, 212]. The exact contribution of these cells to host-microbe interactions and mucus production needs to be elucidated.

### 3D cell culture

In literature on *in vitro* models of the human GIT, the use of the term “3D” is broad. As cited by others, “*in vitro* models are positioned along a continuum between 2-D and 3-D” [698], including inserts with ECM components, synthetic/semisynthetic hydrogels as well as organ-on-chip models. Irrespective of the trigger inducing 3D formation of intestinal cells, the resulting tissue architecture has impact on a wide range of cellular processes, including mucus production. In the 3D intestine model, mucus thickness of Caco-2/HT29-MTX seeded on 3D silk scaffolds was increased relative to cultivation on 2D Transwells [675]. Moreover, the patterned lumens, which increased the number of epithelial cells per unit of length of the lumen, further increased mucus thickness. The human colonic crypt array, a collagen scaffold with crypt-like invaginations in combination with a growth factor gradient supporting, amongst others, mucus production, was employed to improve crypt structures of colonic organoids *in vitro* [699]. In addition, the SWMS method used in this thesis has also been applied to a decellularized 3D small intestinal submucosa, which further increased mucus production (thickness as well as MUC1, 2 and 5AC levels) of HT29-MTX-E12 cells compared to these cells grown on inserts [552]. Creation of *in vivo*-like structures with respective organization of different cell lineages and extracellular matrix, is expected to be further accelerated by 3D bioprinting of biomaterials and/or cells or even 4D bioprinting technologies, in which stimuli-responsive materials are used to address time as a fourth dimension [700, 701].

3D culture is equally relevant for bacteria, given the *in vivo* separation of oxygen-tolerant mucus-associated versus strictly anaerobic luminal bacteria, as discussed above. Moreover, a transcriptomic analysis of *B. theta* and *E. coli* revealed different transcriptional patterns of the same bacterial species in the mucus versus in the intestinal lumen [702]. The use of 3D models was at least deemed crucial for host-pathogen interactions (reviewed in [698]). For instance, using a gut-on-a-chip with 3D architecture, it was shown that *Shigella flexneri* takes advantages

of the microarchitecture as well as mechanical forces to invade the human intestine [703].

Therefore, recreation of the 3D intestinal architecture *in vitro* has consequences for both mucus and microbes. Note that the classical view of the locally produced mucus layer was recently challenged by a mouse study that demonstrated that proximal colon-derived O-glycan-rich mucus encapsulates fecal material including microbiota, to which mucus from the distal colon adheres. Certainly, this adds another dimension to the current 3D models of the human colon. Moreover, this study showed that the microbiota controls the production of mucus by proximal colon goblet cells, whereas *vice versa*, O-glycans affect the structure and function of intestinal bacteria [704]. This brings me to the final obstacle that needs to be tackled when modelling the human colonic mucus layer: the bacteria themselves!

### Bacterial cells

*In vitro* models discussed so far vary in the selection of bacterial species that were investigated, ranging from single species to whole communities. Remarkably, only a handful of models have facilitated co-culture of intestinal cells with mucus-associated bacteria. Moreover, in most of these, the growth of bacteria was relying on exogenously added mucin [212, 247, 248]. In a gut-on-a-chip with Caco-2 cells, increased growth of the genus *Akkermansia* was observed compared to the same inoculum in a conventional liquid suspension culture (without intestinal cells). It was hypothesized that this might be due to the production of mucus by Caco-2 cells [212]. The researchers provide, however, little information on the bacterial growth dynamics in liquid culture, which might not be comparable with growth in a gut-on-a-chip (e.g., bacterial cells have already entered death phase in liquid culture, while just reaching stationary phase on a chip). Moreover, the authors did not consider the difference in community dynamics between chip and liquid culture. This would be interesting to study in more detail, as the spatial segregation of species on a chip is expected to be closer to the *in vivo* situation than in a liquid culture. The iHACS model provided stronger evidence of degradation of MUC2 from statically grown 2D organoids by *A. muciniphila*, resulting in bacterial growth without the addition of exogenous mucin. The mucin layer covering the epithelial cells was, however, degraded after 24 h without replenishment by the host cells [225], so did not sustain long-term co-culture of *A. muciniphila* in the mucus layer.

The other way around, the effect of bacteria on mucus production has only been addressed in a few studies (see **Chapter 5** for examples). While scientists have developed several strategies to co-cultivate intestinal bacteria with intestinal cells as well as sophisticated cultivation methods to enhance mucus production *in vitro*, none to few have employed microbes or microbial compounds to induce a mucus layer representative of the one obtained *in vivo*. An exception is provided by the Colon chip with a bilayered mucus layer, of which the authors hypothesized that the presence



of an impenetrable mucus layer by the epithelial cells could be due to their contact with a complex microbiota before isolation as organoids [224]. Another example is provided by the human colonic crypt array, where a gradient of SCFAs, specifically butyrate, was employed to modulate the size of the crypt's proliferative and differentiated portions [699], resulting in a more *in vivo*-like crypt formation [705]. In both examples, the contribution of bacteria or bacterial metabolites to mucus production, was not quantified. Instead, it seems that in most of the *in vitro* models of the human colon, the intestinal bacteriome (bacteria and their activity) is considered an external exposure or an accessory, rather than a crucial component of the model. This is remarkable, especially given the reciprocal interaction between mucus, bacteria, and model features such as oxygen, shear stress and 3D culture. Obviously, this interplay comprises only a snapshot of a long historical co-evolution of microbes and men, making causality hard to disentangle. But it does raise an important question regarding *in vitro* model design: Should we integrate the intestinal bacteriome in future *in vitro* models, i.e., view it as a model parameter?

## THE FUTURE: INTEGRATION OF MICROBES, MODELS, AND MUCINS

The inclusion of bacteria as a model parameter is hampered by several obstacles, of which a few have already been addressed by recently developed *in vitro* models. Whereas many scientists may have faith in advancing technological developments in the field of *in vitro* models of both host (i.e., gut-on-a-chips, organoids) and microbes (i.e., culturomics and other -omics technologies), this thesis re-emphasizes other limitations as well. These include a lack of standardization in current approaches, incomplete reporting, and the complexity of the colonic mucus layer. Sometimes, we seem to forget that we are only beginning to explore the full potential of the intestinal microbiota. Therefore, it might be time for scientists to take one step back, in order to make two steps forward. I will explain this line of thought by discussing the current challenges one by one and make suggestions on how to overcome each and what it will eventually deliver (Warning: A lot of open doors!).

### One step back...

#### 1. Back to the building blocks

Integration of intestinal bacteria in *in vitro* models of the human GIT: Where do we start? Which bacterial factors should we take into account? And how much does each factor contribute? To answer this question, it might be good to start at the bacterial side.

#### From single species to defined communities

The first prerequisite for integrating bacteria in *in vitro* models of the human GIT is the tractability of the species used. This implies that at least its genome is known and that – if the bacterium is meant to stay alive in the *in vitro* model – its abundance and

activity can be tracked. During the past decade, (meta)genomic approaches have greatly expanded our knowledge of the human gut microbiota and answered a major part of the question “Who is in there?” [6-9], no longer limited to only Westernized populations [12, 17]. In addition, complementary “-omics” approaches including transcriptomics, metabolomics and proteomics have further contributed to answering the question “What can they do?” and “What are they doing?”. Ironically, after transitioning from microbial isolation and cultivation to these culture-independent approaches to characterize and quantify microbial communities of the human GIT, scientists now call again for microbial cultures for experimental validation. So-called culturomics has resulted in the isolation and cultivation of species previously deemed uncultivable [16, 227], and expanded the repertoire of human bacterial and archaeal species in the human GIT [706]. Bacterial strains are deposited, preserved, and catalogued by DMSZ (German Collection of Microorganisms and Cell Cultures) and CCUG (Culture Collection University of Gothenburg) or more specific repositories, including the Human Gastrointestinal Bacteria Culture collection (HBC) and the Culturable Genome Reference (CGR) catalogue. Recent advancements in technologies for single-cell isolation, e.g. Raman microspectroscopy (reviewed in [707]), and cultivation, e.g. nano/microfabricated devices (reviewed in [708]), provide a high-throughput approach to isolate or discover previously uncultured species and are expected to expand the cultivated repertoire even more, especially when these techniques can be applied under anoxic conditions [709, 710].

Next to expanding the array of cultivable species, it is advisable to investigate multiple strains of one species simultaneously, in order to draw a proper conclusion on the bacterial species of interest. Inter-strain differences have been commonly reported, also in the context of host-microbe interactions. It also, however, touches upon another obstacle in modelling the intestinal microbiota: To what extent is the selected bacterial strain as well as its behavior upon cultivation *in vitro* representative of the species *in vivo*? Inter-strain differences may affect bacterial behavior *in vitro* as well as its translation to the *in vivo* human GIT. In addition, growth medium also affects the behavior of bacterial species *in vitro*. However, most cultivable bacterial intestinal species have been grown on traditional, complex media of unknown chemical composition, limiting accurate detection of metabolites and disentangling interactions between species [711]. Recently, the growth of 96 gut bacterial strains from core genera of the human gut, was evaluated across 19 undefined and defined media. About 70% of the tested strains grew successfully on defined media, providing a valuable resource for standardization of cultivation protocols of bacterial strains [712].

Increasing the array of bacterial species and strains is one step to expanding the building blocks of *in vitro* models, but thorough characterization of the bacterial effectors, e.g., metabolites, extracellular vesicles, outer membrane components, is equally important in investigating microbe-host interactions. Additionally, assessing



the effect of these bacterial effectors requires less advanced *in vitro* models, i.e., the creation of an anoxic compartment may not be required in a first round of screening.

The next step would be to study single strains in the context of the whole community: It is widely known that bacteria naturally operate in teams and not in isolation. The use of defined communities *in vivo*, of which a (non-exhaustive) selection was described in **Chapter 2**, provides a first step towards understanding microbe-microbe interactions and thus, also resulting microbe-host interactions. An interesting example is provided by the Oligo-MM<sup>12</sup>, which was originally established and characterized in mice, but recently taken out of the host context to explore the metabolic potential and interactions of its individual members. This study demonstrated competition and changes in community structures under certain *in vitro* (nutrient) conditions, providing key insights for future mechanistic studies “back” *in vivo* [713].

Lastly, as discussed in **Chapter 2**, the definition of a “normal” microbiota is not set in stone. Instead, the human intestinal microbiota has been described as “a landscape representing varying global and/or local levels of stability, species richness and diversity” [23, 714], as opposed to the initial division in enterotypes [715]. In general, a combination of multi-omics, predictive modelling and cultivation experiments of single strains to (defined, tractable) microbial communities is believed critical to advance our understanding of this complex microbiome [716].

Parts of this thesis are primarily centered around one bacterial species: *A. muciniphila*. In the context of aforementioned investment in bacterial building blocks, future experiments with this intestinal symbiont are proposed in more detail in **Box 1**.

BOX 1: A CASE FOR *AKKERMANSIA MUCINIPHILA*

Since the isolation of *A. muciniphila* in 2004, the scientific evidence on this professional mucin degrader has accumulated. Today, the pasteurized form of the bacterium is regarded safe for commercial use in weight gain management [616, 617]. The beneficial effects of *A. muciniphila* as a next-generation beneficial organism have been discussed elaborately elsewhere [162, 166] and were not the main focus of this thesis. Still, **Chapter 5** and **6** provide novel information on the direct and indirect interactions of *A. muciniphila* with its main niche: the colonic mucus layer. In this subsection, we discuss the caveats in the experiments described in this thesis and make suggestions for future research directions regarding *A. muciniphila*. First of all, this thesis was limited to the use of pasteurized bacteria, which has been a pragmatic choice: the use of live bacteria required an anoxic apical environment not achievable under SWMS (**Chapter 5**) and to study the binding, mucin-degrading activity had to be excluded (**Chapter 6**). In **Chapter 5**, however, this is not a major caveat, since pasteurized *A. muciniphila* have demonstrated beneficial effects to the host, similar to live bacteria [164]. Moreover, many of the beneficial effects of both live and pasteurized *A. muciniphila* in high-fat-diet-fed mice could be replicated by administration of recombinantly expressed Amuc\_1100, an outer membrane protein [164]. In addition to this, we and others demonstrated that cell envelope components of *A. muciniphila* (amongst others, peptidoglycan [717], phospholipids [718] and lipo-oligosaccharides (LOS) (manuscript in preparation)) also play a role in host-microbe immune signaling. In brief, regarding our own work, peptidoglycan-derived muropeptides were shown to stimulate NOD1 and NOD2 (Nucleotide-binding and oligomerization domain) in HEK-reporter cell lines [717]. Further, intraperitoneal injection of *A. muciniphila* LOS increased expression of TLR2 in mice, supported by activation of TLR2 of LOS in HEK-reporter cell lines (manuscript in preparation). Thus, investigation of the effects of *A. muciniphila* on the host *in vitro*, does not require live bacteria. Still, it would be interesting to extend the experiments in **Chapter 5** to other compounds of (live) *A. muciniphila* that can influence mucus production across the mucus layer, as has been shown for mucin-degraders *R. gnavus* [156] and *B. thetaiotaomicron* [155]. Examples include aforementioned (released) cell envelope fragments, but also SCFAs and ECVs. The use pasteurized cells in **Chapter 6** provided novel insights into the binding requirements of *A. muciniphila*, due to the use of well-defined mucin TRs. The process of pasteurization, however could have suppressed (additional) binding partners of *A. muciniphila* and the exact binding partners remain to be identified. To fully uncover the binding mechanisms of *A. muciniphila*, its genetic tractability would be desired. So far, however, attempts to genetically manipulate this bacterium have failed. An exception was recently demonstrated by transposon mutagenesis, as





reported in a preprint that provided insight into mucin degradation mechanisms of *A. muciniphila* [719]. It would be interesting to screen the generated mutants to identify the components required for mucin binding, but also in general, other molecular mechanisms that *A. muciniphila* uses to interact with mucus, host cells or other microbes.

Another caveat in the current work is the limitation to a single bacterial strain (Muc<sup>T</sup>). We cannot be sure that the binding specificity of this strain to LacNAc in **Chapter 6** only reflects the behavior of a single strain (isolated from a single human subject). Differences in, amongst others, mucin binding and TLR2 and 4 activation have been observed among and within phylogroups of *A. muciniphila* [643]. Therefore, it would be interesting to expand our findings in our chapters and associated studies using overexpression of TLR and NOD-like receptor in HEK-cells ([717] and manuscripts in preparation, see “List of publications” on page 299) to other strains as well.

Lastly, the cultivation medium used could also have influenced bacterial behavior *in vitro*. Like other mucolytic species, commercially available PGM is still the source of interest for mucin-based medium. As elaborately discussed in this thesis, high inter-batch variability of this compound and its poor representation of human colonic mucin, complicates the quest to study the natural behavior of *A. muciniphila*. Moreover, in **Chapter 6**, we were forced to use a minimal medium, because mucin interfered with the binding assay. Although this medium has been shown to adequately support its growth [647], we can not exclude whether this GlcNAc-based medium influenced its binding preference for LacNAc. The mucin TR reporters used in this chapter (or an extended version) could be considered an alternative to PGM or minimal medium to a more *in vivo*-like growth medium, but would require upscaling of the production and purification of these recombinant proteins.

Overall, this thesis provides novel insights into a key mucin degrader and its interaction with its main colonization niche. Especially its uncovered binding preferences could optimize its (directed) application as a next generation beneficial microbe in human subjects.

### Standardization of cell culture practice

Similar to the bacterial building blocks, standardization on the host side is key. The application of Caco-2 cells knows a long history. In **Chapter 3**, we aimed to make use of this by quantifying the contribution of the microbiome (living (aerobic) bacteria or their products) to the cellular response of Caco-2 cells. Lack of standardization in cell culture practice as well as diversity in bacterial parameters included in this analysis, however, hampered a solid conclusion on the effect of bacterial parameters on Caco-2 gene expression. Although we will never manage to exclude inter-lab variability and labs have the freedom to play with differences in model design, there is still an unnecessarily large heterogeneity in maintenance and application of Caco-2 cells, at

least in the studies included in this analysis. In addition to the experimental parameters already mentioned in **Chapter 3**, this was also reflected in the differences in cultivation medium used, which we decided to exclude from the analysis, as it was hard to quantify and compare. In most cases, approximately 10-20% of medium is fetal bovine serum (FBS) of which the exact composition is highly variable [720], even between batches [721]. Most labs, however, do not dare to make the switch to (generally more standardized) animal-free alternatives due to the costs or the preference not to modify their protocols [722]. A switch to more chemically defined alternatives would promote cell culture standardization – and sustainability, as discussed later in this chapter.

Further, standardization is not limited to model parameters, but also experimental outcomes. In **Chapter 3**, we aimed to compare outcomes between studies, of which transcriptome data was deemed a good starting point due to its comprehensiveness and comparability between studies (if publicly deposited, see next section). TEER is another commonly used outcome in *in vitro* models of the human GIT, but despite that, we detected few similarities in methods and outcomes between studies. In **Chapter 4** and **5**, we decided to include two additional experimental outcomes, which are not often measured in insert studies: medium pH and glucose/lactate levels. Both outcomes provide a relatively easy method to characterize the model more completely, without having to include additional experimental units (wells, in this case). On top of this, it is advisable to invest time and money in monitoring outcomes at multiple timepoints, particularly at the gene expression level due to temporal changes. For **Chapter 4** and **5**, we only included one timepoint in the transcriptome analysis, based on previous studies on SWMS, but it would be interesting to monitor multiple timepoints, for instance by RT-qPCR on a selection of genes to save costs.

With respect to outcomes, (mucin) O-glycosylation is still an overlooked aspect in *in vitro* models of the human GIT, despite being a crucial contributor to mucus layer stability and microbe-mucus interactions. The underexposure of mucin O-glycosylation is mainly due to technical limitations, as briefly discussed in **Chapter 5**. In general, glycomics is a relatively young field, and most of our understanding of the glycome is based on bulk analyses of cell or tissue samples [723], not capturing the heterogeneous expression of different cell types. Single-cell technologies are expected to enhance the knowledge on the glycosylation capacities of cell types [109, 639]. Single-cell transcriptome analysis of Caco-2 cells on-chip (see above, [697]) already demonstrated the power of this approach, even in the context of relatively simple colonic carcinoma cell lines. Eventually, it would be most optimal to capture gene expression, protein expression and glycosylation in one experiment [109].

Standardization of *in vitro* methods requires standard operating procedures (SOPs), as has been called for frequently before and for which guidelines are available [484]. Evidently, it costs a lot of money and time to assess all combinations



of model parameters, especially when including bacterial factors as well, but currently, the opposite is true for *in vitro* models of the human GIT. Therefore, I think it is worth to invest time and money to go back to the basics, e.g., the building blocks of the models. In the context of the bacteria, this implies cultivation and characterization of individual bacterial strains and their products under various conditions, leading to an optimal standard for bacterial cultivation. In the context of the host, similarly, this would mean the standardization of operating procedures and outcomes. Additionally, given the major impact of mucin O-glycosylation, research into microbe-mucin interactions would definitely benefit from investment in the developing field of glycomics. Lastly, standardization would also include the embracement of relatively simple cellular models, as explained in the next section.

### 2. Embracing “simple” cellular *in vitro* models

This thesis demonstrates the power of relatively simple cell lines such as Caco-2, HT29-MTX and HEK293-cells, as well as cheaper methods for their cultivation such as SWMS. Although historically, Caco-2 and HT29-MTX cells have been the model of choice in intestinal *in vitro* modelling, I do not promote the use of these particular cell lines. Other colon carcinoma cell lines have been used and characterized as well [724], which might be worth including for replicating results obtained in one cell line or answering a specific research question. Moreover, similar to established co-cultures of Caco-2 and HT29-MTX, different mixtures can be made with existing or novel cell lines to obtain a more representative intestinal epithelial layer. Although studies involving these cell lines do not easily or directly translate to the human clinical setting, they can provide fundamental insights into cellular behavior by variation of model parameters including mechanical and shear stress, oxygen, co-culture with bacteria or other cell types, as demonstrated by currently available *in vitro* models. Similarly, the increasing availability of gene-editing tools such as Crispr-Cas, can also provide answers to fundamental biological questions in cell lines (and bacteria, see **Box 1**) or address current caveats in (cancer-derived) cell lines. This is elegantly demonstrated by the engineered HEK293-cell library used in **Chapter 6**. The defined nature of these mucin reporters is favorable over mucins purified from humans, pigs or CRC cell lines of which the exact composition is less defined and harder to control. Lastly, simple models allow a cheaper and more high throughput screening of individual (bacterial) compounds as opposed to or prior to organoid(-on-chip) models that are currently costly and require extensive expertise. Thus, the full characterization and standardization of existing cell lines should not be neglected.

### 3. Complete reporting: FAIRification of data

Standardization of *in vitro* models does not imply that scientists are limited in the use of model parameters they want to investigate, as long as they properly report them. The inclusion of bacterial parameters (cultivation, processing and application) comes on top of the existing reporting guidelines for *in vitro* models [484]. As discussed in

**Chapter 3**, however, science is still calling for good *in vitro* reporting standards [524]. Complete reporting does not only refer to increased transparency of the research process and products (e.g. in Guidelines for Transparency and Openness Promotion (TOP) Open Science Framework) [725], but also sharing of data in public repositories. The latter is particularly encouraged by the FAIR Guiding principles (Findable, Accessible, Interoperable and Reusable) for scientific data management and stewardship [525]. Although the guidance document for GIVIMP by the Organization for Economic Co-operation and Development is tailored to cell and tissue-based *in vitro* methods, the FAIR guiding principles are just as relevant for microbiology, where the shift to -omics approaches has resulted in increasing data volume, complexity, and ease of generation. More than ever, it is important to publish the data in its proper context (metadata). Lastly, as recommended in the GIVIMP, reporting of results from an *in vitro* method should also include all related documents and the modifications to improve the method and the rationale for them, which can be facilitated by protocol-based journals such as Nature Protocols and JoVE. Model optimization also includes the modifications that did not work out, which is also gaining more attention with the publication of negative results in journals such as the Journal of Negative Results in Biomedicine, but might as well, in my opinion, deserve a place in publications of the (eventual) positive results.

With the rapid advancement of -omics technologies, as well as the increasing interest in NATs such as microfluidic devices and organoid cultures, scientists should be aware of the reporting issues that are still present in both *in vitro* and *in vivo* research. There are already several guidelines and frameworks in place or under development to ensure proper reporting of (to-be-developed) *in vitro* models and thus, my suggestion is an open door: It is the responsibility of scientists, but also reviewers, editors or users of the *in vitro* models to adhere and give input to existing guidelines, especially when adding another level of complexity by introducing bacteria.

## Two steps forward

### 1. Putting the pieces together

Once the building blocks have been properly defined, it is time to put the pieces together. Such a bottom-up approach to model the host-microbe ecosystem must occur in a systematic manner, in which the bidirectional effects of bacteria and bacterial products on each other and on the cellular *in vitro* model are accurately assessed. Although this may sound like a challenging and costly endeavor, it will certainly deliver additional advantages. First, a systematic approach allows a careful distinction of the effects of model parameters on outcomes of interest, such as mucus production. An example is provided by the bidirectional effects of oxygen levels, non-epithelial cell types, shear stress on both microbe and mucus, as discussed above. Causality of these interactions are hard to dissect *in vivo*. Second, systematic investigation of bacterial effects on host parameters could revisit current cell culture



practice, for instance, by using bacterial products such as SCFAs to enhance differentiation [699]. Another elegant example was demonstrated by a recent study in which the effect of butyrate was assessed on (mucin) O-glycosylation in Caco-2 cells [726]. Third, a systematic approach helps to further refine mathematical models of host-microbe interactions, and *vice versa*, as will be explained in the next section.

Next to disentangling individual interactions of all model parameters, the final aim would be to benchmark the *in vitro* model to the *in vivo* situation of the human GIT, of which our understanding is still far from complete. As discussed in **Chapter 1** and **2**, scientists have relied heavily on animal models that are not representative of the human GIT. Therefore, next to the bottom-up approach in which model parameters (including bacteria) are characterized in a systematic manner, there is still a (future) role for organoid(-on-chip) models and other bio-inspired *in vitro* methods that are currently under development, to provide a complementary top-down approach. Especially the use of donor-derived organoids will broaden our scope on the inter- and intra-individual variability that exists as a consequence of e.g., host diet, age, genetics, environment: model parameters which we, one day, hope to be able to integrate in *in vitro* models of the human GIT. Moreover, more representative models of the GIT provide a more realistic way to foster the growth of bacteria, compared to current *in vitro* bioreactor models, and increase our knowledge of strain- and species-specific behavior *in vitro*. A current obstacle in this, however, is the limitation to sample the respective, residing microbial community from the GIT side of interest, as opposed to the use of fecal samples. Using defined communities of representative species would provide a promising alternative.

Thus, there will eventually be a need for joint efforts: on one hand, relatively simple models of intestinal bacteria (single species, metabolites, cell components) and the host part of the human GIT (cell lines on inserts), as well as, on the other hand, the assembly of representative (defined) microbial communities and more advanced *in vitro* methods. The power of this complementary approach (top-down plus bottom-up) is not only clear from the resulting fundamental understanding of human host-microbe interactions, but also its practical application: before investigation in expensive organoid (on-chip) models, there will still always be need for simpler models to allow high-throughput screening of compounds of interests.

## 2. Integration of data

Standardization and proper reporting of host and microbial model parameters will also improve current mathematical models that are receiving increasing attention to gain system-level understanding of microbe-microbe and host-microbe interactions and generate new hypotheses. Mathematical models exist at the level of individual bacteria and the level of the population, with and without interactions or a host component (Reviewed in [727, 728]). As proposed by others, a continuous feedback loop between mathematic models and experimental models allows refinement of

mathematical models by *in vitro* (bacterial) culture-based models on one hand (e.g. by increasing the repertoire of cultivated species and strains) and the refinement of culture-based models by mathematical predictions on the other hand (e.g. the prediction of defined media to improve cultivation strategies [729]). Together, they can serve as an important step towards predictive modelling of the *in vivo* situation [728].

From the host point of view, the fields of pharmacology and toxicology make use of physiologically based pharmacokinetic (PBPK) models to predict the absorption, distribution, metabolism, and excretion, of chemical substances. After the first major quantification of the metabolic capacity of the intestinal microbiota to transform, accumulate or detoxify drugs and other toxic compounds [730], recently, a PBPK model was developed to disentangle microbial and host contribution to metabolism of a certain drug [731]. Such *in silico* models could also help to refine *in vitro* models of the human GIT commonly used in the field of pharmacology and toxicology.

These examples demonstrate the potential power of mathematical models to support the integration of bacteria in *in vitro* models of the human GIT. (Big) data gets an increasingly prominent role in science, with the rapid development of artificial intelligence-based tools to systematically analyze -omics data without the aid of human intervention and facilitate the discovery of unprecedented patterns. As also expressed by others (e.g. [524, 732]), I feel the field of human intestinal microbiology is not ready for this step yet: As set out in this thesis (**Chapter 2, 3** and this chapter), standardization and proper reporting are prerequisites for in-depth analysis of host-microbe interactions.

### 3. Applications of a complete *in vitro* model of the human colon

Once a complete *in vitro* model of the human colon has been obtained, including a mucus layer as well as an intestinal microbiota, the applications are limitless.

#### Development of host-specific models

The importance of investigating bacterial species in a host-specific context has already been demonstrated by animal models (**Chapter 1** and **2**). Whereas microbiotas of humans and mice are functionally similar, differences at species level exist [178, 179]. As discussed, humanized mice would provide a solution, but based on a recent systematic analysis, researchers warned that care should be taken when inferring causality from these models, and called for more robust animal experiments [733]. For instance, modelling immune-mediated disorders still require host-specific colonization and behavior of the intestinal gut microbiota [734-736]. Recently, the public deposition of murine isolates as well as their taxonomical description have been further expanded by the Mouse Intestinal Bacterial Collection (miBC [737, 738]), Mouse Gut Microbial biobank [739] and Mouse Gastrointestinal Bacteria Catalogue [740], which pave the way for future design of synthetic communities in *in vivo* models.



For instance, based on the metagenomes of newly added strains to the miBC, the Oligo-MM12 was recently expanded to 19 species, improving phenotypic differences between OMM12 and SPF-mice [738]. These developments contribute to the refinement of *in vivo* models that are still deemed necessary. Most importantly, they inspire the development of host-specific *in vitro* models: research in a human gut-on-a-chip demonstrated that the presence of human microbial metabolites more strongly increased the pathogenicity of an enteric pathogen, compared to metabolites from murine microbial species [741]. Next to the human-mouse comparison, there is also inter-individual variability in bacterial colonization. Specifically in the context of the mucus layer, for instance, colonization by *R. gnavus* was found to be host-specific and dependent on mucin glycosylation [742]. In this context, the use of *in vitro* models including an intestinal microbiota would offer another advantage over animal models: in theory, co-cultivation of donor-derived intestinal tissue (organoids) and microbiota would allow a host-specific strategy in which the intestinal tissue is matched with its native microbial strains.

#### Patient-specific preclinical models

In line with host-specificity, organoid(-on-chip)s have also been promoted as a personalized strategy to test patient-specific and GIT-site-specific effects of drugs and have even been proposed to eventually replace animals as preclinical models for a wide range of diseases [743]. Amongst others, organoids from patients with environmental enteric disease, CRC, Crohn's disease, and ulcerative colitis have been successfully cultivated on a chip [743, 744]. As most of these conditions have also been associated with an aberrant microbiota, there is an increasing interest in the use of microbial therapies, including pre-, pro- and postbiotics. To thoroughly investigate their effect, these therapeutics need to be studied in the context of the colonic mucus layer and the residing microbial community, as opposed to current *in vitro* models in which the microbiota or host are separated.

#### Investigation of host components on bacteria

*In vitro* models of the human colon including a living microbiota allow the investigation of host components on bacteria, which is now limited to animal studies. First of all, it is known that host genetics can directly shape the intestinal microbiota [31], but the exact mechanisms remain unclear and are probably a mix of different factors. More specifically regarding the mucus layer, there is also a need to look beyond mucin only: the mucus layer is a reservoir of compounds that contribute to the intestinal barrier, including defensins, proteases, immunoglobulins [104]. Moreover, mucin glycans have not only shown to influence colonization, but also to function as regulatory signals tuning microbial behavior (reviewed in [745]). Next to mucus, several other host-secreted compounds have been shown to influence the intestinal microbiota directly or indirectly, including miRNA [746], lectins (e.g. Intelectin-1 [747]), some of them packaged in ECVs (reviewed in [748]).

### Investigation of host-bacteria adaption

Furthermore, (long-term) co-cultivation of intestinal cells with bacteria and thus, the investigation of host-bacterium adaptation over time should be given more attention. Specifically in the context of the mucus layer, for instance, it would be interesting to investigate how mucin glycosylation evolves over time, upon its first encounter with intestinal bacteria and vice versa, how it selects for certain bacterial populations. One could even think of exploiting long-term co-culture models to study host-bacterium evolution over time, although it was recently argued that co-evolution between mammals and individual bacteria might not be as widespread as thought previously [749]. Moreover, this scientific quest is further complicated by the sterility of the laboratory environment as opposed to constant, natural exposure of human beings to microbes, one of the main drivers for the development of the wildling mouse model, in which the tractable genetics of laboratory mice is combined with the natural environment of wild mice [736].

### Development of bacterial tools

Lastly, another fascinating application of *in vitro* models including live bacteria is the employment of bacteria as a tool to quantify or manipulate the *in vitro* conditions. For instance, an engineered fluorescent reporter strain of the enteric bacterium *Yersinia pseudotuberculosis* was employed to detect oxygen levels *in situ*, in a 3D intestine model [675]. Similarly, the readily cultivable, genetically accessible bacterium *B. thetaiotaomicron* has been proposed as a model organism for intestinal bacteria [234]. Amongst other applications, engineered *B. thetaiotaomicron* has been used *in vivo* to visualize strain-specific crypt localization [750] and regulate bacterial gene expression *in situ* [303]. Recently, engineered strains of *E. coli*, the go-to bacterium in molecular biology, were employed to record transcriptional history of bacterial strains in the murine gastrointestinal tract [751]. Because of the non-invasiveness (as long as safety and biocontainment concerns are addressed) and low costs, genetically engineered bacteria can be employed to diagnose and/or treat diseases such as cancer, diabetes, colitis, cholera (reviewed in e.g. [752]). The application of engineered bacteria or “smart probiotics” ([752] is only expected to increase, since recently, scientists managed to genetically edit complex bacterial communities in a species- and site-specific manner [753]. An elegant example of the application of engineered bacteria in a gut-on-a-chip was recently demonstrated for the treatment of phenylketonuria [754].

### Towards a systems approach

Overall, a complete *in vitro* model of the human GIT, including an intestinal microbiota or microbial compounds, could be considered a “plug & play” model in which both the bacterial side and the host side can be tightly controlled, and the complex puzzle of host-bacteria symbiosis can be completely dis- and re-assembled. The current





challenges as well as the future applications are summarized in **Figure 1**. On one hand, such *in vitro* models provide a basic understanding of the myriad of bidirectional host-microbe interactions taking place in the human intestinal mucosal layer, refining, and reducing the use of current *in vivo* models. Moreover, these *in vitro* models hold great promises for the development of (microbial) therapies in case of intestinal disorders or disorders associated with an aberrant intestinal microbiota composition, eventually replacing the need for animal models. All in all, the integration of the microbiota in *in vitro* models of the human GIT would fit within the systems approach that is needed to fully understand the holobiont [755].

## HOW? A CALL FOR SLOW SCIENCE

Taking a step back, as proposed in this chapter, may be an unpopular message within the highly competitive world of academia. To obtain grants, scientists need to write ambitious proposals for novel *in vitro* models, advanced techniques, and the development of innovative therapeutic strategies. This allows, however, little room for further characterization of existing, simpler *in vitro* models that might be less sexy, but serve as a valuable resource on the long term. In my opinion, it does fit within a bigger trend scientists are calling for: slow science [756, 757].

The number of papers published per scientist has increased during the last decades [758, 759], which is only partially due to more opportunities for collaboration [760, 761]. Most prominently, there has been an increasing selection for productive scientists [762]. This culture of “publish or perish” has already received a lot of criticism in academia [763, 764], upon which scientists have frequently called for “slow science” [756, 757]. Several roads to slow science have been proposed [756], of which one is to take a different perspective on timescales and keep the bigger aim of science in mind (i.e. seeking the truth). To start with, the creation and investigation of big theories requires more than a scientist’s lifetime, which is in sharp contrast with the current speed at which early-career scientists need to accumulate citations and secure grants. Instead, scientists should be encouraged to join long-term projects, which can be offered by another road to slow science: collaboration. Clearly, scientists can learn from microbes as being true team players. Microbial communities – defined or undefined <sup>4</sup> – have already taught us that a diverse community, with various skills and own (sometimes opposing) nutrient preferences, contribute more than the sum of its parts. The same has been demonstrated for human beings with different perspectives [765], strengthening also the value of the current peer review system

---

<sup>4</sup> A defined community – a selection of experts from different fields – sounds like the most logical option. However, when going for an undefined community – a sample from a naturally diverse population – you never know what kind of undiscovered talents you may encounter

[766]. Thus, this concept is not new and, based on absolute numbers [767] and funder's ambitions (e.g. the Dutch Research Council [768]), there is already an upward trend in collaborative research visible during the past years. Another road to slow science could be the assessment of quality over quantity. In fact, it has been shown empirically that the current selection for high output leads to poor methodological practices and increasingly high false discovery rates, resulting in a natural selection for bad science [769]. As opposed to measures such as publication numbers and h-indices, however, both team play and quality are difficult to quantify in, for instance, job and grant application procedures. Solutions would include the description of authors as contributors [770, 771] and the recent proposition to revise the recognition and reward system by the Dutch public knowledge institutions and research funders [772]. The latter does, amongst others, not only include the recognition of quality over quantity (for instance, by requesting a narrative CV), but also recognizes team performance.

In summary, there are multiple roads to slow science, some of which already have been taken. Personally, apart from scientific practice in academia, I think two other disciplines or initiatives would benefit from slow science as well: 1) science education and communication and 2) lab sustainability, which I will explain in the final two sections.

### **Increasing microbiology literacy in society (and science!)**

As mentioned in the first paragraph of this thesis, the human microbiota was long considered an immunological threat. Among the general public, microbes still have a doubtful reputation, which has definitely not improved by the recent COVID-19 pandemic. The importance of (science) communication in the preparation for/control of/measures against a pandemic has been discussed exhaustively (e.g., [773, 774]). Specifically in the context of this thesis, I would like to stress the urgent need for microbiology literacy in society, as proposed by others: "Microbiology literacy needs to become part of the world citizen job description" [775]. The reason is simple: microbes facilitate life on earth globally and locally, by performing an amazingly broad range of (bio)chemical processes. From a more anthropocentric point of view, this is illustrated by the critical functions the human (intestinal) microbiota plays in our health, as demonstrated in this thesis. In addition, microbes have served humankind greatly by providing fermented food and drinks, maintenance of soil fertility and remediation of drinking water. Moreover, microbial biotechnology is expected to promote sustainable economic growth and contribute to the Sustainable Development Goals set by the United Nations [776]. Such expectations, however, also contribute to the other extreme of microbial literacy, i.e., an exaggerated believe that microbes are going to solve all our problems. This too, is best illustrated in the context of the human intestinal microbiota, which has been linked to an overwhelming number of diseases and thus, as we may believe popular science media, could contain the cure



to all these conditions [3]. This belief is further supported by funders and journals, which often require scientists to mention specific potential applications of a scientific discovery, even if this concerns fundamental biology.

Note that the aforementioned (mis)conceptions of microbiology are not limited to non-scientists. Anecdotally, the inventors of the aforementioned 3D intestine used anthropocentric phrases like “infected with human microbiome” when co-cultivating intestinal cells with a human-derived microbiota [683]. In addition, the GIVIMP document only mentions bacteria in the context of cell culture contamination [484]. Similarly, the overoptimistic perspective is presented in the first paper to report the co-culture of intestinal cells with a human-derived microbiota on a chip. The chip is presented as a tool for the development of microbiome-related therapeutics, probiotics, and nutraceuticals, but this is hardly supported by the respective study. Microbial ecology is not mentioned in this paper, *i.e.* the focus is only on the microbial composition, whereas the microbial contribution to the model (such as SCFA production) is not assessed nor addressed [212].

Thus, we need to slow down the general public’s expectations, but also those of funders, journalists, and popular science media, while still feeding the passion for science itself. Here too, a trend towards slow science could play a critical role, by focusing on team play as well as research quality.

The emphasis on team play first applies to the microbiome research field itself, in which microbiome scientists recently acknowledged the fragmentation of microbiome research [755]. Microbiomes are often studied one ecosystem at-a-time, whereas they are actually interconnected [777]. A systems approach was proposed that connects research between scientific fields to enhance our collective understanding of microbiomes. Specifically, regarding awareness and microbiome literacy, the importance to inform and educate policymakers, regulators, farmers citizens and other stakeholders was stressed [755]. Similarly, a systems approach might apply to this thesis, in which I propose a more holobiont view of *in vitro* models of the human GIT. Clearly, there is a role for microbiologists or microbiology educators to cross-fertilize other scientific disciplines in the development of *in vitro* models of an organ that is so densely colonized by intestinal bacteria. It re-emphasizes the need for integration of microbiology courses in biotechnological, (bio)medical, toxicological, and pharmaceutical study programs (*i.e.*, students that are trained to work with or interpret data from *in vitro* models) as well as interdisciplinary research projects involving *in vitro* models of the human GIT.

Second, when engaging the public in (microbiology) science, team play is key as well. To increase microbial literacy in society, researchers have proposed a role for microbiologists, communicators, disseminators, and educators [775, 778]. The contribution of social scientists, with expertise in models of science communication, has also been stressed [779] For instance, whereas many scientists have long adhered to the so-called “deficit model” (*i.e.* a one-way approach in which science communication

is aimed to fill knowledge gaps), there is a call for bi-directional forms of science communication, where dialogue and participation play an equally important role [780, 781]. Thus, *vice versa*, there is also a role for microbiologists to get support in science communication strategies.

Further, the communication of science to both (fellow) scientists and non-scientists could be improved by focusing on quality instead of quantity. As discussed, journals allowing publication of negative results are already in place, but ironically, they add up to the continuously growing pile of publications a scientist has to keep up with [756], possibly escaping the attention of colleagues who perform similar experiments. At the same time, from my experience, (early career) scientists are taught to present their positive results in an attractive way and (over)sell their science, leaving out their failures. The absence of failing role models has been acknowledged in the context of slow science before [756]. The same is true when having to deal with science journalists and popular science media, who want to write an attractive story that is understandable and – most importantly – inspiring to the general public. In some cases, the scientist is presented as a superhero or genius. This may sound boring, but I feel this does not justify the actual common aim of science and adds even more pressure on (early-career) scientists to conduct supposedly “groundbreaking research”.

I acknowledge that the task for microbiologists or scientists in general to exchange knowledge and maintain the dialogue with fellow scientists and the general public, adds up to the already existing responsibilities of a scientist. Here too, however, I feel the term “slow science” is justified, in which we try to stop re-inventing the wheel within our own field but seek opportunities to collaborate with experts from other disciplines and be open on our successes as well as (unavoidable) mistakes. In line with the latter, the true story behind my thesis is explained in **Box 2**.

### Lab sustainability

Lastly, today’s scientists have to deal with another responsibility: their ecological footprint. First of all, this includes the (long distance) travel by plane for international conferences [782], although the recent pandemic has also taught us the benefits of online meetings. Less obvious, at least for outsiders, is the plastic pollution by scientific laboratories, due to the use of plastic disposables for convenience and speed. In 2014, it was estimated that bioscientific research was responsible for 1.8 % of the total global plastic production [783]. Luckily, suppliers of lab consumables are increasingly seeking to reduce packaging materials or promote refill and recycle opportunities (e.g. [784, 785]). When working with microbiologically contaminated materials, however, the alternatives are scarce if you do not want to sacrifice speed, convenience, and reproducibility. Personally, I feel that the speed and convenience scientists got used to, further contributes to overconsumption of consumables, by providing the temptation of trial-and-error rather than careful experimental planning.



Therefore, slow science, including quality-focused, long-term projects, could also have a positive impact on the amount of plastic pollution generated in labs.

On a final note, the ecological footprint of scientists is further increased when animal studies are involved. The 3Rs to improve animal welfare in science have been a cornerstone of this thesis. A still overlooked aspect, however, is the use of animal-derived compounds in (bacterial) cell culture, such as FBS, antibodies and PGM in this thesis. A worldwide survey among (mostly academic) scientists revealed that only half of the participants acknowledged the ethical concerns of animal-derived materials and one third was not aware of the availability of animal-free alternatives [722]. Consequently, FBS is still the supplement of choice, despite high batch-to-batch variability (as discussed); reports from twenty (!) years ago on the development of chemically defined media [786]; and the launch of a serum-free media database (not accessible anymore [787]). On a positive note, however, there is an increasing interest in cultured meat, for which the use of serum-free medium is deemed a prerequisite [788] and hopefully, accelerates the development of serum-free cell culture [789-791]. Potentially, there is also a role for microbes to produce non-animal compounds for cell culture [790].

Did I not mention, microbes are everywhere?

## BOX 2: THE TRUE STORY

It is extremely easy to criticize the work of others, even – or especially – given the fact that published work only reports positive results, omitting or hiding the negative findings that contributed immensely to the final product. In this section, I would like to tell you the true story behind the chapters in this thesis and the mistakes I made, some of which have become essential for the finalization of this thesis. Because a large part of the points in the discussion could not have been raised if I had known all of this in advance...

Literature reviews are a valuable resource for scientists to get an overview of the current state of the art in their field. However, I feel that the literature review in **Chapter 2** did not necessarily have to be written. The main incentive for this review was the fact that I had to dive into literature anyway (during the first months of my PhD) and that our department rewarded bonuses for PhD students that published a review in the first year of their trajectory<sup>5</sup>. Thus, by writing this review, I also contributed to the “publish or perish” culture, just for the sake of multitasking (reading and writing at the same time) and a financial compensation. Still, I hope the comprehensive analysis of gnotobiotic animal models has provided a valuable resource for scientists interested in host-microbe interactions in the human GIT tract.

A similar story exists for **Chapter 3**, which was mainly published because of the pandemic, in which we had to work from home. In hindsight, I really enjoyed writing this chapter, but we could have known that comparison of Caco-2 cellular response between labs is extremely difficult, based on a review already published in 2005 (!). Therefore, the key message of this chapter just re-emphasizes statements made previously by others.

The SWMS method used in **Chapter 4** was published by others in 2013 and later adapted by incorporation of a 3D scaffold. It has not been reported frequently, probably because it was caught up by gut-on-a-chips or similar, more controllable models. Actually, the establishment of a gut-on-a-chip was the original aim of my PhD project, but due to technical and infrastructural limitations, this failed partially and we had to look for an alternative plan. Because of our interest in mucolytic bacteria, we came across the SWMS method. Was it worth to perform an expensive transcriptome analysis on this method? Or should we have informed us

---

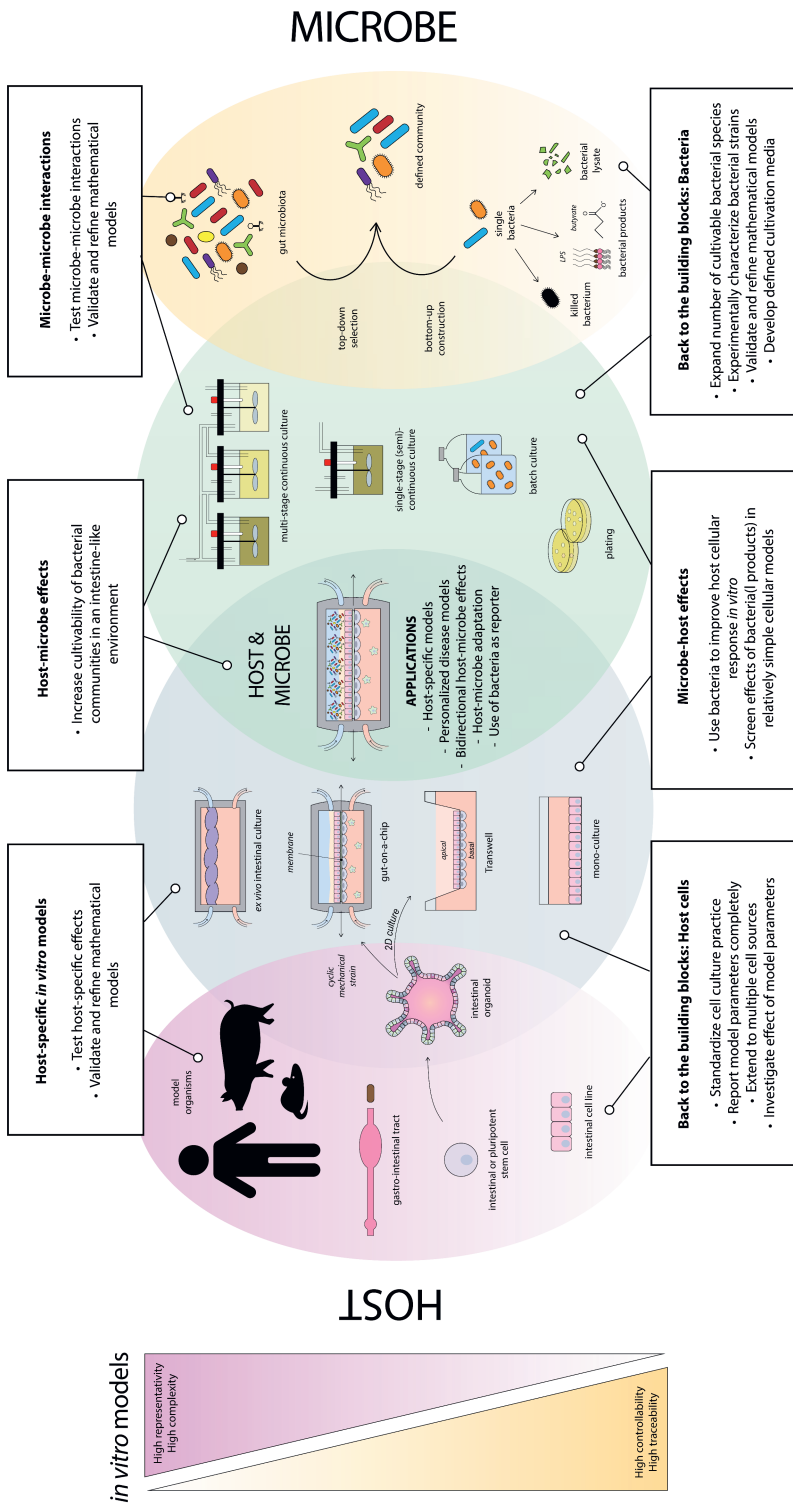
<sup>5</sup> In the meanwhile, this bonus system has been discontinued



better by talking to the scientists that initially published this method? In hindsight, if we had additional money and time, at least it would have been informative to look at gene expression levels at multiple time points to gain a more comprehensive insight in the model.

We decided to continue with the SWMS model in **Chapter 5**, and study the effect of mucus-associated bacteria on the mucus layer. Moreover, transcriptomic analysis would allow us to compare the response with previous findings on the cultivation method. I made the mistake, however, to use a different RNA extraction kit and did not consider including a DNase step. The use of RNAseq instead of microarray, bounced back: Its sensitivity is merciless. The data turned out to be still valid for downstream analysis, but indicated a lack of effect. Therefore, the choice for pasteurized bacteria might not have been the best and should have been backed up with a more systematic pre-screening of microbial compounds in a simpler experimental set-up, as proposed in **Chapter 7**. In addition, the incubation time of 24 h was a shot in the dark and mainly based on others' work with different bacteria. In the same chapter, we also introduced a trimmed down version of the originally proposed gut-on-a-chip, luckily with seemingly less DNA contamination, but we did not manage to analyze the data in time for the thesis deadline.

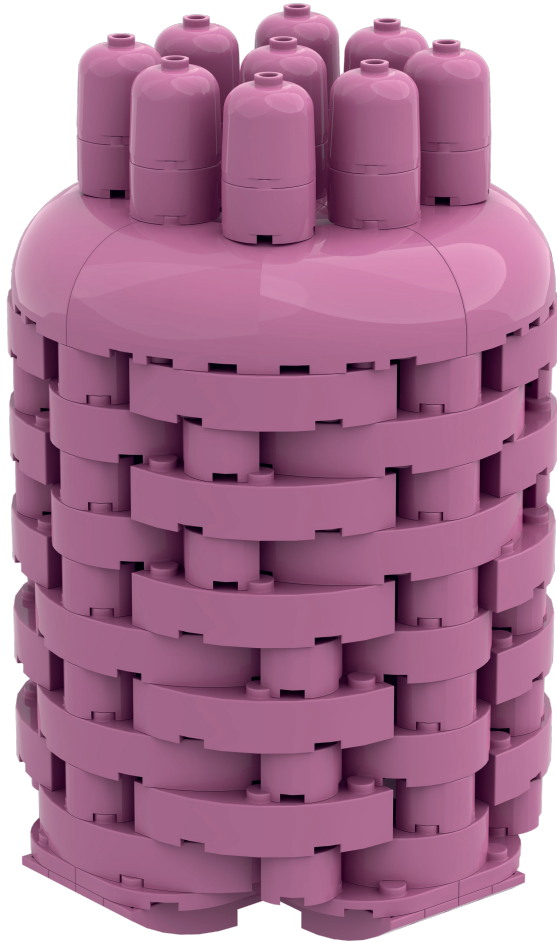
**Chapter 6** is my favorite chapter, but has also been a bumpy road. It turned out I had trouble with inactivating bacterial cells: Initially, I did not manage to properly pasteurize *A. muciniphila* (using a thermoblock instead of a water bath). In addition, despite extensive mammalian cell culture training in Wageningen, I got contamination in my HEK-cell culture during the first week in Copenhagen. You can also put it differently: Although I was trained as an eukaryotic cell biologist, I might be a true microbiologist after all.



**Figure 1.** *In vitro* models of the human GIT, intestinal bacteria and their interactions. Models are organized from bottom to top, with the most representative and complex at the top and the most controllable and traceable - with respect to host parameters or bacterial species - at the bottom. The current challenges and future applications of the integration of host and bacteria *in vitro* are given for each (combination of) model(s)



A Caco-2 cell pretending  
to be an appendix.



(Obviously, the author was too  
lazy to build an appendix)

# APPENDICES

## References

1. Berg, G., et al., *Microbiome definition re-visited: old concepts and new challenges*. *Microbiome*, 2020. **8**(1).
2. Bordenstein, S.R. and K.R. Theis, *Host Biology in Light of the Microbiome: Ten Principles of Holobionts and Hologenomes*. *PLOS Biology*, 2015. **13**(8): p. e1002226.
3. Bik, E.M., *The Hoops, Hopes, and Hypes of Human Microbiome Research*. *Yale J Biol Med*, 2016. **89**(3): p. 363-373.
4. Boron, W.F. and E.L. Boulpaep, *Medical physiology : a cellular and molecular approach*. Updated second edition [updated 2nd ed.] ed. 2012.
5. Sekirov, I., et al., *Gut microbiota in health and disease*. *Physiol Rev*, 2010. **90**(3): p. 859-904.
6. Human Microbiome Project, C., *Structure, function and diversity of the healthy human microbiome*. *Nature*, 2012. **486**(7402): p. 207-14.
7. Qin, J., et al., *A human gut microbial gene catalogue established by metagenomic sequencing*. *Nature*, 2010. **464**(7285): p. 59-65.
8. Quince, C., et al., *Shotgun metagenomics, from sampling to analysis*. *Nat Biotechnol*, 2017. **35**(9): p. 833-844.
9. Rinke, C., et al., *Insights into the phylogeny and coding potential of microbial dark matter*. *Nature*, 2013. **499**(7459): p. 431-7.
10. Nielsen, H.B., et al., *Identification and assembly of genomes and genetic elements in complex metagenomic samples without using reference genomes*. *Nature Biotechnology*, 2014. **32**(8): p. 822-828.
11. Hugon, P., et al., *A comprehensive repertoire of prokaryotic species identified in human beings*. *The Lancet Infectious Diseases*, 2015. **15**(10): p. 1211-1219.
12. Almeida, A., et al., *A new genomic blueprint of the human gut microbiota*. *Nature*, 2019. **568**(7753): p. 499-504.
13. Oren, A. and G.M. Garrity, *Valid publication of the names of forty-two phyla of prokaryotes*. *International Journal of Systematic and Evolutionary Microbiology*, 2021. **71**(10).
14. Dridi, B., et al., *High Prevalence of Methanobrevibacter smithii and Methanospaera stadtmannae Detected in the Human Gut Using an Improved DNA Detection Protocol*. *PLOS ONE*, 2009. **4**(9): p. e7063.
15. Scanlan, P.D. and J.R. Marchesi, *Micro-eukaryotic diversity of the human distal gut microbiota: qualitative assessment using culture-dependent and -independent analysis of faeces*. *The ISME Journal*, 2008. **2**(12): p. 1183-1193.
16. Browne, H.P., et al., *Culturing of 'unculturable' human microbiota reveals novel taxa and extensive sporulation*. *Nature*, 2016. **533**(7604): p. 543-546.
17. Pasolli, E., et al., *Extensive Unexplored Human Microbiome Diversity Revealed by Over 150,000 Genomes from Metagenomes Spanning Age, Geography, and Lifestyle*. *Cell*, 2019. **176**(3): p. 649-662.e20.
18. Barrila, J., et al., *Organotypic 3D cell culture models: using the rotating wall vessel to study host-pathogen interactions*. *Nat Rev Microbiol*, 2010. **8**(11): p. 791-801.

19. Sender, R., S. Fuchs, and R. Milo, *Revised Estimates for the Number of Human and Bacteria Cells in the Body*. PLOS Biology, 2016. **14**(8): p. e1002533.
20. Sender, R., S. Fuchs, and R. Milo, *Are We Really Vastly Outnumbered? Revisiting the Ratio of Bacterial to Host Cells in Humans*. Cell, 2016. **164**(3): p. 337-340.
21. Dominguez-Bello, M.G., et al., *Delivery mode shapes the acquisition and structure of the initial microbiota across multiple body habitats in newborns*. Proceedings of the National Academy of Sciences, 2010. **107**(26): p. 11971-11975.
22. Salonen, A. and W.M. de Vos, *Impact of Diet on Human Intestinal Microbiota and Health*. Annual review of food science and technology, 2014. **5**(1): p. 239-262.
23. Zoetendal, E.G. and W.M. de Vos, *Effect of diet on the intestinal microbiota and its activity*. Current opinion in gastroenterology, 2014. **30**(2): p. 189-195.
24. Flint, H.J., et al., *The role of the gut microbiota in nutrition and health*. Nat Rev Gastroenterol Hepatol, 2012. **9**(10): p. 577-89.
25. Johnson, A.J., et al., *Daily Sampling Reveals Personalized Diet-Microbiome Associations in Humans*. Cell Host Microbe, 2019. **25**(6): p. 789-802.e5.
26. Gibson, G.R., *Dietary Modulation of the Human Gut Microflora Using the Prebiotics Oligofructose and Inulin*. The Journal of Nutrition, 1999. **129**(7): p. 1438S-1441S.
27. Bonder, M.J., et al., *The effect of host genetics on the gut microbiome*. Nat Genet, 2016. **48**(11): p. 1407-1412.
28. Turpin, W., et al., *Association of host genome with intestinal microbial composition in a large healthy cohort*. Nat Genet, 2016. **48**(11): p. 1413-1417.
29. Hughes, D.A., et al., *Genome-wide associations of human gut microbiome variation and implications for causal inference analyses*. Nat Microbiol, 2020. **5**(9): p. 1079-1087.
30. Kurilshikov, A., et al., *Large-scale association analyses identify host factors influencing human gut microbiome composition*. Nat Genet, 2021. **53**(2): p. 156-165.
31. Goodrich, J.K., et al., *Human genetics shape the gut microbiome*. Cell, 2014. **159**(4): p. 789-99.
32. Rothschild, D., et al., *Environment dominates over host genetics in shaping human gut microbiota*. Nature, 2018. **555**(7695): p. 210-215.
33. Zhernakova, A., et al., *Population-based metagenomics analysis reveals markers for gut microbiome composition and diversity*. Science, 2016. **352**(6285): p. 565-9.
34. Xie, H., et al., *Shotgun Metagenomics of 250 Adult Twins Reveals Genetic and Environmental Impacts on the Gut Microbiome*. Cell Syst, 2016. **3**(6): p. 572-584.e3.
35. Yatsunenkov, T., et al., *Human gut microbiome viewed across age and geography*. Nature, 2012. **486**(7402): p. 222-7.
36. David, L.A., et al., *Diet rapidly and reproducibly alters the human gut microbiome*. Nature, 2014. **505**(7484): p. 559-63.



## APPENDICES

37. Godinho-Silva, C., et al., *Light-entrained and brain-tuned circadian circuits regulate ILC3s and gut homeostasis*. *Nature*, 2019. **574**(7777): p. 254-258.
38. Thaïss, Christoph A., et al., *Transkingdom Control of Microbiota Diurnal Oscillations Promotes Metabolic Homeostasis*. *Cell*, 2014. **159**(3): p. 514-529.
39. Blaser, M.J., *Antibiotic use and its consequences for the normal microbiome*. *Science*, 2016. **352**(6285): p. 544-5.
40. Maier, L., et al., *Extensive impact of non-antibiotic drugs on human gut bacteria*. *Nature*, 2018. **555**(7698): p. 623-628.
41. Sousa, T., et al., *The gastrointestinal microbiota as a site for the biotransformation of drugs*. *Int J Pharm*, 2008. **363**(1-2): p. 1-25.
42. Manichanh, C., et al., *Reduced diversity of faecal microbiota in Crohn's disease revealed by a metagenomic approach*. *Gut*, 2006. **55**(2): p. 205-11.
43. Ott, S.J., et al., *Reduction in diversity of the colonic mucosa associated bacterial microflora in patients with active inflammatory bowel disease*. *Gut*, 2004. **53**(5): p. 685-93.
44. Arthur, J.C., et al., *Intestinal inflammation targets cancer-inducing activity of the microbiota*. *Science*, 2012. **338**(6103): p. 120-3.
45. Janney, A., F. Powrie, and E.H. Mann, *Host-microbiota maladaptation in colorectal cancer*. *Nature*, 2020. **585**(7826): p. 509-517.
46. Nougayrède, J.-P., et al., *Escherichia coli Induces DNA Double-Strand Breaks in Eukaryotic Cells*. *Science*, 2006. **313**(5788): p. 848-851.
47. Ou, J., et al., *Diet, microbiota, and microbial metabolites in colon cancer risk in rural Africans and African Americans*. *The American Journal of Clinical Nutrition*, 2013. **98**(1): p. 111-120.
48. Ley, R.E., et al., *Microbial ecology: human gut microbes associated with obesity*. *Nature*, 2006. **444**(7122): p. 1022-3.
49. Turnbaugh, P.J., et al., *An obesity-associated gut microbiome with increased capacity for energy harvest*. *Nature*, 2006. **444**(7122): p. 1027-31.
50. Vrieze, A., et al., *Transfer of intestinal microbiota from lean donors increases insulin sensitivity in individuals with metabolic syndrome*. *Gastroenterology*, 2012. **143**(4): p. 913-6.e7.
51. Scheperjans, F., et al., *Gut microbiota are related to Parkinson's disease and clinical phenotype*. *Movement Disorders*, 2015. **30**(3): p. 350-358.
52. Perez-Pardo, P., et al., *Role of TLR4 in the gut-brain axis in Parkinson's disease: a translational study from men to mice*. *Gut*, 2019. **68**(5): p. 829-843.
53. Schmidt, T.S.B., J. Raes, and P. Bork, *The Human Gut Microbiome: From Association to Modulation*. *Cell*, 2018. **172**(6): p. 1198-1215.
54. Donaldson, G.P., S.M. Lee, and S.K. Mazmanian, *Gut biogeography of the bacterial microbiota*. *Nat Rev Microbiol*, 2016. **14**(1): p. 20-32.
55. Tierney, B.T., et al., *The Landscape of Genetic Content in the Gut and Oral Human Microbiome*. *Cell Host Microbe*, 2019. **26**(2): p. 283-295.e8.
56. Hooper, L.V., et al., *Molecular analysis of commensal host-microbial relationships in the intestine*. *Science*, 2001. **291**(5505): p. 881-4.
57. Bäckhed, F., et al., *Host-Bacterial Mutualism in the Human Intestine*. *Science*, 2005. **307**(5717): p. 1915-1920.

58. Olszak, T., et al., *Microbial Exposure During Early Life Has Persistent Effects on Natural Killer T Cell Function*. Science (New York, N.Y.), 2012. **336**(6080): p. 489-493.
59. Kau, A.L., et al., *Human nutrition, the gut microbiome and the immune system*. Nature, 2011. **474**(7351): p. 327-336.
60. Fukuda, S., et al., *Bifidobacteria can protect from enteropathogenic infection through production of acetate*. Nature, 2011. **469**(7331): p. 543-547.
61. Sommer, F. and F. Bäckhed, *The gut microbiota--masters of host development and physiology*. Nat Rev Microbiol, 2013. **11**(4): p. 227-38.
62. Kaoutari, A.E., et al., *The abundance and variety of carbohydrate-active enzymes in the human gut microbiota*. Nature Reviews Microbiology, 2013. **11**(7): p. 497-504.
63. Bhattacharya, T., T.S. Ghosh, and S.S. Mande, *Global Profiling of Carbohydrate Active Enzymes in Human Gut Microbiome*. PLOS ONE, 2015. **10**(11): p. e0142038.
64. Cummings, J.H., et al., *Short chain fatty acids in human large intestine, portal, hepatic and venous blood*. Gut, 1987. **28**(10): p. 1221-7.
65. Macfarlane, G.T. and S. Macfarlane, *Bacteria, colonic fermentation, and gastrointestinal health*. J AOAC Int, 2012. **95**(1): p. 50-60.
66. Koh, A., et al., *From Dietary Fiber to Host Physiology: Short-Chain Fatty Acids as Key Bacterial Metabolites*. Cell, 2016. **165**(6): p. 1332-1345.
67. Kovatcheva-Datchary, P., et al., *Linking phylogenetic identities of bacteria to starch fermentation in an in vitro model of the large intestine by RNA-based stable isotope probing*. Environ Microbiol, 2009. **11**(4): p. 914-26.
68. Barker, N., et al., *Identification of stem cells in small intestine and colon by marker gene Lgr5*. Nature, 2007. **449**(7165): p. 1003-1007.
69. Sato, T., et al., *Single Lgr5 stem cells build crypt-villus structures in vitro without a mesenchymal niche*. Nature, 2009. **459**(7244): p. 262-265.
70. Karam, S.M., *Lineage commitment and maturation of epithelial cells in the gut*. Front Biosci, 1999. **4**: p. D286-98.
71. van der Flier, L.G. and H. Clevers, *Stem cells, self-renewal, and differentiation in the intestinal epithelium*. Annu Rev Physiol, 2009. **71**: p. 241-60.
72. Pullan, R.D., et al., *Thickness of adherent mucus gel on colonic mucosa in humans and its relevance to colitis*. Gut, 1994. **35**(3): p. 353-9.
73. Atuma, C., et al., *The adherent gastrointestinal mucus gel layer: Thickness and physical state in vivo*. American Journal of Physiology - Gastrointestinal and Liver Physiology, 2001. **280**(5 43-5): p. G922-G929.
74. Ermund, A., et al., *Studies of mucus in mouse stomach, small intestine, and colon. I. Gastrointestinal mucus layers have different properties depending on location as well as over the Peyer's patches*. American Journal of Physiology-Gastrointestinal and Liver Physiology, 2013. **305**(5): p. G341-G347.
75. Sonnenburg, J.L., L.T. Angenent, and J.I. Gordon, *Getting a grip on things: how do communities of bacterial symbionts become established in our intestine?* Nat Immunol, 2004. **5**(6): p. 569-73.
76. Sheng, Y.H., et al., *Mucins in inflammatory bowel diseases and colorectal cancer*. Journal of Gastroenterology and Hepatology, 2012. **27**(1): p. 28-38.



## APPENDICES

77. McGuckin, M.A., et al., *Mucin dynamics and enteric pathogens*. Nature Reviews Microbiology, 2011. **9**(4): p. 265-278.
78. Johansson, M.E., H. Sjövall, and G.C. Hansson, *The gastrointestinal mucus system in health and disease*. Nat Rev Gastroenterol Hepatol, 2013. **10**(6): p. 352-61.
79. Larhed, A.W., P. Artursson, and E. Björk, *The influence of intestinal mucus components on the diffusion of drugs*. Pharm Res, 1998. **15**(1): p. 66-71.
80. Witas, H., et al., *Lipids associated with rat small-intestinal mucus glycoprotein*. Carbohydr Res, 1983. **120**: p. 67-76.
81. Johansson, M.E.V., et al., *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, 2008. **105**(39): p. 15064-15069.
82. Gendler, S.J. and A.P. Spicer, *Epithelial Mucin Genes*. Annual Review of Physiology, 1995. **57**(1): p. 607-634.
83. Perez-Vilar, J. and R.L. Hill, *The structure and assembly of secreted mucins*. J Biol Chem, 1999. **274**(45): p. 31751-4.
84. Johansson, M.E.V., et al., *Composition and functional role of the mucus layers in the intestine*. Cellular and Molecular Life Sciences, 2011. **68**(22): p. 3635.
85. Corfield, A.P., *Mucins: A biologically relevant glycan barrier in mucosal protection*. Biochimica et Biophysica Acta (BBA) - General Subjects, 2015. **1850**(1): p. 236-252.
86. Jonckheere, N. and I. Van Seuningen, *The membrane-bound mucins: From cell signalling to transcriptional regulation and expression in epithelial cancers*. Biochimie, 2009. **92**(1): p. 1-11.
87. Carrington, S., et al., *Microbial interaction with mucus and mucins*, in *Microbial Glycobiology: Structures, Relevance and Applications*. 2009, Elsevier Science.
88. Desseyn, J.-L., D. Tetaert, and V. Gouyer, *Architecture of the large membrane-bound mucins*. Gene, 2008. **410**(2): p. 215-222.
89. Ambort, D., et al., *Function of the CysD domain of the gel-forming MUC2 mucin*. Biochem J, 2011. **436**(1): p. 61-70.
90. van de Bovenkamp, J., et al., *Molecular cloning of human gastric mucin MUC5AC reveals conserved cysteine-rich D-Domains and a putative leucine zipper motif*. Biochemical and Biophysical Research Communications, 1998. **245**(3): p. 853-859.
91. Porchet, N., et al., *Human mucin genes: genomic organization and expression of MUC4, MUC5AC and MUC5B*. Biochem Soc Trans, 1995. **23**(4): p. 800-5.
92. Pigny, P., et al., *Human mucin genes assigned to 11p15.5: identification and organization of a cluster of genes*. Genomics, 1996. **38** 3: p. 340-52.
93. Nguyen, V.C., et al., *Assignment of human tracheobronchial mucin gene(s) to 11p15 and a tracheobronchial mucin-related sequence to chromosome 13*. Human genetics, 1990. **86**(2): p. 167-172.
94. Meezaman, D., et al., *Cloning and analysis of cDNA encoding a major airway glycoprotein, human tracheobronchial mucin (MUC5)*. The Journal of biological chemistry, 1994. **269**(17): p. 12932-12939.
95. Toribara, N.W., et al., *Human gastric mucin. Identification of a unique species by expression cloning*. Journal of Biological Chemistry, 1993. **268**(8): p. 5879-5885.

96. Vinall, L.E., et al., *Variable number tandem repeat polymorphism of the mucin genes located in the complex on 11p15.5*. Hum Genet, 1998. **102**(3): p. 357-66.
97. Gum, J.R., et al., *Molecular cloning of human intestinal mucin cDNAs. Sequence analysis and evidence for genetic polymorphism*. J Biol Chem, 1989. **264**(11): p. 6480-7.
98. Griffiths, B., et al., *Assignment of the polymorphic intestinal mucin gene (MUC2) to chromosome 11p15*. Ann Hum Genet, 1990. **54**(4): p. 277-85.
99. Audie, J.P., et al., *Expression of human mucin genes in respiratory, digestive, and reproductive tracts ascertained by in situ hybridization*. J Histochem Cytochem, 1993. **41**(10): p. 1479-85.
100. Asker, N., et al., *Dimerization of the human MUC2 mucin in the endoplasmic reticulum is followed by a N-glycosylation-dependent transfer of the mono- and dimers to the Golgi apparatus*. J Biol Chem, 1998. **273**(30): p. 18857-63.
101. Asker, N., et al., *The human MUC2 mucin apoprotein appears to dimerize before O-glycosylation and shares epitopes with the 'insoluble' mucin of rat small intestine*. Biochem J, 1995. **308** ( Pt 3)(Pt 3): p. 873-80.
102. Godl, K., et al., *The N terminus of the MUC2 mucin forms trimers that are held together within a trypsin-resistant core fragment*. J Biol Chem, 2002. **277**(49): p. 47248-56.
103. Ambort, D., et al., *Calcium and pH-dependent packing and release of the gel-forming MUC2 mucin*. Proc Natl Acad Sci U S A, 2012. **109**(15): p. 5645-50.
104. Johansson, M.E., et al., *The inner of the two Muc2 mucin-dependent mucus layers in colon is devoid of bacteria*. Proc Natl Acad Sci U S A, 2008. **105**(39): p. 15064-9.
105. Johansson, M.E.V., K.A. Thomsson, and G.C. Hansson, *Proteomic Analyses of the Two Mucus Layers of the Colon Barrier Reveal That Their Main Component, the Muc2 Mucin, Is Strongly Bound to the Fcgbp Protein*. Journal of Proteome Research, 2009. **8**(7): p. 3549-3557.
106. Gustafsson, J.K., et al., *Bicarbonate and functional CFTR channel are required for proper mucin secretion and link cystic fibrosis with its mucus phenotype*. The Journal of experimental medicine, 2012. **209**(7): p. 1263-1272.
107. Nilsson, H.E., et al., *Intestinal MUC2 Mucin Supramolecular Topology by Packing and Release Resting on D3 Domain Assembly*. Journal of Molecular Biology, 2014. **426**(14): p. 2567-2579.
108. Nyström, E.E.L., et al., *Calcium-activated Chloride Channel Regulator 1 (CLCA1) Controls Mucus Expansion in Colon by Proteolytic Activity*. EBioMedicine, 2018. **33**: p. 134-143.
109. Schjoldager, K.T., et al., *Global view of human protein glycosylation pathways and functions*. Nature Reviews Molecular Cell Biology, 2020. **21**(12): p. 729-749.
110. Bennett, E.P., et al., *Control of mucin-type O-glycosylation: A classification of the polypeptide GalNAc-transferase gene family*. Glycobiology, 2012. **22**(6): p. 736-756.
111. Brockhausen I, Schachter H, and S. P., *O-GalNAc Glycans*, in *Essentials of Glycobiology*, Varki A, et al., Editors. 2009, Cold Spring Harbor Laboratory Press: Cold Spring Harbor (NY).





## APPENDICES

112. Larsson, J.M., et al., A complex, but uniform O-glycosylation of the human MUC2 mucin from colonic biopsies analyzed by nanoLC/MSn. *Glycobiology*, 2009. **19**(7): p. 756-66.
113. Robbe, C., et al., Structural diversity and specific distribution of O-glycans in normal human mucins along the intestinal tract. *Biochem J*, 2004. **384**(Pt 2): p. 307-16.
114. Robbe-Masselot, C., et al., Expression of a core 3 disialyl-Le(x) hexasaccharide in human colorectal cancers: a potential marker of malignant transformation in colon. *Journal of proteome research*, 2009. **8**(2): p. 702-711.
115. Xia, L., Core 3-derived O-glycans are essential for intestinal mucus barrier function. *Methods Enzymol*, 2010. **479**: p. 123-41.
116. Klein, A. and P. Roussel, O-acetylation of sialic acids. *Biochimie*, 1998. **80**(1): p. 49-57.
117. Brockhausen, I., Sulphotransferases acting on mucin-type oligosaccharides. *Biochem Soc Trans*, 2003. **31**(2): p. 318-25.
118. Mollicone, R., et al., Immunohistologic pattern of type 1 (Lea, Leb) and type 2 (X, Y, H) blood group-related antigens in the human pyloric and duodenal mucosae. *Lab Invest*, 1985. **53**(2): p. 219-27.
119. Paone, P. and P.D. Cani, Mucus barrier, mucins and gut microbiota: the expected slimy partners? *Gut*, 2020. **69**(12): p. 2232-2243.
120. Johansson, M.E., et al., Normalization of Host Intestinal Mucus Layers Requires Long-Term Microbial Colonization. *Cell Host Microbe*, 2015. **18**(5): p. 582-92.
121. Durbán, A., et al., Assessing Gut Microbial Diversity from Feces and Rectal Mucosa. *Microbial Ecology*, 2011. **61**(1): p. 123-133.
122. Eckburg, P.B., et al., Diversity of the human intestinal microbial flora. *Science*, 2005. **308**(5728): p. 1635-8.
123. Zoetendal, E.G., et al., Mucosa-associated bacteria in the human gastrointestinal tract are uniformly distributed along the colon and differ from the community recovered from feces. *Appl Environ Microbiol*, 2002. **68**(7): p. 3401-7.
124. Albenberg, L., et al., Correlation between intraluminal oxygen gradient and radial partitioning of intestinal microbiota. *Gastroenterology*, 2014. **147**(5): p. 1055-63.e8.
125. Hooper, L.V. and J.I. Gordon, Glycans as legislators of host-microbial interactions: spanning the spectrum from symbiosis to pathogenicity. *Glycobiology*, 2001. **11**(2): p. 1r-10r.
126. Robbe, C., et al., Evidence of Regio-specific Glycosylation in Human Intestinal Mucins: PRESENCE OF AN ACIDIC GRADIENT ALONG THE INTESTINAL TRACT \*. *Journal of Biological Chemistry*, 2003. **278**(47): p. 46337-46348.
127. McLoughlin, K., et al., Host Selection of Microbiota via Differential Adhesion. *Cell Host & Microbe*, 2016. **19**(4): p. 550-559.
128. Szentkúti, L., et al., Pre-epithelial mucus layer in the colon of conventional and germ-free rats. *The Histochemical Journal*, 1990. **22**(9): p. 491-497.
129. Enss, M.L., et al., Changes in colonic mucins of germfree rats in response to the introduction of a "normal" rat microbial flora. *Rat colonic mucin. Journal of experimental animal science*, 1992. **35**(3): p. 110-119.

130. Sharma, R., et al., *Rat intestinal mucosal responses to a microbial flora and different diets*. Gut, 1995. **36**(2): p. 209.
131. Rausch, P., et al., *Colonic mucosa-associated microbiota is influenced by an interaction of Crohn disease and FUT2 (Secretor) genotype*. Proceedings of the National Academy of Sciences, 2011. **108**(47): p. 19030-19035.
132. Larsson, J.M.H., et al., *Altered O-glycosylation profile of MUC2 mucin occurs in active ulcerative colitis and is associated with increased inflammation*. Inflammatory Bowel Diseases, 2011. **17**(11): p. 2299-2307.
133. McGovern, D.P., et al., *Fucosyltransferase 2 (FUT2) non-secretor status is associated with Crohn's disease*. Hum Mol Genet, 2010. **19**(17): p. 3468-76.
134. Parmar, A.S., et al., *Association study of FUT2 (rs601338) with celiac disease and inflammatory bowel disease in the Finnish population*. Tissue Antigens, 2012. **80**(6): p. 488-493.
135. An, G., et al., *Increased susceptibility to colitis and colorectal tumors in mice lacking core 3-derived O-glycans*. Journal of Experimental Medicine, 2007. **204**(6): p. 1417-1429.
136. Cone, R.A., *Barrier properties of mucus*. Advanced Drug Delivery Reviews, 2009. **61**(2): p. 75-85.
137. Owen, C.D., et al., *Unravelling the specificity and mechanism of sialic acid recognition by the gut symbiont Ruminococcus gnavus*. Nat Commun, 2017. **8**(1): p. 2196.
138. Juge, N., *Microbial adhesins to gastrointestinal mucus*. Trends Microbiol, 2012. **20**(1): p. 30-9.
139. Cohen, L.J., et al., *Unraveling function and diversity of bacterial lectins in the human microbiome*. Nat Commun, 2022. **13**(1): p. 3101.
140. Cantarel, B.L., et al., *The Carbohydrate-Active EnZymes database (CAZy): an expert resource for Glycogenomics*. Nucleic Acids Res, 2009. **37**(Database issue): p. D233-8.
141. Ficko-Blean, E. and A.B. Boraston, *Insights into the recognition of the human glycome by microbial carbohydrate-binding modules*. Curr Opin Struct Biol, 2012. **22**(5): p. 570-7.
142. Sonnenburg, J.L., et al., *Glycan Foraging in Vivo by an Intestine-Adapted Bacterial Symbiont*. Science (New York, N.Y.), 2005. **307**(5717): p. 1955-1959.
143. Belzer, C., et al., *Microbial Metabolic Networks at the Mucus Layer Lead to Diet-Independent Butyrate and Vitamin B(12) Production by Intestinal Symbionts*. mBio, 2017. **8**(5).
144. Ng, K.M., et al., *Microbiota-liberated host sugars facilitate post-antibiotic expansion of enteric pathogens*. Nature, 2013. **502**(7469): p. 96-9.
145. Png, C.W., et al., *Mucolytic bacteria with increased prevalence in IBD mucosa augment in vitro utilization of mucin by other bacteria*. Am J Gastroenterol, 2010. **105**(11): p. 2420-8.
146. Carroll, I.M. and N. Maharshak, *Enteric bacterial proteases in inflammatory bowel disease- pathophysiology and clinical implications*. World J Gastroenterol, 2013. **19**(43): p. 7531-43.
147. Vergnolle, N., *Protease inhibition as new therapeutic strategy for GI diseases*. Gut, 2016. **65**(7): p. 1215-24.



## APPENDICES

148. Shon, D.J., et al., *Classification, structural biology, and applications of mucin domain-targeting proteases*. *Biochem J*, 2021. **478**(8): p. 1585-1603.
149. Martens, E.C., et al., *Complex glycan catabolism by the human gut microbiota: the Bacteroidetes Sus-like paradigm*. *J Biol Chem*, 2009. **284**(37): p. 24673-7.
150. Tailford, L.E., et al., *Mucin glycan foraging in the human gut microbiome*. *Frontiers in Genetics*, 2015. **6**.
151. Crost, E.H., et al., *Utilisation of mucin glycans by the human gut symbiont Ruminococcus gnavus is strain-dependent*. *PLoS One*, 2013. **8**(10): p. e76341.
152. Birchenough, G.M., et al., *A sentinel goblet cell guards the colonic crypt by triggering Nlrp6-dependent Muc2 secretion*. *Science*, 2016. **352**(6293): p. 1535-42.
153. Knoop, K.A., et al., *Microbial sensing by goblet cells controls immune surveillance of luminal antigens in the colon*. *Mucosal Immunol*, 2015. **8**(1): p. 198-210.
154. Willemsen, L.E., et al., *Short chain fatty acids stimulate epithelial mucin 2 expression through differential effects on prostaglandin E(1) and E(2) production by intestinal myofibroblasts*. *Gut*, 2003. **52**(10): p. 1442-7.
155. Freitas, M., et al., *A heat labile soluble factor from Bacteroides thetaiotaomicron VPI-5482 specifically increases the galactosylation pattern of HT29-MTX cells*. *Cell Microbiol*, 2001. **3**(5): p. 289-300.
156. Graziani, F., et al., *Ruminococcus gnavus E1 modulates mucin expression and intestinal glycosylation*. *J Appl Microbiol*, 2016. **120**(5): p. 1403-17.
157. Derrien, M., et al., *Akkermansia muciniphila* gen. nov., sp. nov., a human intestinal mucin-degrading bacterium. *Int J Syst Evol Microbiol*, 2004. **54**(Pt 5): p. 1469-1476.
158. Belzer, C. and W.M. de Vos, *Microbes inside--from diversity to function: the case of Akkermansia*. *Isme j*, 2012. **6**(8): p. 1449-58.
159. van Passel, M.W., et al., *The genome of Akkermansia muciniphila, a dedicated intestinal mucin degrader, and its use in exploring intestinal metagenomes*. *PLoS One*, 2011. **6**(3): p. e16876.
160. Ouwerkerk, J.P., et al., *Adaptation of Akkermansia muciniphila to the Oxic-Anoxic Interface of the Mucus Layer*. *Appl Environ Microbiol*, 2016. **82**(23): p. 6983-6993.
161. Collado, M.C., et al., *Intestinal integrity and Akkermansia muciniphila, a mucin-degrading member of the intestinal microbiota present in infants, adults, and the elderly*. *Appl Environ Microbiol*, 2007. **73**(23): p. 7767-70.
162. Cani, P.D. and W.M. de Vos, *Next-generation beneficial microbes : The case of Akkermansia muciniphila*. *Frontiers in microbiology*, 2017. **8**(NA): p. 1765-1765.
163. Depommier, C., et al., *Supplementation with Akkermansia muciniphila in overweight and obese human volunteers: a proof-of-concept exploratory study*. *Nat Med*, 2019. **25**(7): p. 1096-1103.
164. Plovier, H., et al., *A purified membrane protein from Akkermansia muciniphila or the pasteurized bacterium improves metabolism in obese and diabetic mice*. *Nat Med*, 2017. **23**(1): p. 107-113.

165. Depommier, C., et al., *Pasteurized Akkermansia muciniphila increases whole-body energy expenditure and fecal energy excretion in diet-induced obese mice.* Gut Microbes, 2020. **11**(5): p. 1231-1245.
166. Cani, P.D., et al., *Akkermansia muciniphila: paradigm for next-generation beneficial microorganisms.* Nat Rev Gastroenterol Hepatol, 2022. **19**(10): p. 625-637.
167. Belzer, C., et al., *Microbial Metabolic Networks at the Mucus Layer Lead to Diet-Independent Butyrate and Vitamin B<sub>12</sub> Production by Intestinal Symbionts.* mBio, 2017. **8**(5): p. e00770-17.
168. Nuttall, G.H.F. and H. Thierfelder., *Thierisches Leben ohne Bakterien im Verdauungskanal.* 1897. **23**(3): p. 231-235.
169. Trexler, P.C. and R.P. Orcutt, *Development of Gnotobiotics and Contamination Control in Laboratory Animal Science.* 50 years of Laboratory Animal Science, AALAS, 2006: p. 1-8.
170. Thompson, G.R. and P.C. Trexler, *Gastrointestinal structure and function in germ-free or gnotobiotic animals.* Gut, 1971. **12**(3): p. 230-235.
171. Schaedler, R.W., R. Dubos, and R. Costello, *Association of germfree mice with bacteria isolated from normal mice.* The Journal of experimental medicine, 1965. **122**(1): p. 77-82.
172. Treuting, P.M., et al., *Of mice and microflora: considerations for genetically engineered mice.* Vet Pathol, 2012. **49**(1): p. 44-63.
173. Hugenholtz, F. and W.M. de Vos, *Mouse models for human intestinal microbiota research: a critical evaluation.* Cellular and molecular life sciences : CMLS, 2017. **75**(1): p. 149-160.
174. Nguyen, T.L.A., et al., *How informative is the mouse for human gut microbiota research.* Disease models & mechanisms, 2015. **8**(1): p. 1-16.
175. McConnell, E.L., A.W. Basit, and S. Murdan, *Measurements of rat and mouse gastrointestinal pH, fluid and lymphoid tissue, and implications for in-vivo experiments.* J Pharm Pharmacol, 2008. **60**(1): p. 63-70.
176. Sheridan, W.G., R.H. Lowndes, and H.L. Young, *Intraoperative tissue oximetry in the human gastrointestinal tract.* Am J Surg, 1990. **159**(3): p. 314-9.
177. Holmén Larsson, J.M., et al., *Studies of mucus in mouse stomach, small intestine, and colon. III. Gastrointestinal Muc5ac and Muc2 mucin O-glycan patterns reveal a regiospecific distribution.* Am J Physiol Gastrointest Liver Physiol, 2013. **305**(5): p. G357-63.
178. Ley, R.E., et al., *Obesity alters gut microbial ecology.* Proceedings of the National Academy of Sciences of the United States of America, 2005. **102**(31): p. 11070-11075.
179. Krych, L., et al., *Quantitatively Different, yet Qualitatively Alike: A Meta-Analysis of the Mouse Core Gut Microbiome with a View towards the Human Gut Microbiome.* PloS one, 2013. **8**(5): p. e62578-NA.
180. Hazenberg, M.P., M. Bakker, and A. Verschoor-Burggraaf, *Effects of the human intestinal flora on germ-free mice.* J Appl Bacteriol, 1981. **50**(1): p. 95-106.
181. Raibaud, P., et al., *Implantation of bacteria from the digestive tract of man and various animals into gnotobiotic mice.* Am J Clin Nutr, 1980. **33**(11 Suppl): p. 2440-7.



## APPENDICES

182. Stappenbeck, T.S. and H.W. Virgin, *Accounting for reciprocal host-microbiome interactions in experimental science*. *Nature*, 2016. **534**(7606): p. 191-199.
183. Wos-Oxley, M., et al., *Comparative evaluation of establishing a human gut microbial community within rodent models*. *Gut Microbes*, 2012. **3**(3): p. 234-49.
184. Litten-Brown, J.C., A.M. Corson, and L. Clarke, *Porcine models for the metabolic syndrome, digestive and bone disorders: a general overview*. *Animal*, 2010. **4**(6): p. 899-920.
185. Hooda, S., et al., *Current state of knowledge: the canine gastrointestinal microbiome*. *Anim Health Res Rev*, 2012. **13**(1): p. 78-88.
186. Swanson, K.S., et al., *Phylogenetic and gene-centric metagenomics of the canine intestinal microbiome reveals similarities with humans and mice*. *Isme j*, 2011. **5**(4): p. 639-49.
187. Moeller, A.H., et al., *Rapid changes in the gut microbiome during human evolution*. *Proc Natl Acad Sci U S A*, 2014. **111**(46): p. 16431-5.
188. Wettenbestand, B., *Wet op Dierproeven*.
189. EUR-Lex, *Directive 2010/63/EU of the European Parliament and of the Council of 22 September 2010 on the protection of animals used for scientific purposes*.
190. Russell, W.M.S. and R.L. Burch, *The principles of humane experimental technique*. 1959: Methuen.
191. McGonigle, P. and B. Ruggeri, *Animal models of human disease: challenges in enabling translation*. *Biochem Pharmacol*, 2014. **87**(1): p. 162-71.
192. Fogh, J., J.M. Fogh, and T. Orfeo, *One hundred and twenty-seven cultured human tumor cell lines producing tumors in nude mice*. *J Natl Cancer Inst*, 1977. **59**(1): p. 221-6.
193. Chantret, I., et al., *Epithelial polarity, villin expression, and enterocytic differentiation of cultured human colon carcinoma cells: a survey of twenty cell lines*. *Cancer Res*, 1988. **48**(7): p. 1936-42.
194. Matsumoto, H., et al., *Biosynthesis of alkaline phosphatase during differentiation of the human colon cancer cell line Caco-2*. *Gastroenterology*, 1990. **98**(5 Pt 1): p. 1199-207.
195. Pinto, M., et al., *Enterocyte-like differentiation and polarization of the human colon carcinoma cell line Caco-2 in culture*. *Biol.Cell*, 1983. **47**: p. 323-330.
196. Donato, R.P., et al., *Studying Permeability in a Commonly Used Epithelial Cell Line: T84 Intestinal Epithelial Cells*, in *Permeability Barrier: Methods and Protocols*, K. Turksen, Editor. 2011, Humana Press: Totowa, NJ. p. 115-137.
197. Van Klinken, B.J.W., et al., *The human intestinal cell lines Caco-2 and LS174T as models to study cell-type specific mucin expression*. *Glycoconjugate journal*, 1996. **13**(5): p. 757-768.
198. Dexter, D.L., et al., *Heterogeneity of cancer cells from a single human colon carcinoma*. *Am J Med*, 1981. **71**(6): p. 949-56.
199. Lesuffleur, T., et al., *Growth adaptation to methotrexate of HT-29 human colon carcinoma cells is associated with their ability to differentiate into columnar absorptive and mucus-secreting cells*. *Cancer Res*, 1990. **50**(19): p. 6334-43.

200. Hilgendorf, C., et al., *Caco-2 versus Caco-2/HT29-MTX Co-cultured Cell Lines: Permeabilities Via Diffusion, Inside- and Outside-Directed Carrier-Mediated Transport*. Journal of pharmaceutical sciences, 2000. **89**(1): p. 63-75.
201. Araújo, F. and B. Sarmiento, *Towards the characterization of an in vitro triple co-culture intestine cell model for permeability studies*. Int J Pharm, 2013. **458**(1): p. 128-34.
202. Ulluwishewa, D., et al., *Live Faecalibacterium prausnitzii in an apical anaerobic model of the intestinal epithelial barrier*. Cellular Microbiology, 2015. **17**(2): p. 226-240.
203. Sadaghian Sadabad, M., et al., *A simple coculture system shows mutualism between anaerobic faecalibacteria and epithelial Caco-2 cells*. Scientific Reports, 2016. **5**(1): p. 17906-17906.
204. Gayer, C.P. and M.D. Basson, *The effects of mechanical forces on intestinal physiology and pathology*. Cell Signal, 2009. **21**(8): p. 1237-44.
205. Mahler, G.J., et al., *Characterization of a gastrointestinal tract microscale cell culture analog used to predict drug toxicity*. Biotechnology and Bioengineering, 2009. **104**(1): p. 193-205.
206. Sung, J.H., C. Kam, and M.L. Shuler, *A microfluidic device for a pharmacokinetic-pharmacodynamic (PK-PD) model on a chip*. Lab Chip, 2010. **10**(4): p. 446-55.
207. Imura, Y., et al., *A Microfluidic System to Evaluate Intestinal Absorption*. Analytical Sciences, 2009. **25**(12): p. 1403-1407.
208. Kim, H.J., et al., *Human gut-on-a-chip inhabited by microbial flora that experiences intestinal peristalsis-like motions and flow*. Lab on a chip, 2012. **12**(12): p. 2165-2174.
209. Kim, H.J. and D.E. Ingber, *Gut-on-a-Chip microenvironment induces human intestinal cells to undergo villus differentiation*. Integrative biology : quantitative biosciences from nano to macro, 2013. **5**(9): p. 1130-1140.
210. Kim, H.J., et al., *Contributions of microbiome and mechanical deformation to intestinal bacterial overgrowth and inflammation in a human gut-on-a-chip*. Proceedings of the National Academy of Sciences, 2016. **113**(1): p. E7-E15.
211. Kasendra, M., et al., *Development of a primary human Small Intestine-on-a-Chip using biopsy-derived organoids*. Scientific reports, 2018. **8**(1): p. 2871-2871.
212. Jalili-Firoozinezhad, S., et al., *A Complex Human-Gut Microbiome Cultured in an Anaerobic Intestine-on-a-Chip*. Nature biomedical engineering, 2019. **3**(7): p. 520-531.
213. Randall, K.J., J. Turton, and J.R. Foster, *Explant culture of gastrointestinal tissue: a review of methods and applications*. Cell Biol Toxicol, 2011. **27**(4): p. 267-84.
214. Li, M., I.A. de Graaf, and G.M. Groothuis, *Precision-cut intestinal slices: alternative model for drug transport, metabolism, and toxicology research*. Expert Opin Drug Metab Toxicol, 2016. **12**(2): p. 175-90.
215. Ussing, H.H. and K. Zerahn, *Active transport of sodium as the source of electric current in the short-circuited isolated frog skin*. Acta Physiol Scand, 1951. **23**(2-3): p. 110-27.



## APPENDICES

216. Wilson, T.H. and G. Wiseman, *The use of sacs of everted small intestine for the study of the transference of substances from the mucosal to the serosal surface.* J Physiol, 1954. **123**(1): p. 116-25.
217. TNO. Available from: [https://www.tno.nl/media/4327/intestine\\_food.pdf](https://www.tno.nl/media/4327/intestine_food.pdf).
218. Sung, J.H., et al., *Microscale 3-D hydrogel scaffold for biomimetic gastrointestinal (GI) tract model.* Lab on a Chip, 2011. **11**(3): p. 389-392.
219. Sato, T., et al., *Long-term expansion of epithelial organoids from human colon, adenoma, adenocarcinoma, and Barrett's epithelium.* Gastroenterology, 2011. **141**(5): p. 1762-1772.
220. Sato, T., et al., *Single Lgr5 stem cells build crypt-villus structures in vitro without a mesenchymal niche.* Nature 2009 459:7244, 2009. **459**(7244): p. 262-265.
221. Sato, T. and H. Clevers, *Growing self-organizing mini-guts from a single intestinal stem cell: mechanism and applications.* Science, 2013. **340**(6137): p. 1190-4.
222. Spence, J.R., et al., *Directed differentiation of human pluripotent stem cells into intestinal tissue in vitro.* Nature, 2011. **470**(7332): p. 105-9.
223. Vandussen, K.L., et al., *Development of an enhanced human gastrointestinal epithelial culture system to facilitate patient-based assays.* Gut, 2016. **64**(6): p. 911-920.
224. Sontheimer-Phelps, A., et al., *Human Colon-on-a-Chip Enables Continuous In Vitro Analysis of Colon Mucus Layer Accumulation and Physiology.* Cellular and molecular gastroenterology and hepatology, 2019. **9**(3): p. 507-526.
225. Sasaki, N., et al., *Development of a Scalable Coculture System for Gut Anaerobes and Human Colon Epithelium.* Gastroenterology, 2020. **159**(1): p. 388-390.e5.
226. Schloss, P.D. and J. Handelsman, *Metagenomics for studying unculturable microorganisms: cutting the Gordian knot.* Genome Biol, 2005. **6**(8): p. 229.
227. Lagier, J.C., et al., *Culturing the human microbiota and culturomics.* Nat Rev Microbiol, 2018. **16**: p. 540-550.
228. Walker, A.W., et al., *Phylogeny, culturing, and metagenomics of the human gut microbiota.* Trends Microbiol, 2014. **22**(5): p. 267-74.
229. Gibson, G.R., J.H. Cummings, and G.T. Macfarlane, *Use of a three-stage continuous culture system to study the effect of mucin on dissimilatory sulfate reduction and methanogenesis by mixed populations of human gut bacteria.* Appl Environ Microbiol, 1988. **54**(11): p. 2750-5.
230. Van Den Abbeele, P., et al., *Microbial community development in a dynamic gut model is reproducible, colon region specific, and selective for bacteroidetes and Clostridium cluster IX.* Applied and Environmental Microbiology, 2010. **76**(15): p. 5237-5246.
231. Mäkituokko, H., et al., *In vitro effects on polydextrose by colonic bacteria and caco-2 cell cyclooxygenase gene expression.* Nutr Cancer, 2005. **52**(1): p. 94-104.
232. Cinquin, C., et al., *New three-stage in vitro model for infant colonic fermentation with immobilized fecal microbiota.* FEMS Microbiology Ecology, 2006. **57**(2): p. 324-336.

233. Grosskopf, T. and O.S. Soyer, *Synthetic microbial communities*. Curr Opin Microbiol, 2014. **18**(100): p. 72-7.
234. Wexler, A.G. and A.L. Goodman, *An insider's perspective: Bacteroides as a window into the microbiome*. Nat Microbiol, 2017. **2**: p. 17026.
235. Johansson, M.E. and G.C. Hansson, *Preservation of mucus in histological sections, immunostaining of mucins in fixed tissue, and localization of bacteria with FISH*. Methods Mol Biol, 2012. **842**: p. 229-35.
236. Hasegawa, Y., et al., *Preservation of three-dimensional spatial structure in the gut microbiome*. PLoS One, 2017. **12**(11): p. e0188257.
237. Earle, K.A., et al., *Quantitative Imaging of Gut Microbiota Spatial Organization*. Cell Host Microbe, 2015. **18**(4): p. 478-88.
238. Mark Welch, J.L., et al., *Spatial organization of a model 15-member human gut microbiota established in gnotobiotic mice*. Proceedings of the National Academy of Sciences, 2017: p. 201711596-201711596.
239. Gustafsson, J.K., et al., *An ex vivo method for studying mucus formation, properties, and thickness in human colonic biopsies and mouse small and large intestinal explants*. Am J Physiol Gastrointest Liver Physiol, 2012. **302**(4): p. G430-8.
240. Sardelli, L., et al., *Towards bioinspired in vitro models of intestinal mucus*. RSC Adv, 2019. **9**(28): p. 15887-15899.
241. Niv, Y., et al., *Mucin synthesis and secretion in relation to spontaneous differentiation of colon cancer cells in vitro*. Int J Cancer, 1992. **50**(1): p. 147-52.
242. Lesuffleur, T., et al., *Differential expression of the human mucin genes MUC1 to MUC5 in relation to growth and differentiation of different mucus-secreting HT-29 cell subpopulations*. J Cell Sci, 1993. **106** ( Pt 3): p. 771-83.
243. Navabi, N., M.A. McGuckin, and S.K. Lindén, *Gastrointestinal Cell Lines Form Polarized Epithelia with an Adherent Mucus Layer when Cultured in Semi-Wet Interfaces with Mechanical Stimulation*. PloS one, 2013. **8**(7): p. e68761-NA.
244. Vavasseur, F., et al., *Synthesis of O-glycan core 3: characterization of UDP-GlcNAc: GalNAc-R beta 3-N-acetyl-glucosaminyltransferase activity from colonic mucosal tissues and lack of the activity in human cancer cell lines*. Glycobiology, 1995. **5**(3): p. 351-7.
245. Probert, H. and G. Gibson, *Development of a fermentation system to model sessile bacterial populations in the human colon*. Biofilms, 2004. **1**(1): p. 13-19.
246. Van den Abbeele, P., et al., *Incorporating a mucosal environment in a dynamic gut model results in a more representative colonization by lactobacilli*. Microbial Biotechnology, 2012. **5**(1): p. 106-115.
247. Marzorati, M., et al., *The HMI™ module: a new tool to study the Host-Microbiota Interaction in the human gastrointestinal tract in vitro*. BMC microbiology, 2014. **14**(1): p. 133-133.
248. Shah, P., et al., *A microfluidics-based in vitro model of the gastrointestinal human-microbe interface*. Nature communications, 2016. **7**(1): p. 11535-11535.
249. Karlsson, N.G., et al., *Glycosylation differences between pig gastric mucin populations: a comparative study of the neutral oligosaccharides using mass spectrometry*. Biochem J, 1997. **326** ( Pt 3)(Pt 3): p. 911-7.





## APPENDICES

250. Nordman, H., et al., *Gastric MUC5AC and MUC6 are large oligomeric mucins that differ in size, glycosylation and tissue distribution*. *Biochem J*, 2002. **364**(Pt 1): p. 191-200.
251. Werlang, C., G. Cárcarmo-Oyarce, and K. Ribbeck, *Engineering mucus to study and influence the microbiome*. *Nature Reviews Materials*, 2019. **4**(2): p. 134-145.
252. Dunne, C., et al., *Use of Recombinant Mucin Glycoprotein to Assess the Interaction of the Gastric Pathogen Helicobacter pylori with the Secreted Human Mucin MUC5AC*. *Bioengineering (Basel)*, 2017. **4**(2).
253. Ryan, A., et al., *Expression and characterization of a novel recombinant version of the secreted human mucin MUC5AC in airway cell lines*. *Biochemistry*, 2015. **54**(4): p. 1089-99.
254. Nason, R., et al., *Display of the human mucinome with defined O-glycans by gene engineered cells*. *Nat Commun*, 2021. **12**(1): p. 4070.
255. Liu, J., et al., *O-glycan repertoires on a mucin-type reporter protein expressed in CHO cell pools transiently transfected with O-glycan core enzyme cDNAs*. *J Biotechnol*, 2015. **199**: p. 77-89.
256. Bocci, V., *The neglected organ: bacterial flora has a crucial immunostimulatory role*. *Perspect Biol Med*, 1992. **35**(2): p. 251-60.
257. Lynch, S.V. and O. Pedersen, *The human intestinal microbiome in health and disease*. *N Engl J Med*, 2016. **375**(24): p. 2369-2379.
258. Rajilić-Stojanović, M. and W.M. de Vos, *The first 1000 cultured species of the human gastrointestinal microbiota*. *FEMS Microbiology Rev*, 2014. **38**(5): p. 996-1047.
259. Lederberg, J., *Infectious History*. *Science*, 2000. **288**(5464): p. 287.
260. De Roy, K., et al., *Synthetic microbial ecosystems: an exciting tool to understand and apply microbial communities*. *Environmental Microbiol*, 2014. **16**(6): p. 1472-1481.
261. Lagier, J.-C., et al., *Culture of previously uncultured members of the human gut microbiota by culturomics*. *Nat Microbiol*, 2016. **1**: p. 16203.
262. Schaedler, R.W., R. Dubos, and R. Costello, *Association of germfree mice with bacteria isolated from normal mice*. *J Exp Med*, 1965. **122**(1): p. 77-82.
263. Hugenholtz, F. and W.M. de Vos, *Mouse models for human intestinal microbiota research: a critical evaluation*. *Cell Mol Life Sci*, 2018. **75**(1): p. 149-160.
264. Nguyen, T.L., et al., *How informative is the mouse for human gut microbiota research?* *Dis Model Mech*, 2015. **8**(1): p. 1-16.
265. Lundberg, R., et al., *Antibiotic-treated versus germ-free rodents for microbiota transplantation studies*. *Gut Microbes*, 2016. **7**(1): p. 68-74.
266. Science, A.A.f.L., *Development of gnotobiotics and contamination control in laboratory animal science, in 50 years of Laboratory Animal Science, AALAS*. 1999. p. 121-128.
267. Brand, M.W., et al., *The Altered Schaedler flora: continued applications of a defined murine microbial community*. *ILAR J*, 2015. **56**(2): p. 169-178.
268. Rath, H.C., et al., *Normal luminal bacteria, especially Bacteroides species, mediate chronic colitis, gastritis, and arthritis in HLA-B27/human beta2 microglobulin transgenic rats*. *J Clin Invest*, 1996. **98**(4): p. 945-953.

269. Cahill, R.J., et al., *Inflammatory bowel disease: an immunity-mediated condition triggered by bacterial infection with Helicobacter hepaticus*. Infect Immun, 1997. **65**(8): p. 3126-3131.
270. Natividad, J.M.M., et al., *Commensal and probiotic bacteria influence intestinal barrier function and susceptibility to colitis in Nod1<sup>-/-</sup>; Nod2<sup>-/-</sup> mice*. Inflamm Bowel Dis, 2012. **18**(8): p. 1434-46.
271. Wen, L., et al., *Innate immunity and intestinal microbiota in the development of Type 1 diabetes*. Nature, 2008. **455**(7216): p. 1109-1113.
272. Donohoe, D.R., et al., *A gnotobiotic mouse model demonstrates that dietary fiber protects against colorectal tumorigenesis in a microbiota- and butyrate-dependent manner*. Cancer Discov, 2014. **4**(12): p. 1387-1397.
273. Li, X., et al., *Response of germfree mice to colonization by Oxalobacter formigenes and altered Schaedler flora*. Appl Environ Microbiol, 2016. **82**(23): p. 6952-6960.
274. Hapfelmeier, S., et al., *Reversible microbial colonization of germ-free mice reveals the dynamics of IgA immune responses*. Science, 2010. **328**(5986): p. 1705-1709.
275. Brugiroux, S., et al., *Genome-guided design of a defined mouse microbiota that confers colonization resistance against Salmonella enterica serovar Typhimurium*. Nature Microbiol, 2016. **2**(2): p. 1-12.
276. Norin, E. and T. Midtvedt, *Intestinal microflora functions in laboratory mice claimed to harbor a "normal" intestinal microflora. Is the SPF concept running out of date?* Anaerobe, 2010. **16**(3): p. 311-313.
277. Lagkouvardos, I., et al., *The Mouse Intestinal Bacterial Collection (miBC) provides host-specific insight into cultured diversity and functional potential of the gut microbiota*. Nat Microbiol, 2016. **1**(10): p. 16131.
278. Studer, N., et al., *Functional intestinal bile acid 7 $\alpha$ -dehydroxylation by Clostridium scindens associated with protection from Clostridium difficile infection in a gnotobiotic mouse model*. Front Cell Infect Microbiol, 2016. **6**(December): p. 1-15.
279. Li, H., et al., *The outer mucus layer hosts a distinct intestinal microbial niche*. Nat Commun, 2015. **6**: p. 8292.
280. Becker, N., et al., *Human intestinal microbiota: Characterization of a simplified and stable gnotobiotic rat model*. Gut Microbes, 2011. **2**(1): p. 24-33.
281. Ganesh, B.P., et al., *Commensal Akkermansia muciniphila Exacerbates Gut Inflammation in Salmonella Typhimurium-Infected Gnotobiotic Mice*. PLoS ONE, 2013. **8**(9): p. 1-15.
282. Ring, C., et al., *Akkermansia muciniphila strain ATCC BAA-835 does not promote short-term intestinal inflammation in gnotobiotic interleukin-10-deficient mice*. Gut Microbes, 2018: p. 1-16.
283. Slezak, K., et al., *Increased bacterial putrescine has no impact on gut morphology and physiology in gnotobiotic adolescent mice*. Benef Microbes, 2013. **4**(3): p. 253-266.
284. Slezak, K., et al., *Association of germ-free mice with a simplified human intestinal microbiota results in a shortened intestine*. Gut Microbes, 2014. **5**(2): p. 176-182.



## APPENDICES

285. Woting, A., et al., *Clostridium ramosum* promotes High-Fat diet-induced obesity in Gnotobiotic Mouse Models. *mBio*, 2014. **5**(5): p. 1-10.
286. Syed, S.A., G.D. Abrams, and R. Freter, *Efficiency of various intestinal bacteria in assuming normal functions of enteric flora after association with germ-free mice*. *Infect Immun*, 1970. **2**(4): p. 376-386.
287. Freter, R. and G.D. Abrams, *Function of various intestinal bacteria in converting germfree mice to the normal state*. *Infect Immun*, 1972. **6**(2): p. 119-126.
288. Freter, R., R.R. Freter, and H. Brickner, *Experimental and mathematical models of Escherichia coli plasmid transfer in vitro and in vivo*. *Infect Immun*, 1983. **39**(1): p. 60-84.
289. Wells, C.L., H. Sugiyama, and S.E. Bland, *Resistance of mice with limited intestinal flora to enteric colonization by Clostridium botulinum*. *J Infect Dis*, 1982. **146**(6): p. 791-796.
290. Helstrom, P.B. and E. Balish, *Effect of oral tetracycline, the microbial flora, and the athymic state on gastrointestinal colonization and infection of BALB/c mice with Candida albicans*. *Infect Immun*, 1979. **23**(3): p. 764-774.
291. Moreau, M.C., et al., *Increase in the population of duodenal immunoglobulin A plasmocytes in axenic mice associated with different living or dead bacterial strains of intestinal origin*. *Infect Immun*, 1978. **21**(2): p. 532-539.
292. Rezzonico, E., et al., *Bacterial adaptation to the gut environment favors successful colonization: microbial and metabonomic characterization of a simplified microbiota mouse model*. *Gut microbes*, 2011. **2**(6): p. 307-318.
293. McNulty, N.P., et al., *The impact of a consortium of fermented milk strains on the gut microbiome of gnotobiotic mice and monozygotic twins*. *Sci Transl Med*, 2011. **3**(106): p. 1-26.
294. Faith, J.J., et al., *Predicting a Human Gut Microbiota's Response to Diet in Gnotobiotic Mice*. *Science*, 2011. **333**(6038): p. 101-104.
295. Rey, F.E., et al., *Metabolic niche of a prominent sulfate-reducing human gut bacterium*. *Proc Natl Acad Sci U S A*, 2013. **110**(33): p. 13582-13587.
296. Desai, M.S., et al., *A Dietary Fiber-Deprived Gut Microbiota Degrades the Colonic Mucus Barrier and Enhances Pathogen Susceptibility*. *Cell*, 2016. **167**(5): p. 1339-1353.e21.
297. Goodman, A.L., et al., *Identifying genetic determinants needed to establish a human gut symbiont in its habitat*. *Cell Host Microbe*, 2009. **6**(3): p. 279-289.
298. Reyes, A., et al., *Gnotobiotic mouse model of phage-bacterial host dynamics in the human gut*. *Proc Natl Acad Sci U S A*, 2013. **110**(50): p. 20236-20241.
299. McNulty, N.P., et al., *Effects of diet on resource utilization by a model human gut microbiota containing Bacteroides cellulosilyticus WH2, a symbiont with an extensive glycomiome*. *PLoS Biol*, 2013. **11**(8).
300. Wu, M., et al., *Genetic determinants of in vivo fitness and diet responsiveness in multiple human gut Bacteroides*. *Science*, 2015. **350**(6256): p. 1-21.
301. Mark Welch, J.L., et al., *Spatial organization of a model 15-member human gut microbiota established in gnotobiotic mice*. *Proc Natl Acad Sci U S A*, 2017: p. 201711596-201711596.
302. Sugahara, H., et al., *Probiotic Bifidobacterium longum alters gut luminal metabolism through modification of the gut microbial community*. *Sci Rep*, 2015. **5**(June): p. 1-11.

303. Lim, B., et al., *Engineered Regulatory Systems Modulate Gene Expression of Human Commensals in the Gut*. Cell, 2017. **169**(3): p. 547-558.e15.
304. Hibberd, M., *The effects of micronutrient deficiencies on bacterial species from the human gut microbiota*. Sci Transl Med, 2017. **22**(5): p. 733-744.
305. Laycock, G., et al., *A defined intestinal colonization microbiota for gnotobiotic pigs*. Vet Immunol Immunopathol, 2012. **149**(3-4): p. 216-224.
306. Nguyen, T.V., et al., *Transfer of maternal cytokines to suckling piglets: in vivo and in vitro models with implications for immunomodulation of neonatal immunity*. Vet Immunol Immunopathol, 2007. **117**(0): p. 236-248.
307. Butler, J.E. and M. Sinkora, *The isolator piglet: a model for studying the development of adaptive immunity*. Immunol Res, 2007. **39**(1-3): p. 33-51.
308. Rothkotter, H.J., H. Ulbrich, and R. Pabst, *The postnatal development of gut lamina propria lymphocytes: number, proliferation, and T and B cell subsets in conventional and germ-free pigs*. Pediatr Res, 1991. **29**(3): p. 237-42.
309. Wernersson, R., et al., *Pigs in sequence space: a 0.66X coverage pig genome survey based on shotgun sequencing*. BMC Genomics, 2005. **6**(1): p. 70.
310. Gonzalez, L.M., A.J. Moeser, and A.T. Blikslager, *Porcine models of digestive disease: the future of large animal translational research*. Transl Res, 2015. **166**(1): p. 12-27.
311. Jansman, A.J.M., et al., *Effects of a simple or a complex starter microbiota on intestinal microbiota composition in caesarean derived piglets*. J Anim Sci, 2012. **90 Suppl 4**: p. 433-5.
312. Priori, D., et al., *The olfactory receptor OR51E1 is present along the gastrointestinal tract of pigs, co-localizes with enteroendocrine cells and is modulated by intestinal microbiota*. PLoS One, 2015. **10**(6): p. e0129501.
313. Trevisi, P., et al., *The effects of starter microbiota and the early life feeding of medium chain triglycerides on the gastric transcriptome profile of 2- or 3-week-old cesarean delivered piglets*. J Anim Sci Biotechnol, 2017. **8**: p. 82.
314. Butler, J.E., et al., *Antibody repertoire development in fetal and newborn piglets, III. Colonization of the gastrointestinal tract selectively diversifies the preimmune repertoire in mucosal lymphoid tissues*. Immunology, 2000. **100**(1): p. 119-130.
315. Paim, F.C., et al., *Effects of Escherichia coli Nissle 1917 and Ciprofloxacin on small intestinal epithelial cell mRNA expression in the neonatal piglet model of human rotavirus infection*. Gut Pathog, 2016. **8**: p. 66.
316. Huang, H.C., et al., *Effect of antibiotic, probiotic, and human rotavirus infection on colonisation dynamics of defined commensal microbiota in a gnotobiotic pig model*. Benef Microbes, 2017: p. 1-16.
317. Eun, C.S., et al., *Induction of bacterial antigen-specific colitis by a simplified human microbiota consortium in gnotobiotic interleukin-10-/- mice*. Infect Immun, 2014. **82**(6): p. 2239-2246.
318. Kühn, R., et al., *Interleukin-10-deficient mice develop chronic enterocolitis*. Cell, 1993. **75**(2): p. 263-274.
319. Scheinin, T., et al., *Validation of the interleukin-10 knockout mouse model of colitis: antitumour necrosis factor-antibodies suppress the progression of colitis*. Clin Exp Immunol, 2003. **133**(1): p. 38-43.



## APPENDICES

320. Sellon, R.K., et al., *Resident enteric bacteria are necessary for development of spontaneous colitis and immune system activation in interleukin-10-deficient mice.* Infect Immun, 1998. **66**(11): p. 5224-5231.
321. Martin, F.P.J., et al., *A top-down systems biology view of microbiome-mammalian metabolic interactions in a mouse model.* Mol Syst Biol, 2007. **3**(112).
322. Martin, F.P.J., et al., *Probiotic modulation of symbiotic gut microbial-host metabolic interactions in a humanized microbiome mouse model.* Mol Syst Biol, 2008. **4**(157).
323. Martin, F.-P.J., et al., *Panorganismal gut microbiome - Host metabolic crosstalk.* J Proteome Res, 2009. **8**: p. 2090-2105.
324. Luk, B., et al., *Postnatal colonization with human "infant-type" Bifidobacterium species alters behavior of adult gnotobiotic mice.* PLoS One, 2018. **13**(5): p. e0196510.
325. Caballero-Franco, C., et al., *The VSL # 3 probiotic formula induces mucin gene expression and secretion in colonic epithelial cells.* Am J Physiol Gastrointest Liver Physiol, 2007. **292**: p. 315-322.
326. Kim, H.J., et al., *A randomized controlled trial of a probiotic combination VSL# 3 and placebo in irritable bowel syndrome with bloating.* Neurogastroenterol Motil, 2005. **17**(5): p. 687-696.
327. Douillard, F.P., et al., *Comparative genomic analysis of the multispecies probiotic-marketed product VSL#3.* PLoS One, 2018. **13**(2): p. e0192452.
328. Tvede, M. and J. Rask-Madsen, *Bacteriotherapy for chronic relapsing Clostridium difficile diarrhoea in six patients.* Lancet, 1989. **333**(8648): p. 1156-1160.
329. Lawley, T.D., et al., *Targeted restoration of the intestinal microbiota with a simple, defined bacteriotherapy resolves relapsing Clostridium difficile disease in mice.* PLoS Pathog, 2012. **8**(10): p. e1002995.
330. Petrof, E.O., et al., *Stool substitute transplant therapy for the eradication of Clostridium difficile infection: 'RePOOPulating' the gut.* Microbiome, 2013. **1**(1): p. 3-3.
331. Atarashi, K., et al., *Treg induction by a rationally selected mixture of Clostridia strains from the human microbiota.* Nature, 2013. **500**(7461): p. 232-236.
332. Hsiao, A., et al., *Members of the human gut microbiota involved in recovery from Vibrio cholerae infection.* Nature, 2015. **515**(7527): p. 423-426.
333. de Vos, W.M., *Fame and future of faecal transplantations--developing next-generation therapies with synthetic microbiomes.* Microb Biotechnol, 2013. **6**(4): p. 316-25.
334. Natividad, J.M., et al., *Ecobiotherapy rich in Firmicutes decreases susceptibility to colitis in a humanized gnotobiotic mouse model.* Inflamm Bowel Dis, 2015. **21**(8): p. 1883-1893.
335. Norin, E., *Experience with cultivated microbiota transplant: ongoing treatment of Clostridium difficile patients in Sweden.* Microb Ecol Health Dis, 2015. **26**: p. 27638-27638.
336. Faith, J.J., et al., *Identifying gut microbe-host phenotype relationships using combinatorial communities in gnotobiotic mice.* Sci Transl Med, 2014. **6**(220): p. 220ra11-220ra11.

337. Sonnenburg, J.L., C.T.L. Chen, and J.I. Gordon, *Genomic and metabolic studies of the impact of probiotics on a model gut symbiont and host*. PLoS Biol, 2006.
338. Mahowald, M.A., et al., *Characterizing a model human gut microbiota composed of members of its two dominant bacterial phyla*. Proc Natl Acad Sci U S A, 2009. **106**(14): p. 5859-5864.
339. Marcobal, A., et al., *A metabolomic view of how the human gut microbiota impacts the host metabolome using humanized and gnotobiotic mice*. ISME J, 2013. **7**(10): p. 1933-1943.
340. Krych, L., et al., *Quantitatively different, yet qualitatively alike: a meta-analysis of the mouse core gut microbiome with a view towards the human gut microbiome*. PLoS One, 2013. **8**(5): p. e62578.
341. Li, J., et al., *An integrated catalog of reference genes in the human gut microbiome*. Nat Biotechnol, 2014. **32**: p. 834.
342. Falony, G., et al., *Population-level analysis of gut microbiome variation*. Science, 2016. **352**(6285): p. 560-4.
343. Wilson, K.H., et al., *Comparison of fecal biota from specific pathogen free and feral mice*. Anaerobe, 2006. **12**(5-6): p. 249-253.
344. Ley, R.E., et al., *Obesity alters gut microbial ecology*. Proc Natl Acad Sci U S A, 2005. **102**(31): p. 11070-11075.
345. Chung, H., et al., *Gut immune maturation depends on colonization with a host-specific microbiota*. Cell, 2012. **149**(7): p. 1578-1593.
346. Turnbaugh, P.J., et al., *The effect of diet on the human gut microbiome: a metagenomic analysis in humanized gnotobiotic mice*. Sci Transl Med, 2009. **1**(6): p. 6ra14-6ra14.
347. Faith, J.J., et al., *The long-term stability of the human gut microbiota*. Science, 2013. **341**(6141): p. 1237439-1237439.
348. Sarma-Rupavtarm, R.B., et al., *Spatial distribution and stability of the eight microbial species of the altered Schaedler flora in the mouse gastrointestinal tract*. Appl Environ Microbiol, 2004. **70**(5): p. 2791-2800.
349. Freter, R., et al., *Mechanisms that control bacterial populations in continuous-flow culture models of mouse large intestinal flora*. Infect Immun, 1983. **39**(2): p. 676-685.
350. Stecher, B., et al., *Like will to like: abundances of closely related species can predict susceptibility to intestinal colonization by pathogenic and commensal bacteria*. PLoS Pathog, 2010. **6**(1).
351. Pereira, F.C. and D. Berry, *Microbial nutrient niches in the gut*. Environmental Microbiol, 2017. **19**(4): p. 1366-1378.
352. Clavel, T., et al., *Deciphering interactions between the gut microbiota and the immune system via microbial cultivation and minimal microbiomes*. Immunol Rev, 2017. **279**(1): p. 8-22.
353. Shetty, S.A., et al., *Intestinal microbiome landscaping: insight in community assemblage and implications for microbial modulation strategies*. FEMS Microbiology Rev, 2017. **41**(2): p. 182-199.
354. Jalanka-Tuovinen, J., et al., *Intestinal microbiota in healthy adults: temporal analysis reveals individual and common core and relation to intestinal symptoms*. PLoS One, 2011. **6**(7): p. e23035.



## APPENDICES

355. Salonen, A., et al., *The adult intestinal core microbiota is determined by analysis depth and health status*. Clin Microbiol Infect, 2012. **18 Suppl 4**: p. 16-20.
356. Ze, X., et al., *Some are more equal than others: the role of "keystone" species in the degradation of recalcitrant substrates*. Gut Microbes, 2013. **4**(3): p. 236-40.
357. Trosvik, P. and E.J. de Muinck, *Ecology of bacteria in the human gastrointestinal tract--identification of keystone and foundation taxa*. Microbiome, 2015. **3**: p. 44.
358. Qin, J., et al., *A human gut microbial gene catalog established by metagenomic sequencing*. Nature, 2010. **464**(7285): p. 59-65.
359. Kolmeder, C.A., et al., *Comparative metaproteomics and diversity analysis of human intestinal microbiota testifies for its temporal stability and expression of core functions*. PLoS One, 2012. **7**(1): p. e29913.
360. Nielsen, E. and C.W. Friis, *Influence of an intestinal microflora on the development of the immunoglobulins IgG1, IgG2a, IgM and IgA in germ-free BALB/c mice*. Acta Pathol Microbiol Scand C, 1980. **88**(3): p. 121-6.
361. Samuel, B.S. and J.I. Gordon, *A humanized gnotobiotic mouse model of host-archaeal-bacterial mutualism*. Proc Natl Acad Sci U S A, 2006. **103**(26): p. 10011-10016.
362. Manrique, P., et al., *Healthy human gut phageome*. Proc Natl Acad Sci U S A, 2016. **113**(37): p. 10400-10405.
363. Salonen, A. and W.M. de Vos, *Impact of diet on human intestinal microbiota and health*. Annu Rev Food Sci Technol, 2014. **5**(1): p. 239-262.
364. Zoetendal, E.G. and W.M. de Vos, *Effect of diet on the intestinal microbiota and its activity*. Curr Opin Gastroenterol, 2014. **30**(2): p. 189-95.
365. Clavel, T., et al., *Intestinal microbiota in metabolic diseases*. Gut Microbes, 2014. **5**(4): p. 544-551.
366. Esworthy, R.S., D.D. Smith, and F.-F. Chu, *A strong impact of genetic background on gut microflora in mice*. Int J Inflam, 2010. **2010**(2010): p. 986046.
367. Gulati, A.S., et al., *Mouse background strain profoundly influences paneth cell function and intestinal microbial composition*. PLoS One, 2012. **7**(2): p. e32403.
368. Hildebrand, F., et al., *Inflammation-associated enterotypes, host genotype, cage and inter-individual effects drive gut microbiota variation in common laboratory mice*. Genome Biol, 2013. **14**(1): p. R4.
369. Kovacs, A., et al., *Genotype is a stronger determinant than sex of the mouse gut microbiota*. Microb Ecol, 2011. **61**(2): p. 423-428.
370. Toivanen, P., J. Vahtovuori, and E. Eerola, *Influence of major histocompatibility complex on bacterial composition of fecal flora*. Infect Immun, 2001. **69**(4): p. 2372-2377.
371. Deloris Alexander, A., et al., *Quantitative PCR assays for mouse enteric flora reveal strain-dependent differences in composition that are influenced by the microenvironment*. Mamm Genome, 2006. **17**(11): p. 1093-104.
372. Human Microbiome Project Consortium, *Structure, function and diversity of the healthy human microbiome*. Nature, 2012. **486**(7402): p. 207-14.

373. Lay, C., et al., *Colonic microbiota signatures across five northern European countries*. *Appl Environ Microbiol*, 2005. **71**(7): p. 4153-5.
374. Li, M., et al., *Symbiotic gut microbes modulate human metabolic phenotypes*. *Proc Natl Acad Sci U S A*, 2008. **105**(6): p. 2117-22.
375. Mueller, S., et al., *Differences in fecal microbiota in different European study populations in relation to age, gender, and country: a cross-sectional study*. *Appl Environ Microbiol*, 2006. **72**(2): p. 1027-33.
376. Shastri, P., et al., *Sex differences in gut fermentation and immune parameters in rats fed an oligofructose-supplemented diet*. *Biol Sex Differ*, 2015. **6**: p. 13.
377. Markle, J.G.M., et al., *Sex differences in the gut microbiome drive hormone-dependent regulation of autoimmunity*. *Science*, 2013. **339**(6123): p. 1084-8.
378. Yurkovetskiy, L., et al., *Gender bias in autoimmunity is influenced by microbiota*. *Immunity*, 2013. **39**(2): p. 400-12.
379. Org, E., et al., *Sex differences and hormonal effects on gut microbiota composition in mice*. *Gut microbes*, 2016. **7**(4): p. 313-322.
380. Hansen, C.H.F., et al., *Patterns of early gut colonization shape future immune responses of the host*. *PLoS One*, 2012. **7**(3): p. e34043.
381. Olszak, T., et al., *Microbial exposure during early life has persistent effects on natural killer T cell function*. *Science*, 2012. **336**(6080): p. 489-93.
382. El Aidy, S., et al., *The gut microbiota and mucosal homeostasis: colonized at birth or at adulthood, does it matter?* *Gut microbes*, 2013. **4**(2): p. 118-24.
383. El Aidy, S., et al., *The microbiota and the gut-brain axis: insights from the temporal and spatial mucosal alterations during colonisation of the germfree mouse intestine*. *Benef Microbes*, 2012. **3**(4): p. 251-9.
384. Venema, K. and P. van den Abbeele, *Experimental models of the gut microbiome*. *Best Pract Res Clin Gastroenterol*, 2013. **27**(1): p. 115-126.
385. von Martels, J.Z.H., et al., *The role of gut microbiota in health and disease: in vitro modeling of host-microbe interactions at the aerobe-anaerobe interphase of the human gut*, in *Anaerobe*. 2017, Elsevier Ltd. p. 3-12.
386. Payne, A.N., et al., *Advances and perspectives in in vitro human gut fermentation modeling*. *Trends Biotechnol*, 2012. **30**(1): p. 17-25.
387. McDonald, J.A.K., et al., *Evaluation of microbial community reproducibility, stability and composition in a human distal gut chemostat model*. *J Microbiol Methods*, 2013. **95**(2): p. 167-174.
388. Van Den Abbeele, P., et al., *Microbial community development in a dynamic gut model is reproducible, colon region specific, and selective for Bacteroidetes and Clostridium cluster IX*. *Appl Environ Microbiol*, 2010. **76**(15): p. 5237-5246.
389. Marzorati, M., et al., *The HMI™ module: a new tool to study the Host-Microbiota Interaction in the human gastrointestinal tract in vitro*. *BMC Microbiol*, 2014. **14**(1): p. 133-133.
390. Krishnan, M., et al., *VSL#3 probiotic stimulates T-cell protein tyrosine phosphatase-mediated recovery of ifn- $\gamma$ -induced intestinal epithelial barrier defects*. *Inflamm Bowel Dis*, 2016. **22**(12): p. 2811-2823.
391. Martz, S.L., et al., *A human gut ecosystem protects against C. difficile disease by targeting TcdA*. *J Gastroenterol*, 2017. **52**(4): p. 452-465.





## APPENDICES

392. Munoz, S., et al., *Rebooting the microbiome*. Gut microbes, 2016. **7**(4): p. 353-363.
393. Cinque, B., et al., *VSL#3 probiotic differently influences IEC-6 intestinal epithelial cell status and function*. J Cell Physiol, 2017. **232**(12): p. 3530-3539.
394. Cinque, B., et al., *Production conditions affect the in vitro anti-tumoral effects of a high concentration multi-strain probiotic preparation*. PLoS ONE, 2016. **11**(9): p. 1-19.
395. Pagnini, C., et al., *Probiotics promote gut health through stimulation of epithelial innate immunity*. Proc Natl Acad Sci U S A, 2010. **107**(1): p. 454-459.
396. Trinchieri, V., et al., *Efficacy and safety of a multistrain probiotic formulation depends from manufacturing*. Front Immunol, 2017. **8**: p. 1474.
397. Mastrangeli, G., et al., *Effects of live and inactivated VSL#3 probiotic preparations in the modulation of in vitro and in vivo allergen-induced Th2 responses*. Int Arch Allergy Immunol, 2009. **150**(2): p. 133-143.
398. Mariman, R., et al., *The probiotic mixture VSL#3 dampens LPS-induced chemokine expression in human dendritic cells by inhibition of STAT-1 phosphorylation*. PLoS ONE, 2014. **9**(12): p. 1-13.
399. Parlesak, A., et al., *Modulation of cytokine release by differentiated Caco-2 cells in a compartmentalized coculture model with mononuclear leucocytes and nonpathogenic bacteria*. Scand J Immunol, 2004. **60**(5): p. 477-85.
400. Ulluwishewa, D., et al., *Live Faecalibacterium prausnitzii in an apical anaerobic model of the intestinal epithelial barrier*. Cell Microbiol, 2015. **17**(2): p. 226-240.
401. Sato, T., et al., *Single Lgr5 stem cells build crypt villus structures in vitro without a mesenchymal niche*. Nature, 2009. **459**: p. 262.
402. Spence, J.R., et al., *Directed differentiation of human pluripotent stem cells into intestinal tissue in vitro*. Nature, 2011. **470**(7332): p. 105-109.
403. Bartfeld, S., *Modeling infectious diseases and host-microbe interactions in gastrointestinal organoids*. Dev Biol, 2016. **420**(2): p. 262-270.
404. Leslie, J.L., et al., *Persistence and toxin production by Clostridium difficile within human intestinal organoids result in disruption of epithelial paracellular barrier function*. Infect Immun, 2015. **83**(1): p. 138-145.
405. Yissachar, N., et al., *An Intestinal Organ Culture System Uncovers a Role for the Nervous System in Microbe-Immune Crosstalk*. Cell, 2017. **168**(6): p. 1135-1148.e12.
406. Huh, D., et al., *Microengineered physiological biomimicry: organs-on-chips*. Lab Chip, 2012. **12**(12): p. 2156-64.
407. Walsh, D.I., 3rd, et al., *Emulation of Colonic Oxygen Gradients in a Microdevice*. SLAS Technol, 2018. **23**(2): p. 164-171.
408. Kasendra, M., et al., *Development of a primary human Small Intestine-on-a-Chip using biopsy-derived organoids*. Sci Rep, 2018. **8**(1): p. 2871.
409. Villenave, R., et al., *Human gut-on-a-chip supports polarized infection of coxsackie B1 virus in vitro*. PLoS One, 2017. **12**(2): p. e0169412.
410. Kim, H.J., et al., *Human gut-on-a-chip inhabited by microbial flora that experiences intestinal peristalsis-like motions and flow*. Lab Chip, 2012. **12**(12): p. 2165-74.

411. Kim, H.J. and D.E. Ingber, *Gut-on-a-Chip microenvironment induces human intestinal cells to undergo villus differentiation*. Integr Biol (Camb), 2013. **5**(9): p. 1130-40.
412. Kim, H.J., et al., *Contributions of microbiome and mechanical deformation to intestinal bacterial overgrowth and inflammation in a human gut-on-a-chip*. Proc Natl Acad Sci U S A, 2016. **113**(1): p. E7-E15.
413. Shah, P., et al., *A microfluidics-based in vitro model of the gastrointestinal human-microbe interface*. Nat Commun, 2016. **7**(May): p. 11535-11535.
414. Kasendra, M., et al., *Development of a primary human Small Intestine-on-a-Chip using biopsy-derived organoids*. Scientific Reports, 2018. **8**(1): p. 2871.
415. Marx, U., et al., *'Human-on-a-chip' developments: a translational cutting-edge alternative to systemic safety assessment and efficiency evaluation of substances in laboratory animals and man?* Altern Lab Anim, 2012. **40**(5): p. 235-57.
416. Peterson, J., et al., *The NIH Human Microbiome Project*. Genome Res, 2009. **19**(12): p. 2317-23.
417. van der Ark, K.C.H., et al., *More than just a gut feeling: constraint-based genome-scale metabolic models for predicting functions of human intestinal microbes*. Microbiome, 2017. **5**(1): p. 78-78.
418. Bucci, V., et al., *Towards predictive models of the human gut microbiome*. J Mol Biol, 2015. **19**(2): p. 161-169.
419. Macfarlane, G.T. and S. Macfarlane, *Models for intestinal fermentation: association between food components, delivery systems, bioavailability and functional interactions in the gut*. Curr Opin Biotechnol, 2007. **18**(2): p. 156-162.
420. Cani, P.D. and W.M. de Vos, *Next-generation beneficial microbes: the case of Akkermansia muciniphila*. Front Microbiol, 2017. **8**: p. 1765.
421. Fuentes, S. and W.M. de Vos, *How to manipulate the microbiota: fecal microbiota transplantation*. Adv Exp Med Biol, 2016. **902**: p. 143-53.
422. Doolittle, D.J., B.E. Butterworth, and J.M. Sherrill, *Influence of intestinal bacteria, sex of the animal, and position of the nitro group on the hepatic genotoxicity of nitrotoluene isomers in vivo*. Cancer Res, 1983. **43**(6): p. 2836-2842.
423. Jergens, A.E., et al., *Induction of differential immune reactivity to members of the flora of gnotobiotic mice following colonization with Helicobacter bilis or Brachyspira hyodysenteriae*. Microbes Infect, 2006. **8**(6): p. 1602-10.
424. Jergens, A.E., et al., *Helicobacter bilis triggers persistent immune reactivity to antigens derived from the commensal bacteria in gnotobiotic C3H/HeN mice*. Gut, 2007. **56**(7): p. 934-940.
425. Ivanov, I.I., et al., *Specific microbiota direct the differentiation of IL-17-producing T-helper cells in the mucosa of the small intestine*. Cell Host Microbe, 2008. **4**(4): p. 337-349.
426. Cong, Y., et al., *A dominant, coordinated T regulatory cell-IgA response to the intestinal microbiota*. Proc Natl Acad Sci U S A, 2009. **106**(46): p. 19256-19261.
427. Feng, T., et al., *Microbiota innate stimulation is a prerequisite for T cell spontaneous proliferation and induction of experimental colitis*. J Exp Med, 2010. **207**(6): p. 1321-1332.



## APPENDICES

428. Geuking, M.B., et al., *Intestinal Bacterial Colonization Induces Mutualistic Regulatory T Cell Responses*. *Immunity*, 2011. **34**(5): p. 794-806.
429. Natividad, J.M.M., et al., *Differential induction of antimicrobial REGIII by the intestinal microbiota and Bifidobacterium breve NCC2950*. *Appl Environ Microbiol*, 2013. **79**(24): p. 7745-7754.
430. Mosconi, I., et al., *Intestinal bacteria induce TSLP to promote mutualistic T-cell responses*. *Mucosal Immunol*, 2013. **6**(6): p. 1157-1167.
431. Collins, J., et al., *Intestinal microbiota influence the early postnatal development of the enteric nervous system*. *Neurogastroenterol Motil*, 2014. **26**(1): p. 98-107.
432. Peppercorn, M.a. and P. Goldman, *Caffeic acid metabolism by gnotobiotic rats and their intestinal bacteria*. *Proc Natl Acad Sci U S A*, 1972. **69**(6): p. 1413-1415.
433. Ducluzeau, R., et al., *Antagonistic effect of extremely oxygen-sensitive clostridia from the microflora of conventional mice and of Escherichia coli against Shigella flexneri in the digestive tract of gnotobiotic mice*. *Infect Immun*, 1977. **17**(2): p. 415-24.
434. Šinkorová, Z., et al., *Commensal intestinal bacterial strains trigger ankylosing enthesopathy of the ankle in inbred B10.BR (H-2k) male mice*. *Hum Immunol*, 2008. **69**(12): p. 845-850.
435. Wrzosek, L., et al., *Bacteroides thetaiotaomicron and Faecalibacterium prausnitzii influence the production of mucus glycans and the development of goblet cells in the colonic epithelium of a gnotobiotic model rodent*. *BMC Biol*, 2013. **11**: p. 61.
436. Silverthorn, D.U., et al., *Human Physiology: An Integrated Approach*. 2016: Pearson.
437. Sender, R., S. Fuchs, and R. Milo, *Revised Estimates for the Number of Human and Bacteria Cells in the Body*. *PLoS biology*, 2016. **14**(8): p. e1002533-e1002533.
438. Salvador, V., et al., *Sugar composition of dietary fibre and short-chain fatty acid production during in vitro fermentation by human bacteria*. *Br J Nutr*, 1993. **70**(1): p. 189-97.
439. Derrien, M., C. Belzer, and W.M. de Vos, *Akkermansia muciniphila and its role in regulating host functions*. *Microb Pathog*, 2017. **106**: p. 171-181.
440. Parada Venegas, D., et al., *Short Chain Fatty Acids (SCFAs)-Mediated Gut Epithelial and Immune Regulation and Its Relevance for Inflammatory Bowel Diseases*. *Front Immunol*, 2019. **10**: p. 277.
441. Lomer, M.C.E., R.P.H. Thompson, and J.J. Powell, *Fine and ultrafine particles of the diet: influence on the mucosal immune response and association with Crohn's disease*. *Proceedings of the Nutrition Society*, 2002. **61**(1): p. 123-130.
442. Chen, H., L. Meng, and L. Shen, *Multiple roles of short-chain fatty acids in Alzheimer disease*. 2022.
443. Sun, M.F., et al., *Neuroprotective effects of fecal microbiota transplantation on MPTP-induced Parkinson's disease mice: Gut microbiota, glial reaction and TLR4/TNF- $\alpha$  signaling pathway*. *Brain, Behavior, and Immunity*, 2018. **70**: p. 48-60.

444. Mark, M.v.d., et al., *Is Pesticide Use Related to Parkinson Disease? Some Clues to Heterogeneity in Study Results*. *Environmental Health Perspectives*, 2012. **120**(3): p. 340-347.
445. Abreu, M.T. and R.M. Peek, Jr., *Gastrointestinal malignancy and the microbiome*. *Gastroenterology*, 2014. **146**(6): p. 1534-1546 e3.
446. Dihal, A.A., et al., *Pathway and single gene analyses of inhibited Caco-2 differentiation by ascorbate-stabilized quercetin suggest enhancement of cellular processes associated with development of colon cancer*. *Mol Nutr Food Res*, 2007. **51**(8): p. 1031-45.
447. Allen, A.P., et al., *A psychology of the human brain-gut-microbiome axis*. *Soc Personal Psychol Compass*, 2017. **11**(4): p. e12309.
448. Hartstra, A.V., et al., *Insights Into the Role of the Microbiome in Obesity and Type 2 Diabetes*. *Diabetes Care*, 2014. **38**(1): p. 159-165.
449. Sarkar, A., et al., *The Microbiome in Psychology and Cognitive Neuroscience*. *Trends Cogn Sci*, 2018. **22**(7): p. 611-636.
450. Singer-Englar, T., G. Barlow, and R. Mathur, *Obesity, diabetes, and the gut microbiome: an updated review*. *Expert Rev Gastroenterol Hepatol*, 2019. **13**(1): p. 3-15.
451. Etienne-Mesmin, L., et al., *Experimental models to study intestinal microbes-mucus interactions in health and disease*. *FEMS Microbiol Rev*, 2019. **43**(5): p. 457-489.
452. Gustafsson, J.K., et al., *Bicarbonate and functional CFTR channel are required for proper mucin secretion and link cystic fibrosis with its mucus phenotype*. *J Exp Med*, 2012. **209**(7): p. 1263-72.
453. Yissachar, N., et al., *An Intestinal Organ Culture System Uncovers a Role for the Nervous System in Microbe-Immune Crosstalk*. *Cell*, 2017. **168**(6): p. 1135-1148 e12.
454. Mak, I.W., N. Evaniew, and M. Ghert, *Lost in translation: animal models and clinical trials in cancer treatment*. *Am J Transl Res*, 2014. **6**(2): p. 114-8.
455. Thompson, G.R. and P.C. Trexler, *Gastrointestinal structure and function in germ-free or gnotobiotic animals*. *Gut*, 1971. **12**(3): p. 230-5.
456. Kararli, T.T., *Comparison of the gastrointestinal anatomy, physiology, and biochemistry of humans and commonly used laboratory animals*. *Biopharm Drug Dispos*, 1995. **16**(5): p. 351-80.
457. Nguyen, T.L.A., et al., *How informative is the mouse for human gut microbiota research?* *Disease Models & Mechanisms*, 2015. **8**(1): p. 1-16.
458. Turnbaugh, P.J., et al., *The human microbiome project*. *Nature*, 2007. **449**(7164): p. 804-10.
459. Faith, J.J., et al., *Creating and characterizing communities of human gut microbes in gnotobiotic mice*. *ISME J*, 2010. **4**(9): p. 1094-8.
460. Ferdowsian, H.R. and N. Beck, *Ethical and Scientific Considerations Regarding Animal Testing and Research*. *PLOS ONE*, 2011. **6**(9): p. e24059.
461. Russell, W.M.S.B.R.L., *The principles of humane experimental technique*. 1959.
462. Rahman, S., et al., *The Progress of Intestinal Epithelial Models from Cell Lines to Gut-On-Chip*. *Int J Mol Sci*, 2021. **22**(24).



## APPENDICES

463. Wilson, G., et al., *Transport and permeability properties of human Caco-2 cells: An in vitro model of the intestinal epithelial cell barrier*. Journal of Controlled Release, 1990. **11**(1): p. 25-40.
464. Bouwmeester, H., et al., *Characterization of translocation of silver nanoparticles and effects on whole-genome gene expression using an in vitro intestinal epithelium coculture model*. ACS Nano, 2011. **5**(5): p. 4091-103.
465. Brand, W., et al., *Metabolism and transport of the citrus flavonoid hesperetin in Caco-2 cell monolayers*. Drug Metab Dispos, 2008. **36**(9): p. 1794-802.
466. Hidalgo, I.J., T.J. Raub, and R.T. Borchardt, *Characterization of the human colon carcinoma cell line (Caco-2) as a model system for intestinal epithelial permeability*. Gastroenterology, 1989. **96**(3): p. 736-49.
467. Hubatsch, I., E.G. Ragnarsson, and P. Artursson, *Determination of drug permeability and prediction of drug absorption in Caco-2 monolayers*. Nat Protoc, 2007. **2**(9): p. 2111-9.
468. Yamashita, S., et al., *Optimized conditions for prediction of intestinal drug permeability using Caco-2 cells*. European Journal of Pharmaceutical Sciences, 2000. **10**(3): p. 195-204.
469. Sambuy, Y., et al., *The Caco-2 cell line as a model of the intestinal barrier: influence of cell and culture-related factors on Caco-2 cell functional characteristics*. Cell Biology and Toxicology, 2005. **21**(1): p. 1-26.
470. Elzinga, J., et al., *Characterization of increased mucus production of HT29-MTX-E12 cells grown under Semi-Wet interface with Mechanical Stimulation*. PLoS One, 2021. **16**(12): p. e0261191.
471. Hilgendorf, C., et al., *Caco-2 versus Caco-2/HT29-MTX co-cultured cell lines: permeabilities via diffusion, inside- and outside-directed carrier-mediated transport*. J Pharm Sci, 2000. **89**(1): p. 63-75.
472. Lefebvre, D.E., et al., *Utility of models of the gastrointestinal tract for assessment of the digestion and absorption of engineered nanomaterials released from food matrices*. Nanotoxicology, 2015. **9**(4): p. 523-542.
473. Zweibaum, A., et al., *Use of Cultured Cell Lines in Studies of Intestinal Cell Differentiation and Function*, in Comprehensive Physiology. 2011. p. 223-255.
474. Delie, F. and W. Rubas, *A human colonic cell line sharing similarities with enterocytes as a model to examine oral absorption: advantages and limitations of the Caco-2 model*. Crit Rev Ther Drug Carrier Syst, 1997. **14**(3): p. 221-86.
475. Press, B. and D. Di Grandi, *Permeability for intestinal absorption: Caco-2 assay and related issues*. Curr Drug Metab, 2008. **9**(9): p. 893-900.
476. Shah, P., et al., *A microfluidics-based in vitro model of the gastrointestinal human-microbe interface*. Nat Commun, 2016. **7**: p. 11535.
477. Jalili-Firoozinezhad, S., et al., *A complex human gut microbiome cultured in an anaerobic intestine-on-a-chip*. Nat Biomed Eng, 2019. **3**(7): p. 520-531.
478. Kampfer, A.A.M., et al., *Development of an in vitro co-culture model to mimic the human intestine in healthy and diseased state*. Toxicol In Vitro, 2017. **45**(Pt 1): p. 31-43.
479. Kim, H.J., et al., *Contributions of microbiome and mechanical deformation to intestinal bacterial overgrowth and inflammation in a human gut-on-a-chip*. Proc Natl Acad Sci U S A, 2016. **113**(1): p. E7-15.

480. Huch, M., et al., *The hope and the hype of organoid research*. Development (Cambridge), 2017. **144**(6): p. 938-941.
481. Janssen, A.W.F., et al., *Cytochrome P450 expression, induction and activity in human induced pluripotent stem cell-derived intestinal organoids and comparison with primary human intestinal epithelial cells and Caco-2 cells*. Archives of Toxicology, 2020. **95**(3): p. 907-922.
482. McCracken, K.W., et al., *Generating human intestinal tissue from pluripotent stem cells in vitro*. Nature Protocols, 2011. **6**(12): p. 1920-1928.
483. Spence, J.R., et al., *Directed differentiation of human pluripotent stem cells into intestinal tissue in vitro*. Nature, 2010. **470**(7332): p. 105-109.
484. OECD, *Guidance Document on Good In Vitro Method Practices (GIVIMP)*. 2018.
485. Natoli, M., et al., *Good Caco-2 cell culture practices*. Toxicol In Vitro, 2012. **26**(8): p. 1243-6.
486. Angel, P.W., et al., *A simple, scalable approach to building a cross-platform transcriptome atlas*. PLoS Comput Biol, 2020. **16**(9): p. e1008219.
487. Irizarry, R.A., et al., *Exploration, normalization, and summaries of high density oligonucleotide array probe level data*. Biostatistics, 2003. **4**(2): p. 249-64.
488. Blighe, K. and A. Lun, *PCAtools: everything Principal Components Analysis*. 2020.
489. Thorndike, R.L., *Who belongs in the family?* Psychometrika, 1953. **18**(4): p. 267-276.
490. Hoffman, G.E. and E.E. Schadt, *variancePartition: interpreting drivers of variation in complex gene expression studies*. BMC Bioinformatics, 2016. **17**(1): p. 483.
491. Draghici, S., et al., *Global functional profiling of gene expression*. Genomics, 2003. **81**(2): p. 98-104.
492. Kanehisa, M., et al., *KEGG: new perspectives on genomes, pathways, diseases and drugs*. Nucleic Acids Res, 2017. **45**(D1): p. D353-D361.
493. Gene Ontology, C., *The Gene Ontology resource: enriching a GOld mine*. Nucleic Acids Res, 2021. **49**(D1): p. D325-D334.
494. Ashburner, M., et al., *Gene ontology: tool for the unification of biology*. The Gene Ontology Consortium. Nat Genet, 2000. **25**(1): p. 25-9.
495. Wu, T., et al., *clusterProfiler 4.0: A universal enrichment tool for interpreting omics data*. Innovation (Camb), 2021. **2**(3): p. 100141.
496. Balimane, P.V. and S. Chong, *Cell culture-based models for intestinal permeability: a critique*. Drug Discov Today, 2005. **10**(5): p. 335-43.
497. Putaala, H., et al., *Analysis of the human intestinal epithelial cell transcriptional response to Lactobacillus acidophilus, Lactobacillus salivarius, Bifidobacterium lactis and Escherichia coli*. Benef Microbes, 2010. **1**(3): p. 283-95.
498. Ishimoto, Y., et al., *Transient up-regulation of immunity- and apoptosis-related genes in Caco-2 cells cocultured with THP-1 cells evaluated by DNA microarray analysis*. Biosci Biotechnol Biochem, 2010. **74**(2): p. 437-9.
499. Turrioni, F., et al., *Genome analysis of Bifidobacterium bifidum PRL2010 reveals metabolic pathways for host-derived glycan foraging*. Proc Natl Acad Sci U S A, 2010. **107**(45): p. 19514-9.



## APPENDICES

500. Christensen, J., et al., *Defining new criteria for selection of cell-based intestinal models using publicly available databases*. BMC Genomics, 2012. **13**(1): p. 274.
501. Rossi, O., et al., *Vectorial secretion of interleukin-8 mediates autocrine signalling in intestinal epithelial cells via apically located CXCR1*. BMC Research Notes, 2013. **6**(1): p. 431.
502. Sakharov, D., et al., *Towards embedding Caco-2 model of gut interface in a microfluidic device to enable multi-organ models for systems biology*. BMC Syst Biol, 2019. **13**(Suppl 1): p. 19.
503. Lépine, A.F.P., et al., *Lactobacillus acidophilus Attenuates Salmonella-Induced Stress of Epithelial Cells by Modulating Tight-Junction Genes and Cytokine Responses*. Frontiers in Microbiology, 2018. **9**.
504. Kulthong, K., et al., *Transcriptome comparisons of in vitro intestinal epithelia grown under static and microfluidic gut-on-chip conditions with in vivo human epithelia*. Sci Rep, 2021. **11**(1): p. 3234.
505. Kulthong, K., et al., *Comparative study of the transcriptomes of Caco-2 cells cultured under dynamic vs. static conditions following exposure to titanium dioxide and zinc oxide nanomaterials*. Nanotoxicology, 2021. **15**(9): p. 1233-1252.
506. Akoglu, H., *User's guide to correlation coefficients*. Turk J Emerg Med, 2018. **18**(3): p. 91-93.
507. Larregieu, C.A. and L.Z. Benet, *Drug Discovery and Regulatory Considerations for Improving In Silico and In Vitro Predictions that Use Caco-2 as a Surrogate for Human Intestinal Permeability Measurements*. The AAPS Journal, 2013. **15**(2): p. 483-497.
508. Zucco, F., et al., *An inter-laboratory study to evaluate the effects of medium composition on the differentiation and barrier function of Caco-2 cell lines*. Altern Lab Anim, 2005. **33**(6): p. 603-18.
509. Briske-Anderson, M.J., J.W. Finley, and S.M. Newman, *The Influence of Culture Time and Passage Number on the Morphological and Physiological Development of Caco-2 Cells*. Proceedings of the Society for Experimental Biology and Medicine, 1997. **214**(3): p. 248-257.
510. Navabi, N., M.A. McGuckin, and S.K. Lindén, *Gastrointestinal cell lines form polarized epithelia with an adherent mucus layer when cultured in semi-wet interfaces with mechanical stimulation*. PLoS One, 2013. **8**(7): p. e68761.
511. Ma, C., et al., *Organ-on-a-Chip: A New Paradigm for Drug Development*. Trends in Pharmacological Sciences, 2021. **42**(2): p. 119-133.
512. hDMT, *Translational Organ-on-Chip Platform - TOP: Translational Organ-on-Chip Platform*. 2021.
513. Vachon, P.H. and J.-F. Beaulieu, *Transient mosaic patterns of morphological and functional differentiation in the Caco-2 cell line*. Gastroenterology, 1992. **103**(2): p. 414-423.
514. Kim, S.H., et al., *Three-dimensional intestinal villi epithelium enhances protection of human intestinal cells from bacterial infection by inducing mucin expression*. Integrative Biology, 2014. **6**(12): p. 1122-1131.

515. Natoli, M., et al., *Cell growing density affects the structural and functional properties of Caco-2 differentiated monolayer*. J Cell Physiol, 2011. **226**(6): p. 1531-43.
516. ATCC, Caco-2 [Caco2] | ATCC. 2021.
517. Perdigon, G., et al., *Immune system stimulation by probiotics*. J Dairy Sci, 1995. **78**(7): p. 1597-606.
518. Isolauri, E., *Probiotics in human disease*. Am J Clin Nutr, 2001. **73**(6): p. 1142S-1146S.
519. Brazma, A., et al., *Minimum information about a microarray experiment (MIAME)-toward standards for microarray data*. Nat Genet, 2001. **29**(4): p. 365-71.
520. Brazma, A., et al., *MINSEQE: Minimum Information about a high-throughput Nucleotide SeQuencing Experiment - a proposal for standards in functional genomic data reporting*. 2012.
521. Greenhalgh, K., et al., *Integrated In Vitro and In Silico Modeling Delineates the Molecular Effects of a Synbiotic Regimen on Colorectal-Cancer-Derived Cells*. Cell Reports, 2019. **27**(5): p. 1621-1632.e9.
522. Forero, D.A., W.H. Curioso, and G.P. Patrinos, *The importance of adherence to international standards for depositing open data in public repositories*. BMC Research Notes, 2021. **14**(1): p. 405.
523. Ioannidis, J.P.A., et al., *Repeatability of published microarray gene expression analyses*. Nature Genetics, 2009. **41**(2): p. 149-155.
524. Hartung, T., et al., *Toward Good In Vitro Reporting Standards*. ALTEX - Alternatives to animal experimentation, 2019. **36**(1): p. 3-17.
525. Wilkinson, M.D., et al., *The FAIR Guiding Principles for scientific data management and stewardship*. Scientific Data, 2016. **3**(1): p. 160018.
526. de Vries, R. and P. Whaley, *In Vitro Critical Appraisal Tool (IV-CAT): Tool Development Protocol*. 2018.
527. Emmerich, C.H. and C.M. Harris, *Minimum Information and Quality Standards for Conducting, Reporting, and Organizing In Vitro Research*. 2020, Springer. p. 177-196.
528. Vollertsen, A.R., et al., *Facilitating implementation of organs-on-chips by open platform technology*. Biomicrofluidics, 2021. **15**(5): p. 051301.
529. Allen, A., *Mucus—a protective secretion of complexity* Trends Biochem Sci 1983. **8**: p. 169-73
530. Forstner, J. and G. Forstner, *Gastrointestinal mucus in Physiology of the Gastrointestinal Tract*, L.G. Johnson, Editor. 1994, Raven Press New York. p. 1255-83.
531. Deplancke, B. and H.R. Gaskins, *Microbial modulation of innate defense: goblet cells and the intestinal mucus layer*. Am J Clin Nutr, 2001. **73**(6): p. 1131s-1141s.
532. Allen, A. and G. Flemström, *Gastroduodenal mucus bicarbonate barrier: protection against acid and pepsin*. Am J Physiol Cell Physiol, 2005. **288**(1): p. C1-19.
533. Van der Sluis, M., et al., *Muc2-deficient mice spontaneously develop colitis, indicating that MUC2 is critical for colonic protection*. Gastroenterology, 2006. **131**(1): p. 117-29.





## APPENDICES

534. Velcich, A., et al., *Colorectal cancer in mice genetically deficient in the mucin Muc2*. Science, 2002. **295**(5560): p. 1726-9.
535. Atuma, C., et al., *The adherent gastrointestinal mucus gel layer: thickness and physical state in vivo*. Am J Physiol Gastrointest Liver Physiol, 2001. **280**(5): p. G922-9.
536. Simon, G.L. and S.L. Gorbach, *Intestinal flora in health and disease*. Gastroenterology, 1984. **86**(1): p. 174-93.
537. Derrien, M., et al., *Mucin-bacterial interactions in the human oral cavity and digestive tract*. Gut Microbes, 2010. **1**(4): p. 254-268.
538. Bunesova, V., C. Lacroix, and C. Schwab, *Mucin Cross-Feeding of Infant Bifidobacteria and Eubacterium hallii*. Microb Ecol, 2018. **75**(1): p. 228-238.
539. Van Herreweghen, F., et al., *Mucin degradation niche as a driver of microbiome composition and Akkermansia muciniphila abundance in a dynamic gut model is donor independent*. FEMS Microbiol Ecol, 2018. **94**(12).
540. Jany, B.H., et al., *Human bronchus and intestine express the same mucin gene*. J Clin Invest, 1991. **87**(1): p. 77-82.
541. Tytgat, K.M., et al., *Biosynthesis of human colonic mucin: Muc2 is the prominent secretory mucin*. Gastroenterology, 1994. **107**(5): p. 1352-63.
542. Strous, G.J. and J. Dekker, *Mucin-type glycoproteins*. Crit Rev Biochem Mol Biol, 1992. **27**(1-2): p. 57-92.
543. Chang, S.K., et al., *Localization of mucin (MUC2 and MUC3) messenger RNA and peptide expression in human normal intestine and colon cancer*. Gastroenterology, 1994. **107**(1): p. 28-36.
544. Van Klinken, B.J., et al., *Mucin gene structure and expression: protection vs. adhesion*. Am J Physiol, 1995. **269**(5 Pt 1): p. G613-27.
545. Allen, A., *The structure and function of gastrointestinal mucus*, in Physiology of the Gastrointestinal Tract, L.R. Johnson, Editor. 1981, Raven Press: New York. p. 617-39.
546. Browning, T.H. and J.S. Trier, *Organ culture of mucosal biopsies of human small intestine*. J Clin Invest, 1969. **48**(8): p. 1423-32.
547. Gustafsson, J.K., et al., *Bicarbonate and functional CFTR channel are required for proper mucin secretion and link cystic fibrosis with its mucus phenotype*. J Exp Med, 2012. **209**(7): p. 1263-72.
548. van Klinken, B.J., et al., *The human intestinal cell lines Caco-2 and LS174T as models to study cell-type specific mucin expression*. Glycoconj J, 1996. **13**(5): p. 757-68.
549. Behrens, I., et al., *Transport of lipophilic drug molecules in a new mucus-secreting cell culture model based on HT29-MTX cells*. Pharm Res, 2001. **18**(8): p. 1138-45.
550. Costello, C.M., et al., *Synthetic small intestinal scaffolds for improved studies of intestinal differentiation*. Biotechnol Bioeng, 2014. **111**(6): p. 1222-32.
551. Chen, Y., et al., *Robust bioengineered 3D functional human intestinal epithelium*. Sci Rep, 2015. **5**: p. 13708.
552. Reuter, C., et al., *An adherent mucus layer attenuates the genotoxic effect of colibactin*. Cell Microbiol, 2018. **20**(2).
553. Jung, P., et al., *Isolation and in vitro expansion of human colonic stem cells*. Nat Med, 2011. **17**(10): p. 1225-7.

554. Sato, T., et al., *Long-term expansion of epithelial organoids from human colon, adenoma, adenocarcinoma, and Barrett's epithelium*. *Gastroenterology*, 2011. **141**(5): p. 1762-72.
555. Sontheimer-Phelps, A., et al., *Human Colon-on-a-Chip Enables Continuous In Vitro Analysis of Colon Mucus Layer Accumulation and Physiology*. *Cell Mol Gastroenterol Hepatol*, 2020. **9**(3): p. 507-526.
556. Belley, A. and K. Chadee, *Prostaglandin E(2) stimulates rat and human colonic mucin exocytosis via the EP(4) receptor*. *Gastroenterology*, 1999. **117**(6): p. 1352-62.
557. Milano, J., et al., *Modulation of notch processing by gamma-secretase inhibitors causes intestinal goblet cell metaplasia and induction of genes known to specify gut secretory lineage differentiation*. *Toxicol Sci*, 2004. **82**(1): p. 341-58.
558. Hatayama, H., et al., *The short chain fatty acid, butyrate, stimulates MUC2 mucin production in the human colon cancer cell line, LS174T*. *Biochem Biophys Res Commun*, 2007. **356**(3): p. 599-603.
559. Smirnova, M.G., et al., *LPS up-regulates mucin and cytokine mRNA expression and stimulates mucin and cytokine secretion in goblet cells*. *Cell Immunol*, 2003. **221**(1): p. 42-9.
560. Nossol, C., et al., *Air-liquid interface cultures enhance the oxygen supply and trigger the structural and functional differentiation of intestinal porcine epithelial cells (IPEC)*. *Histochem Cell Biol*, 2011. **136**(1): p. 103-15.
561. Klasvogt, S., et al., *Air-liquid interface enhances oxidative phosphorylation in intestinal epithelial cell line IPEC-J2*. *Cell Death Discov*, 2017. **3**: p. 17001.
562. Dickman, K.G. and L.J. Mandel, *Glycolytic and oxidative metabolism in primary renal proximal tubule cultures*. *Am J Physiol*, 1989. **257**(2 Pt 1): p. C333-40.
563. Johnson, L.G., et al., *Enhanced Na<sup>+</sup> transport in an air-liquid interface culture system*. *Am J Physiol*, 1993. **264**(6 Pt 1): p. L560-5.
564. Kondo, M., et al., *Increased oxidative metabolism in cow tracheal epithelial cells cultured at air-liquid interface*. *Am J Respir Cell Mol Biol*, 1997. **16**(1): p. 62-8.
565. Ootani, A., et al., *An air-liquid interface promotes the differentiation of gastric surface mucous cells (GSM06) in culture*. *Biochem Biophys Res Commun*, 2000. **271**(3): p. 741-6.
566. Irizarry, R.A., et al., *Summaries of Affymetrix GeneChip probe level data*. *Nucleic Acids Res*, 2003. **31**(4): p. e15.
567. Dai, M., et al., *Evolving gene/transcript definitions significantly alter the interpretation of GeneChip data*. *Nucleic Acids Res*, 2005. **33**(20): p. e175.
568. Sartor, M.A., et al., *Intensity-based hierarchical Bayes method improves testing for differentially expressed genes in microarray experiments*. *BMC Bioinformatics*, 2006. **7**: p. 538.
569. Oliveros, J.C., Venny. *An interactive tool for comparing lists with Venn's diagrams*. 2007-2015.
570. Demitrack, E.S. and L.C. Samuelson, *Notch regulation of gastrointestinal stem cells*. *J Physiol*, 2016. **594**(17): p. 4791-803.
571. VanDussen, K.L., et al., *Notch signaling modulates proliferation and differentiation of intestinal crypt base columnar stem cells*. *Development*, 2012. **139**(3): p. 488-97.



## APPENDICES

572. Kazanjian, A. and N.F. Shroyer, *NOTCH Signaling and ATOH1 in Colorectal Cancers*. *Curr Colorectal Cancer Rep*, 2011. **7**(2): p. 121-127.
573. Chen, K.Y., et al., *A Notch positive feedback in the intestinal stem cell niche is essential for stem cell self-renewal*. *Mol Syst Biol*, 2017. **13**(4): p. 927.
574. Ghaleb, A.M., et al., *Altered intestinal epithelial homeostasis in mice with intestine-specific deletion of the Krüppel-like factor 4 gene*. *Dev Biol*, 2011. **349**(2): p. 310-20.
575. Rowland, B.D. and D.S. Peeper, *KLF4, p21 and context-dependent opposing forces in cancer*. *Nat Rev Cancer*, 2006. **6**(1): p. 11-23.
576. Ghaleb, A.M. and V.W. Yang, *Krüppel-like factor 4 (KLF4): What we currently know*. *Gene*, 2017. **611**: p. 27-37.
577. Yoon, H.S. and V.W. Yang, *Requirement of Krüppel-like factor 4 in preventing entry into mitosis following DNA damage*. *J Biol Chem*, 2004. **279**(6): p. 5035-41.
578. Hagos, E.G., et al., *Mouse embryonic fibroblasts null for the Krüppel-like factor 4 gene are genetically unstable*. *Oncogene*, 2009. **28**(9): p. 1197-205.
579. Yoon, H.S., et al., *Krüppel-like factor 4 prevents centrosome amplification following gamma-irradiation-induced DNA damage*. *Oncogene*, 2005. **24**(25): p. 4017-25.
580. Chen, Z.Y., J.L. Shie, and C.C. Tseng, *Gut-enriched Kruppel-like factor represses ornithine decarboxylase gene expression and functions as checkpoint regulator in colonic cancer cells*. *J Biol Chem*, 2002. **277**(48): p. 46831-9.
581. Shie, J.L., et al., *Gut-enriched Krüppel-like factor represses cyclin D1 promoter activity through Sp1 motif*. *Nucleic Acids Res*, 2000. **28**(15): p. 2969-76.
582. Shie, J.L., et al., *Role of gut-enriched Krüppel-like factor in colonic cell growth and differentiation*. *Am J Physiol Gastrointest Liver Physiol*, 2000. **279**(4): p. G806-14.
583. Klaewsongkram, J., et al., *Krüppel-like factor 4 regulates B cell number and activation-induced B cell proliferation*. *J Immunol*, 2007. **179**(7): p. 4679-84.
584. Franzén, O., L.M. Gan, and J.L.M. Björkegren, *PanglaoDB: a web server for exploration of mouse and human single-cell RNA sequencing data*. *Database (Oxford)*, 2019. **2019**.
585. Aihara, E., K.A. Engevik, and M.H. Montrose, *Trefoil Factor Peptides and Gastrointestinal Function*. *Annu Rev Physiol*, 2017. **79**: p. 357-380.
586. Taupin, D., et al., *The trefoil gene family are coordinately expressed immediate-early genes: EGF receptor- and MAP kinase-dependent interregulation*. *J Clin Invest*, 1999. **103**(9): p. R31-8.
587. Garcia, M.A., N. Yang, and P.M. Quinton, *Normal mouse intestinal mucus release requires cystic fibrosis transmembrane regulator-dependent bicarbonate secretion*. *J Clin Invest*, 2009. **119**(9): p. 2613-22.
588. Kerem, B., et al., *Identification of the cystic fibrosis gene: genetic analysis*. *Science*, 1989. **245**(4922): p. 1073-80.
589. Moseley, R.H., et al., *Downregulated in adenoma gene encodes a chloride transporter defective in congenital chloride diarrhea*. *Am J Physiol*, 1999. **276**(1): p. G185-92.
590. Gawenis, L.R., et al., *Intestinal NaCl transport in NHE2 and NHE3 knockout mice*. *Am J Physiol Gastrointest Liver Physiol*, 2002. **282**(5): p. G776-84.

591. Schultheis, P.J., et al., *Renal and intestinal absorptive defects in mice lacking the NHE3 Na<sup>+</sup>/H<sup>+</sup> exchanger*. *Nat Genet*, 1998. **19**(3): p. 282-5.
592. Wakabayashi, S., M. Shigekawa, and J. Pouyssegur, *Molecular physiology of vertebrate Na<sup>+</sup>/H<sup>+</sup> exchangers*. *Physiol Rev*, 1997. **77**(1): p. 51-74.
593. Ootani, A., et al., *Sustained in vitro intestinal epithelial culture within a Wnt-dependent stem cell niche*. *Nat Med*, 2009. **15**(6): p. 701-6.
594. Christensen, J., et al., *Defining new criteria for selection of cell-based intestinal models using publicly available databases*. *BMC Genomics*, 2012. **13**: p. 274.
595. Srinivasan, B., et al., *TEER measurement techniques for in vitro barrier model systems*. *J Lab Autom*, 2015. **20**(2): p. 107-26.
596. Hilgendorf, C., et al., *Caco-2 versus Caco-2/HT29-MTX co-cultured cell lines: permeabilities via diffusion, inside- and outside-directed carrier-mediated transport*. *J Pharm Sci*, 2000. **89**(1): p. 63-75.
597. Shields, J.M., R.J. Christy, and V.W. Yang, *Identification and characterization of a gene encoding a gut-enriched Krüppel-like factor expressed during growth arrest*. *J Biol Chem*, 1996. **271**(33): p. 20009-17.
598. Chen, X., et al., *Krüppel-like factor 4 (gut-enriched Krüppel-like factor) inhibits cell proliferation by blocking G1/S progression of the cell cycle*. *J Biol Chem*, 2001. **276**(32): p. 30423-8.
599. Portier, F., et al., *Enhanced sodium absorption in middle ear epithelial cells cultured at air-liquid interface*. *Acta Otolaryngol*, 2005. **125**(1): p. 16-22.
600. Robison, T.W. and K.J. Kim, *Air-interface cultures of guinea pig airway epithelial cells: effects of active sodium and chloride transport inhibitors on bioelectric properties*. *Exp Lung Res*, 1994. **20**(2): p. 101-17.
601. Yamaya, M., et al., *Differentiated structure and function of cultures from human tracheal epithelium*. *Am J Physiol*, 1992. **262**(6 Pt 1): p. L713-24.
602. Kini, A., et al., *Slc26a3 deletion alters pH-microclimate, mucin biosynthesis, microbiome composition and increases the TNF $\alpha$  expression in murine colon*. *Acta Physiol (Oxf)*, 2020. **230**(2): p. e13498.
603. Whitcutt, M.J., K.B. Adler, and R. Wu, *A biphasic chamber system for maintaining polarity of differentiation of cultured respiratory tract epithelial cells*. *In Vitro Cell Dev Biol*, 1988. **24**(5): p. 420-8.
604. Alzheimer, M., et al., *A three-dimensional intestinal tissue model reveals factors and small regulatory RNAs important for colonization with Campylobacter jejuni*. *PLoS Pathog*, 2020. **16**(2): p. e1008304.
605. Delon, L.C., et al., *A systematic investigation of the effect of the fluid shear stress on Caco-2 cells towards the optimization of epithelial organ-on-chip models*. *Biomaterials*, 2019. **225**: p. 119521.
606. Lindner, M., et al., *Physiological shear stress enhances differentiation and mucus-formation of intestinal epithelial cells in vitro*. *Authorea*, 2020.
607. Evans, P.M. and C. Liu, *Roles of Krüpel-like factor 4 in normal homeostasis, cancer and stem cells*. *Acta Biochim Biophys Sin (Shanghai)*, 2008. **40**(7): p. 554-64.
608. Niu, N., et al., *Targeting Mechanosensitive Transcription Factors in Atherosclerosis*. *Trends Pharmacol Sci*, 2019. **40**(4): p. 253-266.



## APPENDICES

609. Zhang, B., et al., *ERK5 negatively regulates Kruppel-like factor 4 and promotes osteogenic lineage cell proliferation in response to MEK5 overexpression or fluid shear stress*. *Connect Tissue Res*, 2021. **62**(2): p. 194-205.
610. Clark, P.R., et al., *MEK5 is activated by shear stress, activates ERK5 and induces KLF4 to modulate TNF responses in human dermal microvascular endothelial cells*. *Microcirculation*, 2011. **18**(2): p. 102-17.
611. van Putten, J.P.M. and K. Strijbis, *Transmembrane Mucins: Signaling Receptors at the Intersection of Inflammation and Cancer*. *J Innate Immun*, 2017. **9**(3): p. 281-299.
612. Jonckheere, N. and I. Van Seuningen, *The membrane-bound mucins: From cell signalling to transcriptional regulation and expression in epithelial cancers*. *Biochimie*, 2010. **92**(1): p. 1-11.
613. Lock, J.Y., T.L. Carlson, and R.L. Carrier, *Mucus models to evaluate the diffusion of drugs and particles*. *Adv Drug Deliv Rev*, 2018. **124**: p. 34-49.
614. Sonnenburg, J.L., et al., *Glycan foraging in vivo by an intestine-adapted bacterial symbiont*. *Science*, 2005. **307**(5717): p. 1955-9.
615. Derrien, M., et al., *Mucin-bacterial interactions in the human oral cavity and digestive tract*. *Gut microbes*, 2010. **1**(4): p. 254-268.
616. Druart, C., et al., *Toxicological safety evaluation of pasteurized Akkermansia muciniphila*. *J Appl Toxicol*, 2021. **41**(2): p. 276-290.
617. Turck, D., et al., *Safety of pasteurised Akkermansia muciniphila as a novel food pursuant to Regulation (EU) 2015/2283*. *Efsa j*, 2021. **19**(9): p. e06780.
618. Everard, A., et al., *Cross-talk between Akkermansia muciniphila and intestinal epithelium controls diet-induced obesity*. *Proc Natl Acad Sci U S A*, 2013. **110**(22): p. 9066-71.
619. van der Lugt, B., et al., *Akkermansia muciniphila ameliorates the age-related decline in colonic mucus thickness and attenuates immune activation in accelerated aging Ercc1 (-/Δ7) mice*. *Immun Ageing*, 2019. **16**: p. 6.
620. Wrzosek, L., et al., *Bacteroides thetaiotaomicron and Faecalibacterium prausnitzii influence the production of mucus glycans and the development of goblet cells in the colonic epithelium of a gnotobiotic model rodent*. *BMC biology*, 2013. **11**(1): p. 61-61.
621. Lesuffleur, T., et al., *Differential expression of the human mucin genes MUC1 to MUC5 in relation to growth and differentiation of different mucus-secreting HT-29 cell subpopulations*. *Journal of cell science*, 1993. **106**(3): p. 771-783.
622. Wang, L., S. Wang, and W. Li, *RSeQC: quality control of RNA-seq experiments*. *Bioinformatics*, 2012. **28**(16): p. 2184-2185.
623. Frankish, A., et al., *GENCODE reference annotation for the human and mouse genomes*. *Nucleic Acids Res*, 2019. **47**(D1): p. D766-d773.
624. Patro, R., et al., *Salmon provides fast and bias-aware quantification of transcript expression*. *Nat Methods*, 2017. **14**(4): p. 417-419.
625. Soneson, C., M.I. Love, and M.D. Robinson, *Differential analyses for RNA-seq: transcript-level estimates improve gene-level inferences*. *F1000Res*, 2015. **4**: p. 1521.
626. Ritchie, M.E., et al., *limma powers differential expression analyses for RNA-sequencing and microarray studies*. *Nucleic Acids Res*, 2015. **43**(7): p. e47.

627. Bourgon, R., R. Gentleman, and W. Huber, *Independent filtering increases detection power for high-throughput experiments*. Proc Natl Acad Sci U S A, 2010. **107**(21): p. 9546-51.
628. Robinson, M.D. and A. Oshlack, *A scaling normalization method for differential expression analysis of RNA-seq data*. Genome Biol, 2010. **11**(3): p. R25.
629. Robinson, M.D., D.J. McCarthy, and G.K. Smyth, *edgeR: a Bioconductor package for differential expression analysis of digital gene expression data*. Bioinformatics, 2010. **26**(1): p. 139-40.
630. Law, C.W., et al., *voom: Precision weights unlock linear model analysis tools for RNA-seq read counts*. Genome Biol, 2014. **15**(2): p. R29.
631. Smyth, G.K., *Linear models and empirical bayes methods for assessing differential expression in microarray experiments*. Stat Appl Genet Mol Biol, 2004. **3**: p. Article3.
632. Li, X., et al., *Genes expressed at low levels raise false discovery rates in RNA samples contaminated with genomic DNA*. BMC Genomics, 2022. **23**(1): p. 554.
633. Song, W.S., et al., *Development of an in vitro coculture device for the investigation of host-microbe interactions via integrative multiomics approaches*. Biotechnol Bioeng, 2021. **118**(4): p. 1612-1623.
634. Madunic, K., et al., *Colorectal cancer cell lines show striking diversity of their O-glycome reflecting the cellular differentiation phenotype*. Cell Mol Life Sci, 2021. **78**(1): p. 337-350.
635. Reuter, C., et al., *An adherent mucus layer attenuates the genotoxic effect of colibactin*. Cellular microbiology, 2017. **20**(2): p. e12812-NA.
636. Vriend, J., et al., *Flow stimulates drug transport in a human kidney proximal tubule-on-a-chip independent of primary cilia*. Biochim Biophys Acta Gen Subj, 2020. **1864**(1): p. 129433.
637. Bossink, E., *Recreating the gut on-chip: Sensors and fabrication technologies for aerobic intestinal host – anaerobic microbiota research 2022*, University of Twente.
638. Nairn, A.V., et al., *Regulation of glycan structures in murine embryonic stem cells: combined transcript profiling of glycan-related genes and glycan structural analysis*. J Biol Chem, 2012. **287**(45): p. 37835-56.
639. Dworkin, L.A., H. Clausen, and H.J. Joshi, *Applying transcriptomics to studyglycosylation at the cell type level*. iScience, 2022. **25**(6): p. 104419.
640. Ringot-Destrez, B., et al., *A Sensitive and Rapid Method to Determin the Adhesion Capacity of Probiotics and Pathogenic Microorganisms to Human Gastrointestinal Mucins*. Microorganisms, 2018. **6**(2).
641. Reunanen, J., et al., *Akkermansia muciniphila Adheres to Enterocytes and Strengthens the Integrity of the Epithelial Cell Layer*. Appl Environ Microbiol, 2015. **81**(11): p. 3655-62.
642. Earley, H., et al., *A Preliminary Study Examining the Binding Capacity of Akkermansia muciniphila and Desulfovibrio spp., to Colonic Mucin in Health and Ulcerative Colitis*. PLoS One, 2015. **10**(10): p. e0135280.
643. Becken, B., et al., *Genotypic and Phenotypic Diversity among Human Isolates of Akkermansia muciniphila*. mBio, 2021. **12**(3).



## APPENDICES

644. Konstantinidi, A., et al., *Exploring the glycosylation of mucins by use of O-glycodomain reporters recombinantly expressed in glycoengineered HEK293 cells*. J Biol Chem, 2022. **298**(4): p. 101784.
645. Taleb, V., et al., *Structural and mechanistic insights into the cleavage of clustered O-glycan patches-containing glycoproteins by mucinases of the human gut*. Nat Commun, 2022. **13**(1): p. 4324.
646. Stams, A.J., et al., *Growth of syntrophic propionate-oxidizing bacteria with fumarate in the absence of methanogenic bacteria*. Appl Environ Microbiol, 1993. **59**(4): p. 1114-9.
647. van der Ark, K.C.H., et al., *Model-driven design of a minimal medium for Akkermansia muciniphila confirms mucus adaptation*. Microb Biotechnol, 2018. **11**(3): p. 476-485.
648. Miller, R.S. and L.C. Hoskins, *Mucin degradation in human colon ecosystems: Fecal population densities of mucin-degrading bacteria estimated by a "most probable number" method*. Gastroenterology, 1981. **81**(4): p. 759-765.
649. Co, J.Y., et al., *Mucins trigger dispersal of Pseudomonas aeruginosa biofilms*. NPJ Biofilms Microbiomes, 2018. **4**: p. 23.
650. Tailford, L.E., et al., *Discovery of intramolecular trans-sialidases in human gut microbiota suggests novel mechanisms of mucosal adaptation*. Nat Commun, 2015. **6**: p. 7624.
651. de Haan, N., et al., *In-Depth Profiling of O-Glycan Isomers in Human Cells Using C18 Nanoliquid Chromatography-Mass Spectrometry and Glycogenomics*. Anal Chem, 2022. **94**(10): p. 4343-4351.
652. Damerell, D., et al., *The GlycanBuilder and GlycoWorkbench glycoinformatics tools: updates and new developments*. Biol Chem, 2012. **393**(11): p. 1357-62.
653. Chandrasekaran, E.V., et al., *Novel interactions of complex carbohydrates with peanut (PNA), Ricinus communis (RCA-I), Sambucus nigra (SNA-I) and wheat germ (WGA) agglutinins as revealed by the binding specificities of these lectins towards mucin core-2 O-linked and N-linked glycans and related structures*. Glycoconjugate Journal, 2016. **33**(5): p. 819-836.
654. Ottman, N., et al., *Genome-Scale Model and Omics Analysis of Metabolic Capacities of Akkermansia muciniphila Reveal a Preferential Mucin-Degrading Lifestyle*. Appl Environ Microbiol, 2017. **83**(18).
655. Huang, K., et al., *Biochemical characterisation of the neuraminidase pool of the human gut symbiont Akkermansia muciniphila*. Carbohydr Res, 2015. **415**: p. 60-5.
656. Luis, A.S., et al., *Sulfated glycan recognition by carbohydrate sulfatases of the human gut microbiota*. Nat Chem Biol, 2022. **18**(8): p. 841-849.
657. Mello, L.V., X. Chen, and D.J. Rigden, *Mining metagenomic data for novel domains: BACON, a new carbohydrate-binding module*. FEBS Lett, 2010. **584**(11): p. 2421-6.
658. Larsbrink, J., et al., *A discrete genetic locus confers xyloglucan metabolism in select human gut Bacteroidetes*. Nature, 2014. **506**(7489): p. 498-502.
659. Nakjang, S., et al., *A novel extracellular metallopeptidase domain shared by animal host-associated mutualistic and pathogenic microbes*. PLoS One, 2012. **7**(1): p. e30287.

660. Ottman, N., et al., *Characterization of Outer Membrane Proteome of Akkermansia muciniphila Reveals Sets of Novel Proteins Exposed to the Human Intestine*. *Front Microbiol*, 2016. **7**: p. 1157.
661. Crouch, L.I., et al., *Prominent members of the human gut microbiota express endo-acting O-glycanases to initiate mucin breakdown*. *Nat Commun*, 2020. **11**(1): p. 4017.
662. Trastoy, B., et al., *Structural basis of mammalian mucin processing by the human gut O-glycopeptidase OgpA from Akkermansia muciniphila*. *Nat Commun*, 2020. **11**(1): p. 4844.
663. Shon, D.J., et al., *An enzymatic toolkit for selective proteolysis, detection, and visualization of mucin-domain glycoproteins*. *Proc Natl Acad Sci U S A*, 2020. **117**(35): p. 21299-21307.
664. Medley, B.J., et al., *A previously uncharacterized O-glycopeptidase from Akkermansia muciniphila requires the Tn-antigen for cleavage of the peptide bond*. *J Biol Chem*, 2022. **298**(10): p. 102439.
665. Malaker, S.A., et al., *The mucin-selective protease StcE enables molecular and functional analysis of human cancer-associated mucins*. *Proc Natl Acad Sci U S A*, 2019. **116**(15): p. 7278-7287.
666. Guo, X., et al., *Genome sequencing of 39 Akkermansia muciniphila isolates reveals its population structure, genomic and functional diversity, and global distribution in mammalian gut microbiotas*. *BMC Genomics*, 2017. **18**(1): p. 800.
667. Ouwerkerk, J.P., et al., *Comparative Genomics and Physiology of Akkermansia muciniphila Isolates from Human Intestine Reveal Specialized Mucosal Adaptation*. *Microorganisms*, 2022. **10**(8).
668. Kirmiz, N., et al., *Comparative Genomics Guides Elucidation of Vitamin B(12) Biosynthesis in Novel Human-Associated Akkermansia Strains*. *Appl Environ Microbiol*, 2020. **86**(3).
669. Karcher, N., et al., *Genomic diversity and ecology of human-associated Akkermansia species in the gut microbiome revealed by extensive metagenomic assembly*. *Genome Biol*, 2021. **22**(1): p. 209.
670. Sun, L., et al., *Installation of O-glycan sulfation capacities in human HEK293 cells for display of sulfated mucins*. *J Biol Chem*, 2022. **298**(2): p. 101382.
671. Omary, M.B., et al., *Not all mice are the same: Standardization of animal research data presentation*. *Hepatology*, 2016. **63**(6): p. 1752-1754.
672. Thomas, P. and T.G. Smart, *HEK293 cell line: a vehicle for the expression of recombinant proteins*. *J Pharmacol Toxicol Methods*, 2005. **51**(3): p. 187-200.
673. Ingber, D.E., *Human organs-on-chips for disease modelling, drug development and personalized medicine*. *Nat Rev Genet*, 2022. **23**(8): p. 467-491.
674. Trietsch, S.J., et al., *Membrane-free culture and real-time barrier integrity assessment of perfused intestinal epithelium tubes*. *Nat Commun*, 2017. **8**(1): p. 262.
675. Chen, Y., et al., *Robust bioengineered 3D functional human intestinal epithelium*. *Scientific Reports*, 2015. **5**: p. 1-11.





## APPENDICES

676. Zhang, J., et al., *Primary human colonic mucosal barrier crosstalk with super oxygen-sensitive Faecalibacterium prausnitzii in continuous culture*. Med (N Y), 2021. **2**(1): p. 74-98 e9.
677. Kim, R., et al., *A Platform for Co-Culture of Primary Human Colonic Epithelium With Anaerobic Probiotic Bacteria*. Front Bioeng Biotechnol, 2022. **10**: p. 890396.
678. Zheng, L., C.J. Kelly, and S.P. Colgan, *Physiologic hypoxia and oxygen homeostasis in the healthy intestine. A Review in the Theme: Cellular Responses to Hypoxia*. Am J Physiol Cell Physiol, 2015. **309**(6): p. C350-60.
679. Kim, R., et al., *An in vitro intestinal platform with a self-sustaining oxygen gradient to study the human gut/microbiome interface*. Biofabrication, 2019. **12**(1): p. 015006.
680. DiMarco, R.L., et al., *Engineering of three-dimensional microenvironments to promote contractile behavior in primary intestinal organoids*. Integr Biol (Camb), 2014. **6**(2): p. 127-142.
681. Williamson, I.A., et al., *A High-Throughput Organoid Microinjection Platform to Study Gastrointestinal Microbiota and Luminal Physiology*. Cell Mol Gastroenterol Hepatol, 2018. **6**(3): p. 301-319.
682. Puschhof, J., C. Pleguezuelos-Manzano, and H. Clevers, *Organoids and organs-on-chips: Insights into human gut-microbe interactions*. Cell Host Microbe, 2021. **29**(6): p. 867-878.
683. Chen, Y., et al., *Bioengineered 3D Tissue Model of Intestine Epithelium with Oxygen Gradients to Sustain Human Gut Microbiome*. Adv Healthc Mater, 2022. **11**(16): p. e2200447.
684. Orrhage, K. and C.E. Nord, *Factors controlling the bacterial colonization of the intestine in breastfed infants*. Acta Paediatr Suppl, 1999. **88**(430): p. 47-57.
685. Friedman, E.S., et al., *Microbes vs. chemistry in the origin of the anaerobic gut lumen*. Proc Natl Acad Sci U S A, 2018. **115**(16): p. 4170-4175.
686. Semenza, G.L., *Hypoxia-inducible factors in physiology and medicine*. Cell, 2012. **148**(3): p. 399-408.
687. Louis, N.A., et al., *Selective induction of mucin-3 by hypoxia in intestinal epithelia*. J Cell Biochem, 2006. **99**(6): p. 1616-27.
688. Furuta, G.T., et al., *Hypoxia-inducible factor 1-dependent induction of intestinal trefoil factor protects barrier function during hypoxia*. J Exp Med, 2001. **193**(9): p. 1027-34.
689. Young, H.W., et al., *Central role of Muc5ac expression in mucous metaplasia and its regulation by conserved 5' elements*. Am J Respir Cell Mol Biol, 2007. **37**(3): p. 273-90.
690. Kelly, C.J., et al., *Fundamental role for HIF-1 $\alpha$  in constitutive expression of human  $\beta$  defensin-1*. Mucosal Immunol, 2013. **6**(6): p. 1110-8.
691. Hassinen, A., et al., *A Golgi-associated redox switch regulates catalytic activation and cooperative functioning of ST6Gal-I with B4GalT-I*. Redox Biol, 2019. **24**: p. 101182.
692. Ilani, T., et al., *The disulfide catalyst QSOX1 maintains the colon mucosal barrier by regulating Golgi glycosyltransferases*. Embo j, 2022: p. e111869.
693. Beaurivage, C., et al., *Development of a human primary gut-on-a-chip to model inflammatory processes*. Sci Rep, 2020. **10**(1): p. 21475.

694. Sunuwar, L., et al., *Mechanical Stimuli Affect Escherichia coli Heat-Stable Enterotoxin-Cyclic GMP Signaling in a Human Enteroid Intestine-Chip Model*. Infect Immun, 2020. **88**(3).
695. Harper, C.E. and C.J. Hernandez, *Cell biomechanics and mechanobiology in bacteria: Challenges and opportunities*. APL Bioeng, 2020. **4**(2): p. 021501.
696. Engle, M.J., G.S. Goetz, and D.H. Alpers, *Caco-2 cells express a combination of colonocyte and enterocyte phenotypes*. J Cell Physiol, 1998. **174**(3): p. 362-9.
697. Shin, W., et al., *Single-cell transcriptomic mapping of intestinal epithelium that undergoes 3D morphogenesis and mechanodynamic stimulation in a gut-on-a-chip*. iScience, 2022. **25**(12): p. 105521.
698. Barrila, J., et al., *Modeling Host-Pathogen Interactions in the Context of the Microenvironment: Three-Dimensional Cell Culture Comes of Age*. Infect Immun, 2018. **86**(11).
699. Wang, Y., et al., *Formation of Human Colonic Crypt Array by Application of Chemical Gradients Across a Shaped Epithelial Monolayer*. Cell Mol Gastroenterol Hepatol, 2018. **5**(2): p. 113-130.
700. Ashammakhi, N., et al., *Advances and Future Perspectives in 4D Bioprinting*. Biotechnol J, 2018. **13**(12): p. e1800148.
701. Ashammakhi, N., et al., *Gut-on-a-chip: Current progress and future opportunities*. Biomaterials, 2020. **255**: p. 120196.
702. Li, H., et al., *The outer mucus layer hosts a distinct intestinal microbial niche*. Nature communications, 2015. **6**(1): p. 8292-8292.
703. Grassart, A., et al., *Bioengineered Human Organ-on-Chip Reveals Intestinal Microenvironment and Mechanical Forces Impacting Shigella Infection*. Cell Host & Microbe, 2019. **26**(4): p. 565.
704. Bergstrom, K., et al., *Proximal colon-derived O-glycosylated mucus encapsulates and modulates the microbiota*. Science, 2020. **370**(6515): p. 467-472.
705. Kaiko, G.E., et al., *The Colonic Crypt Protects Stem Cells from Microbiota-Derived Metabolites*. Cell, 2016. **165**(7): p. 1708-1720.
706. Bilen, M., et al., *The contribution of culturomics to the repertoire of isolated human bacterial and archaeal species*. Microbiome, 2018. **6**(1): p. 94.
707. Wang, Y., et al., *Single cell stable isotope probing in microbiology using Raman microspectroscopy*. Curr Opin Biotechnol, 2016. **41**: p. 34-42.
708. Wu, F. and C. Dekker, *Nanofabricated structures and microfluidic devices for bacteria: from techniques to biology*. Chem Soc Rev, 2016. **45**(2): p. 268-80.
709. Thompson, A.W., et al., *A method to analyze, sort, and retain viability of obligate anaerobic microorganisms from complex microbial communities*. J Microbiol Methods, 2015. **117**: p. 74-7.
710. Bellais, S., et al., *Species-targeted sorting and cultivation of commensal bacteria from the gut microbiome using flow cytometry under anaerobic conditions*. Microbiome, 2022. **10**(1): p. 24.
711. Ponomarova, O. and K.R. Patil, *Metabolic interactions in microbial communities: untangling the Gordian knot*. Curr Opin Microbiol, 2015. **27**: p. 37-44.
712. Tramontano, M., et al., *Nutritional preferences of human gut bacteria reveal their metabolic idiosyncrasies*. Nat Microbiol, 2018. **3**(4): p. 514-522.



## APPENDICES

713. Weiss, A.S., et al., *In vitro* interaction network of a synthetic gut bacterial community. *Isme j*, 2022. **16**(4): p. 1095-1109.
714. Shetty, S.A., et al., *Intestinal microbiome landscaping: Insight in community assemblage and implications for microbial modulation strategies*. *FEMS Microbiology Reviews*, 2017. **41**(2): p. 182-199.
715. Arumugam, M., et al., *Enterotypes of the human gut microbiome*. *Nature*, 2011. **473**(7346): p. 174-80.
716. Shetty, S.A., H. Smidt, and W.M. de Vos, *Reconstructing functional networks in the human intestinal tract using synthetic microbiomes*. *Current Opinion in Biotechnology*, 2019. **58**: p. 146-154.
717. Garcia-Vello, P., et al., *Peptidoglycan from Akkermansia muciniphila MucT: chemical structure and immunostimulatory properties of muropeptides*. *Glycobiology*, 2022. **32**(8): p. 712-719.
718. Bae, M., et al., *Akkermansia muciniphila phospholipid induces homeostatic immune responses*. *Nature*, 2022. **608**(7921): p. 168-173.
719. Davey, L., et al., *Mucin foraging enables Akkermansia muciniphila to compete against other microbes in the gut and to modulate host sterol biosynthesis*. 2022.
720. Honn, K.V., J.A. Singley, and W. Chavin, *Fetal bovine serum: a multivariate standard*. *Proc Soc Exp Biol Med*, 1975. **149**(2): p. 344-7.
721. Price, P.J. and E.A. Gregory, *Relationship between in vitro growth promotion and biophysical and biochemical properties of the serum supplement*. *In Vitro*, 1982. **18**(6): p. 576-84.
722. Cassotta, M., et al., *A worldwide survey on the use of animal-derived materials and reagents in scientific experimentation*. *Eng Life Sci*, 2022. **22**(9): p. 564-583.
723. Levery, S.B., et al., *Advances in mass spectrometry driven O-glycoproteomics*. *Biochimica et Biophysica Acta (BBA) - General Subjects*, 2015. **1850**(1): p. 33-42.
724. Christensen, J., et al., *Defining new criteria for selection of cell-based intestinal models using publicly available databases*. *BMC genomics*, 2012. **13**(1): p. 274-284.
725. Nosek, B.A., Alter, G., Banks, G. C., Borsboom, D., Bowman, S. D., Breckler, S. J., ... DeHaven, A. C. . *Transparency and Openness Promotion (TOP) Guidelines*. 2021 30-12-2022]; Available from: [osf.io/9f6gx](https://osf.io/9f6gx).
726. Madunić, K., et al., *O-glycomic and proteomic signatures of spontaneous and butyrate-stimulated colorectal cancer cell line differentiation*. *Mol Cell Proteomics*, 2023: p. 100501.
727. Bauer, E. and I. Thiele, *From Network Analysis to Functional Metabolic Modeling of the Human Gut Microbiota*. *mSystems*, 2018. **3**(3).
728. Vrancken, G., et al., *Synthetic ecology of the human gut microbiota*. *Nat Rev Microbiol*, 2019. **17**(12): p. 754-763.
729. Magnúsdóttir, S., et al., *Generation of genome-scale metabolic reconstructions for 773 members of the human gut microbiota*. *Nature Biotechnology*, 2017. **35**(1): p. 81-89.
730. Zimmermann, M., et al., *Mapping human microbiome drug metabolism by gut bacteria and their genes*. *Nature*, 2019. **570**(7762): p. 462-467.

731. Zimmermann, M., et al., *Separating host and microbiome contributions to drug pharmacokinetics and toxicity*. Science, 2019. **363**(6427).
732. Tatonetti, N.P., *Translational medicine in the Age of Big Data*. Brief Bioinform, 2019. **20**(2): p. 457-462.
733. Walter, J., et al., *Establishing or Exaggerating Causality for the Gut Microbiome: Lessons from Human Microbiota-Associated Rodents*. Cell, 2020. **180**(2): p. 221-232.
734. Chung, H., et al., *Gut Immune Maturation Depends on Colonization with a Host-Specific Microbiota*. Cell, 2012. **149**(7): p. 1578-1593.
735. Lundberg, R., et al., *Human microbiota-transplanted C57BL/6 mice and offspring display reduced establishment of key bacteria and reduced immune stimulation compared to mouse microbiota-transplantation*. Sci Rep, 2020. **10**(1): p. 7805.
736. Rosshart, S.P., et al., *Laboratory mice born to wild mice have natural microbiota and model human immune responses*. Science, 2019. **365**(6452).
737. Lagkouvardos, I., et al., *The Mouse Intestinal Bacterial Collection (miBC) provides host-specific insight into cultured diversity and functional potential of the gut microbiota*. Nature microbiology, 2016. **1**(10): p. 16131-16131.
738. Afrizal, A., et al., *Enhanced cultured diversity of the mouse gut microbiota enables custom-made synthetic communities*. Cell Host & Microbe, 2022.
739. Liu, C., et al., *The Mouse Gut Microbial Biobank expands the coverage of cultured bacteria*. Nature communications, 2020. **11**(1): p. 1-12.
740. Beresford-Jones, B.S., et al., *The Mouse Gastrointestinal Bacteria Catalogue enables translation between the mouse and human gut microbiotas via functional mapping*. Cell Host & Microbe, 2022. **30**(1): p. 124-138.e8.
741. Tovaglieri, A., et al., *Species-specific enhancement of enterohemorrhagic E. coli pathogenesis mediated by microbiome metabolites*. Microbiome, 2019. **7**(1): p. 43.
742. Wu, H., et al., *The human gut symbiont Ruminococcus gnavus shows specificity to blood group A antigen during mucin glycan foraging: Implication for niche colonisation in the gastrointestinal tract*. PLoS Biol, 2021. **19**(12): p. e3001498.
743. Shin, Y.C., et al., *Three-Dimensional Regeneration of Patient-Derived Intestinal Organoid Epithelium in a Physiodynamic Mucosal Interface-on-a-Chip*. Micromachines (Basel), 2020. **11**(7).
744. Bein, A., et al., *Nutritional deficiency in an intestine-on-a-chip recapitulates injury hallmarks associated with environmental enteric dysfunction*. Nat Biomed Eng, 2022. **6**(11): p. 1236-1247.
745. Wang, B.X., C.M. Wu, and K. Ribbeck, *Home, sweet home: how mucus accommodates our microbiota*. FEBS J, 2021. **288**(6): p. 1789-1799.
746. Liu, S., et al., *The Host Shapes the Gut Microbiota via Fecal MicroRNA*. Cell Host Microbe, 2016. **19**(1): p. 32-43.
747. Wesener, D.A., et al., *Recognition of microbial glycans by human intelectin-1*. Nat Struct Mol Biol, 2015. **22**(8): p. 603-10.
748. Diaz-Garrido, N., et al., *Cell-to-Cell Communication by Host-Released Extracellular Vesicles in the Gut: Implications in Health and Disease*. Int J Mol Sci, 2021. **22**(4).



## APPENDICES

749. Groussin, M., F. Mazel, and E.J. Alm, *Co-evolution and Co-speciation of Host-Gut Bacteria Systems*. Cell Host Microbe, 2020. **28**(1): p. 12-22.
750. Whitaker, W.R., E.S. Shepherd, and J.L. Sonnenburg, *Tunable Expression Tools Enable Single-Cell Strain Distinction in the Gut Microbiome*. Cell, 2017. **169**(3): p. 538-546.e12.
751. Schmidt, F., et al., *Noninvasive assessment of gut function using transcriptional recording sentinel cells*. Science, 2022. **376**(6594): p. eabm6038.
752. Landry, B.P. and J.J. Tabor, *Engineering Diagnostic and Therapeutic Gut Bacteria*. Microbiol Spectr, 2017. **5**(5).
753. Rubin, B.E., et al., *Species- and site-specific genome editing in complex bacterial communities*. Nat Microbiol, 2022. **7**(1): p. 34-47.
754. Nelson, M.T., et al., *Characterization of an engineered live bacterial therapeutic for the treatment of phenylketonuria in a human gut-on-a-chip*. Nat Commun, 2021. **12**(1): p. 2805.
755. Meisner, A., et al., *Calling for a systems approach in microbiome research and innovation*. Curr Opin Biotechnol, 2022. **73**: p. 171-178.
756. Frith, U., *Fast Lane to Slow Science*. Trends Cogn Sci, 2020. **24**(1): p. 1-2.
757. Stengers, I., *Another science is possible: A manifesto for slow science*. 2018: John Wiley & Sons.
758. Bornmann, L. and R. Mutz, *Growth rates of modern science: A bibliometric analysis based on the number of publications and cited references*. Journal of the Association for Information Science and Technology, 2015. **66**(11): p. 2215-2222.
759. Larsen, P.O. and M. von Ins, *The rate of growth in scientific publication and the decline in coverage provided by Science Citation Index*. Scientometrics, 2010. **84**(3): p. 575-603.
760. Nabout, J.C., et al., *Publish (in a group) or perish (alone): the trend from single- to multi-authorship in biological papers*. Scientometrics, 2014. **102**(1): p. 357-364.
761. Wardil, L. and C. Hauert, *Cooperation and coauthorship in scientific publishing*. Physical Review E, 2015. **91**(1): p. 012825.
762. Brischoux, F. and F. Angelier, *Academia's never-ending selection for productivity*. Scientometrics, 2015. **103**(1): p. 333-336.
763. Neill, U.S., *Publish or perish, but at what cost?* J Clin Invest, 2008. **118**(7): p. 2368.
764. Fanelli, D., *Do pressures to publish increase scientists' bias? An empirical support from US States Data*. PLoS One, 2010. **5**(4): p. e10271.
765. Shi, F., et al., *The wisdom of polarized crowds*. Nature Human Behaviour, 2019. **3**(4): p. 329-336.
766. Mercier, H., *The Argumentative Theory: Predictions and Empirical Evidence*. Trends in Cognitive Sciences, 2016. **20**(9): p. 689-700.
767. Bengert, E., et al., *Novel approach for tracking interdisciplinary research productivity using institutional databases*. J Clin Transl Sci, 2022. **6**(1): p. e119.
768. NWO. *Ambitions 'Science works!' - NWO strategy 2023 – 2026* 2023 [cited 2023 04-01-2023]; Available from: <https://www.nwo.nl/en/ambitions-2023-2026#unimpeded-collaboration>.

769. Smaldino, P.E. and R. McElreath, *The natural selection of bad science*. R Soc Open Sci, 2016. **3**(9): p. 160384.
770. Holcombe, A., *Farewell authors, hello contributors*. Nature, 2019. **571**(7764): p. 147.
771. *Authorship policies*. Nature, 2009. **458**(7242): p. 1078-1078.
772. VSNU, N., KNAW, NWO and ZonMw. *Position paper 'Room for everyone's talent'*. 2019 [cited 2023 05-01-2022]; Available from: <https://www.nwo.nl/en/position-paper-room-everyones-talent>.
773. Hyland-Wood, B., et al., *Toward effective government communication strategies in the era of COVID-19*. Humanities and Social Sciences Communications, 2021. **8**(1): p. 30.
774. Matta, G., *Science communication as a preventative tool in the COVID19 pandemic*. Humanities and Social Sciences Communications, 2020. **7**(1): p. 159.
775. Timmis, K., et al., *The urgent need for microbiology literacy in society*. Environ Microbiol, 2019. **21**(5): p. 1513-1528.
776. D'Hondt, K., et al., *Microbiome innovations for a sustainable future*. Nature Microbiology, 2021. **6**(2): p. 138-142.
777. van Bruggen, A.H.C., et al., *One Health - Cycling of diverse microbial communities as a connecting force for soil, plant, animal, human and ecosystem health*. Science of The Total Environment, 2019. **664**: p. 927-937.
778. Timmis, K., J. Timmis, and F. Jebok, *The urgent need for microbiology literacy in society: children as educators*. Microb Biotechnol, 2020. **13**(5): p. 1300-1303.
779. Bradshaw, A., *Microbiological literacy and the role of social science: a response to Timmis et al.* Environ Microbiol, 2021. **23**(11): p. 6350-6354.
780. Nisbet, M.C. and D.A. Scheufele, *What's next for science communication? Promising directions and lingering distractions*. Am J Bot, 2009. **96**(10): p. 1767-78.
781. Reincke, C.M., A.L. Bredenoord, and M.H. van Mil, *From deficit to dialogue in science communication: The dialogue communication model requires additional roles from scientists*. EMBO Rep, 2020. **21**(9): p. e51278.
782. Nathans, J. and P. Sterling, *How scientists can reduce their carbon footprint*. Elife, 2016. **5**.
783. Urbina, M.A., A.J.R. Watts, and E.E. Reardon, *Labs should cut plastic waste too*. Nature, 2015. **528**(7583): p. 479-479.
784. Scientific, T. *Sustainable Product Design*. 05-01-2023]; Available from: <https://www.thermofisher.com/nl/en/home/about-us/product-stewardship.html>.
785. GmbH, S.I. *Tips for a more sustainable lab*. [cited 2023 05-01-2023]; Available from: <https://www.starlabgroup.com/en/sustainability/sustainable-lab-tips.html>.
786. Gstraunthaler, G., *Alternatives to the use of fetal bovine serum: serum-free cell culture*. Altex, 2003. **20**(4): p. 275-81.
787. Brunner, D., et al., *Serum-free cell culture: the serum-free media interactive online database*. Altex, 2010. **27**(1): p. 53-62.



## APPENDICES

- 788. Post, M.J., et al., *Scientific, sustainability and regulatory challenges of cultured meat*. Nature Food, 2020. **1**(7): p. 403-415.
- 789. Stout, A.J., et al., *Simple and effective serum-free medium for sustained expansion of bovine satellite cells for cell cultured meat*. Communications Biology, 2022. **5**(1).
- 790. Okamoto, Y., et al., *Proliferation and differentiation of primary bovine myoblasts using Chlorella vulgaris extract for sustainable production of cultured meat*. Biotechnology Progress, 2022. **38**(3): p. e3239.
- 791. Messmer, T., et al., *A serum-free media formulation for cultured meat production supports bovine satellite cell differentiation in the absence of serum starvation*. Nature Food, 2022. **3**(1): p. 74-85.

## English Summary

*Too long? Go to “Summary for Lego audience” on p. 290*

The human intestinal ecosystem is characterized by a complex interplay between microorganisms and the host. The high variation within the human population with respect to the composition of the intestinal microbiota further complicates the quest towards adequate understanding of this phenomenon so relevant to human health. To study host-microbe interactions, researchers have relied heavily on laboratory animals, but scientific as well as ethical objections exist. In **Chapter 1**, I introduced the human GIT, including its inhabitants, and presented the currently available *in vivo* and *in vitro* models to study host-microbe interactions, with a specific focus on the human colonic mucosal layer.

In general, this thesis can be divided in two parts: In the first part we critically evaluated the use of both *in vivo* and *in vitro* models of the human GIT. In **Chapter 2**, we reviewed the use of defined microbial communities in gnotobiotic animal models in which host parameters are assessed. This literature review reinforced that our limited understanding has often hampered appropriate design of defined communities that represent the human gut microbiota. On top of this, some communities have been applied to *in vivo* models that differ appreciably from the human host. In this chapter, the advantages and disadvantages of using defined microbial communities were outlined, and suggestions for future improvement of host-microbe interaction models were provided. With respect to the microbiota, due to increasing availability of representative cultured isolates and their genomic sequences, our understanding and controllability of the human gut ‘core microbiota’ is likely to increase. With respect to the host, technological advances, such as the development of a gut-on-a-chip and intestinal organoids, may contribute to more accurate *in vitro* models of the human host. These *in vitro* models were critically assessed in **Chapter 3**, centred around the Caco-2 cell line as a well-accepted and highly characterized intestinal barrier model cell line. This cell line has been cultured in different *in vitro* models, ranging from simple static to complex dynamic microfluidic models. We aimed to investigate the effect of these different *in vitro* experimental variables on gene expression. To this end, we systematically collected and extracted data from studies in which transcriptome analyses were performed on Caco-2 cells grown on permeable membranes. A collection of 13 studies comprising 100 samples revealed a weak association of experimental variables with overall as well as individual gene expression. This can be explained by the large heterogeneity in cell culture practice, or the lack of adequate reporting thereof, as suggested by our systematic analysis of experimental parameters not included in the main analysis. Given the rapidly increasing use of *in vitro* cell culture models, including more advanced (micro)fluidic models, our analysis reinforced the need for more standardized reporting protocols





and could serve as a template for future comparative studies on transcriptome and other *in vitro* data.

In the second part of this thesis, we took a closer look at a crucial site for host-microbe interactions: the colonic mucosal layer, which plays a crucial role in human health. To study intestinal mucus function and structure *in vitro*, the mucus-producing intestinal cell line HT29-MTX has been commonly used. However, this cell line produces only low amounts of the intestine-specific MUC2. It has been shown previously that HT29-MTX(-E12) cells cultured under Semi-Wet interface with Mechanical Stimulation (SWMS) produced higher amounts of MUC2, concomitant with a thicker mucus layer, compared to cells cultured conventionally. However, it remains unknown which underlying pathways are involved. Therefore, in **Chapter 4**, we aimed to further explore the cellular processes underlying the increased MUC2 production by HT29-MTX-E12 cells grown under SWMS conditions. Cells grown on Transwell membranes for 14 days under static and SWMS conditions were subjected to transcriptome analysis to investigate underlying molecular pathways at gene expression level. Caco-2 and LS174T cell lines were included as references. We characterized how SWMS conditions affected HT29-MTX-E12 cells in terms of epithelial barrier integrity, by measuring transepithelial electrical resistance, and cell metabolism, by monitoring pH and lactate production per molecule glucose of the conditioned medium. We confirmed higher MUC2 production under SWMS conditions at gene transcript and protein level and demonstrated that this culturing method primarily stimulated cell growth. In addition, we also found evidence for a more aerobic cell metabolism under SWMS conditions, as shown previously for similar models. We suggested different mechanisms by which MUC2 production is enhanced under SWMS and proposed potential applications of this model in future studies. Subsequently, in **Chapter 5**, we used this *in vitro* model to assess the effect of mucus-associated bacterial species on mucus production. Mucus consists mainly of mucin, a heavily O-glycosylated protein, which, amongst others, provides an attachment site and nutrient source for mucus-associated bacteria, such as *Akkermansia muciniphila*. *A. muciniphila* is a next-generation beneficial organism which is able to grow on mucin as the sole source of carbon and nitrogen, and of which the abundance in the human GIT has been correlated with a healthy intestine. Although the impact of this intestinal symbiont on the mucus layer has been studied *in vivo*, the exact effects and underlying mechanisms of this organism on mucus production and glycosylation, remain understudied. Therefore, we used HT29-MTX-E12 cultured under SWMS conditions to test the effect of pasteurized *A. muciniphila* on mucus production *in vitro* by transcriptome analysis. In brief, we detected little to no effects of *A. muciniphila* under the studied conditions. This study provided a first step towards additional experiments to investigate microbe-mucus effects in different *in vitro* models of the human colonic mucus layer, including a gut-on-a-chip and cultivation in flasks. Potentially, these methods could serve as screening methods before investigation in more complex *in vitro* or *in vivo*

models. In **Chapter 6**, we took a closer look at the molecular interactions of *A. muciniphila* with human mucins. Little is known about the mucin-binding properties of *A. muciniphila* despite adhesion to mucins being crucial for its persistence within the human GIT. In this study, we showed that mucin binding of the anaerobic *A. muciniphila* is independent of environmental oxygen levels and barely affected by pasteurization. We further dissected binding of pasteurized *A. muciniphila* to mucins using recombinantly expressed human mucin tandem repeat reporter modules carrying distinct structures and patterns of O-glycans. Using these mucin reporters, we demonstrated that pasteurized *A. muciniphila* selectively recognized the LacNAc (Gal $\beta$ 1-4GlcNAc $\beta$ 1-R) disaccharide on mucin O-glycans, which is abundantly found in human colonic mucins. We further showed that desialylation by endogenous *A. muciniphila* sialidase activity promoted binding. In summary, our study provided novel insights into microbe-mucin interactions important for recognition and colonization by a key mucin-foraging bacterium.

In **Chapter 7**, I evaluated the *in vitro* models used in this thesis, as well as recently developed, more advanced *in vitro* models of the human GIT that incorporated a mucus layer more representative of the one observed *in vivo*. The current caveats with respect to the colonic mucosal layer and co-culture with intestinal bacteria were discussed. Subsequently, I made suggestions on more representative *in vitro* models of the human colon, including the characterization of individual bacterial strains, standardization of cell culture practice, the embracement of relatively simple cellular *in vitro* models as well as complete reporting research methods and results. In addition, the microbial component should not be forgotten. Eventually, this could lead to more complete *in vitro* models of the human GIT, which can be used as host/patient-specific models and allow investigation of the effects of host components on bacteria and host-bacteria adaptation. I defended that this approach fits within a bigger trend of slow science, which, moreover, has the potential to increase microbiology literacy in both society and science and lead to more sustainable lab practices.



## Nederlandse Samenvatting

*Te langdradig? Blader snel door naar “Samenvatting voor Legopubliek” on p. 292*

Het ecosysteem van het maagdarmstelsel van de mens wordt gekenmerkt door een complex samenspel van verschillende micro-organismen en de gastheer. Welke micro-organismen aanwezig zijn, verschilt per persoon, wat het nog moeilijker maakt om deze complexe interacties te begrijpen. Dit terwijl ze juist zo belangrijk zijn voor de gezondheid en het welzijn van de mens. De huidige wetenschappelijke kennis over de interacties tussen mens en microbe is voornamelijk gebaseerd op proefdierstudies, maar dit soort onderzoek loopt tegen steeds meer ethische en wetenschappelijke bezwaren aan. In **hoofdstuk 1** introduceerde ik het menselijke maagdarmstelsel en zijn bewoners, en gaf een overzicht van de beschikbare *in vivo* en *in vitro* modellen om de interacties tussen mens en microbe te bestuderen. Daarbij focuste ik specifiek op de mucuslaag (slijmvlies) die de darmcellen bedekt.

Over het algemeen kan de rest van het proefschrift worden opgedeeld in twee delen: In het eerste deel keken we kritisch naar het huidige gebruik van *in vivo* en *in vivo* modellen die het menselijke maagdarmstelsel representeren. In **hoofdstuk 2** beschreven we het gebruik van gedefinieerde microbiële gemeenschappen in gecontroleerde diermodellen in studies waar ook gekeken werd naar het effect van deze micro-organismen op de gastheer. Dit literatuuronderzoek bevestigde dat de samenstelling van gedefinieerde microbiële gemeenschappen als model voor de menselijke darmmicroben beperkt is geweest door onze geringe kennis over darmmicroben. Bovendien zijn deze gemeenschappen toegepast in diermodellen die substantieel verschillen van de menselijke gastheer. In dit hoofdstuk werden de voor- en nadelen van het gebruik van gedefinieerde microbiële gemeenschappen besproken, en werden suggesties gedaan voor toekomstige verbeteringen van modellen om de interacties tussen mens en microbe te bestuderen. Wat betreft microben zal ons begrip en beheersbaarheid van de menselijke “kernmicrobiota” alleen maar groeien, dankzij de toenemende beschikbaarheid van representatieve, kweekbare stammen en de kennis over hun genetische informatie. Met betrekking tot de gastheer zal technologische vooruitgang, zoals de ontwikkeling van een darm-op-een-chip en darmorganoïden bijdragen aan meer representatieve *in vitro* modellen van de menselijke darm. Dit soort modellen werd kritisch beoordeeld in **hoofdstuk 3**, met een centrale rol voor de cellijn Caco-2, een breed geaccepteerd en goed gekarakteriseerd model om de darmwand te onderzoeken. Deze cellijn is gekweekt in verschillende *in vitro* modellen, variërend van simpele, statische tot complexe, dynamische (microfluidische) modellen. We wilden het effect van deze verschillende *in vitro* experimentele variabelen op genexpressie bestuderen. Op systematische wijze verzamelden en extraheerden we data van studies waarin analyses waren gedaan op het transcriptoom van Caco-2 die gekweekt waren op permeabele

membranen. Een verzameling van 13 studies die 100 monsters omvatten liet zien dat er een zwak verband is tussen experimentele variabelen en zowel totale als individuele expressie van genen. Dit kan verklaard worden door de grote variatie in celkweekmethoden, of het gebrek aan het rapporteren daarvan, zoals gesuggereerd werd door onze systematische analyse van experimentele variabelen die niet in de initiële analyse waren geïnccludeerd. Met het oog op het toenemende gebruik van *in vitro* celkweekmodellen, waaronder meer geavanceerde (micro)fluidische modellen, benadrukte onze analyse de noodzaak voor verbeterde, gestandaardiseerde protocollen om te rapporteren en zou kunnen dienen als een voorbeeld voor toekomstige vergelijkende studies naar *in vitro* data van het transcriptoom of andere experimentele *in vitro* data.

In het tweede deel van dit proefschrift, doken we dieper in op een belangrijke plek voor interacties tussen microbe en mens: de mucuslaag van de dikke darm (colon), die een cruciale rol speelt in de gezondheid van de mens. Voor het bestuderen van de functie en structuur van darm mucus *in vitro*, wordt de mucusproducerende cellijn HT29-MTX vaak gebruikt. Deze cellijn produceert echter weinig van het darmspecifieke MUC2 en de mucuslaag is lang niet zo dik als *in vivo*. Andere wetenschappers hebben laten zien dat HT29-MTX-E12 cellen die gekweekt werden onder omstandigheden met minder vloeistof en mechanische stimulatie (Semi-Wet interface with Mechanical Stimulation) meer MUC2 produceren, wat resulteerde in een dikkere mucuslaag dan cellen die onder normale omstandigheden werden gekweekt. Het bleef echter onduidelijk waarom precies. Daarom verkenden wij in **hoofdstuk 4** de cellulaire processen die verantwoordelijk zouden kunnen zijn voor de verhoogde productie van MUC2 door HT29-MTX-E12 die gekweekt waren onder deze aangepaste omstandigheden. We groeiden de cellen voor 14 dagen onder statische en aangepaste omstandigheden en analyseerden het transcriptoom om onderliggende moleculaire processen op het niveau van de genen te onderzoeken. We includeerden ook Caco-2 en LS174T cellen als vergelijking. We karakteriseerden hoe de aangepaste omstandigheden de HT29-MTX-E12 cellen beïnvloedden door de integriteit van de epitheelcelbarrière te meten via de transepitheel elektrische weerstand, en het celstofwisseling, door pH en de productie van lactaat per molecuul glucose in het kweekmedium te meten. Onze metingen bevestigden dat er meer MUC2 werd geproduceerd onder de aangepaste omstandigheden, op zowel het niveau van het gen als het eiwit en lieten zien dat deze aangepaste kweekmethode vooral invloed had op celgroei. Daarnaast vonden we ook bewijs voor een meer aeroob celstofwisseling onder de aangepaste kweekmethode, wat ook is aangetoond voor soortgelijke modellen. We stelden verschillende mechanismen voor waardoor MUC2 productie onder deze aangepaste omstandigheden is verhoogd en deden voorstellen voor toepassingen van dit model in toekomstige studies. Vervolgens hebben we dit *in vitro* model zelf gebruikt in **hoofdstuk 5**, om het effect te onderzoeken van darmbacteriën die in de mucuslaag wonen op de productie van mucus. Mucus bestaat



voornamelijk uit mucine, een hevig O-glycosyleerd eiwit dat onder andere fungeert als landingsplek en voedsel voor darmbacteriën, zoals *Akkermansia muciniphila*. *A. muciniphila* is een bacterie die in staat is om mucine te gebruiken als enige bron van koolstof en stikstof, en waarvan de aanwezigheid in het menselijke maag-darmstelsel geassocieerd is met een gezonde darm. Hoewel de impact van deze darmsymbiont op de mucuslaag bestudeerd is *in vivo*, zijn de exacte effecten op de mucusproductie en -glycosylering – en de manier waarop – nog niet goed bestudeerd. Daarom kweekten we HT29-MTX-E12 onder de bovengenoemde aangepaste omstandigheden en onderzochten we het effect van gepasteuriseerde *A. muciniphila* op mucusproductie *in vitro* door middel van transcriptoom analyse. Kort samengevat vonden we weinig effect van *A. muciniphila* onder de geteste experimentele omstandigheden. Deze studie vormde de eerste stap naar vervolgexperimenten om het effect van microben op mucus te bestuderen in verschillende modellen van colon mét mucuslaag, waaronder een darm-op-een-chip en kweekflessen. Mogelijk kunnen deze modellen dienen als screenmethoden voorafgaand aan onderzoek in meer complexe *in vitro* en *in vivo* modellen.

In **hoofdstuk 6** doken we dieper in de moleculaire interacties tussen *A. muciniphila* en mucines. Er is weinig bekend over de binding van *A. muciniphila* aan mucine, ondanks dat binding aan mucine zo belangrijk is voor het overleven van sommige bacteriën in het maag-darmstelsel. In deze studie lieten we zien dat binding van *A. muciniphila* – een anaerobe bacterie – niet beïnvloed wordt door de zuurstofconcentraties in de omgeving, en we zagen dat zowel levende als gepasteuriseerde *A. muciniphila* op een soortgelijke manier aan mucus bonden. Daarnaast ontrafelden we de binding van gepasteuriseerde *A. muciniphila* aan mucine met het gebruik van stukjes (reporters) menselijke mucine eiwitten met gedefinieerde patronen van O-glycanen. Met deze mucine reporters lieten we zien dat *A. muciniphila* selectief de disaccharide LacNAc (Gal $\beta$ 1-4GlcNAc $\beta$ 1-R) op mucines van O-glycanen herkent, die veelvuldig aanwezig is in core2 en core3 O-glycanen, een veelvoorkomend glycanenpatroon in mucines in colon van de mens. We lieten ook zien dat het verwijderen van sialzuur door de activiteit van sialidases van *A. muciniphila* de binding bevorderde. Samengevat bood onze studie nieuw inzicht in de interacties tussen microbe en mucine die belangrijk zijn voor de herkenning door en vestiging van een belangrijke mucine-consumerende bacterie.

In **hoofdstuk 7** evalueerde ik de *in vitro* modellen die gebruikt zijn in dit proefschrift, evenals recent ontwikkelde, meer geavanceerde modellen van het menselijke maag-darmstelsel waarin een mucuslaag is gebruikt die meer lijkt op die *in vivo*. Ook werden de huidige valkuilen besproken met betrekking tot de mucuslaag van de colon en de kweek met darmbacteriën. Vervolgens stelde ik voor hoe toe te werken naar meer representatieve *in vitro* modellen van de colon van de mens, zoals het karakteriseren van individuele bacteriestammen, standaardisatie van kweekmethoden, het omarmen van relatief simpele *in vitro* celmodellen en ook het

compleet rapporteren van onderzoeksmethoden en resultaten. Daarbij is het van belang om altijd rekening te houden met de microbiële component. Uiteindelijk zou dit kunnen leiden tot completere *in vitro* modellen van het menselijke maag-darmstelsel die gebruikt kunnen worden als modellen voor een specifieke gastheer of patiënt en onderzoek faciliteren naar de effecten van gastheer op bacteriën en de aanpassing van gastheer en bacterie aan elkaar. Ik verdedigde dat deze aanpak past binnen de bredere trend van “slow science” dat bovendien ook de potentie heeft om de kennis van microbiologie in maatschappij en in de wetenschap zelf te vergroten en meer duurzame laboratoriumpraktijken te stimuleren.



## Summary for Lego Audience

**MICROBES are everywhere.** Our intestines are inhabited by many, of which most are bacteria. Bacteria come in different shapes and sizes, with each microbe its own job. Moreover, they are true team players. We know that the right combination of different bacteria can have a positive effect on our health. However, as we can barely see them with the naked eye and because there are so many, it is difficult to see from the outside what they are actually doing in there.



Three models: A human, a pig and a mouse / Drie modellen: Een wetenschapper (mens), een varken en een muis.

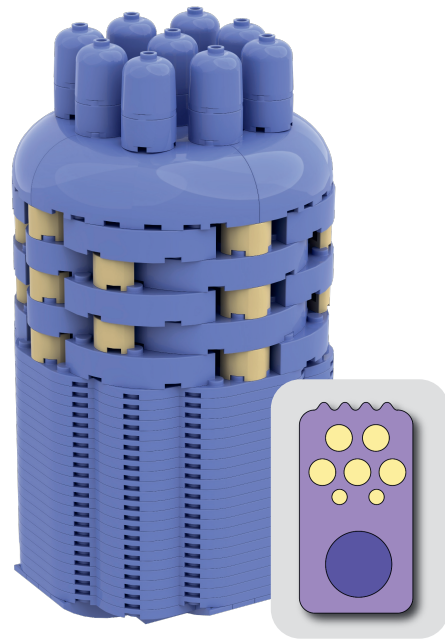
Scientists often use **MODELS** to understand how bacteria behave in our intestines. A model can be an animal, e.g. a mouse or (guinea) pig. We call them *in vivo* models. It has become evident, however, that these animals and their intestinal bacteria, are different from those of humans. Instead, scientists prefer to work with intestinal cells, which are the building blocks of our intestines. Using specific techniques in the lab, they can

grow these cells in such a way, that they behave like they would do in our intestines. We call these techniques "*in vitro* models". In the first part of my thesis, we re-emphasize that the way these animal and cell models are set up and described, can still be improved.

In the second part of my thesis, we zoom in on a specific group of intestinal bacteria, which are the ones living in the slime layer (mucus layer) of our intestines. This mucus layer consists mainly of **MUCINS**, proteins that are covered in a thick layer of sugars. The bacteria living in the mucus layer are fond of these sugars. Moreover, they convert these sugars in to compounds that can be used by other bacteria and our intestinal cells. When we grow these cells in the lab, however, they only produce small amounts of mucus. Except when we place them on a big shaker. We "stole" this idea from other scientists, but no one knew why the cells start behaving like that. Unfortunately we still do not know, but at least we make some suggestions.

In a different study, we introduced the same intestinal cells (in the lab) with an intestinal bacteria that is a professional mucin eater: *Akkermansia muciniphila* (*muciniphila* literally means “mucin-loving”). Under the conditions we tested, however, we observe very little effects of this bacterium on the intestinal cells. In another study, we did demonstrate when *Akkermansia muciniphila* binds to mucins, i.e. when a certain combination of sugars are present. This gives us clues on how this intestinal bacterium manages to survive in the densely populated intestines.

As usual, there is still a lot of research to do. In my thesis, I make several suggestions on how we can use (relatively simple) models to understand the interaction between intestinal bacteria and cells. Additionally, I explain how crucial it is to design and describe these experiments in a standardized way. Another possibility is to use more complex techniques in the lab to mimic the intestines of different people (young and old, healthy and diseased, etc..) to investigate the effect of intestinal bacteria on intestinal cells and the other way around. In any case, we should not forget about the presence of microbes. Lastly, I stress the importance of education on microbes to both scientists and a general audience, because **microbes are everywhere**.



*An intestinal cell producing mucus (yellow parts) / Een darmcel die mucus produceert (gele rondjes)*

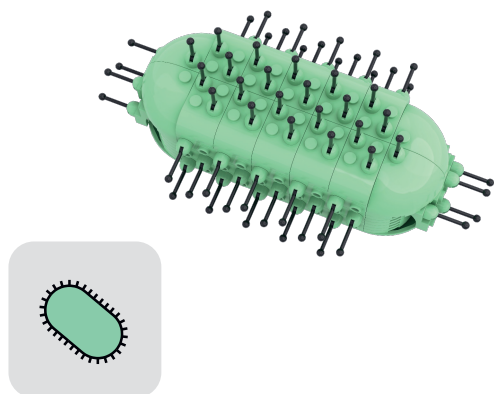




## Samenvatting voor Legopubliek

Zie ook de vorige twee pagina's voor aanvullende illustraties.

**MICROBEN zijn overal.** Alleen al in onze darmen wonen er veel, met name bacteriën. Ze komen voor in alle soorten en maten, met ieder een eigen taak. Bovendien werken ze graag samen. We weten dat de juiste combinatie van verschillende bacteriën, een positief effect heeft op onze gezondheid. Omdat ze echter met het blote oog niet te zien zijn en het er zo veel zijn, kunnen we vanaf de buitenkant lastig zien hoe ze dat precies doen.



*Akkermansia muciniphila, a mucus-loving bacteria / Een bacterie die van mucus houdt.*

Wetenschappers gebruiken vaak **MODELLEN** om na te bootsen hoe bacteriën zich in onze darmen gedragen. Een model kan bijvoorbeeld een proefdier zijn – denk aan een muis of varken. Dit noemen we *in vivo* modellen. Het wordt echter steeds duidelijker dat deze dieren, en hun darmbacteriën, anders zijn dan die van mensen. Daarom werken wetenschappers liever met darmcellen, de bouwstenen van onze darmen. Met bepaalde technieken kunnen ze deze cellen zó laten groeien in het

lab, dat ze zich gaan gedragen zoals in onze darmen. Deze technieken noemen we “*in vitro*”. In het eerste deel van mijn proefschrift benadrukken we (opnieuw) dat er nog wat verbeterpunten zijn in het opzetten en het beschrijven van *in vivo* én *in vitro* modellen.

In het tweede deel van mijn proefschrift kijken we naar een bepaalde groep bacteriën, namelijk degene die in de slijmlaag (mucuslaag) van onze darmen wonen. Deze laag bestaat voor een groot deel uit **MUCINES**, eiwitten die omhuld zijn met een dikke laag suikers. De darmbacteriën vinden met name deze suikers erg lekker en zetten deze suikers om in bruikbare stoffen voor zowel andere darmbacteriën als onze darmcellen. Wanneer we darmcellen in het lab groeien, produceren ze alleen niet zo veel mucus. Behalve wanneer we ze op een grote schudder zetten. Dit trucje hebben we afgekeken van andere wetenschappers, maar niemand wist nog precies waarom deze cellen dit doen. Een kleine domper: dit weten we nog steeds niet, maar we doen wel suggesties.

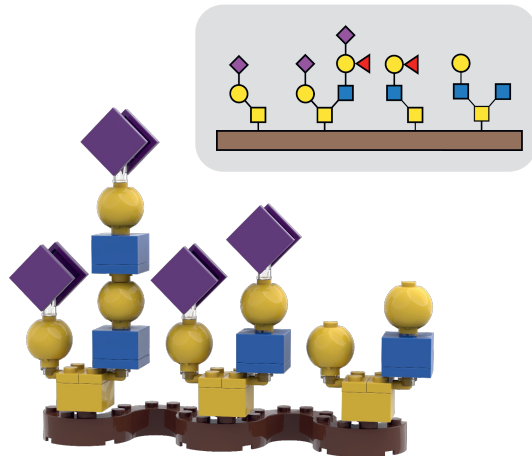
We hebben dezelfde cellen (in het lab) ook laten kennismaken met een darmbacterie die kampioen is in het afbreken van mucine: *Akkermansia muciniphila* (*muciniphila* betekent letterlijk “houdt van mucine”). Onder de omstandigheden die wij hebben

getest, zien we helaas weinig effect van deze bacterie op de darmcellen. In een ander experiment laten we wel zien wanneer *Akkermansia muciniphila* aan mucines bindt, namelijk als er bepaalde suikers aanwezig zijn. Dit geeft ons inzicht over hoe deze darmbacterie zich staande weet te houden in onze dichtbevolkte darmen.

Zoals bij elk onderzoek, is er nog veel werk te doen. In mijn proefschrift doe ik verschillende voorstellen over hoe we (relatief simpele) modellen kunnen inzetten om het samenspel tussen darmbacteriën en -cellen beter te leren begrijpen. Ook benadruk ik hoe belangrijk het is om deze experimenten goed op te zetten en te rapporteren. Een andere

mogelijkheid zou zijn om met wat ingewikkeldere technieken in het lab de darmen na te bootsen van verschillende personen (jong en oud, gezond en ziek, etc.), om te onderzoeken welk effect darmbacteriën hebben op de darmcellen en andersom. In elk geval moeten we niet vergeten om de aanwezigheid van bacteriën niet vergeten! Tot slot verdedig ik het belang van goede voorlichting over microben, voor zowel wetenschappers als een meer algemene publiek.

**Microben zijn immers overal.**



*Mucins, proteins covered in sugar // Mucines, eiwitten omhuld met een dikke laag suikers.*



## APPENDICES

### Abbreviations

ASF = Altered Schaedler Flora

ASC = adult stem cell

CAZyme = carbohydrate-active enzyme

CBM = carbohydrate binding module

CDI = *C. difficile* infection

CRC = colorectal cancer

ECM = extracellular matrix

ECV = extracellular vesicle

EPS = exopolysaccharide

ER = endoplasmic reticulum

FBS/FCS = fetal bovine / calf serum

FISH = fluorescent *in situ* hybridization

FMT = fecal microbiota transplantation

Gal = galactose

GalNAc = N-acetylgalactosamine

NCBI GEO = Gene Expression Omnibus from National Center for Biotechnology Information

GF = germ-free

GH = glycosyl hydrolase

GIT = gastrointestinal tract

Glc = glucose

GlcNAc = N-acetylglucosamine

GIVIMP = Good *In Vitro* Method Practices

HMI = host-microbe interactions

iHACS = Intestinal Hemi-Anaerobic Co-culture System

LOS = lipo-oligosaccharide

LPS = lipopolysaccharide

NAT = non-animal technology

NLR = Nucleotide-binding and oligomerization domain (NOD)-like receptor

PBS = phosphate buffered saline

PGM = porcine gastric mucin

PSC = pluripotent stem cell

PTS = proline, threonine, serine

PUL = polysaccharide-utilization locus

SCFA = short-chain fatty acid

SHIME = Simulator of the Human Intestinal Microbial Ecosystem

SIHUMI = Simplified HUMAN intestinal Microbiota

SPF = specific pathogen free

SWMS = Semi-Wet with Mechanical Stimulation

TEER = transepithelial electrical resistance

TIM = TNO intestinal model

TLR = Toll-like receptor

TR = tandem repeat

VNTR = variable number of tandem repeats



## Acknowledgements

No words can describe how grateful I am for the people that I during my PhD, and who supported me throughout, but... I did an attempt to write them down. Nevertheless, I would like to keep my appreciations personal and will not abuse this section to spend an inexplicable number of pages on inside jokes that others won't understand. Therefore, against all recommendations for transparent and open science, the message below is solely dedicated to you.

Is the message missing? Please contact me.

This image shows a full page of white paper with horizontal dotted lines. The lines are evenly spaced and run across the width of the page, providing a guide for handwriting practice. There are no margins, text, or other markings on the page.

[illegible]

### About the Author



Janneke Elzinga was born on the 11<sup>th</sup> of January, 1994 in Oosterhout, The Netherlands, where she grew up and obtained her Gymnasium degree at the Sint-Oelbertgymnasium (*cum laude*). Next, she moved to Nijmegen and got her BSc in Biomedical Sciences at the Radboud University in Nijmegen (*cum laude*). At the same university, she completed her MSc degree in Molecular Mechanisms of Disease (*cum laude*). For her MSc thesis, she investigated the degradation of human milk oligosaccharides by mucin-degrading specialist *Akkermansia muciniphila* at the Laboratory of Microbiology at Wageningen University &

Research (WUR), in the Molecular Ecology group led by Prof. dr Hauke Smidt. Fascinated by the world of intestinal microbes, she applied for a PhD project at the same department on the establishment of a gut-on-a-chip, which was part of a larger, NWO-funded consortium with researchers from WUR, University of Twente and industrial partners. However, she never lost her interest in *A. muciniphila* – her first microbial date – and collaborated with other scientists in- and outside WUR (see “List of Publications” on p. 299) to feed on mucus-related topics. This included an exchange at the Copenhagen Center for Glycomics during spring 2022, where she investigated the binding of *A. muciniphila* to mucin (**Chapter 6**). In contrast to *A. muciniphila*, Janneke is not a specialist: In summer 2020, she got the opportunity to combine her PhD with her passion for writing, and started a part-time job as a dissemination manager at UNLOCK, an open infrastructure facilitating research on microbial communities. Team play is key! This is something she also learnt in her free time, being active as a marathon ice speed skater. To compensate for the ice rink paradox, she did an attempt to reduce her ecological footprint in the lab as a member of Green Impact Team Helix/AFSG.

In the summer of 2023, Janneke is planning to continue her academic career as a Postdoc at the Copenhagen Center for Glycomics and Department of Veterinary and Animal Sciences in Copenhagen.

## List of Publications

**Janneke Elzinga**, John van der Oost, Willem M. de Vos WM, Hauke Smidt. The Use of Defined Microbial Communities to Model Host-Microbe Interactions in the Human Gut. *Microbiology and Molecular Biol Reviews* 2019;83(2): e00054-18.

<https://doi.org/10.1128%2FMMBR.00054-18>

**Janneke Elzinga\***, Benthe van der Lugt\*, Clara Belzer\*\*, Wilma T. Steegenga\*\* (2021) Characterization of increased mucus production of HT29-MTX-E12 cells grown under Semi-Wet interface with Mechanical Stimulation. *PLoS ONE* 16(12): e0261191.

<https://doi.org/10.1371/journal.pone.0261191>

Pilar Garcia-Vello, Hanne L. P. Tytgat, Joe Gray, **Janneke Elzinga**, Flaviana Di Lorenzo, Jacob Biboy, Daniela Vollmer, Cristina De Castro, Waldemar Vollmer, Willem M. de Vos, Antonio Molinaro, Peptidoglycan from *Akkermansia muciniphila* MucT: chemical structure and immunostimulatory properties of mucopeptides, *Glycobiology* Volume 32, Issue 8, August 2022, 712–719, <https://doi-org.ezproxy.library.wur.nl/10.1093/glycob/cwac027>

Ouwerkerk, Janneke P., Hanne L. P. Tytgat, **Janneke Elzinga**, Jasper Koehorst, Pieter Van den Abbeele, Bernard Henrissat, Miguel Gueimonde, Patrice D. Cani, Tom Van de Wiele, Clara Belzer, and Willem M. de Vos. 2022. Comparative Genomics and Physiology of *Akkermansia muciniphila* Isolates from Human Intestine Reveal Specialized Mucosal Adaptation. *Microorganisms* 10, no. 8: 1605.

<https://doi.org/10.3390/microorganisms10081605>

**Janneke Elzinga**, Menno Grouls, Guido J.E.J. Hooiveld, Meike van der Zande, Hauke Smidt, Hans Bouwmeester. Systematic comparison of transcriptomes of Caco-2 cells cultured under different cellular and physiological conditions. *Archives of Toxicology* 2023, 97(3): 737-753. <https://doi.org/10.1007/s00204-022-03430-y>

**Janneke Elzinga**, Yoshiki Narimatsu, Henrik Clausen, Willem M. de Vos\*, Hanne Tytgat\*. Binding of *Akkermansia muciniphila* to mucin is O-glycan specific. *Manuscript in preparation*.

Pilar Garcia-Vello, Hanne L.P. Tytgat, **Janneke Elzinga**, Matthias van Hul, Hubert Plovier, Ferran Fabregat Nieto, Patrice D. Cani, Simone Nicolardi, Cristina De Castro, Flaviana Di Lorenzo, Alba Silipo, Antonio Molinaro, Willem M. de Vos. Solving the Paradox: The Lipo-oligosaccharide of the Intestinal Symbiont *Akkermansia muciniphila* Shows An Exceptional Structure and Signals Via Toll-like Receptors 2 and 4. *Manuscript in preparation*.





## APPENDICES

Hanne L.P. Tytgat, Daan Swarts, **Janneke Elzinga**, Guus Verver, Thijs Nieuwkoop, Laura Huuskonen, Justus Reunanen, Alejandra Matamoros, Rosa Ester Forgione, Anneleen Segers, Sonsoles Martin-Santamaria, John van der Oost and Willem M. de Vos. Structural Characterization and TLR2-Signaling Capacity of *Akkermansia muciniphila* Amuc\_1100, A Gut-Stable Functional Oligomeric Protein Located on Type IV Pili *Manuscript in preparation*.

**Janneke Elzinga**, Sharon Geerlings, Esmée van Eck, Guido Hooiveld, Elsbeth Bossink, Benthe van der Lugt, Mathieu Odijk, Loes Segerink, Hauke Smidt, Wilma Steegenga, Clara Belzer. The Effect of Mucus-associated Bacteria on Intestinal Mucus Characteristics *In Vitro*. *Manuscript in preparation*.

\* or \*\* = *shared authors*

## Overview of Completed Training Activities

Name course/meeting	Organizing institute(s)	Year
<i>Discipline specific activities</i>		
Centennial symposium	Laboratory of Microbiology, WUR	2017
Human Microbiome Symposium 2018	EMBL	2018
Annual Gut Day	Laboratory of Microbiology, WUR	2018
Wageningen PhD Symposium	VLAG	2019
Annual Gut Day	Microbiotacenter	2019
Intestinal Microbiome of Humans and Animals	VLAG	2019
Consortium meetings BBOL-NWO MicroGut	MicroGut	2018-2021
NOCI/hDMT Gut-Liver-on-Chip meeting	hDMT	2019/2021
Summer Course Glycosciences	VLAG	2021
Symposium Translational Glcosciences	RIMLS Nijmegen	2021
KNVM Spring meeting	KNVM	2022
International secondment	Copenhagen Centre for Glycomics	2022
<i>General courses</i>		
VLAG PhD week	VLAG	2018
Communication with the Media and the General Public	WGS	2018
Reviewing a scientific paper	WGS	2018
Nogepa course 0.8 (H2S)	DTS Opleidingen	2018
Critical Thinking and Argumentation	WGS	2019
Supervising BSc and Msc thesis students	WGS	2020
Nuts and Bolts of Science Outreach course	American Society for Microbiology	2020
Writing Grant Proposals	Radboud into Languages	2022
Designing Effective Science Communications	Coursera, University of Colorado Boulder	2022
<i>Assisting in teaching and supervision activities</i>		
Teaching BSc/MSc course (MIB-30303, MIB-10306, FHM-30806)		2018-2022
Supervising students		2018-2022



## APPENDICES

<i>Other activities</i>		
Preparation of research proposal	Laboratory of Microbiology, WUR	2017
Weekly group meetings	BacGen/MolEco	2017-2022
Journal Club	MolEco	2018-2020
PhD meetings	Laboratory of Microbiology, WUR	2017-2022
PhD Trip	Laboratory of Microbiology, WUR	2019
Blogposts on Darmgezondheid.nl	Stichting Darmgezondheid	2019-2020
Reviewing scientific articles (6x)	Several journals	2020/2021

## About the Cover

Featuring the cover of this thesis is a digital reconstruction of a Lego® mini figure. In particular, it represents the mini fig Willemini, a gift that Janneke received from her mentor (Arjan de Brouwer) when she obtained her MSc degree. The gift was a nod to a poster presentation that Janneke gave during one of her MSc courses, using Lego® bricks to explain the concept of O-glycosylation. The idea was strong, but the performance was disappointing.

The Lego® mini fig became Janneke's mascot throughout her PhD. When teaching a microbiology course during their first year, Catarina Sales E Santos Loureiro and Janneke came up with the idea to launch an Instagram account (@phdpuns) featuring the mini fig. It was named "Willemini the Fox" after Willem de Vos, chair of the Laboratory of Microbiology at the time. Via this account, they made pun of #phdlife\*. The cover shows the Lego® mini fig – as a model of a human being – of which the intestine falls to pieces – in this case microbes and mucins, referring to the contents of this thesis. The microbes, models and mucins are completely built from Lego® bricks. In this context, Lego® can be seen as the building blocks of life, which refers to the title of the NWO-grant supporting Janneke's PhD project. Moreover, although it was not the initial aim of the grant, part of the thesis is concerned with O-glycosylation, completing the life story of Willemini. The fact that Janneke joined an existing collaboration\*\* with a lab in Denmark, the country of Lego®, to work on a project on O-glycosylation, is a welcome coincidence.

The cover design is a co-production of Pieter (Janneke's father) and Janneke and was made using Studio 2.0 by LEGO BrickLink, inc. and Adobe Photoshop.

\*Unfortunately, this account has been discontinued by Instagram due to an erroneous mistake in the account settings, which is Janneke to blame for. Even more unfortunate: this happened when Willemini had just returned home to Denmark, the country of Lego®.

\*\*For which many thanks to Hanne Tytgat!



### Colophon

The research described in this thesis was financially supported by the Netherlands Organisation for Scientific Research (NWO) in the framework of the [Building Blocks of Life](#) programme (737.016.003) and the FEMS Research and Training Grant (FEMS-GO-2021-069).

Financial support from Wageningen University (Laboratory of Microbiology) for printing this thesis is gratefully acknowledged.

Cover and lay-out design: Pieter and Janneke Elzinga

Puns intended

Lego® constructions made in Studio 2.0 (v2.23.3 (1)) by LEGO BrickLink, Inc.

Print: Proefschriftmaken | [www.proefschriftmaken.nl](http://www.proefschriftmaken.nl)



